Ploidy-dependent changes in the epigenome of symbiotic cells correlate with specific patterns of gene expression

Marianna Nagymihály[a,b], Alaguraj Veluchamy[c], Zoltán Györgypál[d], Federico Ariel[e], Teddy Jégu[f], Moussa Benhamed[f], Attila Szűcs[a], Attila Kereszt[a], Peter Mergaert[b,1], and Éva Kondorosi[a,b,1]

[a]Institute of Biochemistry, Biological Research Centre, Hungarian Academy of Sciences, H-6726 Szeged, Hungary; [b]Institute for Integrative Biology of the Cell, UMR9198, CNRS, Université Paris-Sud, Commissariat à l'Énergie Atomique et aux Énergies Alternatives, 91198 Gif-sur-Yvette, France; [c]Division of Biological and Environmental Sciences and Engineering, King Abdullah University of Science and Technology, Thuwal, 23955-6900, Kingdom of Saudi Arabia; [d]Institute of Biophysics, Biological Research Centre, Hungarian Academy of Sciences, H-6726 Szeged, Hungary; [e]Instituto de Agrobiotecnología del Litoral, Facultad de Bioquímica y Ciencias Biológicas, Universidad Nacional del Litoral, Consejo Nacional de Investigaciones Científicas y Técnicas, 3000 Santa Fe, Argentina; and [f]Institute of Plant Sciences Paris-Saclay, CNRS, French National Institute for Agricultural Research, Université Paris-Sud, Université Evry, Université Paris-Diderot, Sorbonne Paris-Cité, Universite Paris-Saclay, 91405, Orsay, France

Contributed by Éva Kondorosi, March 17, 2017 (sent for review November 9, 2016; reviewed by Christian Chevalier and Jens Stougaard)

The formation of symbiotic nodule cells in *Medicago truncatula* is driven by successive endoreduplication cycles and transcriptional reprogramming in different temporal waves including the activation of more than 600 cysteine-rich *NCR* genes expressed only in nodules. We show here that the transcriptional waves correlate with growing ploidy levels and have investigated how the epigenome changes during endoreduplication cycles. Differential DNA methylation was found in only a small subset of symbiotic nodule-specific genes, including more than half of the *NCR* genes, whereas in most genes DNA methylation was unaffected by the ploidy levels and was independent of the genes' active or repressed state. On the other hand, expression of nodule-specific genes correlated with ploidy-dependent opening of the chromatin as well as, in a subset of tested genes, with reduced H3K27me3 levels combined with enhanced H3K9ac levels. Our results suggest that endoreduplication-dependent epigenetic changes contribute to transcriptional reprogramming in the differentiation of symbiotic cells.

Medicago truncatula | symbiosis | nodule-specific cysteine-rich peptides | DNA methylation | chromatin structure

Symbiosis between *Medicago truncatula* and its *Sinorhizobium* bacterium partners results in the formation of nitrogen-fixing root nodules. In these nodules, a persistent apical meristem (ZI) constantly generates the various nodule cell types which in the central region lead to the formation of different nodule zones representing successive developmental stages and continuous aging of nodule cells (Fig. 1*A*). The cell layers below the meristem make up the invasion zone (ZII), where the youngest postmeristematic cells leaving the meristem with 4C genome content can be infected with rhizobia. These symbiotic cells enter a progressive differentiation program by duplicating their genome with repeated endoreduplication (ER) cycles associated with gradual cell enlargement and loss of cell division ability. The DNA content of symbiotic cells is 8C and 16C in ZII whereas it is 32C and 64C in the interzone (IZ) and in the nitrogen-fixing zone (ZIII) (Fig. 1 *A* and *B*) (1). Because not all nodule cells are infected, the symbiotic cells coexist with smaller, uninfected 4C and 8C cells (Fig. 1*A*) (2). Parallel to host cell maturation, the endosymbionts also undergo an irreversible terminal differentiation program (3, 4). During differentiation, the host cells dramatically change their gene-expression programs involving the activation of nodule-specific genes in successive spatial–temporal waves in clusters I to VIII (Fig. 1*C*) (5–7). Activation of the early genes in clusters I–IV occurs in the postmeristematic cells, whereas the genes in clusters V and VI are activated in the distal (ZIId) and proximal (ZIIp) parts of ZII, and the genes in clusters VII and VIII are induced during later stages of nodule development in the IZ and ZIII. Among the activated genes is the nodule-specific cysteine-rich (*NCR*) family, producing hundreds of different NCR peptides that guide terminal differentiation of the endosymbionts, converting them progressively to large polyploid, noncultivable nitrogen-fixing bacteroids. The *NCR* genes are under extremely tight transcriptional control because they are expressed exclusively in the symbiotic nodule cells, and their activation requires the intracellular presence of rhizobia (3, 6–8). Accompanying the maturation process of symbiotic cells, different subsets composed of tens or hundreds of *NCR* genes are expressed during the early, middle, and later stages (groups 1, 2, and 3, respectively, in Fig. 1 *D* and *E*). In situ expression of various *NCR* and nodule-specific genes suggested that their expression could be coupled to specific ploidy levels (6–8).

For nitrogen-fixing nodule development, the polyploidization of the symbiotic cells by repeated ER cycles is essential, because nodules in which ER is blocked are arrested in the nodule primordium stage and undergo early senescence (1). However, it was not known whether and how successive ER cycles and the polyploid state of the genome affect transcriptional reprogramming. The strict nodule specificity and the possible linkage between gene expression and ploidy levels of symbiotic cells suggested that epigenetic modifications along the ER cycles

Significance

Polyploidization of somatic cells is common in angiosperms. The characteristic and inherited developmental pattern of polyploidy in various organs and cell types suggests a role for endoreduplication in differentiation and specialized cell functioning. *Rhizobium*-infected nodule cells provide a unique opportunity to study specific developmental-stage changes in gene expression and in the epigenome during differentiation and at growing ploidy levels. We show that transcriptional reprogramming of hundreds of nodule-specific genes in successive waves correlates with the ploidy levels. Moreover, the dynamic changes observed in the epigenome suggest that chromatin accessibility together with histone tail modifications regulate the transcriptionally active or inert state of the genes.

