creative commons

COMMONS DEED

저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:

저작자표시. 귀하는 원저작자를 표시하여야 합니다.

비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.

변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 이용허락규약(Legal Code)을 이해하기 쉽게 요약한 것입니다.

Disclaimer

dCollection

농 학 석 사 학 위 논 문

전사체 분석에 기반한 질소 공급 조건 변화에 따른 벼 유전자 및 Non-coding RNA의 변화 관찰

Transcriptome-wide Analysis of Protein-coding Genes and Non-coding RNAs in Response to Nitrogen Supply Status in Rice

2014년 8월

서울대학교 대학원

농생명공학부 응용생명화학전공

신 상 윤

Thesis for the Degree of Master of Science

Transcriptome-wide Analysis of Protein-coding Genes and Non-coding RNAs in Response to Nitrogen Supply Status in Rice

Advisor : Chanseok Shin

A thesis submitted to the faculty of the Seoul National University Graduate School in partial fulfillment of the requirement for the degree of Master of Science in the Department of Agricultural Biotechnology.

By

Sang-yoon Shin

Major in Applied Life Chemistry
Department of Agricultural Biotechnology
Seoul National University Graduate School

August 2014
Approved by Major Advisor

농 학 석 사 학 위 논 문

전사체 분석에 기반한 질소 공급 조건 변화에 따른 벼 유전자 및 Non-coding RNA의 변화 관찰

Transcriptome-wide Analysis of Protein-coding Genes and Non-coding RNAs in Response to Nitrogen Supply Status in Rice

지도교수 신 찬 석

이 논문을 석사학위논문으로 제출함.

2014년 8월

서 울 대 학 교 대 학 원

농생명공학부 응용생명화학전공

신 상 윤

신상윤의 논문을 석사학위논문으로 인준함.

2014년 8월

위 원 장 _____ (인)

부 위 원 장 _____ (인)

위 원 _____ (인)

ABSTRACT

Transcriptome-wide Analysis of Protein-coding Genes and Non-coding RNAs in Response to Nitrogen Supply Status in Rice

Sang-Yoon Shin

Major in Applied Life Chemistry

Department of Agricultural Biotechnology

The Graduate School

Seoul National University

Nitrogen is a key nutrient in plant growth and development, and usage of nitrogen fertilizer is directly involved in crop yields. To increase crop yield, the enormous amount of chemical nitrogen fertilizer has been used. However, limitation on nitrogen-use-efficiency of crop plant make amounts of fertilizer to be remained in soil, and leaching of those remaining nitrogen sources to river and ocean occur environmental problems, such as eutrophication.

Non-coding RNAs, including miRNAs and long non-coding RNAs, are a class of RNAs that are not be translated into protein. In recent studies, it is known that some of those miRNAs and long non-coding RNAs function as key regulator and regulate expression level of its target gene through various molecular mechanisms. Moreover, most of those non-coding RNAs also respond to various stimuli and stresses, and they could control the changes of plant physiology in response to those stimuli and stresses. However, roles of non-coding RNAs on nitrogen supply status in rice have not been deeply studied yet.

To comprehensively investigate and understand transcriptome-wide responses of rice to nitrogen starvation, multimodal RNA-Seq methods were applied: strand-specific RNA-Seq, small RNA-Seq, and 2P-Seq. Using strand-specific RNA-Seq, expression profiling and analysis of annotated genes including nitrogen metabolism-related genes were performed. Also, nitrogen-responsive miRNAs were identified in nitrogen-starved rice roots, and the relationship between representative nitrogen supply status-responsive miRNA, miR169, and its target genes, NF-YAs, were analyzed by expression pattern analysis and 5' RLM-RACE. Using strand-specific RNA-Seq dataset, identification of putative lncRNA-deriving loci and analysis of their expression patterns were performed. Moreover, identifying 3' end of transcripts was tried by analyzing 2P-Seq data and those sites detected by 2P-Seq were confirmed by 3' RACE against putative lncRNA

candidate. Dataset analyzed in this study will be used as a fundamental dataset for further studies about roles of genes and non-coding RNAs in response of rice to change of nitrogen supply status.

Keyword: Rice, Nitrogen, Strand-specific RNA-Seq, Small RNA-Seq, 2P-Seq, MicroRNA, Long non-coding RNA

Student Number: 2012-21167

CONTENTS

ABSTRACT.....1

CONTENTS.....4

LIST OF TABLES AND FIGURES.....5

LIST OF ABBREVIATIONS.....7

INTRODUCTION.....9

MATERIALS AND METHODS.....16

RESULTS AND DISCUSSION.....22

SUMMARY.....41

FIGURES AND LEGENDS.....43

TABLES.....68

REFERENCE.....76

SUPPLEMENTARY FIGURES AND LEGENDS.....88

ABSTRACT IN KOREAN.....93

LIST OF FIGURES AND TABLES

Figure 1. Experimental scheme and multimodal transcriptomic analysis used in this study.

Figure 2. Profiling of differentially expressed genes in nitrogen-starved leaves and roots.

Figure 3. Analysis of differentially expressed genes in leaves and roots.

Figure 4. Profiling of nitrogen metabolism-involved genes in leaves and roots.

Figure 5. Normalized small RNA read distribution of 13 small RNA-Seq libraries.

Figure 6. Profiling of nitrogen-responsive miRNAs in roots.

Figure 7. Expression patterns of miR169 family and that of its target genes, NF-YAs.

Figure 8. Identification of putative long non-coding RNAs in rice.

Figure 9. Results of 2P-Seq data analysis.

Figure 10. 2P-Seq could detect 3' end of putative long non-coding RNA.

Figure 11. Expression of *cis*-NAT$_{OsAMT1.1}$.

Figure 12. Genome comparison of *AMT1.1* locus between rice and other plants.

Table 1. List of oligonucleotide sequences used in this study.

Table 2. Read distribution of strand-specific RNA-Seq libraries.

Table 3. Read distribution of small RNA-Seq libraries.

Table 4. Read distribution of 2P-Seq libraries.

Table 5. Summary of putative long non-coding RNA loci identification results.

LIST OF ABBREVIATIONS

Os	*Oryza sativa*
N	Nitrogen
NUE	Nitrogen-use-efficiency
miRNA	MicroRNAs
LncRNA	Long non-coding RNA
LincRNA	Long intergenic non-coding RNA
IncRNA	Intronic non-coding RNA
cis-NAT	*Cis*-natural antisense transcript
Lf	Leaf
Rt	Root
Stv	Starvation of nitrogen
ReN	Resupply of nitrogen
WT	Wild type
FPKM	Fragments per kilobase of exon per million reads
RPKHM	Reads per kilobase of exon per hundred million reads
RP30M	Reads per 30 million reads
GO	Gene Ontology
CPC	Coding potential calculator

ORF	Open reading frame
RLM-RACE	RNA ligase-mediated rapid amplification of cDNA end

INTRODUCTION

Rice and Nitrogen

Rice is one of the major crops worldwide and is important food source for half of people in the world, especially in East Asia. Also, rice is one of important model plants for researching monocot species because of small, fully-sequenced genome. Many researchers in the world are studying and genetically engineering the rice for increasing rice productivity and developing stress-resistant rice, because of its importance as a major food source.

Nitrogen (N) is one of a key macronutrient of crop plant growth and development. In addition, supplement of N is closely involved with crop productivity. Due to the importance of the amount of available N in the growth and yield of crop plant, application of chemical N fertilizers on fields have greatly increased during the past five decades. These have resulted in significant increase of crop yield. However, due to low limitation of crop plant on soil N-use-efficiency (NUE) (Hoque et al., 2006), the proportion of available soil N used by crop plants is less than half of the available amount of N in soil, which leads to the reduction of N-fertilization efficiency of crop productivity (Tilman et al., 2002). In addition, leaching of remaining N into groundwater, river and ocean, causes various aqueous environmental problems such as eutrophication.

Increasing necessities of reducing the chemical N fertilizers usage have

facilitated studies to better understand mechanisms of plant in regulating NUE and to develop NUE-improved crop plant including rice. In recent years, researchers have attempted to make NUE-improved transgenic crops by ectopic over-expression of N metabolism-involved genes or N-responsive genes in rice (Hoque et al., 2006; Bi et al., 2009; Cai et al., 2009; Cai et al., 2010; Brauer et al., 2011; Fang et al., 2013), and to find candidate NUE-involved rice genes in a genome-wide level, microarray technology has been applied. Lian et al. (2006) profiled expression level of 10,422 genes in rice seedlings at early stage of low N stress using microarray, and Bi et al. (2009) performed cDNA microarray against N-starved rice, and showed improved NUE in *OsENOD93-1* overexpressing rice. However, those microarray experiments were only focused on previously known protein-coding genes, and could not detect novel transcripts or non-coding RNAs that might be involved in response of rice on N starvation or NUE. In recent studies, researchers prefer another transcriptomic approach, called RNA-Seq, rather than microarray, because of advantageousness of RNA-Seq relative to microarray; high-throughput dataset, single-nucleotide resolution, low background. Nevertheless, none of studies applying that advantageous technology in identifying rice genes and other novel transcripts responding to N-starvation have been reported yet.

Non-coding RNAs

Recent studies have discovered thousands of novel genomic loci from various organisms that are transcribed into transcripts having little or no protein-coding potential, called non-coding RNAs (Claverie, 2005; Mattick, 2005). These non-coding RNAs can be classified according to its transcript lengths, as small RNAs (18-40 nts), short RNAs (50-200 nts), and long non-coding RNAs (lncRNAs, longer than 200 nts). They are further classified by biogenesis pathways (miRNA/endo-siRNA/piRNA/diRNA) or genomic locations (lincRNA/incRNA/ *cis*-NAT/NAT-miRNA). Most of those non-coding RNAs have function in regulating gene expression level, and they play key roles in manipulating diverse spectrum of biological phenomenon, such as embryonic stem cell pluripotency, cell-cycle regulation, disease and stress-response, with a various molecular mechanisms. Discoveries of functionally important lncRNAs imply that numerous unknown, functional non-coding RNA genes exist in the genomic region referred to as “dark matter” (Johnson et al, 2005; Riddihough, 2005; Kapranov and St. Laurent, 2012). Because of this, a great number of researchers are trying to discover novel non-coding RNAs and identify their mechanisms and functions.

MicroRNAs

MicroRNA (miRNA) is a one of small non-coding RNA classes, which

negatively regulates the expression level of its target genes by destabilizing transcript and inhibiting protein synthesis. In plants, miRNA genes have their own transcription units which are transcribed by RNA polymerase II as primary miRNAs. The stem-loop structures in the primary miRNAs are then cleaved by *DICER-LIKE1* (*DCL1*) into precursor miRNAs (pre-miRNAs), which is cleaved again by *DCL1*, deriving 21-nt-long RNA duplexes, or cleaved by *DCL4*, deriving 24-nt-long RNA duplexes. Those 21 to 24-nt miRNA duplexes are stabilized by 2'-O-methylation of 3' end, which is catalyzed by *Hua Enhancer 1* (*HEN1*), and transported from the nucleus to the cytoplasm by *HASTY*. One of the duplex strands is then incorporated into the Argonaute (*AGO*), and those miRNA-AGO complexes, called RNA-induced silencing complex (RISC), then negatively regulate its target genes by complementary binding between miRNA and target transcript.

Roles of plant miRNAs on plant abiotic stress responses and resistances have been suggested by the identification of linkages between miRNA and its stress response-related target genes, such as miR399-*PHO2* in phosphate deficiency stress (Bari et al., 2006), miR398-*CSD1*/miR398-*CSD2* in oxidative stress (Sunkar et al., 2006), miR395-*SULTR2;1* in sulfate deprivation (Liang and Yu, 2010), and miR393-*AFB3* and miR167-*ARF8* in N starvation stress (Gifford et al., 2008; Vidal et al., 2010). In those previous studies, the expression levels of miRNAs

have dynamically changed in response to stress treatment, and those expressional changes have also affected on the expression level of its target genes which is closely involved in those stresses, suggesting that plant miRNAs serve as core regulator of plant's response to various stresses.

To understand responses and roles of plant miRNAs in N deficient stresses and N homeostasis, expression levels of miRNAs have been examined in plants and crops with using transcriptomic approaches such as microarray (Cai et al., 2012; Nischal et al., 2012) and small RNA-Seq (Jeong et al., 2011; Liang et al., 2012; Zhao et al., 2012). However, previous studies on rice miRNAs responding to N starvations were mainly focused on a few N starvation conditions or on a singular miRNA cases (Yan et al., 2014), thus could not provide sufficient information to understand dynamic, genome-wide changes of the miRNA pool.

Long non-coding RNAs

LncRNAs are generated from various genomic regions including intergenic / intronic region and opposite strand of protein-coding gene regions. According to their genomic origin, these lncRNAs are respectively classified as long intergenic non-coding RNA (lincRNA), intronic non-coding RNA (incRNA) and natural antisense transcript (NAT). Although imperfect action of transcript machineries can generate spurious, non-functional RNAs resembling lncRNAs, recent studies

have shown that some lncRNAs clearly have functions. Because of this, researchers are trying to identify novel lncRNAs in a genome-wide level, and investigate their mechanisms and functions both in animal and plant (Rinn et al., 2007; Franco-zorrilla et al., 2007; Huarte et al., 2010; Heo and Sung, 2011; Ulitsky et al., 2011; Jabnoune et al., 2013; Li et al., 2014). Especially, one of plant lncRNAs, *cis*-NAT$_{PHO1;2}$, which is expressed by phosphate starvation, participates in regulating phosphate homeostasis of rice by enhancing translationn of *OsPHO1;2* genes. This case suggests that lncRNAs could respond to and play a role in plant nutrient-deficient stresses.

Properties of these lncRNAs are similar to common features of mRNA; they are mostly 5' capped, and even are polyadenylated and spliced, implying that the common mRNA biogenesis machineries are shared in the biogenesis of lncRNAs. Because of these structural similarities between mRNA and lncRNA, lncRNAs are discovered in typical cDNA cloning, tiling array, and RNA-Seq datasets used in detecting and quantifying the expression level of mRNA (Okazaki et al., 2002; Carninci et al., 2005). Due to this, genome-wide identification of novel lncRNAs have been tried in Arabidopsis (Liu et al., 2012; Wang et al., 2014), maize (Li et al., 2014), wheat (Xin et al., 2011) and rice (Lu et al., 2012) using RNA-Seq methods. However, none of transcriptomic approaches on N starvation-responsive lncRNAs in rice had been tried yet.

Purpose of this study

Since transcriptomic researches through RNA-Seq on various N supply condition-treated rice had not been tried, this study was aimed to investigate and comprehensively understand responses of rice to N starvation in a transcriptome level using RNA-Seq. To do this, multimodal RNA-Seq methods were applied in this study: strand-specific RNA-Seq, small RNA-Seq, and 2P-Seq. Using strand-specific RNA-Seq, expression profiling and analysis of annotated genes including N metabolism-related genes were performed. Also, N-responsive miRNAs were identified from N-starved rice roots, using small RNA-Seq dataset. Using strand-specific RNA-Seq dataset, I tried to identify novel, putative lncRNA-deriving loci and analyze their expression patterns in roots and leaves of N-starved rice. Moreover, genome-wide identification of polyadenylated transcripts' 3' ends was tried by analyzing 2P-Seq, and 2P-Seq signal-detected sites on putative intergenic lncRNA locus was validated by 3' RACE.

MATERIALS AND METHODS

Plant materials and growth conditions

Rice (*Oryza sativa* cv Nipponbare) was grown under solution described in Yoshida et al. (1976), containing 0.513mM K_2SO_4, 0.998 mM $CaCl_2$, 1.643 mM $MgSO_4$, 0.075 μM $(NH_4)_6Mo_7O_{24}$, 0.25 mM $NaSiO_3$, 0.009 mM $MnCl_2$, 0.019 μM H_3BO_3, 0.155 μM $CuSO_4$, 0.152 μM $ZnSO_4$, and 0.125 mM EDTA-Fe, 0.323 mM NaH_2PO_4, 1.425mM NH_4NO_3. Germinated seeds were grown in soil for 5 days and then transferred into tap water for 3 days before transferred to a NH_4NO_3-containing or NH_4NO_3-removed solution. After growing rice on NH_4NO_3-containing (+N condition) or NH_4NO_3-removed (-N condition) solution for 10 days, those were transferred to NH_4NO_3-removed (-N) and NH_4NO_3-containing (+N) solution, respectively. Rice transferred to NH_4NO_3-removed (-N) were grown for 1 day, 3 days, 5 days and 7 days, and rice transferred to NH_4NO_3-containing (N-resupply) solution were grown for 6 h, 12 h, 24 h, 36 h, 48 h, and 72 h. 3 biological replicates were generated for each time points of N starvation and resupply, and those replicates were mixed into one.

Total RNA isolation

Roots and leaves of N-starved rice sample were harvested and ground with liquid N using mortar and pestle. Total RNA was extracted from ground tissue

using the TRI Reagent (Ambion), according to the manufacturer's instructions. The integrity and quality of the extracted total RNA was checked using NanoDrop 2000 spectrophotometer and non-denaturing agarose gel electrophoresis. Total RNAs that the value of A260/A280 ratio was >1.8 after DNase I treatment were only used in library construction.

Strand-specific RNA-Seq library construction, sequencing and analysis

Strand-specific RNA-Seq libraries were constructed according to a modified protocol of Zhong et al. (2011). Briefly, poly-(A)+ RNA was selected from 5 ug of total RNA using Dynabeads oligo$(dT)_{25}$ (Invitrogen). Poly-(A)-selected RNA was fragmented by incubating at 94 ℃ for 5 min and 1st-strand cDNA was generated using SuperScript III (Invitrogen). 2nd-strand cDNA was generated using dUTP instead of dTTP, which was removed by Uracil DNA Glycosylase (Enzymatics) reaction after adapter ligation. During a final PCR enrichment, each library was indexed using barcode sequences-containing primers (6-mer DNA sequence) for multiplex sequencing. Constructed libraries were analyzed by 101-cycle paired-end sequencing on the Illumina HiSeq 2500 at the NICEM.

Small RNA-Seq library construction, sequencing and analysis

Small RNA libraries were constructed using size-fractionated small RNAs

(from 18 to 30 nt) from 20~25 ug of root total RNA. 1st-strand cDNA was generated using SuperScript III (Invitrogen) after 3' adapter and 5' adapter were sequentially ligated to 3' end and 5' end of fractionated small RNAs, respectively. Synthesized cDNA was amplified by Phusion polymerase (NEB), and amplified small RNA-Seq libraries were purified on 6% native PAGE-gel and analyzed by 51-cycle single-end sequencing on the Illumina HiSeq 2500 at the NICEM.

2P-Seq library construction, sequencing and analysis

2P-Seq libraries were constructed using poly-(A)+ RNA selected from approximately 60 ug of WT root total RNA and WT leaf total RNA, and pooled total RNA mix (including all samples in Figure 1A), respectively. RNA was treated with 0.5 U of RNase T1 (biochemical grade, Ambion) for 20 min at 22 ℃. The reaction was quenched and precipitated using the T1 Precipitation/Inactivation Buffer. Partially digested RNA was reverse-transcribed using 2P-RT1p primer and SuperScript III (Invitrogen). Alkaline hydrolysis was performed to remove remain RNA, and 1st-strand cDNA corresponding to around 200-nt of single-stranded DNA ladder was purified on a 6% urea-PAGE gel. Purified cDNA was circularized using CircLigase II (EpiBio) and amplified by 12 cycles of PCR using Phusion polymerase (NEB). Amplified 2P-Seq library was purified on a 6% native-PAGE gel, and libraries were analyzed by 101-cycle single-end sequencing on the

Illumina HiSeq 2500 at the NICEM, with using custom sequencing primer (2P-seq-PE1.1).

Mapping of sequenced reads and measuring expression levels

To analyze strand-specific RNA-Seq dataset, analyses were performed using the Cufflinks package (Trapnell et al., 2012) and using the rice reference genome (IRGSP-1.0) and gene model annotation file (GFF file) from RAP-DB (http:// http://rapdb.dna.affrc.go.jp/). After aligning strand-specific RNA-Seq reads using TopHat and assembling transcript models using Cufflink, with providing gene model annotation file from RAP-DB, transcript abundance (FPKM, Fragments Per Kilobase of exon per Million reads) was estimated using Cufflinks.

