

IN SILICO CHARACTERISATION OF NOVEL RICE TRANSCRIPTS DIFFERENTIALLY EXPRESSED IN PHOSPHORUS DEFICIENT CONDITIONS SUGGESTS A ROLE OF THESE TRANSCRIPTS IN MULTIPLE ABIOTIC STRESSES

JULIA S. YUMNAM,¹ MAYANK RAI¹ and WRICHA TYAGI¹*

School of Crop Improvement, College of Post-Graduate Studies,
Central Agricultural University (Imphal), Umiam, Meghalaya 793103

(Received: January 19, 2017; accepted: August 18, 2017)

Phosphorus deficiency adversely affects crop productivity. The mechanism of tolerance in plants is not well understood. The current study successfully annotated a set of highly significant (Log_2 RPKM ≥ 3) nine novel sequences up-regulated in P deficient condition identified from a low P tolerant rice genotype. Sequence annotation identified two transcripts (Os01g37260 and Os02g11060) carrying known domains, F-box and WD, respectively. Multiple Expectation maximization for Motif Elicitation (MEME) revealed presence of conserved domains like D[LP][HY][CL]D[CM][DT]C[AP][DQ][IQ]C, [EH][DN]HN[HS][ER][FY][EP]I[HN]H which might play a role in phosphorus deficiency tolerance. Analysis of the upstream regions indicated presence of stress responsive elements like E Box, ABRE, and MYBCORE suggesting regulation of the novel transcripts by DNA binding. Protein localization prediction tool suggests that these novel proteins might be targeted to nucleus, chloroplast and cell wall. Transcripts Os02g03640 and Os02g10250 revealed potential target sites for microRNA binding suggesting role of novel miRNAs in low phosphorus response. Our analysis suggests that an F-box protein, Os01g37260 (OSFBx14) might be a promising candidate gene playing a role in multiple abiotic stresses including P deficiency.

Keywords: *In silico* analysis – phosphorus tolerance – novel transcripts – motif – annotation

INTRODUCTION

Phosphorus (P) is an essential macronutrient required for plant growth and development [12]. It is a constituent of biomolecules and helps regulate many biological processes. Suboptimal levels of P can lead to yield losses in the range of 5–15%. For these reasons, crops are supplied with inorganic P fertilizers. Regions of low P are mostly found in developing countries and procurement of phosphatic fertilizers has significant cost implications for farmers in these countries. Moreover, the source of phosphoric fertilizers i.e., rock phosphate is a non-renewable resource which is likely to be depleted in 50–100 years [6].

*Corresponding author; e-mail address: wtyagi.cau@gmail.com

Plants have developed different physiological and biochemical mechanisms to cope with inadequate supply of inorganic phosphate (Pi) [18]. Previously, we had ascertained that Sahbhagi Dhan (SD), a released Indian rice variety is, a low P tolerant genotype [7]. SD has an ability to maintain high levels of Pi uptake (5.8 ± 2.1 mg/plant) even under soils having low Pi availability (6.25 kg/ha) [26]. Root elongation is observed in phosphorus-deficient soils in various plant species, including rice [11]. Typically, a susceptible genotype shows significantly low uptake of Pi and low P use efficiency and also low root plasticity under low P conditions [18, 26]. SD copes with long-term low P exposure (0.015 mM Pi for 15 days) by increasing root length and root biomass (approx. 1.2 fold increase) [22]. Transcriptome data (SRP058964; PRJNA285664) generated on roots of this genotype in our lab suggests that SD responds to low P levels through a serine-threonine kinase using sugar- and auxin-mediated pathway. This work also resulted in identification of 1620 significantly differentially expressed genes (DEGs) in SD with enriched in biological gene ontology (GO) terms associated with biosynthetic response (19.4%), response to stress (17.5%) and metabolic activity (16.5%) [22].

Although detailed molecular mechanism pertaining to P starvation sensing, signaling and associated target gene activation has not yet been fully elucidated, various studies have suggested the role of key proteins like PHR1, Pht1, etc. in rice [18]. Apart from genes, microRNAs like miR399 and miR6250 are also reported to be involved in regulation of gene expression in Pi starvation [15]. As low P tolerance is a quantitative trait, getting key genes identified and functionally annotated is crucial. However, only about 50% of the genes reported are functionally annotated. Therefore, identifying and annotating novel transcripts expressed under low P conditions from a low P tolerant genotype might be a direct way to understand and achieve low P tolerance at field level. Previously, we had reported two novel genes from another low P tolerant rice genotype and implied their role in imparting tolerance under stress conditions [7]. The objective of this study was to elucidate probable molecular and cellular role of genes significantly expressed only under low P conditions in roots of rice genotype, SD and to select probable candidates for further studies on low P tolerance.

