Bose Institute

Understanding the regulation of a transcription factor *OsDREB2A* induced by salt stress in *Oryza sativa L.*

Ruby Biswas

Under the Supervision of Dr. Shubho Chaudhuri
(Associate Professor, Division of Plant Biology, Bose Institute)

Roll No.: BICU/2016-18/LSC/011

Regd. No.: 223-1221-0086-13

Integrated M.Sc.-Ph.D Program

Special Paper: Plant Molecular Biology and Biotechnology

Dissertation Work

THESIS CERTIFICATE

This is to certify that the work entitled, "**Understanding the regulation of a transcription factor *OsDREB2A* induced by salt stress in *Oryza sativa L.***" submitted by **Ruby Biswas** to the University of Calcutta and Bose Institute for the partial fulfillment of the degree of Master's in Science in Life Sciences is a bona fide record of the project work done by her under my supervision in the duration of five months ending July 16,2018. She has completed the work truthfully and successfully to the best of my knowledge. The content of this thesis, in full or in part has not been submitted to any other Institute or University for the award of any other degree or Diploma.

Dr. Shubho Chaudhuri

Associate Professor

Division of Plant Biology

Bose Institute, Kolkata

ACKNOWLEDGEMENT

It is a genuine pleasure to show my deepest gratitude to my Project Instructor **Dr. Shubho Chaudhuri**, **Associate Professor, Division of Plant Biology, Bose Institute** for providing me the opportunity to work on such an interesting subject for my dissertation. His keen interest and timely advice on the scientific approaches have helped me to successfully accomplish this task.

I would extend my sincere gratitude to the **Dean of Studies and Controller of Examinations** for allowing me to undertake this project work as a part of my Masters in Life Science curriculum. I'm highly indebted to **Prof. Debabrata Basu, the Head of Department and all the faculties of Division of Plant Biology, Bose Institute** for providing me with the necessary facilities instrumental for the implementation of this project.

I am very grateful to **Amit Paul** for his constance guidance and support in the pursue of this project. A debt of gratitude is also owed to **Sayani Dey** for her expertise.

I would also like to thank my lab members **Rwitie Mallik, Pratiti Dasgupta, Jinia Chakraborty** and **Sonal Sachdev** for their valuable advice and encouragement in the execution of this work. I also thank my lab assistant **Mr. Nadiram Kayal** for his ungrudging help.

I express my gratitude to **Mr. Mrinal Das** of the Central Instrument Facility, Bose Institute for his assistance.

Last but not the least I am grateful to my parents, sister and friends for the favors I have received from them in this venture.

RUBY BISWAS
15.07.2018

CONTENTS

Topic **Page Number**

SUMMARY....1

INTRODUCTION....2-4

HYPOTHESIS & OBJECTIVE....5

MATERIALS AND METHODS

- ❖ Databases used for small RNA prediction 6
- ❖ Plant Material and Growth Conditions 6
- ❖ Stress induction conditions 7
- ❖ RNA isolation and cDNA preparation....7
- ❖ Transcript expression analysis using semi quantitative PCR....7
- ❖ RLM-5'RACE 7-15
- ❖ Primer Sequences....16

RESULTS

- ❖ Study of *OsDREB2A* locus using *in silico* approach....17
- ❖ Expression analysis of *OsDREB2A* transcripts and the overlapping anti-sense genes under salt stress conditions....18-22
- ❖ Determination of the cleavage site in *OsDREB2A* transcript 2 of IR-64 by RLM-5'RACE....23

DISCUSSION....24

CONCLUSION....25

REFERENCES....26

SUMMARY

OsDREB2A is a DRE (Dehydration response element)/CRT (C-repeat) sequence binding transcription factor, induced in response to temperature, drought, salinity, etc. Here we report that this gene has two transcripts with identical coding sequences. The expression analysis has shown that the transcript 2 is early salinity stress responsive, whereas transcript 1 is late salinity stress responsive. The 3'UTR region of *OsDREB2A* transcript 2 is overlapped by another gene in the anti-sense strand. This overlapping structure of genes might produce double stranded RNA which, acts as a substrate for endogenous siRNA synthesis. The current study is aimed to investigate the biogenesis and role of endogenous siRNA in the post-transcriptional gene silencing of *OsDREB2A* transcript 2 gene using RLM-5'RACE.

Rice has been studied extensively by molecular genetics and constitutes one of the best characterized crop plants with a fine genetic map(19). The complexity of rice genome increases as we try to elucidate its genome sequence where despite of the presence of large intergenic regions it also contains opposite or anti-sense overlapping genes. Although there have been validated gene overlaps documented, for example in 4%-9% of human genes, 22% of *Drosophila melanogaster genes* and 10% of *Arabidopsis thaliana genes* (1,2,3). Even though the functional significance of these overlapping genes remains unclear, one probable explanation could be the double stranded RNA structures formed from the overlapping transcripts which maybe processed into small RNAs could serve as important regulatory molecules.

These 21-25nt non coding small RNAs depending on their origin have different names(4). Those made artificially or produced *in vivo* from dsRNA precursors are typically called small interfering RNAs (siRNAs) whereas those which are derived from precursor RNAs encoded by miRNA genes expressed in cells are called the micro RNAs (miRNAs) (1,5,6). Both siRNAs and miRNAs are generated from longer RNA molecules by the enzyme Dicer, an RNase-III-like enzyme that recognizes and digests longer dsRNA formed by siRNA precursors or stem loop structures formed by miRNA precursors(8). These small RNAs inhibit expression of homologous target genes in three ways:

i. They trigger destruction of mRNA encoded by the target gene
ii. They inhibit translation of the mRNA
iii. The induce chromatin modification within the target gene and thereby silence its transcription.

Till date, two related machineries have been identified that incorporate siRNA: RISC and RITS. In the RNA Induced Silencing Complex (RISC), siRNAs recognize the target mRNAs and initiate their degradation by endonucleolytic cleavage within the mRNA

region that base paired to siRNA (21). The RNase H domain of Argonaute/PIWI family protein (a sub-unit of RISC) carries out this initial mRNA cleavage event. This is known as the Post Transcriptional Gene Silencing (PTGS) method of siRNA(9). In the nuclear RNA Induced Transcriptional Silencing (RITS) complex, siRNAs target the chromosome regions for chromatin modification, or is the Transcriptional Gene Silencing (TGS) method of gene silencing(11). TGS recruits chromatin modifying enzymes one such example is that of RNA dependent DNA methylation (RdDM) which blocks transcription of the gene itself (7). This involves modification of the promoter region so that further transcription does not take place.