Author contributions: M.N., P.M., and É.K. designed research; M.N. and T.J. performed research; M.N., A.V., Z.G., F.A., M.B., A.S., A.K., P.M., and É.K. analyzed data; and M.N., P.M., and É.K. wrote the paper.

Reviewers: C.C., French National Institute for Agricultural Research Bordeaux-Aquitaine Center; and J.S., Aarhus University.

The authors declare no conflict of interest.

[1]To whom correspondence may be addressed. Email: kondorosi.eva@brc.mta.hu or peter.mergaert@i2bc.paris-saclay.fr.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1704211114/-/DCSupplemental.

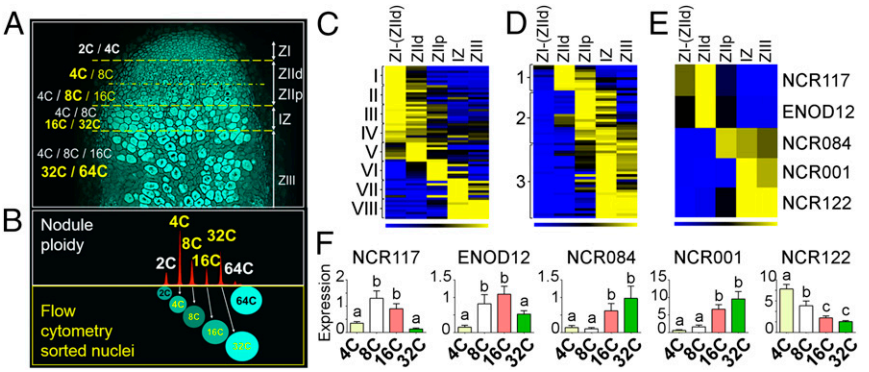

Fig. 1. Differential expression of *M. truncatula* genes during nodule development and at different ploidy levels of the symbiotic cells. (*A*) *M. truncatula* nodule zones ZI–ZIII and the genome copy number (C) of nodule cells in a longitudinal nodule section visualized by autofluorescence. ZI, meristem; ZII, invasion zone (ZIId, distal younger 4C/8C symbiotic cells; ZIIp, proximal older growing symbiotic cells 8C/16C); IZ, interzone large symbiotic cells (16C/32C); ZIII, nitrogen-fixing zone of huge symbiotic cells with large vacuoles (32C/64C) and uninfected cells (4C/8C) in the central region and nodule cortex. The predominant genome copy number of symbiotic cells is in yellow. (*B*) Nodule ploidy profile (*Upper*) and flow cytometry sorting of nuclei with different ploidy levels (*Lower*). Yellow indicates the ploidy of nuclei used for further experiments. (*C–E*) Hierarchical clustering analysis of *M. truncatula* gene sets in clusters I–VIII (*C*), in *NCR* genes, groups 1–3 (*D*), and in selected genes in nodule zones in the RNA-seq experiment (*E*) obtained by laser-capture microdissection [Data from Roux et al. (5)]. Blue-black-yellow colors correspond to low-to-high expression levels as shown in the scale bars. (*F*) Expression levels were determined by RT-qPCR of *NCR117*, *ENOD12*, *NCR084*, *NCR001*, and *NCR122* genes in nuclei of 4C, 8C, 16C, and 32C nodules. The expression levels are relative to the constitutive genes *Mtc27* and *40S*. Values are averages ± SEM from three independent experiments. Letters *a*, *b*, and *c* indicate significant differences (Student–Newman–Keuls test, $P < 0.05$).

might play a critical role in transcriptional regulation of *NCR*s and other nodule-specific genes (7). The recently reported reprogramming of DNA methylation during *M. truncatula* nodule development, regulated by the DNA demethylase gene *DEMETER* (*MtDME*), is in agreement with this hypothesis (9). Changes in 5-cytosine methylation of the DNA (5mC) and in histone tail modifications, alone or in combination, are important determinants of the chromatin structure and gene expression. The sequence context (CG, CHG, or CHH, in which H = A, C, or T) and the location of the DNA methylation in the promoter, gene body, or 3′ regulatory region can have different effects on gene transcription. In general, there is a strong correlation between DNA methylation in the promoter region and a decrease in gene expression (10–13). Among the histone tail modifications, trimethylation of H3 lysine 27 (H3K27me3) (14) also results in low expression level and high tissue-specificity (15), whereas acetylation of lysine 9 of histone H3 (H3K9ac) leads to gene activation (16–19). Here, using nodule nuclei sorted according to their DNA content, we studied genome-wide DNA methylation and chromatin accessibility as well as the presence of repressive H3K27me3 and activating H3K9ac histone tail modifications on selected genes. Patterns of chromatin changes then were matched to published in situ RNA-sequencing (RNA-seq) data of the nodule zones obtained by laser-capture microdissection (5). Our results collectively highlighted a correlation between ploidy level, gene expression, and chromatin structure.