MATERIALS AND METHODS

Selection of transcripts from transcriptome data and validation by quantitative PCR (qPCR)

DEGs expressed only under low P conditions in roots of SD at a high level (\log_2 RPKM ≥ 3.0) were selected from the transcriptome data [Table S1; 22] for further characterization. The qPCR and data analysis was performed as reported previously [22] using gene-specific primers (Table 1). Briefly, amplification was confirmed by melt curve analysis and threshold values (Ct) for each of the transcripts calculated. All data are given in terms of relative mRNA expression under low P condition normalised to ubiquitin gene.

Table 1
Primers used in the present study

Primer	Sequence (5' - 3')
Os01g35830_F	AAGGAATCGTGGAGTGCAAC
Os01g35830_R	TCCAGGACGGATTCTTCAC
Os01g37260_F	CATCGAAAGCAAGGGTTCTC
Os01g37260_R	ATTCTAGGGCAGGGCATTTC
Os02g03640_F	TATTAATGTTTCGAGGCAACTGCT
Os02g03640_R	TCTCCAGTCTGTGTAGCCTAAGCTA
Os02g10250_F	TGGCCAAACAAGAAAACCA
Os02g10250_R	AGCAACCGTCAGGCTTAGAA
Os02g11060_F	CAGCCCCAAGAGAGATTGAC
Os02g11060_R	TTGACACTGGAACCATCAGC
Os03g28250_F	CTTCCTTGCGACCCTTCTT
Os03g28250_R	CCAATTCACATCCCAAATCC
Os03g32330_F	GTGCCTCAGGTTGCAGGT
Os03g32330_R	GCCACTGTCGGTGTTC
Os04g10214_F	AGGTACGGAGGCGTGGA
Os04g10214_R	CACGTGCCATCCACCAC
Os04g26330_F	GGCGGTTTTAACGTCTCCAT
Os04g26330_R	CCTGTACTGCAGGCTTGGAC

Annotation of novel transcripts

Rice Annotation Project (TIGR version 7) was used for annotation at the genome level. To search the homology of the selected genes, BLASTn (Basic Local Alignment Search Tool nucleotide) with three different optimization runs (megablast, discontinuous megablast and BLASTn) was performed. The matches were verified and refined through BLASTX and NCBI open reading frame finder (ORF Finder). Variation at the nucleotide level was identified using available SNP chip data (generated on 400 *O. sativa* genotypes) in GRAMENE database. To find homology at the protein level, protein blast (BLASTp (protein-protein BLAST), PSI-BLAST (Position-Specific Iterated BLAST) and PHI-BLAST (Pattern Hit Initiated BLAST) was performed. The degree of protein hydrophobicity or hydrophilicity was determined using Kyte Doolittle Hydropathic plot [13] using a window size of nine. Function of a novel protein can be predicted by using protein-protein interactions as well as localization of the protein within the cell. The STRING database was used for studying the known and predicted protein interactions. Prediction for protein localization at the subcellular level was done using WOLFPSORT and Plant-mPloc [5].

To search for the *cis* acting elements in upstream sequences of the genes, Plant *cis*-Regulatory DNA Elements database (PLACE) was used.

To investigate the regulatory motifs present in the proteins, MEME (Multiple Expectation maximization for Motif Elicitation) was used (run parameters; minimum width-10, maximum width-50 and maximum number of motifs-20). Three proteins, OsPSTOL1 (Phosphorus Starvation Tolerance 1; AB458444), PTF1 (Pi starvation induced Transcription Factor 1; AY238991) and OsPUPK-20 (Dirigent-like gene; AB458444.1), already reported for tolerance to low P were included as reference. The putative roles of identified motifs were searched using DELTA BLAST.

The potential regulation of the selected transcripts by non-coding RNAs like micro RNA and small RNA was searched by using Rice Functional Genomic Express Database v.6.1 and Cereal small RNAs Database, respectively to identify overlap of reported/annotated non-coding RNA with the transcripts.