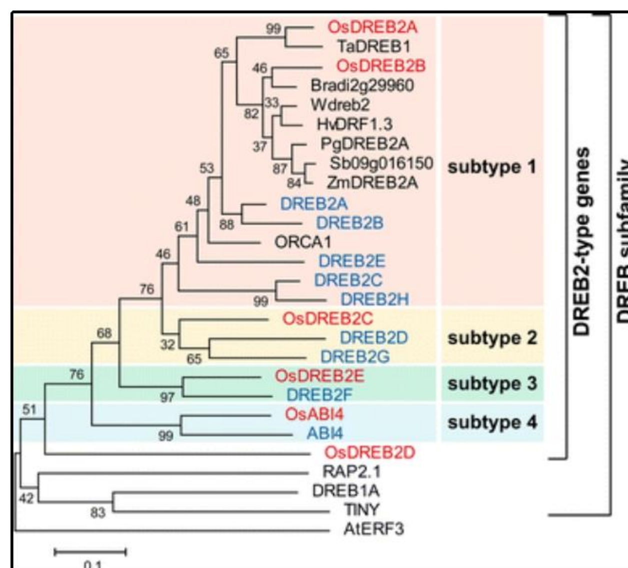


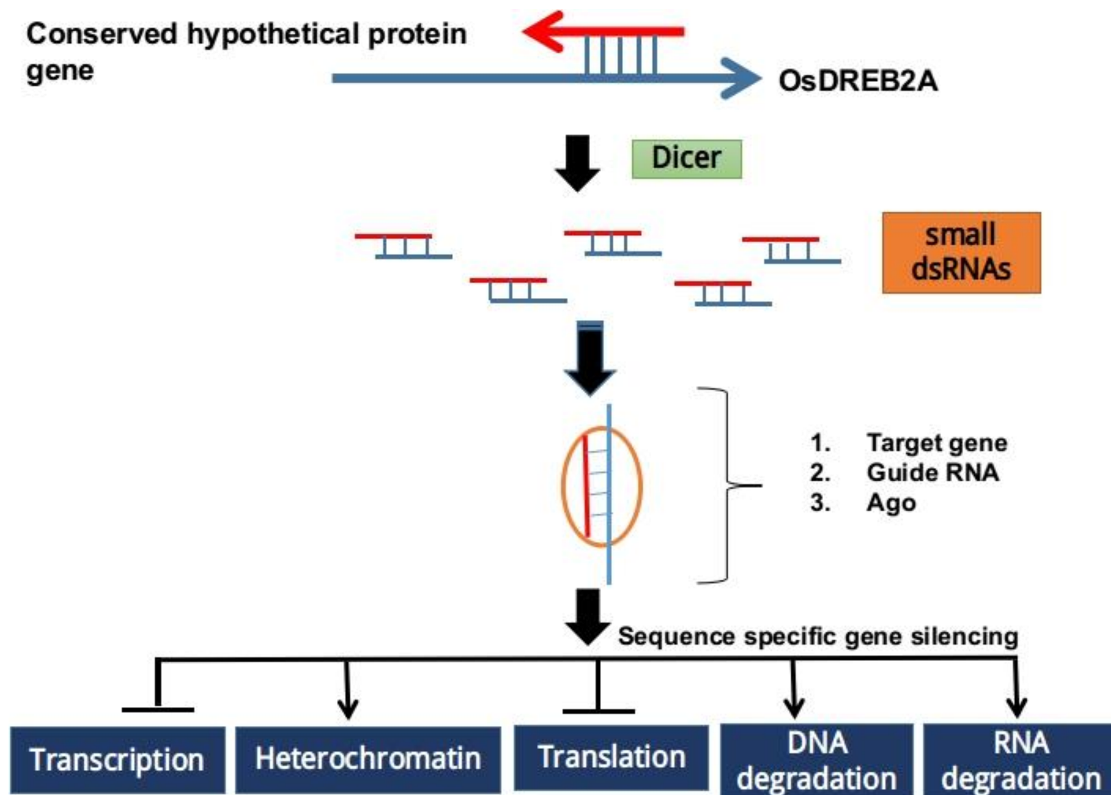
Fig 1: Phylogenetic analysis of the OsDREB2s and related proteins

In this study we report that in chromosome 1 of rice there is an anti-sense gene overlapping *OsDREB2A,* an important transcription factor induced in response to osmotic and high temperature stresses. *DREB2*s (Dehydration Responsive Element Binding protein 2s) play a vital role in plant development and abiotic stress response. It interacts with a *cis* acting DRE (dehydration responsive element) / CRT (C-repeat) sequence and activate the downstream genes responsible to combat osmotic imbalance and high temperature stresses in *Arabidopsis thaliana* (10,12,13,15). *OsDREB2A* an ortholog of *DREB2A* in rice is and is one of the six

DREB2 orthologs in rice (16). Transgenics with over-expressed *OsDREB2A* are better at surviving salinity stress than the wild types(17).

This report was aimed to determine the presence of siRNA produced endogenously from anti-sense overlapping gene pair of *OsDREB2A* and conserved hypothetical protein gene, which probably plays a functional role in regulating the expression of *OsDREB2A*. The presence of siRNA was predicted using bioinformatic approach and siRNA mediated *OsDREB2A* mRNA cleavage was ruled out after significant investigation.

The coding region of *OsDREB2A* transcript 2 is overlapped by another gene which, codes for a conserved hypothetical protein. This overlapping arrangement of genes might lead to the generation of small double stranded RNAs which, may regulate the expression of *OsDREB2A* gene at the transcriptional or translational level.



Proposed Hypothetical model

OBJECTIVES:

1. *In silico* determination of small RNAs in the *OsDREB2A* locus
2. Study of the expression pattern of the two trancripts of OsDREB2A and the overlapping antisense gene
3. Validation of the regulatory role of the sRNA by 5'RLM RACE.

1. **Databases used for smRNA prediction:**

RAP-DB (Rice Annotation Project Database): The gene name(*OsDREB2A*) was used to view the gene locus.

CLUSTALW : It is a tool to align multiple sequences. This was used to align the reverse complementary sequence of *OsDREB2A* sequence with the overlapping gene in the anti-sense strand.

CSRDB (Cereal small RNA Database): The MSU Id of the transcript (LOC_Os01g07120.2) was fed to obtain the smRNA populations, their sequence and targeting locus of the gene.

psRNA Target (plant small RNA Target) :The entire sequence of *OsDREB2A* was submitted and the probable small RNAs from that sequence which had been reported to the database was obtained. This server was also used to check the target location of the small RNAs predicted by CSRDB.

2. **Plant material and growth conditions:**

Seeds of Oryza sativa L. cv. IR-29 and IR-64 were surface sterilized with 0.1% (w/v) $HgCl_2$ for 15 minutes, washed several times with sterile water and allowed to germinate over water-soaked sterile gauge in trays at 37°C in dark for 3 days. The germinated seedlings were grown in water-soaked sterile gauge in trays in presence of 0.25X Murashige and Skoog (MS) complete media at 30 °C under 16 hours light and 8 hours dark photoperiodic cycle with 50% relative humidity and 700 l mol photons m22 s 21 for the desired period in a plant growth chamber.

3. **Stress induction conditions:**

For salt-stress treatment, the 13 days-old seedlings were treated with 250mM NaCl. And tissues were collected for 2hours, 4 hours and 6 hours along with the control samples for both IR-29 and IR-64 rice varieties.

4. **RNA isolation and cDNA preparation:**

RNA samples were isolated from 13 days old rice seedlings of IR-29 and IR-64 varieties using Trizol reagent (Invitrogen) as described by manufacturer protocol followed by DNase1(Invitrogen) digestion. 1 microgram of the total RNA was separated on 1.5% agarose gel in TAE with 0.5μg/ml ethidium bromide added to it . 5μg of total RNA from control and stress treated samples were used for first strand cDNA synthesis by using revertaid reverse transcriptase (ThermoScientific) and oligo dT (ThermoScientific) (18).

5. **Transcript expression analysis using semi quantitative PCR:**

The cDNA obtained for the two varieties and their stress points was first diluted in 1:10 ratio and amplified using *OsDREB2A* specific primers (10microM). Actin gene (Act1, LOC4333919, Os03g0718100) was used as endogenous reference gene. The DNA was extracted and run on 2.5% agarose gel.

6. **Transcript expression analysis using Real Time PCR:**

Real-time PCR was performed using cDNA obtained from the control and treated plants in 20 μl reaction volume containing 1× buffer with SYBR Green (DyNAmo ColorFlash SYBR Green, Thermo Scientific) and 0.25 μM primers. Actin gene (Act1, LOC4333919, Os03g0718100) was used as endogenous reference gene. Average Ct values obtained for the target genes were normalised by the Ct values for Actin (ΔCt). The data were expressed using comparative Ct -(2 −ΔΔCt)

method to obtain relative fold change and comparative Ct -($2^{-\Delta Ct}$) to obtain fold change in control condition between IR-29 and IR-64.

7. RLM-5' RACE:

Background of the experiment: Rapid amplification of cDNA ends (RACE) is a polymerase chain reaction-based technique which facilitates the cloning of full-length cDNA sequences when only a partial cDNA sequence is available. Traditionally, cDNA sequence is obtained from clones isolated from plasmid or phage libraries. Frequently these clones lack sequences corresponding to the 5' ends of the mRNA transcripts. The missing sequence information is typically sought by repeatedly screening the cDNA library in an effort to obtain clones that extended further towards the 5' end of the message. The nature of the enzymatic reactions employed to produce cDNA libraries limits the probability of retrieving extreme 5' sequence even from libraries that are very high quality.

Classic 5'RACE: Classic 5' RACE protocols vary slightly in design, but are essentially equivalent. First strand cDNA is synthesized from either total or poly(A) RNA in a reverse transcription reaction. A defined sequence is then added to the 3' end of the first strand cDNA by tailing with terminal deoxytransferase (TdT), or by ligation of an oligonucleotide adapter. Finally, a gene specific primer is used in conjunction with a primer for the added 3' sequence to amplify the sequence between the adapter and the gene specific primer at the 5' end of the cDNA. Traditional 5' RACE is sometimes successful, but the major limitation of the procedure is that there is no selection for amplification of fragments corresponding to the actual 5' ends of mRNA: all cDNAs are acceptable templates in the reaction. Additionally, the PCR step selects the most efficient amplicons (e.g., the smallest), favoring amplification of less than full-length products. 5' RACE usually produces a heterogeneous population of amplified products.

