様式第１号

論文内容要旨

Temporal Changes in Transcripts of Miniature Inverted-Repeat Transposable Elements during Rice Endosperm Development

イネ胚乳発生時における MITE の転写産物の時系列変化

学位申請者氏名：永田 博基
研究指導教員：木下 哲

Transposable elements (TEs) are repressed by the host epigenetic machinery because of their harmful features capable of transposing from their genomic loci to another. DNA methylation is one of the repressive marks for TEs but its state changes together with the development in plants. Especially, the level of DNA methylation on TEs is reduced in the endosperm tissue of which provides the nutrient to the embryo in angiosperm. Although this suggests that TEs are epigenetically active in the endosperm, their transcription activity and a biological role in the activation are unknown. Here, to obtain molecular clues of TE expression and biological roles of the transcription of TEs in the endosperm development, I performed time-course transcriptome analyses for TEs using the developing endosperm of rice, of which genome is smaller than the other crops but contains TEs in various length and types.

In Chapter 1, I analyzed the temporal change of expression of Miniature Inverted-repeat TEs (MITEs), which is the most representative TEs with short length in rice genome, during the endosperm development using the custom-made microarray rich in rice TE probes. I performed k-means clustering of the temporal expressions of MITEs, and then identified that their expression patterns were dynamically up- and downregulated during the endosperm development. Interestingly, some MITE transcripts were directionally controlled in the endosperms, like protein-coding genes (Fig. 1a). These results suggest that TE transcription is regulated under the temporal or the developmental programs. Furthermore, I discovered the NF-Y binding motif, CCAAT, in the region near the 5' terminal inverted repeat of *Youren* subfamily; identified that the presence of CCAAT in *Yourens* was associated with their expression levels (Fig. 1b,c). This suggests that the directional expression of *Yourens* might be *cis*-regulated by the internal regulatory elements of *Youren* sequences. My results uncover the dynamic change in the transcriptional activity of MITEs and their potential regulatory mechanism during rice endosperm development.

In Chapter 2, I performed a time-course RNA-seq analysis, and then clustering analysis of the expression patterns of differentially expressed (DE) TEs to identify active transcription of TEs during the endosperm development. We found MERMITE18B, which is one of the subfamilies of MULE/MuDR TE superfamily, was highly enriched in limited clusters with the expression peak at the later endosperm development (Fig 2a). As results of epigenomic analyses, DE MERMITE18Bs are DNA-hypomethylated in the endosperm and were accumulated by small interfering RNAs (siRNAs), especially 21-nt in length, compared to non-differentially expressed them (nonDEs) (Fig. 2b-d). My results suggest that the increased transcription level of MERMITE18Bs through DNA hypomethylation results in the production of 21-nt siRNAs. To identify the role of this siRNA biogenesis, we searched target genes of the 21-nt siRNAs in the

rice genome. *Vacuolar Acid Invertase 1* (*VIN1*), a gene encoding sucrose hydrolase, had a putative binding site for a DE MERMITE18B-derived 21-nt siRNA. *VIN1* was predominantly expressed in the early coenocytic endosperm, and then diminished, displaying a negative correlation to the expression of DE MERMITE18B. The accumulation of secondary siRNAs and CHH methylation was also observed in the neighboring region targeted by the siRNA in the endosperm, suggesting that *VIN1* might be a primary candidate of post-transcriptional silencing by DE MERMITE18B-derived 21nt siRNA (Fig. 3). My time-course RNA-seq analysis and multi-omics approaches identified highly expressed MERMITE18B and suppression of *VIN1* potentially via siRNA production from MERMITE18B during rice endosperm development.

In summary, my study supports that the transcription of TEs are orderly regulated like protein-coding genes and might be embedded as host mechanism for controlling the gene expression during rice endosperm development.

