The Natural Antisense Transcript DONE40 Derived from the lncRNA ENOD40 Locus Interacts with SET Domain Protein ASHR3 During Inception of Symbiosis in Arachis hypogaea
Pritha Ganguly, Dipan Roy, Troyee Das, Anindyia Kundu, Fabienne Cartieaux, Zhumur Ghosh, Maitrayee Dasgupta

To cite this version:
Pritha Ganguly, Dipan Roy, Troyee Das, Anindyia Kundu, Fabienne Cartieaux, et al.. The Natural Antisense Transcript DONE40 Derived from the lncRNA ENOD40 Locus Interacts with SET Domain Protein ASHR3 During Inception of Symbiosis in Arachis hypogaea. Molecular Plant-Microbe Interactions, American Phytopathological Society, 2021, 34 (9), pp.1057-1070. 10.1094/MPMI-12-20-0357-R . hal-03597062

HAL Id: hal-03597062
https://hal.inrae.fr/hal-03597062
Submitted on 4 Mar 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution 4.0 International License
The Natural Antisense Transcript DONE40 Derived from the IncRNA ENOD40 Locus Interacts with SET Domain Protein ASHR3 During Inception of Symbiosis in Arachis hypogaea

Pritha Ganguly,1 Dipan Roy,1 Troyee Das,2 Anindya Kundu,1 Fabienne Cartieaux,3 Zhumur Ghosh,2 and Maitrayee DasGupta1,†

1Department of Biochemistry, University of Calcutta, Kolkata, West Bengal, 700019, India
2Division of Bioinformatics, Bose Institute, Kolkata, West Bengal, 700054, India
3LSTM, Université de Montpellier, CIRAD, INRA, IRD, SupAgro, Montpellier, France

†Current address of A. Kundu: NIAB, Cambridge, U.K.

Accepted 29 April 2021.

The long noncoding RNA ENOD40 is required for cortical cell division during root nodule symbiosis (RNS) of legumes, though it is not essential for actinorhizal RNS. Our objective was to understand whether ENOD40 was required for actinorhizal nodule formation in Arachis hypogaea. AhENOD40 express from chromosome 5 (chr5) (AhENOD40-1) and chr15 (AhENOD40-2) during symbiosis, and RNA interference of these transcripts drastically affected nodulation, indicating the importance of ENOD40 in A. hypogaea. Furthermore, we demonstrated several distinct characteristics of ENOD40. (i) Natural antisense transcript (NAT) of ENOD40 was detected from the AhENOD40-1 locus (designated as NAT-AhDONE40). (ii) Both AhENOD40-1 and AhENOD40-2 had two exons, whereas NAT-AhDONE40 was monoexonic. Reverse-transcription quantitative PCR analysis indicated both sense and antisense transcripts to be present in both cytoplasm and nucleus, and their expression increased with the progress of symbiosis. (iii) RNA pull-down from whole cell extracts of infected roots at 4 days post-infection indicated NAT-AhDONE40 to interact with the SET (Su(var)3-9, enhancer of Zeste and Trithorax) domain containing Absent small homeotic disc (ASH) family protein AhASHR3 and this interaction was further validated using RNA immunoprecipitation and electrophoretic mobility shift assay. (iv) Chromatin immunoprecipitation assays indicate deposition of ASHR3-specific histone marks H3K36me3 and H3K4me3 in both of the ENOD40 loci during the progress of symbiosis. ASHR3 is known for its role in optimizing cell proliferation and reprogramming. Because both ASHR3 and ENOD40 from legumes cluster away from those in actinorhizal plants and other nonlegumes in phylogenetic distance trees, we hypothesize that the interaction of DONE40 with ASHR3 could have evolved for adapting the nodule organogenesis program for legumes.

Keywords: ENOD40, DONE40, ASHR3, long noncoding RNA, natural antisense transcript, root nodule symbiosis

In nitrogen-fixing root nodule symbiosis (RNS), a compatible rhizobia–legume interaction activates the Nod-factor-dependent SYM pathway that leads to a local trigger of complex phytohormonal signaling. Together with the SYM pathway, these phytohormonal signals reprogram the cortical cells and regulate their division, ultimately building a nodule meristem for the endocytic accommodation of the symbionts and generation of the nodule primordium (Limpens and Bisseling 2003; Stougaard 2001). During the inception of symbiosis, downstream to the SYM pathway, phytohormonal signaling and nitrate sensing (Mathesius et al. 2000), the long noncoding RNA (IncRNA) ENOD40 is induced to initiate the proliferation of root cortical cells (Crespi et al. 1994). The ENOD40 genes encode polyA containing RNAs of 0.7 kb that contain two short open reading frames (ORFs) and are also present in nonlegumes (Larsen 2003; Sousa et al. 2001). The 10- to 13-mer oligopeptide encoded by ORF1 is conserved among all species (Compaan et al. 2001) (Supplementary Fig. 2A and B). ORF2 possibly does not encode for peptides but may contribute to the folding of the RNA into a highly structured form (Compaan et al. 2001; Girard et al. 2003; Sousa et al. 2001). The region encompassing ORF1 and ORF2 (inter-ORF) shows a high degree of conservation, and this region of ENOD40 RNA sequences tends to form particularly stable secondary structures, indicating these regions to be functionally essential for ENOD40 (Wan et al. 2007). Thus, the ENOD40 RNA forms a special class of RNA in plants known as bifunctional RNAs, with both coding and noncoding characteristics (Kouchi and Hata 1993; Röhrig et al. 2002).

The possible functions of legume ENOD40 have been argued mainly in favor of the induction of cortical cell divisions that lead
to the initiation of nodule primordia and appropriate differentiation and development of nodules (Charon et al. 1997, 1999; Mylona et al. 1995). Expression of infection-related genes is independent of ENOD40 activation, indicating rhizobial invasion to be independent of ENOD40 (Kumagai et al. 2006). Ectopic expression of ENOD40 induces dedifferentiation and extensive division of cortical cells, leading to increased nodules at early time points (Charon et al. 1997). On the other hand, the transcription factor nodule number control 1 directly binds to the ENOD40 promoter to repress it for restricting nodule numbers (Wang et al. 2014). Identification of the interactors has also provided leads in understanding ENOD40 action. For example, ENOD40 interacts with MtRBP1, a homolog of nuclear speckle RNA binding protein (RBP) in Medicago truncatula, where ENOD40 is believed to relocalize RBP1 to the cytoplasm from nucleus speckle (Campalans et al. 2004). This appears similar to the ASCO lncRNA in Arabidopsis, which binds to nuclear speckle RBPs and thereby regulates plant root development (Bardou et al. 2014). The peptide from ORF1 of ENOD40 from Glycine max binds via S-thiolation to sucrose synthase, an important enzyme in sucrose utilization, suggesting a role for ENOD40 peptides in photosynthate accumulation in sink regions (Röhrig et al. 2002; Winter and Huber 2000). The pronounced impact of ENOD40 in the developmental program of nodule organogenesis involving initiation of dedifferentiation and division of cortical cells have been extensively documented. However, the intricate molecular mechanisms associated with the process remain to be understood.