Results

Differential Expression of Nodule-Specific Genes Depends on the Ploidy Levels of the Symbiotic Cells. To confirm the linkage between the expression of nodule-specific genes and ploidy levels, the expression level of selected, differentially expressed genes (Fig. 1*E*) was measured in isolated 4C, 8C, 16C, and 32C nuclei (Fig. 1*B*) with quantitative RT-PCR (RT-qPCR) (Fig. 1*F* and Table S1). The 2C and 64C nuclei were not included in this and our further analyses because they could not be isolated in sufficient amounts. *NCR117* and *ENOD12* (20) were the examples for early ZIId genes, and *NCR084* and *NCR001*, respectively, were the examples for the IZ and ZIII genes. *NCR122* was tested also because, unlike all other *NCR*s, this gene is expressed in the root and in uninfected nodule cells but not in infected ones (7). Both the early genes, *NCR117* and *ENOD12*, were expressed most strongly in the 8C and 16C nuclei, whereas the late genes *NCR084* and *NCR001* were expressed in the 16C and 32C cells, coinciding with their later function in symbiotic cell development. *NCR122* expression was maximal in the 4C cells and then decreased with the increasing ploidy levels, in agreement with the expression of this gene in the low-ploidy cortical cells and uninfected cells of the IZ and ZIII (7). However, the expression of *NCR001* and *NCR122* was significantly higher than that of the other tested genes. Overall, the expression of these nodule-specific genes according to the ploidy level correlated well with their spatial expression pattern in nodules identified previously with in situ hybridization, promoter-reporter gene fusions, and laser-capture microdissection coupled to RNA-seq (**7**, 20).

Only Specific Subsets of Genes Display Ploidy-Dependent Differential DNA Methylation. Because the majority of nodule-specific genes are expressed in the mature symbiotic cells (32C) and are repressed in the uninfected (4C) cells, we analyzed differences in DNA methylation (5mC) genome-wide in isolated 4C and 32C nuclei with reduced representation bisulphite sequencing (RRBS) (21). In total 461,403 methylated cytosines (mCs) were found in the 4C cells: 80% were in CG sequences, 15% were in CHG sequences, and 5% were in CHH sequences, whereas in 32C cells 71% of 519,975 mCs were in CG, 22% were in CHG, and 7% were CHH sequences (Fig. S1*A*). However, the number of detected mCs is an underestimation caused by the RRBS approach, the sequencing coverage, and the stringent parameters used for mC calling (*SI Materials and Methods*). The 71–80% mCs in the CG context was higher than reported in other plants, in which typically 40–55% mCs are in the CG context (13, 22, 23). The reason for this difference is unknown but may be related to the lower GC content of the *M. truncatula* genome (27%) compared with other plant species (32–46%) (24). The overall distribution of mCGs on the 1-kb upstream or downstream regions, exons, introns, intergenic regions, or transposons was similar in 4C and 32C nuclei (Fig. S1*B*).

By grouping mCs at neighboring positions, differentially methylated regions (DmRs) and within them DmR-associated genes (DmRGs) were identified in the 4C and 32C nuclei (25). Most DmRs (7,019; 79%) and DmRGs (4,295; 74%) were found in CG sequences, suggesting the relevance of mCGs in gene regulation (Fig. S1*A*). Of the 39,260 protein-coding genes in the *M. truncatula* genome, 11% (4,295) showed differential methylation, with 6% (2,488) being hypomethylated and 5% (1,807) being hypermethylated in the 32C cells (Fig. 2*A*). Differential methylation was observed on the entire coding region, including the 1-kb upstream and downstream sequences, with predominance in the gene body and primarily in the CG context (Fig. S1 *C* and *D*). Gene Ontology (GO) term distribution analysis of hypomethylated DmRGs revealed a strong overrepresentation of the "nodule morphogenesis" [GO ID: 0009878; P = 1.4E-31; false discovery rate (FDR) = 5.5E-29] and "metal ion binding" (GO ID: 0046872; P = 4.50E-07; FDR = 8.60E-05) GO categories (Table S2). In

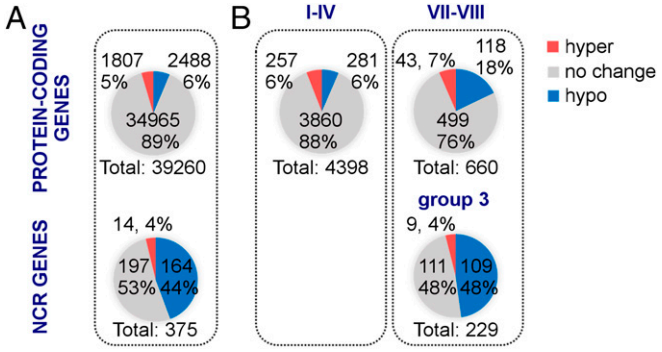

Fig. 2. Differential methylation of nodule-specific and nodule-expressed genes in 32C versus 4C nodule cells. Graphs show the percentage and number of protein-coding genes and *NCR* genes showing hyper- or hypomethylation genome-wide (*A*) and according to the expression profile (*B*). (*B*, *Left*) Genes in clusters I–IV preferentially expressed in 4C. (*Right*) Genes in clusters VII–VIII and the *NCR* genes in group 3 preferentially expressed in 32C. Hypermethylation is shown in red, no change is shown in gray, and hypomethylation is shown in blue.

these two categories of hypomethylated DmRGs, the *NCR*s and other nodule-specific genes accounted for 96.6 and 60.0% of genes, respectively. Of the 375 detected *NCR*s, 44% (164 *NCR*s) were hypomethylated, and only 4% (14 *NCR*s) were hypermethylated at 32C ploidy level (Fig. 2*A*). Differential methylation was predominant in their 1-kb upstream regions and in the CG context (Fig. S1*C*). To investigate the relationship between changes in CG methylation and differential gene expression, protein-coding genes in clusters I–IV that were expressed in 4C and down-regulated in 32C (Fig. 1*C*), genes in clusters VII and VIII that were up-regulated in 32C and were not expressed in 4C (Fig. 1*C*), and *NCR*s expressed in the 32C (group 3; Fig. 1*D*) were analyzed (Fig. 2*B*). In clusters I–IV, 6% (281) of the genes were hypomethylated; unexpectedly, only 6% (257) of the genes in 32C were hypermethylated, suggesting that DNA methylation plays a minor role in their repression. Moreover, these methylation changes were visible mostly in the gene body (Fig. S1*D*). In the case of the genes up-regulated in 32C, the fraction of hypomethylated DmRGs was significantly higher ($P < 0.0001$, Fisher's exact test): 18% (118) in clusters VII and VIII and 48% (109) in group 3 *NCR*s (Fig. 2*B*). These changes occurred mostly in the 1-kb promoter region (Fig. S1*D*). These results suggest that the decrease in DNA methylation might contribute to their activation or up-regulation in 32C.