RESULTS

Selection, validation and functional annotation of transcripts

Selection of DEGs highly expressed only under low P conditions in roots of SD led to a set of 10 transcripts distributed over 5 chromosomes. One of the transcripts, Os12g11390, did not show any hit with any database including TIGR and was therefore, not considered for further gene ontology and functional annotation. The nine transcripts (Table 2) were validated for expression under low P conditions through a) wet lab real-time experiment (Fig. 1), b) search of available rice expression databases and c) candidate gene marker development (unpublished data). The mRNA expression of all the nine transcripts was significantly upregulated after 15 days of low P treatment (Fig. 1a) and the statistical analysis showed significant association ($r^2 = 0.76$) between the results of qPCR and RNA-seq data analyses (Fig. 1b). The analysis of the transcripts revealed variation (276 bp to 5338 bp) in genomic length. Introns were found in all DEGs except *Os02g10250*. The longest predicted protein length was for Os02g03640 (750 amino acid) while Os03g28250 had the shortest (74 amino acids). Conserved domains Fbox and WD domain were found in Os01g37260 and Os02g11060, respectively (Table 2).

Potential functional variants due to the presence of non synonymous SNPs

SNPs are the most common form of genetic variation and rice has around 4000 reported SNPs. GRAMENE tool revealed that nucleotide variations were more in the intronic regions than the exonic regions (Table 2). It has been reported previously that SNPs are more common in the UTR (Untranslated Region) than the exonic region [3]. Out of the total variations possible for the nine transcripts, ten SNPs affect the protein

Table 2
 Annotation of nine rice transcripts based on TIGR database ver. 7.0 and nucleotide variation. Variations (based on GRAMENE database) in the coding region, untranslated regions (5' UTR, intron, 3' UTR) and at the protein level are indicated

Locus ID (LOC_Os)	Log ₂ RPKM	Length (bp and amino acids, resp.)			Putative Domain	SNPs identified				
		Genomic		Protein		Coding Region	Untranslated Region			Number of amino acid changes**
		CDS	Protein				3' UTR	Intron	5' UTR	
01g35830	5.19	2705	699	232	-	-	2	-	-	0
01g37260	3.71	4377	1440	479	F-box	-	-	80	-	1
02g03640	5.87	4292	2253	750	-	-	2	3	-	0
02g10250	4.63	276	276	91	-	-	4	-	-	0
02g11060	3.56	5377	1029	342	WD, G-beta	-	-	-	15	0
03g28250	5.64	340	225	74	-	-	2	3	-	1
03g32330	3.72	5338	375	124	-	-	5	2	28	4
04g10214	4.02	866	294	97	-	-	6	-	-	4
04g26330	6.02	1639	333	110	-	-	-	-	-	0

*F – Cyclin F; WD – Beta-transducin repeats with a tryptophan-aspartic acid (W–D) dipeptide; G – Protein domain activating GTPase; SNP – Single Nucleotide Polymorphism; RPKM – Reads Per Kilobase per Million mapped reads; **Based on non-synonymous mutations.

function (Table 2). Whether these protein variants will show association with abiotic stress in general and with low P tolerance in particular needs to be investigated.

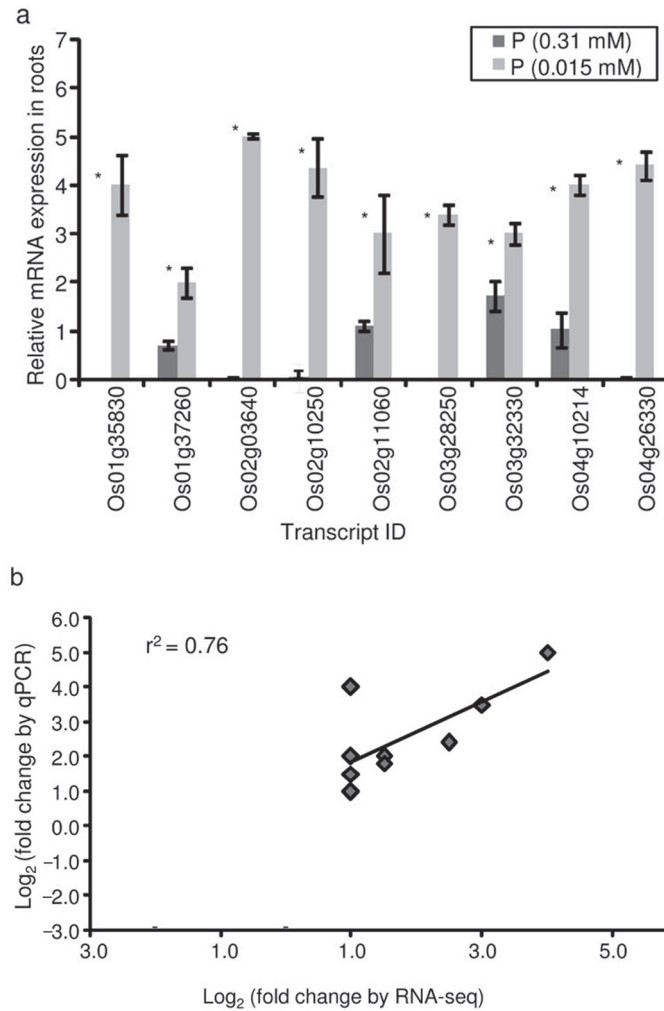


Fig. 1. (a) Fold changes in selected transcripts relative to *ubiquitin* under optimum and low P conditions in genotype SD analyzed by quantitative real-time PCR (qPCR). The values are shown as mean±SE (N = 3). Significance was defined as P<0.05 (Student's *t*-test). (b) The correlation of gene expression results obtained from qPCR analysis and RNA-seq for nine selected genes