In this study, we analyzed ENOD40 genes in a dalbergoid legume, Arachis hypogaea. In these plants, rhizobia invade through natural “cracks” to directly access the cortical cells that divide and develop the characteristic aescynomenoid nodules with uniformly infected central tissue and loss of meristematic activity (Fedorova et al. 2017; Lavin et al. 2001; Tajima et al. 2008). Herein, we show ENOD40 to be encoded from both chromosome 5 (chr5) (designated as AhENOD40-1) and chr15 (designated as AhENOD40-2), respectively.

Additionally, a natural antisense transcript (NAT) of ENOD40 was detected from the AhENOD40-1 locus at chr5, and was designated as NAT-AhDONE40. We demonstrated the interaction of NAT-AhDONE40 with AhASHR3 (ABSENT SMALL HOMEOCTIC DISC1 RELATED3), a member of the Trithorax group of proteins, and have observed significant enrichment of activating histone modification marks at the ENOD40 locus from both chr5 and chr15 of A. hypogaea during the progress of rhizobial infection. Moreover, with ENOD40 RNA being required for initiation of dedifferentiation and division of cortical cells during nodule organogenesis (Charon et al. 1997), and the ASHR3 protein being required for the control of cell division at the root meristem and quiescent center in Arabidopsis (Kumpf et al. 2014), the interaction between AhASHR3 and AhENOD40 may have implications in the regulation of rhizobia-induced cortical cell division at the nodule primordia in legumes.

RESULTS
ENOD40 is essential for symbiosis in A. hypogaea.
ENOD40 is yet to be annotated in the cultivated peanut genome (tetraploid) database or either of its two parental genomes (diploid), A. duranensis (AA) and A. ipaensis (BB). All ENOD40 sequences reported thus far contain two short conserved regions, ORF1 and ORF2, embedded within nonhomologous sequences (Compaan et al. 2001). BLAST search was done against the cultivated peanut genome database (araphy.Tifrunner.gnm1) with M. truncatula ENOD40 sequences (X80262), whereby putative orthologs of ENOD40 in A. hypogaea were identified in chr5 (AA) and chr15 (BB). Additionally, BLAST search using Lupinus luteus ENOD40 sequences (AF352375) identified putative orthologs of ENOD40 in chr3 (AA) and chr13 (BB) although, in this case, only sequences homologous to ORF2 were identified (Supplementary Fig. S1). Alignment of these putative ENOD40 sequences from A. hypogaea indicated the core sequences of ORF1 and ORF2 to be conserved in all four loci (Supplementary Fig. S2A and B). Microsynteny analysis of AhENOD40 loci revealed strongly conserved gene collinearity with other ENOD40s from legumes, indicating them to be orthologous (Fig. 1A). In all legumes, ENOD40 was convergent with YELLOW LEAF SPECIFIC GENE 7-like (YLS7), a member of Trichome birefringence family protein within 2Kb. Intriguingly, in Aeschynomene evenia, a Nod-independent crack entry legume, AeENOD40 loci from chr5 was found to be in synteny with other legumes but, in chr3, YLS7 was absent. Within nonlegumes, synteny of ENOD40 with YLS7 was only conserved in actinorhizal plants although, in most members, only ORF2 sequences were conserved (Supplementary Table S1). In contrast, in nonlegumes such as Populus trichocarpa, Vitis vinifera, and Solanum lycopersicum, only ORF2 sequence was detectable, but it was not syntenic with YLS7. Similarly, in Zea mays and Oryza sativa, ENOD40 and YLS7 were located in different chromosomes and, in Arabidopsis thaliana and Brassica rapa, ENOD40 was completely undetectable, like other SYM gene homologs (Griesmann et al. 2018).

Our first objective was to understand whether ENOD40 was important for RNS in A. hypogaea. For this, RNA interference of AhENOD40 was done targeting the region encompassing ORF1 and ORF2 that was conserved in all four loci detected in A. hypogaea. Quantitative reverse-transcription PCR (RT-qPCR) for AhENOD40 revealed approximately 66 to 83% reduction of the AhENOD40 transcripts in transgenic hairy roots (n = 20) (Fig. 1B). Both RNA interference (RNAs) and empty-vector-transformed composite plants were scored for nodulation at 28 days postinfection (DPI) (Fig. 1C). The downregulation of AhENOD40 resulted in approximately 80% reduction in nodule number in RNAi as compared with empty-vector-transformed hairy roots. The infection zone (IZ) of these matured empty-vector-transformed control nodules showed the presence of uniformly infected plant cells filled with spherical symbiosomes (Fig. 1D). In contrast, the AhENOD40-RNAi nodules appeared to be whitish, and the IZ showed the presence of undifferentiated rod-shaped bacteria. We concluded ENOD40 to have a role in bacteroid differentiation in addition to its known role in cortical cell division during nodule organogenesis. This is similar to what we observed in AhHK1-RNAi and AhCYCLOPS-RNAi conditions, where undifferentiated rhizobia failed to differentiate into spherical symbiosomes (Das et al. 2019; Kundu and DasGupta 2018). For the readout of the symbiotic signaling, we chose to monitor the symbiotic genes such as CYCLOPS and NIN, whose expressions were previously reported to be induced during the progress of symbiosis in A. hypogaea (Karmarkar et al. 2019). CYCLOPS acts as a transcription factor specifically binding to the CYC-box element present in the promoter of NIN to drive its expression in Lotus japonicus during nodulation (Singh et al. 2014). Expression of AhCYCLOPS was approximately fivefold lower in the AhENOD40-RNAi nodules, whereas expression of NIN was unaffected (Fig. 1E). The decrease in AhCYCLOPS could be due to a feed-forward positive effect of ENOD40 on CYCLOPS expression, and the absence of any change in AhNIN expression indicates that AhENOD40 is activated downstream from AhNIN during organogenesis, as noted earlier (Liu et al. 2019).

Both sense and antisense ENOD40 is expressed during symbiosis in A. hypogaea.