RLM RACE: RNA Ligase Mediated Rapid Amplification of cDNA Ends (RLM-RACE) represents a major improvement to the classic RACE technique (20). RLM-RACE is designed to amplify cDNA only from full-length, capped mRNA, usually producing a single band after PCR. The procedure is shown schematically in Figure 2.

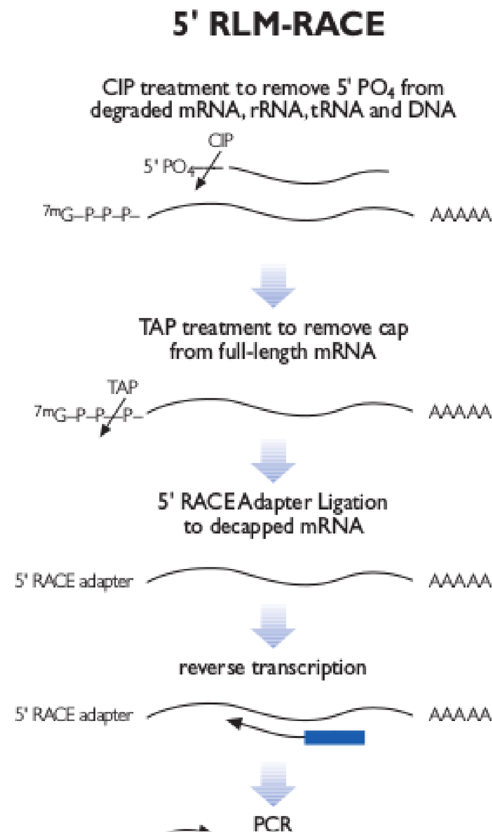


Fig 2: Overview of RLM-5' RACE

Protocol:

A modified RNA ligase mediated rapid amplification of cDNA ends was performed using GeneRacer Kit(Invitrogen).

RNA isolation

Total RNA was extracted from 13 day old seedling of IR-64 using TRIZOL method (Invitrogen).

RNA oligo ligation

Sl no.	Reagent	Control (µL)	6 hrs salt stressed sample (µL)	8 hrs salt stressed sample (µL)
1.	RNA (5µg)	5.160	4.794	9.124
2.	RNA oligo	1	1	1
3.	10X T4 RNA ligase buffer	2	2	2
4.	10mM ATP	2	2	2
5.	50% PEG8000	3	3	3
6.	Ribolock RNase inhibitor	1	1	1
7.	T4 RNA ligase	1	1	1
8.	Nuclease free H_2O	4.84	5.206	0.876

The 20µl reaction mix was incubated at 16°C for 16hrs.

RNA Precipitation:

i. 80µl of RNase free water was added to make the final volume upto 100µl.

ii. To it equal amounts of chloroform and trizol was added and centrifuged at 12000g for 10 mins at 4°C.

iii. The aqueous phase was collected and to it 1/3rd volume of 3M sodium acetate of pH5.2, 3 times the volume of the aqueous phase collected absolute ethanol and 1µl glycogen(20µg/ml) was added.

iv. To pellet the RNA centrifugation was done at 12000g for 10min at 4°C.

v. The supernatant was discarded and to the pellet 75% ethanol was added and inverted several times.

vi. This was centrifuged at 7500g for 5mins after which the supernatant was discarded and the pellet was dried.

vii. The pellet(RNA) was then dissolved in 10µl of RNase free water.

cDNA preparation of the RNA oligo ligated RNA:

The total RNA precipitated was used for cDNA preparation using revertaid reverse transcriptase (ThermoScientific) and oligo dT (ThermoScientific) (18).

The cDNA obtained was then checked with actin specific primers, a constitutive gene for its quality.

Performing touchdown PCR to amplify the 5'ends:

The following table was used to set up the sample reactions:

Sl no.	Reagents	Control sample (µL)	6hrs salt stressed sample (µL)	8hrs salt stressed sample (µL)
1.	Template	1	1	1
2.	dNTP	0.5	0.5	0.5
3.	5 outer primer (10µM)	3	3	3
4.	Gene specific reverse primer (10µM)	1	1	1
5.	Taq.pol	0.5	0.5	0.5
6.	10X buffer	5	5	5
7.	Nuclease free H_2O	37.5	37.5	37.5
8.	$MgCl_2$	1	1	1
TOTAL VOLUME :		**50**	**50**	**50**

PCR cycle:

Sl.no.	Temperature (°C)	Time	Cycles
1.	94	2min	
2.	94	30sec	5
	68	30sec	
3.	94	30sec	5
	66	30sec	
4.	94	30sec	5
	64	30sec	
5.	94	30sec	5
	62	30sec	
6.	94	30sec	20
	60	30sec	
7.	72	10min	
8.	4	∞	

Performing Nested PCR:

1μL of the original amplification reaction was used as template for nested PCR and the reaction was set up using the following table:

Sl no.	Reagent	5'control PCR 1 (µL)	5'control PCR 2 (µL)	5'control PCR 3 (µL)	5'RACE PCR (µL)
1.	Template	-	1	1	1
2.	5' Nested primer (10µM)	0.4	0.4	-	0.4
3.	3' Nested gene specific primer (10µM)	0.4	-	0.4	0.4
4.	10X PCR buffer	2	2	2	2
5.	dNTP (10mM)	0.4	0.4	0.4	0.4
6.	Taq pol	0.2	0.2	0.2	0.2
7.	Nuclease free water	16.6	16	16	15.6
	TOTAL VOLUME	**20**	**20**	**20**	**20**

PCR cycle:

Sl no.	Temperature	Time	Cycle
1.	94°C	2 min	
2.	94°C	30 sec	35 cycles
2.	62°C	30 sec	
3.	72°C	30 sec	
4.	72°C	7 min	
6.	4°C	hold	

Gel Elution and TA ligation:

The product amplified from the nested PCR was run in a 2.5% gel from which the unique bands were cut out and eluted following the manufacturers protocol for gel extraction by Thermo scientific. (13)

The eluted DNA fragment was then ligated to a TA vector, pMD20.

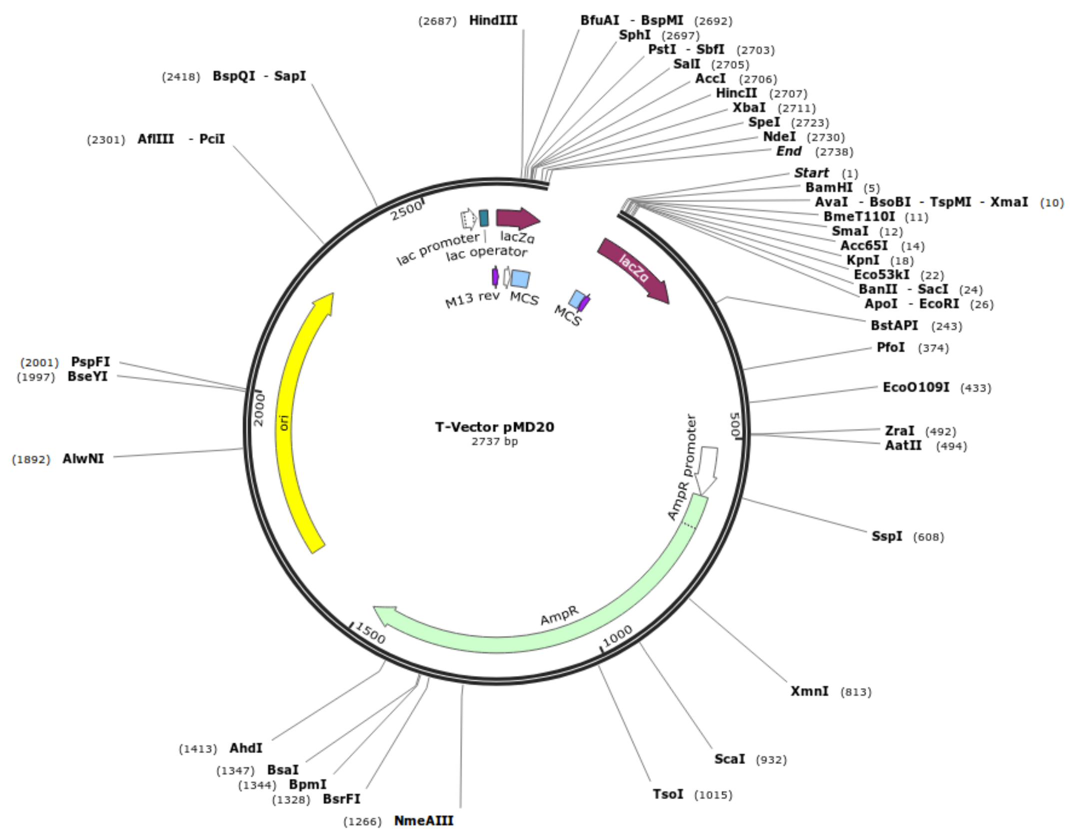


Fig 3: pMD20 vector map

Cloning and Transformation of competent cells:

The TA ligated vectors were then cloned into DH-5α competent E.coli cells which were then transformed by heat shock method.

