4.1 Characterization and Identification of Full Length cDNA of Putative RPK1

4.1.1 Isolation of Full Length cDNA and Sequence Analysis of Putative RPK1

Primers designed from putative RPK1 5' untranslated region (5' UTR) and 3' untranslated region (3' UTR) of Oryza sativa ssp. japonica (gi:18677097) and Oryza rufipogon IRGC105491 Song et al. (2009) successfully amplified a band of 2,419 bp from cDNA of Oryza rufipogon and Oryza sativa ssp. indica cv. MR219 respectively (Figure 4.1). These two fragments of gel extraction product (Figure 4.2) were successfully cloned into pGEM-T Easy vector (Promega, USA). DNA sequencing of Oryza rufipogon putative RPK1 (Oruf RPK1) and Oryza sativa ssp. indica cv. MR219 putative RPK1 (OsI RPK1) were analysed by using GenScan (http://genes.mit.edu/GENSCAN.html) and a 2,055 bp long fragment open reading frame (ORF) was identified in both sequences. Multiple alignment showed that the gene sequences of ORF Oruf RPK1 and OsI RPK1 showed 99 % identity with Oryza sativa ssp. japonica putative RPK1 (OsJ RPK1; gi:18677097; Figure 4.3). A total of eleven single nucleotide polymorphisms (SNPs) were identified among the ORF of Oruf RPK1, OsI RPK1 and OsJ RPK1 (gi:18677097) as shown in Figure 4.3, including ten SNPs between Oruf RPK1 and OsI RPK1, eight SNPs between Oruf RPK1 and OsJ RPK1 (gi:18677097), and five SNPs between OsI RPK1 and OsJ RPK1 (gi:18677097). Five out of the eleven SNPs are non-synonymous substitutions; whereas, the remaining SNPs are synonymous substitutions.

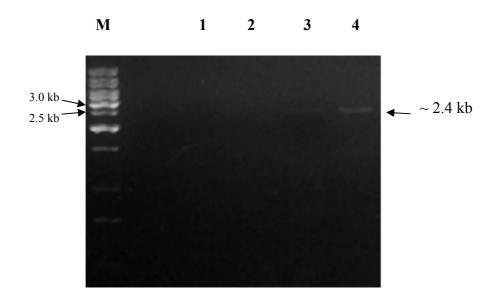


Figure 4.1: RT-PCR amplification of putative *RPK1*. The expected size of *Oruf_RPK1* and *OsI_RPK1* is around 2.4 kb. Lane M: 1 kb marker; Lane 1: DNase I treated total RNA of *Oryza rufipogon*; Lane 2: DNase I treated total RNA of *Oryza sativa ssp. indica* cv. MR219; Lane 3: *Oryza rufipogon* cDNA; Lane 4: *Oryza sativa ssp. indica* cv. MR219 cDNA.

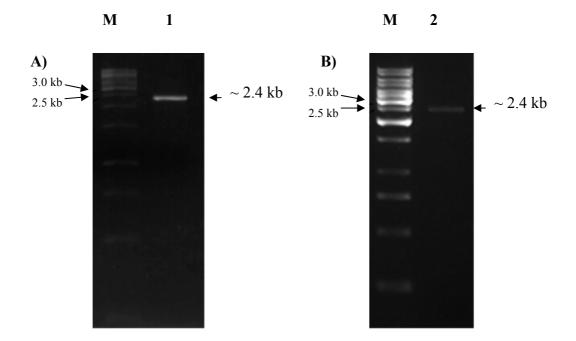


Figure 4.2: Gel extraction products of (A) *Oruf_RPK1* and (B) *OsI_RPK1*. Lane M: 1 kb marker; Lane 1: *Oruf_RPK1*; Lane 2: *OsI_RPK1*.