Transcriptome analysis indicated a significant change of expression of AhENOD40 from chr5 and chr15 with the progress of symbiosis, whereas expression from chr3 and 13 showed no change (Supplementary Fig. S3A). Therefore, we focused our investigation
on AhENOD40 alleles encoded from chr5 and 15. Transcript assembly by Cufflinks followed by the screening of intergenic IncRNAs (as described in Materials and Methods) revealed that both of the alleles for AhENOD40 in chr5 and 15 coded for IncRNA where the nearest protein-coding gene (pcGene) was YLS7 (Fig. 1A). The coding potential for the AhENOD40 sequences was calculated using the Coding Potential Calculator and is shown in Supplementary Figure S3B. AhENOD40 was transcribed from the

---

**Fig. 1.** ENOD40 is essential for symbiosis in *Arachis hypogea*. A, Syntenic relationship of ENOD40. YLS7-ENOD40 enclosed in a dashed box; YLS7 = pink and ENOD40 = yellow (indicated by a red star). Gene IDs from *Lotus japonicus*, *A. hypogaea*, *Glycine max*, *Medicago truncatula*, *Phaseolus vulgaris*, and *Aeschynomene evenia* are shown on the left. B, Quantitative reverse-transcription PCR (RT-qPCR) analysis of AhENOD40 in AhENOD40-RNA interference (RNAi) roots relative to empty-vector-transformed roots (control) at 28 days postinfection (DPI). Primers AhENOD40F-RT and AhENOD40R-RT were used for the RT-qPCR. *AhActin* was used as a reference gene. The histogram represents an average of three biological replicates, each having *n* > 4 root systems, and the error bar represents standard deviation. C, Histogram shows the average number of nodules in transformed (gray) and nontransformed (black) hairy roots in control and AhENOD40-RNAi plants at 28 DPI. Error bar represents the standard deviation (*n* = 30 for control and AhENOD40-RNAi plants). D, Section of nodule in control and AhENOD40-RNAi roots at 28DPI. Image shows propidium iodide (red) and calcofluor (blue) merged. Scale bar = 10 µm. Scanning electron microscopy image of the infection zone in control and AhENOD40-RNAi nodule. Scale bar = 2 µm. E, RT-qPCR analysis of symbiotic marker *AhNIN* and *AhCYCLOPS* in AhENOD40-RNAi roots relative to the control at 28 DPI. *AhActin* was used as a reference gene. The histogram represents an average of three biological replicates, each having *n* > 4 root systems, and the error bar represents standard deviation.
negative strand in the loci arahy.Tifrunner.gnm1.Arathy.05:33566965-33568044 and from the positive strand in the loci arathy.-Tifrunner.gnm1.Arathy.15:24083907-24084972. Sequences of the reads mapped to these loci are shown in Figure 2A, with their detail in Supplementary Table S2. A distinguishing feature of AhENOD40 transcribed from chr5 was an insertion of 28 bp (position 33,567,419 to 33,567,446) within the ORF2, which was absent in ENOD40 transcribed from chr15 (Fig. 2A; Supplementary Fig. S3C). We designate ENOD40 from chr5 as AhENOD40-1 and from chr15 as AhENOD40-2. Such multiple copies of ENOD40 were previously noted in both legumes and nonlegumes (Crespi et al. 1994; Fang and Hirsch 1998; Fletetakis et al. 2000; Roussis et al. 1995). The alignment of mapped reads on the A. hypogaea genome suggested both AhENOD40-1 and AhENOD40-2 to have two exons (Fig. 2A). To verify, the complementary DNAs (cDNAs) for AhENOD40 were amplified using a gene-specific primer and oligodT-adapter primer from messenger RNA isolated from nodulated roots, where AhENOD40-1 and AhENOD40-2 contained 719 bp (accession number MW048740) and 688 bp (accession number MW048741), respectively (Fig. 2B). We verified the presence of a 361-bp intron in AhENOD40-1 and a 377-bp intron in AhENOD40-2 by PCR amplification on genomic DNA (gDNA) and cDNA by exon-specific and primer-specific primers (Fig. 2B). The length of exon1 and exon2 in AhENOD40-1 was 22 and 697 bp, respectively, and in AhENOD40-2 was 69 and 619 bp, respectively (Fig. 2C). This exercise allowed us to verify the exon intron boundary, and the splice sites were identified. The terminal exon 3’ acceptor site is conserved for both AhENOD40-1 and AhENOD40-2, though the 5’ donor site varies (Fig. 2D).

The most distinguishing feature of ENOD40 loci in A. hypogaea was the detection of natural antisense ENOD40 transcripts from our strand-specific transcriptome. Sequences of these antisense read mapped to arathy.Tifrunner.gnm1.Arathy.05:33567054-33567703 overlapping with exon2 and partially with the intron of AhENOD40-1 (Fig. 2A). We have designated the antisense transcript as NAT-AhDONE40. The antisense transcript was amplified from gDNA and cDNA from nodulated roots and was confirmed to be monoxonic of 649 bp (accession number MW048742) (Fig. 2C). Additionally, NAT-AhDONE40 was also predicted to be an IncRNA (Supplementary Fig. S3B). Together, these findings indicate the existence of a natural antisense transcript for AhENOD40-1, where it may function as a sense-antisense (SAS) IncRNA pair. It may be noted that, in ENOD40 RNAi roots, there was significant downregulation of both the sense (AhENOD40-1; approximately 930-fold, AhENOD40-2; approximately 40-fold) and antisense (NAT-AhDONE40; approximately 30-fold) transcripts, indicating that the observed phenotypic effect could be contributed by either or both of the sense and antisense transcripts (Supplementary Fig. S3D). Previously, we have analyzed the transcriptome of five distinct stages of symbiosis in A. hypogaea as follows: 1 DPI = recognition and invasion, 4 DPI = primordia formation, 8 DPI = nodule-like structure, 12 DPI = mature nodules with spherical symbiosomes (Karmakar et al. 2019). To detect the sense (AhENOD40-1 and AhENOD40-2) or antisense (NAT-AhDONE40) transcripts and validate their symbiotic pattern of expression revealed by the transcriptomic data (Supplementary Fig. S3A), we performed strand- and chromosome-specific RT, and the relative levels of all of the RNAs were quantified by qPCR (Fig. 2B and E). The RT-qPCR estimation of these transcripts follows the same expression pattern as was revealed by the raw count of the transcripts (Fig. 2E; Supplementary Fig. S3A) and indicated ENOD40 transcripts (both sense and antisense) to increase from the primordia formation stage at 4 DPI. This symbiosis-associated expression of the NAT-AhDONE40 strongly suggest its importance during the process. We then purified nuclear RNA and, using the same strategy, showed that both sense and antisense transcripts were also detected in the nucleus (Fig. 2E). These data are in accordance with the previous reports where ENOD40 was noted to shuttle between cytoplasm and nucleus (Campalans et al. 2004). The dynamic nature of expression and distribution of both sense and antisense transcripts of AhENOD40 during symbiosis indicate them to be functionally important during nodule organogenesis in A. hypogaea.