Motif number	SEQUENCE OF THE MOTIF	Locus ID of the respective transcript	Reported low P tolerant protein carrying the motif	E-value	Visual representation of the logo	Putative role of the motif as per DELTA BLAST
1	D[L]P[I]H[Y]C[L]D[CM]I[D]T[C]A[P]D[Q]I[Q]C	Os01g37260	AB458444.1	8.6e+000		Aminoaldehyde dehydrogenase
2	[W]Y[L]H[CS]I[D]H[ICQ]I[NS]I[L]R[I]S[S]C[V]I[CH]I[FT]D[IM]	Os02g11060	AB458444	3.5e+001		DNA-binding domain
3	[H]Q[I]R[W]I[L]P[Y]D[Y]N[L]D[K]I[QR]I[A]F[C]C[M]	Os01g37260	AB458444	3.3e+001		Protein kinase Ck2
4	[E]H[I]D[N]H[N]H[S]I[E]R[I]F[Y]I[E]P[I]H[N]H	Os02g10250	AB458444	1.1e+002		Calmodulin-like
5	Y[E]M[I]W[Y]I[M]P[I]G[P]I[N]G[S]I[G]S[I]L[S]I[D]E[I]N[R]Y[I]N[S]I[F]C[V]D	—	AB458444, AY238991	1.5e+002		Aminoaldehyde dehydrogenase
6	P[I]R[C]I[A]I[M]E[E]S[V]I[F]Y[I]L[N]I[L]O[S]I[K]I[K]Q[R]I[E]I[G]R[T]I[S]E[I]N[Q]R[I]W[P]I[K]R[V]I[D]Q[R]I[Y]F	Os02g03640, Os01g37260	AB458444	4.7e+002		Late Embryogenesis Abundant (LEA) protein
7	H[I]P[I]E[N]I[D]R[I]N[E]L[E]F[Y]V[I]F[T]I[S]T[G]I[GL]P	Os01g37260	AB458444.1	2.9e+003		Subtilisin-like

Fig. 2. List of motifs found common between novel rice proteins and low P reported proteins. The *E*-values, sequence and visual representation of the identified common motifs is shown. Height of the individual letters in a stack is the probability of the letter at that position multiplied by the total information content of the stack. Putative role of the common motifs as revealed by DELTA BLAST is also indicated

Homology search at the protein level

BLASTp analysis revealed that Os02g11060 is a WD40 and Os01g37260 an Fbox (homologous to a serine-threonine kinase receptor-associated) protein, respectively. For the rest of the proteins, no homology was obtained. Motifs represent functional regions of a protein and presence of such motifs often gives an insight into the function of the protein. Based on the MEME tool twenty putative motifs with E values ranging from 1.1e-003 to 8.6e+000 were identified. A total of six motifs common between novel rice peptides and two low P tolerant genes (PsTOL1 and dirigent-like protein) were identified (Fig. 2). One motif was found common between PSTOL1 and PTF1. DELTA BLAST revealed that the six identified motifs may have roles similar to various proteins like aminoaldehyde dehydrogenase, protein kinase, LEA protein and DNA and calcium binding (calmodulin) proteins. All the above mentioned proteins are known to play a role in stress response in plants [19] and therefore, the five proteins too might play a role in response to other abiotic stresses apart from phosphorus starvation.

Protein characterisation at cellular level

Kyte Doolittle hydrophathy plot predicts nature of the protein i.e., globular or transmembrane. All the proteins seem to be globular in nature except Os01g37260 and Os02g03640 (Fig. 3a and 3b) which have transmembrane regions and might be associated with membranes. Most of the proteins were predicted to be localized in the nucleus, chloroplast, cytoplasm, mitochondria or shuttle between chloroplast

Table 3
Gene ontology (GO) of the nine transcripts highly expressed under low phosphorus conditions