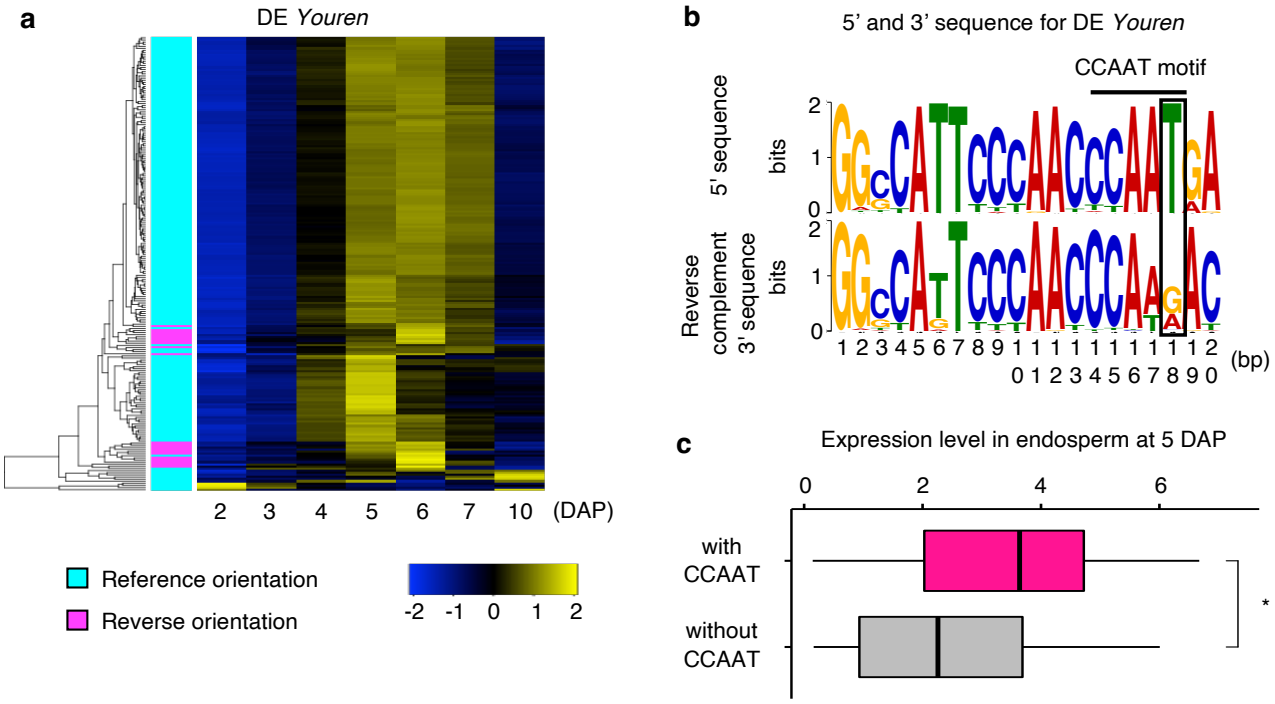

Fig. 1 Enrichment of CCAAT motif at the 5' region of in DE *Yourens* that showed unique expression pattern and direction among their probes. **a** Heatmap representation of the relationship between expression patterns and transcriptional direction of DE *Yourens* (n=214). Color key indicates row Z-score of $\log_2$-transformed normalized signal intensity. Cyan, *Youren* in reference orientation; magenta, *Youren* in reverse orientation. **b** Sequence logos for the 5′ (upper) and 3′ reverse complementary (bottom) sequences that include the terminal inverted repeats (TIRs) and the inner regions of DE *Youren* (n=198), drawn by MEME suite. The

horizontal line above the logos indicates the conserved CCAAT motif, known as the NF-Y binding site. The box indicates the last T of the CCAAT motif, which is not conserved in the 3′ sequences of DE *Youren*. **c** Box plots of the expression level of *Yourens* with (magenta, n=165) and without (gray, n=33) the CCAAT motif in endosperm at 5 DAP. **P* < 0.05 using a Wilcoxon rank sum test.

