Analysis of the function of a rice microRNA, *miR820*

Misuzu Nosaka

Laboratory of Plant Genetics and Breeding
Division of Bioresource Functions
Department of Biological Mechanisms and Functions
Graduate School of Bioagricultural Sciences
Nagoya University, Furo-cho, Chikusa-ku, Nagoya, Aichi, Japan

TABLE OF CONTENTS

ABBREVIATIONS

DCL1	dicer-like protein 1
dsRNA	double-stranded RNA
GFP	green fluorescent protein
miR820	*microRNA820*
miRNA	microRNA
pri-miRNA	miRNA primary transcript
PTGS	post transcriptional gene silencing
RdDM	RNA-directed DNA methylation
RISC	RNA-induced silencing complex
siRNA	small interfering RNA
SNP	single-nucleotide polymorphism
TE	transposable element

ABSTRACT

In many eukaryotes, the major component of their genome is repetitive sequences like transposable elements (TE) that are also called genomic parasites. Transposition of TE often induces harmful mutations such as inserted mutations and aberrant chromosomes. Thus, most TEs are kept silence by host's defense machinery through small RNA-mediated RNA silencing. On the other hand, although TEs are kept silence, most of the eukaryotic genomes are composed by TEs and their remnants. This suggests the existence of a mechanism by which TE can suppress or avoid host's silencing, however, nothing is substantially known about this mechanism.

In Chapter 1 of the main thesis, I summarized the accumulated knowledge of the interplay between TEs and host until now. In the following Chapter 2, I reported the analysis of a pathway that TE avoids the silencing against itself by using host's RNA silencing machinery in rice. A rice microRNA (miRNA), *miR820*, is produced from CACTA DNA transposon and predicted to target *OsDRM2* encoding DNA methyltransferase. In general, DNA methylation by DNA methyltransferase has an important role in silencing TE. Therefore, I set up a hypothesis that *miR820* prevents DNA methylation of its own locus by suppressing the expression of DNA methyltransferase gene and helps the transposon to avoid host's silencing, and verified the hypothesis. Firstly, I confirmed experimentally that *miR820* represses DNA methylation of transposons through suppressing the expression level of DNA methyltransferase gene, *OsDRM2*. Secondly, to clarify whether the *OsDRM2* regulation

by *miR820* is conserved in the genus *Oryza* including wild rice species, I checked the sequences of *miR820* and its target site in *OsDRM2* among wild rice species. As a result, the *OsDRM2* regulation by *miR820* was conserved in all the genus *Oryza* tested. In this process, I shed light on one of the wild rice species, *Oryza punctata*. In *Oryza punctata*, I found that the sequences of *miR820* and its target site had co-evolved to maintain their recognition each other, and CACTA transposons carrying *miR820* proliferated extremely in these lineages. These findings suggest that *miR820* is a component of an anti-host machinery for transposon, and moreover, they suggest the possibility of regulation by *miR820* contributes not only to transposon but also to adaptation of host.

In Chapter 3, I reported the analysis of a transcription mechanism of *miR820* precursor. *miR820* precursor is transcribed from *MIR820* locus and processed into *miR820*. Since *MIR820* is located inside the transposon, it exists in heterochromatic region unlike the ordinary miRNAs exist in euchromatic regions. Thus, the transcription mechanism of *MIR820* possibly differs from that of ordinary miRNAs. In this research, firstly, I found that, among the five *MIR820* loci in the Nipponbare genome, only the one located on chromosome 7 was transcribed. Histone modification and DNA methylation status around the *miR820* coding region of *MIR820* locus on chromosome 7 differed from that of other four loci. Together, these observations suggest that *miR820* regulates the DNA methylation status of its own locus and *MIR820* transcription may depend on this epigenetic modification status. This research revealed that *miR820* is a unique miRNA which functions both in *cis* and in *trans*.

In Chapter 4, I analyzed the relationship between the copy number of CACTA transposon carrying *miR820* and the expression level of *miR820* in various rice cultivars and wild rice species. In *Oryza punctata*, I confirmed that the copy number of *MIR820*

has increased and more than 18 copies exist, while *miR820* is not expressed. This suggested that exploding proliferation of CACTA transposon strongly induced the silencing of CACTA transposon and suppression of *miR820* transcription. Thus, I checked the copy number of CACTA transposon with *miR820* in rice cultivars that are expected to express *miR820*. As a result, among the rice cultivars, the copy number of CACTA transposon with *miR820* at the maximum was 11, and most cultivars had five copies or so. In these rice cultivars examined, there was no obvious correlation between the copy number of CACTA transposon with *miR820* and the expression level of *miR820*. Therefore, I concluded that transposition of CACTA transposon is affected by another regulatory mechanism in addition to the suppression of *OsDRM2* by *miR820*.

In Chapter 5, base on the results in Chapter 2 to Chapter 4, I considered the overview of this thesis. In this thesis, I revealed that two types of mechanism both exist; one is host's mechanism silencing TE and the other is TE's mechanism avoiding or suppressing host's silencing. Until now, most of the research focused on the host's defense machinery against TE. This is the first research threw light on anti-host machinery of TEs, which are genomic parasites. This research demonstrated that the interplay between these two mechanisms is related to composition of host genome.

CHAPTER 1

GENERAL INTRODUCTION

General introduction

Most transposable elements (TEs) are kept silence by their host's defense machinery since active TE, which induces inserted mutations and aberrant chromosomes by transposition, is harmful to host genome stability. On the other hand, TEs have increased their copy number on their host genome by their selfish behavior. Consequently, most of the eukaryotic genome is mainly composed of TEs and their remnants. Although most TEs are silenced by their host, TEs are major component of their host genome. This suggests the existence of a running battle between the host's defense machinery suppressing transposition of TEs and TE's countermeasures against host-mediated silencing. However, until now, nothing is known about the strategies that TEs have evolved to avoid host-mediated machinery suppressing TE's transposition. In this chapter, my research on plant sheds light on a close battle between host genome and their parasitic TEs, and the function of small RNAs concerning to this battle.

RNA silencing as defense machinery against exogenous genes

1) Genomic immune system against genomic parasites

The major component of eukaryotic genomes is TEs and their remnants (Kidwell, 2002). TEs have increased their copy number in the host genome through self-replication faster than the host genome by using host's cellular replication system. Thus, TEs are referred to as ultimate parasites that proliferate selfishly on the genome (Doolittle and Sapienza, 1980; Orgel and Crick, 1980). On the other hand, looking from the host's side, TE's

transposition that induces inserted mutations and aberrant chromosomes cause instability of host genome. Therefore, most of TEs are kept silence by their host. A number of groups are carrying out research on the host's defense machinery against genomic parasites, such as TEs, in plants and animals. Until today, small RNA-mediated RNA silencing and its corresponding mechanism is known to participating in suppression of TE, and this mechanism is known to have a common role in plants and animals to some degree (Zilberman and Henikoff, 2004; Brennecke et al., 2007; Kuramochi-Miyagawa et al., 2008). Particularly, in plants, many mutants of host's components concerning to RNA silencing are analyzed, and these components are known to act in the background of a phenomenon called Post Transcriptional Gene Silencing (PTGS) or Co-suppression (Chan et al., 2004). PTGS or Co-suppression is a phenomenon that suppresses the expression of both exogenous and endogenous gene with the same sequence when the exogenous gene is introduced to the plant cell. Thus, it is thought that a genomic invader, such as TE, is also recognized as exogenous gene and suppressed its expression by this mechanism (Slotkin et al., 2005).

As well as TEs, viruses also invade host genome and proliferate by using host's cellular replication system. Against the infection of viruses, host has defense machinery, such as the immune system. In plants, most pathogenic viruses are using RNA for their genome, and the host's defense machinery against these viruses are, alike the case of TEs, concerned with RNA silencing (Waterhouse et al., 2001).

As mentioned above, it is thought that one of the biological significance of RNA silencing is the host's defense machinery against the exogenous genes. In other words, RNA silencing can be regarded as genomic immune system since it can distinguish the exogenous genes like genomic parasites from endogenous genes, and suppress the

exogenous gene specifically (Waterhouse et al., 2001; Plasterk, 2002; Aravin et al., 2007; Saito and Siomi, 2010).

2) Countermeasures of genomic parasites against host-mediated silencing

As long as the relationship between the host and the parasites exists, it is not one-sided. That is to say, while the host defends itself against the attack from the parasites constantly, the parasites acquire the system to overcome the host's defense machinery with high speed. Indeed, most RNA viruses infecting plant cell have mechanism that avoid or suppress host's defense machinery. For example, many of the RNA viruses are coding RNA silencing suppressor protein (suppressor) on their genome. Suppressor affects RNA silencing, which works as host's defense machinery, at various points; some suppressors block the loading of small RNA, which induces RNA silencing, onto a RNA-induced silencing complex (RISC), and others prevent the RNA silencing to start up by binding to the small RNA (Voinnet et al., 2000; Llave et al., 2000; Kasschau et al., 2003). These findings suggest that RNA silencing is just in the front line of arms race between the host and the parasites.

On the other hand, TEs have increased their copy number on host genome by their selfish behavior. As a result, TEs and their remnants are the major component of eukaryotic genome. This suggests the existence of a running battle between the host's defense machinery and the TE's countermeasures against the host to evade the host silencing. In TEs, so far, the mechanism to escape or suppress host's defense machinery has not been cleared as in RNA viruses. However, some strategies are presumed to work as TE's countermeasures against the host (Lisch and Slotkin, 2011). For instance, it is known that one kind of TE tends to be inserted near the host gene. This tendency

suggests that TE avoids being a target for host's silencing together with the gene adjacent to its inserted position (Liu, 2004). There are also some researches suggesting the existence of TE's aggressive mechanism capturing host gene that contributes to its own proliferation. TEs in plants and animals are capturing a gene coding a protease, which is called ULP1, frequently. Although the function of ULP1 in TE's activity is not yet clear, the fact that TEs are proliferating by capturing the same gene beyond the species is presumed to lead the possibility that this gene somehow contributes to TE's activity (Hoen et al., 2006; van Leeuwen et al., 2007).

In sum, even though the existence of TE's countermeasures against the host is strongly suggested, there is hardly any example making clear of its functional mechanism concretely. My colleagues have analyzed the small RNA in rice comprehensively (Lacombe et al., 2008). In this process, I focused on a microRNA (miRNA) produced from the TE's loci and targeting a *de novo* DNA methyltransferase gene, which is related to suppression of TEs in the downstream of host's RNA silencing machinery. In the next chapter, I will demonstrate the running battle between the host and the parasitic TE through the function of small RNA, which come to light by my research.

CHAPTER 2

Role of transposon-derived small RNAs in the interplay between genomes and parasitic DNA in rice.

Abstract

RNA silencing is a defense system against "genomic parasites" such as transposable elements (TE), which are potentially harmful to host genomes. In plants, transcripts from TEs induce production of double-stranded RNAs (dsRNAs) and are processed into small RNAs (small interfering RNAs, siRNAs) that suppress TEs by RNA-directed DNA methylation. Thus, the majority of TEs are epigenetically silenced. On the other hand, most of the eukaryotic genome is composed of TEs and their remnants, suggesting that TEs have evolved countermeasures against host-mediated silencing. Under some circumstances, TEs can become active and increase in copy number. Knowledge is accumulating on the mechanisms of TE silencing by the host; however, the mechanisms by which TEs counteract silencing are poorly understood. Here, I show that a class of TEs in rice produces a microRNA (miRNA) to suppress host silencing. Members of the *microRNA820 (miR820)* gene family are located within CACTA DNA transposons in rice and target a *de novo* DNA methyltransferase gene, *OsDRM2*, one of the components of epigenetic silencing. I confirmed that *miR820* negatively regulates the expression of *OsDRM2*. In addition, I found that expression levels of various TEs are increased quite sensitively in response to decreased *OsDRM2* expression and DNA methylation at TE loci. Furthermore, I found that the nucleotide sequence of *miR820* and its recognition site within the target gene in some *Oryza* species have co-evolved to maintain their base-pairing ability. The co-evolution of these sequences provides evidence for the functionality of this regulation. My results demonstrate how parasitic

elements in the genome escape the host's defense machinery. Furthermore, my analysis of the regulation of *OsDRM2* by *miR820* sheds light on the action of transposon-derived small RNAs, not only as a defense mechanism for host genomes but also as a regulator of interactions between hosts and their parasitic elements.

Introduction

RNA silencing is a mechanism mediated by small RNAs that regulates gene expression in eukaryotes at both the transcriptional and post-transcriptional levels. RNA silencing has a wide range of essential functions in cellular processes necessary for development of animals and plants, and it also has a role in defense against "genomic parasites" such as transposable elements (TEs) and viruses (Plasterk 2002; Almedia and Allshire 2005; Aravin et al., 2007). Silencing of TEs is triggered by small RNAs derived from the TE loci themselves. These small RNAs are usually 24 nt long in plants and are called small interfering RNAs (siRNAs). siRNAs are produced from TE transcripts by an enzyme called Dicer. Dicer acts on double-stranded RNA generated either by the action of RNA-dependent RNA polymerases or by transcription from both DNA strands. The TE-derived siRNAs are loaded onto Argonaute proteins, which degrade TE transcripts or repress translation by means of base-pairing between the transcripts and siRNAs (Saito and Siomi 2010). In plants, TE-derived siRNAs also induce RNA-directed DNA methylation (RdDM), resulting in epigenetic inactivation of the TEs (Zilberman and Henikoff 2004; Lisch 2009; Matzke et al., 2009).

Although the majority of TEs are epigenetically silenced, most of the eukaryotic genome is composed of TEs and their remnants (Feschotte et al., 2002; Kidwell 2002). This suggests that TEs have evolved countermeasures against host silencing (Lisch 2009), but the mechanisms by which TEs counteract silencing are poorly understood. In this paper, I demonstrate that a small RNA derived from certain TE loci suppresses the

host silencing machinery. Generally, siRNAs produced from TEs trigger silencing of those same TEs; however, in this case, TEs escape host silencing by producing a class of miRNAs that acts on the host silencing machinery. My analysis provides evidence for a novel mechanism by which transposons reduce host silencing, and it provides a glimpse of "front line" of host genome–parasitic DNA interactions through the action of small RNAs produced from the transposon.

Materials and Methods

Plant materials

Wild-type Nipponbare and *waf1* mutant rice plants were grown in soil or in tissue culture boxes at 29°C under continuous light. DNA, plants, and seeds of *Oryza* species were kindly provided by the National Institute of Genetics (Mishima, Japan).

Plasmid construction and production of transgenic plants

OsDRM2 cDNA was kindly provided by Dr. S. Iida, Shizuoka Prefectural University (Shizuoka, Japan). The *p35S:OsDRM2 intact:GFP*, *p35S:OsDRM2 mutation1:GFP*, and *p35S:OsDRM2 mutation2:GFP* vectors were constructed by introducing mutations using the GeneTailor Site-Directed Mutagenesis System (Invitrogen). Next, the part of each *OsDRM2* cDNA that included the *miR820* target site was amplified and cloned into the pENTR/D-TOPO vector (Invitrogen). The resultant vectors containing the cDNA fragments were introduced into the pGWB5 binary vector (Nakagawa et al., 2007), which carries a GFP reporter gene driven by the 35S promoter, by using Gateway technology (Invitrogen). For *pAct:pre-miR820:Nos* construction, a 0.5-kb *pre-miR820* fragment was amplified and inserted into the pCRII vector (Invitrogen). A *pre-miR820* fragment was then excised with *Xba*I and *Sma*I, and cloned into the binary vector carrying the rice *Actin* gene promoter and *Nos* terminator. For *pAct:OsDRM2 RNAi:Nos* construction, a 0.9-kb *OsDRM2* cDNA fragment with *Pst*I and *Xba*I linkers was cloned into the *Pst*I and *Xba*I sites of the pBS-SK vector containing a partial *GUS*

fragment at its *EcoR*V site. Similarly, a cDNA fragment with *Hind*III and *Sma*I/*Apa*I linkers was inserted into the *Hind*III and *Apa*I sites of the vector. The resultant vector was cloned into the *Xba*I and *Sma*I sites of a binary vector carrying the rice *Actin* gene promoter and *Nos* terminator. These binary vectors were introduced into *Agrobacterium* strain EHA101 and used for transformation of rice by the standard method (Hiei et al., 1994). The primers used for vector construction are listed in Table 2.1.

RNA analysis

Total RNA was isolated from shoots of *waf1* and various tissues of Nipponbare wild-type non-transgenic plants; shoots of *p35S:OsDRM2 intact:GFP*, *p35S:OsDRM2 mutation1:GFP*, and *p35S:OsDRM2 mutation2:GFP* T_2 plants; and calli of *pAct:pre-miR820* and *pAct:OsDRM2 RNAi* by using TRIzol reagent (Invitrogen). For analysis of *waf1* and wild-type plants and of *pAct:pre-miR820:Nos* and *pAct:OsDRM2 RNAi:Nos* callus, 10 μg of each RNA sample was loaded onto an agarose or acrylamide gel (for analysis of *OsDRM2* and *miR820a/b/c*, respectively), separated by electrophoresis, and blotted onto nylon membranes. The membranes were probed with oligo DNA complementary to *miR820a/b/c* or *OsDRM2* cDNA, depending on the experiment.

5′ RACE

Total RNA was purified with the RNeasy Mini Kit (QIAGEN) according to the manufacturer's instructions. 3 μg of purified total RNA was subjected to RNA Oligo ligation with the GeneRacer Kit (Invitrogen) according to the manufacturer's instructions. The oligo-ligated RNA was reverse-transcribed using Omniscript Reverse

Transcriptase (QIAGEN) with random primers (N_9). PCR and nested PCR were performed using *Ex Taq* DNA polymerase (TaKaRa). Primers used for 5′ RACE PCR are listed in Table 2.1. Amplified bands were gel-purified, cloned, and sequenced.

RT-PCR

Relative expression levels were quantified using the StepOnePlus Real-Time PCR system (Applied Biosystems) and the One Step SYBR PrimeScript RT-PCR Kit II (TaKaRa). The quantitative RT-PCR reactions contained 5 µl 2× One Step SYBR RT-PCR Buffer 4, 0.5 µl DMSO, 0.4 µl PrimeScript 1 step Enzyme Mix 2, 0.2 µl 50× ROX reference dye, 50 ng total RNA, and 400 nM of each primer, and were run in triplicate. The mixtures were first reverse-transcribed at 42°C for 5 min, then amplified via PCR using a two-step cycling program (95°C for 5 s, 60°C for 20 s) for 40 cycles. Quantitative RT-PCR specificity was checked for each run with a dissociation curve, at temperatures ranging from 95°C to 60°C. Data from quantitative RT-PCR were analyzed using the standard-curve method. The housekeeping genes *OsActin* and *OsGAPDH* were used to normalize the quantitative RT-PCR output. Primers used for quantitative RT-PCR are listed in Table 2.1.

***McrBC*-PCR**

Genomic DNAs were isolated from wild-type nontransgenic and *pAct:OsDRM2 RNAi:Nos* calli. For *McrBC*-PCR analysis, 500 ng of genomic DNAs were digested with or without 40 units of *McrBC* restriction enzyme (New England Biolabs) for 12 hr. PCR was performed using *Ex Taq* DNA polymerase (TaKaRa). Primers used for PCR are listed in Table 2.1. *OsActin* and *Centromere 8* are controls for regions with low and

high DNA methylation, respectively.

Sequence analysis

Genomic DNA samples from various *Oryza* species were kindly provided by the National Institute of Genetics (Mishima, Japan). I amplified both *miR820* and its target site in *DRM2* by PCR using the primers listed in Table 2.1. The amplified DNA fragments were gel-purified and used as templates for direct sequencing. The miRNA target score was calculated for each *miR820*:*DRM2* duplex based on the method described in (Allen et al., 2005). To detect the copy number of CACTA TEs carrying *miR820* by Southern blot analysis, genomic DNA samples were extracted from leaves of Nipponbare (AA), W1514 (BB), W1331 (BBCC), and W1805 (CC), treated with RNase A, and digested with restriction enzymes. These samples were loaded onto an agarose gel, separated by electrophoresis, blotted onto a nylon membrane, and probed with the *pre-miR820* DNA fragment.

Mapping of CACTA carrying *pre-miR820*

My strategy to map *miR820*-CACTA from BB-genome species was based on the synteny between AA and BB *Oryza* species (Kim et al., 2007). Briefly, by screening the BAC library of a BB-genome species, I identified BAC clones carrying *miR820*-CACTA from BB. Then, using the BAC end sequences of these clones deposited to database, I identified the corresponding physical position of these clones in the Nipponbare genome. This strategy is advantageous over other methods, such as transposon display, to monitor the varieties of transposon, especially long transposons with specific internal sequences, because transposon display identifies only the ends of

transposon sequences. The precise method used for this experiment was as follows: A BAC filter and library of *Oryza punctata* (genome BB) genomic DNA were purchased from the Arizona Genomics Institute (Tucson, AZ). By screening these libraries using a labeled *pre-miR820* DNA fragment, I identified 48 BAC clones carrying *miR820*-CACTA. I confirmed that these clones carried *miR820*-CACTA by PCR amplification and sequencing of the region around *pre-miR820* in CACTA. Using the BAC end sequence obtained from http://www.omap.org/, I located those BACs on a physical map of the Nipponbare rice genome. Multiple sequence alignment for the phylogenetic analysis was constructed using Clustal X, and an unrooted tree was made by the neighbor-joining method (Saitou and Nei 1987) using PAUP 4.0 software (Sinauer Associates).

In situ mRNA hybridization

In situ hybridization was performed as previously described by Kouchi and Hata (1993). For the OsDRM2 probe, the full-length cDNA clone was used as a template for in vitro transcription. Hybridizations were conducted at 55°C overnight; slides were then washed four times at 50°C for 10 min each. An excess amount of sense transcript was used as negative control.

Results and Discussion

Transposon-derived *miR820* targets *de novo* DNA methyltransferase gene *OsDRM2*

miRNAs are produced from stem structures formed within noncoding transcripts (Meyers et al., 2008) and negatively regulate the expression of a range of plant genes, mainly by mRNA cleavage (Voinnet 2009). *miR820* is a small-RNA species with sizes of 22 and 24 nt (Chellappan et al., 2010; Wu et al., 2010). *miR820* is produced from transcripts originating from a region inside a class of CACTA DNA transposons in rice (Fig. 2.1A). There are five copies of the CACTA transposon containing the *miR820* precursor (*pre-miR820*) in the rice (*Oryza sativa* L.) Nipponbare genome (Rice Annotation Project 2008) (Fig. 2.2A-B). Three of the *pre-miR820*s (*miR820a, -b, and –c*) encode the identical miRNA sequence (Luo et al., 2006), whereas *miR820d* and *miR820e* differ from the other three by one and two nucleotides, respectively (Fig. 2.2C). The nucleotide sequences of the fold-back region of all five *pre-miR820* sequences show high sequence similarity to parts of *Os03g0110800* and the homologous region extends into the second exon and third intron of *Os03g0110800* (Fig. 2.1A, 2.3, 2.4). Thus, *pre-miR820* possibly originated from *Os03g0110800*, and the number of *pre-miR820* copies increases as the CACTA TEs propagate.

Because of this homology, *miR820* is predicted to target *Os03g0110800* (*OsDRM2*), which encodes a *de novo* DNA methyltransferase orthologous to *Arabidopsis DRM1/2* (Luo et al., 2006; Cao etal., 2000; Cao and Jacobsen 2002; Sharma et al., 2009;

Henderson et al., 2010) (Fig. 2.5A). It has been reported that the 24-nt species of *miR820* acts as a guide for DNA methylation at its target site, possibly through RdDM (Wu et al., 2010). Indeed, I also confirmed the function of the 24-nt *miR820* species by detecting a high level of cytosine methylation specific to its presumed target site (Fig. 2.5B). Because *pre-miR820* loci simultaneously produce both 22-nt and 24-nt miRNA species (Fig. 2.1B), I investigated whether the 22-nt *miR820* species regulates *OsDRM2* expression through mRNA degradation by mapping the 22-nt *miR820* cleavage site of *OsDRM2*. I found a cleavage site at the predicted position for miRNA-based target gene cleavage (Fig. 2.5C).

I further confirmed that this cleavage depends on the presence of *miR820* by using the *waf1* mutant in rice (Abe et al., 2010) (Fig. 2.1B-D). In *waf1,* accumulation of small RNAs is greatly decreased because of a mutation in *HEN1*, a gene encoding an RNA methyltransferase that is required for the stability of small RNAs (Li et al., 2005; Yu et al., 2005; Yang et al., 2006). In *waf1*, the expression levels of both the 22-nt and 24-nt species of *miR820* decreased compared to the wild-type (Fig. 2.1B). To confirm that *OsDRM2* mRNA cleavage depends on the presence of *miR820,* I checked for the cleavage product in *waf1* mutants and in the wild-type. In the *waf1* mutants, there was no detectable cleavage of *OsDRM2* mRNA by *miR820* (Fig. 2.1C). I also confirmed that the expression level of *OsDRM2* increased in *waf1* compared to the wild-type (Fig. 2.1D). It is possible that this increase was not due solely to the loss of *miR820* because in *waf1*, the levels of most other small RNAs are also reduced (Abe et al., 2010). However, considering that *OsDRM2* gave the highest hit score when *miR820* was used in BLAST searches against the entire rice genome (IRGSP Pseudomolecules 1.0) other than *miR820* itself, it is very likely that *miR820* negatively regulates the expression of

OsDRM2 at least in part.

Negative regulation of *OsDRM2* by *miR820* activates TE expression

To test whether the expression level of *OsDRM2* depends on recognition by *miR820*, I made transgenic rice plants that express a fusion of a green fluorescent protein (GFP) gene and *OsDRM2* with or without synonymous mutations within the *miR820* recognition site; I then observed the GFP fluorescence and measured *GFP* mRNA levels (Fig. 2.6A, B). As expected, the expression level of the *OsDRM2:GFP* fusion gene with an intact *miR820* recognition site was much lower than for those with synonymous mutations. In wild-type plants, both *miR820* and *OsDRM2* were expressed in all the tissues tested, although their expression levels differed between tissues (Fig. 2.5D, G).

Next, I tested whether the expression patterns of *OsDRM2* and *miR820* overlapped. Northern analysis using total RNA extracted from vegetative shoots from two wild-type rice cultivars demonstrated that both genes were expressed within this tissue (Fig. 2.5E). *In situ* hybridization experiments revealed that *OsDRM2* is ubiquitously expressed in vegetative shoots (Fig. 2.5F). This suggests that the expression patterns of *miR820* and *OsDRM2* overlap at the cellular level, supporting the idea that *miR820* regulates *OsDRM2*. On the other hand, I did not observe a clear inverse relationship between the levels of *miR820* and *OsDRM2* expression. This might be because the expression levels of *miR820* and *OsDRM2* differed between tissues, and because *miR820* might reduce the amount of *OsDRM2* expression but not abolish it completely. Indeed, I found that overexpression of *pre-miR820* under the control of a strong constitutive promoter mildly reduced but did not eliminate the expression of *OsDRM2* (Fig. 2.7A-D).

Because *de novo* DNA methyltransferase is a component of the host's silencing machinery (Cao et al., 2000; Cao and Jacobsen 2002), I tested whether reduced *OsDRM2* expression would affect the transcription of TEs by using transgenic rice plants in which *OsDRM2* expression was reduced by RNAi. I found that the expression levels of several TEs were increased in *DRM2* RNAi transgenic lines; furthermore, the expression levels of TEs such as *RIRE7* and CACTA carrying *pre-miR820* were inversely related to the degree of *DRM2* suppression (Fig. 2.6C, 2.8). Next, I observed the DNA methylation status at several TE loci by *McrBC*-PCR analysis (Fig. 2.6D). In *OsDRM2* RNAi lines, DNA methylation within CACTA (including the *pre-miR820* region) and *RIRE7* is clearly reduced compared to the wild-type. Furthermore, I also observed elevated expression of *RIRE7* in the same *pre-miR820* overexpression experiment in which *OsDRM2* expression was found to be mildly reduced (Fig. 2.7E). These experimental data are consistent with the idea that *OsDRM2* is involved in TE silencing through DNA methylation.

The sequences of *miR820* and its target site in *OsDRM2* have co-evolved in BB/BBCC *Oryza* species

I did not find *miR820* or its precursor sequence in the Arabidopsis or maize genome, suggesting that *miR820* is not widely conserved in plants. I then tested whether regulation by *miR820* is conserved among various *Oryza* species. I successfully amplified and sequenced both *miR820* and its recognition site in *DRM2* from the genomic DNAs of various accessions of *Oryza* (Ge et al., 1999) (Fig. 2.9; Table 2.2), strongly suggesting the conservation of this regulation mechanism among *Oryza* species. I recovered sequences identical to *miR820a/b/c* from all the *Oryza* genomes tested

except for the BB and BBCC genomes (Fig. 2.9A; Table 2.2). In species with BB or BBCC genomes, *miR820*-related sequences had three nucleotide substitutions compared with *miR820a/b/c*. Considering the phylogenetic relationships among *Oryza* species (Ge et al., 1999), the *miR820* sequence recovered from BB/BBCC *Oryza* species has diverged from *miR820a/b/c* (Fig. 2.10A).