Locus ID	Putative gene ontology terms		
	Cellular	Biological	Molecular
01g35830	–	–	Nucleotide binding
01g37260	Cytosol, Chloroplast	FBX14	Response to abiotic stress
02g03640	Chloroplast, nucleus	Late Embryogenesis Abundant (LEA) protein	Response to abiotic stress
02g10250	Nucleus	Calmodulin-Like Calcium-Binding protein	Signaling in abiotic stress response
02g11060	Nucleus	WD40-38	Response to abiotic stress
03g28250	Chloroplast, cytoplasm mitochondria	–	–
03g32330	Nucleus	–	–
04g10214	Cell wall, cytoplasm, mitochondria	–	–
04g26330	Nucleus, chloroplast	Cyclin like F-box containing	Response to abiotic stress; tissue specificity

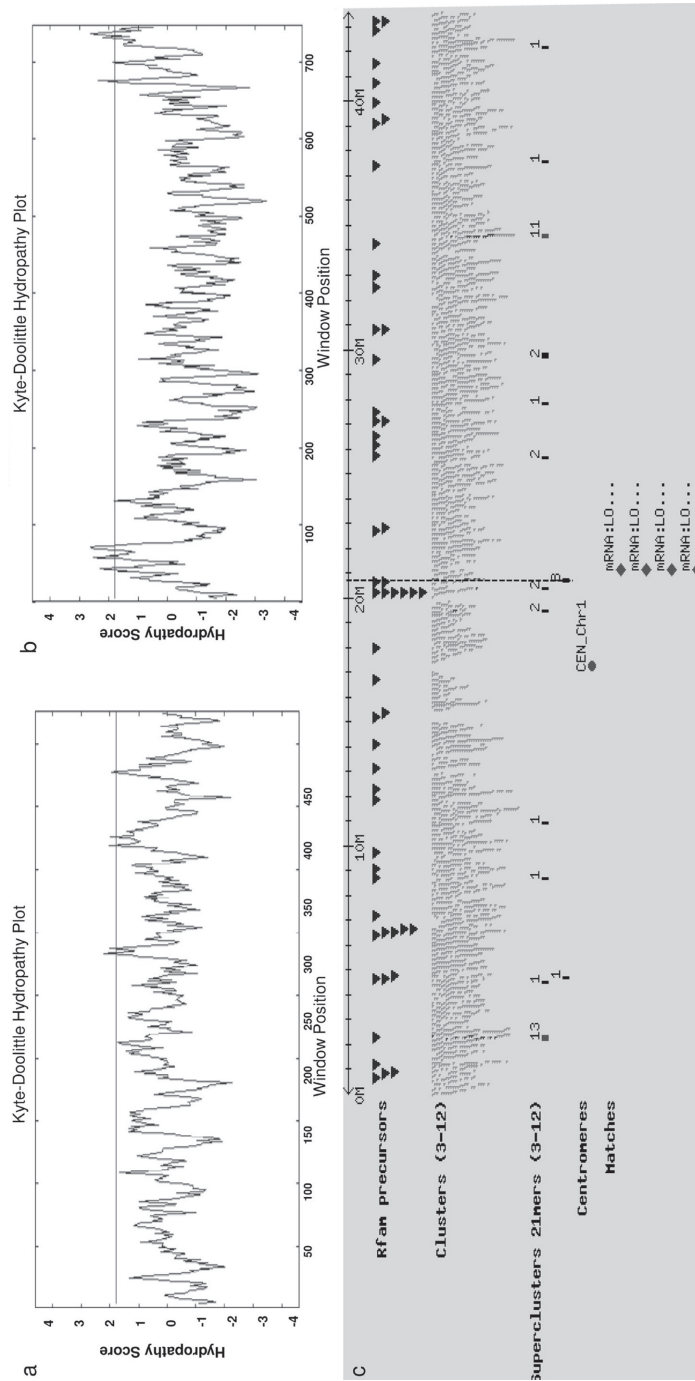


Fig. 3. Kyte Doolittle Hydropathy plots for Os01g37260 (a) and Os02g03640 (b), respectively. X axis represents amino acid sequence whereas the Y axis represents degree of hydrophobicity and hydrophilicity. (c) A snapshot showing small RNA supercluster overlapping Os01g37260. The dotted vertical line and diamond shaped boxes indicate small RNA supercluster and the target gene (also referred as “matches”), respectively on the genome browser (top arrow). Triangles indicate RNA family precursor (Pfam); grey and black vertical lines indicate RNA cluster and superclusters, respectively while white circle indicates the position of centromere (cen)

Table 4
List of *cis*-acting elements along with their sequence and putative function present in the upstream regulatory region of the nine transcripts