RNA pull-down indicates NAT-AhDONE40 RNA to interact with AhASHR3.

Previous investigations have identified RBP1 from nodular extracts to be an interactor of sense ENOD40 (Campalans et al. 2004). To understand the biological significance of NAT-AhDONE40, we explored the use of RNA pull-down assays coupled with mass spectrometry (MS) analysis to identify the potential interactor proteins of NAT-AhDONE40. We attempted to identify such interactors in whole-cell extracts of infected roots at 4 DPI because the expression of both sense and antisense AhENOD40 significantly increases from 4 DPI onward (Fig. 2E). Moreover, our previous investigation revealed an onset of a major transcriptional program at 4 DPI during the primordia formation (Karmakar et al. 2019). For RNA pull-down, we used the ORF1-interORF-ORF2 region (298 bp) because earlier reports have indicated that both ORF1 and ORF2 and the interORF region are involved in the regulation of ENOD40 activity in Medicago roots (Sousa et al. 2001). Additionally, the interORF region of ENOD40s was previously noted to have conserved secondary structure (Sousa et al. 2001), and our analysis indicated that both sense (AhENOD40-1 and AhENOD40-2) and antisense (NAT-AhDONE40) RNA has that secondary structure conserved in the interORF region (Supplementary Fig. S4). Following RNA pull-down assays, matrix-assisted laser desorption-ionization time of flight (MALDI-TOF) analysis was performed, where we considered proteins having at least two unique peptides with a Mascot Ion Score above 20 as potential interactors of AhDONE40 (Supplementary Table S3). As a control, MALDI-TOF analysis was also performed using AhENOD40-1 as bait to parallelly profile the potential interactor proteins of AhENOD40-1 (Supplementary Table S4). The best hit with AhDONE40 was a putative SU(var)3-9, enhancer of Zeste and Trithorax (SET) domain containing protein of the Trithorax family from Arabidopsis thaliana, which was not found among the interactors of AhENOD40-1. Taking this as a clue, the RNA pull-down experiment was repeated, where the interacting proteins were analyzed using liquid chromatography tandem MS (LC-MS/MS). In this case, the peak lists were searched against a forward and reverse Arabidopsis thaliana UniProt database appended with all 50 SET domain-containing proteins in A. hypogaea. We could detect three spectra representing unique peptides from AhXDS33H1 (chr1, AA) and AhBQLE9W (chr11, BB) (Fig. 3A). We developed a distance tree with all of the SET domain-containing proteins from A. hypogaea and Arabidopsis thaliana, where AhXDS33H1 and AhBQLE9W clustered with AtASHR3 within the ASH group and, therefore, was predicted to be an ortholog of AtASHR3 (Supplementary Fig. S5A). We profiled the relative (log-twofold) expression values of all 50 SET domain-containing proteins of A. hypogaea from our symbiotic transcriptome (GSE98997), where the expression of Arabu.4C25F (AhXDS33H1, chr1) and Araip.VL728 (AhBQLE9W, chr11) was significantly high at 4 and 8 DPI (Fig. 3B). However, the absolute values of expression indicated only AhXDS33H1- but not AhBQLE9W-encoded AhASHR3 to be highly expressed during the progress of symbiosis (Fig. 3C). RT-qPCR analysis then confirmed the expression of AhXDS33H1-encoded AhASHR3 to be maximum at 4 DPI (Fig. 3D). Interestingly, in AhENOD40-RNAi, there was no significant change in the expression of AhASHR3, which shows that ENOD40 transcripts may not be
Fig. 2. Both sense and antisense ENOD40 is expressed during symbiosis in Arachis hypogaea. A, Alignment of the Sequence Read Archive contigs with the genome sequence of Arahy.05 (33,566,964 to 33,568,044) and Arahy.15 (24,083,907 to 24,084,972). Contigs from positive-strand coding for AhDONE40 are marked with green (33,567,054 to 33,567,703). Contigs from negative-strand coding for AhENOD40-1 are marked with red (33,566,964 to 33,567,661 and 33,568,022 to 33,568,044). Contigs from positive-strand coding for AhENOD40-2 are marked with green (24,083,907 to 24,083,976 and 24,084,353 to 24,084,972). All contigs are aligned in the 5’ to 3’ direction. Conserved boxes are aligned to the 5’ to 3’ direction of the peptides translated. Open reading frame 1 (ORF1) is marked with yellow and ORF2 with orange in both AhENOD40-1 and AhENOD40-2. The blue shaded region represents the 28-nucleotide insertion in Arahy.05.

B, Schematic representation of the primer positions in AhDONE40, AhENOD40-1, and AhENOD40-2 transcripts. Exons are marked with black boxes, and ORF1 and ORF2 are denoted as in A. The 28-bp insertion in Arahy.05 is denoted by a blue box and the region of insertion in Arahy.15 by a blue triangle. The transcriptional start site is denoted by a black arrow. Primer combinations are marked on each lane. D, Alignment of splicing boundaries of AhENOD40-1 and AhENOD40-2 with splice sites highlighted in blue. The black box represents exon position. E, Relative transcript abundance of AhENOD40-1, AhENOD40-2, and AhDONE40 in the total RNA and nuclear RNA in different time points (1, 4, 8, 12, and 21 days postinfection). Transcript levels were normalized relative to the histone H3 gene. Error bars represent the standard deviation for three biological replicates. Two-way analysis of variance was used to assess significant differences, where *P < 0.05.
Fig. 3. RNA pull-down indicates that AhDONE40 RNA interacts with AhASHR3. A, Product ion spectrum of indicated peptides of AhASHR3. B, Phylogenetic analysis of putative Su(var)3-9, enhancer of Zeste and Trithorax (SET) domain-containing proteins from Arachis hypogaea database. The Atx, E(z), SMYD, Ash, and SUV groups are marked. C, Absolute counts of AhXD33H1 and AhBQLE9W transcripts are plotted. D, Relative transcript abundance of AhASHR3 was determined by quantitative reverse-transcription PCR at different time points post rhizobial infection. Transcript levels were calculated relative to the expression of histone H3. Error bars represent standard errors for three biological replicates. E, Schematic representation of the AhASHR3 protein (accession number MW167779) with its respective domains. F, Phylogenetic analysis of ASHR3 and ENOD40 genes. Symbiotic leguminous plants are marked in green, actinorhizal plants in red, and the nonsymbiotic plants in blue.
required for ASHR3 expression (Supplementary Fig. S5B). The 1,563-bp coding sequence of putative AhASHR3 (GenBank accession number MW167779) was generated by PCR amplification from cDNA prepared from infected roots at 4 DPI. AhASHR3 contains dual PHD domains at the N-terminal region (residues 145 to 193 and 194 to 239), an associated-with-SET (AWS) domain (residues 306 to 351), SET domain (residues 350 to 473), and post-SET domain (residues 473 to 489) (Fig. 3E).