		10	20	30	40	50 60
OsI_RPK1 OsJ_RPK1 Oruf_RPK1 Clustal Consensus	ATGGCCGG ATGGCCGG	CGTCGTGACO CGTCGTGACO CGTCGTGACO	GAGGGCGGTGC GAGGGCGGTGC	GCGGCGGCGGT GCGGCGGCGGT GCGGCGGCGGT	GCTGGTGGTG GCTGGTGGTG	GTCGTGGTCGTG GTCGTGGTCGTG GTCGTGGTCGTG
	1	70	80	90	100	110 120
OsI_RPK1 OsJ_RPK1 Oruf_RPK1 Clustal Consensus	GCTGCTGC GCTGCTGC	CGAGCTCGT CGAGCTCGT CGAGCTCGT	GCCGCGGAGC GCCGCGGAGC GCCGCGGAGC	CCGCCGCCGAG CCGCCGCCGAG	CGAGCGGTCG CGAGCGGTCG CGAGCGGTCG	GCGCTGCTGGCG GCGCTGCTGGCG GCGCTGCTGGCG
		130	140	150	160	170 180 .
OsI_RPK1 OsJ_RPK1 Oruf_RPK1 Clustal Consensus	TTCCTGGC TTCCTGGC	GGCGACGCCC GGCGACGCCC	GCACGAGCGGC GCACGAGCGGC	GTCTCGGGTG GTCTCGGGTG	GAACTCCTCG. GAACTCCTCG.	ACGTCGGCGTGC ACGTCGGCGTGC ACGTCGGCGTGC
		190	200	210	220	230 240
OsI_RPK1 OsJ_RPK1 Oruf_RPK1 Clustal Consensus	GGGTGGGT GGGTGGGT GGGTGGGT	CGGGGGTGAC CGGGGTGAC CGGGGTGAC	GTGCGACGCCG GTGCGACGCCG GTGCGACGCCG	GGAACGCCAC GGAACGCCAC GGAACGCCAC	GGTGGTGCAG GGTGGTGCAG GGTGGTGCAG	GTGCGGCTCCCC GTGCGGCTCCCC GTGCGGCTCCCC
	10 202 10 102000	250	260	270	280	290 300
OsI_RPK1 OsJ_RPK1 Oruf_RPK1 Clustal Consensus	GGCGTGGG GGCGTGGG GGCGTGGG	GCTCATCGG GCTCATCGG GCTCATCGG	CGCCATCCCGC CGCCATCCCGC	CGGGGCACGCT CGGGGCACGCT CGGGGGACTCT	CGGCCGCCTC. CGGCCGCCTC. CGGCCGCCTC.	ACCAACCTGCAG ACCAACCTGCAG ACCAACCTGCAG
	ũ.	310	320	330	340	350 360
OsI_RPK1 OsJ_RPK1 Oruf_RPK1 Clustal Consensus	GTGCTCTC	CCTCCGCTC	CAACCGCATCC CAACCGCATCC CAACCGCATCC	CTCGGCGGCAT CTCGGCGGCAT CTCGGCGGCAT	CCCCGACGAC CCCCGACGAC	GTGCTCCAGCTC GTGCTCCAGCTC GTGCTCCAGCTC
		370	380	390	400	410 420
	53 - 556 ⁻ 38 - 1756-1556	181 Cont 1921 Las 202				
OsI_RPK1 OsJ_RPK1 Oruf_RPK1 Clustal Consensus	CCCCAGCT CCCCAGCT	CCGCCTCCTC	CTTCCTCCAGA CTTCCTCCAGA CTTCCTCCAGA	ACAACCTCCT ACCAACCTCCT	CTCCGGCGCC. CTCCGGCGCC.	ATCCCGCCGGCG ATCCCGCCGGCG ATCCCGCCGGAG ATCCCGCCGGAG
OsJ_RPK1 Oruf_RPK1	CCCCAGCT CCCCAGCT	CCGCCTCCT CCGCCTCCT CCGCCTCCT 430	CTTCCTCCAGA CTTCCTCCAGA CTTCCTCCAGA CTTCCTCCAGA	AACAACCTCCT AACAACCTCCT AACAACCTCCT **********	CTCCGGCGCC CTCCGGCGCC CTCCGGCGCC CTCCGGCGCC 460	ATCCCGCCGGCG ATCCCGCCGGCG
OsJ_RPK1 Oruf_RPK1	CCCCAGCT CCCCAGCT CCCCAGCT ******** GTCAGCAA GTCAGCAA	CCGCCTCCT CCGCCTCCT CCGCCTCCT 430 GCTCGCCGCC GCTCGCCGCC	CTTCCTCCAGA CTTCCTCCAGA CTTCCTCCAGA 440 CTCCGAGAGGGC CCTCGAGAGGGC	AACAACCTCCT AACAACCTCCT AACAACCTCCT 450 	CTCCGGCGCC CTCCGGCGCCC CTCCGGCGCCC **********	ATCCCGCCGGCG ATCCCGCCGGCG ATCCCGCCGGAG ***********
OsJ RPK1 Oruf RPK1 Clustal Consensus OsI RPK1 OsJ RPK1 Oruf RPK1	CCCCAGCT CCCCAGCT CCCCAGCT ******** GTCAGCAA GTCAGCAA GTCAGCAA *******	CCGCCTCCT(CCGCCTCCT(********************************	CTTCCTCCAGA CTTCCTCCAGA CTTCCTCCAGA 440 CTCCGAGAGGGC CCTCGAGAGGGC	AACAACCTCCT AACAACCTCCT AACAACCTCCT 450 	CTCCGGCGCCC CTCCGGCGCCC CTCCGGCGCCC ********* 460 CAGCAACAAC CAGCAACAAC CAGCAACAAC	A T C C C G C C G G C G A T C C C G C C G G C G A T C C C G C C G G A G *******************************
OsJ RPK1 Oruf RPK1 Clustal Consensus OsI RPK1 OsJ RPK1 Oruf RPK1	CCCCAGCT CCCCAGCT CCCCAGCT ******** GTCAGCAA GTCAGCAA GTCAGCAA GTCAGCAA ATCCCCTT ATCCCCTT	CCGCCTCCT CCGCCTCCT CCGCCTCCT ********* 430 GCTCGCCGCC GCTCGCCGCC GCTCGCCGCC ********* 490 CACGCTCAAC	440 CTTCCAGA CTTCCTCCAGA CTTCCTCCAGA 440 CTTCGAGAGGC CCTCGAGAGGC CCTCGAGAGGC CCTCGAGAGGC CCTCGAGAGC CCTCGACACCC ACCTCACCT CAACCTCACCT	AACAACCTCCT AACAACCTCCT AACAACCTCCT 450 	CTCCGGCGCC. CTCCGGCGCC. CTCCGGCGCC. ********** 460 CAGCAACAAC CAGCAACAAC CAGCAACAAC ********* 520 TCTCCGCCTCC	A T C C C G C C G G C G A T C C C G C C G G C A T C C C G C C G G A A T C C C G C C G G A A T C C C G C C C G A A T C C C G C C C G G G C C C C T C T C C G G G G C C C C T C T C C G G G G C C C C T C T C C G G G G C C C C T C T C C G G G G C C C C T C T C C G G G G C C C C T C T C G G G G C C C C T C T C G G G G C C C C T C T C G G G G C C C C T C T C G G G G C C C C T C T C G G G G C C C C T C T C G G G G C C C C T C T C G G G G C C C C T C T C G G G G C C C C T C T C G G G G C C C C T C T C G G G G C C C C C T C T C G G G G C C C C C T C T C G G G G C C C C C C C C C C C
OsJ RPK1 Oruf RPK1 Clustal Consensus OsI RPK1 Oruf RPK1 Clustal Consensus OsI RPK1 OsJ RPK1 OsJ RPK1 Oruf RPK1	CCCCAGCT CCCCAGCT CCCCAGCT ******** GTCAGCAA GTCAGCAA GTCAGCAA GTCAGCAA ATCCCCTT ATCCCCTT	CCGCCTCCT CCGCCTCCT (CCGCCTCCT ********* 430 GCTCGCCGCC GCTCGCCGCC (CTCGCCGCC (CTCGCCGCC (CTCGCCGCC (CTCGCCGCC (CTCGCCGCC (CTCGCCGCC (CTCGCCGCC (CCCCCCCCC) (CCCCCCCCCCC) (CCCCCCCCCC	CTTCCTCCAGA CTTCCTCCAGA CTTCCTCCAGA 440 	AACAACCTCCT AACAACCTCCT AACAACCTCCT 450 	CTCCGGCGCC. CTCCGGCGCCC. CTCCGGCGCCC. ********** 460 CAGCAACAACC CAGCAACAACACC CAGCAACAACACC CAGCAACAACAAC **********	A T C C C G C C G G C G A T C C C G C C G G C G A T C C C G C C G G G G A T C C C G C C G C G G G 470 480
OsJ RPK1 Oruf RPK1 Clustal Consensus OsI RPK1 Oruf RPK1 Clustal Consensus OsI RPK1 OsJ RPK1 OsJ RPK1 Oruf RPK1	CCCCAGCT CCCCAGCT CCCCAGCT ******** GTCAGCAA GTCAGCAA GTCAGCAA GTCAGCAA ******** ATCCCCTT ATCCCCTT ATCCCCTT ATCCCCTT CCCCCT CTCCCGG CTCTCCGG	CCGCCTCCT CCGCCTCCT CCGCCTCCT ********* 430 	440 440 CTTCCAGAGAGGC CTTCGAGAGAGGC CCTCGAGAGAGGC CCTCGAGAGGGC CCTCGAGAGGGC CCTCGAGAGGC CCTCGACAGAGGC CAACCTCACCT CAACCTCACCT CAACCTCACCT 560 CACCATCACCA	AACAACCTCCT AACAACCTCCT AACAACCTCCT 450 	CTCCGGCGCC. CTCCGGCGCC. CTCCGGCGCC. ***********************************	A T C C C G C C G G C G A T C C C G C C G G C G A T C C C G C C G G G G A T C C C G C C G C G G G 470 480
OsJ RPK1 Oruf RPK1 Clustal Consensus OsJ RPK1 Oruf RPK1 Oruf RPK1 Clustal Consensus OsJ RPK1 Oruf RPK1 Oruf RPK1 Oruf RPK1 Oruf RPK1 OsJ RPK1 OsJ RPK1 Oruf RPK1 Oruf RPK1	CCCCAGCT CCCCAGCT CCCCAGCT ******** GTCAGCAA GTCAGCAA GTCAGCAA GTCAGCAA ******** ATCCCCTT ATCCCCTT ATCCCCTT ******** CTCTCCGG CTCTCCGG CTCTCCGG ******	CCGCCTCCT(CCGCCTCCT(********************************	CTTCCTCCAGA CTTCCTCCAGA CTTCCTCCAGA 440 	AACAACCTCCT AACAACCTCCT AACAACCTCCT 450 	CTCCGGCGCC. CTCCGGCGCC. CTCCGGCGCC. ********** 460 CAGCAACAACC CAGCAACAACACC CAGCAACAACACC **********	ATCCCGCCGGCG ATCCCGCCGGCG ATCCCGCCGGCG ATCCCGCCGGCG ATCCCGCCGCGGAG ************************************
OsJ RPK1 Oruf RPK1 Clustal Consensus OsJ RPK1 Oruf RPK1 Oruf RPK1 Clustal Consensus OsJ RPK1 Oruf RPK1 Oruf RPK1 Oruf RPK1 Oruf RPK1 OsJ RPK1 OsJ RPK1 Oruf RPK1 Oruf RPK1	CCCCAGCT CCCCAGCT CCCCAGCT ******** GTCAGCAA GTCAGCAA GTCAGCAA ATCCCCTT ATCCCCTT ATCCCCTT ATCCCCTT ******* CTCTCCGG CTCTCCGG CTCTCCGG CTCTCCGG	CCGCCTCCT CCGCCTCCT CCGCCTCCT ********* 430 GCTCGCCGCC GCTCGCCGCC CCCCCCCC ********* 490 CACGCTCAAC CACGCTCAAC CACGCTCAAC CACGCTCAAC CACGCTCAAC GAACATCCCC GAACATCCCC GAACATCCCC GAACATCCCC GAACATCCCC	CTTCCTCCAGA CTTCCTCCAGA CTTCCTCCAGA CTTCCTCCAGA 440 	AACAACCTCCT AACAACCTCCT AACAACCTCCT 450 	CTCCGGCGCC. CTCCGGCGCC. CTCCGGCGCC. CTCCGGCGCC. 460 	ATCCCGCCGGCG ATCCCGCCGGCG ATCCCGCCGGGGG ATCCCGCCGGGGGG ATCCCGCCGGGGGGG ATCCCGCCGCGGGGGG 470 480
OsJ RPK1 Oruf RPK1 Clustal Consensus OsJ RPK1 Oruf RPK1 Oruf RPK1 Clustal Consensus OsJ RPK1 Oruf RPK1 Clustal Consensus OsJ RPK1 Oruf RPK1 Oruf RPK1 Clustal Consensus OsJ RPK1 Oruf RPK1 Oruf RPK1 Oruf RPK1 Oruf RPK1 Oruf RPK1 Oruf RPK1 Oruf RPK1 Oruf RPK1	CCCCAGCT CCCCAGCT CCCCAGCT ******** GTCAGCAA GTCAGCAA GTCAGCAA ATCCCCTT ATCCCCTT ATCCCCTT ATCCCCTT ******* CTCTCCGG CTCTCCGG CTCTCCGG CTCTCCGG	CCGCCTCCT CCGCCTCCT CCGCCTCCT ********* 430 GCTCGCCGCC GCTCGCCGCC CCCCCCCC ********* 490 CACGCTCAAC CACGCTCAAC CACGCTCAAC CACGCTCAAC CACGCTCAAC GAACATCCCC GAACATCCCC GAACATCCCC GAACATCCCC GAACATCCCC	CTTCCTCCAGA CTTCCTCCAGA CTTCCTCCAGA CTTCCTCCAGA 440 	AACAACCTCCT AACAACCTCCT AACAACCTCCT 450 	CTCCGGCGCC. CTCCGGCGCC. CTCCGGCGCC. CTCCGGCGCC. 460 	ATCCCGCCGGCG ATCCCGCCGGCG ATCCCGCCGGGGG ATCCCGCCGGGGGG ATCCCGCCGGGGGGG ATCCCGCCGCGGGGGG 470 480