To analyze small RNA-Seq dataset, small RNA-Seq dataset were analyzed using FASTX toolkit, to remove 3' adapter sequences from small RNA reads and low-quality reads. Refined reads were then compared with sequences of rRNA, tRNA, snRNA and snoRNA to remove those non-coding RNA-matched reads, then remaining reads were aligned to rice reference genome using MicroRazerS program. Expression level of mapped small RNA reads were calculated as RP30M (reads per 30 million reads).

To analyze 2P-Seq dataset, 2P-Seq reads ranging from 20 to 95-nt were selected after additional oligo(dT) and 'VN' sequences were sequentially removed. Those

reads were mapped onto rice reference genome (IRGSP-1.0) using TopHat. Estimation of gene expression level were performed using RAP-DB gene model and calculated in RPKHM (reads per kilobase of exon per hundred million reads).

Gene Ontology (GO) analysis with AgriGO

Gene Ontology analysis was performed using singular analysis enrichment (SEA, Fisher's test, $P < 0.05$, False Discovery Rate (FDR) < 0.05) on AgriGO website (http://bioinfo.cau.edu.cn/agriGO/). Since AgriGO support the form of both IRGSP gene locus ID and TIGR-MSU gene locus ID but GO term description of TIGR-MSU version is more substantial than IRGSP GO term description, TIGR-MSU gene locus ID corresponding to IRGSP gene locus ID was used in analysis.

5' RLM-RACE and 3' RACE

To identify the both 5' and 3' end of the putative lncRNA, transcript specific 5' RNA ligase-mediated rapid amplification of cDNA ends (5' RLM-RACE) and 3' RACE was performed, simultaneously, using GeneRacer Kit (Invitrogen). 5 ug of total RNA extracted from the rice roots or leaves was ligated to 0.25 ug of GeneRacer RNA oligo 5' adapter. 5' adapter-ligated total RNA was then reverse-transcribed using SuperScript III (Invitrogen) and oligo$(dT)_{24}$-containing RT

primer (60-mer) supplied in GeneRacer Kit. To amplify the 5'/3' end of the cDNA, touchdown PCR was performed using GeneRacer 5'/3' Primer and gene-specific touchdown PCR primer, and PCR product was further amplified by nested PCR using GeneRacer 5'/3' Nested Primer and another gene-specific nested PCR primer, with 2X PCR pre-mix (Solgent). Amplified PCR product was then subcloned into TA vector (RBC) and transformed to DH10B. Plasmid was extracted from transformed DH10B and then sequenced.

Strand-specific RT-PCR of *cis*-NAT$_{OsAMT1.1}$

Transcripts of *OsAMT1.1* and *cis*-NAT$_{OsAMT1.1}$ were reverse-transcribed with qPCR0026 and qPCR0010, respectively, using SuperScript III (Invitrogen) from 1 ug of total RNA extracted from ReN-0hr leaves. Synthesized 1st-strand cDNA of *OsAMT1.1* and *cis*-NAT$_{OsAMT1.1}$ were then respectively amplified using 3 pairs of primers (P1 (qPCR0009-qPCR00010), P2 (qPCR0027-qPCR0028), P3 (qPCR0005-qPCR0006) in Figure 11A), and amplicons generated by 30-cycle of PCR were resolved in 1.5% agarose gel.

RESULTS AND DISCUSSION

Profiling the changes of annotated gene expression level in the nitrogen-starved leaves and roots using strand-specific RNA-Seq

To analyze the effects of N starvation on the transcriptome of rice, strand-specific RNA-Seq libraries were constructed using roots and leaves of N-starved (Stv) and re-supplied (ReN) rice (Figure 1). By analyzing 24 libraries using Illumina HiSeq 2500, approximately 1 billion reads were obtained (Table 2), and reads were mapped onto the rice genome (International Rice Genome Sequencing Project (IRGSP)-1.0). Reference annotation-based transcript (RABT) assembly was performed using the Cufflink package to count expression level of IRGSP-annotated genes (Trapnell et al., 2012). Abundance of assembled transcripts was estimated by FPKM (fragment per kilobase of exon per million reads) using the Cufflink package.

To analyze the expression patterns of annotated genes respond to N starvation treatment, differential gene expression analysis was performed; expression levels of genes at four N-starved samples (Stv-1D, Stv-3D, Stv-5D, Stv-7D) were compared with that of Stv-0D (non-starvation-treated). Annotated genes expressed more than 3 FPKM in at least one of two compared samples, and genes up-regulated or down-regulated by more than 2-fold comparing to Stv-0D were used

in analysis.

In the leaves, 1,547 and 1,244 genes were up-regulated and down-regulated, respectively, in response to various N starvation conditions (Figure 2A, 2B, 2C). Of those, 200 and 240 of genes were respectively up-regulated and down-regulated in Stv-1D leaves, whereas 640 and 541 of genes were respectively up-regulated and down-regulated in Stv-3D leaves. Associated with the results that there were little overlaps of up-regulated and down-regulated genes between Stv-1D leaves and Stv-3D leaves (Figure 2B), and 849 genes were up-regulated and 651 genes were down-regulated in a Stv-3D leaves compared to Stv-1D leaves (Supplementary Figure 1A), these results suggest that dramatic changes occurred between Stv-1D leaves and Stv-3D leaves. In the up-regulated genes, several genes encoded glutamine amidotransferase, peptidase, pectinesterase, cellulase, β-1,3-glucanase, xyloglucan endotransglycosylase and chitinase. It is known that the leaves are main storage place of N during vegetative stage, and N sources stored in the form of proteins are remobilized through protein degradation pathway to seed and other organs during leaf senescence, mainly in the form of amino acids (Hirel et al., 2007; Gregersen et al., 2008; Kant et al., 2011). Peptidase genes are involved in protein degradation during leaf senescence to release amino acids (Gregersen et al., 2008), and glutamine amidotransferase removes ammonia group from glutamine, releasing ammonium ion. Consistent with the up-regulation of

those genes, several amino acid transporters and two ammonium transporters (*OsAMT1.1* and *OsAMT1.2*), which is related in translocation of amino acid and ammonium ion, were also up-regulated in Stv-3D leaves. The transcript level of ribulose-1,5-bisphosphate carboxylase/oxygenase (RuBisCO) is also down-regulated. Since RuBisCO is major protein where leaf N is stored (20~30% of total leaf N in C3 plant) and whose degradation is important to N remobilization (Xu et al., 2012), down-regulated transcript level of RuBisCO would reduce its protein level, indicating storage of N in leaf would be inhibited and N sources derived from degradation of RuBisCO by above-described pathway would be remobilized. These results suggest that dismantling of various cell components and remobilization of N sources occurred on Stv-3D leaves. Other hydrolytic enzymes mentioned above could promote these actions by depolymerizing and hydrolyzing cell wall components, resulting the modification and degradation of cell wall structure during leaf senescence (Lee et al., 2007; Breeze et al., 2011), which also could be involved in leaf senescence process for remobilization of N sources.

Discoloration of leaf and reduced photosynthesis, meaning destabilization of chloroplasts in leaf, are major phenotypes caused by N starvation (Gregersen et al., 2008). Consistent with these N starvation phenotypes, Gene Ontology (GO) analysis performed against subset of down-regulated genes in Stv-3D leaves revealed a significant overrepresentation of genes related in thylakoid and plastid

components, associated with down-regulation of photosynthesis-related genes such as PsbP(Photosynthesis II complex assembly/stability-related), Photosystem II, thioredoxin and ferredoxin (Figure 3A).

Interestingly, 562 (47.8%) of 1,181 differentially expressed genes in Stv-3D leaves, including some of above-mentioned genes involved in N remobilization and photosynthesis, were also differentially expressed after 5-day and/or 7-day nitrogen starvation (Figure 3B). GO analysis of up-regulated and down-regulated genes in Stv-5D leaves and/or Stv-7D leaves showed that transport-related genes and plastid-related genes (including Photosystem II, chlorophyll a/b-binding protein and ATP synthase) were significantly overrepresented (Supplementary Figure 1B, 1C, 1D). These results suggest that the responses starting from 3-day after nitrogen starvation were continued and even were amplified in proportion to increases of N starvation period.

In the roots, 2,838 and 1,670 genes were up-regulated and down-regulated, respectively (Figure 2A, 2D, 2E). Relative to expression level changes in Stv-1D leaves, 713 and 556 genes were up- and down-regulated, respectively, which means the gene set in the roots responded to N starvation earlier than leaves, which is not so surprising because of the importance of root in sensing and uptake of external N. In the Stv-1D roots, several genes encoding Cytochrome P450 and peroxidase were induced (Figure 3C, 3D). It is reported that reactive oxygen

species (ROS), which play important roles in signal transduction pathways, have been increased in response to N starvation in Arabidopsis roots, and along with increase of ROS, peroxidase and *AtNRT2.1* has also been induced (Shin et al., 2005). Moreover, OH production mediated by peroxidase could induce cell wall loosening, which is necessary for cell extension, has been reported in maize roots (Liszkay et al, 2004). These previous reports and induction of these genes in 1-day nitrogen-starved roots could be coupled with the elongation of primary root, which is one of the known major phenotype of N-starved roots.

Analysis of expression patterns of nitrogen metabolism-involved genes in roots

The intake and assimilation of external N sources primarily occurs in the roots, and involved genes are known to dynamically respond to N supply state (Sonoda et al., 2003a; Sonoda et al., 2003b; Ishiyama et al., 2004; Araki and Hasegawa, 2006; Feng et al., 2011; Li et al., 2012). To understand the response of N metabolic pathway-involved genes in the N-starved roots, expression profiling of these genes was performed.

Most of genes involved in N uptake and assimilation were highly expressed in the roots of Stv-0D than in the leaves, except for *OsGS2*, *Fd-GOGAT* and *NADH-GOGAT2* (Figure 4B). Moreover, expression level of these genes in the roots changed more dramatic than that in the leaves (Figure 4B). Because of these points,

expressional analysis was mainly focused on changes in the roots.

Of the 7 nitrate transporter genes in rice, *OsNRT2.1*, *OsNRT2.2*, *OsNRT2.3a*, *OsNRT2.3b*(splicing variant of *OsNRT2.3a*) and *OsNAR2.2* were induced in response to N starvation, whereas *NRT/CHL* (*AtNRT1.1* homolog) and CLCa genes were down-regulated (Figure 4C). Two nitrate transporters, *OsNRT2.1* and *OsNRT2.2*, were up-regulated until 5 days after N starvation. Unexpectedly, however, expression level of these genes was recovered in 7-day N-starved roots, whereas isoforms of *OsNRT2.3* were gradually up-regulated until 7 days after nitrogen starvation. *AtCLCa*, nitrate/proton antiporter specifically localized in vacuole membrane, is known to accumulate nitrate on vacuole in Arabidopsis (Angeli et al., 2006). Downregulation of rice *CLCa* homolog would release the vacuole-accumulated nitrate to the cytoplasm, which could be one of the resistant responses upon nitrogen starvation. Interestingly, rice *CLCa* homolog was down-regulated both in the roots and the leaves, indicating that *CLCa* was regulated both in the roots and the leaves.

Among the 9 ammonium-transporter genes, *OsAMT1.2*, *OsAMT1.3*, *OsAMT2.2*, *OsAMT3.1* and *OsAMT3.3* were differentially expressed in at least one of four N starvation time points (Figure 4C). *OsAMT1.2* (Os05g0620500) and *OsAMT1.3* (Os05g0620600) showed reciprocal expression patterns; *OsAMT1.2* was induced whereas *OsAMT1.3* was reduced in the N-starved roots. Previous study revealed

that *OsAMT1.1* and *OsAMT1.3* are positively regulated by internal glutamine concentration, whereas *OsAMT1.2* is negatively regulated (Sonoda et al., 2003b) and induced by N starvation (Sonoda et al., 2003a; Li et al., 2012). These results and previous reports indicate that internal glutamine concentration might be reduced. Similar with the reciprocal responses of *OsAMT1.2-OsAMT1.3*, *OsGS1.1-OsGS1.2* (glutamine synthetase) also showed reciprocal responses. *OsGS1.1*, which exhibits lower *Km* and higher *Vmax* value on ammonium ion than *OsGS1.2* (Ishyama et al., 2004), was gradually induced ranging from 1-day nitrogen starvation to 7-day nitrogen starvation. Whereas, less-efficient *OsGS1.2* was dramatically down-regulated in 1-day nitrogen-starved roots and increased again in 3-day and 5-day nitrogen-starved roots (Figure 4C). Expressional inversion from less-efficient form to more-efficient form in the roots suggests that the rate and efficiency of ammonium ion assimilation increased in the nitrogen-starved roots to capture available ammonium ion as much as possible. Although uptake efficiencies of *OsAMT1* family members remains to be elucidated, increase of the *OsAMT1.3* and decrease of the *OsAMT1.2* would improve ammonium uptake efficiency when the available ammonium ion does not exist or exist in a low level. Overall, these results suggest that N uptake and assimilation system in roots were modulated in a transcript level, in response to N starvation.

Expression profiling of miRNAs and its target genes in nitrogen-starved and re-supplied rice roots

To understand the responses of miRNAs on N starvation and N re-supply, small RNA-Seq libraries were constructed using rice roots of 13 N-starved and re-supplied samples, and constructed libraries were analyzed using Illumina HiSeq 2500. Raw sequenced reads ranging from 11,206,654 to 49,709,205 were obtained from each library, and 168,054,637 of total reads were mapped onto the International Rice Genome Sequencing Project-1.0 (IRGSP-1.0) after trimming 3' adapter sequence and removing small RNA reads matching structural non-coding RNAs (rRNA, tRNA, snRNA) (Table 3). The abundance of small RNA reads of each library was estimated by RP30M (reads per 30 million reads).

Distribution of mapped small RNA reads obtained from roots revealed that most of the reads had a length of 21 and 24 nucleotides. Major peaks were formed in 21-U and 24-A, which represent miRNA and siRNA pool, respectively (Figure 5). Previous reports have showed that 20 nucleotide-length small RNA reads were significant in rice roots (Jeong et al., 2011; Secco et al., 2013). However, significant expression of 20 nucleotide-length small RNA reads was not observed in our dataset.

To analyze the response of miRNA pool to N supply status, differential expression analysis between N-starved roots and N-supplied (or re-supplied) roots

was performed. 4 timepoints (Stv-0D, Stv-7D, ReN-0h, ReN-72h) described in Figure 1A were chosen for analysis and the expression level of miRNAs were compared except for the comparison between Stv-7D and ReN-72h.

Three combination of differential expression analysis revealed that, of 147 distinct miRNA sequences expressed more than 150 RP30M at least 1 of 4 libraries, 82 distinct miRNAs were differentially expressed in at least 1 of 3 combination of differential expression analysis (Figure 6A). As expected, miRNAs down-regulated in response to N starvation treatment were significantly overlapped with up-regulated miRNAs in ReN-72h (N re-supplied condition) compared with ReN-0h (N-starved condition) (Figure 6B). Consistent with above-mentioned results, most of up-regulated miRNAs in Stv-7D and/or ReN-0h roots were also overlapped with down-regulated miRNAs in ReN-72h roots (Figure 6C). These results indicate that the great portion of differentially expressed miRNAs was directly regulated by N supply status.

A number of miRNAs were mainly down-regulated upon N starvation condition (Stv-7D and ReN-0hr). Of those, most of miR169 family members were significantly down-regulated in the roots of Stv-7D and ReN-0hr, whereas up-regulated in ReN-72hr roots (Figure 6B, 7A). miR169, one of the conserved miRNAs found in diverse plant species, has been reported as representative N-responsive miRNAs in Arabidopsis (Zhao et al., 2011; Liang et al., 2012) and

maize (Zhao et al., 2012), which has not been reported in rice. Expression profiling against miR169 family members showed that the expression level was changed in proportion to treatment period of N starvation and re-supply, except for the change between ReN-0hr and ReN-6hr. Down-regulation of miR169 expression level on ReN-6h was unexpected. However, considering that the expression levels of miR169 family members increased gradually after 6 hr re-supply of N, this could be the result of extreme early time response, which was not expected and had not been previously reported.

It is known that miR169 mainly regulates the expression level of Nuclear transcription Factor Y subunit A (NF-YA), the conserved target gene of miR169 in plant species including rice, by complementary pairing with 3' UTR binding site of NF-YA transcript. In rice, 10 annotated and 1 putative NF-YA genes exist, and 8 of those 11 NF-YA genes were predicted as target of miR169 in miRNA target prediction result (Figure 7B). To identify whether miR169 cleaves NF-YA genes, 5' RLM-RACE was performed against *NF-YA5*, the most up-regulated NF-YA gene in Stv-7D (6.9-fold up-regulated). 5' end sequence tags were significantly detected between 10 and 11-nt distant from 5' end of predicted miRNA sequence, indicating that miR169 regulates *NF-YA5* by cleaving *NF-YA5* transcript (Figure 7D). Consistent with this result, expression level of *NF-YA5* was dramatically and gradually induced in N-starved roots, which was reciprocal to expression patterns

of miR169 family members (Figure 7C). Moreover, this reciprocal expression pattern between NF-YA gene and miR169 was observed in most of other NF-YA genes targeted by miR169 (Figure 7C), suggesting that the changes on the level of miR169 upon N starvation is involved in increase of NF-YA genes.

It is interesting that the expression of miR169 is regulated not only by N supply status but also by various abiotic stresses such as high salinity and drought (Li et al., 2008; Zhao et al., 2009; Zhang et al., 2011), and the regulation of NF-YA genes by miR169 is closely involved in the resistance or sensitivity on those abiotic stresses. These connections between miR169, NF-YA genes and multiple abiotic stresses suggest the importance of miR169 - NF-YA regulatory pathway on the abiotic stress responses of plants', including N-starvation.

Identification of putative lncRNA candidates in rice using strand-specific RNA-seq

To identify putative lncRNA deriving loci, reference annotation-based transcript (RABT) assembly was performed using the Cufflink package. To assemble accurate transcript models and identify N starvation stress-regulated transcripts, transcript assembly was performed against RNA-Seq data from leaves and roots, respectively.

69,848 and 70,322 of transcript models were assembled from leaves and roots,

respectively. Of those, 4,161 and 4,635 of novel transcript models that did not overlap RAP-DB annotated gene loci were identified from leaves and roots, respectively. To refine novel, putative lncRNA loci, assembled transcript models were filtered using following criteria: (1) remove transcripts overlapping MSU(v7) gene loci, another rice gene annotation differ from RAP-DB; (2) remove transcripts overlapping known miRNAs and structural ncRNAs (tRNA, rRNA, snRNA, snoRNA, and other structural RNAs); (3) remove transcripts that Coding Potential Calculator (CPC) score is > -1.0; (4) remove transcripts shorter than 200-nt (Figure 8A, Table 5). 5,261 of transcript isoforms passing through (1) and (2) were obtained from the roots and leaves. Those isoforms derived from the same locus were then integrated (Table 5, Step 5). After integration, CPC score of isoforms were calculated (Kong et al., 2007). Considering that CPC compares predicted open reading frame (ORF) of query sequence with annotated ORF sequences in protein database (Swiss-Prot database in this study), and most of rice-only annotated ORF sequences were just computationally predicted and had not been confirmed by experiment, transcript isoforms representing CPC score ≥ 0 in 1st-round CPC analysis were further analyzed using rice ORF sequences-excluded Swiss-Prot database. After 2nd-round CPC analysis, loci not containing isoforms having CPC score ≤ -1.0 were filtered out, and filtering out of short transcript-deriving loci (< 200-nt) was followed. In total, 1,840 putative lncRNA loci

deriving 2,869 transcript isoforms were finally obtained (Table 5, Step 8). Of those filtered putative lncRNA loci, 1,458 loci were located on intergenic region (unannotated region), and 382 loci were located on the antisense strand of RAP-DB annotated gene loci (Figure 9B). Sequence composition analysis showed that the average A/U content of putative lncRNAs was placed between that of CDS/5'UTR and that of 3'UTR, and was larger than those of introns of protein-coding genes and lncRNAs, which were previously reported in other lncRNA studies (Ulitsky et al, 2011; Nam et al, 2012). Of those 1840 loci, 1,173 loci contained repeat sequence including simple repeat of nucleotide sequences, and 844 (45.8%) of those were containing transposable element sequence.