To investigate if differential methylation has quantitative effects on gene expression, we compared the transcript abundance of hypermethylated and hypomethylated DmRGs using the published in situ RNA-seq data of the nodule zones (Fig. S2 and Dataset S1). We found that hypomethylation, hypermethylation, or no change in methylation affected the expression levels of the protein-coding and *NCR* genes differently (Fig. S2). The lower expression levels of protein-coding genes in 32C cells correlated with their hypermethylation or no change in their methylation status, whereas the transcript abundance of hypomethylated protein-coding genes was significantly higher in these cells. This finding is in agreement with the prevailing view that demethylation is a positive regulator of gene expression. However, such a correlation was not observed for the group 3 *NCR* genes, and, unexpectedly, the expression levels of the few hypermethylated *NCR*s were particularly high. Moreover, hypomethylated *NCR* genes exhibited expression levels similar to those of the unmethylated ones. Nevertheless, the expression of the group 3 *NCR* genes, independent of their methylation status, was much higher than the average expression of the cluster VII and VIII genes. Taken together, these data suggest that demethylation might be involved in the activation of a large subset of *NCR* genes but is not correlated with their extremely high expression.

Approximately 53% of the *NCR* genes did not differ in DNA methylation at 4C and 32C (Fig. 2*A*). These genes might be active in the 8C and 16C cells and display hypomethylation in these nuclei, or, alternatively, DNA methylation might not play a pivotal role in their regulation. In our selected gene set, *NCR117*, *ENOD12*, *NCR001*, and *NCR122* belonged to this category, whereas *NCR084* was in the hypomethylated DmRG category, allowing us to test these possibilities. By carrying out methylated DNA immunoprecipitation (MeDIP) coupled to qPCR (MeDIP-qPCR) in 4C, 8C, 16C, and 32C nuclei and measuring DNA methylation of the five selected genes along the regions 1 kb upstream and 500–600 bp downstream from their translational start site (TSS), we found strikingly different patterns and extents of DNA methylation (Fig. S3 and Table S3). Despite its high expression, the highest level of DNA methylation was observed in the *NCR001* gene between −1,000 and −600 bp, peaking at −800 bp, decreasing at −400 bp and even further at −200 bp, and not detectable in the coding sequence. This profile was unaffected by the ploidy levels between −1,000 and −800 bp, and only a slight decrease in methylation was observed at −600 bp in the higher ploidy nuclei as compared with 4C nuclei. *NCR117* had a lower level of methylation between −600 bp and −200 bp with the highest level at −400 bp at all ploidy levels. A significant decrease was observed in 8C and 16C at −200 bp. *ENOD12* was methylated at low levels in the gene body but not in the 1-kb upstream region. Neither the upstream region nor gene body was methylated in *NCR122*. The strongly methylated gene body of *NCR084* also was unaffected by the ploidy levels, unlike its 1,000-bp upstream region where DNA methylation decreased significantly with the growing ploidy levels in line with the bisulphite sequencing analysis and its expression in 16C and 32C cells.

Our findings show that clusters VII and VIII and *NCR* genes are overrepresented in hypomethylated regions, suggesting that a decrease in DNA methylation, particularly in the 1-kb upstream region, contributes to their expression, whereas other genes with unaltered DNA methylation in these groups might require chromatin modifications.

Chromatin Accessibility Varies with Ploidy and Correlates with Gene Expression. To know how chromatin compaction and accessibility change during the course of nodule differentiation and how this change influences gene expression, we studied the chromatin accessibility in 4C, 8C, 16C, and 32C nuclei genome-wide using the assay for transposase-accessible chromatin sequencing

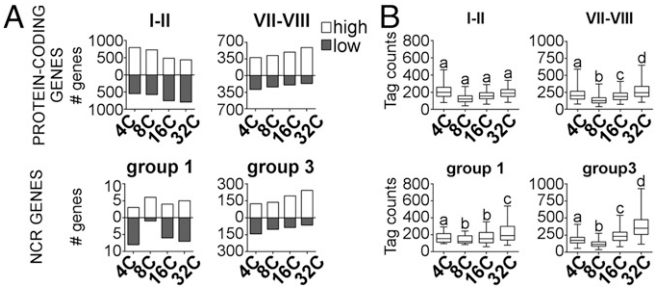

Fig. 3. Differential chromatin accessibility of nodule-specific and nodule-expressed genes in 4C, 8C, 16C, and 32C cells. (*A*) Number of early (I–II and group 1) and late (VII–VIII and group 3) protein-coding and *NCR* genes showing high (white) or low (dark gray) chromatin accessibility in 4C, 8C, 16C, and 32C nuclei. The number of genes is shown on the *y* axis. (*B*) ATAC-seq tag counts in the 1-kb upstream region of protein-coding and *NCR* genes from early (I–II and group 1) and late (VII–VIII and group 3) gene clusters in 4C, 8C, 16C, and 32C nuclei. Box plots show the upper, and lower quartiles; the median is shown by a horizontal line; and whiskers show the 2.5 and 97.5 percentiles of the accessibility values. Different letters indicates significant differences (Wilcoxon's rank sum test, $P < 0.05$).