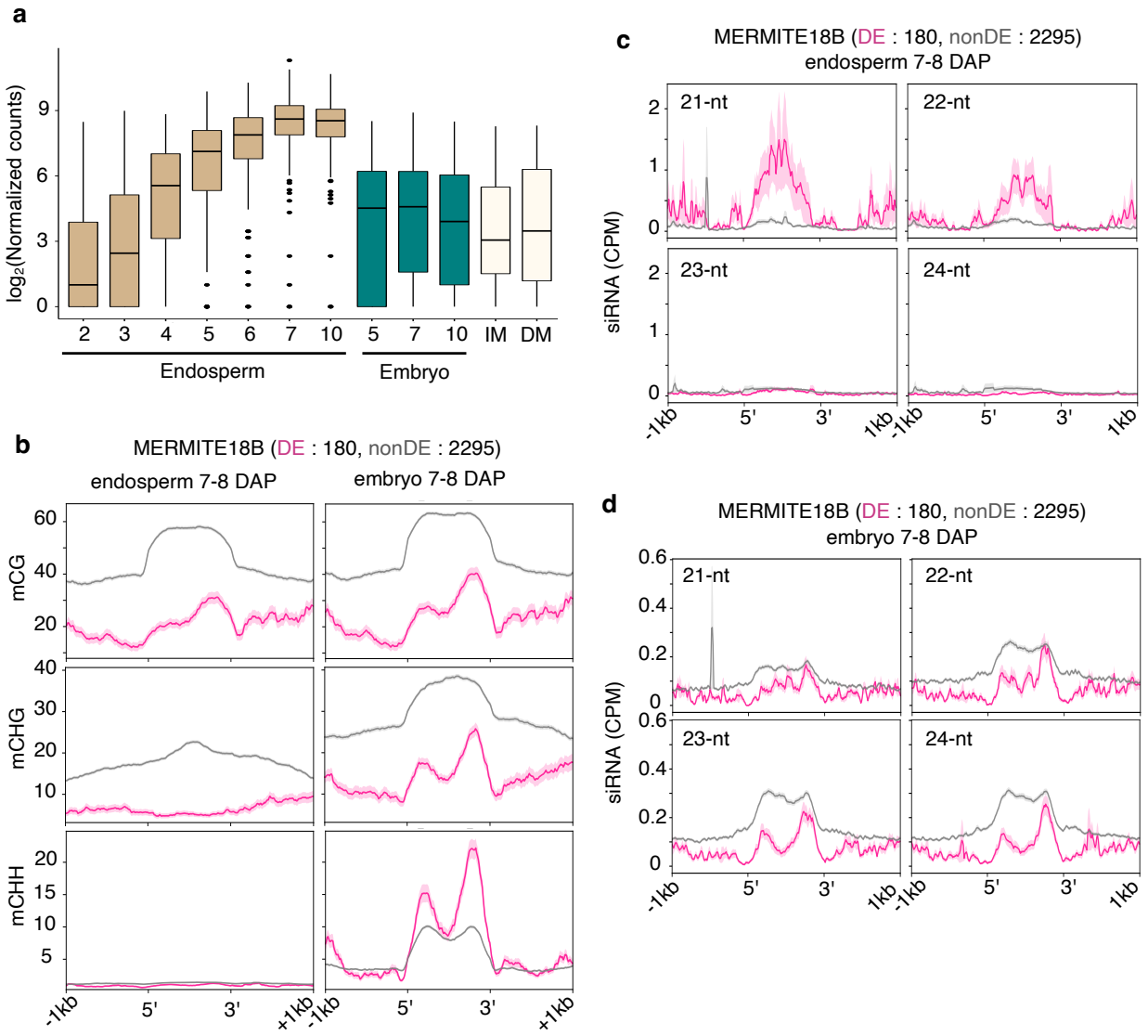

Fig. 2 The levels of transcription, DNA methylation and siRNAs of MERMITE18B in rice endosperm. **a** Transcription levels of DE MERMITE18B in various rice tissues. “IM” and “DM” indicate inflorescences at the indeterminant and determinant meristem stages. **b** DNA methylation pattern of DE (dark magenta) and nonDE (gray) MERMITE18B in rice endosperm (left) and embryo (right) at 7-8 DAP. DNA methylation at CG, CHG, and CHH contexts are shown in the top, middle, and bottom panels, respectively. **c,d** 21, 22, 23 and 24-nt siRNA levels of MERMITE18B in rice endosperm (c) and embryo (d). **b-d** DE and nonDE MERMITE18B

displayed by dark magenta and gray lines, respectively. The levels of siRNA and DNA methylation were examined at TE bodies and their adjacent +/- 1 kb regions.