Expression profiling of putative lncRNA candidates

To analyze expression patterns of putative lncRNAs, expression level was calculated by FPKM (fragment per kilobase of exon per million reads) using the Cufflink package. Expression analysis was performed using expression data obtained from roots and leaves of Stv-0D, Stv-1D, Stv-3D, Stv-5D, Stv-7D, and wild type. Most of putative lncRNA candidates were expressed in a low level relative to protein-coding gene group (Figure 8D). 38% (703) and 62% (1,141) of putative lncRNA-deriving loci were expressed more than 1 and 0.5 FPKM, respectively, in at least 1 of 12 libraries.

Organ-preferential expression of putative lncRNAs (2-fold cutoff) was observed both in Stv-0D and four N starvation-treated samples. Of those loci expressed more than 0.5 FPKM, 309 and 252 loci were preferentially expressed, respectively in leaves and roots of Stv-0D, and 252 and 250 loci were preferentially expressed respectively in leaves and roots of N-starved set ((average FPKM of four N-starved samples) ≥ 0.5 FPKM).

To analyze the expression patterns of putative lncRNAs in N-starved condition, FPKM-estimated expression level of four N-starved time points (Stv-1D, Stv-3D, Stv-5D, Stv-7D) were compared with that of Stv-0D. Considering low expression level of putative lncRNA candidates relative to protein-coding genes, putative lncRNA loci expressed ≥ 0.5 FPKM in at least one of two compared samples were used in expression analysis. In total, 838 (including 694 intergenic loci and 144 antisense loci) of putative lncRNA loci were differentially expressed (≥ 2-fold-change) in leaves and roots. Of those, 553 intergenic and 119 antisense loci were differentially expressed in organ-preferential pattern (Supplementary Figure 2).

Analysis of 2P-Seq results

To identify 3' end of transcripts and find alternative polyadenylation pattern upon changes of N supply status in a genome-wide level, three 2P-Seq libraries were constructed using pooled total RNA from 22 of roots and leaves (Total-Pool)

of N-starved/resupplied rice, and total RNA of wild type roots and leaves (WT (Lf+Rt)), respectively. In total, 161,280,481 reads were obtained by analyzing libraries with Illumina HiSeq 2500 (Table 4). Of those, reads between 20-nt and 95-nt were aligned to genome, and 123,264,021 (49,924,828 of Total-Pool and 73,339,193 of WT(Lf+Rt)) reads were mapped to rice genome (IRGSP-1.0) (Figure 9A, Table 4). Most of genome-mapped reads were aligned on genic region (Figure 9B), and reads ranging from 6.86% to 7.39% were aligned on intergenic region. Those reads aligned to genic region were significantly detected in the 3' UTR region (31,723 of 44,540 RAP-DB transcript models), followed by intronic and exonic region (Figure 9C).

To confirm the correlation of transcript abundance estimated using strand-specific RNA-Seq and 2P-Seq, those reads mapped to RAP-DB gene models were calculated by RPKHM (reads per kilobase of exon per hundred million reads). Comparison between average RPKHM of 24 libraries and Total-Pool 2P-Seq library showed that they were correlated (r^2 = 0.82, Figure 9D). These results suggest that 2P-Seq results could partly represent transcript expression level.

Identification of putative long non-coding RNA's 3' end using 2P-Seq and validation of 2P-Seq signal-detected sites using 3' RACE

7.39% and 6.86% of 2P-Seq reads derived from Total-Pool and WT (Lf+Rt)

were mapped onto intergenic region (Figure 9B), representing possibilities of 2P-Seq data application on defining 3' end of poly-(A)-tailed putative long non-coding RNA. Of those selected 1,840 putative lncRNA loci, 2P-Seq signals were detected in 55% (1,020) loci (Figure 10A). Of those 2P-Seq signal-detected loci, N starvation-responsive Chr04G0191 (Chr4:27219529-27220526) were selected. Chr04G0191 was expressed specifically in roots, and was down-regulated in response to N starvation (Figure 10B). 2P-Seq signal of Chr04G0191 were detected only in Total-Pool 2P-Seq library, and two distinct clusters of 2P-Seq signal were detected (2P-1 and 2P-2, Figure 10C). To confirm whether those two putative 3' ends could be detected by other method, 3' RACE was performed against Chr04G0191 (Figure 10C). Since 2P-Seq use oligo(dT)-priming on reverse transcription step like 3' RACE, 2P-Seq signals would be also detected by 3' RACE. As expected, 1 and 5 of 3' RACE tags were confirmed where 2P-1 and 2P-2 sites were detected (Figure 10C, 10D). These results indicate that 2P-Seq dataset could be applicable method in defining 3' ends of poly- (A)-tailed putative lncRNAs in a genome-wide level.

Identification of *cis*-NAT$_{OsAMT1.1}$ using strand-specific RNA-Seq and 2P-Seq

The ability to detect and measure expression level of two independent transcripts derived from different strands on the same genomic loci is one of

powerful advantages of strand-specific RNA-Seq. Also, 2P-Seq sequencing was performed with 101-cycle single-end analysis, sequenced reads of 2P-Seq also has orientation information of transcript. Relying on these strand-specificity of two RNA-Seq methods, I found that *cis*-NAT$_{OsAMT1.1}$ (Os04g0509500) was expressed on the opposite strand of *OsAMT1.1* (Os04g0509600) (Figure 11A, 11C).

The expression level of *cis*-NAT$_{OsAMT1.1}$ was low relative to that of *OsAMT1.1*, but interestingly, the expression pattern of *cis*-NAT$_{\mathrm{OsAMT1.1}}$ in leaves was concordant with that of *OsAMT1.1*, and this concordance was observed exclusively in leaves (Figure 11B). It is possible that the generation of *cis*-NAT$_{OsAMT1.1}$ and the expressional concordance between *OsAMT1.1* and *cis*-NAT$_{OsAMT1.1}$ are due to RNA-dependent RNA transcription by RNA-dependent RNA Polymerases (RdRPs), which leads to production of siRNA by *DCL3b* and *DCL4* and down-regulation of its template RNA by inducing cleavage on template RNA (Song et al., 2012a; Song et al., 2012b). However, transcripts derived by the actions of RdRPs do not have poly-(A)-tail structure in general. Since 2P-Seq signals were detected exactly on the 3' end region of *cis*-NAT$_{\mathrm{OsAMT1.1}}$ both in WT (Lf+Rt) and Total-Pool (red box on 2P-Seq signal in Figure 11A), it looks like *cis*-NAT$_{OsAMT1.1}$ was transcribed by one of DNA-dependent RNA polymerase, possibly by RNA polymerase II.

Since RNA-Seq read-mapped region on minus strand (colored by red in Figure

11A) was broader than the genomic location of annotated *cis*-NAT$_{OsAMT1.1}$ (Os04g0509500), 5' RLM-RACE was performed to identify 5' transcription start site precisely. Also, to confirm the expression of *cis*-NAT$_{OsAMT1.1}$ on minus strand, strand-specific RT-PCR was performed using gene-specific primers. 5' end of *cis*-NAT$_{OsAMT1.1}$ confirmed by 5' RLM-RACE located ~290-bp downstream of Os04g0509500, which matched with 5' end of strand-specific RNA-Seq read-mapped region (Figure 11D). In addition, strand-specific RT-PCR results also showed that strand-specific RNA-Seq result was valid. These results suggest that strand-specific RNA-Seq and 2P-Seq could contribute to predict precise gene models of *cis*-NATs.

Since the sequence of *OsAMT1.1* is well-conserved in diverse plant species, it is natural that the antisense strand of *OsAMT1.1* is also conserved. To check the possibility whether *cis*-NAT$_{AMT1.1}$ is annotated on or expressed in other plant species, genomic synteny between rice and other plants was investigated by genomic region comparison serviced by EnsemblPlants (http://plants.ensembl.org/). Comparing *OsAMT1.1* – *cis*-NAT$_{OsAMT1.1}$ region with 8 *Oryza* species and 7 other plants, I found that *cis*-NAT of AMT1.1 homologs were annotated in Arabidopsis (At4g13510 (*AtAMT1.1*)–At4g13505 (*cis*-NAT)), *Oryza brachyantha* (OB04g25450 – OB04G25460 (*cis*-NAT)) and *Zea mays* (GRMZM2G118950 – GRMZM2G420059 (*cis*-NAT) and GRMZM2G175140 –

GRMZM2G474905 (*cis*-NAT)) (Figure 12). Although evidences supporting their expression were only found in case of GRMZM2G420059 (supported by expressed sequence tag), possibilities of their expression in N starvation condition have not been examined before, and still exist.

SUMMARY

By applying multimodal RNA-Seq methods and analyzing data derived from RNA-Seq, I could obtain large-scale transcriptomic dataset about N-starved and resupplied rice. Using strand-specific RNA-Seq, expression profiling and analysis of annotated genes including nitrogen metabolism-related genes were performed. Numbers of annotated genes were up-regulated and/or down-regulated in response to nitrogen starvation stress. Of those, genes involved in transport and assimilation of nitrogen were dynamically changed in roots rather than leaves, in response to nitrogen starvation stresses. Also, reciprocal expression changes between *OsGS1;1* and *OsGS1;2* and between *OsAMT1;2* and *OsAMT1;3* were observed in Stv-1D roots.

By analyzing data from small RNA-Seq, 82 nitrogen-responsive miRNAs were identified in nitrogen-starved rice roots. Also, relationship between one of N supply status-responsive miRNA, miR169, and its target genes, NF-YA, were analyzed by comparison of expression pattern and 5' RLM-RACE.

Using strand-specific RNA-Seq dataset, genome-wide identification of novel, putative lncRNA-deriving loci were performed. In total, 1,458 intergenic and 382 antisense putative lncRNA-deriving loci were identified. I analyzed their expression patterns in roots and leaves of nitrogen-starved rice, and observed their N-responsive expression patterns and organ-preferential patterns.

Moreover, identyfing 3’ end of putative long non-coding RNA was tried by analyzing 2P-Seq dataset. To validate signals detected by 2P-Seq, one N-responsive putative intergenic lncRNA was chosen and 2P-Seq signals were validated by 3’ RACE against putative lncRNA candidate. Together, analysis results of dataset derived by multimodal RNA-Seq methods gave a lot of information about the rice transcriptome pool in response to change of nitrogen supply status.

FIGURES AND LEGENDS

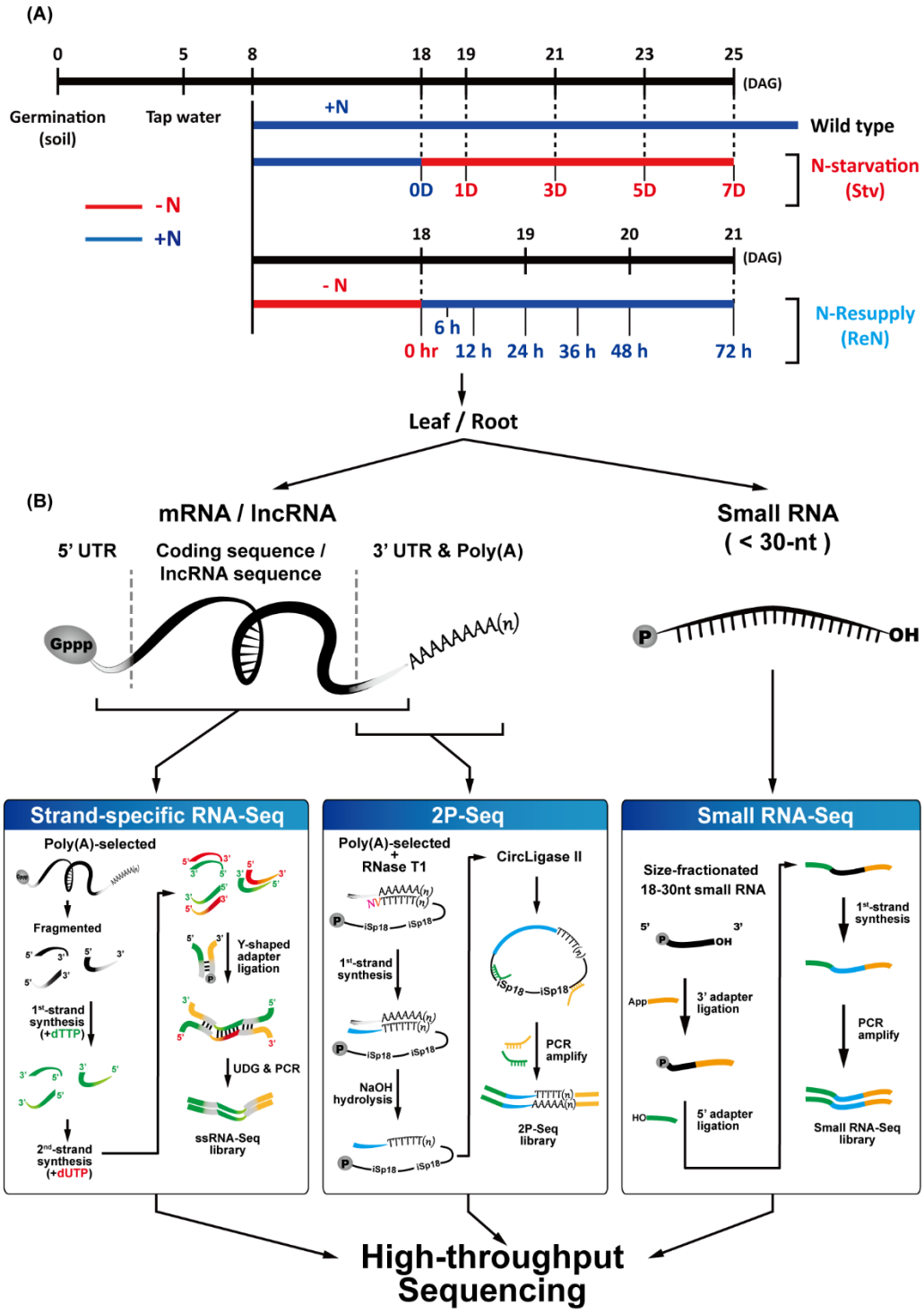

(A)
0
5
8
18
19
21
23
25
(DAG)
Germination (soil)
Tap water
+N
Wild type
N-starvation (Stv)
0D
1D
3D
5D
7D
- N
+N
18
19
20
21
(DAG)
- N
N-Resupply (ReN)
6 h
0 hr
12 h
24 h
36 h
48 h
72 h
Leaf / Root
(B)
mRNA / lncRNA
Small RNA (< 30-nt)
5' UTR
Coding sequence / lncRNA sequence
3' UTR & Poly(A)
Gppp
AAAAAAAA(n)
P
OH
Strand-specific RNA-Seq
Poly(A)-selected
Fragmented
1st-strand synthesis (+dTTP)
2nd-strand synthesis (+dUTP)
Y-shaped adapter ligation
UDG & PCR
ssRNA-Seq library
2P-Seq
Poly(A)-selected + RNase T1
1st-strand synthesis
NaOH hydrolysis
CircLigase II
PCR amplify
2P-Seq library
Small RNA-Seq
Size-fractionated 18-30nt small RNA
3' adapter ligation
5' adapter ligation
1st-strand synthesis
PCR amplify
Small RNA-Seq library
High-throughput Sequencing

Figure 1. Experimental scheme and multimodal transcriptomic analysis used in this study. (A) Nitrogen starvation and re-supply treatment scheme. (B) Multimodal transcriptomic methods used in this study.

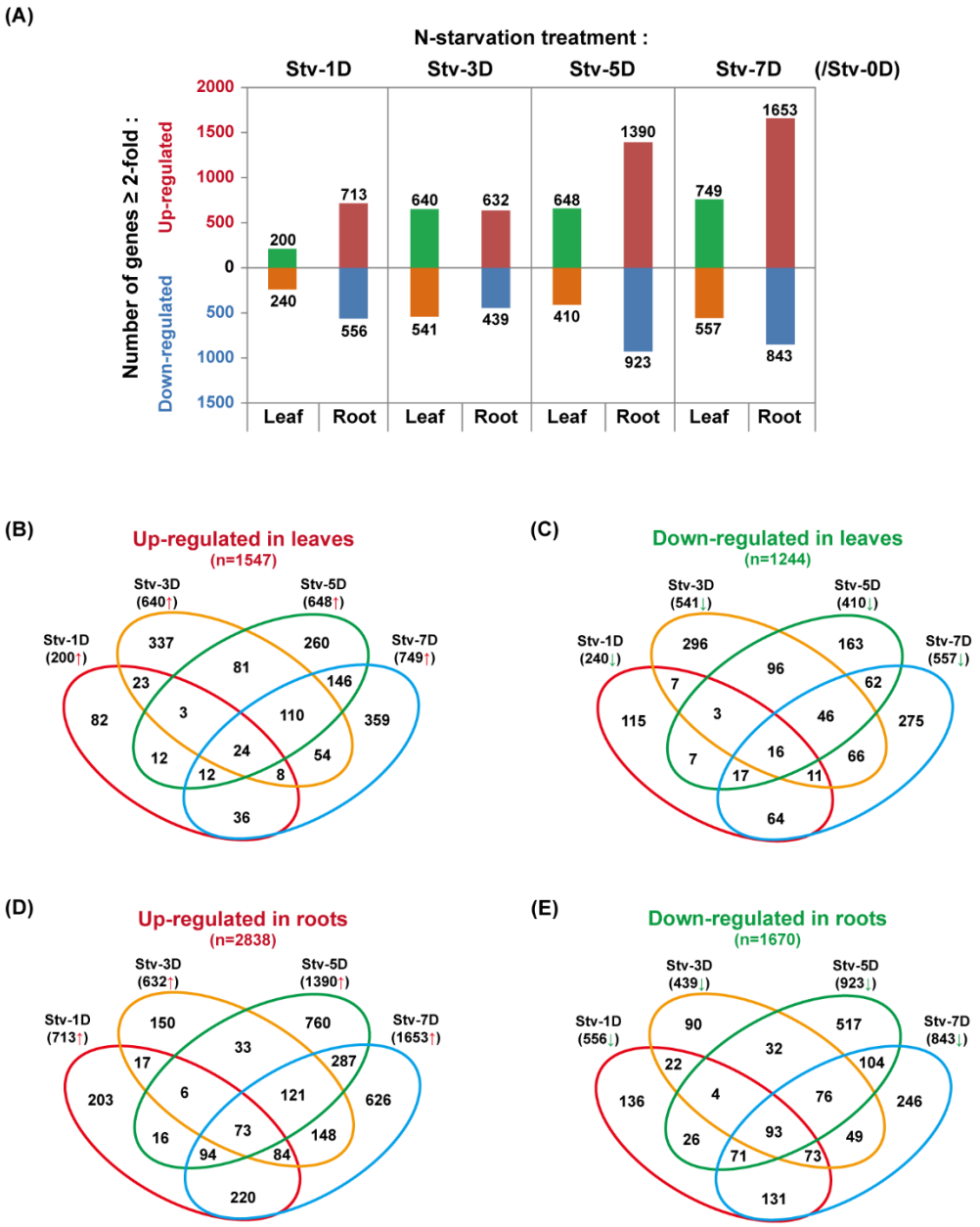

(A)
N-starvation treatment :
Stv-1D
Stv-3D
Stv-5D
Stv-7D
(/Stv-0D)
Number of genes ≥ 2-fold :
Up-regulated
Down-regulated
2000
1500
1000
500
0
500
1000
1500
200
240
713
556
640
541
632
439
648
410
1390
923
749
557
1653
843
Leaf
Root
Leaf
Root
Leaf
Root
Leaf
Root
(B)
Up-regulated in leaves
(n=1547)
Stv-3D (640↑)
Stv-5D (648↑)
Stv-1D (200↑)
Stv-7D (749↑)
337
260
81
23
146
82
3
110
359
12
24
54
12
8
36
(C)
Down-regulated in leaves
(n=1244)
Stv-3D (541↓)
Stv-5D (410↓)
Stv-1D (240↓)
Stv-7D (557↓)
296
163
96
7
62
115
3
46
275
7
16
66
17
11
64
(D)
Up-regulated in roots
(n=2838)
Stv-3D (632↑)
Stv-5D (1390↑)
Stv-1D (713↑)
Stv-7D (1653↑)
150
760
33
17
287
203
6
121
626
16
73
148
94
84
220
(E)
Down-regulated in roots
(n=1670)
Stv-3D (439↓)
Stv-5D (923↓)
Stv-1D (556↓)
Stv-7D (843↓)
90
517
32
22
104
136
4
76
246
26
93
49
71
73
131

Figure 2. Profiling of differentially expressed genes in nitrogen-starved leaves and roots. (A) Number of ≥ 2-fold up-regulated and down-regulated genes in roots and leaves. Red / green bars indicate up-regulated in roots and leaves, blue / orange bars indicate down-regulated genes in roots and leaves. (B)(C)(D)(E) Overlaps between (B)(D) up-regulated genes, and overlaps between (C)(E) down-regulated genes.