There are also several nucleotide substitutions in the *miR820* recognition site in *DRM2* in some *Oryza* genomes (Fig. 2.9B). Remarkably, in the BB and BBCC genomes, there are five nucleotide substitutions in *DRM2*. Thus, in the BB and BBCC genomes, there are eight nucleotide substitutions in *miR820* and its recognition site in *DRM2*, compared with the corresponding *miR820* and target sequences in Nipponbare. This number of substitutions could greatly affect the capability of *miR820* to regulate *DRM2* in species with those genomes; however, the degree of base-pairing between *miR820* and its target site in *DRM2* in the BB and BBCC genomes is conserved (Fig. 2.10B; Table 2.2). This indicates that, in BB/BBCC *Oryza* species, the sequences of *miR820* and its target site in *DRM2* have co-evolved to maintain the ability to form a stable RNA–RNA duplex. The co-evolution of these sequences strongly suggests that the regulation of *DRM2* by *miR820* is functional and that those nucleotide changes have accumulated as a result of the interplay between the host genome and the parasitic elements in these species.

TEs carrying *pre-miR820* have proliferated in BB-genome species

To see whether co-evolution of the nucleotide sequences of *DRM2* and *miR820* affected the behavior of TEs carrying *pre-miR820* in the BB genome, I performed Southern blot analysis to detect the copy number of CACTA carrying *pre-miR820* (Fig. 2.11A). I found that the copy number of CACTA with *pre-miR820* was much higher in the

BB/BBCC *Oryza* species than in the AA species Nipponbare. I also successfully determined the genomic locations of CACTA with *pre-miR820* in the BB genome (see Materials and Methods for details) and found that at least 18 copies of CACTA with *pre-miR820* are dispersed throughout this genome (Fig. 2.11B). I also sequenced the CACTA with *pre-miR820* in BB-genome species and conducted phylogenetic analysis using *pre-miR820* sequences from Nipponbare and BB-genome species. This analysis revealed a sudden increase in copy number of CACTA carrying *pre-miR820*, in which identical sequences around the *pre-miR820* region were recovered from multiple loci (Fig. 2.11C). Because the *miR820* sequence in the BB species shown in Fig. 2.10B was obtained by direct sequencing of PCR products, it should be representative of the *miR820* sequence in BB species. In fact, the majority (11 out of 18 copies) of *pre-miR820* found in the BB genome carries the same *miR820* sequence as the one recovered by direct sequencing, which is also the sequence that would form the most stable hybrid with the *DRM2* sequence found in the BB genome. These results suggest that the CACTA transposon with this *miR820* sequence was predominantly proliferated or maintained, and became the predominant *miR820* in BB species.

I hypothesize the following scenario as a mechanism connecting the co-evolution of *miR820* and *DRM2* and the rapid increase in the copy number of CACTA carrying *pre-miR820*. When *OsDRM2* expression decreases, possibly because of nucleotide substitutions within *miR820* that enable it to form more stable hybrids with *OsDRM2* or for other reasons, more *miR820* can be produced, possibly because host-mediated silencing is suppressed efficiently. Indeed, RNAi-mediated suppression of *OsDRM2* increased *pre-miR820* expression (Fig. 2.8C). This is expected to drive the selection of nucleotide substitutions at *miR820* or at its target site because drastic reduction of

OsDRM2 levels could be lethal. This hypothesis is supported by the fact that I recovered *OsDRM2* RNAi transgenic plants with about half the normal expression level of *OsDRM2* (Fig. 2.6C). Thus, there should be selection pressure for mutations within the miRNA target site of *OsDRM2*. In turn, TE would favor changes in the *miR820* sequence that correspond to the changes in the target site. This evolutionary "arms race", in which hosts and parasitic DNA co-evolve, allows nucleotide substitutions to accumulate within both the miRNA and its target sequence, which maintains the ability to form stable hybrids between them. This might account for the fact that, in BB species, the most predominant CACTAs with *pre-miR820* were those that could form the most stable hybrid with the target sequence.

A model for the regulation of *DRM2* by *miR820* sequences is shown in Fig. 2.11D. In general, TE-derived small RNAs act as a trigger for silencing (Zilberman and Henikoff 2004; Lisch 2009; Matzke et al., 2009). However, in this case, transposons that incorporate miRNA genes that target the host's silencing machinery are able to counteract the host's defense system. Similar examples of "arms races" between hosts and parasites are well documented in studies of plant RNA viruses and their hosts (Waterhouse et al., 2001), in which RNA viruses that encode silencing-suppressor proteins are able to escape silencing by the host. My model for the regulation of *DRM2* by *miR820* predicts that this regulation might affect not only CACTA carrying *pre-miR820* but also other TEs. Indeed, in *DRM2* RNAi lines, I observed upregulation of expression from TEs other than CACTA (Fig. 2.6C). However, considering that BB-containing species have relatively small genomes compared with other *Oryza* species (Kim et al., 2007), the downregulation of *DRM2* by *miR820* would not be expected to affect a large number of TEs in BB-containing species. Rather the effect

might be relatively specific to particular TEs or their lineage, as has been observed for the *Arabidopsis ddm1* mutation (Tsukahara et al., 2009).

So far, *miR820* has been found only in rice, suggesting a recent origin. The primary and secondary structures of *pre-miR820* also support this idea, because *pre-miR820* still shows high homology within its stem parts to the intron sequence of *OsDRM2*. In general, non-conserved miRNA genes evolve very fast and they often appear and disappear from the genome. Considering that *miR820* is encoded by parasitic DNA and its primary function seems to be as an anti-host agent, it is possible that *miR820* might be lost in the future, as is often the case for non-conserved miRNA genes. However, it is intriguing to speculate that *miR820* might function not only as an anti-host mechanism for parasites but also in a way that is beneficial for the host. The co-evolution of *miR820* and its recognition site in BB/BBCC species supports this idea. It is possible that, in order to adapt against genomic stresses such as climate or environmental changes, the host maintained or created genome flexibility by keeping or allowing *DRM2* under the regulation of *miR820* in BB species in the past. Thus, my analysis of the regulation of *DRM2* by *miR820* sheds light on the action of two types of transposon-derived small RNAs, siRNA and miRNA, in the battles and possibly even the cooperation between plant genomes and their parasites.

Table 2.1. Primers used in this study.

Target gene	Primer name	Sequence (5' to 3')	Purpose
OsDRM2	qDRM1aF1	CAGAGGTAGCTCTCCACAGAC	RT-PCR
	qDRM1aR2	CAGCTCTTCAGAATCGTTCTG	RT-PCR
OsActin	ActAK060893Fw	GAGTATGATGAGTCGGGTCCAG	RT-PCR
	ActAK060893Rv	ACACCAACAATCCCAAACAGAG	RT-PCR
sGFP	sGFPF2	AAGCTGGAGTACAACTACAAC	RT-PCR
	sGFPR1	TTGTGGCGGATCTTGAAGTTC	RT-PCR
CACTA (ORF1)	CAChprF10	GTATGAATATGGCAAGCCGTT	RT-PCR
	CAChprR10	GAAACTGAAGGCGAAGTTTGC	RT-PCR
Tos17 (TPase)	Tos17F3	CACCAGGTGTGGAAGCTCCAC	RT-PCR
	Tos17R3	TACCACTGAGCTGAAGCGTGC	RT-PCR
RIRE7 (TPase)	qRIRE7F2	TCGCCAATGATCGCCTTGGTC	RT-PCR
	qRIRE7R2	AGACGATCCATCATGACCATC	RT-PCR
pre-miR820	SatofN	TTACGTTCCTTACCGATMTTG	RT-PCR
	R1SATON	TGATAACGTAMGAACTACACCTCC	RT-PCR
OsDRM2 (cleaved)	GeneRacer 5'	GGACACTGACATGGACTGAAGGAGTA	5'RACE
	AK065147R2	TTGCTCTGTATCTGTATCCCCAATCTCCTT	5'RACE
OsDRM2 (uncleaved)	qDRM2F1	GATAGCGACAATGATAAGTTCGAG	5'RACE
	qDRM2R3	TCCTCGTCGACTTCAAGAGTC	5'RACE
OsActin	ActU	TCCATCTTGGCATCTCTCAG	5'RACE
	ActL	GTACCCTCATCAGGCATCTG	5'RACE
OsDRM2 mutation1	miRJ-mutF1	CTAAGGAACATAGATGCTCCGGGCCCCTCAACACGATTACCGCAGGATGC	GFP fusion
	miRJ-mutR1	GGAGCATCTATGTTCCTTAGGGCCGGGGCA	GFP fusion
OsDRM2 mutation2	miRJ-mutF2	AACATAGATGCTCCTGGTCCCTCAACACGATTACCGCAGGATGC	GFP fusion
	miRJ-mutR2	GGACCAGGAGCATCTATGTTCCTTAGGGCC	GFP fusion
OsDRM2 intact PCR	DRM2CACCF6	CACCATGGTGGACTGGGCTTCAG	GFP fusion
	DRM2MIRjR4	CGGCAGCCTCGTGGACGGACC	GFP fusion
OsDRM2 mutation1 PCR	DRM2CACCF6	CACCATGGTGGACTGGGCTTCAG	GFP fusion
	DRM2miRjR4m1	CGGTAATCGTGTTGAGGGGCC	GFP fusion
OsDRM2 mutation2 PCR	DRM2CACCF6	CACCATGGTGGACTGGGCTTCAG	GFP fusion
	DRM2miRjR4m2	CGGTAATCGTGTTGAGGGACC	GFP fusion
OsDRM2 RNAi 1	DRM1a-PstIFw	CTGCAGAGCAATTGCTTGAGTTACTTC	RNAi construction
	DRM1a-XbaIRv	TCTAGACAAGTTTGCAAGCAGTTGAAC	RNAi construction
OsDRM2 RNAi 2	DRM1a-HindIIIFw	AAGCTTCAGAGCAATTGCTTGAGTTACTTC	RNAi construction
	DRM1a-SmaIApaIRv	GGGCCCGGGCAAGTTTGCAAGCAGTTGAAC	RNAi construction
pre-miR820	CACTApremiRJF2	TGATGAATATCCTTACCAATCTTG	sequencing
	CACTApremiRJR6	CATGTTTTGATCAMATGGCTAGCT	sequencing
OsDRM2	qDRM2F1	GATAGCGACAATGATAAGTTCGAG	sequencing
	qDRM2R3	TCCTCGTCGACTTCAAGAGTC	sequencing
pre-miR820	CACTApremiRJF1	CACATAAGAACAGACCATCTACAC	Probe amplification
	CACTApremiRJR3	TCTTTCATACTGCAATTGCGCTAG	Probe amplification
CACTA (ORF1)	HPR F	GCACAGAACCGATGTCACTAAC	McrBC-PCR
	HPR R	CTTCGCCAGGTCCGACATCTTC	McrBC-PCR
RIRE7 (LTR)	RIRE7 F1	AGGACATCCCTTCCAACGATACAAC	McrBC-PCR
	RIRE7 R1	TCTTGCCGTGCCAAGAACAACCTTG	McrBC-PCR
pre-miR820	MIRJ F	CACATAAGAACAGACCATCTACAC	McrBC-PCR
	MIRJ R	ATGTTTTGATCACATGGCTAGCTC	McrBC-PCR
OsActin	ActU	TCCATCTTGGCATCTCTCAG	McrBC-PCR
	ActL	GTACCCTCATCAGGCATCTG	McrBC-PCR
Centromere 8	Cen8-301 F1	CCGATATGCCAAAGAGCGAGTC	McrBC-PCR
	Cen8-301 R1	CAAATCATCTATCCTCAAGTCC	McrBC-PCR
pre-miR820	Satof	CTTASGTTCYTTRCYGATCTTG	Phylogeny
	R1SATO	TGACARCRTAYRAACTACACCTC	Phylogeny

Table 2.2. Sequencing analysis of *miR820* and its target site in *DRM2* among various *Oryza* species.

Acc. No.	Genome type	Species	*miR820* sequence identity *	*DRM2* sequence identity*	miRNA target score **
Nipponbare	AA	*O. sativa*	*miR820a/b/c* 22/22	22/22	2.5
			miR820d 21/22		3.5
			miR820e 20/22		5.5
W 0106	AA	*O. rufipogon*	22/22	22/22	2.5
W 1514	BB	*O. punctata (2X)*	19/22	17/22	2
W 1024	BBCC	*O. punctata (4X)*	19/22	17/22	2
W 1213	BBCC	*O. minuta*	19/22	17/22	2
W 1331	BBCC	*O. minuta*	19/22	17/22	2
W 1805	CC	*O. eichingeri*	22/22	20/22	2.5
W 0002	CC	*O. officinalis*	22/22	20/22	2.5
W 1830	CC	*O. officinalis*	22/22	20/22	2.5
W 1166	CCDD	*O. latifolia*	22/22	20/22	0.5
W 1197	CCDD	*O. latifolia*	22/22	20/22	0.5
W 0008	EE	*O. australiensis*	22/22	20/22	0.5
W 1401	FF	*O. brachyantha*	22/22	22/22	2.5
W 1711	FF	*O. brachyantha*	22/22	22/22	2.5
W 0003	GG	*O. granulata*	22/22	22/22	2.5
W 1220	HHJJ	*O. longiglumis*	22/22	22/22	2.5
W 0604	HHJJ	*O. ridleyi*	22/22	22/22	2.5

*The number of identical nucleotides of *miR820* and its target site in *DRM2*, respectively, between Nipponbare and various *Oryza* are shown, followed by dash and total length of each sequence.

**miRNA target score (Allen et al., 2005) indicates the degree of the mismatches as perfect match as zero.

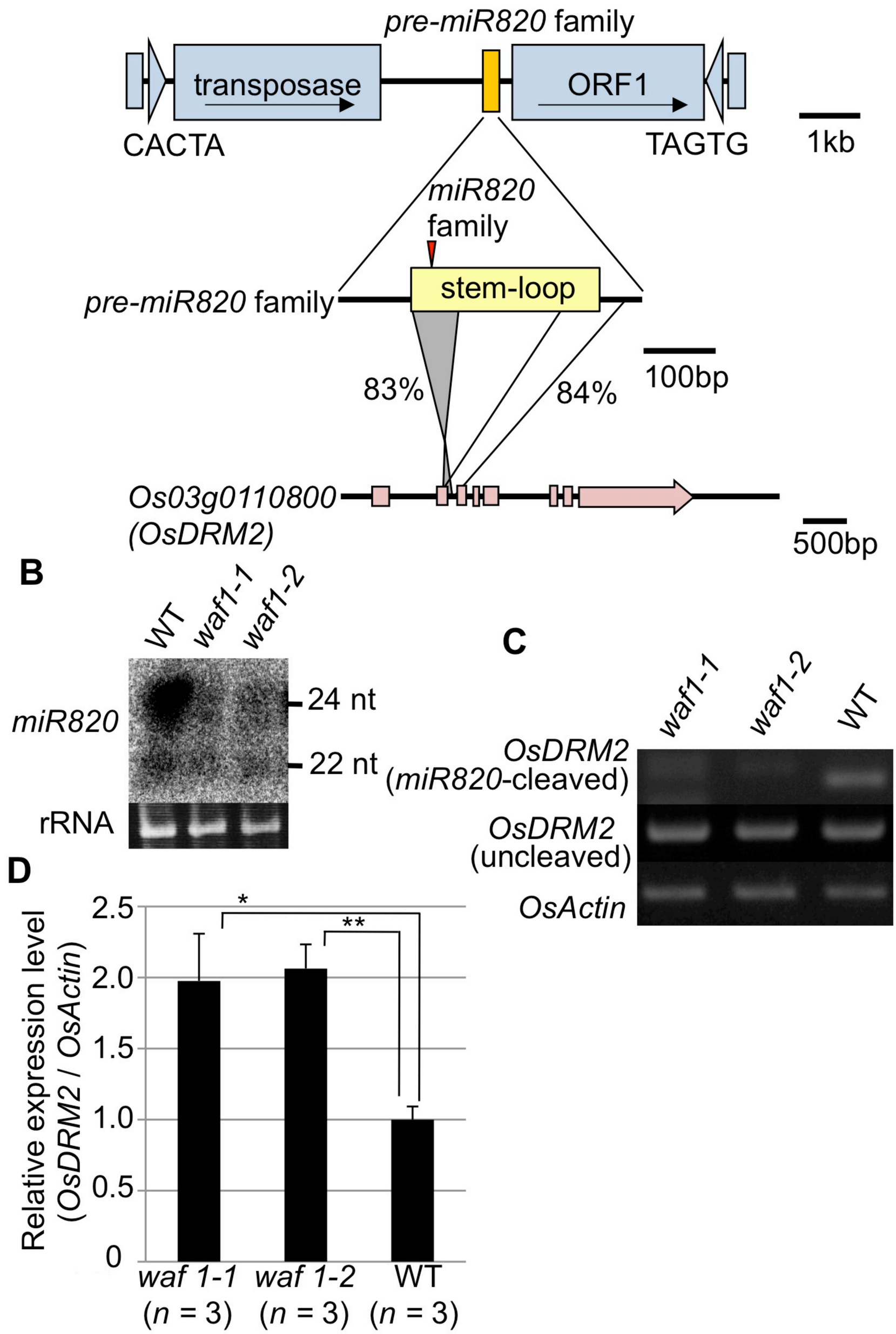
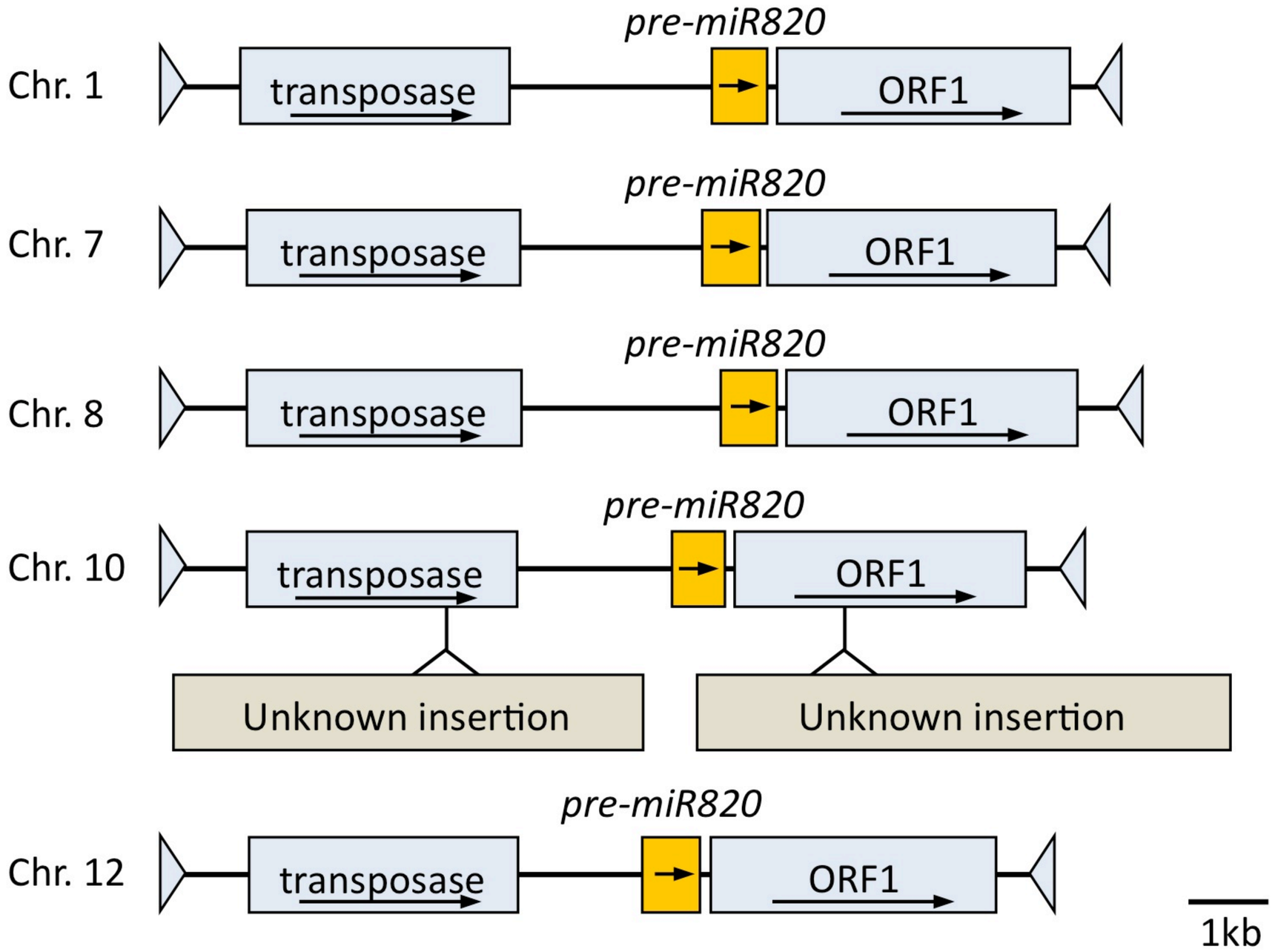
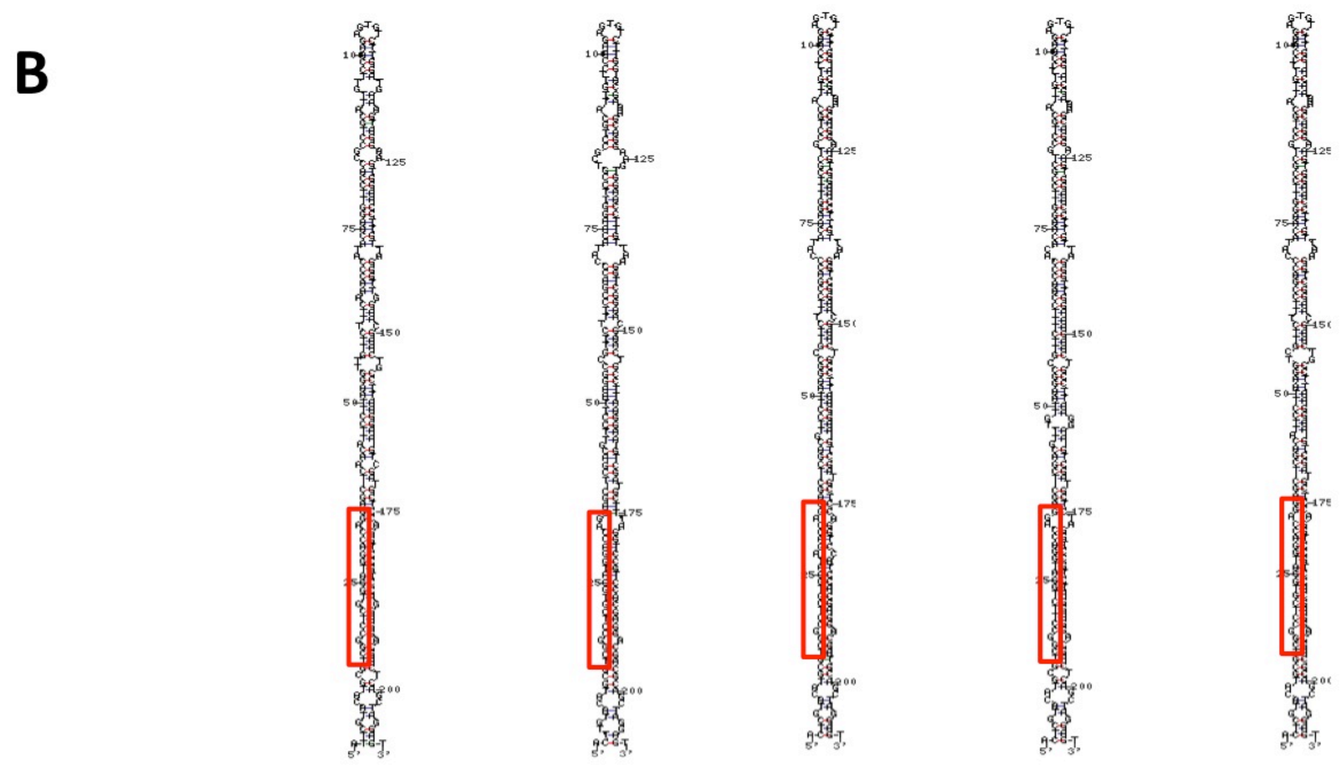


Figure 2.1. *miR820* family members are located within CACTA transposons and target the DNA methyltransferase gene *OsDRM2*.

(A) The location of *miR820* within the CACTA DNA transposon is shown in the first row. The second and third rows indicate the similarity between the sequences of the *miR820* precursor (*pre-miR820*) and *Os03g0110800* (*OsDRM2*). The numbers beside the lines indicate the nucleotide identities between the regions. The red triangle indicates the location of *miR820* within the stem-loop region. (B) Northern blot analysis of *miR820* expression in the wild-type (WT) and in *waf1* mutants. (C) Detection of *miR820*-cleaved *OsDRM2* mRNA by RNA ligation–mediated 5′ RACE (upper panel). The same cDNA templates were used for PCR to amplify *OsDRM2* (middle panel) and *OsActin* (bottom panel) as controls for *OsDRM2* expression and RNA integrity. (D) qRT-PCR analysis measuring the expression level of *OsDRM2* in *waf1-1*, *waf1-2*, and WT. The expression level of the WT was set as 1. Values are means, with bars showing standard errors. Significance was assessed by a two-tailed Student's *t*-test; (**) significant at the 1% level; (*) significant at the 5% level. The n value represents the number of mutant or wild-type individuals.





C

miR820a/b/c (Chr. 1, 7, 10)	1	UCGGCCUCGUGGAUGGACCAGG	22
miR820d (Chr.8)	1	. . . . . A	22
miR820e (Chr.12)	1	. . . . . UC	22

Figure 2.2. Structures and sequences of the five copies of *miR820* in Nipponbare rice. (A) Schematic representations of the structures of the five copies of CACTA transposons carrying *miR820*. (B) Stem-loop structures of the five copies of *pre-miR820* in Nipponbare, predicted by the mfold program. The *miR820* sequences are designated by red rectangles. (C) Sequence alignment of members of the *miR820* family, *miR820a–e,* in Nipponbare rice. Dots indicate identical nucleotides.

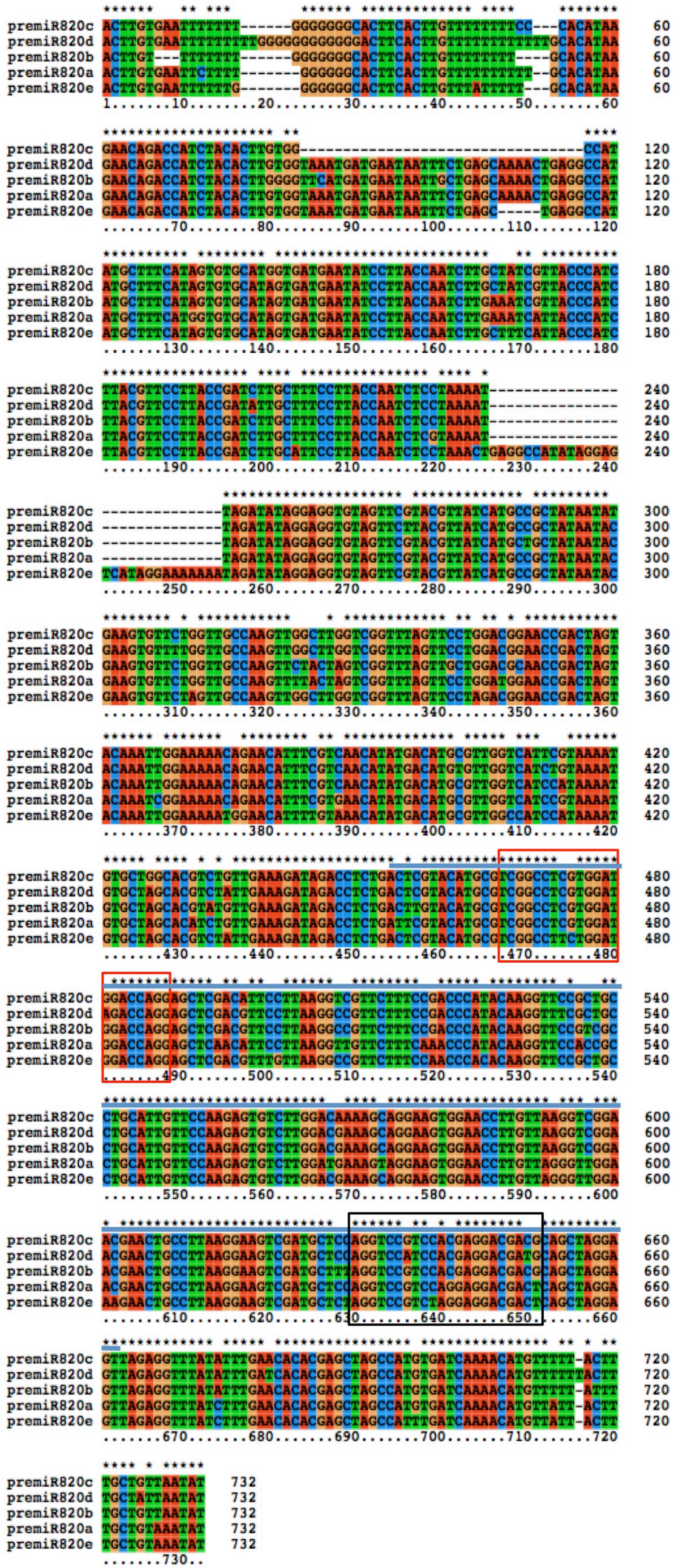


Figure 2.3. Multiple sequence alignment of the five copies of *pre-miR820* in the Nipponbare genome.