Od. BPR1 TGGCCGTCGCCCGGGGTAAGCCCCGCACCTCCCCGGGGCGGCGCCCCTCCTCCAAGAAG Od_BPR1 TGGCCGTCGCCCGGGGTAAGCCCCGCCACCTCCCCGGGGCGGCGCGCGC		730	740 75	0 760	770 780
Odl PERLI CGGAGGCTCTCTGEGCGCCGCATTCGCCGCGCGATCGTCGTCGTGGCGCCGC Odl PERLI CGGAGGCTCTCTGCGCGCCGCATTCGCCGCGGCATCGTCGTGGCGCGCGC	OsJ_RPK1 Oruf_RPK1	TCGCCGTCGCCCGG TCGCCGTCGCCCGG TCGCCGTCGCCCGG	GGTGAGCCCCGCCGAC GGTGAGCCCCGCCGAC GGTGAGCCCCGCCGAC	GTCCCCGGGGCGGCA GTCCCCGGGGCGGCA GTCCCCGGGGCGCGCA	TCCTCCTCCAAGAAG TCCTCCTCCAAGAAG TCCTCCTCCAAGAAG
OIL PERL OUL PERL	OsJ_RPK1 Onuf_RPK1	CGGAGGCTCTCTGG CGGAGGCTCTCCGG	CGCCGCAATCGCCGGC CGCCGCAATCGCCGGC CGCCGCAATCGCAGGC	ATCGTCGTGGGCGCCC ATCGTCGTGGGCGCCC ATCGTCGTGGGCGCCC	GTCGTGCTGGCGCTG GTCGTGCTGGCGCTG
OLI ERKI OLI ERKI	OsJ_RPK1 Oruf_RPK1	CTCCTCCTCGTCGC CTCCTCCTCGTCGC	CGCCGTGCTCTGCGCG CGCCGTGCTCTGCGCG	GTGTCCAAGCGGCGGG GTGTCCAAGCGGCGGG GTGTCCAAGCGGCGGG	CGAGGCGCCAGCGAG CGAGGCGCCAGCGAG
Od. RPK1 CCCCCGCCGGGGTCCGGCGAGGGGACGGGCATGACGTCGTCGTCCAAGGAGGACATGGGC Omit PRK1 CCCCCGCCGGGGTCCGGCGAGGGGACGGGCATGACGTCGTCGTCCAAGGAGGACATGGGC Clustal Consensus 1880 1640 1650 1660 1675 1680 Od. RPK1 GCCGCGGTCCGGCGGCGCGCGGCGGCGGCGGGGGGGGGG	OsJ_RPK1 Onuf_RPK1	GGACCGAAGAGCAC GGACCGAAGAGCAC GGACCGAAGAGCAC GGACCGAAGAGCAC	GACGGCGGCGGCGGCGGCG GACGGCGGCGGCGGCGGCG GACGGCGGCGGCGGCGGCG	GGTGCGGGCGCCGCCC GGTGCGGGGCGCCGCCC GGTGCGGGCGCCGCCC	GCCGCGAGAGGTGTC GCCGCGAGAGGTGTC GCCGCGAGAGGTGTC GCCGCGAGAGGTGTC
Osl RPK1 Osl RPK1 Osl RPK1 Chustal Consensus GGCGCGTCCGGGTCGGCCGCGGCGGCGGGGGGGGGGGG	OsJ_RPK1 Oruf_RPK1	CCCCCGCCGGGGTC CCCCCGCCGGGGTC CCCCCGCCGGGGTC	 CGGCGAGGGGACGGGC CGGCGAGGGGACGGGC CGGCGAGGGGACGGGC	ATGACGTCGTCGTCCA ATGACGTCGTCGTCCA ATGACGTCGTCGTCCA	AAGGAGGACATGGGC AAGGAGGACATGGGC AAGGAGGACATGGGC
Osl RPK1 GT GT T C GT GG GG A A GG GG C C GG GT A C A G C T T C G A C C T G G A GG A G	OsJ_RPK1 Oruf_RPK1	GGCGCGTCCGGGTC GGCGCGTCCGGGTC GGCGCGTCCGGGTC GGCGCGTCCGGGTC	GGCCGCGGCGGCGGCGGTG GGCCGCGGCGGCGGTG GGCCGCGGCGGCGGTG	GCGGCGGTGGCGGCGG GCGGCGGTGGCGGCGG GCGGCGGTGGCGGCGG GCGGCGGTGGCGGCGG	GAGCCGAGCAGGCTG GAGCCGAGCAGGCTG GAGCCGAGCAGGCTG GAGCCGAGCAGGCTG
Osl RPK1 GCGGAGGTGCTCGGGAAGGGGAGCGTGGGGACGTCGTACAAGGCGGTGCTGGAGGAAGGG OsJ RPK1 GCGGAGGTGCTCGGGAAGGGGAGCGTGGGGACGTCGTACAAGGCGGTGCTGGAGGAAGGG Omf RPK1 GCGGAGGTGCTCGGGAAGGGGAGCGTGGGGACGTCGTACAAGGCGGTGCTGGAGGAAGGG Clustal Consensus 1210 1220 1240 1250 1260 Osl RPK1 ACGACGGTGGTGGTGAAGCGGCGCGCGCGCGCGCGGGGCGCGGAGTTCGACGCC ACGACGGTGGTGGTGAAGCGGCTCAAGGACGTGGCGGTGGCGGCGGCGGAGTTCGACGCC Image: Consensus	OsJ_RPK1 Oruf_RPK1	GTGTTCGTGGGGAA GTGTTCGTGGGGGAA	GGGGGCCGGGTACAGC GGGGGCCGGGTACAGC	TTCGACCTGGAGGACC	CTGCTGCGGGGGGTCG CTGCTGCGGGGCGTCG
Osl RPK1 ACGACGGTGGTGGTGGAGCGGCGCGCGCGCGCGGCGGCGCGCGGCGCGCGC	OsJ_RPK1 Oruf_RPK1	GCGGAGGTGCTCGG GCGGAGGTGCTCGG	GAAGGGGAGCGTGGGG GAAGGGGAGCGTGGGG	ACGTCGTACAAGGCGC ACGTCGTACAAGGCGC	GTGCTGGAGGAAGGG GTGCTGGAGGAAGGG
OsLRPK1 CACATGGACGCGCTCGGCAAGGTGGAGCACCGCAACGTCCTCCCCGTCCGCGCCTACTAC OsLRPK1 CACATGGACGCGCTCGGCAAGGTGGAGCACCGCAACGTCCTCCCCGTCCGCGCCTACTAC	OsJ_RPK1 Oruf_RPK1	ACGACGGTGGTGGT	 GAAGCGGCTCAAGGAC GAAGCGGCTCAAGGAC GAAGCGGCTCAAGGAC	GTGGCGGTGGCGCGGG GTGGCGGTGGCGCGGG	CGCGAGTTCGACGCC CGCGAGTTCGACGCC
Omf_RPK1 CACATGGACGCGCTCGGCAAGGTGGAGCACCGCAACGTCCTCCCCGTCCGCGCCTACTAC Clustal Consensus ************************************	OsJ_RPK1 Oruf_RPK1	CACATGGACGCGCT CACATGGACGCGCT CACATGGACGCGCT	CGGCAAGGTGGAGCAC CGGCAAGGTGGAGCAC	CGCAACGTCCTCCCCC CGCAACGTCCTCCCCC	GTCCGCGCCTACTAC GTCCGCGCCTACTAC
1330 1340 1350 1360 1370 1380 OsL RPK1 TTCTCCAAGGACGAGAAGCTCCTCGGTCTTCGACTACCT TCCCCAACGGCAGCCTCTCCGCC Image: Construction of the constructio	OsJ_RPK1 Onuf_RPK1	TTCTCCAAGGACGA TTCTCCAAGGACGA TTCTCCAAGGACGA	GAAGCTCCTCGTCTTC GAAGCTCCTCGTCTTC GAAGCTCCTCGTCTTC	GACTACCTTCCCAACC GACTACCTCCCCAACC GACTACCTCCCCAACC	GGCAGCCTCTCCGCC GGCAGCCTCTCCGCC GGCAGCCTCTCCGCC GGCAGCCTCTCCGCC
1390 1400 1410 1420 1430 1440 OsL RPK1 A TGCTCCACGGGAGCCGGGGGGTCCGGCAAGACGCCGCTGGACTGGGACGCGCGAGATGCGG OsL RPK1 A TGCTCCACGGGAGCCGGGGGTCCGGCAAGACGCCGCTGGACTGGGACGCGCGGGATGCGGG OsL RPK1 A TGCTCCACGGGAGCCGGGGGTCCGGCAAGACGCCGCTGGACTGGGACGCGCGGGATGCGGG A TGCTCCACGGGAGCCGGGGGTCCGGCAAGACGCCGCTGGACTGGGACGCGCGGATGCGG Omf RPK1 A TGCTCCACGGGAGCCGGGGGTCCGGCAAGACGCCGCTGGACTGGGACGCGCGGGATGCGG A TGCTCCACGGGAGCCGGGGGTCCGGCAAGACGCCGCTGGACTGGGACGCGCGGGATGCGGG	OsJ_RPK1	ATGCTCCACGGGAG ATGCTCCACGGGAG	CCGGGGGTCCGGCAAG	ACGCCGCTGGACTGG ACGCCGCTGGACTGG	GACGCGCAGATGCGG GACGCGCGGGATGCGG