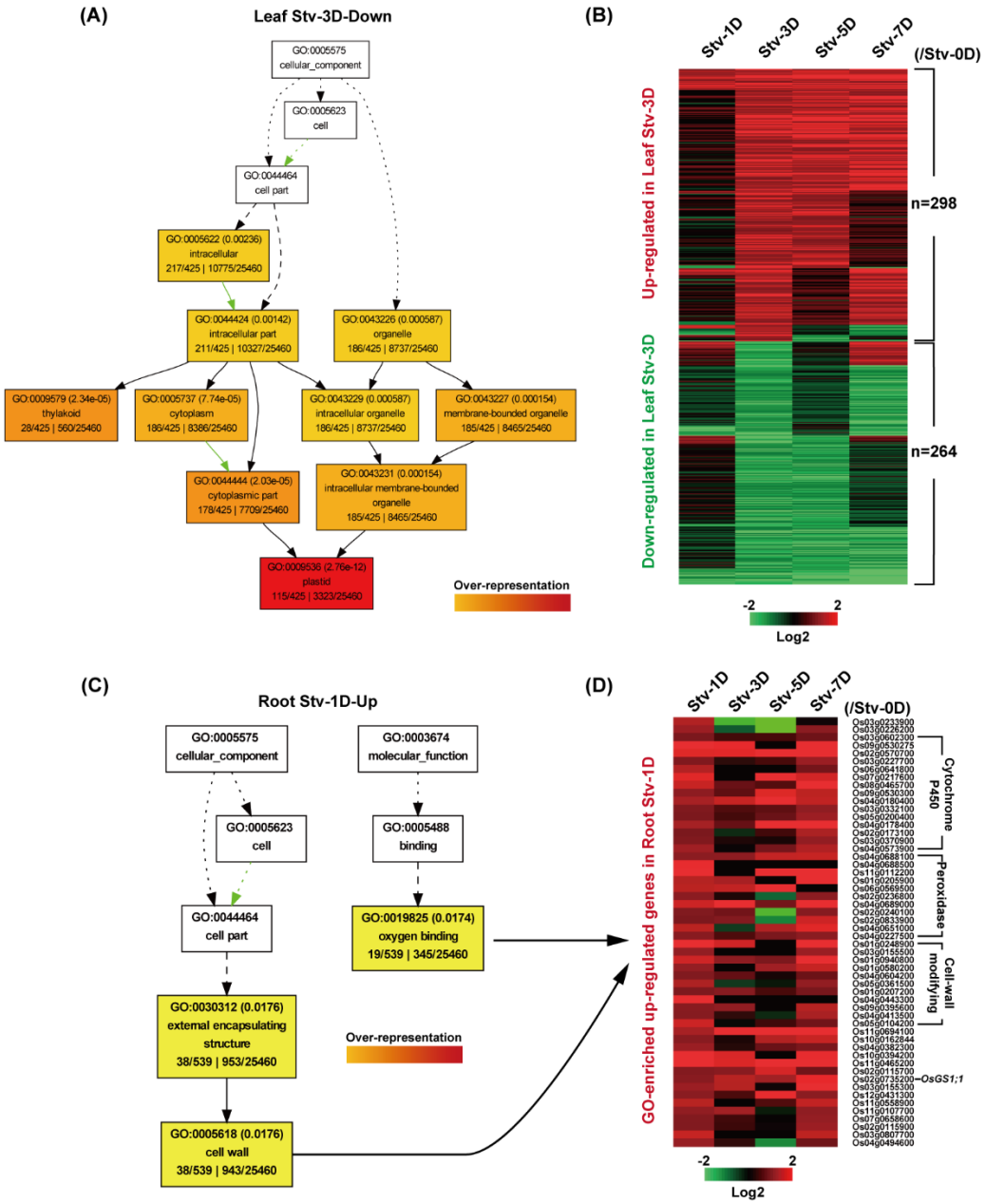

(A)
Leaf Stv-3D-Down
GO:0005575
cellular_component
GO:0005623
cell
GO:0044464
cell part
GO:0005622 (0.00236)
intracellular
217/425 | 10775/25460
GO:0044424 (0.00142)
intracellular part
211/425 | 10327/25460
GO:0043226 (0.000587)
organelle
186/425 | 8737/25460
GO:0009579 (2.34e-05)
thylakoid
28/425 | 560/25460
GO:0005737 (7.74e-05)
cytoplasm
186/425 | 8386/25460
GO:0043229 (0.000587)
intracellular organelle
186/425 | 8737/25460
GO:0043227 (0.000154)
membrane-bounded organelle
185/425 | 8465/25460
GO:0044444 (2.03e-05)
cytoplasmic part
178/425 | 7709/25460
GO:0043231 (0.000154)
intracellular membrane-bounded
organelle
185/425 | 8465/25460
GO:0009536 (2.76e-12)
plastid
115/425 | 3323/25460
Over-representation
(B)
Stv-1D
Stv-3D
Stv-5D
Stv-7D
(/Stv-0D)
Up-regulated in Leaf Stv-3D
Down-regulated in Leaf Stv-3D
n=298
n=264
-2
2
Log2
(C)
Root Stv-1D-Up
GO:0005575
cellular_component
GO:0003674
molecular_function
GO:0005623
cell
GO:0005488
binding
GO:0044464
cell part
GO:0019825 (0.0174)
oxygen binding
19/539 | 345/25460
GO:0030312 (0.0176)
external encapsulating
structure
38/539 | 953/25460
Over-representation
GO:0005618 (0.0176)
cell wall
38/539 | 943/25460
(D)
Stv-1D
Stv-3D
Stv-5D
Stv-7D
(/Stv-0D)
GO-enriched up-regulated genes in Root Stv-1D
Cytochrome P450
Peroxidase
Cell-wall modifying
OsGS1;1
-2
2
Log2

Figure 3. Analysis of differentially expressed genes in leaves and roots. (A) AgriGO representation of the overrepresented GO terms in the down-regulated genes in Stv-3D leaves was generated using singular analysis enrichment (Fisher's test, $P < 0.05$, FDR < 0.05). Numbers in parenthesis represents the FDR value. (B) Heatmap representation of gene expression patterns of differentially expressed genes in Stv-3D. (C) AgriGO representation of the overrepresented GO terms in the up-regulated genes in Stv-1D roots. (D) Heatmap representation of genes' expression patterns overrepresentated in (C).

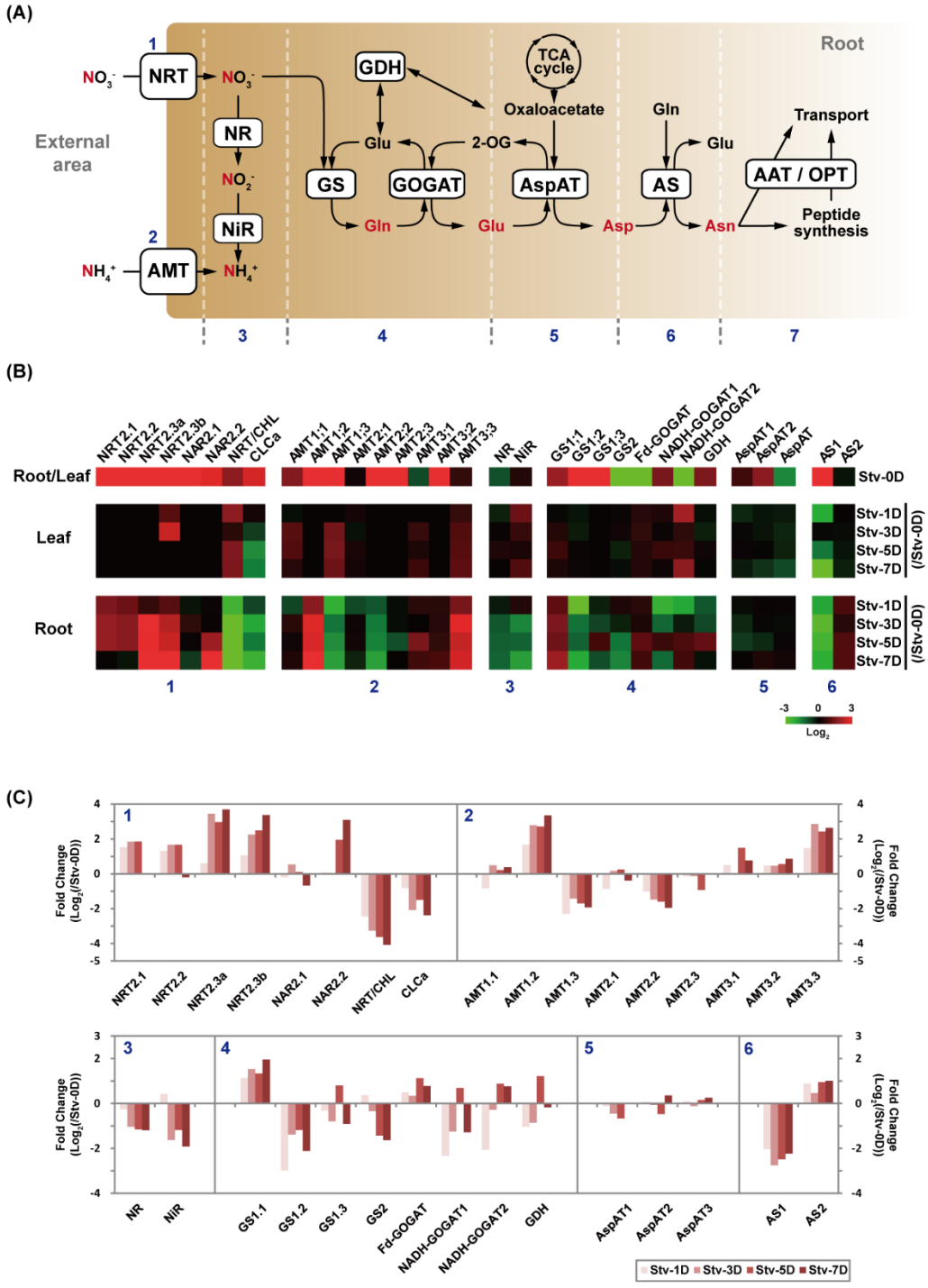
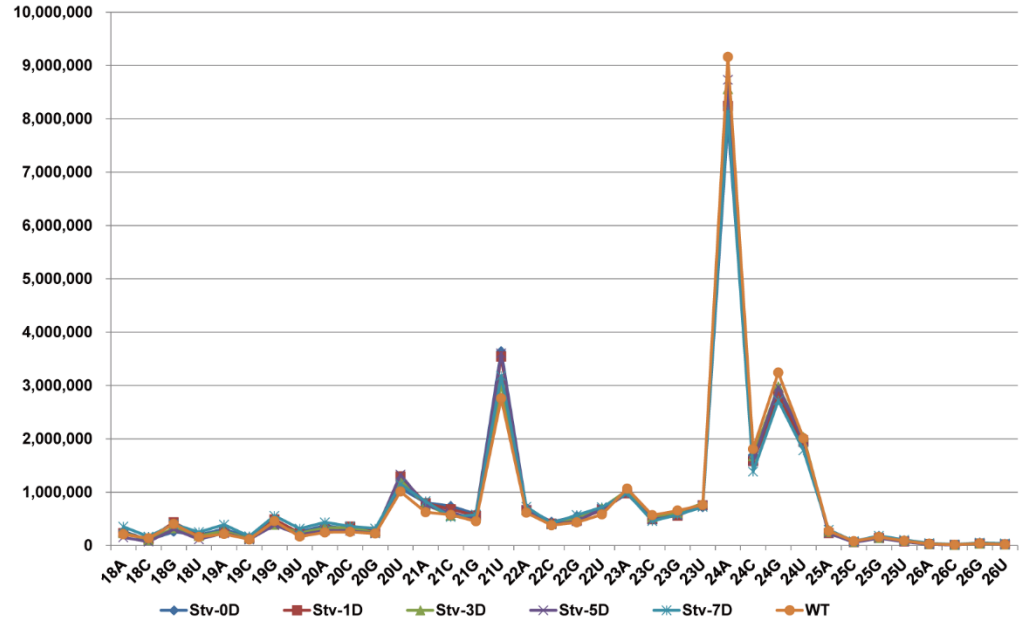

(A)
Root
External area
NO3-
NRT
NO3-
NR
NO2-
NiR
NH4+
AMT
NH4+
GDH
TCA cycle
Oxaloacetate
Glu
2-OG
GS
GOGAT
AspAT
Gln
Glu
Asp
Gln
AS
Glu
Asn
AAT / OPT
Transport
Peptide synthesis
(B)
NRT2.1
NRT2.2
NRT2.3a
NRT2.3b
NAR2.1
NAR2.2
NRT/CHL
CLCa
AMT1.1
AMT1.2
AMT1.3
AMT2.1
AMT2.2
AMT2.3
AMT3.1
AMT3.2
AMT3.3
NR
NiR
GS1.1
GS1.2
GS1.3
GS2
Fd-GOGAT
NADH-GOGAT1
NADH-GOGAT2
GDH
AspAT1
AspAT2
AspAT
AS1
AS2
Root/Leaf
Leaf
Root
Stv-0D
Stv-1D
Stv-3D
Stv-5D
Stv-7D
(/Stv-0D)
Log2
(C)
Fold Change (Log2(/Stv-0D))
Stv-1D
Stv-3D
Stv-5D
Stv-7D

Figure 4. Profiling of nitrogen metabolism-involved genes in leaves and roots. (A) Transportation and assimilation pathwas of nitrogen in roots. (B) Organ-preferentail expression and response to nitrogen starvation of nitrogen metabolism-invloved genes. Blue numbers indicate genes involved in each step of nitrogen assimilation pathway in (A). (C) Expressional changes of genes in nitrogen-starved roots relative to Stv-0D.

(A)


10,000,000
9,000,000
8,000,000
7,000,000
6,000,000
5,000,000
4,000,000
3,000,000
2,000,000
1,000,000
0
18A 18C 18G 18U 19A 19C 19G 19U 20A 20C 20G 20U 21A 21C 21G 21U 22A 22C 22G 22U 23A 23C 23G 23U 24A 24C 24G 24U 25A 25C 25G 25U 26A 26C 26G 26U
Stv-0D
Stv-1D
Stv-3D
Stv-5D
Stv-7D
WT

(B)

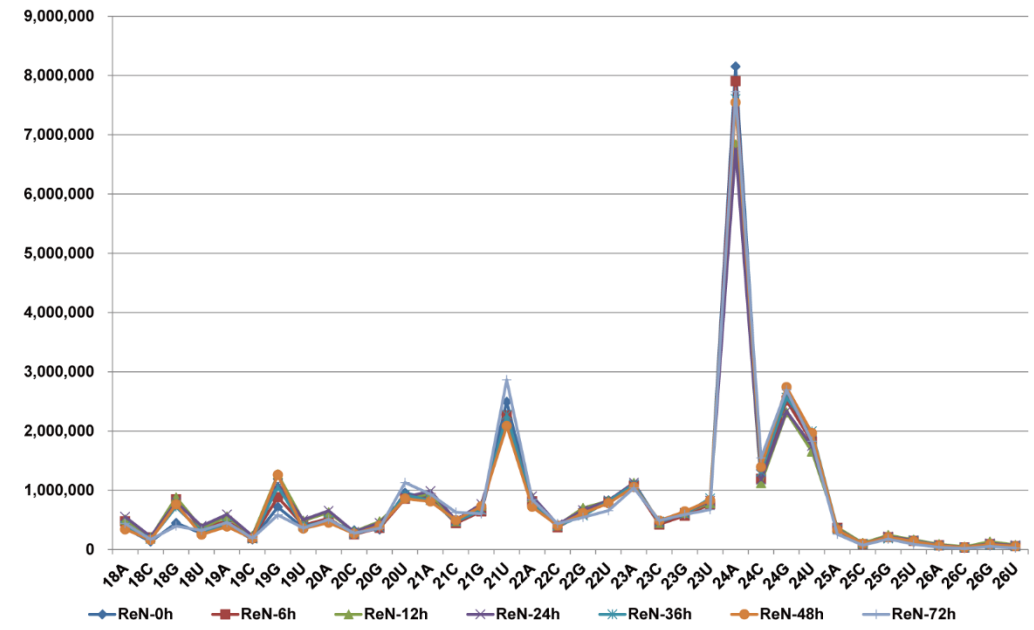

9,000,000
8,000,000
7,000,000
6,000,000
5,000,000
4,000,000
3,000,000
2,000,000
1,000,000
0
18A 18C 18G 18U 19A 19C 19G 19U 20A 20C 20G 20U 21A 21C 21G 21U 22A 22C 22G 22U 23A 23C 23G 23U 24A 24C 24G 24U 25A 25C 25G 25U 26A 26C 26G 26U
ReN-0h
ReN-6h
ReN-12h
ReN-24h
ReN-36h
ReN-48h
ReN-72h

Figure 5. Normalized small RNA read distribution of 13 small RNA-Seq libraries. (A) Distribution of RP30M-normalized small RNA reads on nitrogen-starved (Stv) roots and (B) nitrogen-resupplied (ReN) roots.