Selection of Transformants:

The transformed cells were selected in presence of ampicilin and blue white screening was done to select the transformants. The transformants were then cultured overnight in 2ml of Luria Broth at 37°C.

Plasmid isolation and confirmation of insert by restriction digestion:

Plasmids were isolated using the manufacturers protocol for plasmid isolation by Thermo Scientific.(14) 300ng of the isolated plasmid was digested using the restriction enzymes BamHI and HindIII. The digested product was then run in a 2.5% gel to confirm whether the insert was actually inserted within the plasmid or not.

Sequencing:

The plasmids isolated from the positive clones were sequenced using M13 forward and M13 reverse primers.

Sequencing PCR cycle:

<table>
<tr><th>Sl no.</th><th>Temperature</th><th>Time</th><th>Cycles</th></tr>
<tr><td>1.</td><td>96°C</td><td>2 mins</td><td></td></tr>
<tr><td>2.</td><td>96°C</td><td>30 sec</td><td rowspan="3">25 cycles</td></tr>
<tr><td>3.</td><td>55°C</td><td>15 sec</td></tr>
<tr><td>4.</td><td>60°C</td><td>1 min</td></tr>
<tr><td>5.</td><td>4°C</td><td>hold</td><td></td></tr>
</table>

PRIMER SEQUENCE:

Sl no.	Primer	Sequence	Tm	GC content
1.	*OsDREB2A* transcript1 F.	5' GTG GCA CAA CTC GTG TGT TGA 3'	58.2 °C	52.40%
2.	*OsDREB2A* transcript1 R.	5' CCT CTC TAA CAG CCG CC 3'	55.3°C	64%
3.	*OsDREB2A* transcript2 F.	5' GAT GAA TCC ACT GCG AGC 3'	53.6°C	55.60%
4.	*OsDREB2A* transcript2 R.	5' GGG AGC AAA CCT TGA GGA GC 3'	58.5°C	60%
5.	Antisense exon1 F.	5' CCG TTC TCC CAA TGG CGC 3'	59.6ºC	67%
6.	Antisense exon 1 R.	5' GAT CTC AGC ACC CAG TCC AC 3'	57.6°C	60%
7.	Antisense exon 2 F.	5' GGT GGA CAA CAT AAG GTG TGA TGT GG 3'	59.6°C	50%
8.	Antisense exon 2 R.	5' GCT GAC CGA TGC CTT TAG TCC 3'	58°C	57.10%
9.	Outer Gene Specific Primer 1	5' GAT GGT GTA AAC GGA AGG GAG G 3'	68°C	54.50%
10.	Nested Gene Specific Primer 1	5' GGT TTC GTG ACA GGT ACT CG 3'	62°C	55%
11.	Outer Gene Specific Primer 2	5'- GAG GCT CCC CAT GCT CTT CG 3'	66°C	65%
12.	Nested Gene Specific Primer 2	5' GCT GAC TGG GCC AAA GTG C 3'	62°C	63.15%
13.	RNA Oligo outer primer	5 'CGA CTG GAG CAC GAG GAC ACT GA3'	74°C	60.8%
14.	RNA Oligo nested primer	5' GGA CAC TGA CAT GGA CTG AAG GAG TA 3'	78°C	50%
15.	M13 Forward	5' CGC CAG GGT TTT CCC AGT CAC GAC 3'	67.8°C	71.42%
16.	M13 Reverse	5' TCA CAC AGG AAA CAG CTA TGA C 3'	58.4°C	45.45%
17	RNA oligo adapter	5' CGA CUG GAG CAC GAG GAC ACU GAC AUG GAC UGA AGG AGU AGA AA 3'		

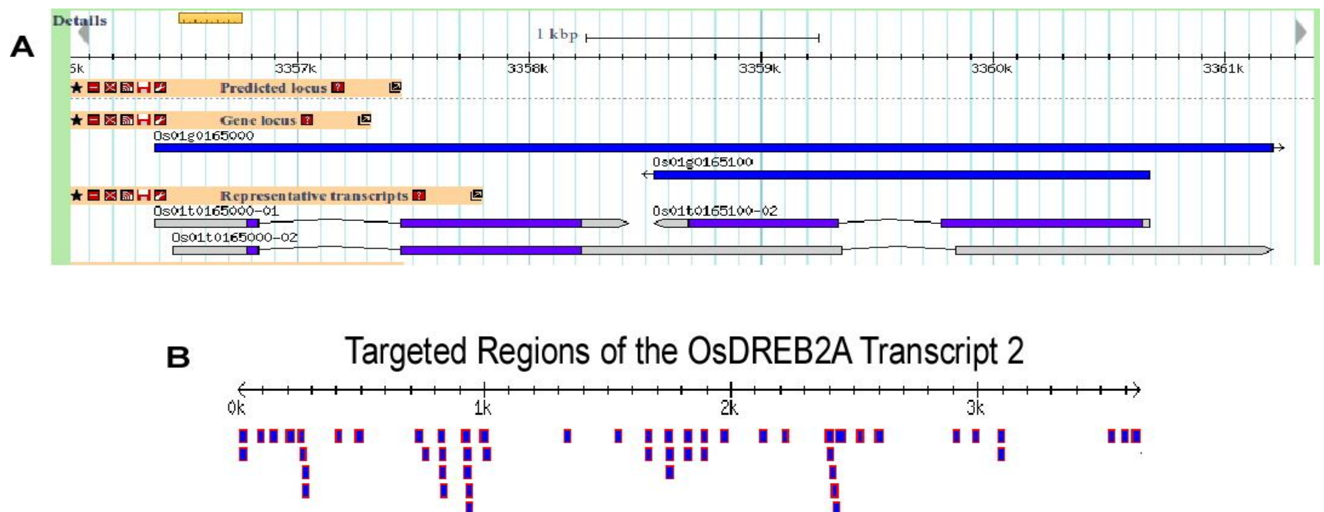


Fig 4: Bioinformatic analysis of the *OsDREB2A* locus

A. The genome browser view of the *OsDREB2A* locus using RAP-DB

B. smRNA targeting the *OsDREB2A* transcript 2 map taken from CSRDB. The transcript does not contain any intronic regions.

Study of *OsDREB2A* locus using *in silico* approach

The sequence information of *OsDREB2A* and the Conserved Hypothetical Protein genes obtained from RAP-DB (Fig4 A) shows that the there are two transcripts for the *OsDREB2A* gene (Os01g165000) and their CDS are identical and the 3'UTR of the transcript 2 (Os01t165000-02) was longer than the transcript 1 (Os01t165000-02). The 5'UTR of the transcript 1 was longer than transcript 2 by 78nt. The map also shows the presence of another gene (Os01g0165100) present in the anti-sense strand of the transcript2. The sequence similarity between the overlapping gene in the anti-sense strand and the reverse complementary sequence of transcript 2 using CLUSTALW showed 100% alignment. The search for small RNA was carried out in the CSRDB and psRNA Target databases and the list of small RNAs targeting the *OsDREB2A* transcript 2 are listed in the Table.