(ATAC-seq) (Fig. 3) (26). At the whole-genome scale the chromatin accessibility showed considerable differences with growing ploidy levels. First, the early and late protein-coding and *NCR* genes were classified according to the accessibility of their chromatin (Fig. 3*A*). We found that the number of genes with high or low compactness in their 1-kb upstream region correlated well with the up- or down-regulation of genes in the respective gene clusters. The number of cluster I and II genes with high chromatin accessibility was largest in 4C and 8C and was lower in 16C and 32C, whereas it increased with growing ploidy levels in clusters VII and VIII. The number of early group 1 *NCR* genes with high chromatin accessibility was highest in the 8C nuclei, whereas in group 3 of the late *NCR* genes chromatin was highly accessible in 16C and even more in 32C. The distribution of the ATAC-seq tag counts displayed only small variations at the four ploidy levels in the genes in clusters I and II, whereas the genes in clusters VII and VIII had significantly increased numbers of reads in the 1-kb upstream region in 32C nuclei (Fig. 3*B*). For the *NCR*s in group 3, the read counts increased in 16C and were even more pronounced in the 32C nuclei (Fig. 3*B*).

To evaluate the correlation between expression and accessibility, Pearson correlation coefficients (PCC) were determined between the reads per kilobase of transcript per million reads mapped (RPKM) values and ATAC-seq tag counts for protein-coding genes and *NCR* genes belonging to different expression groups from 4C to 32C ploidy levels. No correlation was found (PCC = 0.28, SEM = 0.01, SD = 0.51) in early-expressed protein-coding genes (clusters I and II), but the correlation was evident in case of the late-expressed genes in clusters VII and VIII (PCC = 0.52, SEM = 0.02, SD = 0.44) and in group 3 *NCR* genes (PCC = 0.84, SEM = 0.01, SD = 0.20).

We also assessed the correlation between the methylation status of genes and chromatin accessibility. Genes of clusters VII and VIII and *NCR* genes of group 3 showed an increase in chromatin accessibility in 32C compared with 4C, independently of their hypermethylated, hypomethylated, or unchanged methylation state (Fig. S4), indicating that chromatin opening also can occur in the absence of demethylation of the genes.

The local chromatin state of the five selected genes is shown in Fig. 4. Based on the ATAC-seq tag frequency, the chromatin shows relatively low accessibility in nuclei in which the corresponding gene is not expressed but opens strongly in nuclei in which the gene is transcriptionally active. In the *NCR117* region the chromatin was relatively most open in 8C, in line with its early expression. *ENOD12* is present in a rather open chromatin region where the accessibility was the highest in the 8C and 16C cells. For *NCR084* and *NCR001* the chromatin became highly open in 16C and 32C, but in *NCR122* it was most accessible in 4C and 8C. The genome viewer at the five selected loci revealed that changes in the local chromatin state are not limited to the promoter regions but can extend over the neighboring regions. Thus, taken together, the differential accessibility of chromatin regions at various ploidy levels was in line with the activity of the genes and with the expression of several neighboring genes in the same chromatin microenvironment.

Ploidy-Dependent Variation in H3K27me3 and H3K9ac Chromatin Marks Correlates with the Expression Pattern of *NCR* Genes. To analyze the correlation between chromatin modifications and gene expression, we used ChIP-qPCR to investigate how the antagonistic chromatin marks H3K27me3 and H3K9ac correlate with the activation of weakly (*NCR117*, *ENOD12*, and *NCR084*) and strongly (*NCR001* and *NCR122*) expressed genes in different ploidy nuclei (Fig. 5 and Table S3). On the *NCR117* locus the level of H3K27me3 was relatively high in 4C and 32C nuclei, where this gene is repressed, but was very low, especially close to the TSS, in 8C and was low in 16C, where the gene is active ($P < 0.05$). H3K9ac was observed in the 4C, 16C, and 32C nuclei. In the case of *ENOD12*, *NCR084*, and *NCR001*, the H3K27me3 level was high in the 4C cells and decreased throughout the entire region in the 8C cells, decreased even further in the 16C cells, and increased again in the 32C cells. Similarly to *NCR117*, H3K9ac modifications were observed at all three loci in 4C, 16C, and 32C nuclei, but 8C was devoid of H3K9ac modifications. The extent of H3K9ac modifications in the gene body of *NCR084* and *NCR001* was particularly important in 16C and 32C nuclei. In the case of *NCR122*, a low level of H3K27me3 together with high H3K9ac marks in the 4C and 8C nuclei coincided well with gene activity, whereas decreasing H3K9ac and increasing H3K27me3 correlated with gradual repression of *NCR122* in 16C and 32C nuclei.

Because both H3K27me3 and H3K9ac contribute to gene regulation, we calculated the ratio of H3K27me3 and H3K9ac for these five genes upstream and downstream from the TSS (Fig. 5*B*). In *NCR117*, the presence of H3K27me3 in the −400–TSS region in 4C cells could be responsible for gene silencing. In *NCR084* and *NCR001*, the high H3K27me3/H3K9ac ratio in 4C and 8C cells coincides with their repression, and lower values correlate with their activation. The same tendency, although less pronounced, also was valid for *ENOD12*. In contrast, for *NCR122* the H3K27me3/H3K9ac ratio was low in 4C and 8C cells, in line with high gene expression, and decreased with an increased H3K27me3/H3K9ac ratio in 16C and 32C cells.

In general, modification of H3K27me3 showed dynamic changes in both the promoter and the gene body as a function of ploidy levels, whereas changes in H3K9ac were most important in the gene body. The repressed state of the genes coincided with a high level of H3K27me3 and chromatin compaction, whereas gene activation coincided with a reduced level of H3K27me3 and an increased H3K9ac/H3K27me3 ratio. The increased H3K9ac levels correlated with higher gene expression and an open chromatin conformation. Collectively these data suggest that the activities of the *NCR*s are dynamically regulated by their chromatin modifications.