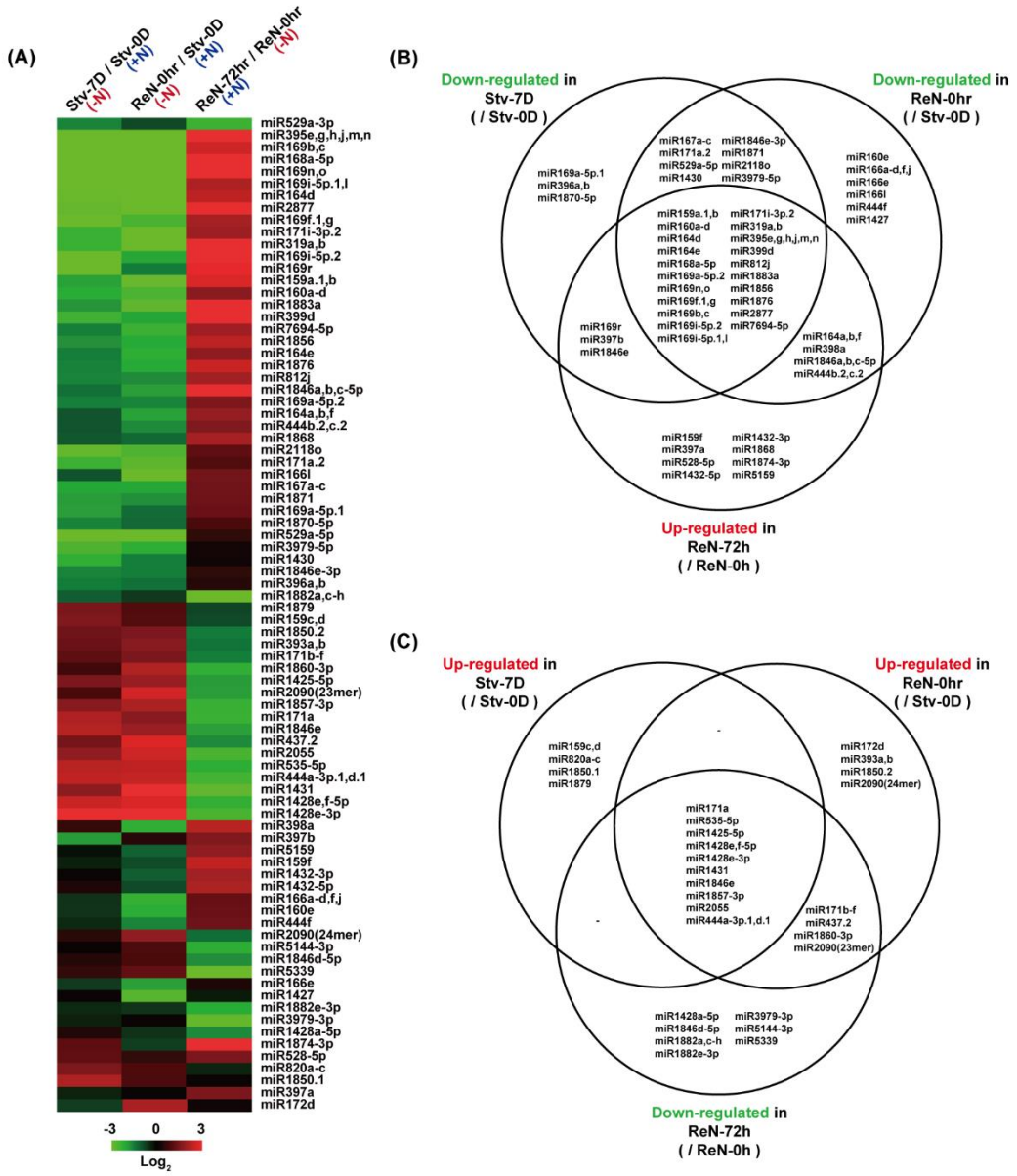
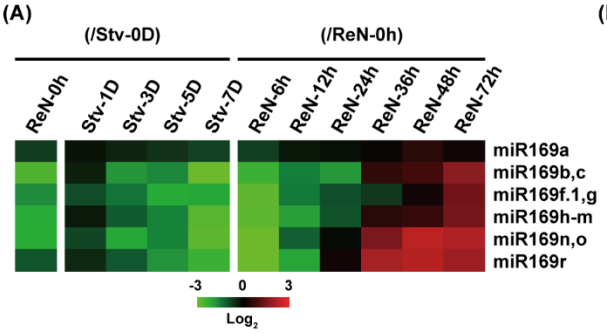

(A)
Stv-7D / Stv-0D
(-N) (+N)
ReN-0hr / Stv-0D
(-N) (+N)
ReN-72hr / ReN-0hr
(+N) (-N)
miR529a-3p
miR395e,g,h,j,m,n
miR169b,c
miR168a-5p
miR169n,o
miR169i-5p.1,l
miR164d
miR2877
miR169f.1,g
miR171i-3p.2
miR319a,b
miR169l-5p.2
miR169r
miR159a.1,b
miR160a-d
miR1883a
miR399d
miR7694-5p
miR1856
miR164e
miR1876
miR812j
miR1846a,b,c-5p
miR169a-5p.2
miR164a,b,f
miR444b.2,c.2
miR1868
miR2118o
miR171a.2
miR166l
miR167a-c
miR1871
miR169a-5p.1
miR1870-5p
miR529a-5p
miR3979-5p
miR1430
miR1846e-3p
miR396a,b
miR1882a,c-h
miR1879
miR159c,d
miR1850.2
miR393a,b
miR171b-f
miR1860-3p
miR1425-5p
miR2090(23mer)
miR1857-3p
miR171a
miR1846e
miR437.2
miR2055
miR535-5p
miR444a-3p.1,d.1
miR1431
miR1428e,f-5p
miR1428e-3p
miR398a
miR397b
miR5159
miR159f
miR1432-3p
miR1432-5p
miR166a-d,f,j
miR160e
miR444f
miR2090(24mer)
miR5144-3p
miR1846d-5p
miR5339
miR166e
miR1427
miR1882e-3p
miR3979-3p
miR1428a-5p
miR1874-3p
miR528-5p
miR820a-c
miR1850.1
miR397a
miR172d
-3 0 3
Log2
(B)
Down-regulated in
Stv-7D
(/ Stv-0D)
Down-regulated in
ReN-0hr
(/ Stv-0D)
miR169a-5p.1
miR396a,b
miR1870-5p
miR167a-c miR1846e-3p
miR171a.2 miR1871
miR529a-5p miR2118o
miR1430 miR3979-5p
miR160e
miR166a-d,f,j
miR166e
miR166l
miR444f
miR1427
miR159a.1,b miR171i-3p.2
miR160a-d miR319a,b
miR164d miR395e,g,h,j,m,n
miR164e miR399d
miR168a-5p miR812j
miR169a-5p.2 miR1883a
miR169n,o miR1856
miR169f.1,g miR1876
miR169b,c miR2877
miR169l-5p.2 miR7694-5p
miR169l-5p.1,l
miR169r
miR397b
miR1846e
miR164a,b,f
miR398a
miR1846a,b,c-5p
miR444b.2,c.2
miR159f miR1432-3p
miR397a miR1868
miR528-5p miR1874-3p
miR1432-5p miR5159
Up-regulated in
ReN-72h
(/ ReN-0h)
(C)
Up-regulated in
Stv-7D
(/ Stv-0D)
Up-regulated in
ReN-0hr
(/ Stv-0D)
miR159c,d
miR820a-c
miR1850.1
miR1879
-
miR172d
miR393a,b
miR1850.2
miR2090(24mer)
miR171a
miR535-5p
miR1425-5p
miR1428e,f-5p
miR1428e-3p
miR1431
miR1846e
miR1857-3p
miR2055
miR444a-3p.1,d.1
-
miR171b-f
miR437.2
miR1860-3p
miR2090(23mer)
miR1428a-5p miR3979-3p
miR1846d-5p miR5144-3p
miR1882a,c-h miR5339
miR1882e-3p
Down-regulated in
ReN-72h
(/ ReN-0h)

Figure 6. Profiling of nitrogen-responsive miRNAs in roots. (A) Heatmap representation of nitrogen-responsive miRNAs. (B) MiRNAs down-regulated in nitrogen starvation and up-regulated in nitrogen re-supply (2-fold cutoff). (C) MiRNAs up-regulated in nitrogen starvation and down-regulated in nitrogen re-supply (2-fold cutoff).

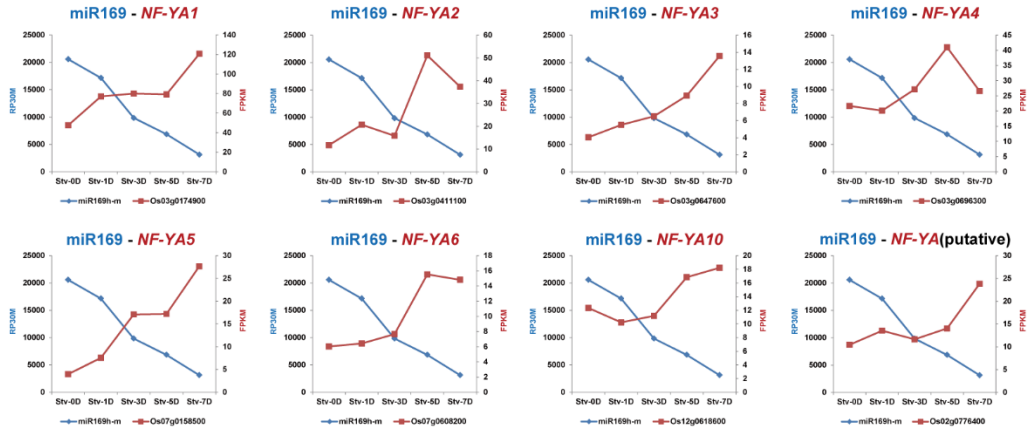
(A)



(B)

```
3′   GGCCGUUCAGUAGGAACCGAC   5′   miR169b,c
3′    UCCGUUCAGUAGGAACCGAU   5′   miR169f.1,g
3′   GUCCGUUCAGUAGGAACCGAU   5′   miR169h-m
3′   AUCCGUUCAGUAAGAACCGAU   5′   miR169n,o
3′  UGUCCGUUUAGUAGGAACCGAU   5′   miR169r
3′   AGCCGUUCAGUAGGAACCGAC   5′   miR169a
     |∘|||||  ||||||||||||
5'  CAUGGCAAAUCAUCCUUGGCUU   3'   NF-YA1
5'  CGUGGCAAUUCAUCCUUGGCUU   3'   NF-YA2
5'  GUAGGCAAAUCAUCCUUGGCUC   3'   NF-YA3
5'  UCAGGCAAUUCAUUCUUGGCUU   3'   NF-YA4
5'  GGUGGCAAUUCAUCCUUGGCUG   3'   NF-YA5
5'  GGUGGCAAUUCAUCCUUGGCUU   3'   NF-YA6
5′  GUAGGCAAAUCAUUCUUGGCUC   3′   NF-YA10
5′  UUAGGCAAUUCAUCCUUGGCUU   3`   NF-Y(putative)
```

(C)



(D)

```
                 6/7
                  ↓
5'   GUGGCAAUUCAUCCUUGGCUG 3'  NF-YA5
      .:::::  :::::::::::::
3'   AGCCGUUCAGUAGGAACCGAC 5'  miR169a
     21 20 19 18 17 16 15 14 13 12 11 10 9 8 7 6 5 4 3 2 1
```

Figure 7. Expression patterns of miR169 family and that of its target genes, NF-YAs. (A) Heatmap representation of miR169 family expression patterns (log-2 scale). (B) Target prediction result of miR169 and its target gene, NF-YA genes. (C) Reciprocal patterns of miR169h-m and NF-YAs. (D) 5' RLM-RACE results of *NF-YA5*.

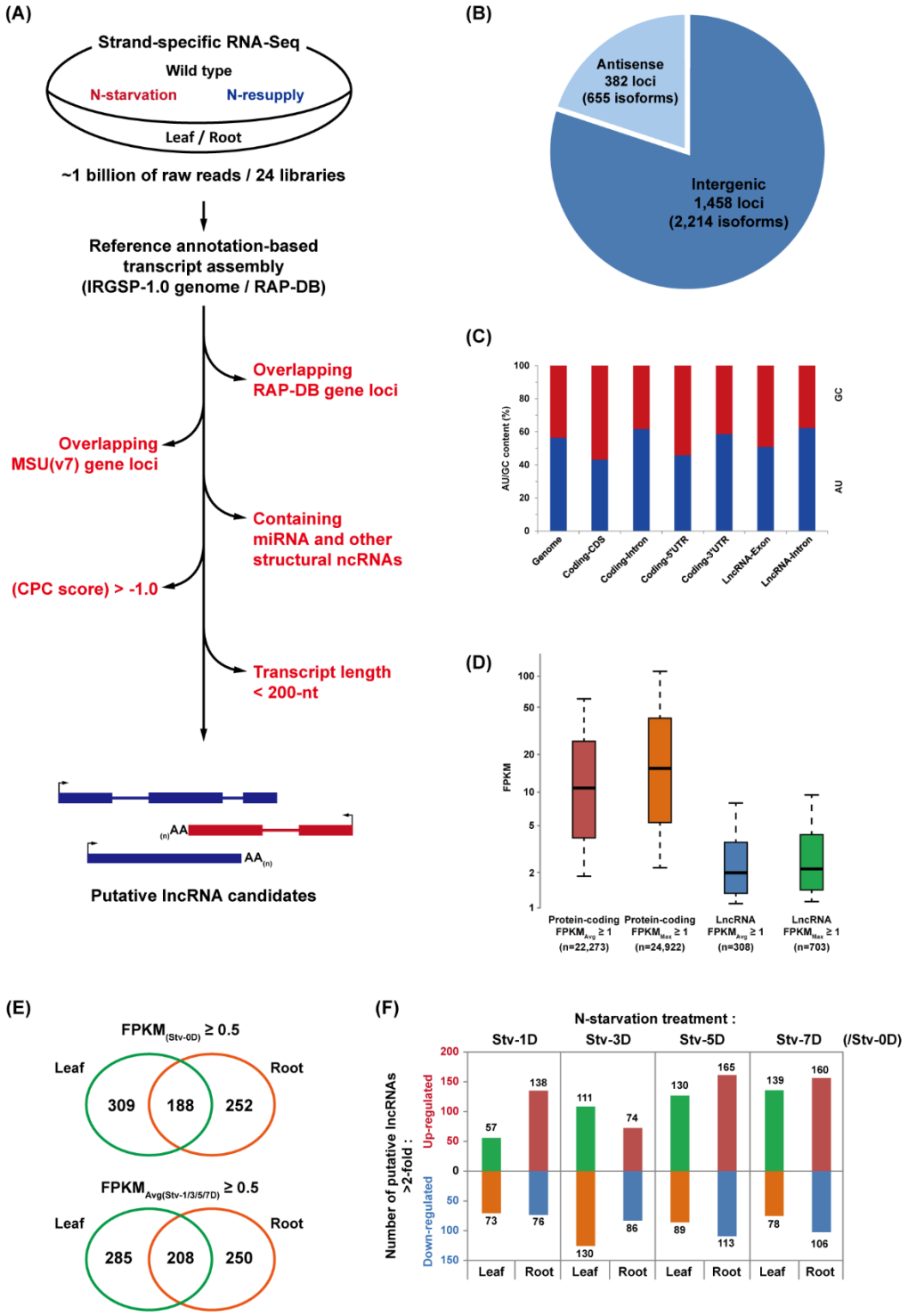

(A)
Strand-specific RNA-Seq
Wild type
N-starvation
N-resupply
Leaf / Root
~1 billion of raw reads / 24 libraries
Reference annotation-based transcript assembly (IRGSP-1.0 genome / RAP-DB)
Overlapping RAP-DB gene loci
Overlapping MSU(v7) gene loci
Containing miRNA and other structural ncRNAs
(CPC score) > -1.0
Transcript length < 200-nt
Putative lncRNA candidates
(B)
Antisense 382 loci (655 isoforms)
Intergenic 1,458 loci (2,214 isoforms)
(C)
AU/GC content (%)
GC
AU
Genome
Coding-CDS
Coding-Intron
Coding-5'UTR
Coding-3'UTR
LncRNA-Exon
LncRNA-Intron
(D)
FPKM
Protein-coding FPKM Avg ≥ 1 (n=22,273)
Protein-coding FPKM Max ≥ 1 (n=24,922)
LncRNA FPKM Avg ≥ 1 (n=308)
LncRNA FPKM Max ≥ 1 (n=703)
(E)
FPKM (Stv-0D) ≥ 0.5
Leaf
Root
309
188
252
FPKM Avg(Stv-1/3/5/7D) ≥ 0.5
285
208
250
(F)
N-starvation treatment :
Stv-1D
Stv-3D
Stv-5D
Stv-7D
(/Stv-0D)
Number of putative lncRNAs >2-fold :
Up-regulated
Down-regulated
57
138
73
76
111
74
130
86
130
165
89
113
139
160
78
106
Leaf
Root

Figure 8. Identification of putative long non-coding RNA loci in rice. (A) Pipeline for identification of lncRNA using strand-specific RNA-Seq data. (B) Genomic position of identified putative lncRNAs. (C) AU/GC content of exon and intron of putative lncRNA loci, compared to that of whole genome and CDS / 5'UTR / 3'UTR / intron of annotated protein-coding gene. (D) FPKM value distribution of annotated protein-coding genes and putative lncRNAs. Box plots and bars respectively indicate 25th / 50th (median) / 75th percentiles and 10th / 90th percentiles. (E) Organ-preferenrtial expression of putative lncRNAs. (F) Number of $\geq$ 2-fold up-regulated and down-regulated putative lncRNAs in roots and leaves ($\geq$ 0.5-FPKM cutoff at least one of two compared time points). Red / green bars indicate up-regulated in roots and leaves, blue / orange bars indicate down-regulated genes in roots and leaves.

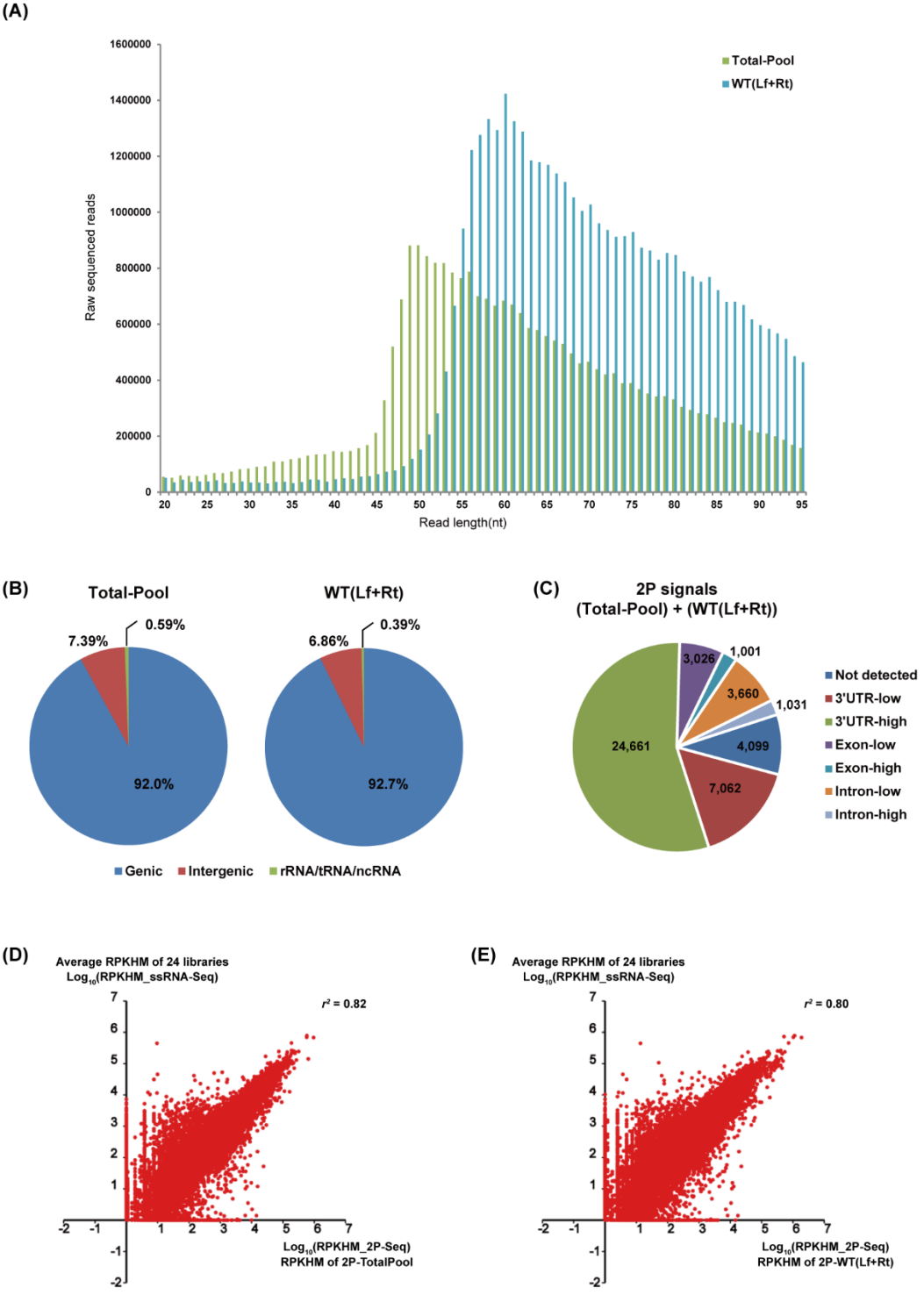

(A)
Total-Pool
WT(Lf+Rt)
Raw sequenced reads
Read length(nt)
(B)
Total-Pool
WT(Lf+Rt)
7.39%
0.59%
92.0%
6.86%
0.39%
92.7%
Genic
Intergenic
rRNA/tRNA/ncRNA
(C)
2P signals
(Total-Pool) + (WT(Lf+Rt))
3,026
1,001
3,660
1,031
4,099
7,062
24,661
Not detected
3'UTR-low
3'UTR-high
Exon-low
Exon-high
Intron-low
Intron-high
(D)
Average RPKHM of 24 libraries
Log10(RPKHM_ssRNA-Seq)
r² = 0.82
Log10(RPKHM_2P-Seq)
RPKHM of 2P-TotalPool
(E)
Average RPKHM of 24 libraries
Log10(RPKHM_ssRNA-Seq)
r² = 0.80
Log10(RPKHM_2P-Seq)
RPKHM of 2P-WT(Lf+Rt)

Figure 9. Results of 2P-Seq data analysis. (A) Read length distribution of sequenced 2P-Seq libraries. (B)(C) Genomic positional distribution of 2P-Seq reads (cutoff = 50 normalized 2P signal). (D)(E) Expressional correlation between two different RNA-Seq, strand specific RNA-Seq and 2P-Seq (RPKHM = reads per kilobase of exon per hundred million reads).