Sequence alignment was made using ClustalX at default settings. The blue lines, red box, and black box indicate the regions corresponding to the stem-loop structure, *miR820*, and *miR820**, respectively.

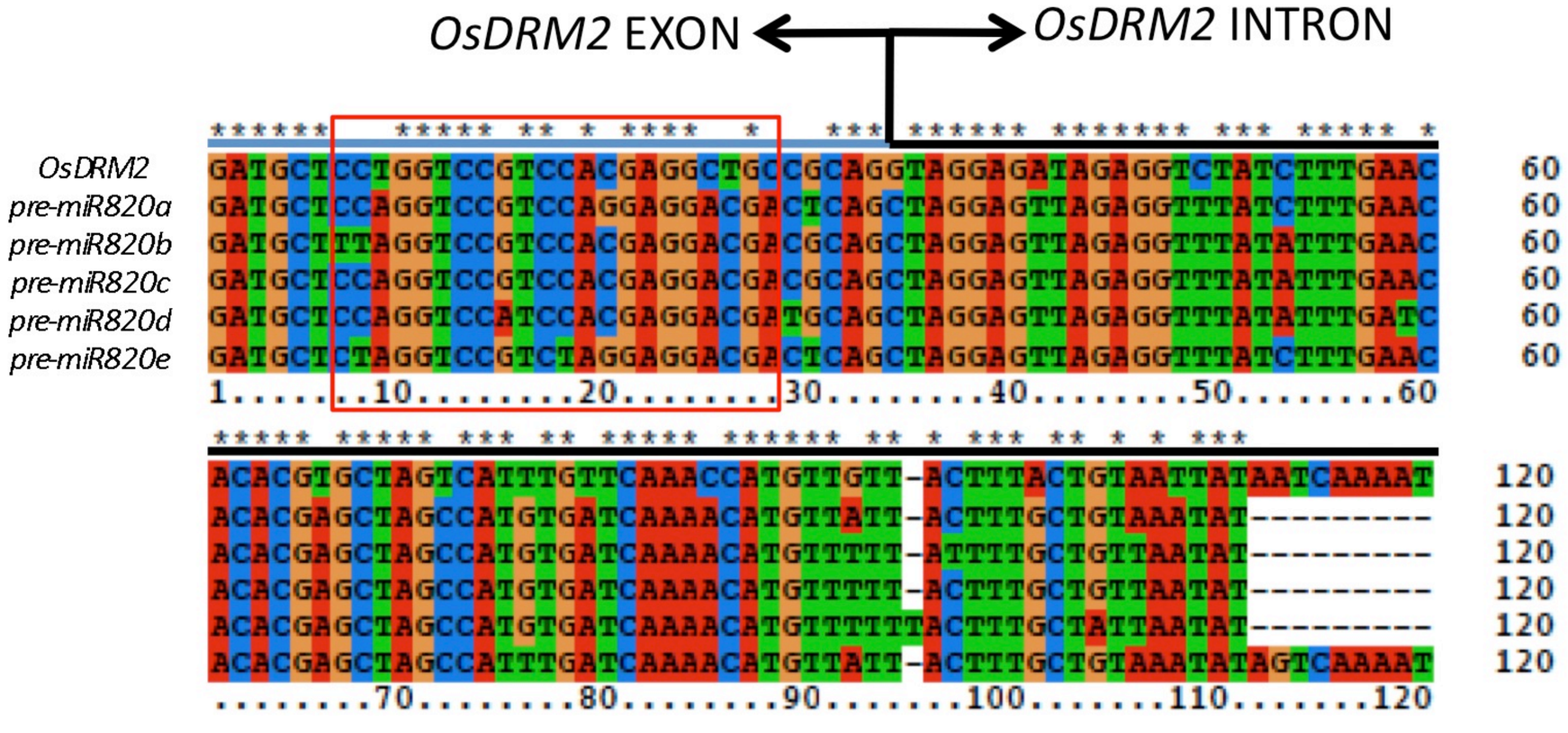
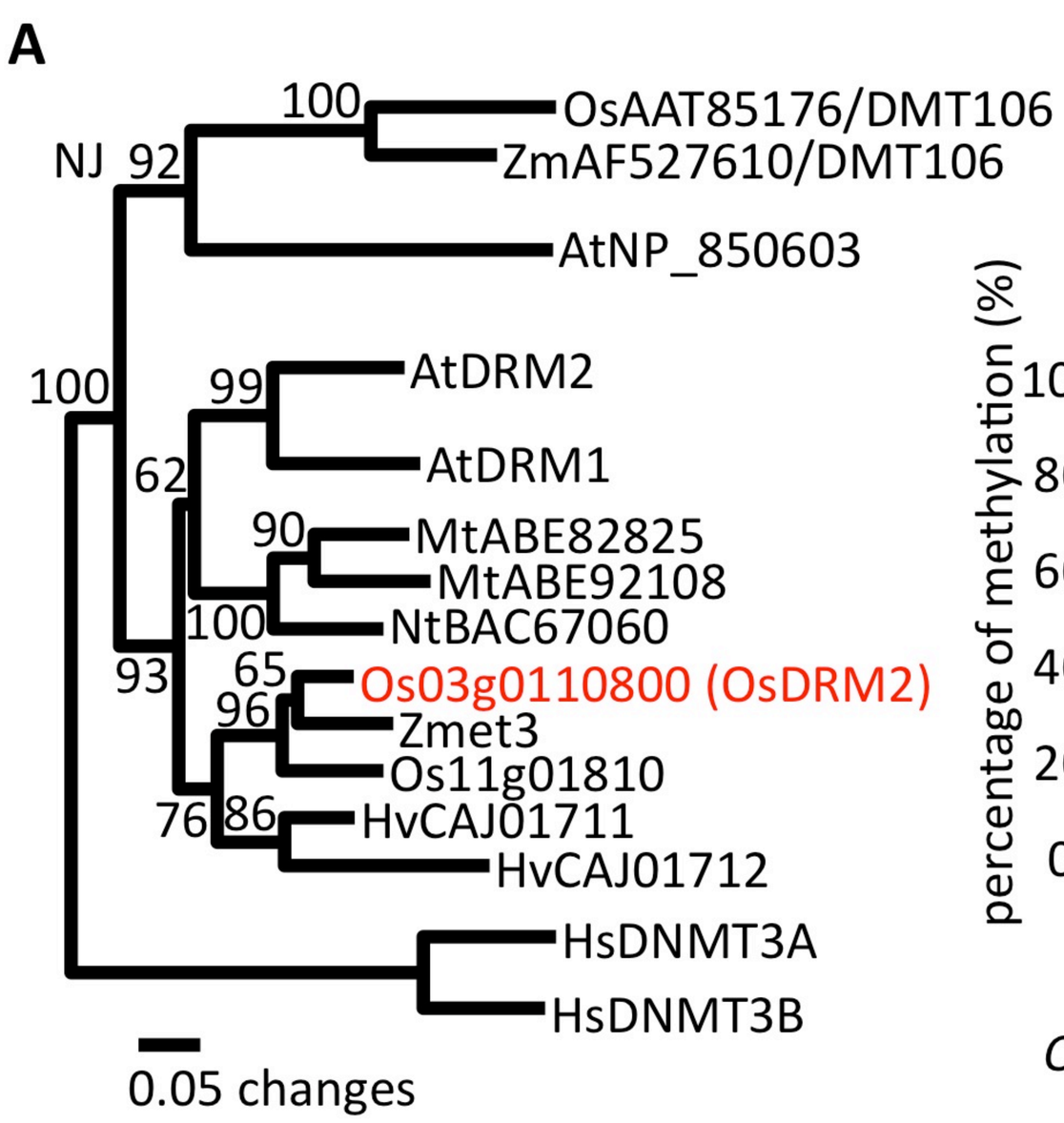
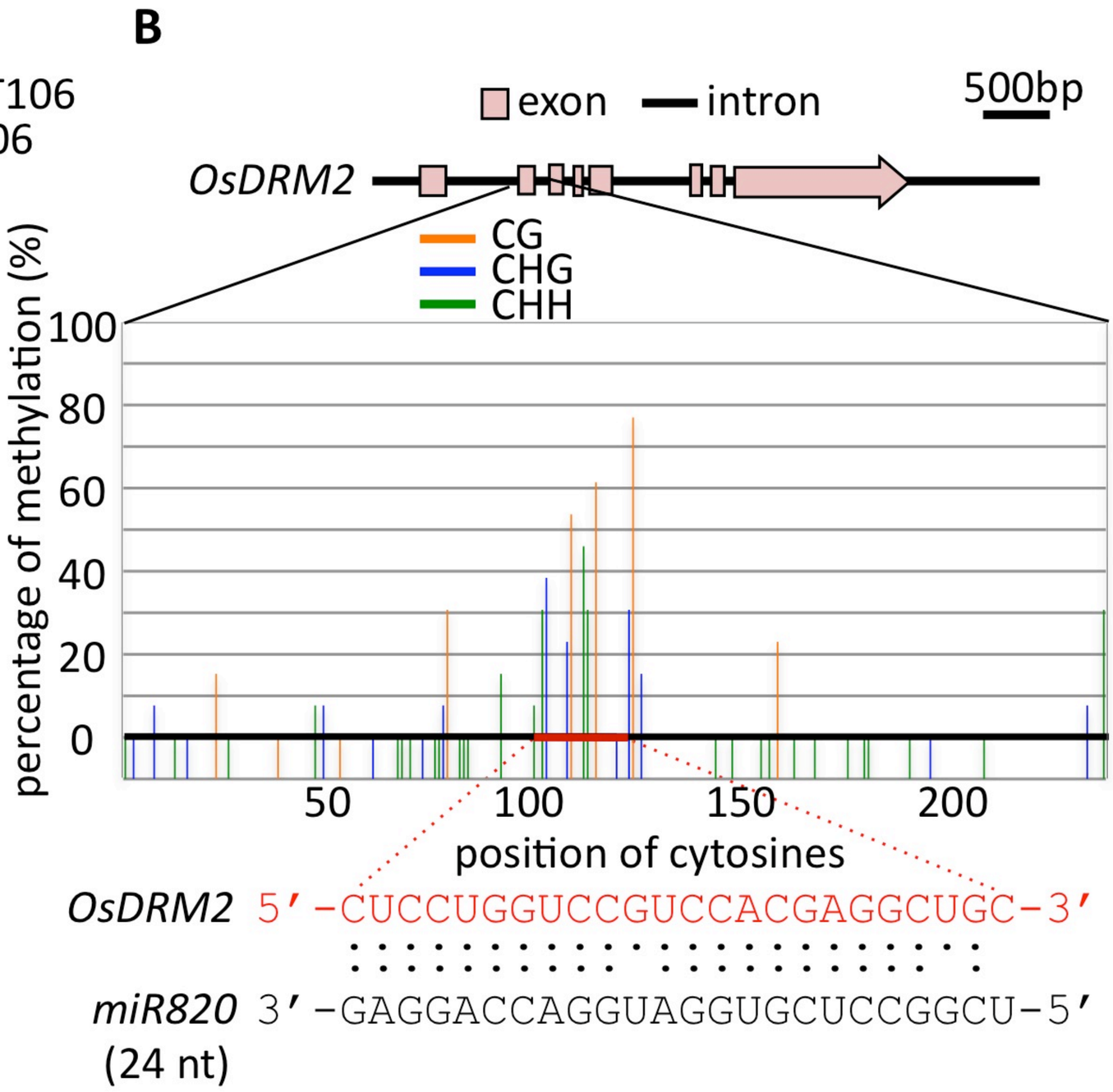
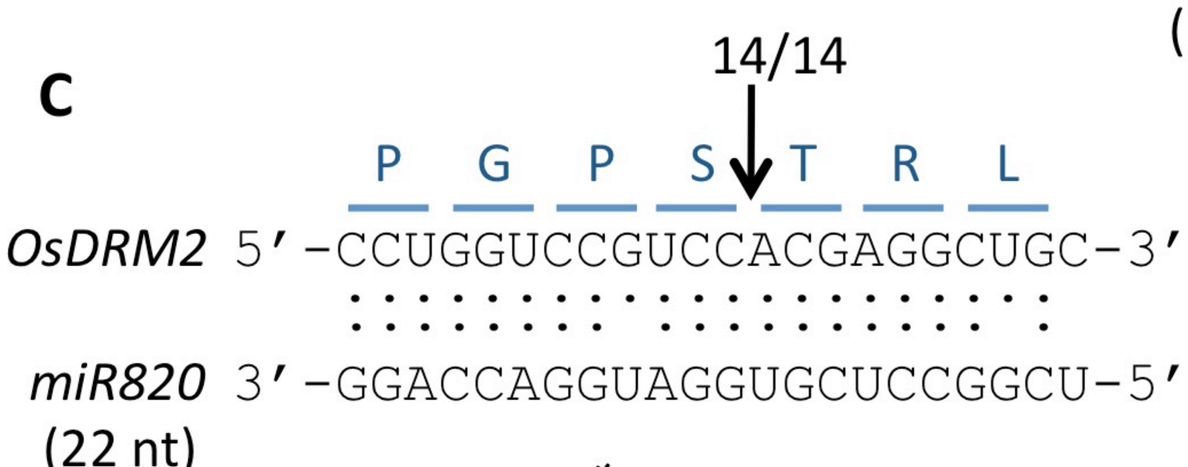
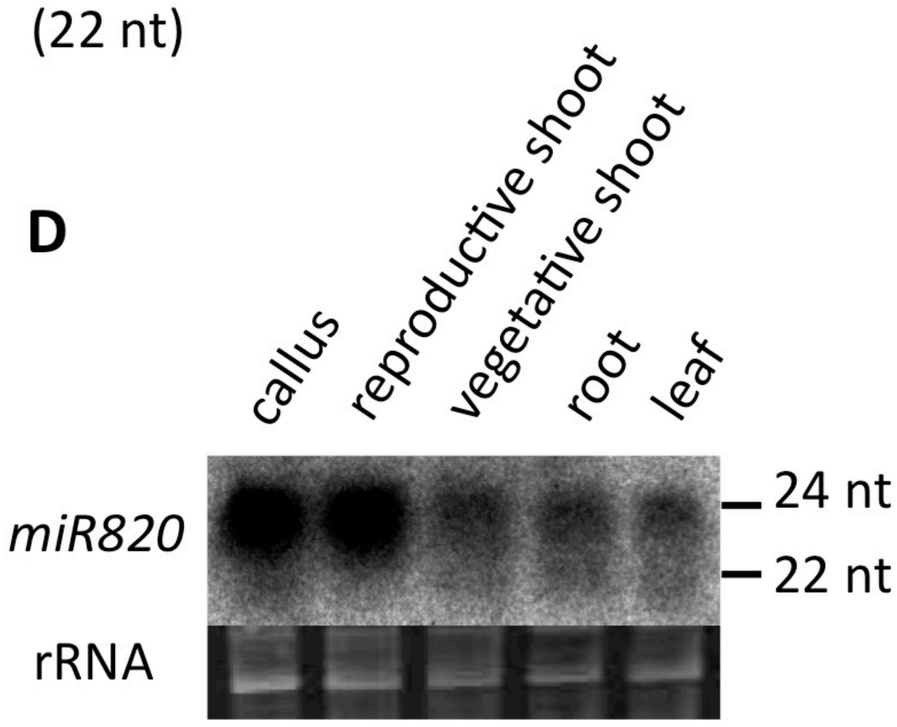


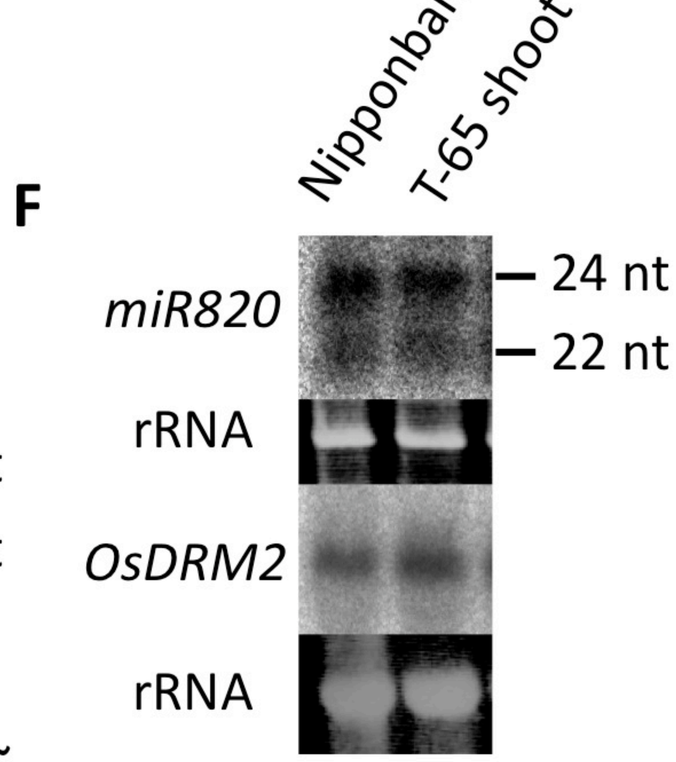
Figure 2.4. Multiple sequence alignment of the five copies of *pre-miR820* and a part of the *OsDRM2* sequence in the Nipponbare genome.
Sequence alignment was made using ClustalX at default settings. The blue line, black lines, and red box indicate the regions corresponding to the *OsDRM2* second exon, third intron, and *miR820**, respectively, in the *pre-miR820* sequences and the *miR820* recognition site in *OsDRM2*.

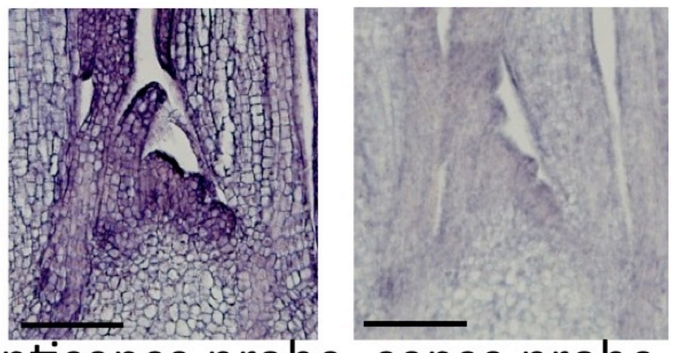

A
100
OsAAT85176/DMT106
NJ 92
ZmAF527610/DMT106
AtNP_850603
100
99
AtDRM2
AtDRM1
62
90
MtABE82825
MtABE92108
100
NtBAC67060
93
65
Os03g0110800 (OsDRM2)
96
Zmet3
Os11g01810
76
86
HvCAJ01711
HvCAJ01712
HsDNMT3A
HsDNMT3B
0.05 changes


B
exon
intron
500bp
OsDRM2
CG
CHG
CHH
percentage of methylation (%)
100
80
60
40
20
0
50
100
150
200
position of cytosines
OsDRM2 5′-CUCCUGGUCCGUCCACGAGGCUGC-3′
miR820 3′-GAGGACCAGGUAGGUGCUCCGGCU-5′
(24 nt)


C
14/14
P G P S T R L
OsDRM2 5′-CCUGGUCCGUCCACGAGGCUGC-3′
miR820 3′-GGACCAGGUAGGUGCUCCGGCU-5′
(22 nt)


D
callus
reproductive shoot
vegetative shoot
root
leaf
miR820
24 nt
22 nt
rRNA


F
Nipponbare
T-65 shoot
miR820
24 nt
22 nt
rRNA
OsDRM2
rRNA


G
antisense probe
sense probe

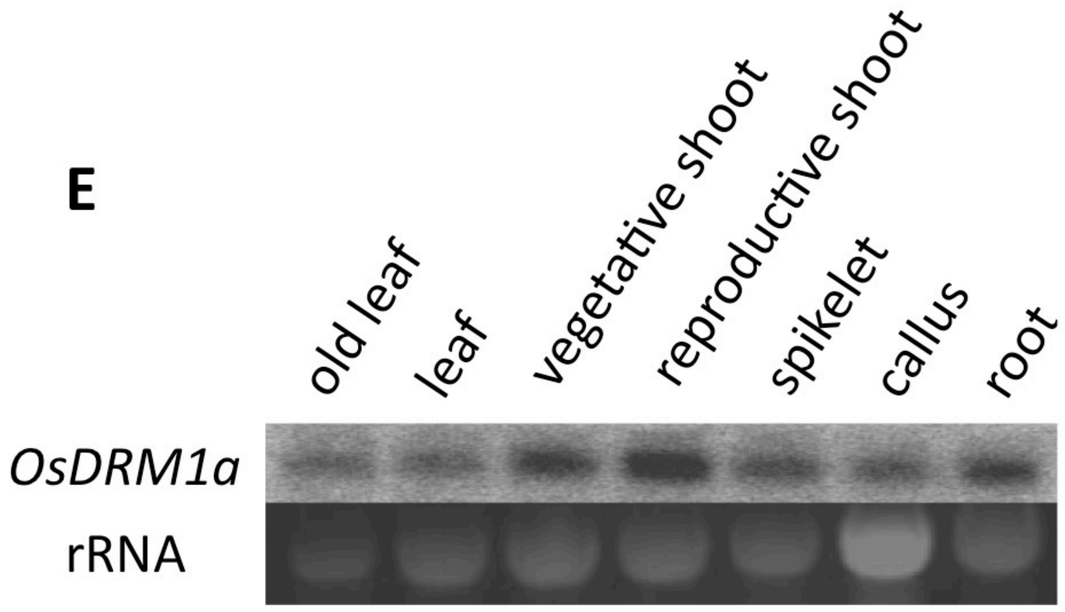

E
old leaf
leaf
vegetative shoot
reproductive shoot
spikelet
callus
root
OsDRM1a
rRNA

Figure 2.5. The target of *miR820* is *OsDRM2*.
(A) Phylogenetic tree of *de novo* DNA methyltransferases. Amino acid sequences in the Cyt-C5 DNA methylase domain were used for this analysis. The bootstrap values from 1000 replicates are indicated at each node. Os, rice; Z and Zm, maize; At, Arabidopsis; Nt, tobacco; Mt, Medicago; Hv, barley; Hs, human. (B) Analysis of DNA methylation status of *OsDRM2* in wild-type rice by bisulfite sequencing. The colored vertical lines above and below the bold black bar show the percentage of methylation and the position of individual cytosine sites, respectively. The target site of *miR820* is indicated by the red bar. (C) Mapping of the *miR820* cleavage site in *OsDRM2*. The arrow indicates the position of the cleaved end. The numbers above the arrow denote the number of clones ending at this position (left) and the total number of clones sequenced (right). (D) Northern blot analysis of *miR820* expression in various tissues. (E) Northern blot analysis of *OsDRM2* expression in various tissues. (F) Northern blot analysis of *miR820* and *OsDRM2* expression in vegetative shoots of two wild-type (WT) strains, Nipponbare and T-65. (G) *In situ* mRNA localization of *OsDRM2* in the vegetative shoot of Nipponbare using anti-sense probe (left panel) and the excess amount of sense probe as control (right panel), respectively. Bars = 50 um.

A

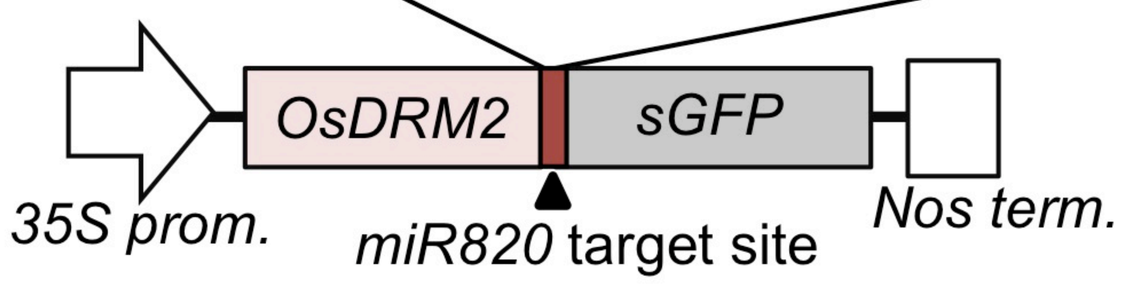
miR820 3′-GGACCAGGUAGGUGCUCCGGCU-5′

OsDRM2 intact 5′-CCUGGUCCGUCCACGAGGCUGC-3′

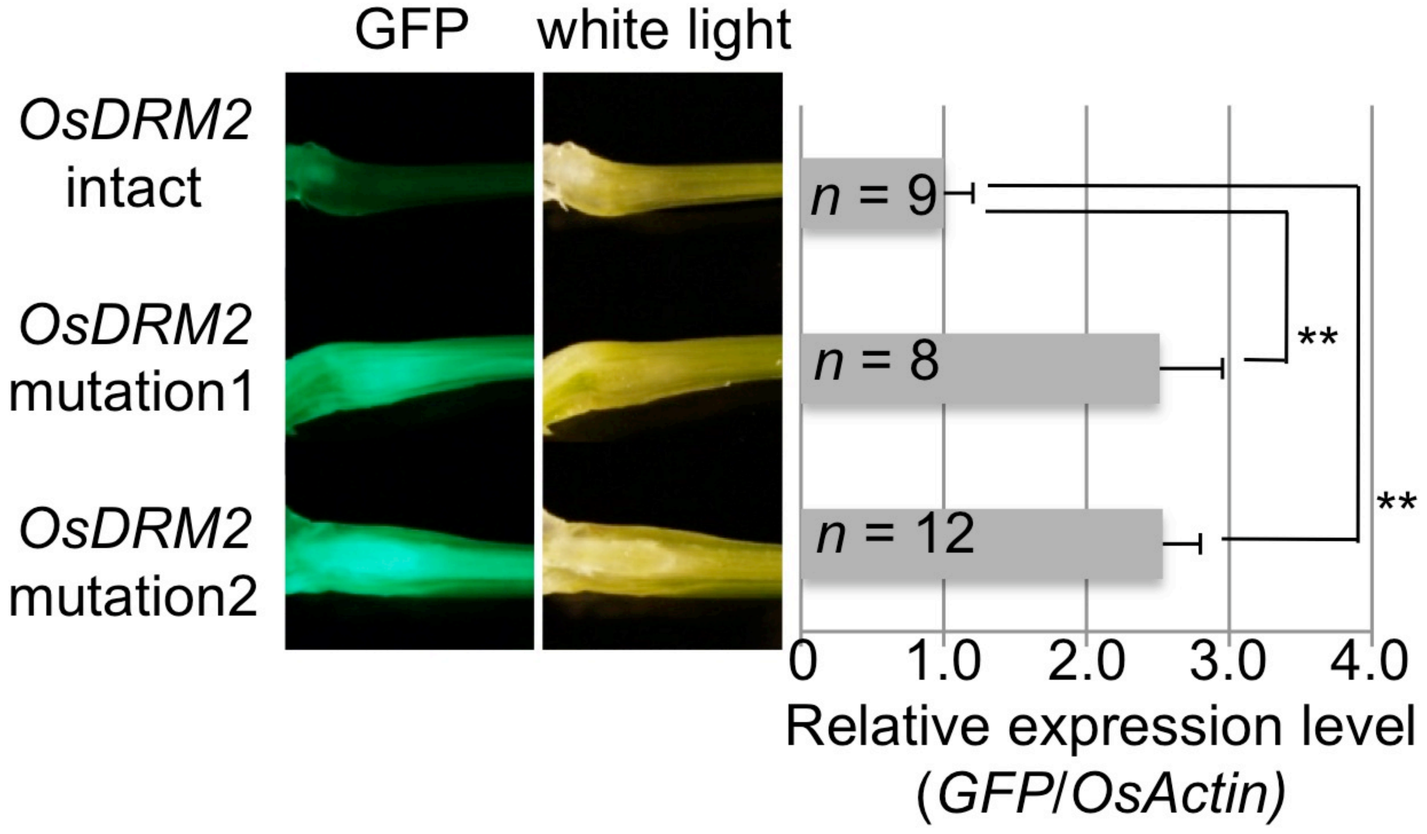
OsDRM2 mutation1 5′-CCGGGCCCCUCAACACGAUUAC-3′