Name of <i>cis</i> -element	Locus ID	Sequence of <i>cis</i> -elements	Putative functional role at the molecular level	Reference
EBOXBNNAPA	01g35830, 01g37260, 02g03640, 02g10250, 02g11060, 03g28250, 03g32330, 04g10214, 04g26330	CNNTG	Light-responsive and tissue-specific phenylpropanoid biosynthesis	9
LTRECOREATCOR15	01g35830, 02g03640, 03g28250, 03g32330, 04g26330	CCGAC	Low temperature stress response	1
MYB1AT	02g10250, 03g32330, 04g26330	WAACAA	Dehydration response	21
MYB2AT	04g26330	TAACTG	Dehydration response	21
MYBCORE	01g37260, 02g03640, 02g10250, 02g11060, 03g28250, 03g32330, 04g10214, 04g26330	TNGTTR	Dehydration response	21
MYB2CONSENSUSAT	01g37260, 02g03640, 02g10250, 02g11060, 03g28250, 04g10214, 04g26330	YAACKG	Dehydration response	21
MYBCOREATCYCBI	04g10214	AACGG	Cell cycle	16
MYBST1	01g37260, 02g03640, 02g11060, 03g28250, 04g26330	GGATA	Dehydration response	2
POLLENILELAT52	02g03640, 02g10250, 02g11060, 03g28250, 03g32330, 04g26330	AGAAA	Pollen specific expression	8
WRKY71OS	01g37260, 01g35830, 02g11060, 02g03640, 02g10250, 03g28250, 04g10214, 04g26330,	TGAC	Gibberellin signaling	27
CBFHV	02g03640, 02g10250, 04g10214	RYCGAC	Low temperature stress response	23
DRERTCOREAT	03g28250, 04g10214, 04g26330	RCCGAC	Dehydration response	17
ABRERATCAL	01g37260, 02g10250, 02g11060, 03g28250, 03g32330	ACGTG	Dehydration stress and senescence	10
ABRELATERDI	03g32330	MACGYGB	Ca ²⁺ responsive	10

and nucleus except for Os04g10214, which appears to be localized in the cell wall (Table 3). Protein interaction prediction revealed that Os02g11060 interacts with components of phenazine biosynthesis. Phenazines are versatile secondary metabolites of bacterial origin that function in biological control of plant pathogens [14].

Promoter level motif search suggests role of transcripts in abiotic stress

An *in silico* analysis of 1kb upstream region from the transcription start site of each one of the nine transcripts was performed using the PLACE. The analysis revealed differences in the type, number and position of *cis* elements in upstream regions of the nine peptides (Table 4). At least 15 different motifs including TATA box were identified. The rest of the 14 motifs are involved in abiotic stress and plant development. The most over-represented motifs were EBOXBNNAPA, MYCCONSENSUSAT, WRKY71OS and MYBCORE. EBOXBNNAPA is an E-box sequence [9] which is recognized by transcription factors like R2R3-MYB (MYB), bZIP (Basic Leucine Zipper), and basic helix loop helix (bHLH) and directs tissue-specific expression. MYCCONSENSUSAT is a MYC recognition site found in the promoters of dehydration-responsive genes and regulates gene expression under cold stress [4]. MYBCORE provides binding site for plant MYB proteins, which are involved in regulating gene expression in response to various stresses [20]. MYC and MYB proteins are involved in various signalling pathways involving hormone abscisic acid [21] and other abiotic stresses [24]. Role of members of WRKY and MYB family is well elucidated in low P stress [25]. POLLENILELAT52 as a regulatory element responsible for pollen-specific gene regulation is well characterised in tomato [8].

Regulation by miRNA and small RNA

Gene expression can be modulated by non-coding RNAs. *In silico* analysis predicted role of microRNA in regulating expression of Os02g03640 and Os02g10250. Os01g37260 appears to be regulated by small RNA supercluster (Fig. 3c) suggesting role of small RNA.

DISCUSSION

We selected DEGs significantly expressed only under low P condition for further annotation by setting a cut-off of \log_2 RPKM ≥ 3 . This robust criterion reduced the number of putative targets (number of selected DEGs decreased from 127 (\log_2 RPKM ≥ 2) to 10 (\log_2 RPKM ≥ 3)). Based on this selection criterion, two transcripts with known domains viz. Fbox and WD were identified in SD. Fbox domain containing proteins have been reported to be differentially regulated in different abiotic stress

conditions. WD domain containing proteins are involved in many functions including signal transduction, pre-mRNA processing and cytoskeleton assembly. Six motifs common between two reported proteins for P tolerance and five novel proteins. These motifs suggest functional involvement in DNA binding, stress responses (LEA) and as secondary messenger (Calmodulin-Like Calcium-Binding protein). Protein targeting analysis suggests that four proteins are targeted to the chloroplast, three to the nucleus, and one to the cell wall.