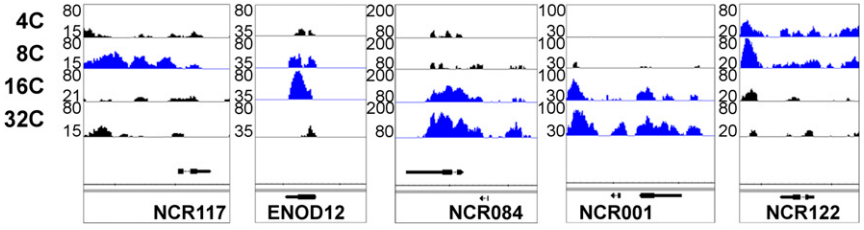

Fig. 4. ATAC-seq tag frequency profile along the *NCR117*, *ENOD12*, *NCR084*, *NCR001*, and *NCR122* genes at different ploidy levels. Black indicates the repressed state, and blue represents the active state of the genes.

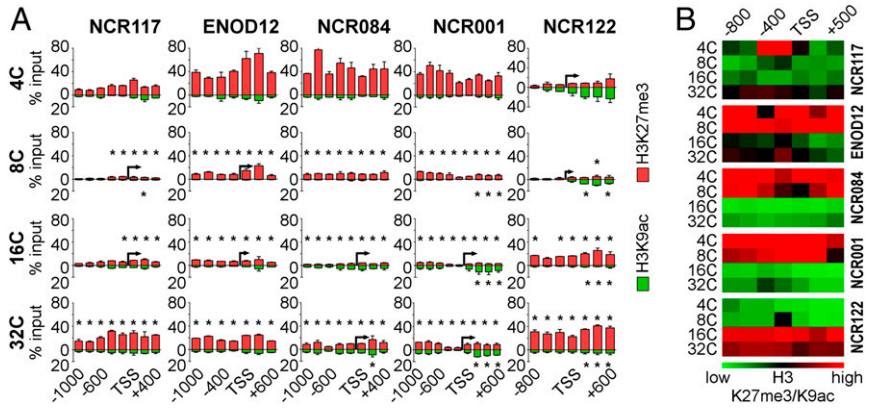

Fig. 5. H3K27me3 and H3K9ac modifications of the *NCR117*, *ENOD12*, *NCR084*, *NCR001*, and *NCR122* genes at different ploidy levels (4C–32C) detected with ChIP-qPCR. (*A*) H3K27me3 (red) and H3K9ac (green) levels are shown as percentages of input chromatin. The *x* axis represents the distance from the TSS. Each bar represents a given primer pair used for this analysis. Black arrows at the TSS indicate ploidy levels at which the gene is active. The percent of input values represents the average ± SEM of three independent experiments. Asterisks indicate a significant difference compared with 4C: $*P < 0.05$, two-way ANOVA with Tukey's multiple comparison test. (*B*) The ratio of high (repression) to low (activation) H3K27me3 and H3K9ac levels for the five selected loci (at the indicated gene positions relative to the TSS) is indicated by the red-to-green scale.

Discussion

Symbiotic nodule cells represent a unique model for studying epigenetic changes in respect to growing ploidy and transcriptional reprogramming. Our data provide insight into the dynamic changes of epigenetic modifications during the development of this specific cell type. By identifying ploidy-specific DNA methylation, repressive H3K27me3 and activating H3K9ac histone marks, and the chromatin accessibility profiles of differentially expressed genes, we get a view of the multilayered epigenetic control of symbiotic cell differentiation. Moreover, our findings provide a step forward in the understanding of the extremely tight regulation of *NCR* genes.

Differential DNA Methylation Is Rare but Predominant in the *NCR* Gene Family. Nodules derive from cortical cells in the root elongation–differentiation zone. Induced by the *Rhizobium* signal molecules—the Nod factors—cortical cells dedifferentiate and enter the mitotic cycle. Proliferation of these cells leads to the establishment of the meristem. DNA methylation of the meristematic cells represents the initial and likely the most methylated status of DNA. Cells leaving the meristem and infected with rhizobia undergo consecutive waves of transcriptional reprogramming, and here we show that differential gene expression is coupled to given ploidy levels. Methylation of promoters usually provokes gene silencing, whereas demethylation was expected to play a major role in gene activation. However, only a subset of early and late genes exhibited changes in methylation in 4C and 32C cells, indicating that the DNA methylation state of the 4C cells has been copied during the repeated ER cycles. *NCR*s behaved differently from most protein-coding genes, because 160 of 295 genes became hypomethylated. This finding is in line with the recently reported down-regulation of the *MtMET* methyltransferase gene (Medtr6g065580) and up-regulation of the nodule-specific *MtDME* DNA demethylase in the interzone cells and with down-regulation of *NCR* genes in the MtDME RNAi lines (9). The differential methylation patterns that we observed could be, in large part, the consequence of the action of MtDME. Why the activation of certain *NCR* genes, but not others, is associated with demethylation requires further investigation. An attractive reason why differential DNA methylation is important for the expression of a subset of *NCR* genes might be related to the presence of transposable elements (TEs) in the vicinity of many *NCR* genes (9). TEs in plant genomes usually are maintained transcriptionally silent. This silencing of TEs also may affect neighboring genes, and *NCR* gene activation thus may require demethylation mediated by MtDME. Accordingly, many transposon genes are activated concomitantly with the *NCR*s during symbiotic cell differentiation (9).

Changes in Chromatin Accessibility with the Ploidy Levels. Growing evidence suggests that changes in DNA methylation can influence other epigenetic modifications and can directly affect the chromatin structure (27, 28). We found that dynamic opening and closing of the chromatin at different ploidy levels of the cells correlated well with the active/repressed expression state of the genes. Chromatin reorganization was particularly important close to or at the TSS. However, open chromatin alone was not sufficient for gene activation. The change in chromatin accessibility was not tightly linked to DNA methylation, because it also occurred when the DNA methylation status was not affected by the ploidy. Moreover, the degree of chromatin opening did not always show a direct correlation with the level of DNA methylation.