OsDRM2 mutation2 5′-CCUGGUCCCUCAACACGAUUAC-3′

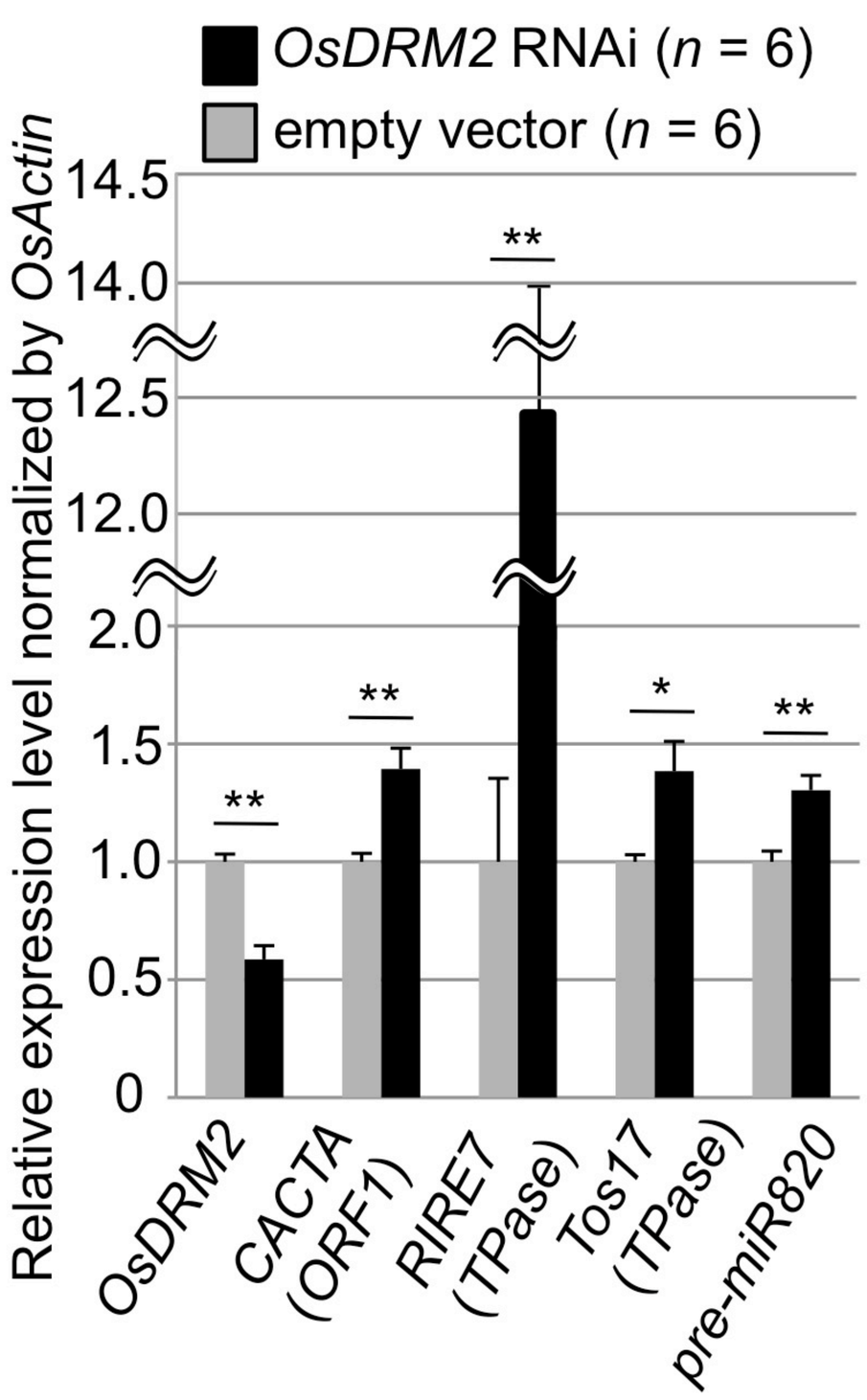
P G P S T R L



B



C



D

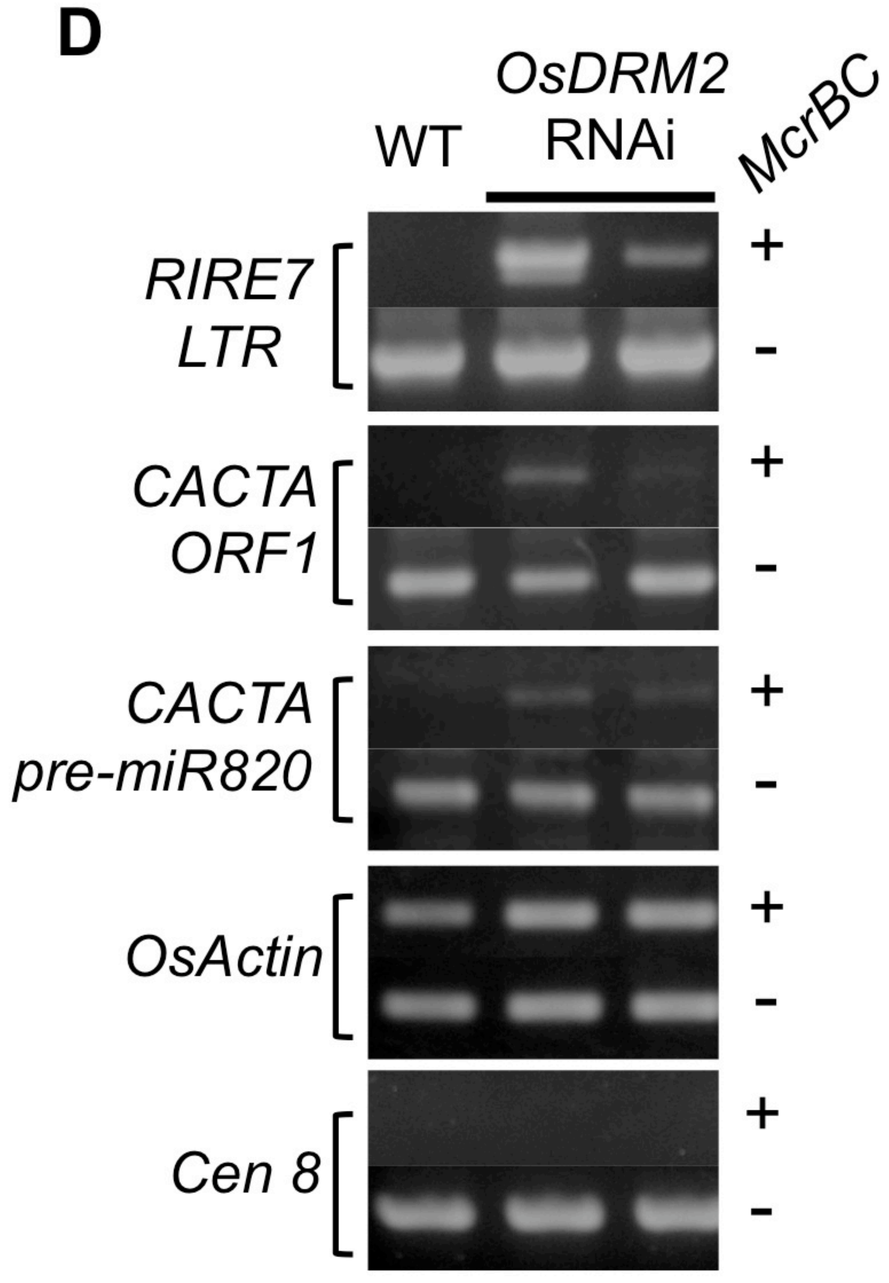


Figure 2.6. *OsDRM2* is negatively regulated by *miR820*.
(A) The general structure of the *OsDRM2::GFP* fusion constructs is shown at the bottom, and the sequences of *miR820a/b/c* and the target sites in the *35S:OsDRM2 intact:GFP*, *35S:OsDRM2 mutation1:GFP*, and *35S:OsDRM2 mutation2:GFP* constructs are shown at the top. The changed nucleotides in the mutant genes are shown in red letters. (B) The panels on the left show GFP fluorescence and white-light observations of longitudinal sections of shoots of transgenic plants transformed with the constructs in (A). The graph at the right shows relative expression levels of *GFP* mRNA in the corresponding transgenic lines as measured by quantitative RT-PCR. The expression level of *OsDRM2*-intact lines was set as 1. (C) Relative expression levels of *OsDRM2* and TEs in *OsDRM2* RNAi transgenic callus (black bars) measured by qRT-PCR. The expression level of empty-vector lines was set as 1. (D) *McrBC*-PCR analysis of genomic DNA from callus of WT and two independent transgenic lines of *OsDRM2* RNAi. Two of the six *OsDRM2* RNAi transgenic lines analyzed in (C) were used. In (B) and (C), values are means, with bars showing standard errors. Significance was assessed by a two-tailed Student's *t*-test; (**) significant at the 1% level; (*) significant at the 5% level. The n value represents the number of independent transformants.

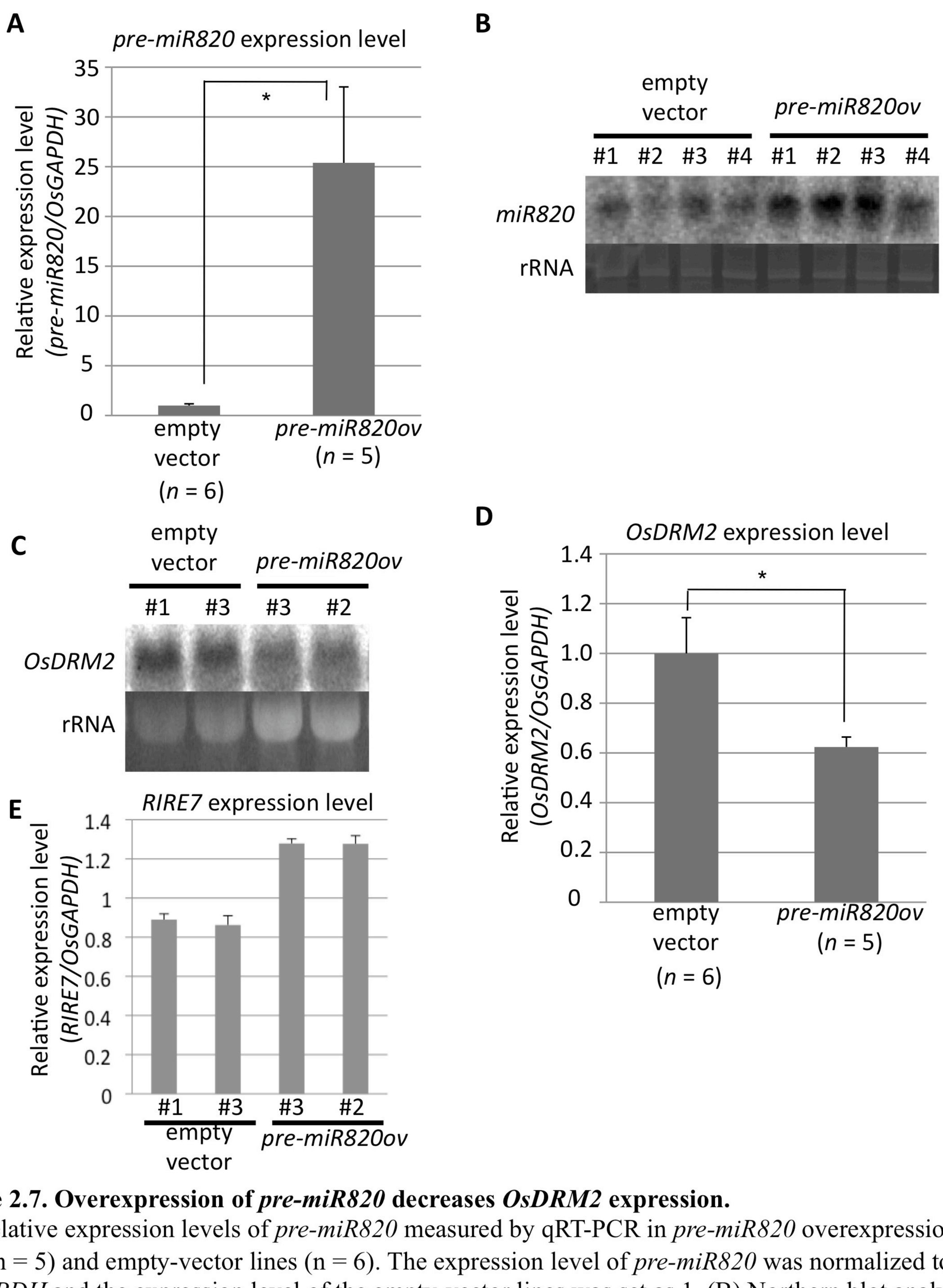


Figure 2.7. Overexpression of *pre-miR820* decreases *OsDRM2* expression.
(A) Relative expression levels of *pre-miR820* measured by qRT-PCR in *pre-miR820* overexpression lines (n = 5) and empty-vector lines (n = 6). The expression level of *pre-miR820* was normalized to *OsGAPDH* and the expression level of the empty-vector lines was set as 1. (B) Northern blot analysis showing increased *miR820* expression in two independent transgenic lines transformed with the *pre-miR820* overexpression construct compared to empty-vector controls. (C) Northern blot analysis showing decreased *OsDRM2* expression in two *pre-miR820* overexpression lines compared to empty-vector controls; line numbers correspond to those in (B). (D) Relative expression levels of *OsDRM2* measured by qRT-PCR in the same transgenic lines as in (A). The expression level of *OsDRM2* was normalized to *OsGAPDH*. The expression level of empty-vector lines was set as 1. (E) Relative expression levels of *RIRE7* measured by qRT-PCR in the same transgenic lines as in (C). The expression level of *RIRE7* was normalized to *OsGAPDH*. The expression level of empty-vector lines was set as 1. In (A), (D) and (E), values are means, with bars showing standard errors. In (A) and (D), significance was assessed by a two-tailed Student's *t*-test; (*) significant at the 5% level.

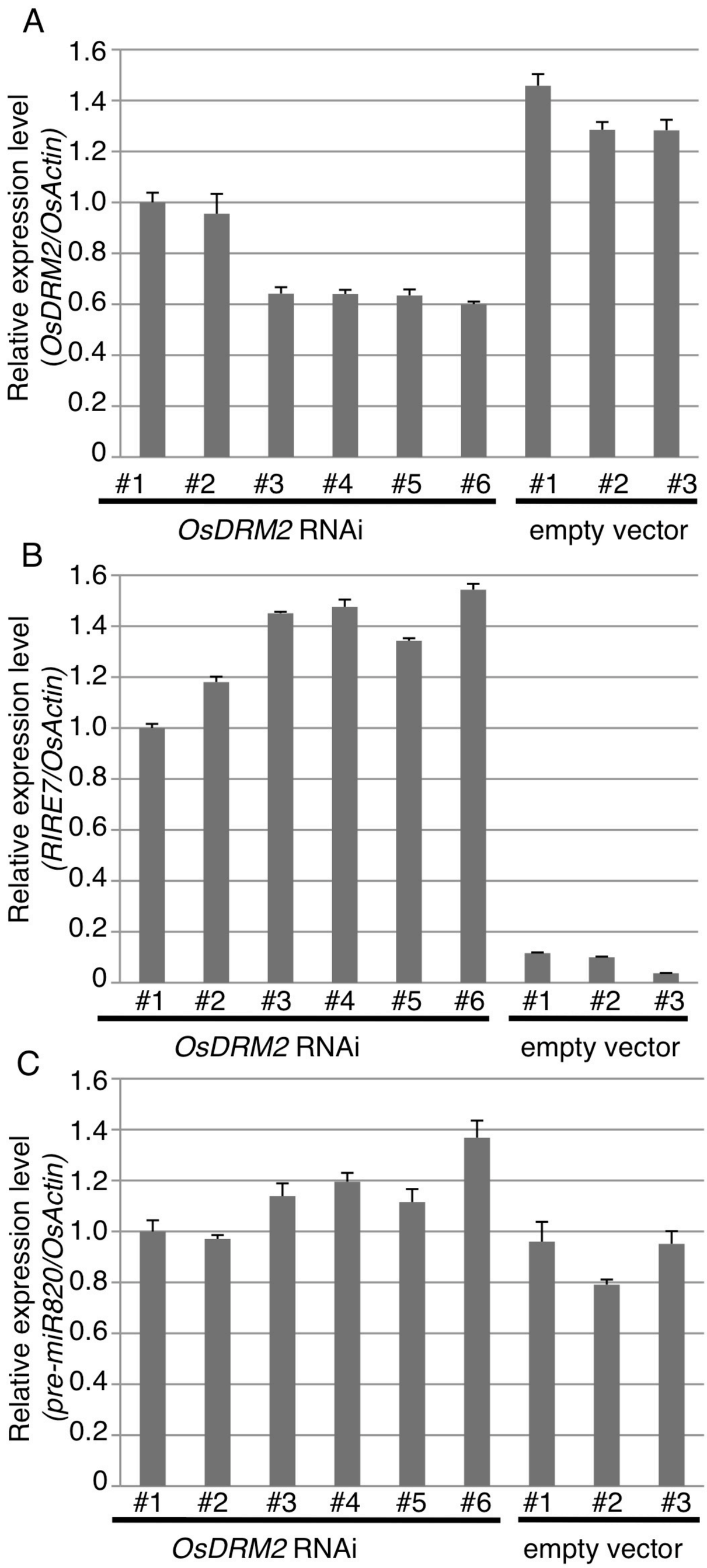


Figure 2.8. Decreased expression of *OsDRM2* by RNAi is associated with increased TE expression.

(A–C) Relative expression levels of *OsDRM2* (A), *RIRE7* (B), and the *pre-miR820* region of CACTA (C) in independent transgenic lines transformed with an *OsDRM2* RNAi construct (n = 6) or an empty vector (n = 3). Relative expression levels were measured by qRT-PCR and normalized to *OsActin*. Data shown are means of three technical replicates, with bars representing the standard errors. The relative expression level of *OsDRM2* RNAi #1 was set to 1.

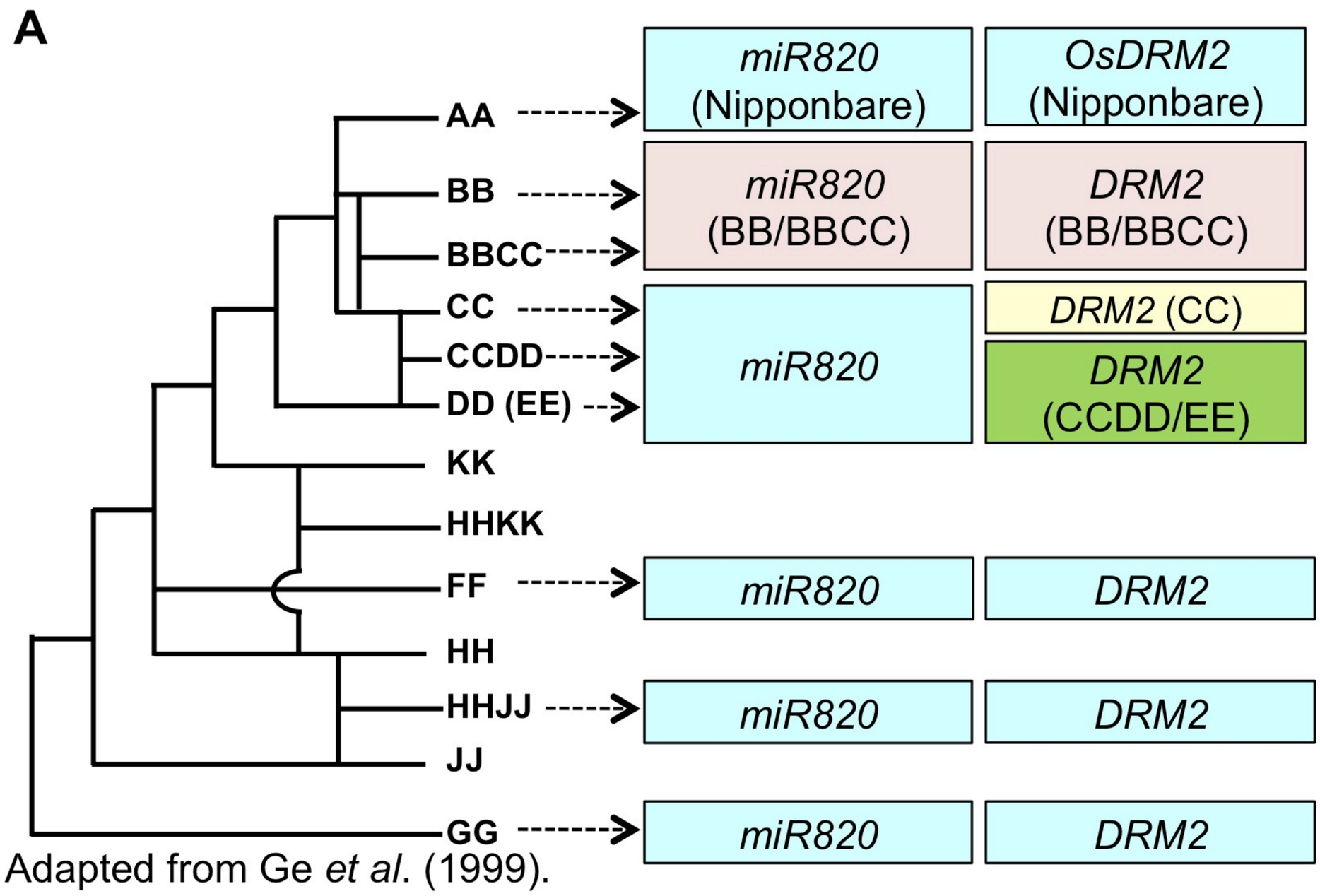
A

Acc. No.	Genome type	Species	Sequence (5'-3')
NB (*miR820a, b, c*)	AA	*O. sativa*	UCGGCCUCGUGGAUGGACCAGG
NB (*miR820d*)	AA	*O. sativa*	..............A.....
NB (*miR820e*)	AA	*O. sativa*	.......UC..........
W0106	AA	*O. rufipogon*	.....
W1514	BB	*O. punctata (2X)*	...AU...C.....
W1024	BBCC	*O. punctata (4X)*	...AU...C.....
W1213	BBCC	*O. minuta*	...AU...C.....
W1331	BBCC	*O. minuta*	...AU...C.....
W1805	CC	*O. eichingeri*	.....
W0002	CC	*O. officinalis*	.....
W1361	CC	*O. officinalis*	.....
W1830	CC	*O. officinalis*	.....
W0017	CCDD	*O. alta*	.....
W1182	CCDD	*O. alta* or *O. latifolia*	.....
W1166	CCDD	*O. latifolia*	.....
W1197	CCDD	*O. latifolia*	.....
W0613	CCDD	*O. grandiglumis*	.....
W1194	CCDD	*O. grandiglumis*	.....
W0008	EE	*O. australiensis*	.....
W1401	FF	*O. brachyantha*	.....
W1711	FF	*O. brachyantha*	.....
W0003	GG	*O. granulata*	.....
W0067 (B)	GG	*O. granulata*	.....
W1356	GG	*O. meyeriana*	.....
W1220	HHJJ	*O. longiglumis*	.....
W0001	HHJJ	*O. ridleyi*	.....
W0604	HHJJ	*O. ridleyi*	.....

B

Acc. No.	Genome type	Species	Sequence (5'-3')
Nipponbare	AA	*O. sativa*	CCUGGUCCGUCCACGAGGCUGC
W0106	AA	*O. rufipogon*	.....
W1514	BB	*O. punctata (2X)*	.....A.....AUC.A
W1024	BBCC	*O. punctata (4X)*	.....A.....AUC.A
W1213	BBCC	*O. minuta*	.....A.....AUC.A
W1331	BBCC	*O. minuta*	.....A.....AUC.A
W1527	CC	*O. eichingeri*	.....U..A
W1805	CC	*O. eichingeri*	.....U..A
W0002	CC	*O. officinalis*	.....U..A
W1830	CC	*O. officinalis*	.....U..A
W1166	CCDD	*O. latifolia*	.....C.A
W1197	CCDD	*O. latifolia*	.....C.A
W2200	CCDD	*O. latifolia*	.....U.....C.A
W2220	CCDD	*O. grandiglumis*	.....C.A
W0008	EE	*O. australiensis*	.....C.A
W1628	EE	*O. australiensis*	.....C.A
W1401	FF	*O. brachyantha*	.....
W1711	FF	*O. brachyantha*	.....
W0003	GG	*O. granulata*	.....
W1220	HHJJ	*O. longiglumis*	.....
W0604	HHJJ	*O. ridleyi*	.....

Figure 2.9. Sequence alignments of *miR820* and its target site in *DRM2* among *Oryza* species.
(A) Alignment of sequences of *miR820* among various *Oryza* species. (B) Alignment of sequences of the *miR820* target site in *DRM2* among various *Oryza* species. Dots indicate nucleotides identical to those in Nipponbare *miR820a/b/c* (A) or *DRM2* (B).