Transcription factors (TFs) are known to interact with *cis*-regulatory elements in the promoter regions of genes. In this study, the upstream region of the novel transcripts revealed the presence of many stress responsive elements. Motif EBOXBNNAPA (CCGAC), previously reported to be involved in low phosphorus tolerance, was present in all the promoter regions. It implies that these nine novel transcripts may belong to a similar metabolic/cellular pathway having role in low P conditions. The putative promoter regions also harbour elements that respond to ABA such as WRKY71OS (TGAC) and are regulated by sugar and gibberellic acid (GA). This *cis*-acting element was found in all the transcripts except Os03g32330. LTRECOREATCOR15 (present in Os04g26330, Os02g03640, Os03g28250, Os01g35830 and Os03g32330) plays a role in response to ABA and is known to be involved in cold induction of BN115 gene in winter *Brassica napus*. It also mediates light signaling by phytochrome and is necessary for cold- or drought- induced gene expression through the C/DRE in *Arabidopsis* [1]. EBOXBNNAPA, MYCCONSENSUSAT, WRKY71OS and MYBCORE motifs were also abundantly distributed in the nine transcripts. This suggests that these motifs may have a vital role to play in P deficiency mechanism. In fact, we had previously suggested SD relies on signalling components like transcription factors as well as members of dehydrin and LEA protein family to trigger root growth in response to low P conditions [22].

Potential target sites for microRNA were found in Os02g03640 and Os02g10250. It has been reported before that micro RNAs like mir399 are key regulators controlling gene expression under low P conditions [15]. The regulatory role of these newly identified micro RNAs needs to be further studied. Based on GO five transcripts (Os01g37260, Os02g03640, Os02g10250, Os02g11060 and Os04g26330) involved in various types of biological functions are associated with other abiotic stress responses while Os01g35830 might be involved in nucleotide binding (Table 3).

In summary, our results suggest that the novel rice transcripts might have varied role in abiotic stresses. The nucleotide level variation is more in the non-coding region. Motifs like EBOXBNNAPA, MYCCONSENSUSAT, WRKY71OS and MYBCORE are abundantly distributed in the upstream regulatory region of these transcripts. Annotation at protein level revealed that two proteins have transmembrane regions while others are globular in nature. Amino acid analysis revealed presence of seven common motifs associated with stress responses. Our analysis suggests that Os01g37260 is a Fbox protein involved in response to various abiotic stresses and is a promising candidate for future studies.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the funding from National Agricultural Innovative Project (NAIP) project grant number C30033/415101-036 to WT. JSY would like to thank Central Agricultural University (Imphal) for scholarship.