Unlike legumes, ENOD40 expression does not increase during actinorhizal symbiosis and, in both actinorhizal plants and nonlegumes, ENOD40 expression is restricted to vascular bundles (Kouchi et al. 1999; Santi et al. 2003). Again, actinorhizal nodules appear to be modified lateral roots with central vascular tissue, whereas legume nodules represent stem-like organs with a peripheral vascular system (Pawlowski et al. 2003). It may be noted that, in distance trees, both ENOD40 and ASHR3 were completely segregated between legumes, nonlegumes, and actinorhizal plants (Fig. 3F). Such congruence between phylogeny and phenotype indicates the biological implication of ENOD40–ASHR3 interaction during the origin of nodule primordia in legumes.

AhASHR3 directly binds NAT-AhDONE40.

Our next objective was to validate AhENOD40–AhASHR3 interaction in vivo by RNA immunoprecipitation (RIP) assays. For this, we overexpressed the conjugate protein AhASHR3-green fluorescent protein (GFP) in hairy roots of A. hypogaea, and interacting RNAs were precipitated using an anti-GFP antibody. Because both sense and antisense transcripts of AhENOD40 were noted in the nucleus and cytoplasm, we used total cellular extracts of 35S::AhASHR3-GFP-transformed roots at 4 DPI for the RIP assay. As a control, empty-vector-transformed roots expressing GFP were analyzed under identical conditions. Interestingly, only NAT-AhDONE40 was significantly enriched by immunoprecipitation of AhASHR3-GFP, whereas detection of sense transcripts AhENOD40-1 and AhENOD40-2 was insignificant (Fig. 4A). It may be noted that, in vector-transformed roots, there was no enrichment of both sense and antisense ENOD40 transcripts in anti-GFP immunoprecipitate, confirming the specificity of the observed NAT-AhDONE40–ASHR3-GFP interaction.

Finally, our objective was to check whether AhASHR3 directly binds to NAT-AhDONE40. For this, electrophoretic mobility shift assays (EMSAs) were employed using the ORF1-interORF-ORF2 region (298 bp) of NAT-AhDONE40 that was used for our pull-down experiment. For the AhASHR3 protein, we used the AWS-SET-postSET region (184 amino acids) because several factors indicated that the preSET-SET-postSET boundaries in the SET family proteins contain motifs that tightly bind single-strand DNA and RNA (Krajewski and Vassiliev 2011; Krajewski et al. 2005) (Fig. 3E; Supplementary Fig. S6). The representative gel in Figure 4B shows the change in NAT-AhDONE40 migration on the native gel in the presence of AhASHR3. The free RNA migrates as multiple slow-moving bands and a single fast-moving band, indicating different conformers of NAT-AhDONE40. In the presence of AhASHR3, the mobility shift of these bands indicated that NAT-AhDONE40 could directly bind with AhASHR3. Unlabeled NAT-AhDONE40 in 50-fold molar excess successfully competed for the binding and release of the free probe, indicating the specificity of the interaction. Unlabeled AhENOD40–1 could also compete for the binding at 200-fold molar excess but the release of the free RNA probe was not noted, further indicating the specificity of interaction with NAT-AhDONE40 and AhASHR3 (Fig. 4B). AhASHR3 does not interact with a highly structured 18S rRNA, indicating the specificity of the observed interaction between AhASHR3 and NAT-AhDONE40 transcripts (Supplementary Fig. S6).
Fig. S7A). When we tested the direct interaction of AhASHR3 with AhENOD40-1, we could detect shifted bands that were outcompeted by both unlabeled AhENOD40-1 and NAT-AhDONE40 (Supplementary Fig. S7B). It may be noted that NAT-AhDONE and AhENOD40 have approximately 34% sequence identity, which may explain the interaction of AhENOD40-1 in vitro (Supplementary Fig. S7C). Overall, our results indicate that both sense and antisense transcripts of AhENOD40 have the potential to interact with AhASHR3 but only the interaction between the antisense transcripts could be biologically significant.

**ENOD40 loci from both chr5 and chr15 undergo dynamic histone H3 modifications during the progress of symbiosis in A. hypogaea.**

To understand whether there is any link between the NAT-AhDONE40–AhASHR3 interaction and epigenetic regulation of the expression of ENOD40, we followed the dynamic changes in the histone modification landscape of the ENOD40 loci. We chose to monitor the levels of H3K36me3 and H3K4me3, which are specific marks associated with ASH1 homologs (Cartagena et al. 2008). Additionally, we monitored H3K27ac that is linked to Trithorax group (TrxG)-mediated activation (Geisler and Paro 2015; Schuettengruber et al. 2007) and H3K9ac as a general activation marker (Schuettengruber et al. 2011). A repressive H3K27me3 mark linked to Polycomb complex group (PcG)-mediated silencing (Mikulski et al. 2017) was also monitored. We used the same set of primers to distinguish between the ENOD40 locus from chr5 and chr15 and monitored the above marks in both loci during the early stages of symbiosis in A. hypogaea roots, where AhASHR3 expression was maximally upregulated (Fig. 5A). At 4 DPI in both ENOD40 loci, there was a significant enrichment of H3K36me3, the ASHR3-specific mark (Lee et al. 2020; Zhang and Ma 2012), over the other TrxG-mediated activation marks H3K4me3 and H3K27ac (Fig. 5B). At 8 DPI, all of these activation marks were similarly enriched, though enrichment of H3K27ac remained insignificant at the ENOD40 loci in chr15. The general activation mark H3K9ac was also significantly upregulated at both loci at 8 DPI, and the repressive mark H3K27me3 was absent in all of the stages. Our results reveal a robust enrichment of H3K36me3 at both ENOD40 loci that is associated with the significant increase in expression of AhENOD40-1, AhENOD40-2, and NAT-AhDONE40 in A. hypogaea roots during the onset of symbiosis. This evidence further indicates the functional implications of the interaction between AhASHR3 and NAT-AhDONE40, where the interaction may facilitate a feed-forward loop that promotes the expression of ENOD40 via the recruitment of ASHR3 to the ENOD40 loci.