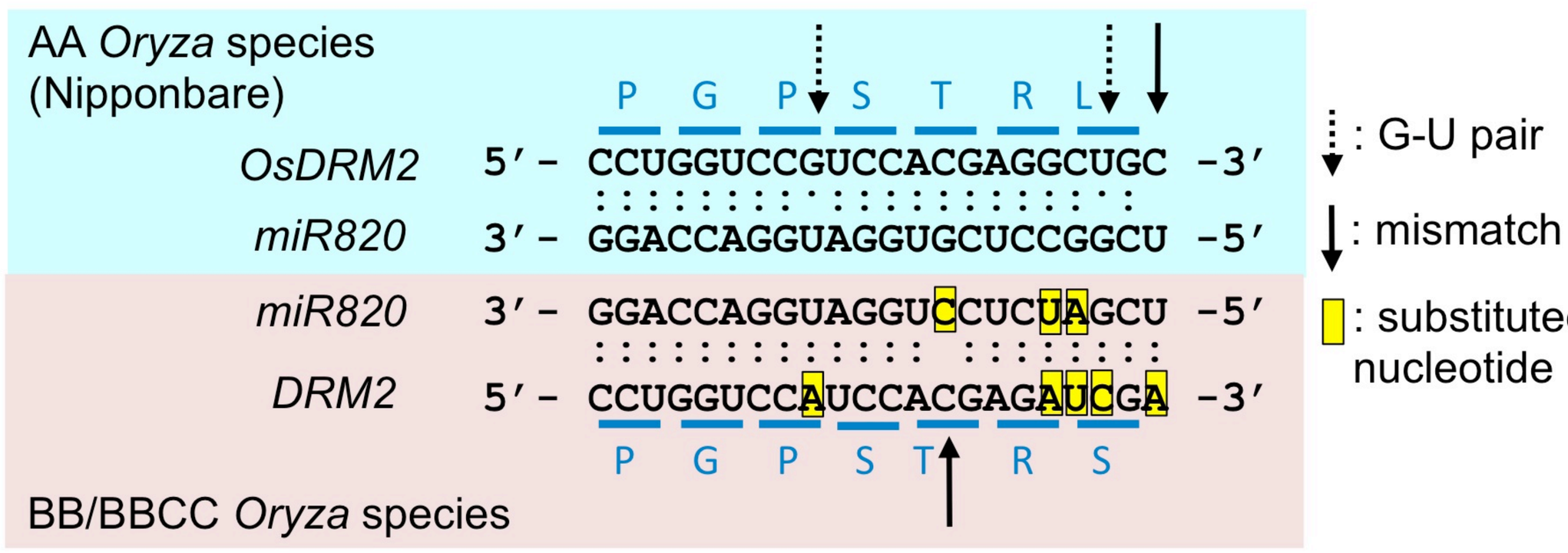
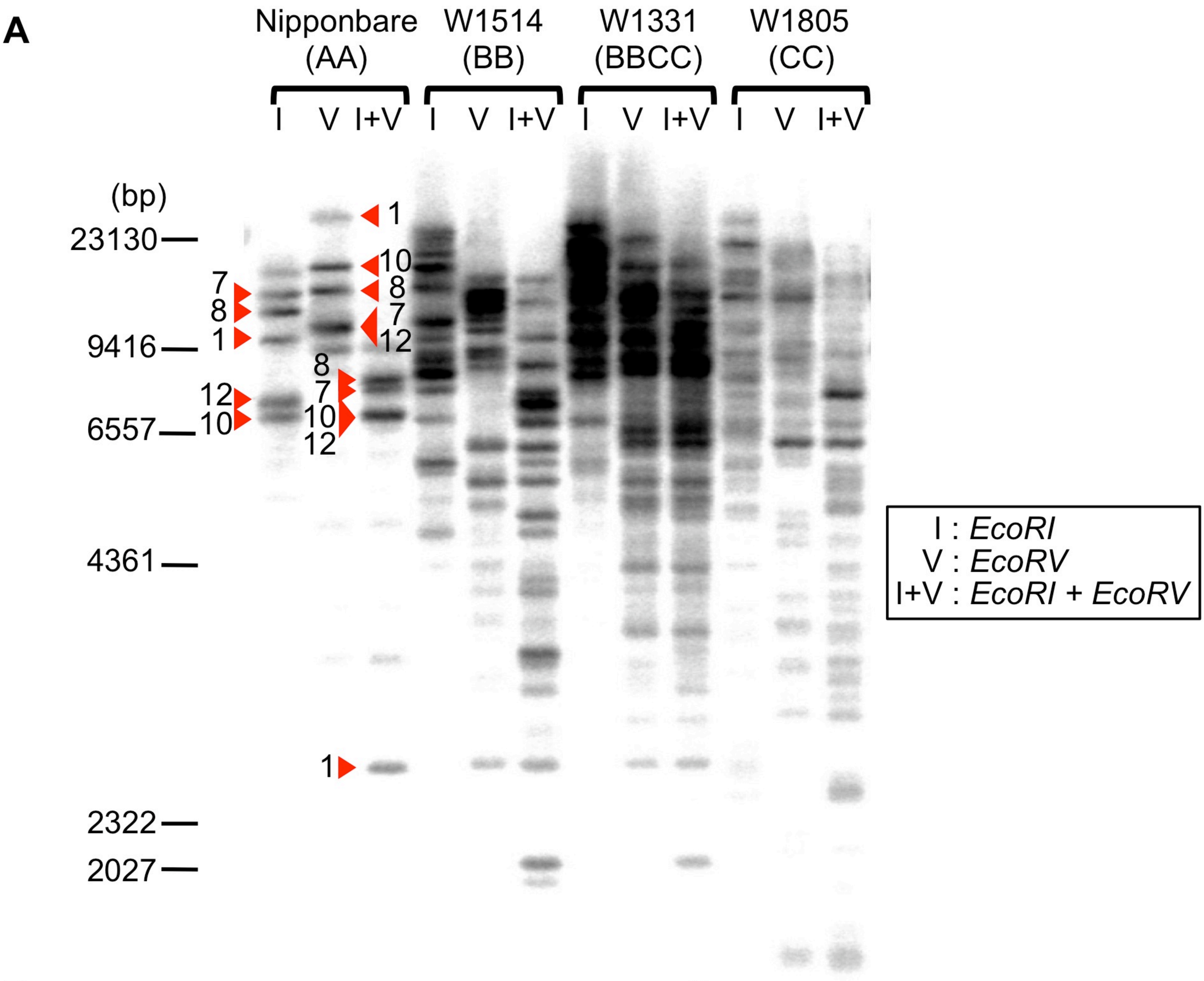
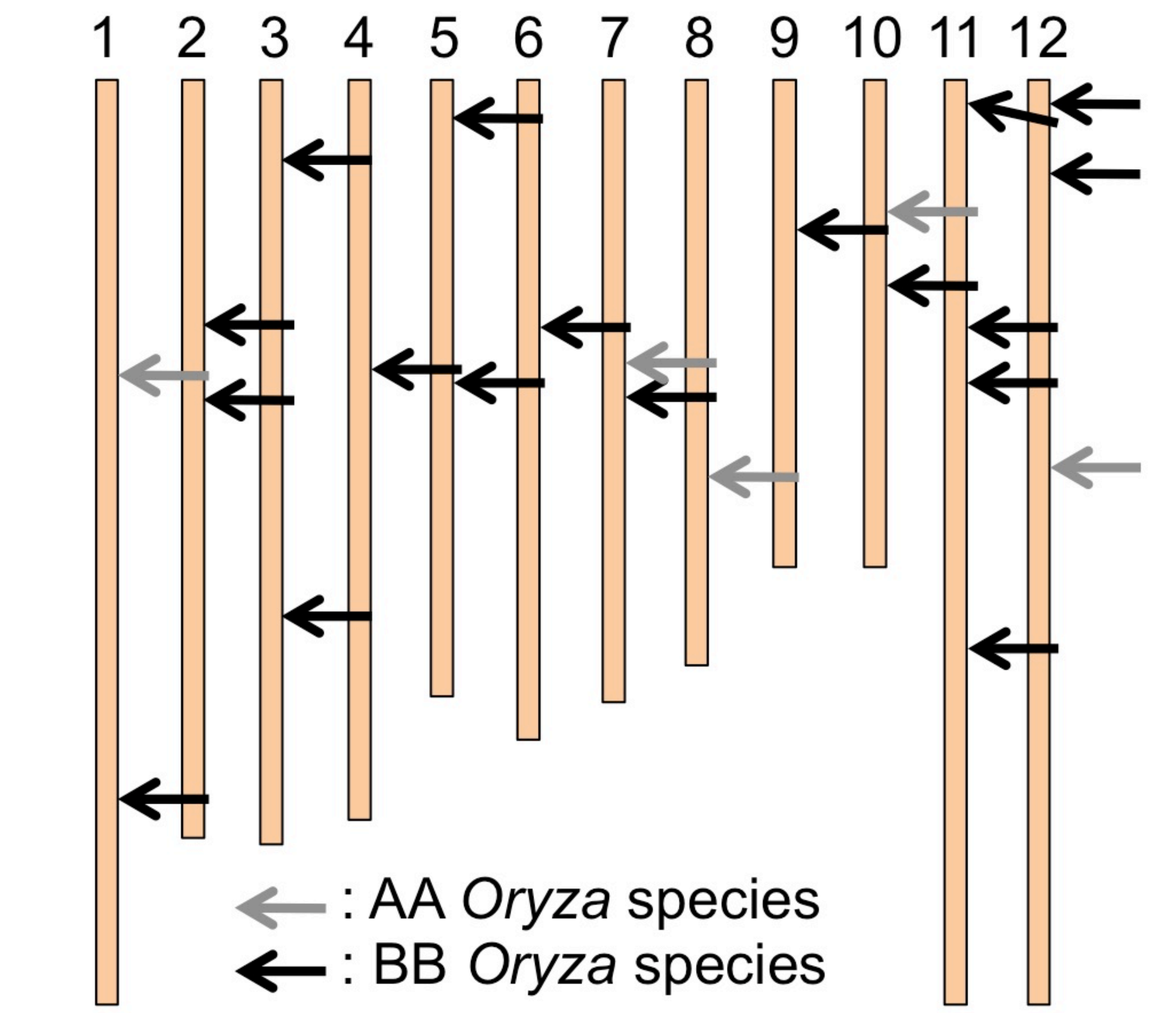


Figure 2.10. Regulation of *DRM2* by *miR820* is conserved among *Oryza* species.
(A) The phylogenetic tree shows the evolutionary relationships between *Oryza* species (left). Boxes indicate the types of *miR820* sequences and *DRM2* target site sequences identified in each genome (right). Within each column, boxes of the same color indicate identical sequences. The genome origins of the sequences in the boxes are indicated by arrows. (B) Sequence alignments of *miR820* and its target site in *DRM2* in the AA (Nipponbare; blue box) and BB/BBCC genomes (pink box). Dots between nucleotides indicate the type of nucleotide pair: a double dot indicates an A-U or G-C pair, a single dot indicates a G-U pair, and no dot indicates a mismatch. The positions of G-U pairs and mismatches are shown with broken and solid arrows, respectively. The eight nucleotide substitutions found between AA and BB/BBCC *Oryza* species are highlighted in yellow. Blue lines above and below the sequence of *DRM2* indicate the codons. Letters above and below the lines indicate the amino acids encoded by *DRM2* in AA and BB/BBCC *Oryza* species, respectively. Phylogenetic tree in (A) adapted from Ge et al. (1999).

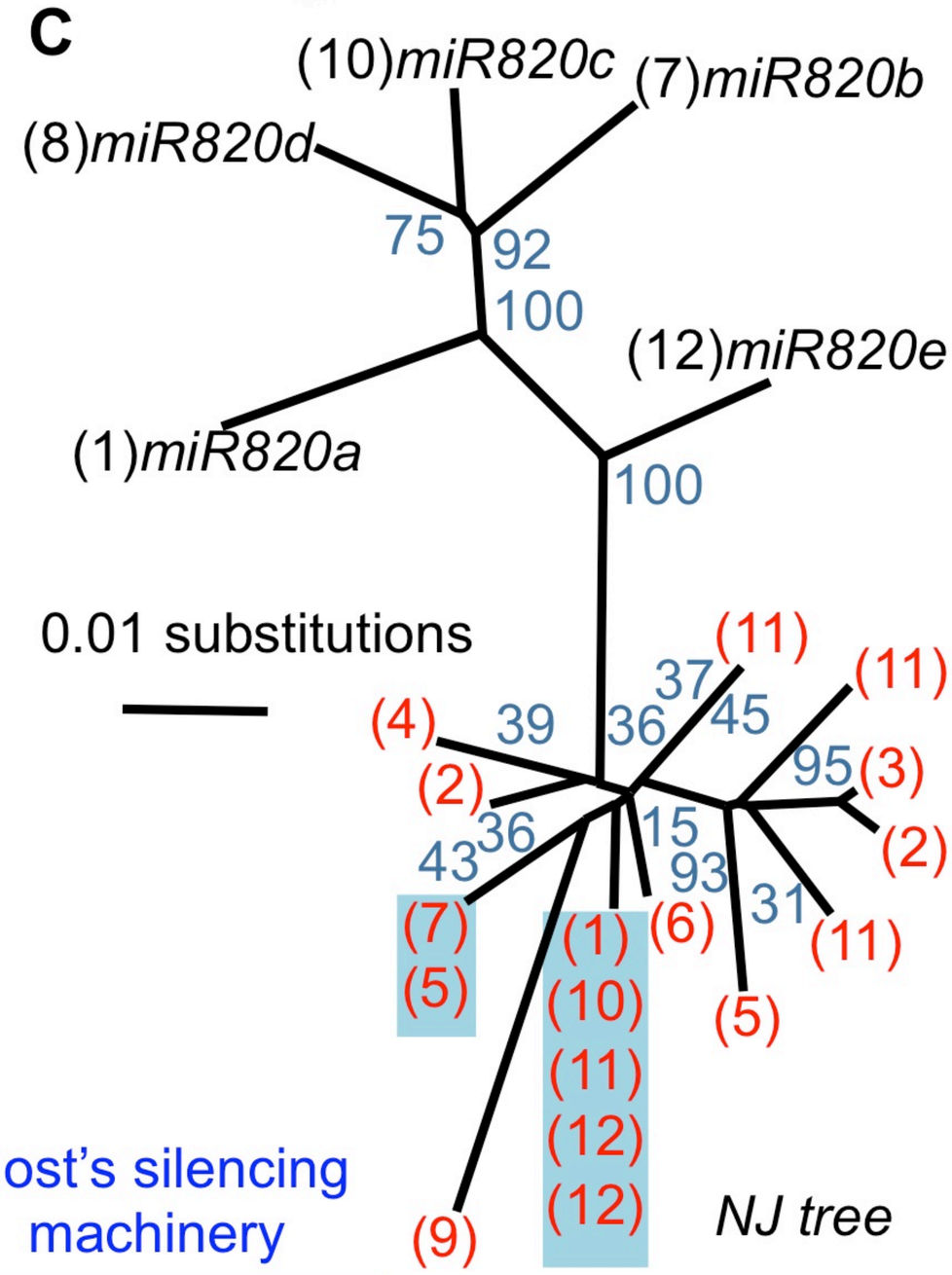
A


Nipponbare (AA)
W1514 (BB)
W1331 (BBCC)
W1805 (CC)
I V I+V I V I+V I V I+V I V I+V
(bp)
23130
9416
6557
4361
2322
2027
I : EcoRI
V : EcoRV
I+V : EcoRI + EcoRV

B


1 2 3 4 5 6 7 8 9 10 11 12
: AA Oryza species
: BB Oryza species

C


(10)miR820c
(7)miR820b
(8)miR820d
(12)miR820e
(1)miR820a
0.01 substitutions
NJ tree

D

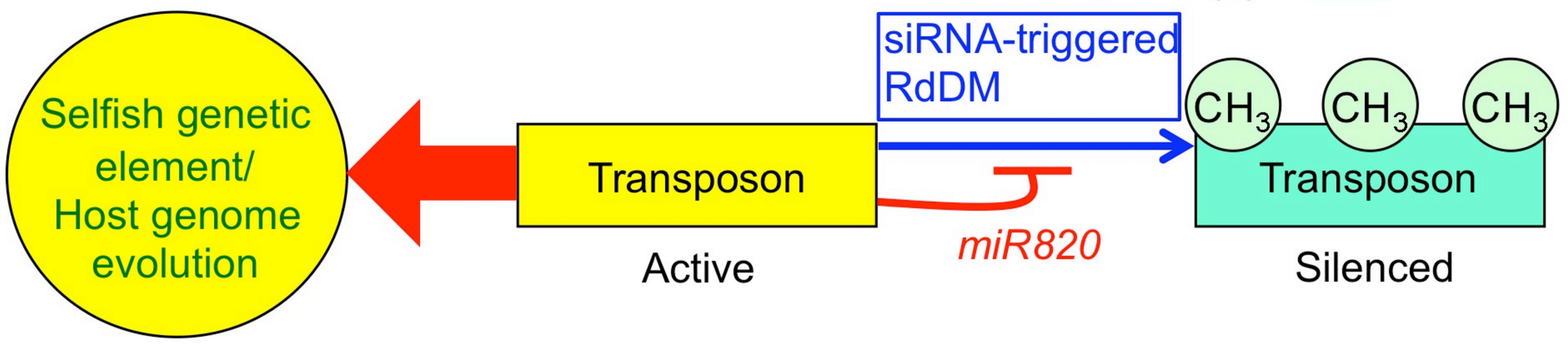

Host's silencing machinery
siRNA-triggered RdDM
Selfish genetic element/ Host genome evolution
Transposon
Active
miR820
CH3 CH3 CH3
Transposon
Silenced

Figure 2.11. Increased copy number of CACTA carrying *pre-miR820* in the BB/BBCC genome.

(A) Detection of CACTA TEs carrying *pre-miR820* by Southern blot analysis. Genomic DNA from AA, BB, BBCC, and CC *Oryza* species were digested with the enzymes indicated and probed with *pre-miR820*. Red triangles indicate the bands corresponding to the five copies of *pre-miR820* in Nipponbare. The number next to each triangle indicates the chromosome location of that copy. (B) Mapping of CACTA carrying *pre-miR820* in the rice genome. The genomic locations of CACTA carrying *pre-miR820* in AA and BB *Oryza* species are shown by gray arrows and black arrows, respectively. (C) Phylogenetic analysis of *miR820*-CACTA sequences. Bootstrap values (1000 replicates) are given for branch nodes. Black and red numbers in parentheses indicate the chromosome locations of *pre-miR820* sequences in the AA and BB genomes, respectively. Multiple copies on one branch, indicating that the identical sequence was found at multiple loci, are highlighted in blue. (D) A model for the regulation of *DRM2* by *miR820*. Active transposons behave as "selfish" genetic elements. This characteristic is counteracted by the host's silencing machinery (blue arrow), which acts to methylate and silence transposon loci. This action can be blocked by *miR820* (thin red line), which suppresses the host's silencing machinery and can drive host genome evolution (thick red arrow).

CHAPTER 3

Expression of a rice microRNA, *miR820*, is associated with the epigenetic modifications at its own locus.

Abstract

Small RNAs, such as small interfering RNAs (siRNAs) or microRNAs (miRNAs), regulate gene expression at transcriptional and posttranscriptional levels in eukaryotes. miRNAs are processed from duplexes formed on single-stranded RNA. They regulate expression of their target gene either by cleaving mRNA or supressing translation. In general, the primary miRNA transcripts are synthesized by RNA polymerase II and processed similarly to mRNAs. *MIRNA* genes are usually located in transcriptionally active euchromatic regions. In contrast, siRNAs are processed from duplexes made of two RNA molecules. One of them is often derived from a transposable element (TE) or from repetitive sequences that reside in heterochromatic regions. The other strand is synthesized by the RNA-dependent RNA polymerase on the first strand as a template. siRNAs establish epigenetic marks in parasitic DNA such as TEs, thus they usually act in *cis*. The rice miRNA *miR820,* encoded by CACTA TEs (five copies, located on different chromosomes), reduces the expression of the *de novo* DNA methyltransferase gene *OsDRM2.* Because *miR820* is derived from silent TEs, in which the heterochromatic histone modifications are enriched, the mechanism of *MIR820* transcription could be expected to differ from typical miRNAs. Here I show that the primary transcript of *MIR820* is mainly derived from the CACTA TE copy on chromosome 7. Histone modification and DNA methylation status around the *MIR820* region on

chromosome 7 differed from that of the other four loci. These unique epigenetic modifications in *MIR820* on chromosome 7 were only found around the *miR820* coding region. I conclude that *MIR820* transcription may depend on the unique epigenetic modifications, which in turn may be established by the action of *miR820* in *cis*. This suggests a dual function of *miR820* in *cis* and in *trans*.

Introduction

microRNAs (miRNAs) and small interfering RNAs (siRNAs) are small RNAs that regulate gene expression at both transcriptional and post-transcriptional levels in eukaryotes (Carthew and Sontheimer, 2009). The small RNAs are essential for plant development, environmental responses, and defense against genomic parasites such as transposable elements (TEs) and viruses (Plasterk, 2002; Almeida and Allshire, 2005; Aravin et al., 2007).

miRNAs are 21- or 22-nt-long RNA molecules processed from a duplex (formed on a single-stranded RNA) by Dicer-like protein 1 (DCL1) in plants (Kurihara and Watanabe, 2004; Kurihara et al, 2006). miRNAs selectively recognize their target mRNAs based on the base pairing, and cleave them or suppress translation (Voinnet, 2009; Sun, 2012). miRNA primary transcripts (pri-miRNAs) are synthesized by RNA polymerase II, and are processed into capped, spliced, and polyadenylated pri-miRNAs (which is similar to mRNA processing) (Lee at al., 2004). Actively transcribed *MIRNA* genes are usually located in euchromatic regions (similar to protein coding genes).

siRNAs (usually 20–30 nt long) originate from TEs, repetitive regions, other intergenic regions or transgenes. The initial siRNA transcripts are synthesized by RNA polymerase IV, a member of the plant-specific DNA-dependent RNA polymerases. These transcripts are converted into double-stranded RNA either by RNA-dependent RNA polymerases or by transcription from both strands, and are processed by DCL2–4 (Vazquez et al., 2008).

siRNAs are responsible for defense against genomic parasites such as TEs or viruses (Voinnet, 2008). siRNAs establish epigenetic marks (such as DNA cytosine methylation) in the parasitic DNA from which they originate, thus they usually act in *cis*. In plants, TE-derived siRNAs are loaded to Argonaute proteins. The complex is recruited to the locus from which the siRNA has originated by the guide RNAs transcribed by another plant-specific DNA-dependent RNA polymerase, RNA polymerase V (Cao et al., 2000; Cao and Jacobsen 2002; Wierzbicki, 2012). Then the complex further recruits the downstream effectors, such as *de novo* DNA methyltransferase, DRM2, to establish and/or maintain epigenetic marks. Thus, siRNAs from TE loci induce RNA-directed DNA methylation, which results in epigenetic inactivation of TEs (Zilberman and Henikoff, 2004; Lisch, 2009; Matzke et al., 2009). This mechanism results in epigenetic silencing of most TEs (Feschotte et al., 2002). Thus, there are many differences between miRNAs and siRNAs (Ambros et al., 2003; Meyers et al., 2008): (1) miRNAs are made from single transcripts, whereas siRNAs are produced from two RNA molecules. (2) The length of most miRNAs is 21 nt, whereas that of siRNAs is 24 nt in plants. (3) miRNAs act on *trans* targets, whereas siRNAs act in *cis* (although there are some exceptions such as ta-siRNAs). (4) miRNAs regulate gene expression post-transcriptionally, whereas siRNAs induced transcriptional silencing. (5) Because siRNAs arise from various genomic locations, their molecular diversity is much higher than that of miRNAs. (6) The expression levels of miRNAs, especially those conserved among various plant species, are usually higher than those of most siRNAs. (7) Small RNA profiling in many plant species revealed that the content of siRNA is much higher than that of miRNA.

Several small RNAs cannot be easily classified as miRNAs or siRNAs. One such

example is rice *miR820* (Chellappan et al., 2010; Wu et al., 2010; Nosaka et al., 2012). It is classified as a miRNA because it is produced from a single transcript with potential fold back structure and because it cleaves its *trans* target mRNA encoding DRM2. In addition, *miR820* is highly expressed as with other conserved miRNAs (miRBase; http://www.sanger.org/). However, it also has similar character to siRNAs. For example, it is originated from transposons, its size is 22 or 24 nt, and these two forms are processed by DCL1 and DCL3, respectively (Cao et al., 2000; Cao and Jacobsen 2002; Sharma et al., 2009; Henderson et al., 2010). In a previous study, I have shown that the role of *miR820* is to enable TEs to suppress *OsDRM2*, the major effector of host defense (Nosaka et al., 2012). However, it is still unknown how *MIR820* is transcribed.

In this study, I show that most of CACTA TEs carrying *MIR820* are transcriptionally inactive and harbor chromatin relatively enriched in silent histone marks. High expression of *miR820* solely depends on transcription from one CACTA TE on chromosome 7. Epigenetic marks, such as DNA methylation and histone modification in the *MIR820* region in CACTA TE on chromosome 7, show a unique pattern different from four other copies. I conclude that *MIR820* transcription may depend on epigenetic modifications, which may be in turn established by the action of *miR820* in *cis*. Thus, *miR820* appears to act both on *cis* and *trans* targets, and its own transcription is under epigenetic control.

Materials and Methods

Plant materials and growth conditions

Rice cultivars Nipponbare and Yukihikari were used. Plants were grown in soil or in tissue culture boxes at 29°C under continuous light.

RNA purification, PCR and sequencing

Total RNA was purified from seedlings with the RNeasy Mini Kit (QIAGEN) according to the manufacturer's instructions. Total RNA (50 ng) was reverse-transcribed by using Omniscript Reverse Transcriptase (QIAGEN) with random primers (N_9). PCR was performed by using *Ex Taq* DNA polymerase (TaKaRa). Primers used for RT-PCR are listed in Table 3.1. Amplified bands were gel-purified, cloned, and sequenced.

5′ RACE

Total RNA (3 μg) was subjected to RNA oligo ligation with the GeneRacer Kit (Invitrogen) according to the manufacturer's instructions. The oligo-ligated RNA was reverse-transcribed by using Omniscript Reverse Transcriptase (QIAGEN) with random primers (N_9). PCR and nested PCR were performed by using *Ex Taq* DNA polymerase (TaKaRa). Primers used for 5′ RACE PCR are listed in Table 3.1. Amplified bands were gel-purified, cloned, and sequenced.

ChIP-qPCR

Chromatin immunoprecipitation was performed as described previously (Miura et al., 2009) with the anti-H3K4me2 (ab1012), anti-H3K4me2me3 (ab6000), anti-H3K9ace (ab12179), and anti-H3K9me2 (ab1220) antibodies (Abcam). The immunoprecipitates were analyzed by quantitative PCR (qPCR) on a StepOnePlus Real-Time PCR system (Applied Biosystems). The qPCR reactions contained 5 µl 2× SYBR Premix *Ex Taq* II (Tli RNase H Plus) (TaKaRa), 0.5 µl DMSO, 0.2 µl 50× ROX reference dye, 1 µl immunoprecipitated DNA, and 400 nM of each primer. Each reaction was run in triplicate using a three-step cycling program (95°C for 15 s, 55°C for 30 s, 72°C for 30 s; 50 cycles). qPCR specificity was checked for each run with a dissociation curve at 95°C–60°C. The data were analyzed by using the standard curve method. The enrichment relative to input DNA was used to normalize the qPCR output. Primers used for qPCR are listed in Table 3.1. The *C-kinase substrate* and *Centromere 8-30* (*Cen 8-30*) genes were used as controls for euchromatic and heterochromatic genes, respectively (Nagaki et al., 2004); PCR primers were as designed by these authors.

Bisulfite sequencing

Total RNA (10 µg) from Yukihikari seedlings was subjected to bisulfite treatment with the EpiTect Bisulfite Kit (QIAGEN) according to the manufacturer's instructions. Primers used for bisulfite sequencing are listed in Table 3.1.

qPCR

Relative expression levels were quantified by qPCR performed by using the One Step SYBR PrimeScript RT-PCR Kit II (TaKaRa) on a StepOnePlus Real-Time PCR system

(Applied Biosystems). The reactions contained 5 µl 2× One Step SYBR RT-PCR Buffer 4, 0.5 µl DMSO, 0.4 µl PrimeScript One step Enzyme Mix 2, 0.2 µl 50× ROX reference dye, the equivalent of 50 ng total RNA, and 400 nM of each primer. Each reaction was run in triplicate. The mixtures were reverse-transcribed at 42°C for 5 min, and amplified by using a two-step cycling program (95°C for 5 s, 60°C for 20 s; 40 cycles). qRT-PCR specificity was checked for each run with a dissociation curve at 60°C–95°C. The data were analyzed by the standard curve method. The housekeeping gene *OsGAPDH* was used to normalize the qRT-PCR output. Primers used for qPCR are listed in Table 3.1.

Results

CACTAs carrying *MIR820* are epigenetically silenced

miR820 is a TE-encoded gene that targets *OsDRM2*, a gene required for the host defense against parasitic DNA such as TEs (Nosaka et al., 2012). No transpositionally active TEs have been reported in the Nipponbare genome under standard cultivation conditions. Histones associated with inactive TEs bear modifications typical for heterochromatin, and transcription in these regions is low. However, *miR820* is highly expressed from CACTA TEs. In order to clarify the mechanism of MIR820 transcription from silent TEs, I determined the chromatin state in the transposase regions of CACTAs carrying *MIR820* by using ChIP-qPCR with primers that amplify all five copies simultaneously (Fig. 3.1). The recovery of chromatin by the antibodies recognizing the euchromatic marks (H3K4me2, H3K4me2me3, H3K9ace) was low. In contrast, high recovery was observed for the heterochromatic mark (H3K9me2) (Fig. 3.1). Thus, all five copies of CACTAs carrying *MIR820* seem to be in the silenced chromatin context at least in this region.

MIR820 is mostly transcribed from one of the five CACTA copies

Although the alignment of the sequences of CACTAs carrying *MIR820* showed that they are very similar, I found several single-nucleotide polymorphisms (SNPs) and insertions/deletions among the five copies (Fig. 3.2). I used these polymorphisms to distinguish the primary transcripts of *MIR820* of different origin. I amplified the

corresponding fragments by RT-PCR with primers that recognize all five copies, followed by cloning and sequencing. Among 31 clones obtained, 29 and 2 clones had sequences identical to the *MIR820* regions on chromosomes 7 and 10, respectively (Fig. 3.2). Thus, *MIR820* is mainly transcribed from CACTA on chromosome 7.

Histone modifications in the *MIR820* loci

Because *miR820* is actively produced from CACTA on chromosome 7, it is possible that the local histone modifications in the promoter region and/or in the vicinity of the *MIR820* transcriptional unit could be different from the other four copies. Therefore, I tried to identify the transcription start site of *MIR820* on chromosome 7 by 5′ RACE analysis. I could not detect any PCR product amplified from 5′-capped or triphosphate RNAs (data not shown), but I successfully amplified the 5′ monophosphate RNA. Three positions of 5′ ends of *MIR820* transcripts were detected at approximately 400 bp upstream of the *miR820* coding region (Fig. 3.3A). Because the 5′ monophosphate end is produced by nuclease digestion, it is possible that these amplified fragments are not primary transcripts but are processed. However, these regions amplified by modified 5′ RACE are at least in the transcriptional units.