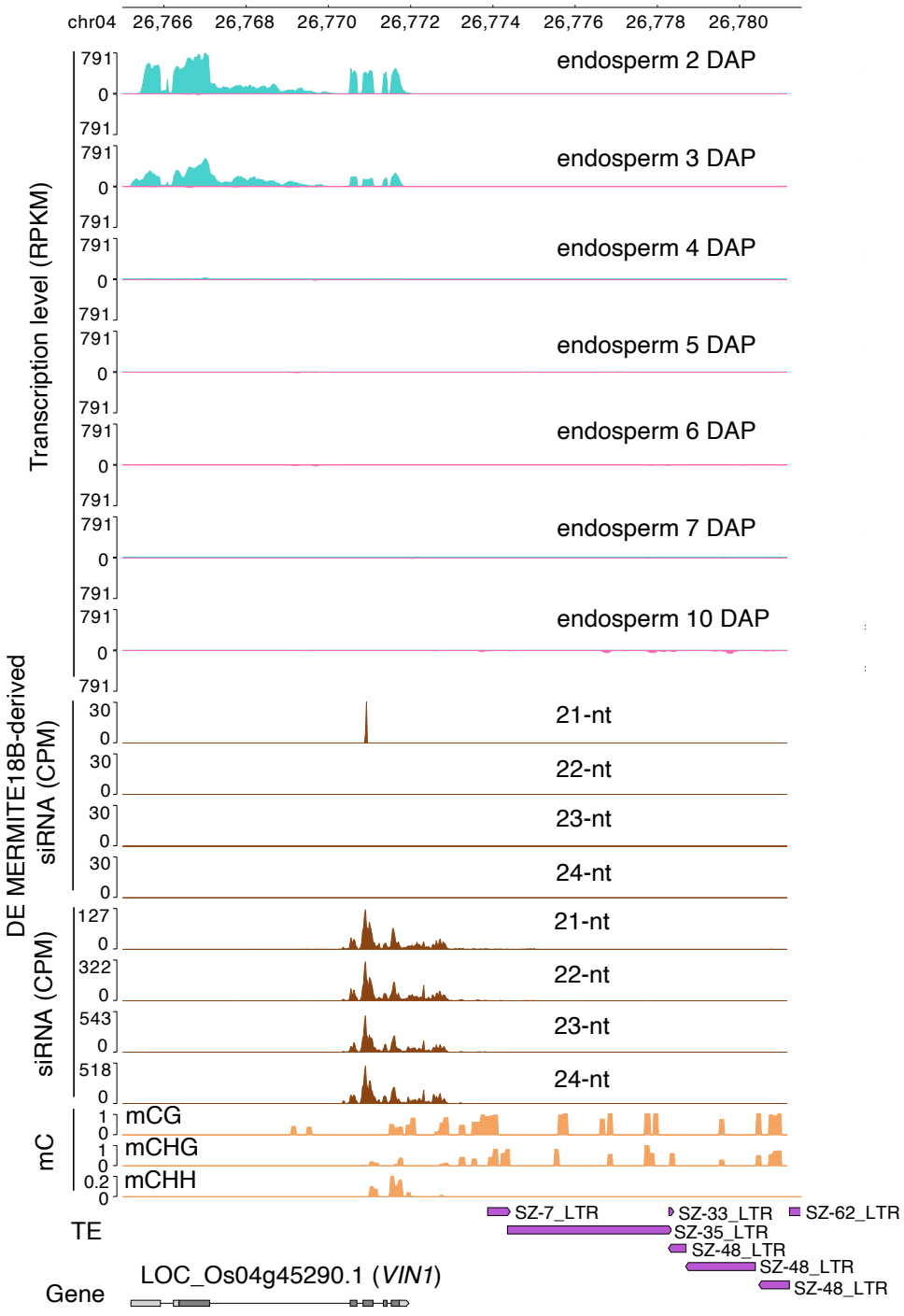

Fig. 3 Screenshot of transcription, siRNA enrichment, and DNA methylation at the *VIN1* locus. The overview illustrates transcription (Watson and Crick strand for each DAP), DE MERMITE18B-derived siRNA (21 to 24-nt), siRNA (21 to 24-nt), DNA methylation (CG, CHG and CHH contexts), gene, and TE loci.