	1450	1460 1470	1480	1490 1500
OsI_RPK1 OsJ_RPK1 Oruf_RPK1 Clustal Consensus	TCGGCGCTGTCGGCGGCG TCGGCGCTGTCGGCGGCG	CGCGGGCTGGCGCACCTG CGCGGGCTGGCGCACCTG CGCGGGCTGGCGCGCCTG	CACACGGTGCAC CACACGGTGCAC CACACGGTGCAC CACACGGTGCAC	AGCCTCGTGCAC AGCCTCGTGCAC AGCCTCGTGCAC
	1510	1520 1530	1540	1550 1560
OsI_RPK1 OsJ_RPK1 Oruf_RPK1 Clustal Consensus	GGCAACGTCAAGTCCTCC GGCAACGTCAAGTCCTCC	CAACGTCCTGCTCCGGCCG CAACGTCCTGCTCCGGCCG CAACGTCCTGCTCCGGCCG	GACGCCGACGCG GACGCCGACGCG GACGCCGACGCG	GCGGCGCTCTCG GCGGCGCTCTCG
	1570	1580 1590	1600	1610 1620
OsI_RPK1 OsJ_RPK1 Oruf_RPK1 Clustal Consensus	GACTTCTGCCTCCACCCC GACTTCTGCCTCCACCCC	GATCTTCGCGCCGTCGTCG GATCTTCGCGCCGTCGTCG GATCTTCGCGCCGTCGTCG	GCGCGCCCGGGC GCGCGCCCGGGC GCGCGCCCGGGC	GCCGGCGGGTAC GCCGGCGGGTAC
	1630	1640 1650	1660	1670 1680 .
OsI_RPK1 OsJ_RPK1 Oruf_RPK1 Clustal Consensus	CGCGCGCCGGAGGTGGTG CGCGCGCCGGAGGTGGTG CGCGCGCCGGAGGTGGTG	3 G A C A C G C G G C G G C C G A C G 3 G A C A C G C G G C G G C C G A C G 3 G A C A C G C G G C G G C C G A C G 4 F A F A F A F A F A F A F A F A F A F	GTACAAGGCGGAC GTACAAGGCGGAC GTACAAGGCGGAC	GTCTACTCCCTC GTCTACTCCCTC GTCTACTCCCTC
	1690	1700 1710	1720	1730 1740
OsI_RPK1 OsJ_RPK1 Oruf_RPK1 Clustal Consensus	GGCGTGCTCCTCCTAGAG	GCTCCTCACCGGCAAGTCA GCTCCTCACCGGCAAGTCG GCTCCTCACCGGCAAGTCG	CCGACGCACGCG	TCCCTGGAAGGG TCCCTGGAAGGG
	1750	1760 1770	1780	1790 1800
OsI_RPK1 OsJ_RPK1 Oruf_RPK1 Clustal Consensus	GACGGCACGCTGGACCTC GACGGCACGCTGGACCTC	C C A C G G T G G G T G C A G T C C C C A C G G T G G G T G C A G T C C C C A C G G T G G G T G C A G T C C	GTGGTGCGGGAG GTGGTGCGGGAG GTGGTGCGGGAG	GAGTGGACGGCG GAGTGGACGGCG GAGTGGACGGCG
	1810	1820 1830	1840	1850 1860
OsI_RPK1 OsJ_RPK1 Oruf_RPK1 Clustal Consensus	GAGGTGTTCGACGTCGAC GAGGTGTTCGACGTCGAC	GCTGGTGCGGCTCGGCGCG GCTGGTGCGGCTCGGCGCG GCTGGTGCGGCTCGGCGCG GCTGGTGCGGCTCGGCGCG	GAGCGCGGGAGGAG GAGCGCGGGAGGAG	GAGATGGTGGCG GAGATGGTGGCG
	1870	1880 1890	1900	1910 1920 .
OsI_RPK1 OsJ_RPK1 Oruf_RPK1 Clustal Consensus	CTGCTCCAGGTGGCCATC CTGCTCCAGGTGGCCATC	GCGTGCGTCGCCACCGTG GCGTGCGTCGCCACCGTG GCGTGCGTCGCCACCGTG	CCCGACGCGCGG CCCGACGCGCGG CCCGACGCGCGG	CCCGACGCCCCC
	1930 	1940 1950 	1960 • • • • • • • • • • •	1970 1980 .
OsI_RPK1 OsJ_RPK1 Oruf_RPK1 Clustal Consensus	GACGTGGTCAGGATGATC GACGTGGTCAGGATGATC	C G A G G A G A T C G G C G G C G G C C G A G G A G A T C G G C G G C G G C C G A G G A G A T C G G C G G C G G C C G A G G A G A T C G G C G G C G G C C A G G A G A T C G G C G G C G G C C A G G A G A G A T C G G C G G C G G C C A G G A G A G A T C G G C G G C G G C G G C C G A G G A G A T C G G C G G C G G C G G C C G A G G A G A T C G G C G G C G G C G G C C G A G G A G A T C G G C G G C G G C C G A G G A G A T C G G C G G C G G C C G A G G A G A T C G G C G G C G G C C G A G G A G A T C G G C G G C G G C C G A G G A G A T C G G C G G C G G C C G A G G A G A T C G G C G G C G G C C G A G G A G A T C G G C G G C G G C C G A G G A G A T C G G C G G C G G C C G A G G A G A T C G G C G G C G G C G G C C G A G G A G A T C G G C G G C G G C G G C C G A G G A G A T C G G C G G C G G C G G C C G A G G A G A T C G G C G G C G G C G G C G G C C G A G G A G A T C G G C G C G G C G G C G G C G G C G G C G G C G G C G G C G G C G G C G G C G G C G G C G	CACGGCCGGACG CACGGCCGGACG	A C G A C G G A G G A G A C G A C G G A G G A G
	1990	2000 2010	2020	2030 2040
OsI_RPK1 OsJ_RPK1 Oruf_RPK1 Clustal Consensus	AGCGAGGAGGGGCGTGCGG AGCGAGGAGGGCGTGCGG	3 G G A A C G T C G G A G G A G G A G 3 G G A A C G T C G G A G G A G G A G 3 G G A A C G T C G G A G G A G G A G 4 G G A A C G T C G G A G G A G G A G 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4	CGGTCGCGGGGC	ACGCCTCCGGCC ACGCCTCCGGCC
	2050			
OsI_RPK1 OsJ_RPK1 Oruf_RPK1 Clustal Consensus	GCCCCGACGCCGTGA GCCCCGACGCCGTGA GCCCCGACGCCGTGA			