Next, I examined the histone modification state around the transcriptional unit of each *MIR820* on chromosomes 1, 7, 8, 10, and 12 by ChIP-qPCR (Fig. 3.3B–E). I designed three pairs of PCR primers that amplify the 3′ end of the transposase ORF, the region recovered as the 5′ end of *MIR820*, and the region just upstream of the *miR820* coding region (Fig. 3.3A; regions 1–3, respectively). The recovery of chromatin by the antibodies recognizing the euchromatic marks (H3K4me2, H3K4me2me3, H3K9ace) was low in these regions for all five *MIR820* copies compared to a region in the

C-kinase substrate gene (an actively transcribed gene), and mostly higher than that of *Cen8-30* (an authentic heterochromatic sequence in rice) (Fig. 3.3B–D). The recovery of chromatin in these three regions by the H3K9me2 antibody was higher than (in particular, in region 1 from all five copies) or similar to that for *C-kinase substrate* (Fig. 3.3E). Only region 3 on chromosome 7 showed low active marks compared to all other regions in all CACTAs carrying *MIR820*. There was no obvious change in the silent marks of histone modification in this region among five copies. Thus, the histones in region 3 on chromosome 7 have a unique modification pattern.

MIR820 on chromosome 7 has a high level of asymmetric cytosine methylation around the _miR820_ coding region

Because the histone modification state around the *miR820* coding region on chromosome 7 was different from those on other chromosomes, I suspected that the DNA methylation in this region could also be different. I analyzed the DNA methylation of each *MIR820* (including *miR820* and *miR820** coding regions) on chromosomes 1, 7, 8, and 10 by bisulfite sequencing (Fig. 3.4). In all copies tested, the ratio of cytosine methylation at most of CG and CHG sites around *miR820* and *miR820** coding regions was 80–100%. On chromosomes 1, 8, and 10, methylation at CHH sites was much lower (or even undetectable) than at the adjacent CG or CHG sites, whereas on chromosome 7 methylation at CHH sites adjacent to the *miR820* coding region was higher than on other chromosomes.

Thus, both the histone modifications and DNA methylation around the *miR820* coding region on chromosome 7 are different from copies on other chromosomes.

Discussion

Because TEs are potentially harmful for their host, most of them are silenced by the host defense machinery, and they are located in heterochromatic regions. This defense machinery is activated by siRNAs derived from genomic parasites such as TEs. These siRNAs guide the silencing machinery to the genomic region where they originated, or to the homologous sequence(s) in the genome to epigenetically silence them. My previous report showed that *miR820* acts against the host silencing machinery by attenuating *OsDRM2* (Nosaka et al., 2012). Thus, it seemed plausible that CACTA TEs carrying *MIR820* could be active; however, so far there has been no evidence of active DNA transposons in Nipponbare under standard cultivation conditions. Here I found that *MIR820* is actively transcribed from a CACTA TE copy on chromosome 7. This finding raised the question of why only this copy is capable of transcription, and how it differs from the four other copies.

The ChIP-qPCR analysis and bisulfite sequencing around *MIR820* regions gave me some hints on this issue. Histones in region 3 of the *MIR820* locus on chromosome 7 have low levels of active marks (H3K4 di/tri methylation and H3K9 acetylation). Despite this, *MIR820* on chromosome 7 is the major transcribed copy. I found that the low level of active marks in this copy correlates with the high level of asymmetric cytosine methylation (CHH) in the same region. Although I am unable to depict the molecular framework of how this pattern of epigenetic modifications in region 3 allows transcription of *MIR820*, it is plausible that these unique modifications allow

transcription from epigenetically silenced TE.

A unique feature of *miR820* is that it exists as both 22- and 24-nt species. I and another group have shown that *miR820* cleaves *OsDRM2* mRNA, and also induces DNA methylation at the *miR820* recognition site in the *OsDRM2* locus. In the present study, the high level of CHH methylation was observed mostly within and around the *miR820* coding region on chromosome 7. This pattern of DNA methylation may also be induced by *miR820*. The high level of CHH methylation is only observed in *miR820* encoded on chromosome 7, whereas the corresponding regions on chromosomes 1, 8, and 10 should be also recognized by *miR820*. One possible explanation is that *miR820*-directed CHH methylation is coupled with transcription, and this is why *miR820* acts in *cis* but has no effect on *miR820* coding regions on other chromosomes.

I propose a model for the mode of action of *miR820* (Fig. 3.5). *Pri-miR820* transcripts derived from chromosome 7 are processed into mature *miR820*. This induces both *OsDRM2* mRNA cleavage and DNA methylation at the *miR820* target site in the *OsDRM2* genomic region (in *trans*). *miR820* also induces CHH methylation in the *miR820* genomic region on chromosome 7 (in *cis*) and epigenetically regulates its own transcription. Thus, *miR820* is a unique miRNA that acts both in *cis* and in *trans*, and its own transcription is under epigenetic control. Typically, TE-derived small RNAs are siRNAs that act in *cis*, whereas miRNAs act in *trans*. My data confirms that *miR820* has features resembling both miRNA and siRNA, possibly because it is encoded by a parasitic gene that uses the host machinery to counteract silencing.

Table 3.1. Primers and probes used in this study. Regions 1–3 (see Fig. 3.3A) are indicated after the chromosome numbers.

Target genes, chromosomes, and regions	Sequence (5′ to 3′)	Purpose
pre-miR820	TGATGAATATCCTTACCAATCTTG	sequencing
	TGATAACGTAMGAACTACACCTCC	sequencing
C-kinase substrate	CGACTAAACCACTCCAATCATC	ChIP-qPCR
	CCAATCAAAACTTCTCCTGTAA	ChIP-qPCR
Centromere 8	CCGATATGCCAAAGAGCGAGTC	ChIP-qPCR
	CAAATCATCTATCCTCAAGTCC	ChIP-qPCR
CACTA (ORF1)	GTATGAATATGGCAAGCCGTT	ChIP-qPCR
	GAAACTGAAGGCGAAGTTTGC	ChIP-qPCR
pre-miR820	GGACACTGACATGGACTGAAGGAGTA	5′ RACE
	TGATAACGTAMGAACTACACCTCC	5′ RACE
pre-miR820 (chr.1-1)	AATGGGTCAGAGACACAACAA	ChIP-qPCR
	TCTTATCGTCGGAAGGGTCAC	ChIP-qPCR
pre-miR820 (chr.1-2)	GTTTGATAACAGAACAACTAC	ChIP-qPCR
	GAAATTATTCATCATTTACCA	ChIP-qPCR
pre-miR820 (chr.1-3)	TAGTTCGTACGTTATCATGCC	ChIP-qPCR
	AGGAACTAAACCGACTAGTAA	ChIP-qPCR
pre-miR820 (chr.7-1)	AATGGGTCAGAGACACAACAT	ChIP-qPCR
	TCTTATCATCGGAAGGGTCAT	ChIP-qPCR
pre-miR820 (chr.7-2)	GCTTGAGAACAGAACAACTAA	ChIP-qPCR
	GCAATTATTCATCATGAACCC	ChIP-qPCR
pre-miR820 (chr.7-3)	TAGTTCGTACGTTATCATGCT	ChIP-qPCR
	AGCAACTAAACCGACTAGTAG	ChIP-qPCR
pre-miR820 (chr.8-1)	AATGGGTCAGAGACACAACAG	ChIP-qPCR
	TCTTATCGTCGAAAGGGTCAA	ChIP-qPCR
pre-miR820 (chr.8-2)	GCTTGAGAACAGAACAGCTAC	ChIP-qPCR
	GAAATTATTCATCATTTACCA	ChIP-qPCR
pre-miR820 (chr.8-3)	TAGTTCTTACGTTATCATGCC	ChIP-qPCR
	AGGAACTAAACCGACCAAGCC	ChIP-qPCR
pre-miR820 (chr.10-1)	AATGGATCAGAGAACAGAACAG	ChIP-qPCR
	TCTTATCGTCGGAAGGGTCAC	ChIP-qPCR
pre-miR820 (chr.10-2)	GCTTGAGAACAGACAACTAC	ChIP-qPCR
	TGGTAAGGATATTCATCACC	ChIP-qPCR
pre-miR820 (chr.10-3 & chr.12-3)	TAGTTCGTACGTTATCATGCC	ChIP-qPCR
	AGGAACTAAACCGACCAAGCC	ChIP-qPCR
pre-miR820 (chr.12-1)	AATGGGTTAGAGACACAACAG	ChIP-qPCR
	TCTTATCGTCAGAAGGGTCAC	ChIP-qPCR
pre-miR820 (chr.12-2)	GCTTGAGAACAGAACAACTAC	ChIP-qPCR
	GAAATTATTCATCATTTACCA	ChIP-qPCR

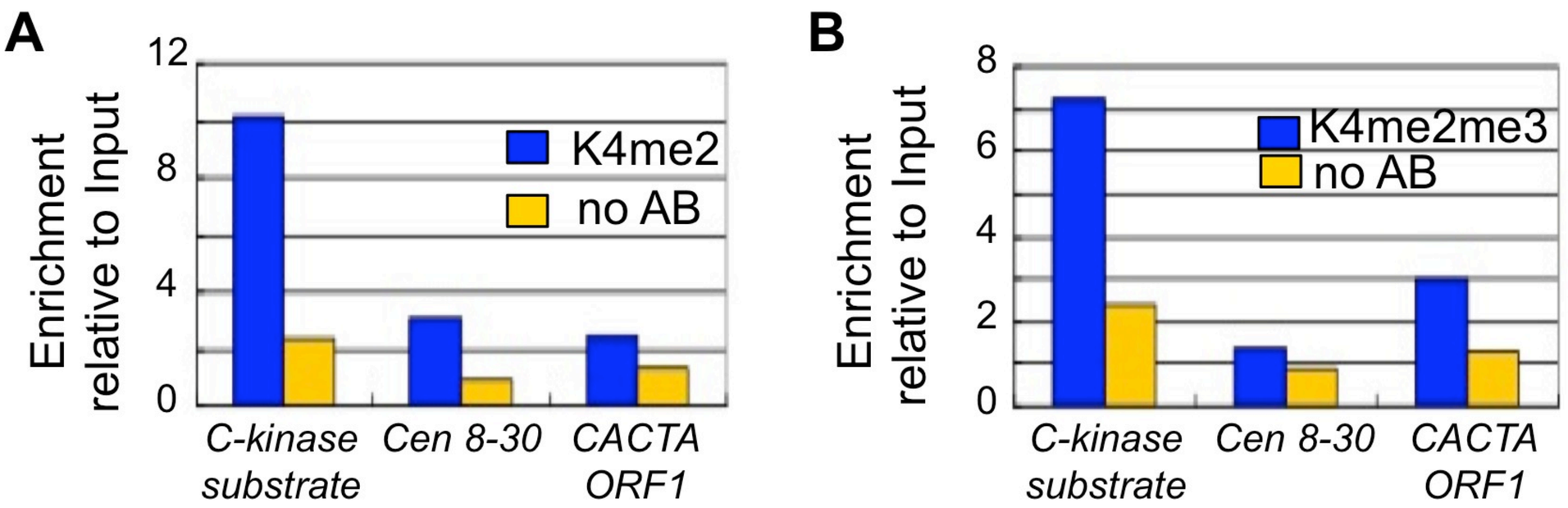


Figure 3.1. *miR820*-CACTA is located in a heterochromatic region.
Chromatin immunoprecipitation was performed with antibodies that detect active euchromatic marks (H3K4me2 (A), H3K4me2 and H3K4me3 (B), and H3K9ace (C)) or a constitutive heterochromatic mark (H3K9me2 (D)). As a negative control, immunoprecipitation was performed without an antibody (yellow boxes). *C-kinase substrate* and *Centromere 8-30* were used as controls for euchromatic and heterochromatic regions, respectively. Values are means of three technical replicates.

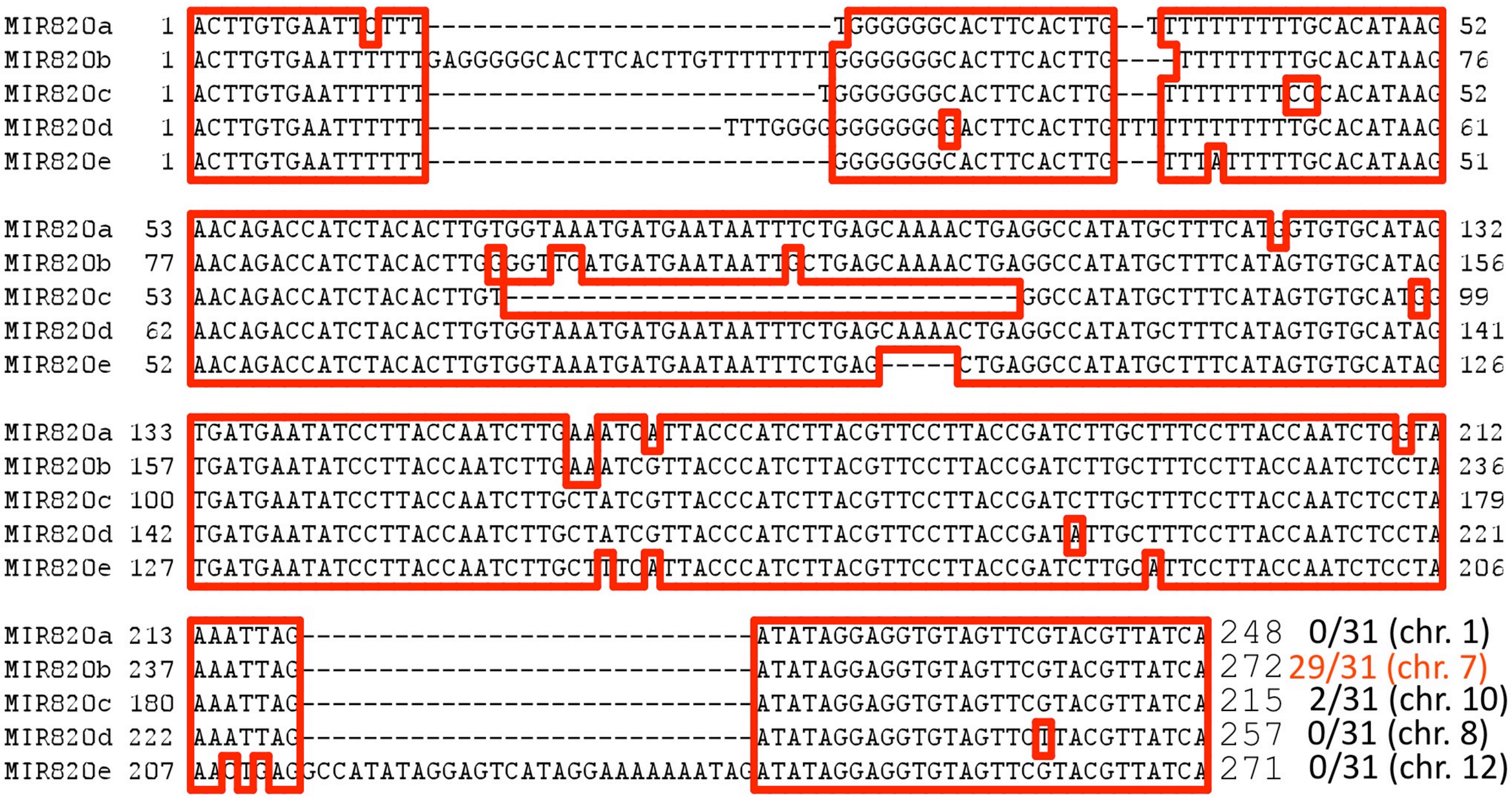
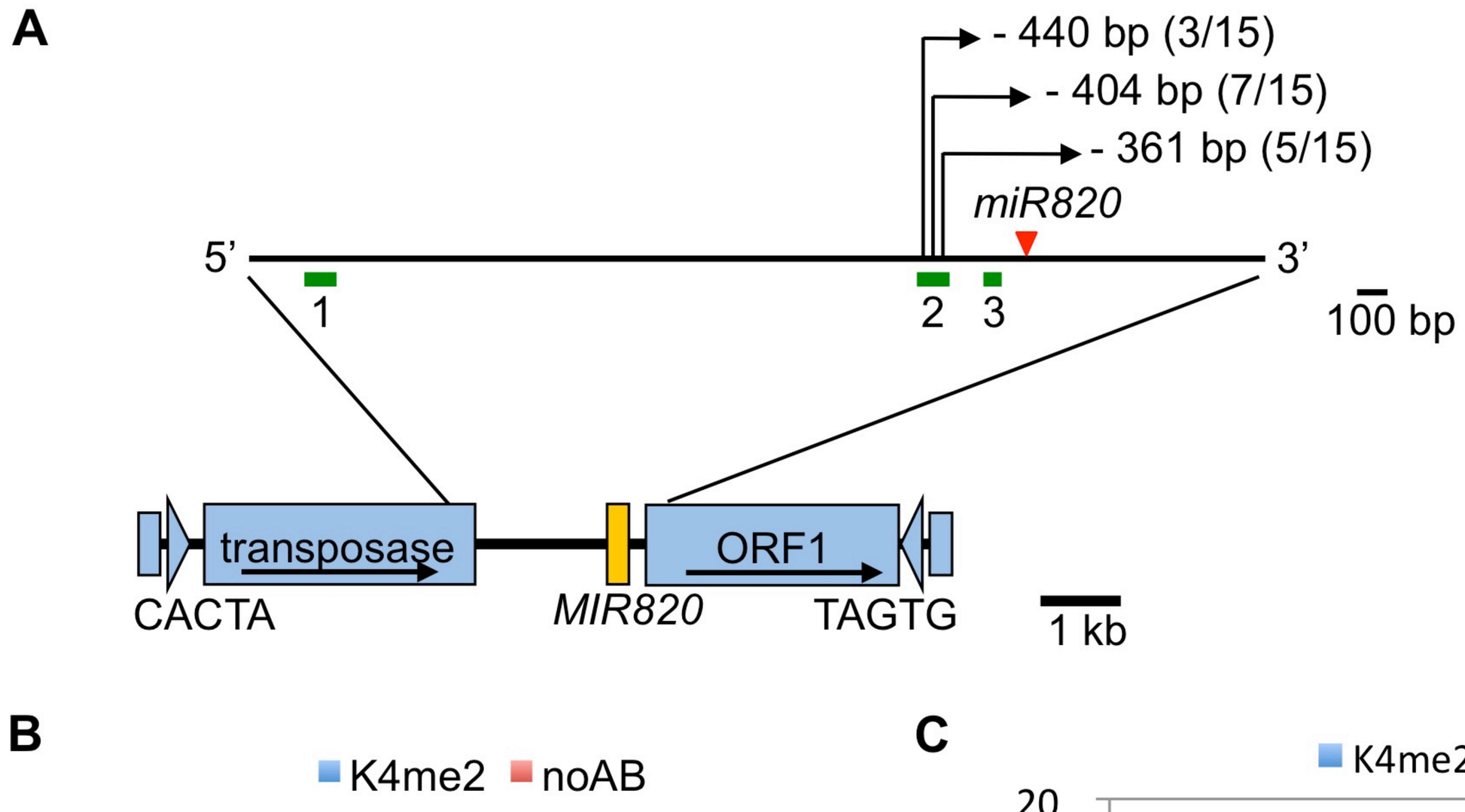
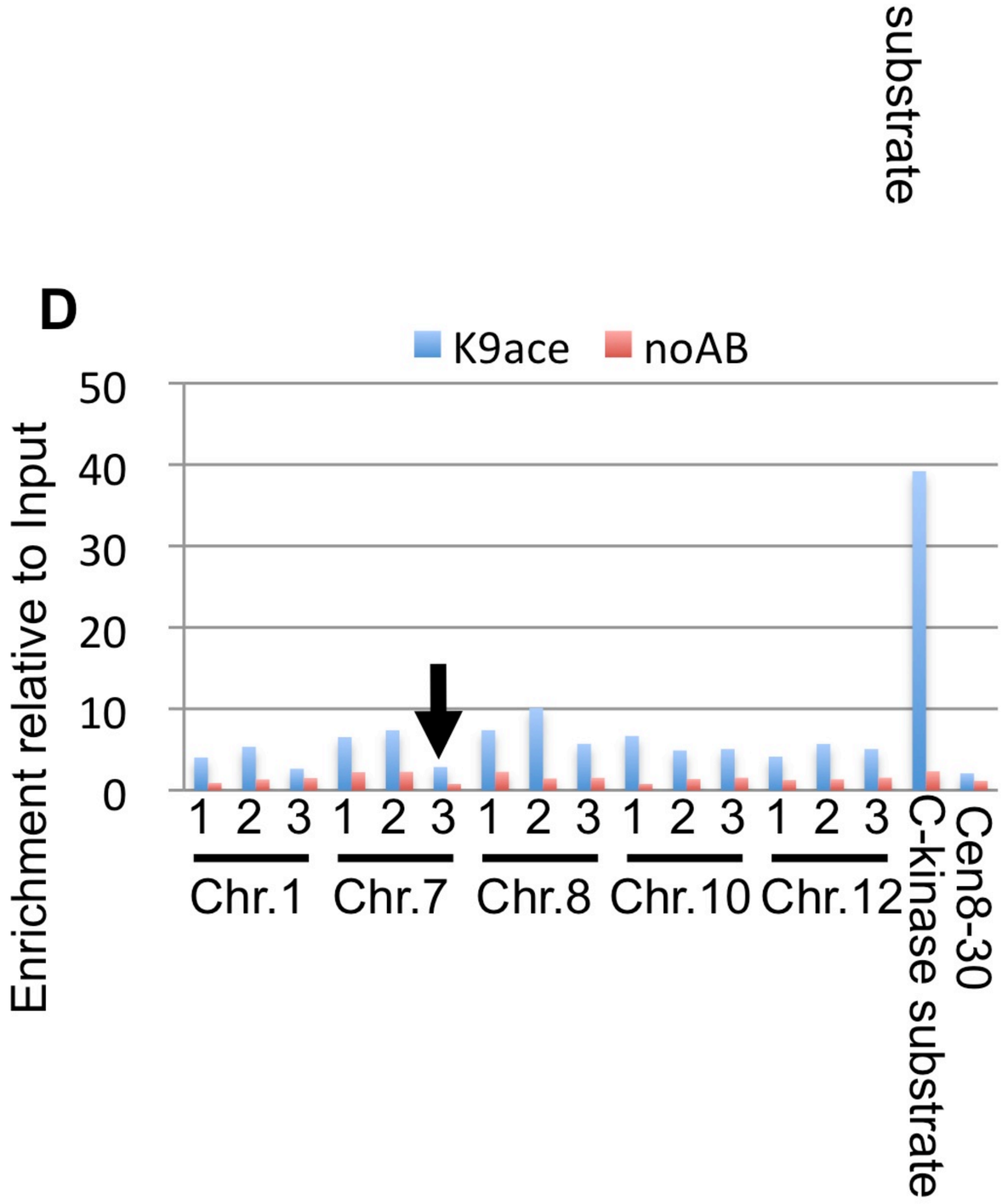


Figure 3.2. *pre-miR820* is mainly transcribed from the locus on chromosome 7.
Alignment of *pre-miR820* transcribed from different chromosomes. The numbers next to the arrows denote the numbers of clones with the nucleotide sequence of each *pre-miR820* out of 31 clones sequenced.


A
- 440 bp (3/15)
- 404 bp (7/15)
- 361 bp (5/15)
miR820
5'
3'
1
2
3
100 bp
transposase
ORF1
CACTA
MIR820
TAGTG
1 kb

B
K4me2
noAB
Enrichment relative to Input
15
10
5
0
1 2 3 1 2 3 1 2 3 1 2 3 1 2 3
Chr.1
Chr.7
Chr.8
Chr.10
Chr.12
C-kinase substrate
Cen8-30

C
K4me2me3
noAB
Enrichment relative to Input
20
15
10
5
0
1 2 3 1 2 3 1 2 3 1 2 3 1 2 3
Chr.1
Chr.7
Chr.8
Chr.10
Chr.12
C-kinase substrate
Cen8-30


D
K9ace
noAB
Enrichment relative to Input
50
40
30
20
10
0
1 2 3 1 2 3 1 2 3 1 2 3 1 2 3
Chr.1
Chr.7
Chr.8
Chr.10
Chr.12
C-kinase substrate
Cen8-30

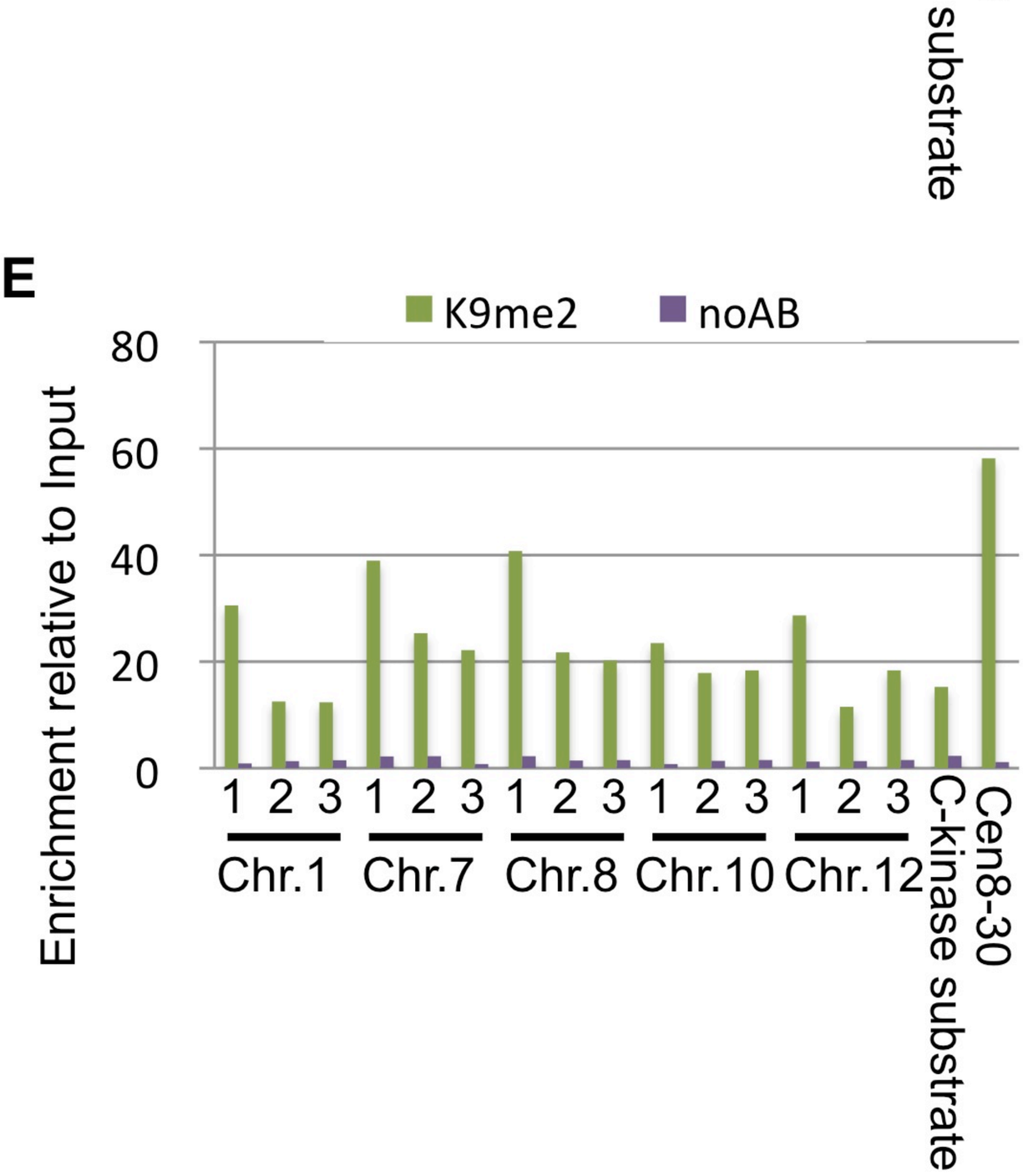

E
K9me2
noAB
Enrichment relative to Input
80
60
40
20
0
1 2 3 1 2 3 1 2 3 1 2 3 1 2 3
Chr.1
Chr.7
Chr.8
Chr.10
Chr.12
C-kinase substrate
Cen8-30

Figure 3.3. Characterization of *pre-miR820* transcripts.
(A) Location of *miR820* within the *CACTA* transposon (lower panel) and the position of the 5′ ends of *pre-miR820* transcripts (upper panel). The black line indicates the *CACTA* transposon genomic region. The red triangle indicates the position of *miR820*. The right-angled arrows show the position of the 5′ ends of *pre-miR820* transcripts. The numbers following the arrows show the position of the 5′ end of each transcript relative to the 5′ end of *miR820*. The numbers of clones sharing the same 5′ ends are shown in parentheses (out of 15 clones analyzed). ChIP-qPCR was performed in the regions shown by the green lines; the numbers (1–3) correspond to the numbers under the *x* axes in (B–E). ChIP-qPCR was performed with antibodies against H3K4me2 (B), H3K4me2 and H3K4me3 (C), and H3K9ace (D), which detect active euchromatic marks, and H3K9me2 (E), which detects a constitutive heterochromatic mark. Equal amounts of input DNA and the immunoprecipitates with or without antibodies (no AB) were analyzed and normalized against input DNA. Values are means of three technical replicates. *C-kinase substrate* and *Centromere 8-30* are controls for euchromatic and heterochromatic regions, respectively.