	TABLE FOR THE SMALL RNAs MAPPED FROM LEFT TO RIGHT				
Locus details	OsDREB2A sequence	small RNA sequence	Length(nt)	Alignment Score	Free Energy (kcal/mol)
	5 GCACAACUCGUGU-GUUGAC-CCGCCUCGCGC 3	3 UGUUGGGCUGAGCAACUGCCGCGGAGC 5	27	21	-36.9
	5 CAGCAGCCGCCCCCGCCCGCGGUUGCU 5	3 UGGUUGGU-UUACGAGGAUACU 5	21	18	-29
	5 AGCUGCAGCGGCGGCGGCGGCUGUUAG 3	3 ACGUUGUUGCCGUCACCGACC 5	21	21	-34.5
	5 ACCGACACCGACACGCUCGCCAUGCCA 3	3 AGGUGGUUAUGCGAGUGGUAC 5	21	23.5	-31.3
	5 ACCGACACGCUCGCCAUGCCACCAAGC 3	3 GUGUGUGAGUGGUACAGUGGU 5	21	27.5	-32.6
	5 CGCCAUGCCACCAAGCGAGGCGGCGGC 3	3 AUUGUGGUGGUUCG-UUCGCCU 5	21	16	-31.5
	5 CUCGAUGGAGCGGGGGGAGGGGAGGAG 3	3 CCUUCGCCCCACUCUCC-CCUC 5	21	20	-35.4
	5 CUGAUUCA-AUCGCUGAAACCAUCAAGU 3	3 AGUUGUGGUAACUUUGGUAGU 5	21	15	-25.8
not in complementary region	5 CGCUGGAGGCUGCGCAUGCAUACGAUG 3	3 AUCUUUCUACGCGUACGUACA 5	21	18	-27.1
	5 GGCGGCAAGGGCAAUGUAUGGUCCCAC 3	3 CCGCCGUU---GUUA-AUACCAGGU 5	21	16	-34.7
	5 GCCAACUCUGG-CUGCA-CAUCAGCACCU 3	3 GAGACCGGAUGUAGUAGUCGC 5	21	18.5	-28.7
	5 CUCUGGCUGCACAUC-AGCACCUUCAUU 3	3 GUCGAGUAGUUUGUGGAAGUC 5	21	18	-28.3
	5 GCACAUCAGCACCU-UCAUUGAUGAUGU 3	3 GUAGUCGUGGGUAAUAAUAAC 5	21	15	-23.2
	5 CAUCGUGGCUAAUGGG-CCAGCUGUGUUGU 3	3 UACUGUUUACCCGGGUAGACACA 5	23	23	-31.5
	5 GCUAAUGGGCCAGCUGUGUUGUAUCAG 3	3 CCGUCCGGUAUACACAACAUC 5	21	17	-26.8
	5 AGCUGUGUUGUA-UCAGCCUGAUAAGAA 3	3 AGCAGCAUGAGUCGGACUAGC 5	21	17	-27.6
	5 GCAGGAUGUUAAAACAGAAGGGAGCAA 3	3 CGUCCUAUUGUUUCUUCCCUU 5	21	18	-31.2
	5 GUUUUCGUCUUACUGGGUAAUAGCUUU 3	3 AAGCAGAAAUGUCCGUUAUCA 5	21	17	-24.7
	5 UAAAUUCUCAAACAGGUAUAUUAGGAA 3	3 GAAGGGUUUGUUC--AUGAUCCU 5	21	15	-28.6
	5 AUAGCGCUCUAUGGUCAUUCAAAGUAC 3	3 CGUGAGAUUAAAG-AAGUUUCA 5	21	17	-24.1
	5 CAAGCCUCCU-UGUUUCACUGUGAAAGG 3	3 AGGAGGCAAAGUGUUACAUUC 5	21	17	-25.7
	5 UCCUUGUUUCACUGUGAAAGGUCUUCUGUU 3	3 GGCAAAGUGUUACAUUCAGUAAGA 5	24	17	-23.5
	5 UCCUUGUUUCACUGUGAAAGGUCUUCUGUU 3	3 GGCAAAGUGUUACAUUCCGAAAGA 5	24	22	-24.3
	5 UCCAUGGUCUUCCCUA-AUGCUGCACCA 3	3 AUUGGAAGGGAUGUACGAUGU 5	21	20	-27.1
	5 ACGUGCGAAAAUGACACCACCCUUCUU 3	3 CCUAUUUUACUGUGGUAAGAG 5	21	19.5	-24.4
	5 ACGUGCGAAAAUGACACCACCCUUCUU 3	3 ACGUUGACUGUGAUGGG-AGAU 5	21	17	-28.3
	5 CAGGGAGAGCUUGGAAGGCUUUGAAGA 3	3 CUCCCUCAACUUCCGUAACUU 5	21	20	-28.9
	5 AAAUGUGUUUG--GGGGGUCGAAUGAA 3	3 ACGCAAACGGCCCUCAGCU 5	19	14	-28.6
	5 AAACCGACAGG-ACAUCUGAUUGUUGCA 3	3 GUUGUCCGUGUAGCUUAACAU 5	21	19.5	-24.7
	5 GUAAGCGAGAAGAGAAGGCGCUCAUCC 3	3 AUGUUUGAUUCCGCGAGUAUA 5	21	17	-25.9
	5 AGCGAGAAGAGAAGGCGCUCAUCCUGU 3	3 UUUUUUUCCCGUGAAUAGGAU 5	21	19	-26.8
	5 CGCUCAUCCUGUGUUGGCACUGCAGGA 3	3 AAGUAGGAAACAA-CGUGACAC 5	21	20	-27.5
All in the same locus	5 CAUCCUGUGUUGGCACUGCAGGAUCUU 3	3 AGAUACAAGUGUGACGUUCAC 5	21	16	-25.4
	5 GUUGGCACUGCAGGAUCUUCCAUGCUC 3	3 CUCGUGACGU-CAAGGAGGUAG 5	21	18	-32.2
	5 CUUCCAUGCUCAUGGCAAUGAGUGGGCAGU 3	3 GGUGUGAAUACCGUUUC-AAUCCGU 5	24	17	-32.2
	5 GCGAUGGCACCCC-AAAUGAAUGGUGUG 3	3 ACCGUGGGGCGUUACUUAGUA 5	21	19	-28.4
	5 GCAUCCAUUGCUCCUCAAGGUUUGCUC 3	3 AGGUGAUGA-G-GUUACAAACGU 5	21	15.5	-27.4
	5 UAUGAUUUCAGGUGGACUGGUUUUGAC 3	3 GGAGUUUCCAUGGCCAAAACU 5	21	18	-25.6
	5 GGAGGGCUGGGUGGGAGGGGUUUUGAC 3	3 ACCGACCCUGCCUUCCUAAGC 5	21	15	-30.5
	5 GCGCCAUUGGGAGAACGGAGGCAG 3	3 GGUAACCC-GUUGCCUCAG 5	18	19	-26
	5 UGGUUGGGUGGAAUGGAUUAGUUUGGA 3	3 CACUCGUGACUUAAUUAAACC 5	21	18	-22.6
not in complementary region	5 CCUAACGACACCAAAUUUGUGACUUGA 3	3 AUGUUGUGGUUUAAACA-UAAU 5	21	22	-24.3
	5 CGACUGGGAAUUGGAACCUGAUGCAUCUCA 3	3 GACUCUGUAAUUUGGACU-CGUAGU 5	24	20	-32.1
	5 CGUUCCGUCAAGCUUCAAGAUUUACCUCUC 5	3 AGGUAGUUUGAAACCAUAGAUGGA 5	24	18	-27.9

** the *OsDREB2A* sequence marked yellow was the probable site which was thought to have amplified in the nested PCR

Expression analysis of *OsDREB2A* transcripts and the overlapping anti-sense genes under salt stress conditions

The expression patterns for both IR-29 and IR-64 was examined at 250mM NaCl for control 2 hour, 4 hour and 6 hour intervals with semi quantitative and real time PCR analysis and Actin was used as an internal control to normal the expression data obtained for each gene. To specifically identify the *OsDREB2A* transcripts primers were designed specific for each transcript (Primer Sequence list). The fold change in the *OsDREB2A* transcript 2 for IR-29 was not significant (Fig 5E) and fluctuated at every time point. The expression of the overlapping anti-sense gene also negligible and at each time point the expression of the two gene were antagonistic to each other(Fig5 E). In the IR-64 rice variety the fold change in the expression of *OsDREB2A* transcript 2 increased with time and was 10 times at 4 hours salt stress