REFERENCES

1. Baker, S. S., Wilhelm, K. S., Thomashow, M. F. (1994) The 5'-region of *Arabidopsis thaliana cor15a* has cis-acting elements that confer cold, drought and ABA-regulated gene expression. *Plant Mol Biol.* *24*, 701–713.
2. Baranowskij, N., Frohberg, C., Prat, S., Willmitzer, L. (1994) A novel DNA binding protein with homology to Myb oncoproteins containing only one repeat can function as a transcriptional activator. *EMBO J.* *13*, 5383–5392.
3. Castle, J. C. (2011) SNPs occur in regions with less genomic sequence conservation. *PLoS ONE* *6*, 1–12.
4. Chinnusamy, V., Ohta, M., Kanrar, S., Lee, B. H., Hong, X., Agarwal, M., Zhu, J. K. (2003) ICE1: a regulator of cold-induced transcriptome and freezing tolerance in *Arabidopsis*. *Genes Dev.* *17*, 1043–1054.
5. Chou, K. C., Shen, H. B. (2010) Plant-mPLoc: a top-down strategy to augment the power for predicting plant protein subcellular localization. *PLoS ONE* *5*, e11335.
6. Cordell, D., Drangert, J., White, S. (2009) The story of phosphorus: Global food security and food for thought. *Glob. Environ. Change* *19*, 292–305.
7. Dkhar, F., Rai, M., Tyagi, W. (2014) Looking beyond *PtTOL1*: marker development for two rice genes showing differential expression in P deficient conditions. *J. Genet.* *93*, 573–577.
8. Filichkin, S. A., Leonard, J. M., Monteros, A., Liu, P. P., Nonogaki, H. (2004) A novel endobeta-mannanase gene in tomato LeMAN5 is associated with anther and pollen development. *Plant Physiol.* *134*, 1080–1087.
9. Hartmann, U., Sagasser, M., Mehrtens, F., Stracke, R., Weisshaar, B. (2005) Differential combinatorial interactions of cis-acting elements recognized by R2R3-MYB, BZIP, and BHLH factors control light-responsive and tissue-specific activation of phenylpropanoid biosynthesis genes. *Plant Mol. Biol.* *57*, 155–171.
10. Kaplan, B., Davydov, O., Knight, H., Galon, Y., Knight, M. R., Fluhr, R. (2006) From rapid transcriptome changes induced by cytosolic Ca²⁺ transients reveal ABRE-related sequences as Ca²⁺-responsive cis elements in *Arabidopsis*. *Plant Cell* *18*, 2733–2748.
11. Kirk, G. J. D., Du, L. V. (1997) Changes in rice root architecture, porosity, and oxygen and proton release under phosphorus deficiency. *New Phytol.* *135*, 191–200.
12. Kochian, L. (2012) Rooting for more phosphorus. *Nature* *488*, 466–467.
13. Kyte, J., Doolittle, R. (1982) A simple method for displaying the hydrophobic character of a protein. *J. Mol. Biol.* *157*, 105–132.
14. Mavrodi, D. V., Peever, T. L., Mavrodi, O. V., Parejko, J. A., Raaijmakers, J. M., Lemanceau, P., Mazurier, S., Heide, L., Blankenfeldt, W., Weller, D. M., Thomashow, L. S. (2010) Diversity and evolution of the phenazine biosynthesis pathway. *Appl. Environ. Microbiol.* *76*, 866–879.
15. Pant, B. D., Buhtz, A., Kehr, J., Scheible, W. (2008) MicroRNA399 is a long-distance signal for the regulation of plant phosphate homeostasis. *Plant J.* *53*, 731–738.
16. Planchias, S., Samland, A. K., Murray, J. A. H. (2004) Differential stability of *Arabidopsis* D-type cyclins: CYCD3;1 is a highly unstable protein degraded by a proteasome-dependent mechanism. *Plant J.* *38*, 616–625.
17. Qin, F., Sakuma, Y., Li, J., Liu, Q., Li, Y. Q., Shinozaki, K., Yamaguchi-Shinozaki, K. (2004) Cloning and functional analysis of a novel DREB1/CBF transcription factor involved in cold-responsive gene expression in *Zea mays* L. *Plant Cell Physiol.* *45*, 1042–1052.

18. Raghothama, K. G. (1999) Phosphate acquisition. *Annu. Rev. Plant. Physiol.* 50, 665–693.
19. Secco, D., Jabnourne, M., Walker, H. Wu, P., Shou, H., Poirier, Y., Whelan, J. (2013) Spatio-temporal transcript profiling of rice roots and shoots in response to phosphate starvation and recovery. *Plant Cell* 25, 4285–4304.
20. Todaka, D., Nakashima, K., Shinozaki, K., Yamaguchi-Shinozaki, K. (2012) Towards understanding transcriptional regulatory networks in abiotic stress responses and tolerance in rice. *Rice* 5, 1–9.
21. Tuteja, N. (2007) Abscisic acid and abiotic stress signaling. *Plant Signal. Behav.* 2, 135–138.
22. Tyagi, W., Rai, M. (2017) Root transcriptomes of two acidic soil adapted Indica rice genotypes suggest diverse and complex mechanism of low phosphorus tolerance. *Protoplasma* 254, 725–736.
23. Xue, G. P. (2002) Characterisation of the DNA-binding profile of barley HvCBF1 using an enzymatic method for rapid, quantitative and high-throughput analysis of the DNA-binding activity. *Nucleic Acids Res.* 30, 77.
24. Yang, A., Dai, X., Zhang, W. A. (2012) R2R3-type MYB gene, *OsMYB2*, is involved in salt, cold, and dehydration tolerance in rice. *J. of Exp. Bot.* 63, 2541–2556.
25. Yu, S., Ligang, C., Liping, Z., Diqiu, Y. (2010) Overexpression of *OsWRKY72* gene interferes in the abscisic acid signal and auxin transport pathway of *Arabidopsis*. *J. Biosci.* 35, 459–71.
26. Yumnam, J. S., Rai, M., Tyagi, W. (2015) Allele mining across two low-P tolerant genes PSTOL1 and PupK20-2 reveals novel haplotypes in rice genotypes adapted to acidic soils. *Plant Genet. Resour.* 1, 1–9.
27. Zhang, Z. L., Xie, Z., Zou, X., Casaretto, J., Ho, T. H., Shen, Q. J. (2004) A rice WRKY gene encodes a transcriptional repressor of the gibberellin signaling pathway in aleurone cells. *Plant Physiol.* 134, 1500–1513.