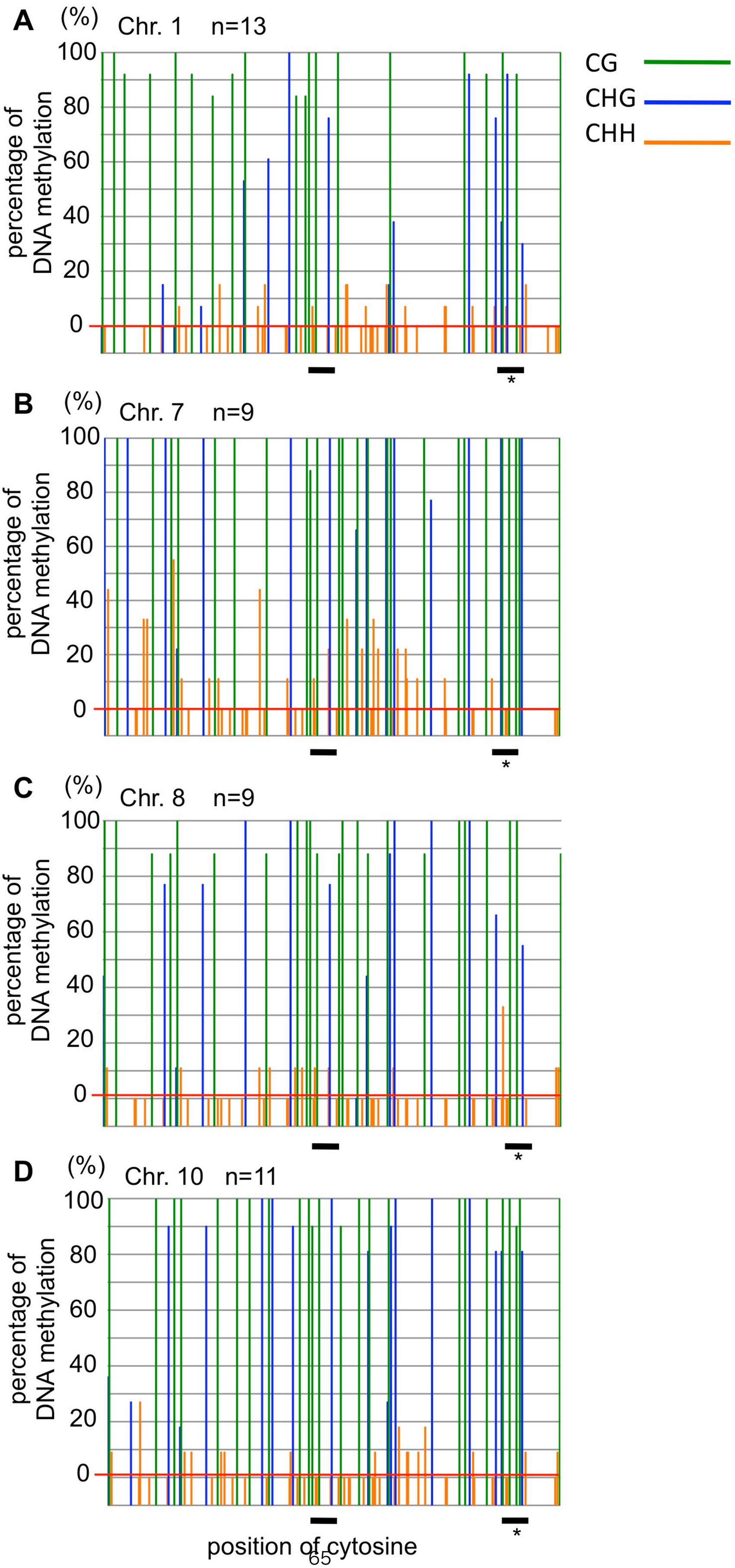

A
(%)
Chr. 1 n=13
100
80
60
40
20
0
percentage of DNA methylation
CG
CHG
CHH
*
B
(%)
Chr. 7 n=9
100
80
60
40
20
0
percentage of DNA methylation
*
C
(%)
Chr. 8 n=9
100
80
60
40
20
0
percentage of DNA methylation
*
D
(%)
Chr. 10 n=11
100
80
60
40
20
0
percentage of DNA methylation
*
position of cytosine

Figure 3.4. DNA methylation in the *pre-miR820* region.
Bisulfite sequencing analysis of *pre-miR820a* (A), *pre-miR820b* (B), *pre-miR820d* (C), and *pre-miR820c* (D) regions on their respective chromosomes. The colored lines show the positions of individual cytosines; those extending above the horizontal line indicate methylated sites. The height of the colored lines indicates the percentage of methylation. *miR820* and *miR820** are indicated by black bars without and with an asterisk, respectively. The n value represents the number of clones analyzed.

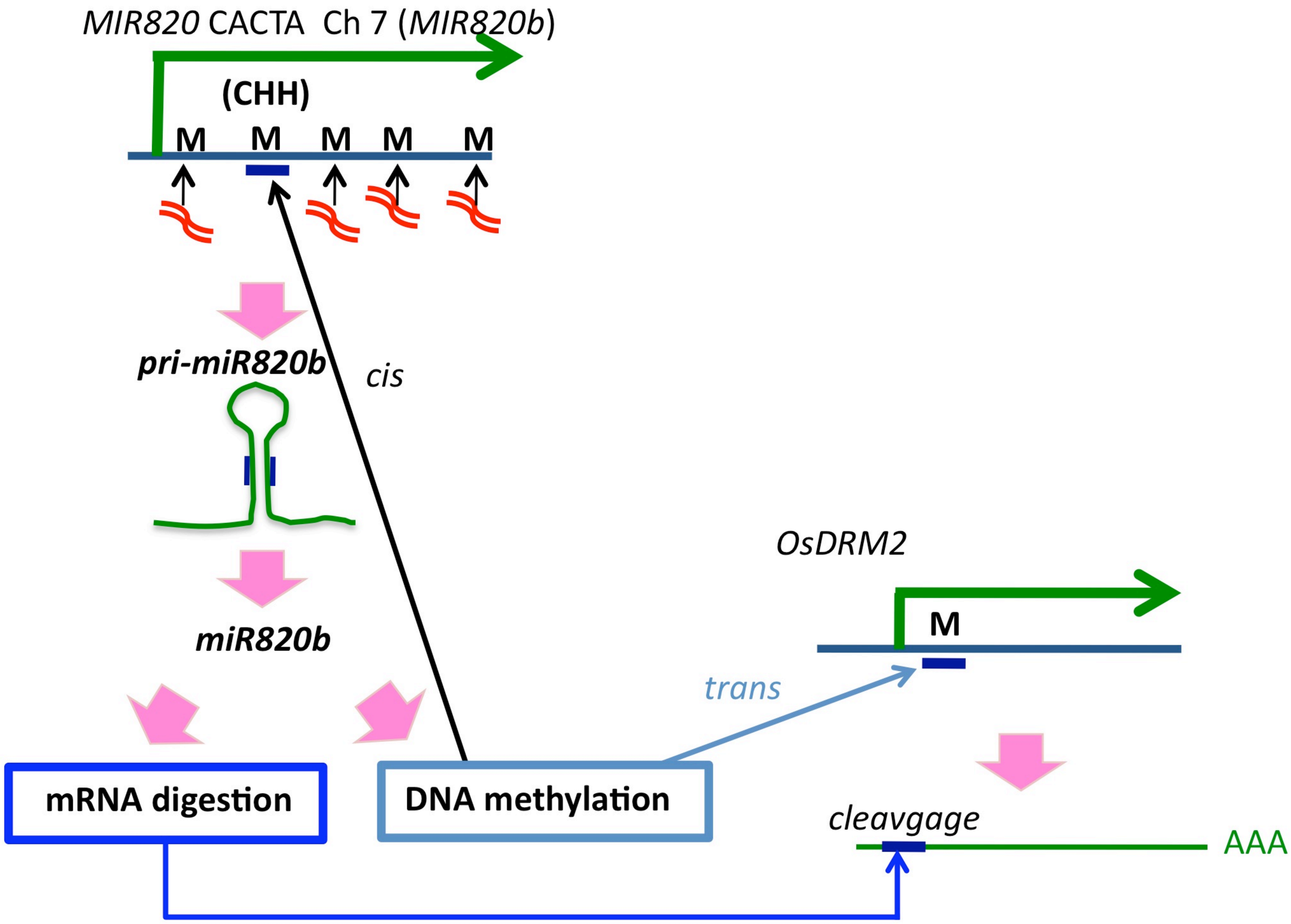


Figure 3.5. A model for the action of *miR820*.
pri-miR820 transcripts are mainly transcribed from the *MIR820b* locus on chromosome 7. *miR820* processed from *pri-miR820b* induce CHH DNA methylation (M) in the *MIR820b* genomic region (in *cis*). This could allow *miR820* transcription from this lous. *miR820* also induces *OsDRM2* mRNA digestion and DNA methylation at the *miR820* target site in the *OsDRM2* genomic region (in *trans*). Thus, *miR820* has two functions, regulating its own transcription in *cis* and *OsDRM2* expression in *trans*.

CHAPTER 4

The copy number variation of CACTA DNA transposon carrying *MIR820* in rice does not correlate with *MIR820* expression.

Abstract

miR820 is a small-RNA species with sizes of 22 and 24 nt. *miR820* is produced from transcripts originating from a region inside a class of CACTA DNA transposons in rice. Because *MIR820* is a transposon gene, its expression could depend on the copy number of transposon carrying *MIR820*. Here, I investigated the copy number of CACTA carrying *MIR820* and the expression level of *MIR820* in various cultivars and wild-species of rice. I found that there is no positive correlation between the copy number and the expression level, suggesting that transcription of *MIR820* is regulated not by the dose of the copy but by the epigenetic state of each copy.

Introduction

The major component of eukaryotic genomes is TEs and their remnants (Kidwell, 2002). TEs have increased their copy number in the host genome through their replication faster than the host genome. Thus, TEs are referred to as ultimate parasites that proliferate selfishly on the genome (Doolittle and Sapienza, 1980; Orgel and Crick, 1980). On the other hand, looking from the host's side, TE's transposition that induces inserted mutations and aberrant chromosomes cause instability of host genome. Therefore, most of TEs are kept silent by their host. A number of groups are carrying out research on the host's defense machinery against genomic parasites, such as TEs, in plants and animals. Until today, small RNA-mediated RNA silencing and its corresponding mechanism is known to participating in suppression of TE, and this mechanism is known to have a common role in plants and animals to some degree (Zilberman and Henikoff, 2004; Brennecke et al., 2007; Kuramochi-Miyagawa et al., 2008).

Although most TEs are silenced by their host, TEs are major component of their host genome. This suggests the existence of a running battle between the host's defense machinery suppressing transposition of TEs and TE's countermeasures against host-mediated silencing. However, until now, little is known about the strategies that TEs have evolved to avoid host-mediated machinery suppressing TE's transposition. In my previous work, I showed that members of the *microRNA820* (*miR820*) gene family negatively regulate *OsDRM2*, a *de novo* DNA methyltransferase gene, and that by this

regulation transposons escape from the host silencing. I also found that, in some of the wild-rice accession, such as BB and BBCC genome species, CACTA carrying *MIR820* is drastically proliferated (Nosaka et al., 2012).

Materials and Methods

Plant materials and growth conditions

Nipponbare, Yukihikari, 9311 and T65 rice plants were grown in tissue culture boxes or in soil with at 29 °C under continuous light. Seeds of *Oryza* species and NIAS rice core collection were kindly provided by the National Institute of Genetics (Mishima, Japan) and the National Institute of Agrobiological Sciences (Tsukuba, Japan), respectively.

Northern blot analysis

Total RNA of rice seedling was purified with the RNeasy Mini Kit (QIAGEN) according to the manufacturer's instructions. 10 μg of each RNA sample was loaded onto acrylamide gels for analysis of *miR820,* separated by electrophoresis, blotted to membranes. The membranes were probed with oligo DNA complementary to *miR820*.

5' RACE

Total RNA of rice seedling was purified with the RNeasy Mini Kit (QIAGEN) according to the manufacturer's instructions. 3 μg of total RNA purified as described above was subjected to RNA Oligo ligation with the GeneRacer Kit (Invitrogen) according to the manufacturer's instructions. GeneRacer RNA Oligo-ligated RNA was reverse-transcribed using Omniscript Reverse Transcriptase (QIAGEN) with random primers (N_9). PCR and nested PCR were performed using *Ex Taq* DNA polymerase (TaKaRa). Primers used for 5' RACE PCR are listed in Table 4.1. Amplified bands

were gel-purified, cloned, and sequenced.

Quantitative RT-PCR

Relative expression levels were quantified using the StepOnePlus Real-Time PCR system (Applied Biosystems) and One Step SYBR PrimeScript RT-PCR Kit II (TaKaRa). The quantitative RT-PCR reactions contained 5 µl 2× One Step SYBR RT-PCR Buffer 4, 0.5 µl DMSO, 0.4 µl PrimeScript 1 step Enzyme Mix 2, 0.2 µl 50× ROX reference dye, the equivalent of 50 ng total RNA, and 400 nM of each primer, and were run in triplicate. The mixtures were first reverse-transcribed at 42 °C for 5 min, then amplified via PCR using a two-step cycling program (95 °C for 5 s, 60 °C for 20 s) for 40 cycles. Quantitative RT-PCR for *Oryza* species was performed in two ways; with reverse-transcription and without reverse-transcription. Quantitative RT-PCR specificity was checked for each run with a dissociation curve, at temperatures ranging from 95 °C to 60 °C. Data from quantitative RT-PCR were analyzed using the standard-curve method. The housekeeping gene, *OsActin,* was used to normalize the quantitative RT-PCR output. Primers used for quantitative RT-PCR are listed in Table 4.1.

Southern blot analysis

Genomic DNA samples were extracted from leaves of NIAS rice core collection, treated wuth RNase A, and digested with *EcoRI* and *EcoRV*. These samples were loaded onto an agarose gel, separated by electrophoresis, blotted onto a nylon membrane, and probed with *pre-miR820* DNA fragment.

Results and Discussion

I assumed that the expression of *miR820* would be extremely high. To confirm this, I conducted northern blot analysis to detect *miR820* accumulation in 3 different cultivated rice and one wild-rice species (W1514: *Oryza punctata*) (Fig. 4.1A). Surprisingly, no expression was detected in W1514 in which more than 18 copies of *MIR820* exist (Nosaka et al., 2012), and, in accordance with this, no cleavage product of *OsDRM2* by *miR820* was detected in W1514 (Fig. 4.1B). Next I determined the expression of *pre-miR820* by the quantitative RT-PCR in two cultivated rice species and four wild-rice species containing BB genome, in which CACTA carrying *MIR820* is drastically proliferated (Nosaka et al., 2012) (Fig. 4.1C). It turned out that the absence of *miR820* in wild-rice species containing BB genome is due to reduction of transcription of *MIR820* rather than that of processing of *pre-miR820* transcript. Because the copy number of *MIR820* in wild-rice species containing BB genome is much higher than Nipponbare, it is possible that there is a strong bias of silencing CACTA carrying *MIR820* and silencing the expression of *MIR820* in these species.
I then explored rice accessions with moderate copy numbers of *MIR820* to see if the expression level of *MIR820* depends on its copy number. I determined the copy number of *MIR820* by Southern hybridization using the collections of cultivated rice, Japanese Rice Core Collection (JRC) and World Rice Core Collection (WRC) including 45 and 56 cultivars, respectively (Kojima et al., 2005; Ebana et al., 2008) (Tables 4.2 and 4.3). In 45 JRC cultivars, the average, minimum and maximum copy numbers of *MIR820*

were 4.5, 2 and 6, respectively. In 56 WRC cultivars, the average, minimum and maximum copy numbers of *MIR820* were 6.3, 3 and 11, respectively. Then, I determined the expression of *pre-miR820* by the quantitative RT-PCR in these collections and turned out that there is no correlation between *MIR820* copy number and in expression in moderate level of *MIR820* copies (Fig. 4.2B).

Previously, I have shown that *miR820* acts as a countermeasure of transposons against the host's defense machinery by down-regulating the expression of *de novo* DNA methyltransferase which is responsible for the inactivation of transposons. Therefore, I can assume that once *miR820* effectively suppress the function of *OsDRM2*, silencing of transposons including CACTA carrying *MIR820* would be released and this could result in more transcription of *MIR820* to form the feed forward loop, enforcing the function of *miR820*. However, in this study, I found that the expression of *MIR820* is relatively constant in many cultivated rice species even though the copy number *MIR820* varies two to eleven copies. This suggests that transcription of *MIR820* is regulated not by the dose of the copy but rather by the epigenetic state of each locus. This mechanism would be evolved during "arm-race" between the host and parasite, and allow the host to prevent the overwhelming victory of parasites by stopping the feed forward loop triggered by *miR820*.

Table 4.1. Primers used in this study.

Target gene	Primer name	Sequence (5' to 3')	Purpose
OsDRM2 (cleaved)	GeneRacer 5'	GGACACTGACATGGACTGAAGGAGTA	5'RACE
	AK065147R2	TTGCTCTGTATCTGTATCCCCAATCTCCTT	5'RACE
OsDRM2 (uncleaved)	qDRM2F1	GATAGCGACAATGATAAGTTCGAG	5'RACE
	qDRM2R3	TCCTCGTCGACTTCAAGAGTC	5'RACE
OsActin	ActU	TCCATCTTGGCATCTCTCAG	5'RACE
	ActL	GTACCCTCATCAGGCATCTG	5'RACE
pre-miR820 (*Oryza* species)	Satof	CTTASGTTCYTTRCYGATCTTG	RT-PCR
	R1SATO	TGACARCRTAYRAACTACACCTC	RT-PCR
OsActin	ActAK060893Fw	GAGTATGATGAGTCGGGTCCAG	RT-PCR
	ActAK060893Rv	ACACCAACAATCCCAAACAGAG	RT-PCR
pre-miR820 (core collection)	SatofN	TTACGTTCCTTACCGATMTTG	RT-PCR
	R1SATON	TGATAACGTAMGAACTACACCTCC	RT-PCR
OsDRM2	qDRM1aF1	CAGAGGTAGCTCTCCACAGAC	RT-PCR
	qDRM1aR2	CAGCTCTTCAGAATCGTTCTG	RT-PCR
pre-miR820	CACTApremiRJF1	CACATAAGAACAGACCATCTACAC	Probe amplification
	CACTApremiRJR3	TCTTTCATACTGCAATTGCGCTAG	Probe amplification

Table 4.2. Copy numbers of *MIR820* in the Japanese Rice Core Collection (JRC).

JRC No.	Name	Origin	Copy Number
	Gaisen	Japan(unknown)	
JRC01	Mochi		4
JRC03	Hinode	Kinki	5
JRC04	Senshou	Tokyo	4
	Yamada		
JRC05	Bake	Kagoshima	4
		Kantou	
JRC06	Kaneko B	Touzan	5
	Iruma		
JRC07	Nishiki	Saitama	5
	Okka	Japan(unknown)	
JRC08	Modoshi		4
JRC10	Hirayama	Tokyo	4
JRC11	Kahei	Kagoshima	4
JRC12	Oiran	Kumamoto	5
	Bouzu		
JRC13	Mochi	Ooita	5
JRC17	Akage	Akita	4
JRC19	Wataribune	Shiga	3
JRC20	Hosogara	Aomori	4
JRC21	Akamai	Kouchi	5
JRC22	Mansaku	Nagano	5
JRC23	Ishijiro	Toyama	5
JRC24	Joushuu	Yamagata	5
		Japan(unknown)	
JRC25	Dango		5
JRC26	Aikoku	Fukui	5
JRC27	Ginbouzu	Ishikawa	5
	Shinriki		
JRC28	Mochi	Kumamoto	5
	Shichimen		
	chou	Japan(unknown)	
JRC29	Mochi		5
	Morita		
JRC30	Wase	Yamagata	5
JRC31	Kameji	Shimane	6
JRC32	Omachi	Okayama	5
JRC33	Shinriki	Hyougo	5
	Kyoutoasa		
JRC34	hi	Kyoto	5
JRC35	Kabashiko	Miyazaki	5
	Shinyamad		
JRC37	aho 2	Hyougo	5
	Nagoya		
JRC38	Shiro	Akita	2
JRC39	Shiroine	Tokushima	4
JRC40	Akamai	Nagasaki	5
JRC41	Akamai	Tokushima	3
JRC42	Touboshi	Kagoshima	5
		Kantou	
JRC43	Akamai	Touzan	3
JRC44	Karahoushi	Kagoshima	3
JRC46	Fukoku	Hokkaidou	5
		Japan(unknown)	
JRC47	Okabo		4
JRC48	Hakamuri	Kagoshima	4
	Rikutou	Japan(unknown)	
JRC49	Rikuu 2		5
JRC51	Shinshuu	Nagano	5
JRC52	Aichiasahi	Aichi	3
		Kantou	
JRC53	Raiden	Touzan	5
	Houmanshi		
JRC54	nden Ine	Kagoshima	4

Table 4.3. Copy numbers of *MIR820* in the World Rice Core Collection (WRC).

WRC No.	Name	Origin	Copy Number
WRC01	Nipponbare	Japan	5
WRC02	Kasalath	India	6
WRC03	Bei Khe	Cambodia	6
WRC04	Jena 035	Nepal	5
WRC05	Naba	India	8
WRC06	Puluik Arang	Indonesia	3
WRC07	Davao 1	Philippines	3
WRC09	Ryou Suisan Koumai	China	5
WRC10	Shuusoushu	China	9
WRC11	Jinguoyin	China	5
WRC12	Dahonggu	China	7
WRC13	Asu	Bhutan	6
WRC14	IR 58	Philippines	4
WRC15	Co 13	India	7
WRC16	Vary Futsi	Madagascar	7
WRC17	Keiboba	China	11
WRC18	Qingyu (Seiyu)	China	8
WRC19	Deng Pao Zhai	China	4
WRC20	Tadukan	Philippines	6
WRC21	Shwe Nang Gyi	Myanmar	6
WRC22	Calotoc	Philippines	8
WRC23	Lebed	Philippines	5
WRC24	Pinulupot 1	Philippines	5
WRC25	Muha	Indonesia	10
WRC26	Jhona 2	India	10
WRC27	Nepal 8	Nepal	9
WRC28	Jarjan	Bhutan	7
WRC29	Kalo Dhan	Nepal	7
WRC30	Anjana Dhan	Nepal	5
WRC31	Shoni	Bangladesh	6
WRC32	Tupa 121-3	Bangladesh	6
WRC33	Surjamukhi	India	9
WRC34	ARC 7291	India	9
WRC35	ARC 5955	India	9
WRC36	Ratul	India	7
WRC37	ARC 7047	India	8
WRC39	Badari Dhan	Nepal	8
WRC40	Nepal 555	India	6
WRC41	Kaluheenati	Sri Lanka	5
WRC42	Local Basmati	India	5
WRC43	Dianyu 1	China	4
WRC44	Basilanon	Philippines	6
WRC45	Ma sho	Myanmar	6
WRC46	Khao Nok	Laos	5
WRC47	Jaguary	Brazil	8
WRC48	Khau Mac Kho	Vietnam	8
WRC49	Padi Perak	Indonesia	8
WRC50	Rexmont	USA	5
WRC51	Urasan 1	Japan	7
WRC52	Khau Tan Chiem	Vietnam	3
WRC53	Tima	Bhutan	5
WRC55	Tupa729	Bangladesh	7
WRC57	Milyang 23	Korea	4
WRC98	Deejiaohualuo Hong Cheuh	China	4
WRC99	Zai	China	5
WRC100	Vandaran	Sri Lanka	5

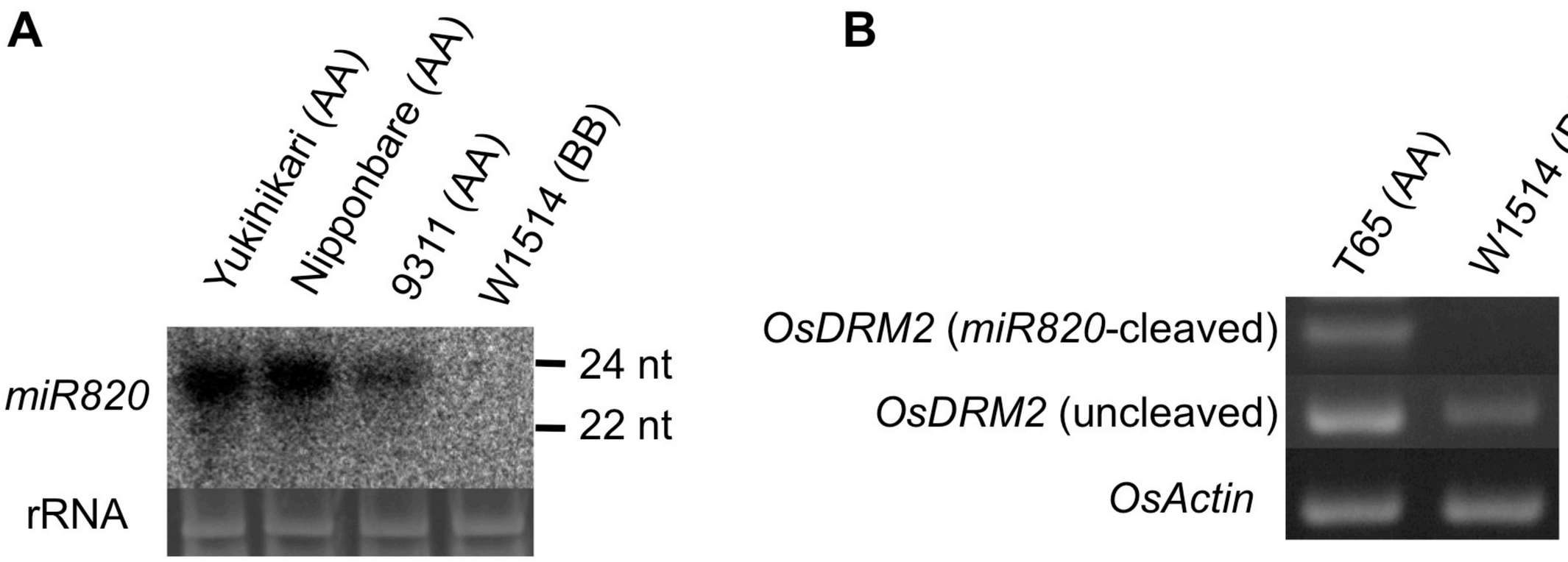


Figure 4.1. Expression and functional analysis of *miR820* in cultivated rice and wild-rice species containing BB genome.

(A) Northern blot analysis of *miR820* expression in AA and BB *Oryza* species. (B) Detection of *miR820*-cleaved *OsDRM2* mRNA by RNA ligation–mediated 5' RACE (upper panel) in AA and BB *Oryza* species. The same cDNA templates were used for PCR to amplify *OsDRM2* (middle panel) and *OsActin* (bottom panel) as controls for *OsDRM2* expression and RNA integrity. (C) qRT-PCR analysis measuring the expression level of *MIR820* in AA, BB and BBCC *Oryza* species. Relative expression level of *MIR820* is calculated by dividing the value with RT reaction by the value without RT reaction. Values are means of triplicate experiments, with bars showing standard errors.