Figure 4.3: ClustalW ORF sequence alignment of *Oruf_RPK1*, *OsI_RPK1* and *OsJ_RPK1* (gi:18677097). Boxes indicate SNPs.

The genotype variation of putative *RPK1* of *Oryza rufipogon*, *Oryza sativa* ssp. *indica* cv. MR219, BC₂F₇ line 7 and BC₂F₇ line 23 were identified using a direct sequencing method. Purified RT-PCR products of putative *RPK1* were sequenced directly. The sequencing data of *Oryza rufipogon*, *Oryza sativa* ssp. *indica* cv. MR219, BC₂F₇ line 7 and BC₂F₇ line 23 shown in Figure 4.4. Putative *RPK1* of *Oryza rufipogon* and BC₂F₇ line 23 were identical to each other; whereas, *Oryza sativa* ssp. *indica* cv. MR219 and BC₂F₇ line 7 were identical to each other. Sequencing chromatogram of putative *RPK1* revealed that two bases of cytosine (C) and adenine (A) at position 791 and 807 respectively in *Oryza rufipogon* and BC₂F₇ line 23 were exchanged completely to thymine (T) and cytosine (C) in *Oryza sativa* ssp. *indica* cv. MR219 and BC₂F₇ line 7. Thus, BC₂F₇ line 23 is homozygous for the putative *RPK1* allele from *Oryza rufipogon*; whereas, BC₂F₇ line 7 is homozygous for the putative *RPK1* allele from *Oryza sativa* ssp. *indica* cv. MR219.