**DISCUSSION**

In dalbergoid legumes such as A. hypogaea and A. evenia, rhizobia directly invade cortical cells through epidermal cracks to generate the aeschynomenoid nodules (Sprent and James 2007). Our objective was to functionally characterize ENOD40 during the aeschynomenoid nodule development in A. hypogaea. During symbiosis, ENOD40 is expressed from chr5 (AhENOD40–1) and chr15 (AhENOD40–2), which differ by insertion of 28 bp in the ORF2 region of AhENOD40–1 (Fig. 2A; Supplementary Fig. 5).
S3C). As in all other legumes, ENOD40-RNAs drastically reduced the nodule numbers in *A. hypogaea* (Fig. 1C), which is in accordance with its role in nodule primordia formation (Kumagai et al. 2006). Thus, the function of ENOD40 appeared to be conserved in legumes. In all tested legumes, ENOD40 was found to be in synteny with YLS7, a Trichome birefringence-like gene, in a convergent manner (Fig. 1A). Trichome birefringence genes have a role in plant–pathogen interaction by acetylation of the cell wall and can also play a role in rhizobial accommodation during symbiosis (Escudero et al. 2017; Sun et al. 2020). Because sense and antisense lncRNAs are defined according to the nearest pcGene position (RIKEN Genome Exploration Research Group et al. 2005), sense AhENOD40 appears to be an antisense lncRNA for YLS7 because their transcriptions are convergent. The only exception was chr3 of *A. evenia*, where YLS7 was not detectable, and ENOD40 was convergent with a leucine-rich repeat receptor-like kinase gene. Transcriptomic analysis of *A. evenia* during symbiosis revealed *AceENOD40* to be expressed from chr3 and not chr5, where ENOD40 was syntenic with YLS7 (Supplementary Fig. S8). It may be noted that, similar to actinorhizal plants, *A. evenia* was a nod-factor-independent legume. Though the significance of proximity of YLS7 and ENOD40 is not understood, its association with nod factor dependence is intriguing.

Our results reveal the presence of a natural antisense transcript of *ENOD40* (NAT-AhDONE40) encoded from the *AhENOD40-1* locus in *A. hypogaea*, which is unique and has never been reported earlier. This makes *AhENOD40-1*–NAT-AhDONE40 a SAS gene pair, where opposite genomic strands within the same locus get transcribed (Galante et al. 2007). Because *AhENOD40-1* and *AhDONE40-2* differ by a 28-bp insertion, AhDONE40 may also act as a trans-NAT for AhENOD40. Transcriptome analysis of numerous model organisms reveals that a significant subset represents SAS partners involving both pcGenes and lncRNA, most often one coding and one noncoding in each pair (Babak et al. 2007; Georg and Hess 2011; Grinchuk et al. 2010; Hongay et al. 2013). Transcriptome analysis of *A. evenia* during symbiosis revealed *AceENOD40* to be expressed from chr3 and not chr5, where ENOD40 was syntenic with YLS7 (Supplementary Fig. S8). It may be noted that, similar to actinorhizal plants, *A. evenia* was a nod-factor-independent legume. Though the significance of proximity of YLS7 and ENOD40 is not understood, its association with nod factor dependence is intriguing.

From a functional standpoint, both ENOD40 RNA and ASHR3 have been previously shown to be implicated in the control of cell division in different independent studies (Charon et al. 1997; Kumpf et al. 2014). Mutations in ASHR3 in *Arabidopsis* disrupts the pattern of coordinated DNA replication and cell division and increase the cell division rate in the quiescent center of roots (Kumpf et al. 2014), whereas overexpression of *ENOD40* in nitrogen-deprived transgenic *Medicago* plants resulted in extensive cortical cell division in the roots (Charon et al. 1997). Considering the involvement of both ASHR3 and ENOD40 RNA in the control of cell division in roots, it can be envisaged that the observed interaction of NAT-AhDONE40 and ASHR3 may act as a mechanism for the regulation of cortical cell division during RNS. Regardless of the downstream functional implications, the interaction between AhASHR3 and NAT-AhDONE40 connects, for the first time, the antisense transcripts of an lncRNA implicated in the development of a novel organ (nodule) with master epigenetic controllers of developmental processes (PcG/TrxG complexes) that are ancient and evolutionarily conserved across kingdoms. Such interactions between lncRNA and chromatin modifying complexes are well documented in both plant and animal systems. For example, in plants, the lncRNAs COOLAIR and COLDWRAP generated from the FLC locus, and AG-ncRNA4 transcribed from the second intron of AGAMOUS (AG), interact with CURLY LEAF, a plant E(z) homolog in *PcG* (Berry and Dean 2015; Heo and Sung 2011; Kim and Sung 2017; Wu et al. 2018). Another example is the interaction of lncRNA AUXIN-REGULATED PROMOTER LOOP RNA with LIKE HETEROMERIC PROTEIN 1, a PcG protein that regulates the auxin response in roots (Ariel et al. 2014). The MAF4 antisense RNA (MAS-NAT) produced from the MADS AFFECTING FLOWERING 4 (MAF4) locus interacts with WDR5a, a TrxG member that activates the MAF4 locus (Zhao et al. 2018). In the animal system, the *Xist* and HOTAIR lncRNA interact with the PcG protein EZH2 (Rinn et al. 2007; Zhao et al. 2008), and the lncRNA DBR-T interacts with the TrxG protein ASH1-like (Cabianca et al. 2012).

Intronic interactions with the TrxG and PcG complexes are associated with epigenetic regulation. For example, the lncRNAs COOLAIR and AG-ncRNA4 interaction with CURLY LEAF leads to deposition of the repressive mark H3K27me3 in the FLC and AG locus, respectively (Csorba et al. 2014; Wu et al. 2018). Both COOLAIR and TWISTED LEAF cause synchronized replacement of activating histone marks (H3K36me3 and H3K4me3) with the repressive ones (H3K27me3), highlighting the role of lncRNAs in coordinating the switching of chromatin states (Csorba et al. 2014; Liu et al. 2018). On the other hand, MAS-NAT interaction with a TrxG member (WDR5a) leads to deposition of activation mark (H3K4me3) at the MAF4 locus (Zhao et al. 2018). This is similar to our results, where we demonstrate a significant increase in deposition of activation marks (H3K36me3, H3K4me3, H3K27ac, and H3K9ac) on ENOD40 loci. Among these modifications, H3K36me3 and H3K4me3 are considered to be the signature H3 modification marks deposited
transcriptome analysis and filtering out
et al. (2009). done according to Kundu and DasGupta (2018) and Sinharoy
GACCAAGTAAATTCTCCAC-3

AhDONE40

understanding the mode of action of
Plant material and rhizobial strain.
MATERIALS AND METHODS
of sense and antisense
AhDONE40

In summary, our investigation provides significant clues toward understanding the mode of action of ENOD40 during the initiation of cortical cell division by demonstrating the interaction of NAT-AhDONE40 with AhASHR3 and dynamic alterations of ENOD40 locus that may be required for cell fate reprogramming. Further experimentation is required for understanding the individual roles of sense and antisense ENOD40s and their mechanistic relationships. It remains to be seen whether NAT-DONE40 expression is universal in legumes and whether they function as the ENOD40–NAT-DONE40 pair for nodule organogenesis.