Histone Modifications: The Level of H3K27me3 Is Crucial for Gene Expression and Is Epistatic over H3K9ac. We analyzed the repressive H3K27me3 and the activating H3K9ac profile of selected symbiosis-specific genes from different spatial and temporal expression classes in growing ploidy nuclei. The H3K27me3 marks were predominant in the 4C cells and are likely essential for the repressed state of *NCR*s and other symbiotic genes. The H3K27me3 marks were dramatically reduced in the 8C and 16C cells and increased again in the 32C cells. Such a reduction of the H3K27me3 marks could be sufficient for activation of the *NCR117* and *ENOD12* genes, which are already activated in 8C cells. For the *NCR084* and *NCR001* genes, which are activated from 16C on, reduced H3K27me3 marks together with the presence of the active H3K9ac marks and accessible chromatin structure contributed to gene activation. *NCR*s usually are expressed at very high levels in 32C cells that, in addition to reduced H3K27me3, have also the active H3K9ac marks in the gene body and the promoter and an open chromatin structure. On the other hand, in the case of early genes, the increase of H3K27me3 marks correlated with their repression in the 32C cells. The primary role of the H3K27me3 mark in the regulation of these genes is in line with previous observations that the deposition and dynamic regulation of the H3K27me3 mark are important in controlling tissue-specific gene expression and plant differentiation (29).

Epigenome Reconfiguration in ER-Driven Cell Differentiation. Our results are consistent with studies of somatic cell reprogramming and cellular differentiation in plants and animals, showing that gene expression is controlled by multiple regulatory processes from nucleosome positioning through histone modification and DNA methylation (27, 28, 30). We propose that the dynamic changes in the chromatin environment that we describe here determine the cell type-specific transcriptional program by turning on or off the temporally and spatially expressed nodule-specific genes and therefore contribute to symbiotic cell differentiation. Importantly, this remodeling is associated with the ER process that drives this differentiation. Cellular differentiation and morphogenesis through ER cycles is widespread in plants (31–33). Well-studied examples are the seed endosperm, the root hairs, and trichomes or the root cortical cells, all of which are formed through ER (34–37). Also the establishment of many plant-biotrophic interactions, and notably the formation of the plant cells that are at the interface with the microbial partner, require ER. These interactions include the nodules discussed here and also mycorrhiza and pathogenic interactions with powdery mildews and nematodes (31–33). Epigenetic processes similar to the ones we described here during the differentiation of nodule symbiotic cells also could be important in these other ER-mediated differentiation programs. On the other hand, the formation of symbiotic nodule cells in *M. truncatula* involves the activation of a tremendous and extremely specific transcriptional program involving genes exclusively expressed in these cells (5–7). It is unclear if other cellular differentiation processes involve a similarly profound transcriptional reprogramming and if they therefore require the same strict regulatory controls as the symbiotic nodule cells.

Materials and Methods

Plant Material, Bacterial Strain, and Growth Conditions. *M. truncatula* A17 seedlings were inoculated with *Sinorhizobium medicae* strain WSM419, and 28 d post inoculation nodules from three independent biological replicates were collected in dry ice and kept at −80 °C. For ChIP experiments nodules were collected on ice in 1% formaldehyde, and cross-linking was performed under vacuum for 15 min. Formaldehyde was quenched with 125 mM glycine for 5 min under vacuum, and nodules were washed twice with milliQ water, dried, and stored at −80 °C until further use.

Purification of Nuclei and Downstream Analyses. Nodules were chopped with a razor blade, and nuclei were obtained by a Beckman MoFlow Astrios (Beckman Coulter) cell sorter. Comparable nuclear and total cellular mRNA pools (21, 22) made possible the gene expression analysis in purified nuclei that was performed with RT-qPCR on total RNA purified from nuclei with Tri-Reagent (Sigma). Methylation analysis was performed with RRBS and MeDIP on nuclear DNA purified with the DNeasy Blood and Tissue kit (Qiagen). For ChIP, formaldehyde-fixed nuclei were used. For ATAC-seq, unfixed nuclei were treated as described (28). Details on all procedures are provided in *SI Materials and Methods*.

ACKNOWLEDGMENTS. We thank Célia Casset for help with the qPCR experiments, Dr. Mickael Bourge for the flow cytometry sorting of nodule nuclei, and Dr. István Nagy for providing the MBD2 protein for the MeDIP experiment. This work benefitted from the facilities and expertise of the Imagif Cell Biology Unit of the Gif campus (which is supported by the Infrastructures en Biologie Santé et Agronomie), the Agence Nationale de la Recherche (ANR) under Investments for the Future programs France-BioImaging Infrastructure Grant ANR-10-INSB-04-01 and Saclay Plant Sciences Grant ANR-10-LABX-0040-SPS, and the Conseil Général de l'Essonne. This work was supported by ANR Grant ANR-13-BSV7-0013 and by European Research Council SYM-BIOTICS Advanced Grant 269067 (to É.K.). M.N. was the recipient of a PhD grant from CAMPUS France and was supported by the SYM-BIOTICS grant.