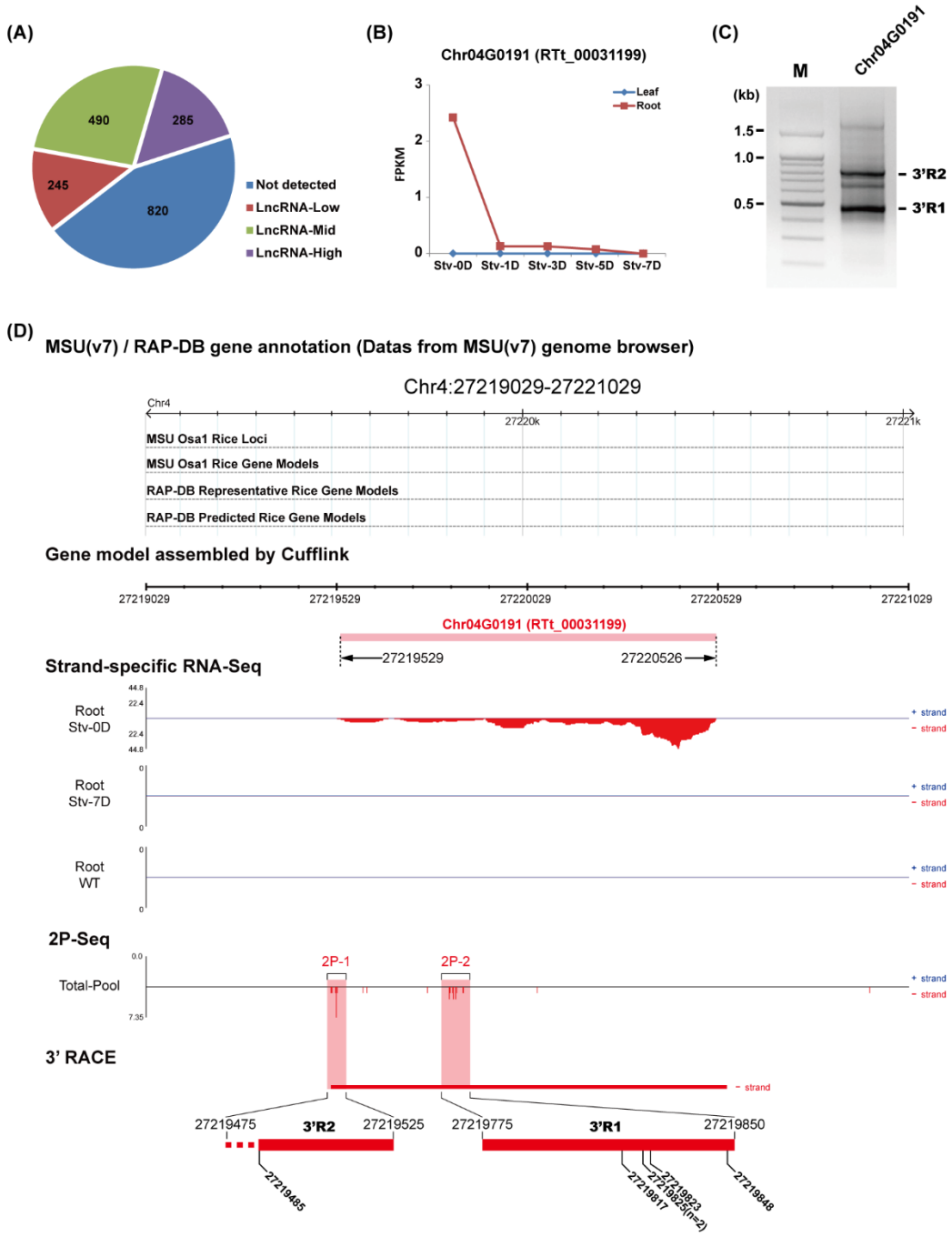

(A)
490
285
245
820
Not detected
LncRNA-Low
LncRNA-Mid
LncRNA-High
(B)
Chr04G0191 (RTt_00031199)
FPKM
Leaf
Root
Stv-0D
Stv-1D
Stv-3D
Stv-5D
Stv-7D
(C)
M
Chr04G0191
(kb)
1.5
1.0
0.5
3'R2
3'R1
(D)
MSU(v7) / RAP-DB gene annotation (Datas from MSU(v7) genome browser)
Chr4:27219029-27221029
Chr4
27220k
27221k
MSU Osa1 Rice Loci
MSU Osa1 Rice Gene Models
RAP-DB Representative Rice Gene Models
RAP-DB Predicted Rice Gene Models
Gene model assembled by Cufflink
27219029
27219529
27220029
27220529
27221029
Chr04G0191 (RTt_00031199)
27219529
27220526
Strand-specific RNA-Seq
Root Stv-0D
Root Stv-7D
Root WT
+ strand
- strand
2P-Seq
Total-Pool
2P-1
2P-2
3' RACE
27219475
3'R2
27219525
27219775
3'R1
27219850
27219485
27219817
27219823
27219825(n=2)
27219848

Figure 10. 2P-Seq could detect 3' end of putative long non-coding RNA. (A) Number of 2P-tag detected putative lncRNA loci; High : (normalized 2P signal) ≥ 50; Mid : 10 ≤ (normalized 2P signal) < 50; Low : (normalized 2P signal) < 10 (B) Expression pattern of Chr04G0191 in N-starved leaves and roots. (C) 3' RACE result. (D) ssRNA-Seq and 2P-Seq signal distribution on putative lncRNA deriving loci.

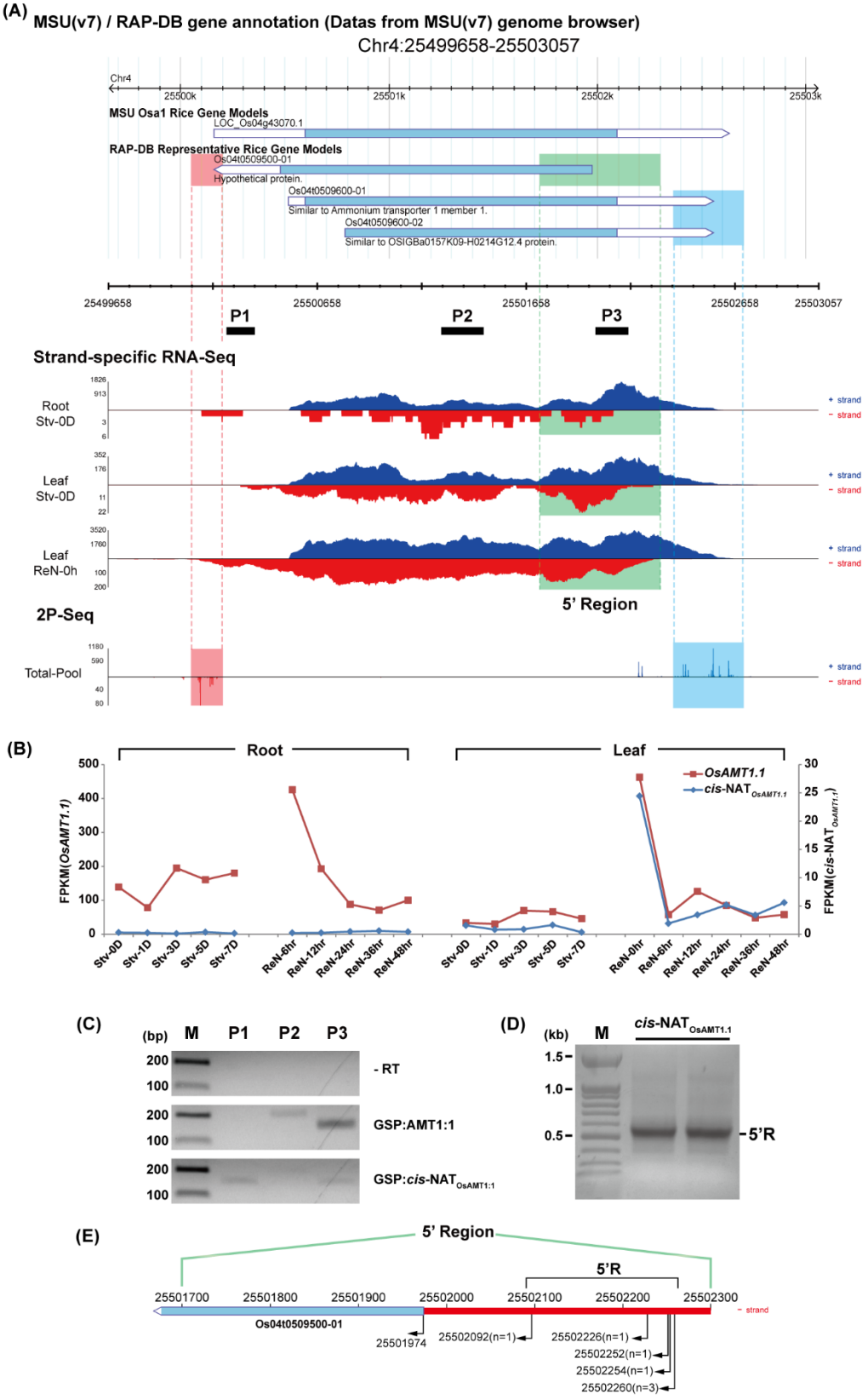
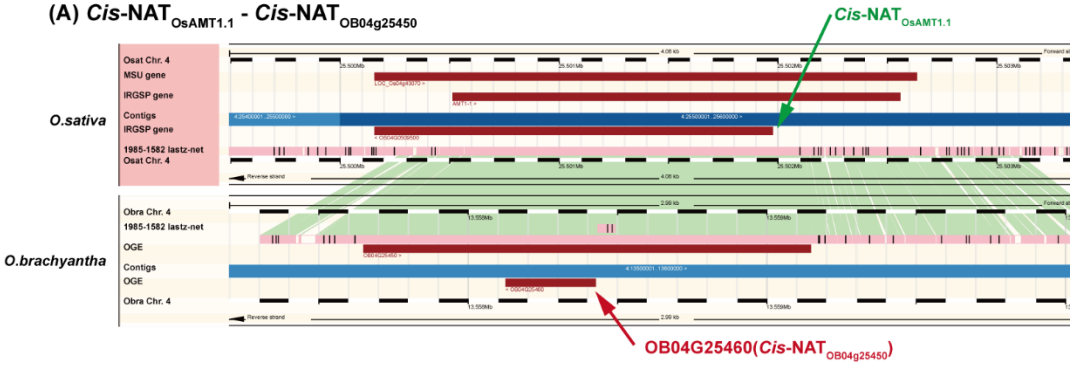
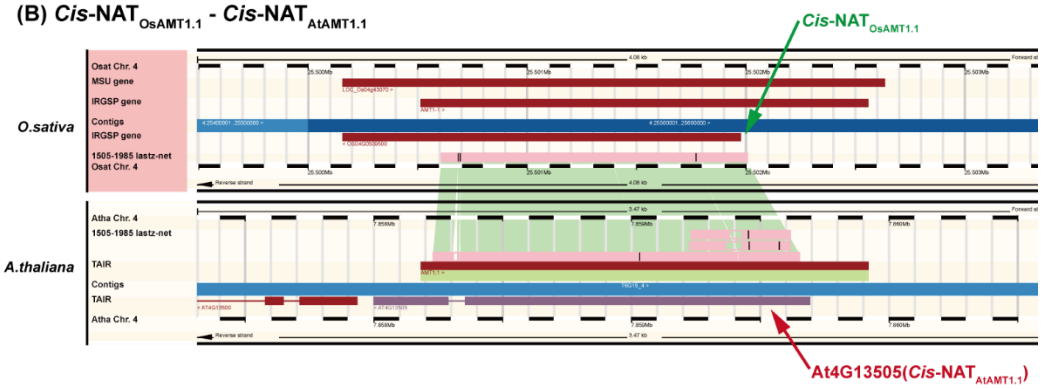

(A)
MSU(v7) / RAP-DB gene annotation (Datas from MSU(v7) genome browser)
Chr4:25499658-25503057
MSU Osa1 Rice Gene Models
LOC_Os04g43070.1
RAP-DB Representative Rice Gene Models
Os04t0509500-01
Hypothetical protein.
Os04t0509600-01
Similar to Ammonium transporter 1 member 1.
Os04t0509600-02
Similar to OSIGBa0157K09-H0214G12.4 protein.
P1
P2
P3
Strand-specific RNA-Seq
Root Stv-0D
Leaf Stv-0D
Leaf ReN-0h
5' Region
2P-Seq
Total-Pool
+ strand
- strand
(B)
Root
Leaf
OsAMT1.1
cis-NAT OsAMT1.1
FPKM(OsAMT1.1)
FPKM(cis-NAT OsAMT1.1)
Stv-0D
Stv-1D
Stv-3D
Stv-5D
Stv-7D
ReN-6hr
ReN-12hr
ReN-24hr
ReN-36hr
ReN-48hr
ReN-0hr
(C)
(bp)
M
P1
P2
P3
- RT
GSP:AMT1:1
GSP:cis-NAT OsAMT1:1
(D)
(kb)
cis-NAT OsAMT1.1
5'R
(E)
5' Region
5'R
Os04t0509500-01
25501974
25502092(n=1)
25502226(n=1)
25502252(n=1)
25502254(n=1)
25502260(n=3)

Figure 11. Expression of *cis*-NAT$_{OsAMT1.1}$. (A) Distribution of strand-specific RNA-Seq reads and 2P-Seq reads on the *OsAMT1.1* genomic locus. (C) Strand-specific RT-PCR result. (D) 5' RLM-RACE PCR result. (E) Position of sequenced 5' RLM-RACE tags.

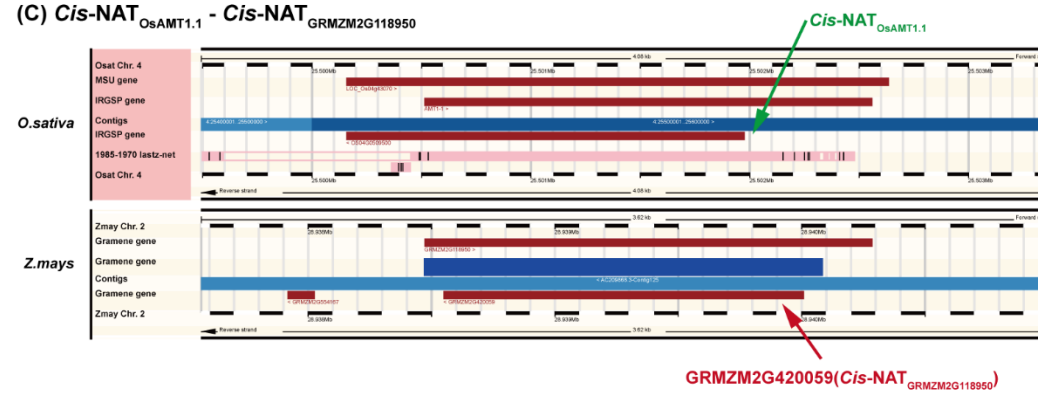
(A) *Cis*-NAT$_{OsAMT1.1}$ - *Cis*-NAT$_{OB04g25450}$



(B) *Cis*-NAT$_{OsAMT1.1}$ - *Cis*-NAT$_{AtAMT1.1}$



(C) *Cis*-NAT$_{OsAMT1.1}$ - *Cis*-NAT$_{GRMZM2G118950}$



(D) *Cis*-NAT$_{OsAMT1.1}$ - *Cis*-NAT$_{GRMZM2G175140}$

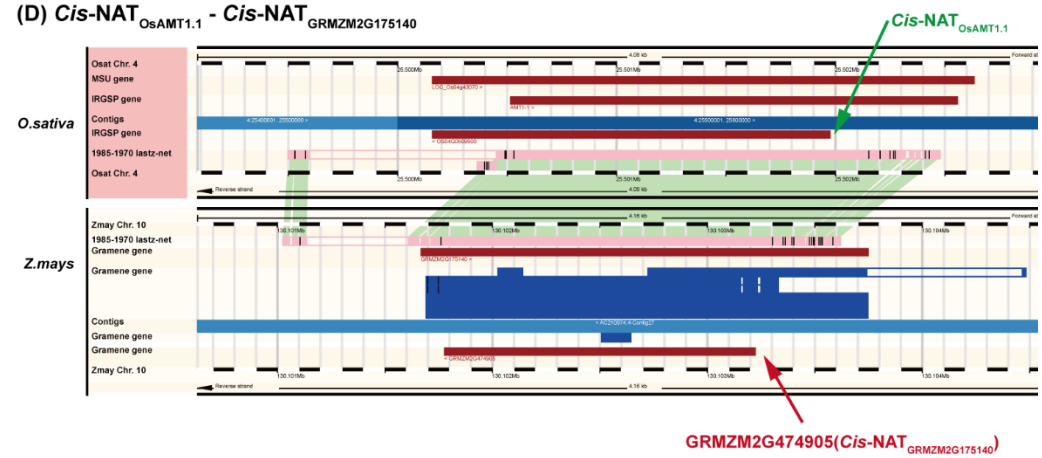


Figure 12. Genome comparison of *AMT1.1* locus between rice and other plants. (A) *Cis*-NAT$_{AMT1.1}$ (OB04G25460) in *Oryza brachyantha*. (B) *Cis*-NAT$_{AMT1.1}$ (AT4G13505) in *Arabidopsis thaliana*. (C) *Cis*-NAT$_{AMT1.1}$ (GRMZM2G420059) and (D) *Cis*-NAT$_{AMT1.1}$ (GRMZM2G474905) in *Zea mays*.

TABLES

Table 1. List of oligonucleotide sequences used in this study.