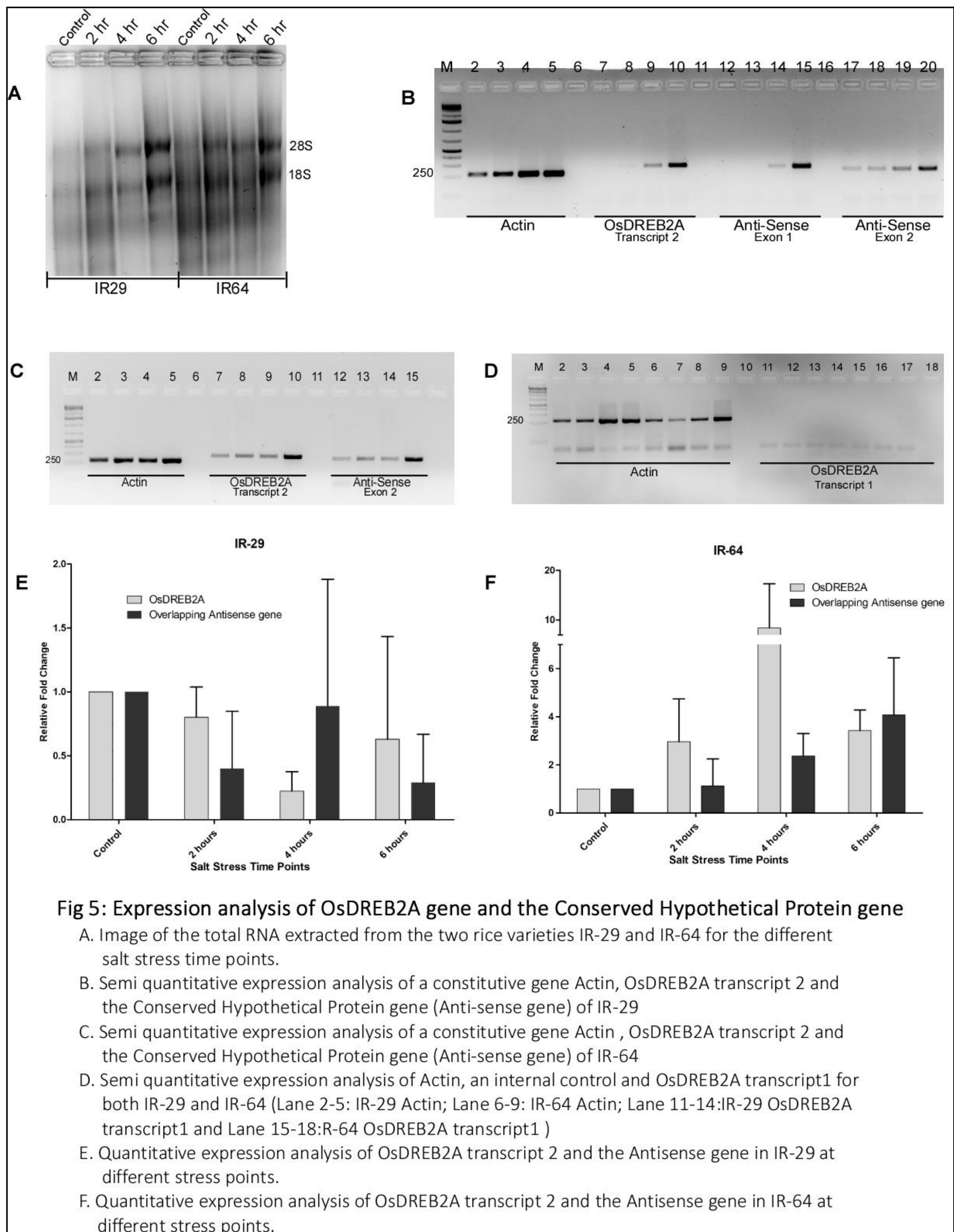


Fig 5: Expression analysis of OsDREB2A gene and the Conserved Hypothetical Protein gene

A. Image of the total RNA extracted from the two rice varieties IR-29 and IR-64 for the different salt stress time points.
B. Semi quantitative expression analysis of a constitutive gene Actin, OsDREB2A transcript 2 and the Conserved Hypothetical Protein gene (Anti-sense gene) of IR-29
C. Semi quantitative expression analysis of a constitutive gene Actin , OsDREB2A transcript 2 and the Conserved Hypothetical Protein gene (Anti-sense gene) of IR-64
D. Semi quantitative expression analysis of Actin, an internal control and OsDREB2A transcript1 for both IR-29 and IR-64 (Lane 2-5: IR-29 Actin; Lane 6-9: IR-64 Actin; Lane 11-14:IR-29 OsDREB2A transcript1 and Lane 15-18:R-64 OsDREB2A transcript1)
E. Quantitative expression analysis of OsDREB2A transcript 2 and the Antisense gene in IR-29 at different stress points.
F. Quantitative expression analysis of OsDREB2A transcript 2 and the Antisense gene in IR-64 at different stress points.

after which its expression started to decrease. The expression of the the anti-sense gene also increased with time (Fig 5F).

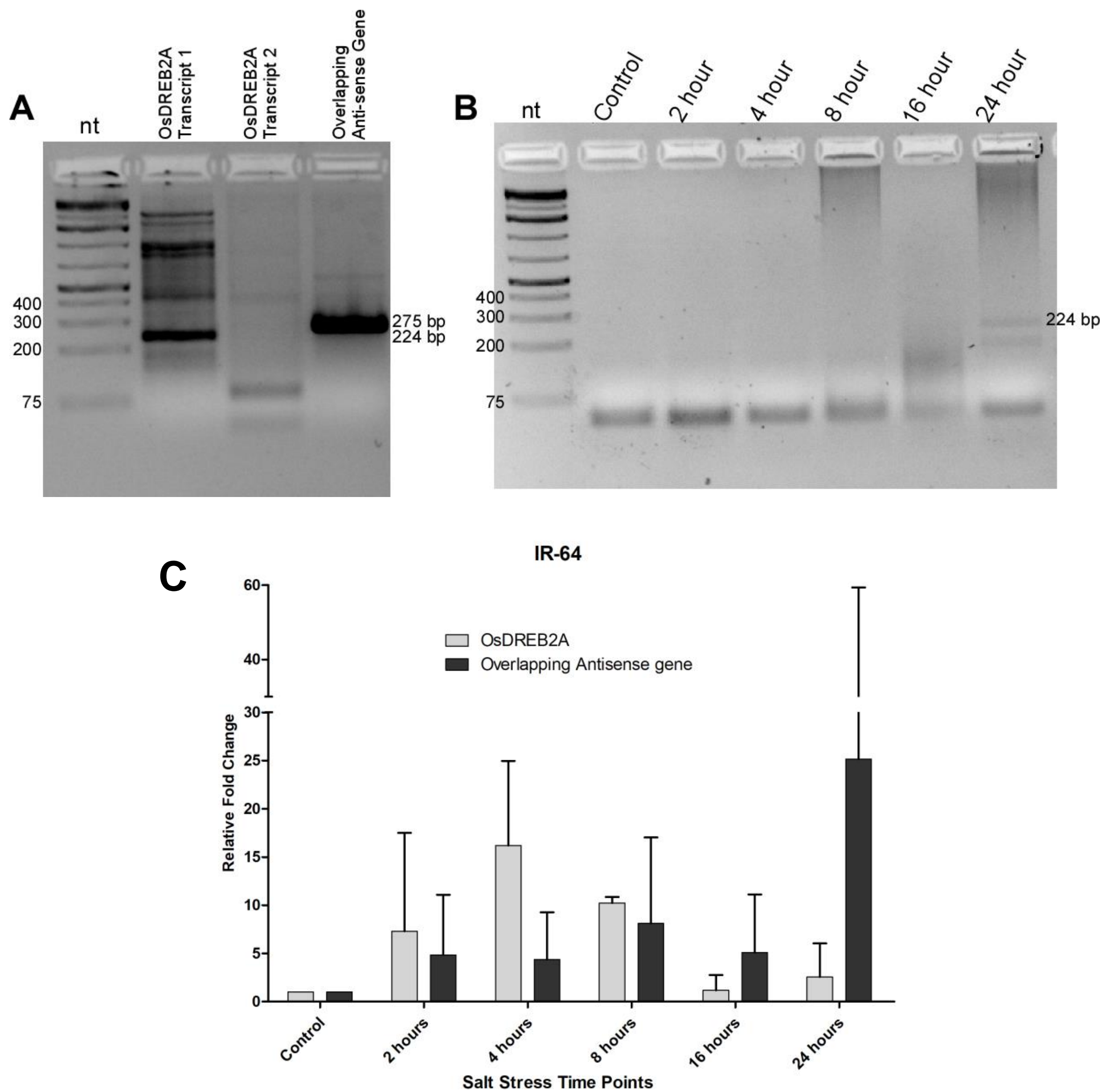


Fig 6: Expression analysis of the genes in IR-64.