1. Vinardell JM, et al. (2003) Endoreduplication mediated by the anaphase-promoting complex activator CCS52A is required for symbiotic cell differentiation in Medicago truncatula nodules. *Plant Cell* 15:2093–2105.
2. Xiao TT, et al. (2014) Fate map of Medicago truncatula root nodules. *Development* 141:3517–3528.
3. Van de Velde W, et al. (2010) Plant peptides govern terminal differentiation of bacteria in symbiosis. *Science* 327:1122–1126.
4. Mergaert P, et al. (2006) Eukaryotic control on bacterial cell cycle and differentiation in the Rhizobium-legume symbiosis. *Proc Natl Acad Sci USA* 103:5230–5235.
5. Roux B, et al. (2014) An integrated analysis of plant and bacterial gene expression in symbiotic root nodules using laser-capture microdissection coupled to RNA sequencing. *Plant J* 77:817–837.
6. Maunoury N, et al. (2010) Differentiation of symbiotic cells and endosymbionts in Medicago truncatula nodulation are coupled to two transcriptome-switches. *PLoS One* 5:e9519.
7. Guefrachi I, et al. (2014) Extreme specificity of NCR gene expression in Medicago truncatula. *BMC Genomics* 15:712.
8. Mergaert P, et al. (2003) A novel family in Medicago truncatula consisting of more than 300 nodule-specific genes coding for small, secreted polypeptides with conserved cysteine motifs. *Plant Physiol* 132:161–173.
9. Satgé C, et al. (2016) Reprogramming of DNA methylation is critical for nodule development in Medicago truncatula. *Nat Plants* 2:16166.
10. Zhang X, et al. (2006) Genome-wide high-resolution mapping and functional analysis of DNA methylation in arabidopsis. *Cell* 126:1189–1201.
11. Zilberman D, Gehring M, Tran RK, Ballinger T, Henikoff S (2007) Genome-wide analysis of Arabidopsis thaliana DNA methylation uncovers an interdependence between methylation and transcription. *Nat Genet* 39:61–69.
12. Cokus SJ, et al. (2008) Shotgun bisulphite sequencing of the Arabidopsis genome reveals DNA methylation patterning. *Nature* 452:215–219.
13. Garg R, Narayana Chevala V, Shankar R, Jain M (2015) Divergent DNA methylation patterns associated with gene expression in rice cultivars with contrasting drought and salinity stress response. *Sci Rep* 5:14922.
14. Feng S, Jacobsen SE (2011) Epigenetic modifications in plants: An evolutionary perspective. *Curr Opin Plant Biol* 14:179–186.
15. Zhang X, et al. (2007) Whole-genome analysis of histone H3 lysine 27 trimethylation in Arabidopsis. *PLoS Biol* 5:e129.
16. Kurdistani SK, Tavazoie S, Grunstein M (2004) Mapping global histone acetylation patterns to gene expression. *Cell* 117:721–733.
17. Schübeler D, et al. (2004) The histone modification pattern of active genes revealed through genome-wide chromatin analysis of a higher eukaryote. *Genes Dev* 18: 1263–1271.
18. Roh TY, Cuddapah S, Zhao K (2005) Active chromatin domains are defined by acetylation islands revealed by genome-wide mapping. *Genes Dev* 19:542–552.
19. Zhou J, et al. (2010) Genome-wide profiling of histone H3 lysine 9 acetylation and dimethylation in Arabidopsis reveals correlation between multiple histone marks and gene expression. *Plant Mol Biol* 72:585–595.
20. Pichon M, et al. (1992) Rhizobium meliloti elicits transient expression of the early nodulin gene ENOD12 in the differentiating root epidermis of transgenic alfalfa. *Plant Cell* 4:1199–1211.
21. Meissner A, et al. (2005) Reduced representation bisulfite sequencing for comparative high-resolution DNA methylation analysis. *Nucleic Acids Res* 33:5868–5877.
22. Shangguan L, et al. (2013) Evaluation of genome sequencing quality in selected plant species using expressed sequence tags. *PLoS One* 8:e69890.
23. Song QX, et al. (2013) Genome-wide analysis of DNA methylation in soybean. *Mol Plant* 6:1961–1974.
24. Lister R, et al. (2008) Highly integrated single-base resolution maps of the epigenome in Arabidopsis. *Cell* 133:523–536.
25. Saito Y, Tsuji J, Mituyama T (2014) Bisulfighter: Accurate detection of methylated cytosines and differentially methylated regions. *Nucleic Acids Res* 42:e45.
26. Buenrostro JD, Giresi PG, Zaba LC, Chang HY, Greenleaf WJ (2013) Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat Methods* 10:1213–1218.
27. Hussein SM, et al. (2014) Genome-wide characterization of the routes to pluripotency. *Nature* 516:198–206.
28. Lee DS, et al. (2014) An epigenomic roadmap to induced pluripotency reveals DNA methylation as a reprogramming modulator. *Nat Commun* 5:5619.
29. Lafos M, et al. (2011) Dynamic regulation of H3K27 trimethylation during Arabidopsis differentiation. *PLoS Genet* 7:e1002040.
30. Huang K, et al. (2015) Dynamically reorganized chromatin is the key for the reprogramming of somatic cells to pluripotent cells. *Sci Rep* 5:17691.
31. Kondorosi E, Roudier F, Gendreau E (2000) Plant cell-size control: Growing by ploidy? *Curr Opin Plant Biol* 3:488–492.
32. Larkins BA, et al. (2001) Investigating the hows and whys of DNA endoreduplication. *J Exp Bot* 52:183–192.
33. Bramsiepe J, et al. (2010) Endoreplication controls cell fate maintenance. *PLoS Genet* 6:e1000996.
34. Vanstraelen M, et al. (2009) APC/C-CCS52A complexes control meristem maintenance in the Arabidopsis root. *Proc Natl Acad Sci USA* 106:11806–11811.
35. Sugimoto-Shirasu K, et al. (2005) RHL1 is an essential component of the plant DNA topoisomerase VI complex and is required for ploidy-dependent cell growth. *Proc Natl Acad Sci USA* 102:18736–18741.
36. Bourdon M, et al. (2012) Evidence for karyoplasmic homeostasis during endoreduplication and a ploidy-dependent increase in gene transcription during tomato fruit growth. *Development* 139:3817–3826.
37. Wildermuth MC (2010) Modulation of host nuclear ploidy: A common plant biotroph mechanism. *Curr Opin Plant Biol* 13:449–458.