Name	Sequence (5' to 3')
TruSeq_PCR1	**AATGATACGGCGACCACCGAGATCTACAC--TCTTTCCCTACACGACGCTCTTCCGATC*T**
TruSeq PCR2.1	**CAAGCAGAAGACGGCATACGAGATCGTGATGTG--ACTGGAGTTCAGACGTGTGCTCTTCCGATC*T**
TruSeq PCR2.2	**CAAGCAGAAGACGGCATACGAGATACATCGGTG--ACTGGAGTTCAGACGTGTGCTCTTCCGATC*T**
TruSeq PCR2.3	**CAAGCAGAAGACGGCATACGAGATGCCTAAGTG--ACTGGAGTTCAGACGTGTGCTCTTCCGATC*T**
TruSeq PCR2.4	**CAAGCAGAAGACGGCATACGAGATTGGTCAGTG--ACTGGAGTTCAGACGTGTGCTCTTCCGATC*T**
TruSeq PCR2.5	**CAAGCAGAAGACGGCATACGAGATCACTGTGTG--ACTGGAGTTCAGACGTGTGCTCTTCCGATC*T**
TruSeq PCR2.6	**CAAGCAGAAGACGGCATACGAGATATTGGCGTG--ACTGGAGTTCAGACGTGTGCTCTTCCGATC*T**
TruSeq PCR2.7	**CAAGCAGAAGACGGCATACGAGATGATCTGGTG--ACTGGAGTTCAGACGTGTGCTCTTCCGATC*T**
TruSeq PCR2.8	**CAAGCAGAAGACGGCATACGAGATTCAAGTGTG--ACTGGAGTTCAGACGTGTGCTCTTCCGATC*T**
TruSeq PCR2.9	**CAAGCAGAAGACGGCATACGAGATCTGATCGTG--ACTGGAGTTCAGACGTGTGCTCTTCCGATC*T**
TruSeq PCR2.10	**CAAGCAGAAGACGGCATACGAGATAAGCTAGTG--ACTGGAGTTCAGACGTGTGCTCTTCCGATC*T**
TruSeq PCR2.11	**CAAGCAGAAGACGGCATACGAGATGTAGCCGTG--ACTGGAGTTCAGACGTGTGCTCTTCCGATC*T**
TruSeq PCR2.12	**CAAGCAGAAGACGGCATACGAGATTACAAGGTG--ACTGGAGTTCAGACGTGTGCTCTTCCGATC*T**
TruSeq PCR2.13	**CAAGCAGAAGACGGCATACGAGATTTGACTGTG--ACTGGAGTTCAGACGTGTGCTCTTCCGATC*T**

(p : monophosphate / * : phosphorothioate)

(Continue)

Table 1. (Continue)

Name	Sequence (5' to 3')
TruSeq PCR2.14	CAAGCAGAAGACGGCATACGAGATGGAACTGTG--ACTGGAGTTCAGACGTGTGCTCTTCCGATC*T
TruSeq PCR2.15	CAAGCAGAAGACGGCATACGAGATTGACATGTG--ACTGGAGTTCAGACGTGTGCTCTTCCGATC*T
TruSeq PCR2.16	CAAGCAGAAGACGGCATACGAGATGGACGGGTG--ACTGGAGTTCAGACGTGTGCTCTTCCGATC*T
TruSeq PCR2.17	CAAGCAGAAGACGGCATACGAGATCTCTACGTG--ACTGGAGTTCAGACGTGTGCTCTTCCGATC*T
TruSeq PCR2.18	CAAGCAGAAGACGGCATACGAGATGCGGACGTG--ACTGGAGTTCAGACGTGTGCTCTTCCGATC*T
TruSeq PCR2.19	CAAGCAGAAGACGGCATACGAGATTTTCACGTG--ACTGGAGTTCAGACGTGTGCTCTTCCGATC*T
TruSeq PCR2.20	CAAGCAGAAGACGGCATACGAGATGGCCACGTG--ACTGGAGTTCAGACGTGTGCTCTTCCGATC*T
TruSeq PCR2.21	CAAGCAGAAGACGGCATACGAGATCGAAACGTG--ACTGGAGTTCAGACGTGTGCTCTTCCGATC*T
TruSeq PCR2.22	CAAGCAGAAGACGGCATACGAGATCGTACGGTG--ACTGGAGTTCAGACGTGTGCTCTTCCGATC*T
TruSeq PCR2.23	CAAGCAGAAGACGGCATACGAGATCCACTCGTG--ACTGGAGTTCAGACGTGTGCTCTTCCGATC*T
TruSeq PCR2.24	CAAGCAGAAGACGGCATACGAGATGCTACCGTG--ACTGGAGTTCAGACGTGTGCTCTTCCGATC*T
TruSeq_Upper adapter	ACACTCTTTCCCTACACGACGCTCTTCCGATC*T
TruSeq_Lower adapter	pGATCGGAAGAGCACACGTCTGAACTCCAGTCA*C
Small RNA marker-18nt	GGUGGACAUCACUUACGC
Small RNA marker-30nt	AGUAUGGGCAUUUCGCAGCCUACCGUGGUG

(p : monophosphate / * : phosphorothioate / underline : complementary sequences)

(Continue)

Table 1. (Continue)

Name	Sequence (5' to 3')
Small RNA 3' Adapter	(rApp)TGGAATTCTCGGGTGCCAAGG(ddC)
Small RNA 5' Adapter	GUUCAGAGUUCUACAGUCCGACGAUC
Small RNA RT primer	CCTTGGCACCCGAGAATTCCA
RP_Forward	AATGATACGGCGACCACCGAGATCTACAC- -GTTCAGAGTTCTACAGTCCGA
RPI_01	CAAGCAGAAGACGGCATACGAGATCGTGATGTG- -ACTGGAGTTCCTTGGCACCCGAGAATTCCA
RPI_02	CAAGCAGAAGACGGCATACGAGATACATCGGTG- -ACTGGAGTTCCTTGGCACCCGAGAATTCCA
RPI_03	CAAGCAGAAGACGGCATACGAGATGCCTAAGTG- -ACTGGAGTTCCTTGGCACCCGAGAATTCCA
RPI_04	CAAGCAGAAGACGGCATACGAGATTGGTCAGTG- -ACTGGAGTTCCTTGGCACCCGAGAATTCCA
RPI_05	CAAGCAGAAGACGGCATACGAGATCACTGTGTG- -ACTGGAGTTCCTTGGCACCCGAGAATTCCA
RPI_06	CAAGCAGAAGACGGCATACGAGATATTGGCGTG- -ACTGGAGTTCCTTGGCACCCGAGAATTCCA
2P-RT1p	pTGGAATTCTCGGGTGCCAAGGAACTCCAGT- -CAC(iSp18)CACTCA(iSp18)CTTTCCCTACACGAC- -GCTCTTCCGATCTTTTTTTTTTTTTTTTTTTTVN
2P-PCR-F.1	AATGATACGGCGACCACCGAGATCTACAC- -TCTTTCCCTACACGACGCTCTTCCGATCT
2P-PCR-RPI1	CAAGCAGAAGACGGCATACGAGATCGTGATGTG- -ACTGGAGTTCCTTGGCACCCGAGAATTCCA
2P-PCR-RPI2	CAAGCAGAAGACGGCATACGAGATACATCGGTG- -ACTGGAGTTCCTTGGCACCCGAGAATTCCA

(rApp : adenylated / ddC : dideoxycytidine / iSp18 : hexaethyleneglycol)

(Continue)

Table 1. (Continue)

Name	Sequence (5' to 3')
2P-Seq-PE1.1	ACACTCTTTCCCTACACGACGCT--CTTCCGATCTTTTTTTTTTTTTTTTTTTT
GeneRacer 5' Adapter	CGACUGGAGCACGAGGACACUGA--CAUGGACUGAAGGAGUAGAAA
GeneRacer RT primer	GCTGTCAACGATACGCTACGT-AACGGCATGACAGTG(T)24
GeneRacer 5´ Primer	CGACTGGAGCACGAGGACACTGA
GeneRacer 5´ Nested Primer	GGACACTGACATGGACTGAAGGAGTA
GeneRacer 3´ Primer	GCTGTCAACGATACGCTACGTAACG
GeneRacer 3´ Nested Primer	CGCTACGTAACGGCATGACAGTG
3'RTP(Chr04G0191)	CATTAGTGGCAGACGATGTGTC
3'RNP(Chr04G0191)	CTAGCTGCTGCTTAAGCCCT
5miRTP(*OsNFYA5*)	ACATGAGCAACTGAGAAGCAGTTTGC
5miRNP(*OsNFYA5*)	GCAGTTTGCATCAGTTGCATGAAATTG
5CapRTP (*cis*-NAT$_{OsAMT1.1}$)	GCGCTCTTCGCGAGGCAGAAGTACGT
5CapRNP (*cis*-NAT$_{OsAMT1.1}$)	GCGAGGCAGAAGTACGTCGAGGAGATCT
qPCR0005	GGGTTGGTGGGTTCATGCT

(Continue)

Table 1. (Continue)

Name	Sequence (5' to 3')
qPCR0006	CGAGATCGGAGAGGGAGAGAA
qPCR0009	GAATCTTCCGTGGTGGAGGG
qPCR0010	GACATGACACAGGGAGCGAA
qPCR0026	CTAACCTCTGATTCTTCCCACCT
qPCR0027	TGAATCCGTACCAGCCGAAC
qPCR0028	GCTCCGGTGTCATCGACTTC

Table 2. Read distribution of strand-specific RNA-Seq libraries

Library	Raw Reads	Genome-alined
Root_Wt	44,982,276	41,256,577
Leaf_Wt	40,340,692	38,708,625
Root_Stv-0D_Rep1	46,305,580	42,353,950
Root_Stv-0D_Rep2	47,098,224	44,077,358
Root_Stv-1D	39,422,438	36,753,294
Root_Stv-3D	40,470,104	37,454,123
Root_Stv-5D	36,666,360	31,713,237
Root_Stv-7D	46,180,760	40,346,856
Leaf_Stv-0D	43,182,458	41,372,279
Leaf_Stv-1D	48,712,784	46,806,348
Leaf_Stv-3D	42,041,466	40,433,735
Leaf_Stv-5D	35,885,362	34,529,262
Leaf_Stv-7D	49,096,674	46,999,841
Root_ReN-6h	44,282,622	39,087,824
Root_ReN-12h	48,919,356	42,526,145
Root_ReN-24h	43,026,228	38,924,707
Root_ReN-36h	32,133,870	27,768,488
Root_ReN-48h	40,900,256	36,502,573
Leaf_ReN-0h	45,592,994	43,504,013
Leaf_ReN-6h	45,748,360	43,518,529
Leaf_ReN-12h	37,412,294	35,767,362
Leaf_ReN-24h	35,258,524	33,647,410
Leaf_ReN-36h	41,749,172	39,975,434
Leaf_ReN-48h	46,389,370	44,257,668
Total	**1,021,798,224**	**948,285,638**

Table 3. Read distribution of small RNA-Seq libraries

Library	Raw Reads	Adapter-clipped	Genome-alined
Root_Wt	47,428,219	47,339,177	24,396,955
Root_Stv-0D	38,497,507	38,433,809	19,205,706
Root_Stv-1D	39,380,343	39,314,156	24,152,785
Root_Stv-3D	38,403,526	38,338,837	23,262,781
Root_Stv-5D	40,219,110	40,147,608	21,379,565
Root_Stv-7D	49,797,334	49,709,205	20,976,033
Root_ReN-0h	21,492,871	21,457,033	7,663,338
Root_ReN-6h	18,654,336	18,623,048	4,532,445
Root_ReN-12h	18,022,328	17,991,395	3,732,108
Root_ReN-24h	18,129,332	18,098,595	4,103,845
Root_ReN-36h	16,187,683	16,160,256	3,752,082
Root_ReN-48h	21,411,453	21,375,750	5,982,864
Root_ReN-72h	11,225,723	11,206,654	4,914,130
Total	**378,849,765**	**378,195,523**	**168,054,637**

Table 4. Read distribution of 2P-Seq libraries

Library	Raw reads	20nt ≤ (Read) ≤ 95nt	Genome-alined
Mixed total	**69,184,708**	**59,082,152**	**49,924,828**
WT_LF/RT	**92,095,773**	**78,823,319**	**73,339,193**
Total	**161,280,481**	**137,905,471**	**123,264,021**

Table 5. Summary of putative long non-coding RNA loci identification results

Step	Filter Condition			Leaves	Roots
1	Assembled transcript models			69,848	70,322
2	Not matched with RAP-DB gene models			4,161	4,635
3	Not overlapping MSU(v7) gene models			2,628	2,718
4	Not overlapping miRNAs / structural RNAs			2,587	2,674
5	Integration	Identical intron-exon transcript models between leaves and roots		1,029	
		Unique transcript models		4,232	
				Intergenic	Antisense
		Genomic location of transcript models		3,025	1,207
		Integrate putative lncRNA loci		2,446	990 [1)]
6	Coding Potential Calculator (CPC)	CPC analysis with rice gene-included Swiss-Prot DB	Loci (CPC value) ≥ 0 (1)	701	344
			Loci '-1.0 < (CPC value) < 0 (1)	570	356
			Loci (CPC value) ≤ -1.0	1,175 (a)	290 (b)
7		CPC analysis against (1) with rice gene-excluded Swiss-Prot DB	Loci (CPC value) ≥ 0	151	133
			Loci '-1.0 < (CPC value) < 0	815	475
			Loci (CPC value) ≤ -1.0	305 (c)	92 (d)
6-7	CPC-cleaned putative lncRNA loci (≤ -1.0)			1,480 (a+c)	382 (b+d)
8	Loci containing transcript ≥ 200-nt			1,458	382

REFERENCE

Araki, R., and Hasegawa, H. (2006). Expression of rice (Oryza sativa L.) genes involved in high-affinity nitrate transport during the period of nitrate induction. Breeding Sci *56*, 295-302.

Bari, R., Pant, B.D., Stitt, M., and Scheible, W.R. (2006). PHO2, microRNA399, and PHR1 define a phosphate-signaling pathway in plants. Plant Physiol *141*, 988-999.

Bi, Y.M., Kant, S., Clark, J., Gidda, S., Ming, F., Xu, J.Y., Rochon, A., Shelp, B.J., Hao, L.X., Zhao, R., *et al.* (2009). Increased nitrogen-use efficiency in transgenic rice plants over-expressing a nitrogen-responsive early nodulin gene identified from rice expression profiling. Plant Cell Environ *32*, 1749-1760.

Brauer, E.K., Rochon, A., Bi, Y.M., Bozzo, G.G., Rothstein, S.J., and Shelp, B.J. (2011). Reappraisal of nitrogen use efficiency in rice overexpressing glutamine synthetase1. Physiol Plantarum *141*, 361-372.

Breeze, E., Harrison, E., McHattie, S., Hughes, L., Hickman, R., Hill, C., Kiddle, S., Kim, Y.S., Penfold, C.A., Jenkins, D., *et al.* (2011). High-

Resolution Temporal Profiling of Transcripts during Arabidopsis Leaf Senescence Reveals a Distinct Chronology of Processes and Regulation. Plant Cell *23*, 873-894.

Cai, H.M., Lu, Y.G., Xie, W.B., Zhu, T., and Lian, X.M. (2012). Transcriptome response to nitrogen starvation in rice. J Biosciences *37*, 731-747.

Cai, H.M., Xiao, J.H., Zhang, Q.F., and Lian, X.M. (2010). Co-suppressed glutamine synthetase2 gene modifies nitrogen metabolism and plant growth in rice. Chinese Sci Bull *55*, 823-833.

Cai, H.M., Zhou, Y., Xiao, J.H., Li, X.H., Zhang, Q.F., and Lian, X.M. (2009). Overexpressed glutamine synthetase gene modifies nitrogen metabolism and abiotic stress responses in rice. Plant Cell Rep *28*, 527-537.

Carninci, P., Kasukawa, T., Katayama, S., Gough, J., Frith, M.C., Maeda, N., Oyama, R., Ravasi, T., Lenhard, B., Wells, C., *et al.* (2005). The transcriptional landscape of the mammalian genome. Science *309*, 1559-1563.

Claverie, J.M. (2005). Fewer genes, more noncoding RNA. Science *309*, 1529-1530.

De Angeli, A., Monachello, D., Ephritikhine, G., Frachisse, J.M., Thomine,

S., Gambale, F., and Barbier-Brygoo, H. (2006). The nitrate/proton antiporter AtCLCa mediates nitrate accumulation in plant vacuoles. Nature *442*, 939-942.

Fang, Z.M., Xia, K.F., Yang, X., Grotemeyer, M.S., Meier, S., Rentsch, D., Xu, X.L., and Zhang, M.Y. (2013). Altered expression of the PTR/NRT1 homologue OsPTR9 affects nitrogen utilization efficiency, growth and grain yield in rice. Plant Biotechnol J *11*, 446-458.

Feng, H.M., Yan, M., Fan, X.R., Li, B.Z., Shen, Q.R., Miller, A.J., and Xu, G.H. (2011). Spatial expression and regulation of rice high-affinity nitrate transporters by nitrogen and carbon status. Journal of experimental botany *62*, 2319-2332.

Franco-Zorrilla, J.M., Valli, A., Todesco, M., Mateos, I., Puga, M.I., Rubio-Somoza, I., Leyva, A., Weigel, D., Garcia, J.A., and Paz-Ares, J. (2007). Target mimicry provides a new mechanism for regulation of microRNA activity. Nature genetics *39*, 1033-1037.

Gifford, M.L., Dean, A., Gutierrez, R.A., Coruzzi, G.M., and Birnbaum, K.D. (2008). Cell-specific nitrogen responses mediate developmental plasticity. Proceedings of the National Academy of Sciences of the United States of America *105*, 803-808.

Gregersen, P.L., Holm, P.B., and Krupinska, K. (2008). Leaf senescence and nutrient remobilisation in barley and wheat. Plant Biology *10*, 37-49.

Heo, J.B., and Sung, S. (2011). Vernalization-Mediated Epigenetic Silencing by a Long Intronic Noncoding RNA. Science *331*, 76-79.

Hirel, B., Le Gouis, J., Ney, B., and Gallais, A. (2007). The challenge of improving nitrogen use efficiency in crop plants: towards a more central role for genetic variability and quantitative genetics within integrated approaches. Journal of experimental botany *58*, 2369-2387.

Hoque, M.S., Masle, J., Udvardi, M.K., Ryan, P.R., and Upadhyaya, N.M. (2006). Over-expression of the rice OsAMT1-1 gene increases ammonium uptake and content, but impairs growth and development of plants under high ammonium nutrition. Functional Plant Biology *33*, 153-163.

Huarte, M., Guttman, M., Feldser, D., Garber, M., Koziol, M.J., Kenzelmann-Broz, D., Khalil, A.M., Zuk, O., Amit, I., Rabani, M., *et al.* (2010). A Large Intergenic Noncoding RNA Induced by p53 Mediates Global Gene Repression in the p53 Response. Cell *142*, 409-419.

Ishiyama, K., Inoue, E., Tabuchi, M., Yamaya, T., and Takahashi, H. (2004). Biochemical background and compartmentalized functions of cytosolic

glutamine synthetase for active ammonium assimilation in rice roots. Plant Cell Physiol *45*, 1640-1647.

Jabnoune, M., Secco, D., Lecampion, C., Robaglia, C., Shu, Q., and Poirier, Y. (2013). A rice cis-natural antisense RNA acts as a translational enhancer for its cognate mRNA and contributes to phosphate homeostasis and plant fitness. The Plant Cell Online *25*, 4166-4182.

Jeong, D.H., Park, S., Zhai, J., Gurazada, S.G., De Paoli, E., Meyers, B.C., and Green, P.J. (2011). Massive analysis of rice small RNAs: mechanistic implications of regulated microRNAs and variants for differential target RNA cleavage. Plant Cell *23*, 4185-4207.

Johnson, J.M., Edwards, S., Shoemaker, D., and Schadt, E.E. (2005). Dark matter in the genome: evidence of widespread transcription detected by microarray tiling experiments. Trends Genet *21*, 93-102.

Kant, S., Bi, Y.M., and Rothstein, S.J. (2011). Understanding plant response to nitrogen limitation for the improvement of crop nitrogen use efficiency. Journal of experimental botany *62*, 1499-1509.

Kapranov, P., and St Laurent, G. (2012). Dark Matter RNA: Existence, Function, and Controversy. Front Genet *3*, 60.

Kong, L., Zhang, Y., Ye, Z.Q., Liu, X.Q., Zhao, S.Q., Wei, L., and Gao, G. (2007). CPC: assess the protein-coding potential of transcripts using sequence features and support vector machine. Nucleic Acids Research *35*, W345-W349.

Lee, E.J., Matsumura, Y., Soga, K., Hoson, T., and Koizumi, N. (2007). Glycosyl hydrolases of cell wall are induced by sugar starvation in Arabidopsis. Plant Cell Physiol *48*, 405-413.

Li, L., Eichten, S.R., Shimizu, R., Petsch, K., Yeh, C.T., Wu, W., Chettoor, A.M., Givan, S.A., Cole, R.A., Fowler, J.E., *et al.* (2014). Genome-wide discovery and characterization of maize long non-coding RNAs. Genome Biol *15*.

Li, S.M., Li, B.Z., and Shi, W.M. (2012). Expression Patterns of Nine Ammonium Transporters in Rice in Response to N Status. Pedosphere *22*, 860-869.

Li, W.X., Oono, Y., Zhu, J.H., He, X.J., Wu, J.M., Iida, K., Lu, X.Y., Cui, X.P., Jin, H.L., and Zhu, J.K. (2008). The Arabidopsis NFYA5 transcription factor is regulated transcriptionally and posttranscriptionally to promote drought resistance. Plant Cell *20*, 2238-2251.