MATERIALS AND METHODS
Plant material and rhizobial strain.
A. hypogaea seed JL-24 were acquired from ICRISAT, Telangana, India. The seed were germinated and inoculated with SEMIA 6144 according to Karmakar et al. (2019) and Sinharoy

AhENOD40

The AhENOD40-RNAi construct was generated by amplification of ENOD40 from A. hypogaea root cDNA using primers 5'CACCATGAAGCTTGTTGGCC-3' and 5'GCAAGTT GACCAAGTAAATTCTCCAC-3'. Amplified fragments were cloned into pENTR/D-TOPOR (Life Technologies) and then into binary vectors pK7GW1WG2D(CC) (Karimi et al. 2002) by gateway technology through the LR clonase reaction kit (Life Technologies). Plant transformation and analysis was done according to Kundu and DasGupta (2018) and Sinharoy et al. (2009).

Transcriptome analysis and filtering out intergenic IncRNAs.
Sequence Read Archives of A. hypogaea from previously published strand-specific RNA transcriptome GEO data having accession ID GSE98997 have been analyzed (Karmakar et al. 2019). The data, corresponding to six time points in triplicate, were mapped to the A. hypogaea whole genome downloaded from PeanutBase. Alignment was performed with Hisat2 using the option “–rna-strandness R” (Pertea et al. 2016). Transcript assembly was performed with Cufflinks v2.1 (Trapnell et al. 2012) followed by Cuffcompare (Trapnell et al. 2012) using the A. hypogaea genome annotation file, and intergenic IncRNAs were filtered out. (i) Intergenic IncRNAs were initially filtered out using the –u class code option of Cuffcompare. In order to check whether the transcripts obtained actually lie within the intergenic region of A. hypogaea, intergenic positions were obtained by bedtools intersect (Quinlan and Hall 2010) using the A. hypogaea genome annotation file. (ii) At least three reads getting mapped to transcript had to be represented in read count. (iii) Transcript coverage of the mapped reads should be more than 70% (both read counts and read coverage are checked with bedtools coverage (Quinlan and Hall 2010). (iv) Transcripts should be at least 500 bp away from nearest coding gene (checked by in house perl script). Finally, the sequence was obtained for the filtered set of transcripts using bedtools getfasta (Quinlan and Hall 2010) and coding potential scores were calculated with the CPC2 tool (Kang et al. 2017). Transcripts whose CPC2 score indicated noncoding were chosen finally.

Sequence analyses.
Multiple sequence alignment was performed using MUSCLE with default parameters (Edgar 2004). The phylogenetic maximum-likelihood tree was generated using MEGA X (Kumar et al. 2018) with the bootstrap test of 1,000 replicates. The synteny analysis was done using PeanutBase. The distance and position of ENOD40 were manually analyzed and aligned.

Detection of AhENOD40 sense and antisense transcripts by RT-qPCR-based assay.
Total RNA was prepared using the Macherey Nagel Nucleo-Spin RNA plant kit according to the manufacturer’s protocol. For nuclear RNA isolation, the nucleus was first isolated using the Cell Lytic Plant Nuclei Isolation/Extraction Kit (Sigma), and RNA was isolated using Trizol/chloroform (Simms et al. 1993). The purified total RNA and nuclear RNA were reverse transcribed separately with the forward primer (ENOD40 F RT) designed to hybridize specifically with NAT-AhDONE40, yielding F-cDNA, and the reverse primer (ENOD40 R RT) that specifically hybridizes with both AhENOD40-1 and AhENOD40-2, yielding R-cDNA. Both RT reactions were performed at 50°C for 1 h using the Superscript III RT (Invitrogen) and, for each sample, a negative control reaction without Superscript III RT was also performed. During RT-qPCR analysis, to distinguish between chr5 and chr15, the 28-bp insertion in chr5 was used as a marker. Primers were designed for chr5 encompassing the 28-bp and were marked as L1F and L1R. Similarly, the flanking regions of the 28-bp insertion in chr5 was selected for chr15-specific primer L2R which, in the case of chr5, will not bind to it. To assess the presence of AhENOD40-1 and AhENOD40-2, the R-cDNA was amplified using primers ENOD40 R RT and L1R for chr5 and L2F for chr15. Similarly, to assess the presence of NAT-AhDONE40, the F-cDNA was amplified using primers ENOD40 F RT and L1F (Rosok and Sioud 2004). Quantification of the respective amounts of AhENOD40-1, AhENOD40-2, and NAT-AhDONE40 from the uninfected and the infected samples was done by the absolute quantification method using standard curves generated for the ampiclon corresponding to the region amplified by the respective primer pairs (Boulter et al. 2016).

In vitro transcription reaction.
The AhENOD40-1, AhENOD40-2, and NAT-AhDONE40 clones were used as the template for the PCR-based amplification of ENOD40 with the forward primer harboring the T7 polymerase binding site at the 5’ end for the generation of sense and antisense transcripts. PCR product (1 µg) was mixed with 10× transcription buffer, rNTP mix, radiolabeled or biotinylated UTP, RNase inhibitor (Promega Corp.), and T7 RNA polymerase (New England Biolabs) according to the supplier’s instructions, and the reaction mixture was incubated at 37°C for 1 h followed by DNaseI digestion for 15 min. The reaction mixture was separated in 5% native
polyacrylamide gel electrophoresis, and the ENOD40 RNA was extracted.

RNA pull-down assay.
Total cellular protein extracts from A. hypogaea root tissues were prepared by homogenization of the root tissues in binding buffer containing 50 mM Tris-HCl (pH 8.0), 10 mM EDTA (pH 8.0), 50 mM KCl, 150 mM NaCl, 1 mM phenylmethylsulfonyl fluoride, and 1% Nonidet P40. The homogenate was vortexed intermittently for 45 min and kept on ice between the pulses. The homogenate was centrifuged at 13,000 rpm for 30 min at 4°C, and the supernatant was collected. The supernatant (1 ml) representing 1 to 2 mg of the total protein extract was incubated overnight at 4°C with 50 pmol of in vitro transcribed biotinylated AhENOD40 or NAT-AhDONE40 RNA (encompassing ORF1 to ORF2) in the presence of 1 mM dithiothreitol and 5 mM MgCl2 on a rotation wheel at 10 rpm. Streptavidin-coated magnetic beads (Dynal) were washed according to the manufacturer's protocol, then equilibrated in binding buffer for 15 min. In all, 100 μl of the beads was added to the binding reaction mixture for 5 h to isolate the RNA protein complexes. Following incubation, the beads were washed twice with binding buffer and resuspended in 50 mM ammonium bicarbonate buffer; then, bead tryptic digestion, MALDI-MS/MS, and LC-MS/MS proceeded according to Fukuyama et al. (2012) and Kaiser et al. (2008). The generated peptide masses were searched against the UniProt protein sequence of AhASHR3 using the Progenesis QI search engine. The sequences of SET-domain-containing proteins were appended to the forward and reverse Arabidopsis thaliana UniProt database.