A. Genomic DNA expression analysis (No band in Transcript2 as the primers were designed from exon- exon junction, 224bp band for Transcript1 and 275bp band for overlapping anti-sense gene)

B. cDNA expression analysis of *OsDREB2A* transcript 1.

C. Quantitative expression analysis of *OsDREB2A* transcript 2 and the anti-sense gene for longer stress points.

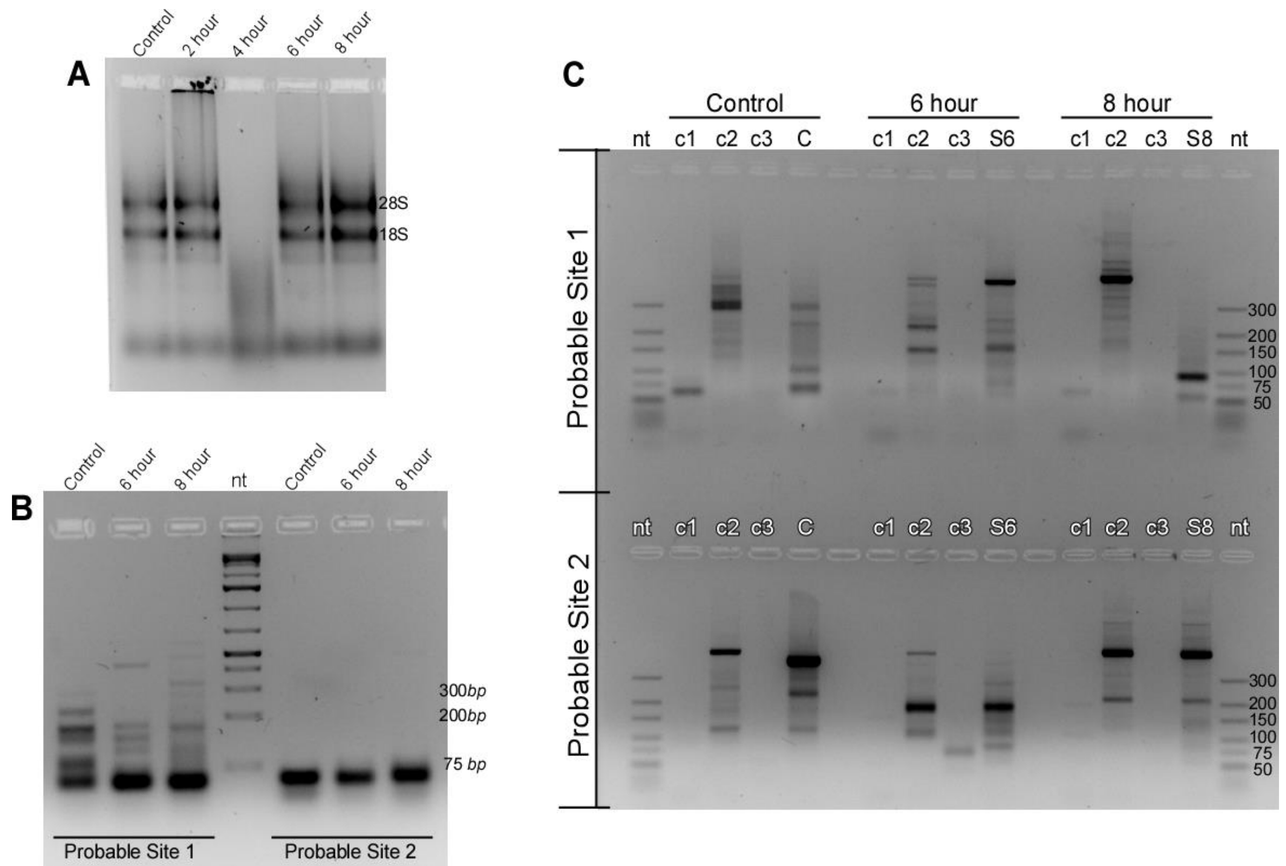


Fig 7: RLM 5' RACE data analysis for IR-64

A. RNA gel image of the different salt stressed samples.

B. Touchdown PCR gel image for the amplification of the 5'ends. Two sites having high population of smRNA targeting the *OsDREB2A* transcript2 were chosen and were named Probable site 1 and Probable site 2 respectively for each time points namely 6hours and 8 hours.

C. Nested PCR for the two sites. c1: control amplification without any template; c2: control amplification with only template and RNA oligo specific forward primer; c3: control amplification with only template and gene specific reverse primer; LaneC, S6, S8: nested amplification of the previous PCR product.

PROBABLE SITE 1 SEQUENCE:

*TCGCATTATAGCGCTCTATGGTCATTCAAAGTACCAAAAGATACATCATGGAGATTTCATAAGGGGAAGTTGCAGACATTCTCTGAATTCAAGCCTCCTTGTTTCACTGTGAA AGGTCTTCTGTTTGTTAGGCCAGTAGGTCAGCCAAATCCAGCCTCTAATCTCCCAGTCCATGGTCTTCCCTAATGCTGCACCACAGATCTTCCACCAGACGCTGTCTCCATATTC ATCACGTGCGAAAATGACACCACCCTTCTTCGTGAATGGGCGATACACGGT***CGAGTACCTGTCACGAAACC***AGGCCCGTGGCTGAAGCAGGGAGAGCTTGGAAGGCTTTG AAGAATAAA***CCTCCCTTCCGTTTACACCATC***ATGGAG*

PROBABLE SITE 2 SEQUENCE:

*AGTTGCTGATAAGTAAGCGAGAAGAGAAGGCGCTCATCCTGTGTTGGCACTGCAGGATCTTCCATGCTCATGGCAATGAGTGGGCAGTATAGCTGGACCTTCCAGAGCCCA CTTGCTGAAAGAGCATATGAGAACAGTGGCGATGGCACCCCAAATGAATGGTGTGAGGATCGTTGCTCAACAATCCAGTTAGTCACTGCAAGATTCATTGTCTGCATCCATT GCTCCTCAAGGTTTGCTCCCAACATTCTGACAAAGGTCTTTGCAGCATCTTT***GCACTTTGGCCCAGTCAGC***TGAGTCCTCAGGGATTCGAGGCATCGACAG***CGAAGAGCATG GGGAGCCTC***ATAGAT*

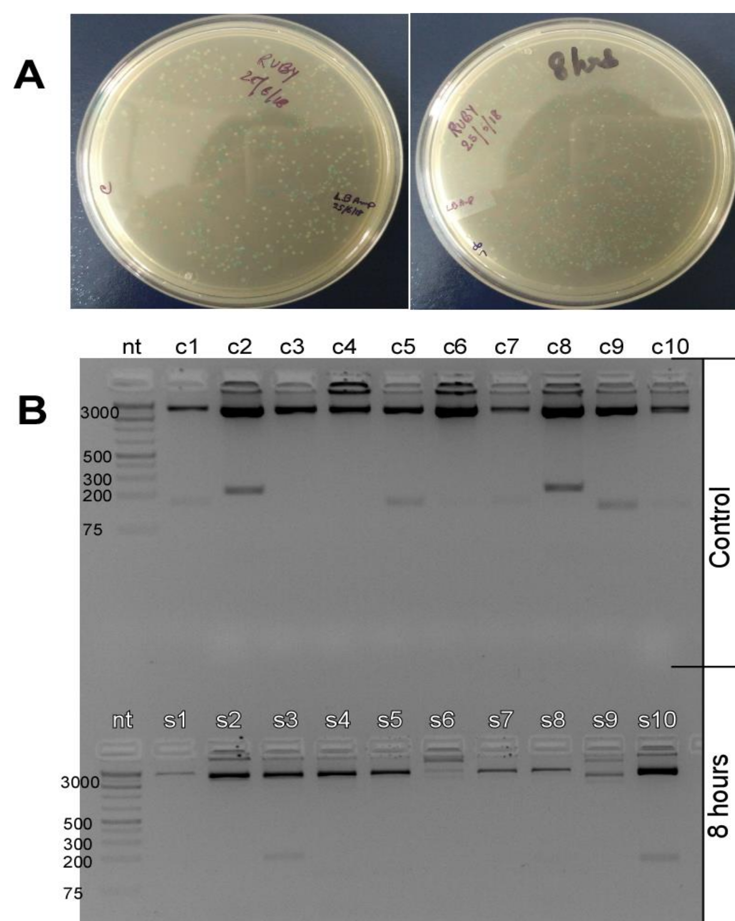


Fig 8: Transformant selection and conformation of insertion

A. Blue white screening for transformants in Ampicillin plates

B. Gel image of the *BamH* I and *Hind* III restriction enzyme digested vector to confirm the insertion of the fragment.

The expression of *OsDREB2A* transcript 1 was not detected in both IR-29 and IR-64 (Fig 5D). In order to verify whether the primers specific for *OsDREB2A* transcript 1 were working PCR was done using genomic DNA of IR-64 and band was obtained for the transcript (Fig 6A). To determine the exact time point at which the transcript 1 starts expressing we performed a semi quantitative PCR for control, 2hour, 4hour, 8hour, 16 hour and 24 hour salt stressed samples. Amplification of transcript 1 was obtained in 24 hour salt stressed sample (Fig 6B). Furthermore, a quantitative expression analysis of transcript 2 and the overlapping gene of the late stress hours

in IR-64 shows that fold change in transcript 2 rapidly declines after 4 hour salt stressed sample and this decline in fold change for the overlapping gene is observed after 8hours salt stressed sample (Fig 6C).