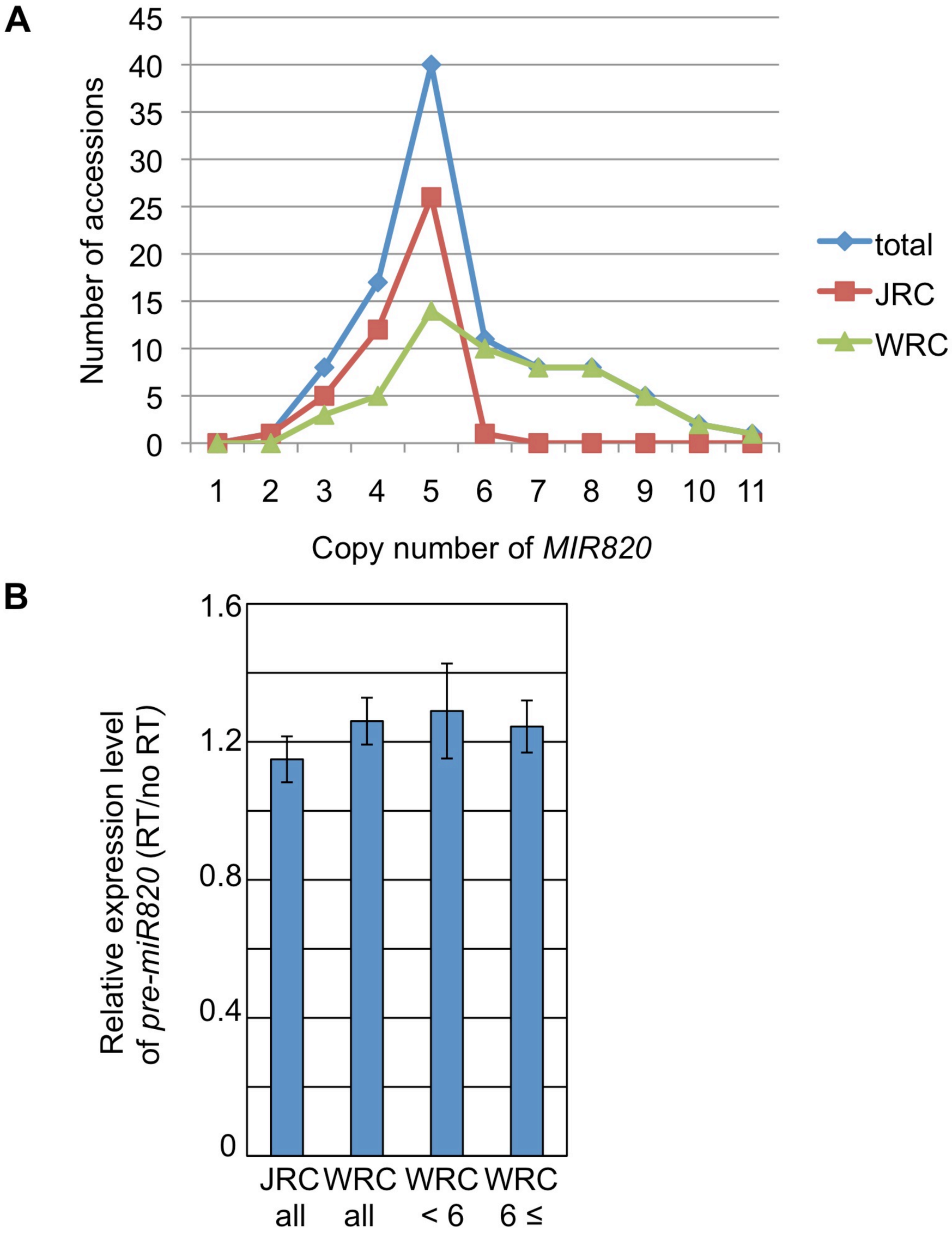


Figure 4.2. Copy number and Expression level of *MIR820* in the collections of cultivated rice accessions, JRC and WRC.

(A) Variation of the copy number of CACTA transposons carrying *MIR820* was determined by Southern blot analysis. Genomic DNAs from NIAS rice core collection were digested with *EcoRI* and *EcoRV*, and probed with *pre-miR820* DNA fragment. Distribution of the number of bands in Japanese rice collection (JRC), World rice collection (WRC) and the combination of both collection (total) are indicated by red, green and blue lines, respectively. (B) qRT-PCR analysis measuring the expression level of *MIR820* in the collections of cultivated rice accessions, JRC and WRC. To see the correlation between the copy number and the expression level of *MIR820*, the mean expression level of JRC, WRC, WRC (≤ 5 copies) and WRC (6 copies >) are indicated. Relative expression level of *MIR820* is calculated by dividing the value with RT reaction by the value without RT reaction. Values are means of triplicate experiments, with bars showing standard errors.

CHAPTER 5

GENERAL DISCUSSION

General discussion

Interplay between Transposons and host genome

1) Running battle between TEs and host genome through *miR820* in rice

Suppressing mechanisms of TE by host can be mainly divided into two groups, post-transcriptional suppression and transcriptional suppression. In both mechanisms, small RNA derived from TE's loci is related to starting up the suppression. In plants, double-stranded RNA (dsRNA) derived from TE's loci and small interfering RNA (siRNA) produced from this dsRNA by DICER take part in both mechanisms. These small RNAs can suppress the TE's activity by degrading the transcripts derived from TE's loci (Slotkin et al., 2005). Furthermore, inside the nucleus, the small RNAs recognize the RNA transcribing from TE's loci, recruit *de novo* DNA methyltransferase to the TE's loci where themselves were produced for inducing methylation of the cytosine residue at this loci, and make the TE silent at transcriptional level (Pikaard et al., 2008; Law and Jacobsen, 2010). By the mechanism mentioned above, small RNA (siRNA) produced from TE's loci take part in suppression of TEs.

I ascertained that miRNAs called *miR820* were produced from TE's loci during the comprehensive analysis of rice small RNA (Nosaka et al., 2012). Five copies of *miR820* are found on the reference genome of Nipponbare cultivar, whose genome annotation is completed among the *Oryza sativa* variety. All five copies of *miR820* are located at the same position inside the CACTA DNA transposon. This suggests that, while the CACTA transposon, where *miR820* is located, proliferated, *miR820* increased its copy

number, simultaneously. *miR820* is predicted to target *OsDRM2* encoding *de novo* DNA methyltrasnferase. Since *de novo* DNA methyltrasnferase has a pivotal role in host's defense machinery against exogenous genes, I set up the following hypothesis. Generally, small RNA (siRNA) derived from TE starts up the host's defense machinery against exogenous genes. However, I supposed that, in case of *miR820*, although it derived from TE's loci, it suppressed the host's silencing machinery against TEs. Actually, I confirmed; *miR820* cleaves mRNA of *OsDRM2*, *OsDRM2* with mutation in *miR820* recognition site increases its expression, and over expression of *miR820* decreases the expression of *OsDRM2*. Moreover, through the analysis using transgenic rice plants in which *OsDRM2* expression was reduced by RNAi, I confirmed that *OsDRM2* was related to transcriptional suppression of various transposon's gene in rice. All of these results support the hypothesis mentioned above. In Nipponbare, whose genome annotation is completed, both *miR820* and *OsDRM2* were expressed within all the tissues tested. Thus, the regulation of *OsDRM2* expression by *miR820* is likely to be a rheostat-like system rather than an on-off binary switch-like system. On the other hand, TE transposing actively has not discovered in Nipponbare until now. Then, is there a meaning for reducing *OsDRM2* expression by *miR820* in wild-type rice? Comparing the countermeasures of TEs with the silencing suppressor of viruses well known, several differences can be assumed. In case of viruses, even if the blocking of RNA silencing machinery by silencing suppressor reduces the adaptation rate of host, it is not a big issue. Because once they are replicated, they can infect another individual. In fact, RNA silencing machinery plays a vital role in expression of host's endogenous genes, such as miRNA-mediated gene expression. While in the case of TEs, as horizontal transmission hardly occurs, reducing the adaptation rate has huge effect on

their proliferation. Therefore, powerful suppresser of RNA silencing like virus's silencing suppressor is not a suitable countermeasure for TE, whereas the regulation of *OsDRM2* expression by *miR820* is precisely a suitable countermeasure for TE. This is because *miR820* only has effect on *de novo* DNA methylation, which is a restricted point for regulation, and does not block the whole machinery of RNA silencing. Moreover, the regulation by *miR820* is a rheostat-like regulation system. From these points, the regulation by *miR820* has avoided the problem mentioned above.

2) Emergence of reciprocal relationship as a result of the running battle

The regulation by *miR820* is extremely reasonable as a countermeasure of TE. However, *miR820* is not found in other plants except rice. Thus, it is thought that the regulation of *OsDRM2* expression by *miR820* is evolved uniquely, at least, in CACTA transposon of *Oryza sativa* including Nipponbare. How about wild rice in genus *Oryza*? I examined the sequence of *miR820* and its target site on *OsDRM2* in various kinds of wild rice. As a result, among all the wild rice examined in genus *Oryza*, the regulation of *OsDRM2* by *miR820* was conserved. However, in some of the wild rice, *Oryza punctata*, several nucleotide substitutions were found in the sequence of *miR820* and its target site. I focused on the position and the effect of these substitutions. In *O. punctata* comparing with Nipponbare, there were three and five nucleotide subtitutions in *miR820* and its target site sequence, respectively. However, the degree of base-pairing between *miR820* and its target site was conserved even in *O. punctata*. This suggests that, in *O. punctata*, the sequences of *miR820* and its target site have co-evolved to maintain the ability to recognize each other, that is, it indicates the regulation of *OsDRM2* by *miR820* is functional. At the moment, it is difficult to explain clearly why both sequences of

miR820 encoded by TE and *OsDRM2* encoded by host have accumulated these particular nucleotide substitutions. At least, I can suggest that these substitutions accumulated because not only TE but also host was put under the selection pressure against the increase of *OsDRM2* expression. Then, in what kind of situation does the host undergo the selection against the increase of *OsDRM2* expression?

So far, since TEs selfishly increase their copy on the host genome, for host, TE's activity was taken as a neutral factor or a negative factor that threaten the genome stability of host (Slotkin and Martienssen, 2007). However, knowledge is accumulating, which indicates that the role of TEs in host is more close to positive. For example, it is reported that several transcription factors possessing obvious function in host used to be genes related to transposition of TEs (Joly-Lopez et al., 2012). This suggests that host has manipulated the genes in TEs. There are also reports suggesting that TE can add sequences that regulate the expression of genes related to environmental response and so on to host's genes. Thus, it is no wonder that *miR820,* which looks like taking part in selfish activity of TE, also has a function of benefiting the host side. Particularly in *O. punctata*, since the sequences of *miR820* and its recognition site in *OsDRM2* have co-evolved, it is expected that there was a situation that suppressing *de novo* methylation increased the adaptation rate of host. I hypothesize such situation as the following scenario. When TE or its related sequence is inserted near into the host's essential gene and the DNA methylation of that TE for silencing also suppresses the expression of the adjacent host's gene, the host might suppress the *OsDRM2* expression by using *miR820*. Speaking in a more positive meaning, in the history of *O. punctata,* it may have undergone selection pressure that forced to change the expression of host genes in an epigenetic way through *OsDRM2* for corresponding to the changes in its

living environment and so on. In this manner, while the host manipulates the gene in TEs, TEs can seize the opportunity to increase their copies. In fact, I confirmed that CACTA transposon with *miR820* has greatly increased its copy number in *O. punctata*. Thus, as long as TEs are genomic parasites, it is quite possible that the interplay between host and TEs will exist, and in some cases, the interplay between them may develop into a reciprocal relationship.

Conclusion

Until now, discussions on RNA silencing were mainly about its role in host's defense machinery. However, in fact, RNA silencing is a front line of the running battle between host genome and genomic parasites such as viruses and TEs. It is also clear from the findings that, in viruses, silencing suppressors, which have extremely various effective points to oppose RNA silencing, have evolved uniquely. As for TE's countermeasures against the host, many things are still unknown. The regulation by *miR820* might be a rare case that can be explained in a simple schema. However, the functions of many genes encoded by TEs are unknown. In future, functional analysis of these genes from the viewpoint of countermeasure against the host, especially, analysis from the viewpoint of interplay with the host will be more important.

REFERENCES

Abe, M., Yoshikawa, T., Nosaka, M., Sakakibara, H., Sato, Y., Nagato, Y., and Itoh, J.-I. (2010) *WAVY LEAF 1*, an ortholog of *Arabidopsis HEN1*, regulates shoot development through maintaining microRNA and *trans*-acting siRNA accumulation in rice. Plant Physiol. **154**, 1335-1346.

Allen, E., Xie, Z., Gustafson, A.M., and Carrington, J.C. (2005) microRNA-directed phasing during trans-acting siRNA biogenesis in plants. Cell **121**, 207-221.

Almeida, R., and Allshire, R.C. (2005) RNA silencing and genome regulation. Trends Cell Biol. **15**, 251-258.

Ambros, V., Bartel, B., Bartel, D.P., Burge, C.B., Carrington, J.C., Chen, X.M., Dreyfuss, G., Eddy, S.R., Griffiths-Jones, S., Marshall, M., et al. (2003) A uniform system for microRNA annotation. RNA **9**, 77-279.

Aravin, A.A., Hannon, G.J., and Brennecke, J. (2007) The Piwi-piRNA pathway provides an adaptive defense in the transposon arms race. Science 318, 761-764.

Brennecke, J., Aravin, A.A., Stark, A., Dus, M., Kellis, M., Sachidanandam, R., and Hannon, G.J (2007) Discrete small RNA-generating loci as master regulators of transposon activity in Drosophila. Cell 128, 1089-1103.

Cao, X.F., Springer, N.M., Muszynski, M.G., Phillips, R.L., Kaeppler, S., and Jacobsen, S.E. (2000) Conserved plant genes with similarity to mammalian *de novo* DNA methyltransferases. Proc. Natl. Acad. Sci. USA **97**, 4979-4984.

Cao, X.F., and Jacobsen, S.E. (2002) Role of the *Arabidopsis* DRM methyltransferases

in de novo DNA methylation and gene silencing. Curr. Biol. **12**, 1138-1144.

Carthew, R.W., and Sontheimer, E.J. (2009) Origins and mechanisms of miRNAs and siRNAs. Cell **136**, 642-655.

Chan, S.W.L., Zilberman, D., Xie, Z.X., Johansen, L.K., Carrington, J.C., and Jacobsen, S.E. (2004) RNA silencing genes control *de novo* DNA methylation. Science 303, 1336.

Chellappan, P., Xia, J., Zhou, X., Gao, S., Zhang, X., Coutino, G., Vazquez, F., Zhang, W., and Jin, H. (2010) siRNAs from miRNA sites mediate DNA methylation of target genes. Nucleic Acids Res. **38**, 6883-6894.

Doolittle, W.F., and Sapienza, C. (1980) Selfish genes, the phenotype paradigm and genome evolution. Nature **284**, 601-603.

Ebana, K., Kojima, Y., Fukuoka, S., Nagamine, T., and Kawase, M. (2008) Development of mini core collection of Japanese rice landrace. Breed. Science **58**, 281-291.

Feschotte, C., Jiang, N., and Wessler, S.R. (2002) Plant transposable elements: Where genetics meets genomics. Nat. Rev. Genet. **3**, 329-341.

Ge, S., Sang, T., Lu, B.R., and Hong, D.Y. (1999) Phylogeny of rice genomes with emphasis on origins of allotetraploid species. Proc. Natl Acad. Sci. USA **96**, 14400-14405.

Henderson, I.R., Deleris, A., Wong, W., Zhong, X., Chin, H.G., Horwitz, G.A., Kelly, K.A., Pradhan, S., and Jacobsen, S.E. (2010) The *De novo* cytosine methyltransferase DRM2 requires intact UBA domains and a catalytically mutated paralog DRM3 during RNA-directed DNA methylation in *Arabidopsis thaliana*. PLoS Genet. **6**, e1001182.

Hiei, Y., Ohta, S., Komari, T., and Kumashiro, T. (1994) Efficient transformation of rice (*Oryza sativa L.*) mediated by *Agrobacterium* and sequence analysis of the boundaries of the T-DNA. Plant J. **6**, 271-282.

Hoen, D.R., Park, K.C., Elrouby, N., Yu, Z.H., Mohabir, N., Cowan, R.K., and Bureau, T.E. (2006) Transposon-mediated expansion and diversification of a family of ULP-like genes. Mol. Biol. Evol. **23**, 1254-1268.

Joly-Lopez, Z., Forczek, E., Hoen, D.R., Juretic, N., and Bureau, T.E. (2012) A gene family derived from transposable elements during early angiosperm evolution has reproductive fitness benefits in *Arabidopsis thaliana*. PLoS Genet. **8**, e1002931.

Kasschau, K.D., Xie, Z.X., Allen, E., Llave, C., Chapman, E.J., Krizan, K.A., and Carrington, J.C. (2003) P1/HC-Pro, a viral suppressor of RNA silencing, interferes with Arabidopsis development and miRNA function. Dev. Cell **4**, 205-217.

Kidwell, M.G. (2002) Transposable elements and the evolution of genome size in eukaryotes. Genetica **115**, 49-63.

Kim, H., Miguel, P.S., Nelson, W., Collura, K., Wissotski, M, Walling J.G., Kim, J.P., Jackson, S.A., Soderlund, C., and Wing, R.A. (2007) Comparative physical mapping between *Oryza sativa* (AA genome type) and *O. punctata* (BB genome type). Genetics **176**, 379-390.

Kojima, Y., Ebana, K., Fukuoka, S., Nagamine, T., and Kawase, M. (2005) Development of an RFLP-based rice diversity research set of germplasm. Breed. Science **55**, 431-440.

Kouchi, H., and Hata, S. (1993) Isolation and characterization of novel nodulin cDNAs representing genes expressed at early stages of soybean nodule development. Mol.

Gen. Genet. **238**, 106-119.

Kuramochi-Miyagawa, S., Watanabe, T., Gotoh, K., Totoki, Y., Toyoda, A., Ikawa, M., Asada, N., Kojima, K., Yamaguchi, Y., Ijiri, T.W. et al. (2008) DNA methylation of retrotransposon genes is regulated by Piwi family members MILI and MIWI2 in murine fetal testes. Genes Dev. **22**, 908-917.

Kurihara, Y., and Watanabe, Y. (2004) *Arabidopsis* micro-RNA biogenesis through Dicer-like 1 protein functions. Proc. Natl Acad. Sci. USA **101**, 12753-12758.

Kurihara, Y., Takashi, Y., and Watanabe, Y. (2006) The interaction between DCL1 and HYL1 is important for efficient and precise processing of pri-miRNA in plant microRNA biogenesis. RNA **12**, 206-212.

Lacombe, S., Nagasaki, H., Santi, C., Duval, D., Piegu, B., Bangratz, M., Breitler, J.C., Guiderdoni, E., Brugidou, C., Hirsch, J. et al. (2008) Identification of precursor transcripts for 6 novel miRNAs expands the diversity on the genomic organization and expression of miRNA genes in rice. BMC Plant Biol. **8**, 123.

Law, J. A., and Jacobsen, S. E. (2010) Establishing, maintaining and modifying DNA methylation patterns in plants and animals. Nat. Rev. Genet. **11**, 204-220.

Lee, Y., Kim, M., Han, J.J., Yeom, K.H., Lee, S., Baek, S.H. and Kim, V.N. (2004) MicroRNA genes are transcribed by RNA polymerase II. EMBO J. **23**, 4051-4060.

van Leeuwen, H., Monfort, A., and Puigdomenech P. (2007) Mutator-like elements identified in melon, Arabidopsis and rice contain ULP1 protease domains. Mol. Genet. Genomics **277**, 357-364.

Li, J., Yang, Z., Yu, B., Liu, J., and Chen, X. (2005) Methylation protects miRNAs and siRNAs from a 3′-end uridylation activity in *Arabidopsis*. Curr. Biol. **15**,

1501-1507.

Lisch, D. (2009) Epigenetic regulation of transposable elements in plants. Annu. Rev. Plant Biol. **60**, 43-66.

Lisch, D., and Slotkin, R.K. (2011) Strategies for silencing and escape: The ancient struggle between transposable elements and their hosts. Intl Rev. Cell Mol. Biol. **292**, 119-152.

Liu, J., He, Y.H., Amasini, R., and Chen, X.M. (2004) siRNA targeting an intronic transposon in the regulation of natural flowing behavior in Arabidopsis. Genes Dev. **18**, 2873-2878.

Llave, C., Kasschau K.D., and Carrington, J.C. (2000) Virus-encoded suppressor of posttranscriptional gene silencing targets a maintenance step in the silencing pathway. Proc. Natl Acad. Sci. USA **97**, 13401-13406.

Luo, Y.C., Zhou, H., Li, Y., Chen, J.Y., Yang, J.H., Chen, Y.Q., and Qu, L.H. (2006) Rice embryogenic calli express a unique set of microRNAs, suggesting regulatory roles of microRNAs in plant post-embryogenic development. FEBS Lett. **580**, 5111-5116.

Matzke, M., Kanno, T., Claxinger, L., Huettel, B., and Matzke, A.J.M. (2009) RNA-mediated chromatin-based silencing in plants. Curr. Opin. Cell Biol. **21**, 367-376.

Meyers, B.C., Axtell, M.J., Bartel, B., Bartel, D.P., Baulcombe, D., Bowman, J.L., Cao, X., Carrington, J.C., Chen, X., Green, P.J., et al. (2008) Criteria for annotation of plant microRNAs. Plant Cell **20**, 3186-3190.

Miura, K., Agetsuma, M., Kitano, H., Yoshimura, A., Matsuoka, M., Jacobsen, S.E., and Ashikari, M. (2009) A metastable *DWARF1* epigenetic mutant affecting plant

stature in rice. Proc. Natl Acad. Sci. USA **106**, 11218-11223.

Nagaki, K., Cheng, Z.K., Ouyang, S., Talbert, P.B., Kim, M., Jones, K.M., Henikoff, S., Buell, C.R., and Jiang, J.M. (2004) Sequencing of a rice centromere uncovers active genes. Nat. Genet. **36**, 138-145.

Nakagawa, T., Suzuki, T., Murata, S., Nakamura, S., Hino, T., Maeo, K., Tabata, R., Kawai, T., Tanaka, K., Niwa, Y. et al. (2007) Improved gateway binary vectors: High-performance vectors for creation of fusion constructs in transgenic analysis of plants. Biosci. Biotech. Biochem. **71**, 2095-2100.

Nosaka, M., Itoh, J.-I., Nagato, Y., Ono, A., Ishiwata, A., and Sato, Y. (2012) Role of transposon-derived small RNAs in the interplay between genomes and parasitic DNA in rice. PLoS Genet. **8**, e1002953.

Orgel, L.E., and Crick, F.H.C. (1980) Selfish DNA – The ultimate parasite. Nature **284**, 604-607.

Pikaard, C.S., Haag, J.R., Ream, T., and Wierzbicki, A.T. (2008) Roles of RNA polymerase IV in gene silencing. Trends Plant Sci. **13**, 390-397.

Plasterk, R.H.A. (2002) RNA silencing: The genome's immune system. Science **296**, 1263-1265.

Rice Annotation Project (2008) The rice annotation project database (RAP-DB): 2008 update. Nucleic Acids Res. **36**, D1028-D1033.

Saito, K., Siomi, M.C. (2010) Small RNA-mediated quiescence of transposable elements in animals. Dev. Cell **19**, 687-697.

Saitou, N., and Nei, M. (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. **4**, 406-425.

Sharma, R., Singh, R.K.M., Malik, G., Deveshwar, P., Tyagi, A.K., Kapoor, S., and

Kapoor, M. (2009) Rice cytosine DNA methyltransferases - gene expression profiling during reproductive development and abiotic stress. FEBS J. **276**, 6301-6311.

Slotkin, R.K., Freeling M, and Lisch, D. (2005) Heritable transposon silencing initiated by a naturally occurring transposon inverted duplication. Nature Genet. **37**, 641-644.

Slotkin, R.K., and Martienssen, R. (2007) Transposable elements and the epigenetic regulation of the genome. Nat. Rev. Genet. **8**, 272-285.

Sun, G. (2012) MicroRNAs and their diverse functions in plants. Plant Mol. Biol. **80**, 17-36.

Vazquez, F., Blevins, T., Ailhas, J., Boller, T., and Meins, F., Jr. (2008) Evolution of *Arabidopsis MIR* genes generates novel microRNA classes. Nucleic Acids Res. **36**, 6429-6438.

Tsukahara, S., Kobayashi, A., Kawabe, A., Mathieu, O., Miura, A., and Kakutani, T. (2009) Bursts of retrotransposition reproduced in *Arabidopsis*. Nature **461**, 423-426.

Voinnet, O., Lederer C., and Baulcombe, D.C. (2000) A viral movement protein prevents spread of the gene silencing signal in *Nicotiana benthamiana*. Cell **103**, 157-167.

Voinnet, O. (2008) Use, tolerance and avoidance of amplified RNA silencing by plants. Trends Plant Sci. **13**, 317-328.

Voinnet, O. (2009) Origin, biogenesis, and activity of plant microRNAs. Cell **136**, 669-687.

Waterhouse, P.M., Wang, M.B., and Lough, T. (2001) Gene silencing as an adaptive

defence against viruses. Nature **411**, 834-842.

Wierzbicki, A.T. (2012) The role of long non-coding RNA in transcriptional gene silencing. Curr. Opin. Plant Biol. **15**, 517-522.

Wu, L., Zhou, H., Zhang, Q., Zhang, J., Ni, F., Liu, C., and Qi, Y. (2010) DNA methylation mediated by a microRNA pathway. Mol. Cell **38**, 465-475.

Yang, Z., Ebright, Y.W., Yu, B., and Chen, X. (2006) HEN1 recognizes 21-24 nt small RNA duplexes and deposits a methyl group onto the 2' OH of the 3' terminal nucleotide. Nucleic Acids Res. **34**, 667-675.

Yu, B., Yang, Z.Y., Li, J.J., Minakhina, S., Yang, M.C., Padgett R.W., Steward, R., and Chen, X.M. (2005) Methylation as a crucial step in plant microRNA biogenesis. Science **307**, 932-935.

Zilberman, D., and Henikoff, S. (2004) Silencing of transposons in plant genomes: kick them when they're down. Genome Biol. **5**, 249.

ACKNOWLEDGEMENT

I would like to express my deepest appreciation to my supervisor, Dr. Yutaka Sato for his elaborated guidance, considerable encouragement and invaluable discussion that make my research of great achievement and my study life unforgettable.

I would like to express my sincere gratitude to my supervisors, Drs. Mikio Nakazono, Yoshiaki Inukai and Hidemi Kitano for their encouragement, insightful comments and unstinted affection.

Finally, I am very grateful to past and present members of the Plant Genetics and Breeding Laboratory for their sincere assistance and supports during my study.

List of Publications

1) **Nosaka, M., Itoh, J-I., Nagato, Y., Ono, A., Ishiwata, A., and Sato, Y.** (2012) Role of transposon-derived small RNAs in the interplay between genomes and parasitic DNA in rice., *PLoS Genet.* **8**, e1002953.

2) **Nosaka, M., Ono, A., Ishiwata, A., Shimizu-Sato, S., Ishimoto, K., Noda, Y., and Sato, Y.** (2013) Expression of the rice microRNA *miR820* is associated with epigenetic modifications at its own locus., *Genes Genet. Syst.* In press.

3) **Nosaka, M., Ono, A., Ishiwata, A., Shimizu-Sato, S., Ishimoto, K., Noda, Y., and Sato, Y.** (2013) The copy number of rice CACTA transposons caryying *MIR820* in does not correlate with *MIR820* expression., *Plant Signal Behav.* In press.

List of Publications (appendix)

1. Original Papers

1) **Nagasaki, H., Itoh, J-I., Hayashi, K., Hibara, K-I., Satoh-Nagasawa, N., Nosaka, M., Mukohata, M., Ashikari, M., Kitano, H., Matsuoka, M., Nagato, Y., and Sato Y.** (2007) The small interfering RNA production pathway is required for shoot meristem initiation in rice. *Proc. Natl. Acad. Sci. USA* **104**, 14867-14871.

2) **Abe, M., Yoshikawa, T., Nosaka, M., Sakakibara, H., Sato, Y., Nagato, Y., Itoh, J-I.** (2010) *WAVY LEAF 1*, an ortholog of Arabidopsis *HEN1*, regulates shoot development by maintaining microRNA and trans-acting siRNA accumulation in rice. *Plant Physiol.* **154**, 1335-1346.

3) **Ishiwata, A., Ozawa, M., Nagasaki, H., Kato, M., Noda, Y., Yamaguchi, T., Nosaka, M., Shimizu-Sato, S., Nagasaki, A., Maekawa, M., Hirano, H. Y., and Sato, Y.** (2013) Two *WUSCHEL-related homeobox* genes, *narrow leaf2* and *narrow leaf3*, control leaf width in rice. *Plant Cell Physiol.* **54**, 779-792.

2. Others

1) **野坂実鈴, 佐藤豊 (2008)** 植物の内在性小分子RNA. 無敵のバイオテクニカルシリーズRNA実験ノート, 羊土社 **下巻**, 22-23.

2) **佐藤豊, 野坂実鈴, 伊藤純一** (2008) small RNAによる茎頂分裂組織形成制御. 細胞工学別冊植物細胞工学シリーズ・植物のエピジェネティクス, 秀潤社 **24巻**, 106-114.

3) **佐藤豊, 野坂実鈴, 伊藤純一** (2008) RNAサイレンシングの経路を介したイネの茎頂分裂組織構築機構. ブレインテクノニュース, 独立行政法人農業・食品産業技術総合研究機構 **128号**, 6-10.