EMSA.
The DNA fragment corresponding to the coding sequences of the AWS, SET, and postSET domains of the A. hypogaea homolog of ASHR3 protein, AhASHR3 (residues 306 to 489), was cloned in pET32a, expressed in BL21 (DE3) (Thorsten et al. 2008). EMSA was done according to Roy et al. (2019) using 1.5 fmol of RNA and 0.5 and 3 μM recombinant AhASHR3 for binding and 3 μM for competition assays.

RIP.
The AhASHR3 overexpression construct was generated by amplification from A. hypogaea cDNA template with primers 5’-CACCATGGCCGATTTGGGGAATTTGTCTCTTTCCGAG-3’ and 5’-TCATACATGTTTCTATGGCAGGACATGCAGCT G-3’ complementary to the full length of AhASHR3. Amplified fragments were cloned into pENTR/D-TOPO (Life Technologies) and then into binary vectors pK7GWF2 (Karimi et al. 2002) by gateway technology through the LR Clonase reaction kit (Life Technologies). Constructs are then transformed into Agrobacterium rhizogenes strain R1000. The full-length AhASHR3 was overexpressed in A. hypogaea, and the transformed plants were inoculated with SEMIA 6144. At 4 DPI, the transformed roots were crosslinked by 1% formaldehyde under vacuum and were inoculated with SEMIA 6144. At 4 DPI, the transformed bacterium rhizogenes G-3 was inoculated with anti-GFP antibody (11814460001; Roche) overnight on a. Instead of GFP-Trap_MA beads from ChromoTek, we used a magnetic beads (Dynal) were washed according to the manufacturer’s protocol, then equilibrated in binding buffer for 15 min. In all, 100 μl of the beads was added to the binding reaction mixture for 5 h to isolate the RNA protein complexes. Following incubation, the beads were washed twice with binding buffer and resuspended in 50 mM ammonium bicarbonate buffer; then, bead tryptic digestion, MALDI-MS/MS, and LC-MS/MS proceeded according to Fukuyama et al. (2012) and Kaiser et al. (2008). The generated peptide masses were searched against the UniProt protein sequence of AhASHR3 using the Progenesis QI search engine. The sequences of SET-domain-containing proteins were appended to the forward and reverse Arabidopsis thaliana UniProt database.

EMSA.
The DNA fragment corresponding to the coding sequences of the AWS, SET, and postSET domains of the A. hypogaea homolog of ASHR3 protein, AhASHR3 (residues 306 to 489), was cloned in pET32a, expressed in BL21 (DE3) (Thorsten et al. 2008). EMSA was done according to Roy et al. (2019) using 1.5 fmol of RNA and 0.5 and 3 μM recombinant AhASHR3 for binding and 3 μM for competition assays.

RIP.
The AhASHR3 overexpression construct was generated by amplification from A. hypogaea cDNA template with primers 5’-CACCATGGCCGATTTGGGGAATTTGTCTCTTTCCGAG-3’ and 5’-TCATACATGTTTCTATGGCAGGACATGCAGCT G-3’ complementary to the full length of AhASHR3. Amplified fragments were cloned into pENTR/D-TOPO (Life Technologies) and then into binary vectors pK7GWF2 (Karimi et al. 2002) by gateway technology through the LR Clonase reaction kit (Life Technologies). Constructs are then transformed into Agrobacterium rhizogenes strain R1000. The full-length AhASHR3 was overexpressed in A. hypogaea, and the transformed plants were inoculated with SEMIA 6144. At 4 DPI, the transformed roots were crosslinked by 1% formaldehyde under vacuum and quenched by the addition of 0.125 M glycine. RIP was performed as described by Mermaz et al. (2018), with some modifications. Instead of GFP-Trap_MA beads from ChromoTek, we used Magna ChIP Protein A+G Magnetic Beads (16-663; Sigma-Aldrich, Merck). For precleaving, the protein sample was incubated with equilibrated beads, and the beads were collected using the magnet. The supernatant was collected and was allowed to mix with anti-GFP antibody (11814460001; Roche) overnight on a rotator at 4°C for the immunoprecipitation of GFP-tagged proteins. Equilibrated beads (20 μl) were added to the supernatant and mixed for 3 h at 4°C. The beads bound with antibody-AhASHR3-RNA were collected using a magnet and the process according to Mermaz et al. (2018). The immunoprecipitated RNA was reverse transcribed to cDNA and was quantified with RT-qPCR for detection of enrichment of RNAs. The enrichment of AhDONE40, AhENOD40-1, and AhENOD40-2 was normalized to the input using the standard %Input method.

Chromatin immunoprecipitation assays.
Chromatin immunoprecipitation (ChIP) assays were done according to Roy et al. (2014) with root tissue samples collected at three different stages of rhizobial infection (0, 4, and 8 DPI), using antibodies against unmodified histone H3 (number ab1791; Abcam, Cambridge), H3K27me3 (number ab6002; Abcam, Cambridge), H3K27ac (number ab4729; Abcam, Cambridge) H3K4me3 (number ab8580; Abcam, Cambridge), H3K9ac (number ab0812; Abcam, Cambridge), and H3K36me3 (number ab9050; Abcam, Cambridge). qPCRs were carried out using Luna Universal qPCR Master Mix (New England Biolabs) with primers specific for the conserved region of the ENOD40 locus from chr5 and 15. The AhActin gene was used for normalization in each case. Respective amounts of the immunoprecipitated DNA in uninfected and rhizobia infected samples were quantified according to Pavanagdakar et al. (2010), and the data were normalized as the levels of histone H3 modifications relative to unmodified H3 for each independent sample. Postnormalization, the data were represented as fold change over uninfected.

ACKNOWLEDGMENTS
We thank IISER Kolkata, India for access to sequencing facility; the Central Instrument Facility, Bose Institute, India for access to MALDI-TOF and LC-MS facilities; S. Sinharoy and S. Chaudhuri for critical reading of the manuscript; and H. Dey for excellent technical support.

AUTHOR-RECOMMENDED INTERNET RESOURCES
GEO data: https://www.ncbi.nlm.nih.gov/geo/
PeanutBase: https://www.peanutbase.org/

LITERATURE CITED