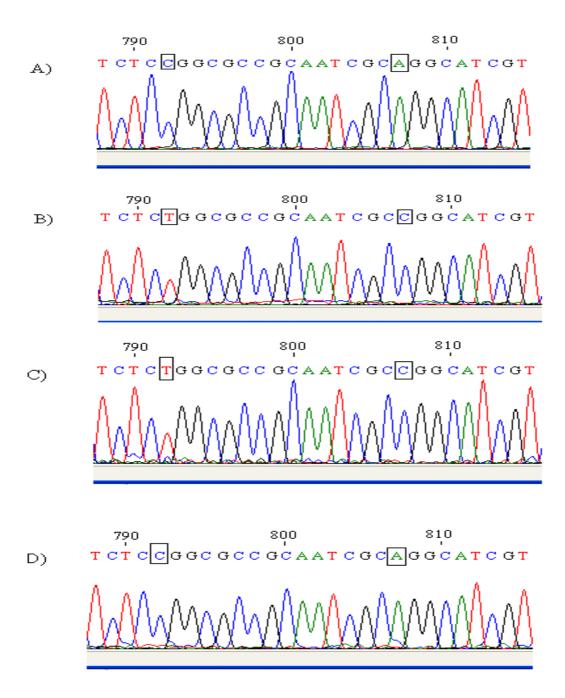


Figure 4.4: Base substitution of partial putative *RPK1* sequences revealed by sequencing from (A) *Oryza rufipogon*, (B) *Oryza sativa* ssp. *indica* cv. MR219, (C) BC_2F_7 line 7 and (D) BC_2F_7 line 23. Base substitutions are boxed. Sequencing chromatogram of *Oryza rufipogon* and BC_2F_7 line 23 were identical to each other; whereas, sequencing chromatogram of *Oryza sativa* ssp. *indica* cv. MR219 and BC_2F_7 line 7 were identical to each other.

The cDNA sequences of Oruf RPK1 (Figure 4.5) and OsI RPK1 (Figure 4.6) were successfully amplified and cloned into cloning vector pJET1.2/blunt (CloneJETTM PCR Cloning Kit, Fermentas, Lithuania) by RACE. The nucleotide sequence obtained for the Oruf RPK1 cDNA is 2,832 bp long with a 404 bp 5' UTR and a 373 bp 3' UTR. The nucleotide sequence of the OsI RPK1 cDNA is 2,576 bp long with a 103 bp 5' UTR and a 418 bp 3' UTR. The sequences of Oruf RPK1 and OsI RPK1 were compared with OsJ RPK1 (gi:18677097) as shown in Table 4.1. The sequences from these three rice types have some similar features: Their ORF lengths were 2,055 bp with 73 % GC content. Genomic DNA sequence of Oruf RPK1 (Song et al., 2009), Oryza sativa ssp. indica variety 9311 putative RPK1 (gi:57015219) and OsJ RPK1 (gi:18677097) were included in the analysis. The 5' and 3' splice sites of the introns were identified according to GT-AG rules (Sheth et al., 2006; Figure 4.7). Two exons and one intron were identified. The putative RPK1 of the three different species are interrupted by various lengths of intron sequences: The length of the Oruf RPK1 intron is 3,752 bp, while the OsI RPK1 (gi:57015219) intron is 3,345 bp and the OsJ RPK1 (gi:18677097) intron is 3,357 bp.

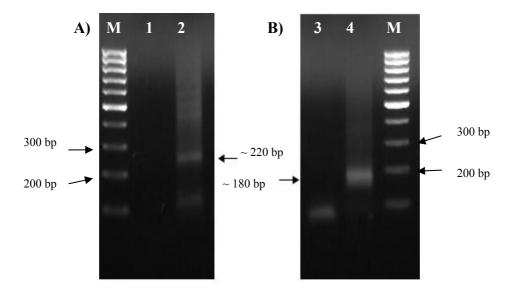


Figure 4.5: The inner 5' and 3' RLM-RACE PCR amplification of *Oruf_RPK1*. (A) The expected size of inner 5' RLM-RACE PCR *Oruf_RPK1* is around 220 bp. (B) Inner 3' RLM-RACE PCR *Oruf_RPK1* give a band of the expected size of around 180 bp. Lane M: 100 bp marker; Lane 1 and 3: Minus-template control; Lane 2: Outer 5' RLM-RACE PCR of *Oryza rufipogon*; Lane 4: Outer 3' RLM-RACE PCR of *Oryza rufipogon*.

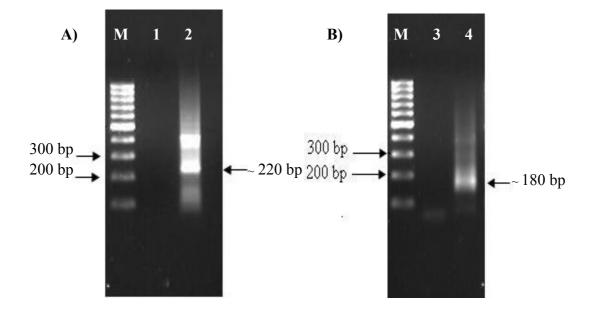


Figure 4.6: The inner 5' and 3' RLM-RACE PCR amplification of *OsI_RPK1*. (A) The expected size of inner 5' RLM-RACE PCR *OsI_RPK1* is around 220 bp. (B) Inner 3' RLM-RACE PCR *OsI_RPK1* give a band of the expected size of around 180 bp. Lane M: 100 bp marker; Lane 1 and 3: Minus-template control; Lane 2: Outer 5' RLM-RACE PCR of *Oryza sativa* ssp. *indica* cv. MR219; Lane 4: Outer 3' RLM-RACE PCR of *Oryza sativa* ssp. *indica* cv. MR219.

Feature	Oryza rufipogon	<i>Oryza sativa</i> ssp. <i>indica</i> cv. MR219	Oryza sativa ssp. japonica
Genomic DNA	6,584 bp ^a	5,400 bp ^b	5,411 bp ^c
cDNA	2,832 bp	2,576 bp	2,535 bp ^c
5' UTR	404 bp	103 bp	120 bp
3' UTR	386 bp	418 bp	359 bp
ORF	2,055 bp	2,055 bp	2,055 bp
GC content	73 %	73 %	73 %
Intron	1	1	1
	(3,752 bp) ^a	(3,345 bp) ^b	(3,357 bp) ^c
Exon	2	2	2
	(1,390 bp + 665 bp)	(1,390 bp + 665 bp)	(1,390 bp + 665 bp)
Amino acid	684 residues	684 residues	684 residues

Table 4.1: Comparative sequence analysis of *Oruf_RPK1*, *OsI_RPK1* and *OsJ_RPK1* (gi:18677097) cDNA.

^a The gene sequence for genomic DNA and predicted intronic region were retrieved from previously reported *Oruf_RPK1* by Song *et al.* (2009).

^b The gene sequence for genomic DNA and predicted intronic region were retrieved

from Oryza sativa ssp. indica variety 9311 putative RPK1 (gi:57015219).

^c The gene sequence for genomic DNA and cDNA were retrieved from OsJ_RPK1

(gi:18677097).