Lian, X.M., Wang, S.P., Zhang, J.W., Feng, Q., Zhang, L.D., Fan, D.L., Li,

X.H., Yuan, D.J., Han, B., and Zhang, Q.F. (2006). Expression profiles of 10,422 genes at early stage of low nitrogen stress in rice assayed using a cDNA microarray. Plant Mol Biol *60*, 617-631.

Liang, G., He, H., and Yu, D.Q. (2012). Identification of Nitrogen Starvation-Responsive MicroRNAs in Arabidopsis thaliana. PLoS ONE *7*, e48951.

Liang, G., and Yu, D. (2010). Reciprocal regulation among miR395, APS and SULTR2;1 in Arabidopsis thaliana. Plant signaling & behavior *5*, 1257-1259.

Liszkay, A., van der Zalm, E., and Schopfer, P. (2004). Production of reactive oxygen intermediates ($O_2^{\bullet-}$, H_2O_2, and •OH) by maize roots and their role in wall loosening and elongation growth. Plant Physiol *136*, 3114-3123.

Liu, J., Jung, C., Xu, J., Wang, H., Deng, S.L., Bernad, L., Arenas-Huertero, C., and Chua, N.H. (2012). Genome-Wide Analysis Uncovers Regulation of Long Intergenic Noncoding RNAs in Arabidopsis. Plant Cell *24*, 4333-4345.

Lu, T.T., Zhu, C.R., Lu, G.J., Guo, Y.L., Zhou, Y., Zhang, Z.Y., Zhao, Y., Li, W.J., Lu, Y., Tang, W.H., *et al.* (2012). Strand-specific RNA-seq reveals widespread occurrence of novel cis-natural antisense transcripts in rice. BMC Genomics *13*.

Mattick, J.S. (2005). The functional genomics of noncoding RNA. Science *309*, 1527-1528.

Nam, J.W., and Bartel, D.P. (2012). Long noncoding RNAs in C. elegans. Genome Res *22*, 2529-2540.

Nischal, L., Mohsin, M., Khan, I., Kardam, H., Wadhwa, A., Abrol, Y.P., Iqbal, M., and Ahmad, A. (2012). Identification and comparative analysis of microRNAs associated with low-N tolerance in rice genotypes. PLoS ONE *7*, e50261.

Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S., Nikaido, I., Osato, N., Saito, R., Suzuki, H., *et al.* (2002). Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs. Nature *420*, 563-573.

Riddihough, G. (2005). In the forests of RNA dark matter. Science *309*, 1816-1816.

Rinn, J.L., Kertesz, M., Wang, J.K., Squazzo, S.L., Xu, X., Brugmann, S.A., Goodnough, L.H., Helms, J.A., Farnham, P.J., Segal, E., *et al.* (2007). Functional demarcation of active and silent chromatin domains in human HOX loci by Noncoding RNAs. Cell *129*, 1311-1323.

Secco, D., Jabnoune, M., Walker, H., Shou, H.X., Wu, P., Poirier, Y., and Whelan, J. (2013). Spatio-Temporal Transcript Profiling of Rice Roots and Shoots in Response to Phosphate Starvation and Recovery. Plant Cell *25*, 4285-4304.

Shin, R., Berg, R.H., and Schachtman, D.P. (2005). Reactive oxygen species and root hairs in Arabidopsis root response to nitrogen, phosphorus and potassium deficiency. Plant Cell Physiol *46*, 1350-1357.

Song, X.W., Li, P.C., Zhai, J.X., Zhou, M., Ma, L.J., Liu, B., Jeong, D.H., Nakano, M., Cao, S.Y., Liu, C.Y., *et al.* (2012). Roles of DCL4 and DCL3b in rice phased small RNA biogenesis. Plant Journal *69*, 462-474.

Song, X.W., Wang, D.K., Ma, L.J., Chen, Z.Y., Li, P.C., Cui, X., Liu, C.Y., Cao, S.Y., Chu, C.C., Tao, Y.Z., *et al.* (2012). Rice RNA-dependent RNA polymerase 6 acts in small RNA biogenesis and spikelet development. Plant Journal *71*, 378-389.

Sonoda, Y., Ikeda, A., Saiki, S., von Wiren, N., Yamaya, T., and Yamaguchi, J. (2003a). Distinct expression and function of three ammonium transporter genes (OsAMT1;1-1;3) in rice. Plant Cell Physiol *44*, 726-734.

Sonoda, Y., Ikeda, A., Saiki, S., Yamaya, T., and Yamaguchi, J. (2003b).

Feedback regulation of the ammonium transporter gene family AMT1 by glutamine in rice. Plant Cell Physiol *44*, 1396-1402.

Sunkar, R., Kapoor, A., and Zhu, J.K. (2006). Posttranscriptional induction of two Cu/Zn superoxide dismutase genes in Arabidopsis is mediated by downregulation of miR398 and important for oxidative stress tolerance. Plant Cell *18*, 2051-2065.

Tilman, D., Cassman, K.G., Matson, P.A., Naylor, R., and Polasky, S. (2002). Agricultural sustainability and intensive production practices. Nature *418*, 671-677.

Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D.R., Pimentel, H., Salzberg, S.L., Rinn, J.L., and Pachter, L. (2012). Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nat Protoc *7*, 562-578.

Ulitsky, I., Shkumatava, A., Jan, C.H., Sive, H., and Bartel, D.P. (2011). Conserved Function of lincRNAs in Vertebrate Embryonic Development despite Rapid Sequence Evolution. Cell *147*, 1537-1550.

Vidal, E.A., Araus, V., Lu, C., Parry, G., Green, P.J., Coruzzi, G.M., and Gutierrez, R.A. (2010). Nitrate-responsive miR393/AFB3 regulatory module

controls root system architecture in Arabidopsis thaliana. Proceedings of the National Academy of Sciences of the United States of America *107*, 4477-4482.

Wang, H., Chung, P.J., Liu, J., Jang, I.C., Kean, M.J., Xu, J., and Chua, N.H. (2014). Genome-wide identification of long noncoding natural antisense transcripts and their responses to light in Arabidopsis. Genome Res *24*, 444-453.

Xin, M.M., Wang, Y., Yao, Y.Y., Song, N., Hu, Z.R., Qin, D.D., Xie, C.J., Peng, H.R., Ni, Z.F., and Sun, Q.X. (2011). Identification and characterization of wheat long non-protein coding RNAs responsive to powdery mildew infection and heat stress by using microarray analysis and SBS sequencing. BMC plant biology *11*.

Xu, G.H., Fan, X.R., and Miller, A.J. (2012). Plant Nitrogen Assimilation and Use Efficiency. Annu Rev Plant Biol *63*, 153-182.

Xu, Z., Zhong, S., Li, X., Li, W., Rothstein, S.J., Zhang, S., Bi, Y., and Xie, C. (2011). Genome-wide identification of microRNAs in response to low nitrate availability in maize leaves and roots. PLoS ONE *6*, e28009.

Yan, Y., Wang, H., Hamera, S., Chen, X., and Fang, R. (2014). miR444a has multiple functions in the rice nitrate-signaling pathway. Plant Journal *78*, 44-55.

Yoshida, S., Forno, D.A., Cock, J.H. and Gomez, K.A. (1976). Laboratory manual for physiological studies of rice. Philippines: IRRI; 83.

Zhang, X.H., Zou, Z., Gong, P.J., Zhang, J.H., Ziaf, K., Li, H.X., Xiao, F.M., and Ye, Z.B. (2011). Over-expression of microRNA169 confers enhanced drought tolerance to tomato. Biotechnol Lett *33*, 403-409.

Zhao, B.T., Ge, L.F., Liang, R.Q., Li, W., Ruan, K.C., Lin, H.X., and Jin, Y.X. (2009). Members of miR-169 family are induced by high salinity and transiently inhibit the NF-YA transcription factor. BMC molecular biology *10*.

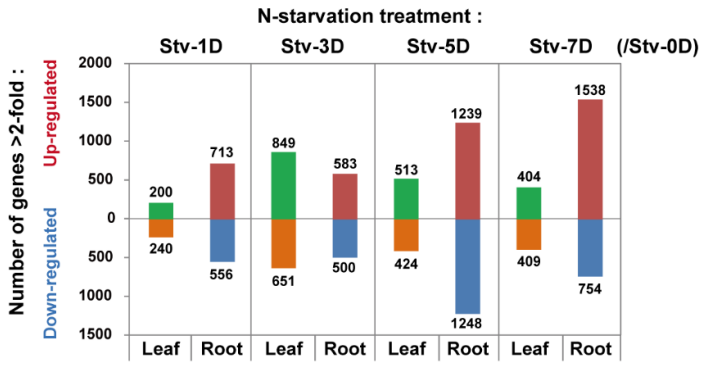
Zhao, M., Ding, H., Zhu, J.K., Zhang, F.S., and Li, W.X. (2011). Involvement of miR169 in the nitrogen-starvation responses in Arabidopsis. New Phytol *190*, 906-915.

Zhao, M., Tai, H.H., Sun, S.Z., Zhang, F.S., Xu, Y.B., and Li, W.X. (2012). Cloning and Characterization of Maize miRNAs Involved in Responses to Nitrogen Deficiency. PLoS ONE *7*, e29669.

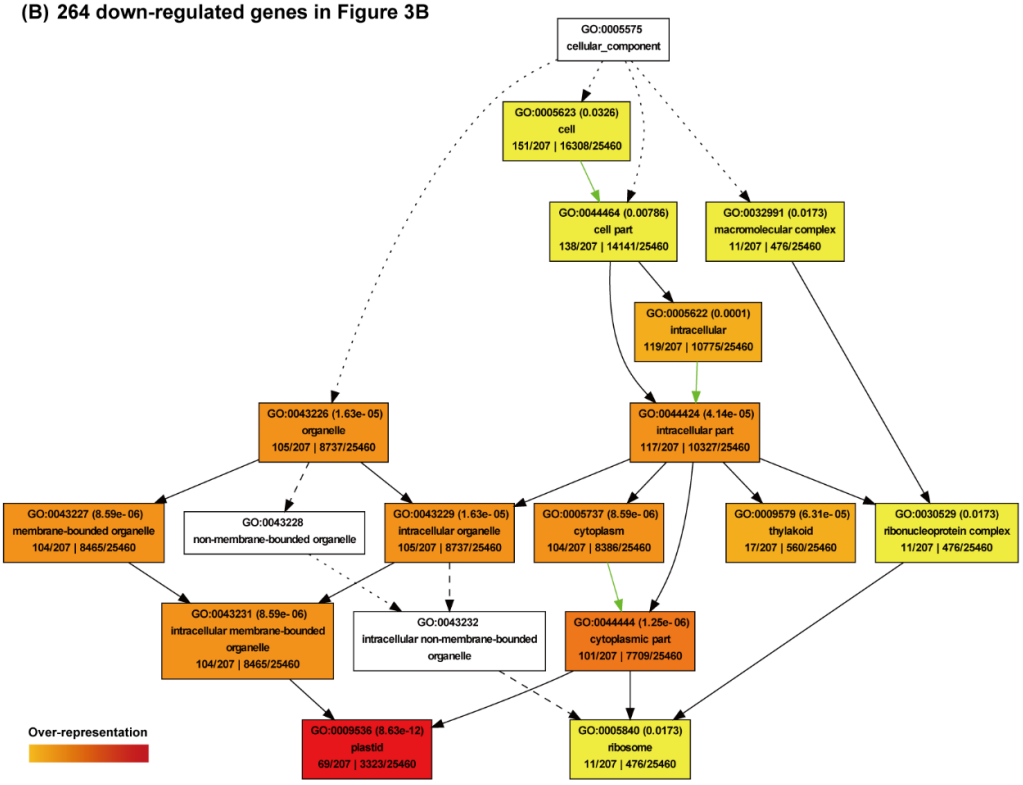
Zhong, S., Joung, J.G., Zheng, Y., Chen, Y.R., Liu, B., Shao, Y., Xiang, J.Z., Fei, Z., and Giovannoni, J.J. (2011). High-throughput illumina strand-specific RNA sequencing library preparation. Cold Spring Harbor protocols *2011*, 940-949.

SUPPLEMENTARY FIGURES AND LEGENDS

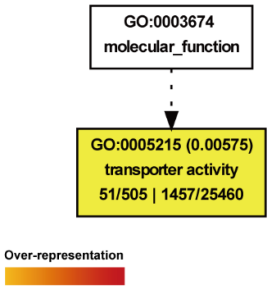
(A)



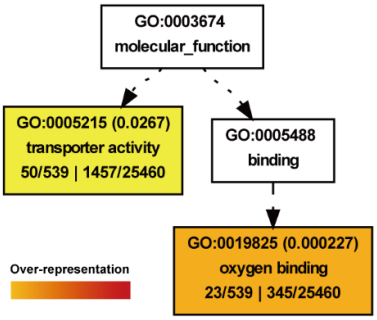
(B) 264 down-regulated genes in Figure 3B



(C) Up-regulated genes in Leaf Stv-5D

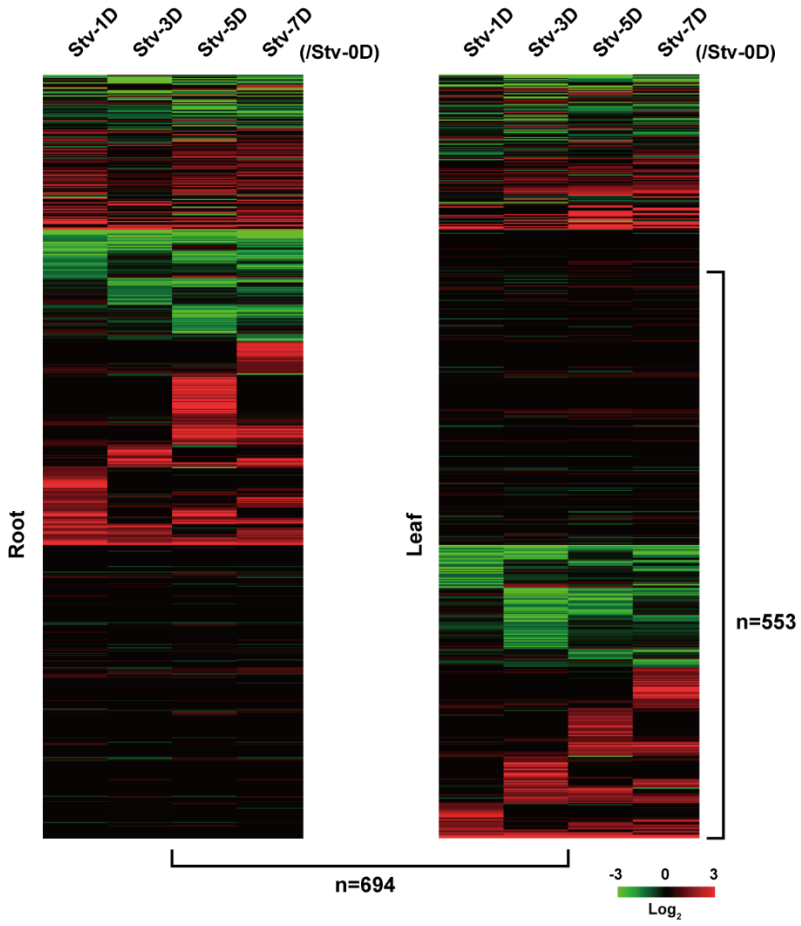


(D) Up-regulated genes in Leaf Stv-7D

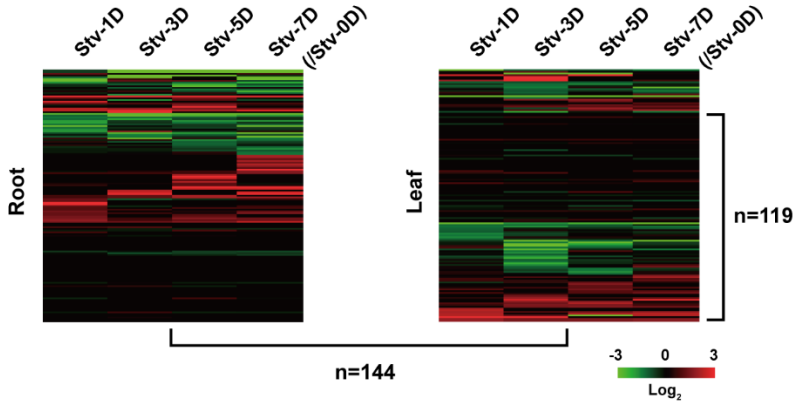


Supplementary figure 1. Additional analysis results of differentially expressed genes in leaves and roots. (A) Number of genes differentially expressed between two adjacent N-starved roots and leaves. (B) AgriGO representation of the overrepresented GO terms in the 264 down-regulated genes in Figure 3B was generated using singular analysis enrichment (Fisher's test, $P < 0.05$, $FDR < 0.05$). Numbers in parenthesis represents the FDR value. (C) AgriGO representation of the overrepresented GO terms in the up-regulated genes in Stv-5D leaves and (D) in Stv-7D leaves.

(A) N-responsive intergenic lncRNA loci



(B) N-responsive antisense lncRNA loci



Supplementary figure 2. Expression profiling of N-responsive putative lncRNAs. (A) Expression patterns of 694 N-responsive putative intergenic lncRNAs in N-starved roots and leaves. (B) Expression patterns of 144 N-responsive putative antisense lncRNAs in N-starved roots and leaves.

ABSTRACT IN KOREAN

전사체 분석에 기반한 질소 공급 조건 변화에 따른 벼 유전자 및 Non-coding RNA의 변화 관찰

질소는 식물의 생장 · 발달에 필수적인 영양요소이며, 작물 생산량을 결정하는 중요 요소로서 작물 생산량 증진을 위한 질소 비료의 사용이 증가하고 있다. 그러나 작물의 낮은 질소 이용 효율로 인한 질소 비료 사용 효율성의 저하 및 질소 비료의 용출로 인한 수질 악화 등의 문제로 인하여, 작물의 질소 이용 효율 관련 연구의 필요성 및 연구 사례가 증가하고 있다.

Non-coding RNA는 단백질을 암호화하지 않는 RNA 전사체로, 최근 다양한 연구를 통해 miRNA 및 long non-coding RNA 등의 non-coding RNA가 표적 유전자의 상위 조절 인자로서, 다양한 메커니즘을 통해 동 · 식물 내 중요 현상에 관여하는 사례가 다수 보고되고 있다. 특히 식물의 non-coding RNA들은 영양 결핍 스트레스를 포함한 다양한 스트레스 환경에 대한 반응 및 대응 기작에 관여하는 것으로 알

려지고 있다. 그러나 질소 공급 조건에 대한 벼 non-coding RNA 발현 및 역할에 대한 transcriptome-wide한 연구는 미진한 상태이다.

본 연구는 질소 공급 조건의 변화에 따른 벼의 전사체 수준의 반응을 종합적으로 관찰 및 이해하고자 하였으며, 이를 위하여 다양한 형태의 RNA-Seq 기법이 사용되었다. Strand-specific RNA-Seq을 이용하여 유전자의 발현 수준 확인 및 분석을 수행하였으며, putative long non-coding RNA deriving loci 발굴을 수행하였다. 또한 small RNA-Seq 분석을 통하여 miRNA 발현 수준 확인을 통해 질소 공급 조건 의존적으로 변하는 miRNA를 확인하였으며, 이들 중 miR169에 대해 타겟 유전자와의 발현 패턴 분석 및 5' RLM-RACE를 통한 타겟 검증을 수행하였다. 마지막으로, 2P-Seq 데이터 분석 수행을 통해 strand-specific RNA-Seq을 통해 발굴된 putative long non-coding RNA 후보 중 한 transcript의 3' end에 signal을 확인하였으며, 3' RACE를 수행하여 데이터의 유효성을 검증하였다. 본 연구를 통해 분석 및 확보한 데이터는 벼의 질소 공급 상태 변화에 따른 전사체 수준의 변화에 대한 이해, 그리고 유전자와 non-coding RNA의 역할에 대한 추가적인 연구를 진행하는 데에 있어 기초 자료로 활용될 것으로 기대한다.