Determination of the cleavage site in *OsDREB2A* transcript 2 of IR-64 by RLM-5'RACE

To map the cleavage site in the transcript 2 we performed RLM-5'RACE for control, 6hour and 8hour salt stressed samples. Two probable sites where the small RNA population was high were chosen and the reverse gene specific primers for those regions were designed (Primer sequence list). The amplification using outer primers for the small RNA targeting probable site 1 did give bands of interest but no such bands were obtained for the probable site 2 (Fig 7B). The data obtained from nested PCR showed the presence of a unique band in the control sample at 90 nucleotide fragment position which appeared enriched in the 8hour sample of the first probable site (Fig 7C). The enrichment of the band in the 8hour stressed sample was a clear indication that the cleavage of the transcript 2 at 8hour time point was high. This data can be correlated to the decrease in the expression of the *OsDREB2A* transcript2 and the overlapping anti-sense gene. It is possible that the guide RNA originating from the overlapping anti-sense gene was getting loaded in the RISC complex and was targeting the 3'UTR of the transcript 2 of *OsDREB2A* gene. No such unique band was noted in the probable site 2 (Fig 7C). This 90 nucleotide positioned band was one of the small RNA targeting predicted cleaved site product (yellow marked in the sequence). The fragment was inserted in the pMD20 vector and the transformants were selected by blue white screening in ampicillin plates (Fig 8A). However, on sequencing, the predicted cleavage site was not obtained.

The sequence analysis of the two transcripts of *OsDREB2A* show that they both have identical CDS which means that the two transcripts code for the same protein (Fig 3 A). The only difference lies in their 3'UTR which is longer in case of the second transcript. This elongated region of the second transcript overlaps the Conserved Hypothetical Protein gene in the anti-sense strand.

In salt sensitive variety IR29, the expression pattern of the overlapped anti-sense gene and the transcript2 of *OsDREB2A* are antagonistic to each other (Fig4 E). In IR64, the expression of *OsDREB2A* transcript 2 increases in phase with the expression of the overlapping anti-sense gene (Fig 5 C). Expression study has shown that Transcript1 is a late stress specific transcript in IR64 (Fig4 D), whereas transcript2 is early salinity stress responsive. The phasing in the expression pattern of transcript2 and the overlapped anti-sense gene in IR-64 led to the investigation of probable regulatory mechanism involved in generation of siRNA (Fig5 C).

RLM 5'-RACE was performed in order to see the role of endogenous siRNA in the post-transcriptional gene silencing (PTGS) of *OsDREB2A* transcript2. The idea was to search for a cleavage site in the *OsDREB2A* transcript 2. However, after sequencing no cleavage site was found specific to *OsDREB2A* transcript 2 or any region of the locus as such.

This study paves the way for further validation of a siRNA and its regulation via TGS which, remains unexplored in this work. From the study it can be concluded that the expression of the two transcripts of *OsDREB2A* are correlated for proper rescuing of the plant from salt stress and their regulation via siRNA mediated PTGS is not the mechanism.

CONCLUSION

- *OsDREB2A* transcript1 is a late salt stress responsive gene whereas *OsDREB2A* transcript 2 is an early salt stress responsive gene.
- The hypothesized regulation of *OsDREB2A* transcript 2 by siRNA is not mediated by post transcriptional gene silencing.

REFERENCE

1. Ambros, V. (2003). MicroRNA pathways in flies and worms: growth, death, fat, stress, and timing. *Cell* 113, 673–676.

2. Boi, S., Solda, G., and Tenchini, M.L. (2004). Shedding light on the darkside of the genome: overlapping genes in higher eukaryotes. *Curr.Genomics* 5, 509–524.

3. Jen, C.-H., Michalopoulos, I., Westhead, D.R., and Meyer, P. (2005). Natural antisense transcripts with coding capacity in *Arabidopsis* may have a regulatory role that is not linked to double-stranded RNA degradation. *Genome Biol*. 6, R51

4. Bartel, D.P. (2004). MicroRNAs: genomics, biogenesis, mechanism, and function. Cell 116, 281–297.

5. Baulcombe, D. (2004). RNA silencing in plants. *Nature* 431, 356–363.

6. Carrington, J.C., and Ambros, V. (2003). Role of microRNAs in plant and animal development. *Science* 301, 336–338.

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

8. Gasciolli, V., Mallory, A.C., Bartel, D.P., and Vaucheret, H. (2005). Partially redundant functions of Arabidopsis DICER-like enzymes and a role for DCL4 in producing trans-acting siRNAs. *Curr. Biol*. 15, 1494–1500.

9. Hamilton, A.J., and Baulcombe, D.C. (1999). A species of small antisense RNA in post transcriptional gene silencing in plants. *Science* 286, 950–952.

10. Liu Q, Kasuga M, Sakuma Y, Abe H, Miura S, Yamaguchi-Shinozaki K, Shinozaki K (1998) Two transcription factors, DREB1 and DREB2, with an EREBP/AP2 DNA binding domain separate two cellular signal transduction pathways in drought- and low temperature-responsive gene expression, respectively, in *Arabidopsis*. *Plant Cell* 10:1391–1406

11. M. Bühler, A. Verdel, and D. Moazed, "Tethering RITS to a nascent transcript initiates RNAi- and heterochromatin-dependent gene silencing," *Cell*, vol. 125, no. 5, pp. 873–886, Jun. 2006.

12. Nakashima K, Shinwari ZK, Sakuma Y, Seki M, Miura S, Shinozaki K, Yamaguchi-Shinozaki K (2000) Organization and expression of two *Arabidopsis* DREB2 genes encoding DRE-binding proteins involved in dehydration- and high-salinity-responsive gene expression. *Plant Mol Biol* 42:657–665

13. B. Vogelstein and D. Gillespie, "Preparative and analytical purification of DNA from agarose," *Proc. Natl. Acad. Sci.* U. S. A., vol. 76, no. 2, pp. 615–619, Feb. 1979.

14. Nakashima K, Ito Y, Yamaguchi-Shinozaki K (2009) Transcriptional regulatory networks in response to abiotic stresses in *Arabidopsis* and grasses. *Plant Physiol* 149:88–95

15. Sakuma Y, Maruyama K, Qin F, Osakabe Y, Shinozaki K, Yamaguchi-Shinozaki K (2006b) Dual function of an *Arabidopsis* transcription factor DREB2A in water-stress-responsive and heat-stress responsive gene expression. *Proc Natl Acad Sci USA* 103:18822–18827

16. Dubouzet JG, Sakuma Y, Ito Y, Kasuga M, Dubouzet EG, Miura S, Seki M, Shinozaki K, Yamaguchi-Shinozaki K (2003) OsDREB genes in rice *Oryza sativa L.*, encode transcription activators that function in drought-, high-salt- and cold-responsive gene expression. *Plant J* 33:751–763

17. Cui M, Zhang W (2011) Induced over-expression of the transcription factor OsDREB2A improves drought tolerance in rice. *Plant Physiology and Biochemistry* 49:1384-1391

18. Chomczynski P, Mackey K. (1995) Short technical report. Modification of the TRIZOL reagent procedure for isolation of RNA from Polysaccharide-and proteoglycan rich sources. *Biotechniques* 19(6): 942-5

19. Takuji Sasaki, Takashi Matsumoto (2002) The genome sequence and structure of rice chromosome 1 *Nature* 420: 312–316

20. K. Maruyama and S. Sugano, "Oligo-capping: a simple method to replace the cap structure of eukaryotic mRNAs with oligoribonucleotides," *Gene, vol.* 138, no. 1–2, pp. 171–174, Jan. 1994

21. Hannon, G.J. (2002). RNA interference. *Nature* 418, 244–251.