List of my articles

1. Tonosaki K, Ono A, Kunisada M, Nishino M, Nagata H, Sakamoto S, Kijima ST, Furuumi H, Nonomura KI, Sato Y, Ohme-Takagi M, Endo M, Comai L, Hatakeyama K, Kawakatsu T, Kinoshita T. Mutation of the imprinted gene *OsEMF2a* induces autonomous endosperm development and delayed cellularization in rice. *Plant Cell* 33: 85-103, 2021

2. Hiroki Nagata, Akemi Ono, Kaoru Tonosaki, Taiji Kawakatsu, Kentaro Yano, Yuji Kishima, Tetsu Kinoshita. Temporal Changes in Transcripts of MITE Transposable Elements during Rice Endosperm Development. *The Plant Journal*, **109**(5): 1035-1047, 2022.

3. Leonie Verhage. Can transposable elements rewire transcriptional networks in the developing rice endosperm? *The Plant Journal*, **109**(5): 1033-1034, 2022. (Research Highlight)

4. Hiroki Nagata, Akemi Ono, Kaoru Tonosaki, Taiji Kawakatsu, Tetsu Kinoshita. Relationship between expression of *Vacuolar acid invertase 1* and accumulation of a 21-nt siRNA corresponding to highly expressed Mutator-like transposable elements, MERMITE18B, in rice endosperm. *submitted.*

トランスポゾン (TE) は、ゲノム上の遺伝子座から別の遺伝子座に転移することができる有害な特性を持つため、宿主のエピジェネティックな機構によって抑制される。DNA メチル化は TE に対する抑制マークの 1 つであり、特に被子植物の胚乳組織で TE に対する DNA メチル化レベルが低下することが知られている。このことは、胚乳において TE がエピジェネティックに活性化していることを示唆しているが、胚乳における転写活性化については不明な点が多い。そこで、胚乳発生時における TE 発現とその生物学的な役割について分子的な手がかりを得るために、多様な長さと種類の TE をゲノムにもつイネの胚乳組織を用いて、TE の時系列トランスクリプトーム解析を行った。

第 1 章では、イネ TE プローブを豊富に含むマイクロアレイを用いて、イネゲノムの中で最も多くのコピー数存在し長さの短い TE である Miniature Inverted-repeat TE (MITE) の胚乳発生時における発現量の時間変化を調べた。その結果、MITE は胚乳発生時において動的に発現を増減させることが明らかになった。興味深いことに、MITE サブファミリーである Youren を含むいくつかの MITE 転写産物は、胚乳において方向性をもって制御されていた。さらに、Youren の 5' 末端逆位反復配列の近傍領域に NF-Y 転写因子の結合モチーフ CCAAT を発見し、このモチーフの存在と発現レベルに関連があることを明らかにした。以上の結果は、イネの胚乳発生時における MITE の転写活性の動的な変化を明らかにするものである。

第 2 章では、RNA-seq による時系列解析と、DE (Differentially Expressed) TE の発現変化のクラスタリング解析により、胚乳形成時ににおける比較的配列の長い TE の動的な転写を明らかにした。その結果、MULE/MuDR TE スーパーファミリーのサブファミリーの 1 つである MERMITE18B が、DE TEs のいくつかのクラスターに高度に濃縮されていることを見出した。DE MERMITE18B は nonDE と比較して胚乳中で DNA 低メチル化されており、特に 21-nt siRNA が蓄積していることがわかった。この siRNA の生合成の役割を明らかにするために、イネゲノム中の標的遺伝子を探索した。その結果、スクロースの加水分解酵素をコードする *Vacuolar Acid Invertase 1* (*VIN1*) は、DE MERMITE18B 由来の 21-nt siRNA と潜在的な結合部位を有していた。*VIN1* は初期の多核体期の胚乳で優位に発現し、その後減少し、DE MERMITE18B の発現と負の相関を示した。このことから、*VIN1* は DE MERMITE18B 由来の 21-nt siRNA による転写後サイレンシングにおける主要な候補遺伝子である可能性が示唆された。以上の結果から、RNA-seq 解析とマルチオミクス解析により、イネ胚乳発生過程で MERMITE18B が高発現し、MERMITE18B 由来の siRNA により *VIN1* が抑制される可能性が示唆された。

これらの結果は、TE の転写がタンパク質をコードする遺伝子と同様に秩序立って制御されており、イネ胚乳発生時に遺伝子発現を制御する宿主の機構に TE の転写産物が組み込まれている可能性を支持するものである。