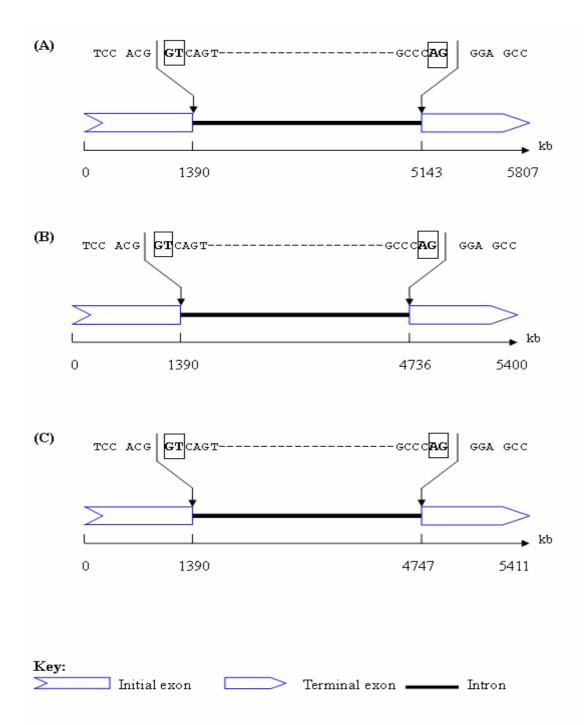


Figure 4.7: Predicted 5' and 3' splice sites of putative *RPK1*. Genomic DNA sequence of (A) *Oruf_RPK1* (Song *et al.*, 2009), (B) *Oryza sativa* ssp. *indica* putative *RPK1* (gi:57015219) and (C) *OsJ_RPK1* (gi:18677097) were aligned with cDNA sequence of *Oruf_RPK1*, *OsI_RPK1* and *OsJ_RPK1* (gi:18677097) respectively. Donor and acceptor sites are boxed.

4.1.2 Structure Analysis of Putative RPK1

The Oruf RPK1, OsI RPK1 and OsJ RPK1 (gi:18677097) ORF sequences were predicted to encode a polypeptide of 684 amino acids using GenScan (Appendix B). Amino acid sequences of Oruf RPK1 and OsI RPK1 showed 99 % identity with OsJ RPK1 (gi:18677097) as shown in Figure 4.8. The Simple Modular Architecture Research Tool (SMART) analysis showed that the Oruf RPK1, OsI RPK1 and OsJ RPK1 (gi:18677097) contained seven conserved domains, namely signal peptide, leucine-rich repeat N terminus (LRRNT 2), LRR motif, transmembrane helix and serine/threonine protein kinase. The LRRNT 2 and LRR motif belong to extracellular receptor domains; whereas, serine/threonine protein kinase belongs to cytoplasmic protein kinase domains (Figure 4.8). The SMART is used as a relational database management system (RDBMS; http://www.PostgreSQL.org), which is able to cross-references to other domain databases. The signal peptide was predicted using SignalP 3.0 Server (http://www.cbs.dtu.dk/services/SignalP; Nielsen et al., 1997; Bendtsen et al., 2004; Emanuelsson et al., 2007). The transmembrane helix was predicted using Transmembrane Protein Topology with a Hidden Markov Model (TMHMM) Server 2.0 (http://www.cbs.dtu.dk/services/TMHMM/; Krogh et al., 2001). The LRRNT 2 was predicted using protein domain families (PFAM) database (Bateman et al., 2004). The remaining conserved domains were identified using Basic Local Alignment Search Tool (BLAST) database (Karlin and Altschul, 1993). Interestingly, amino acid sequence of signal peptide, LRRNT 2, LRR1, LRR2 and LRR3 showed 100 % identity among these three species. However, two amino acid

	Signal peptide Pfam: LRRNT 2
OsI RPK1	MAGVVTRAVAAAVLVVVVVAAAELVAAEPPPSERSALLAFLAATPHERRLGWNSSTSAC
OsJ RPK1	MAGVVTRAVAAAVLVVVVVAAAELVAAEPPPSERSALLAFLAATPHERRLGWNSSTSAC
Oruf RPK1	MAGVVTRAVAAAVLVVVVVAAAELVAAEPPPSERSALLAFLAATPHERRLGWNSSTSAC

	LRR 1
OSI_RPK1	GWVGVTCDAGNATVVQVRLPGVGLIGAIPPGTLGRLTNLQVLSLRSNRILGGIPDDVLQL
OsJ RPK1	GWVGVTCDAGNATVVQVRLPGVGLIGAIPPGTLGRLTNLQVLSLRSNRILGGIPDDVLQL
Oruf RPK1	GWVGVTCDAGNATVVQVRLPGVGLIGAIPPGTLGRLTNLQVLSLRSNRILGGIPDDVLQL

	LRR 2 LRR 3
OsI RPK1	POLRLLFLONNLLSGAIPEAVSKLAALERLVLSSNNLSGPIPFTLNNLTSLRALRLDGNK
OsJ RPK1	POLRLLFLONNLLSGAIPEAVSKLAALERLVLSSNNLSGPIPFTLNNLTSLRALRLDGNK
Oruf_RPK1	PQLRLLFLQNNLLSGAIPFEVSKLAALERLVLSSNNLSGPIPFTLNNLTSLRALRLDGNK
0-T	LSGNIPSISIQSLAVFNVSDNNLNGSIPASLARFPAEDFAGNLQLCGSPLPPCKSFFPSP
OsI_RPK1 OsJ RPK1	LSGNIPSISIQSLAVENVSDNNLNGSIPASLAREPALDEAGNLOLCGSPLPPCKSEEPSP
Oruf_RPK1	LSGNIPSISIQSLOVFNVSDNNLNGSIFASDALFFALDFAGNLQLCGSPDFFCKSFFPSP LSGNIPSISIQSLAVFNVSDNNLNGSIFASDALFFALDFAGNLQLCGSPDFPCKSFFPSP
	Transmembrane
OsI RPK1	SPSPGVSPADVPGAASSSKKRRLSGAAIAGIVVGAVVLALLLLVAAVLCAVSKRRRGASE
OsJ_RPK1	SPSPGVSPADVPGAASSSKKRRLSGAAIAGIVVGAVVLALLLLVAAVLCAVSKRRRGASE
Oruf_RPK1	SPSPGVSPADVPGAASSSKKRRLSGAAIAGIVVGAVVLALLLLVAAVLCAVSKRRRGASE

OsI_RPK1	GPKSTTAAAAGAGAAAARGVPPPGSGEGTGMTSSSKEDMGGASGSAAAAVAAVAAEPSRL
OsJ RPK1	GPKSTTAAAAGAGAAAARGVPPPGSGEGTGMTSSSKEDMGGASGSAAAAVAAVAAEPSRL
Oruf RPK1	GPKSTTAAAAGAGAAAARGVPPPGSGEGTGMTSSSKEDMGGASGSAAAAVAAVAAEPSRL
511	***************************************
	Serine / threonine kinase protein domain
OSI_RPK1	VFVGKGAGYSFDLEDLLRASAEVLGKGSVGTSYKAVLEEGTTVVVKRLKDVAVARREFDA
OsJ_RPK1	VFVGKGAGYSFDLEDLLRASAEVLGKGSVGTSYKAVLEEGTTVVVKRLKDVAVARREFDA
OSJ_RPK1 Oruf_RPK1	VFVGKGAGYSFDLEDLLRASAEVLGKGSVGTSYKAVLEEGTTVVVKRLKDVAVARREFDA
Oruf_RPK1	VFVGKGAGYSFDLEDLLRASAEVLGKGSVGTSYKAVLEEGTTVVVKRLKDVAVARREFDA
	VFVGKGAGYSFDLEDLLRASAEVLGKGSVGTSYKAVLEEGTTVVVKRLKDVAVARREFDA ************************************
Oruf_RPK1 OsI_RPK1	VFVGKGAGYSFDLEDLLRASAEVLGKGSVGTSYKAVLEEGTTVVVKRLKDVAVARREFDA
Oruf_RPK1 Osl_RPK1 OsJ_RPK1	VFVGKGAGYSFDLEDLLRASAEVLGKGSVGTSYKAVLEEGTTVVVKRLKDVAVARREFDA ************************************
Oruf_RPK1 OsI_RPK1 OsJ_RPK1 Oruf_RPK1	VFVGKGAGYSFDLEDLLRASAEVLGKGSVGTSYKAVLEEGTTVVVKRLKDVAVARREFDA ************************************
Oruf_RPK1 Osl_RPK1 OsJ_RPK1 Oruf_RPK1 Osl_RPK1	VFVGKGAGYSFDLEDLLRASAEVLGKGSVGTSYKAVLEEGTTVVVKRLKDVAVARREFDA
Oruf_RPK1 Osl_RPK1 OsJ_RPK1 Oruf_RPK1 Osl_RPK1 OsJ_RPK1	VFVGKGAGYSFDLEDLLRASAEVLGKGSVGTSYKAVLEEGTTVVVKRLKDVAVARREFDA
Oruf_RPK1 Osl_RPK1 OsJ_RPK1 Oruf_RPK1 Osl_RPK1	VFVGKGAGYSFDLEDLLRASAEVLGKGSVGTSYKAVLEEGTTVVVKRLKDVAVARREFDA
Oruf_RPK1 Osl_RPK1 OsJ_RPK1 Oruf_RPK1 Osl_RPK1 OsJ_RPK1 Oruf_RPK1	VFVGKGAGYSFDLEDLLRASAEVLGKGSVGTSYKAVLEEGTTVVVKRLKDVAVARREFDA
Oruf_RPK1 OsI_RPK1 Oruf_RPK1 OsI_RPK1 OsJ_RPK1 Oruf_RPK1 Oruf_RPK1 OsI_RPK1	VFVGKGAGYSFDLEDLLRASAEVLGKGSVGTSYKAVLEEGTTVVVKRLKDVAVARREFDA
Oruf_RPK1 OSI_RPK1 OSJ_RPK1 Oruf_RPK1 OSJ_RPK1 Oruf_RPK1 Oruf_RPK1 OSJ_RPK1 OSJ_RPK1	VFVGKGAGYSFDLEDLLRASAEVLGKGSVGTSYKAVLEEGTTVVVKRLKDVAVARREFDA ************************************
Oruf_RPK1 OsI_RPK1 Oruf_RPK1 OsI_RPK1 OsJ_RPK1 Oruf_RPK1 Oruf_RPK1 OsI_RPK1	VFVGKGAGYSFDLEDLLRASAEVLGKGSVGTSYKAVLEEGTTVVVKRLKDVAVARREFDA
Oruf_RPK1 OSI_RPK1 OSJ_RPK1 Oruf_RPK1 OSJ_RPK1 Oruf_RPK1 Oruf_RPK1 OSJ_RPK1 Oruf_RPK1 Oruf_RPK1	VFVGKGAGYSFDLEDLLRASAEVLGKGSVGTSYKAVLEEGTTVVVKRLKDVAVARREFDA
Oruf_RPK1 OSI_RPK1 OSJ_RPK1 Oruf_RPK1 OSJ_RPK1 Oruf_RPK1 Oruf_RPK1 OSJ_RPK1 Oruf_RPK1 Oruf_RPK1 Oruf_RPK1	VFVGKGAGYSFDLEDLLRASAEVLGKGSVGTSYKAVLEEGTTVVVKRLKDVAVARREFDA
Oruf_RPK1 OsI_RPK1 OsJ_RPK1 Oruf_RPK1 OsJ_RPK1 Oruf_RPK1 Oruf_RPK1 OsJ_RPK1 Oruf_RPK1 Oruf_RPK1 OsJ_RPK1 OsJ_RPK1	VFVGKGAGYSFDLEDLLRASAEVLGKGSVGTSYKAVLEEGTTVVVKRLKDVAVARREFDA
Oruf_RPK1 OSI_RPK1 OSJ_RPK1 Oruf_RPK1 OSJ_RPK1 Oruf_RPK1 Oruf_RPK1 OSJ_RPK1 Oruf_RPK1 Oruf_RPK1 Oruf_RPK1	VFVGKGAGYSFDLEDLLRASAEVLGKGSVGTSYKAVLEEGTTVVVKRLKDVAVARREFDA
Oruf_RPK1 OsI_RPK1 OsJ_RPK1 Oruf_RPK1 OsJ_RPK1 OsJ_RPK1 Oruf_RPK1 OsJ_RPK1 Oruf_RPK1 OsJ_RPK1 OsJ_RPK1 OsJ_RPK1 Oruf_RPK1	VFVGKGAGYSFDLEDLLRASAEVLGKGSVGTSYKAVLEEGTTVVVKRLKDVAVARREFDA
Oruf_RPK1 OSI_RPK1 OSJ_RPK1 Oruf_RPK1 OSJ_RPK1 OSJ_RPK1 Oruf_RPK1 OSJ_RPK1 Oruf_RPK1 OSJ_RPK1 OSJ_RPK1 OSJ_RPK1 Oruf_RPK1 OSJ_RPK1 OSI_RPK1	VFVGKGAGYSFDLEDLLRASAEVLGKGSVGTSYKAVLEEGTTVVVKRLKDVAVARREFDA
Oruf_RPK1 OsI_RPK1 OsJ_RPK1 Oruf_RPK1 OsJ_RPK1 OsJ_RPK1 Oruf_RPK1 OsJ_RPK1 Oruf_RPK1 OsJ_RPK1 OsJ_RPK1 OsJ_RPK1 Oruf_RPK1	VFVGKGAGYSFDLEDLLRASAEVLGKGSVGTSYKAVLEEGTTVVVKRLKDVAVARREFDA
Oruf_RPK1 OSI_RPK1 OSJ_RPK1 Oruf_RPK1 OSJ_RPK1 OSJ_RPK1 Oruf_RPK1 OSJ_RPK1 OSJ_RPK1 OSJ_RPK1 OSJ_RPK1 Oruf_RPK1 OSJ_RPK1 OSJ_RPK1 OSJ_RPK1	VFVGKGAGYSFDLEDLLRASAEVLGKGSVGTSYKAVLEEGTTVVVKRLKDVAVARREFDA

Figure 4.8: ClustalW alignment of amino acid sequence and gene structure prediction of Oruf_RPK1, OsI_RPK1 and OsJ_RPK1 (gi:18677097). Functional domains identified by SMART analysis are indicated above the sequence. Amino acid substitutions are boxed.

substitutions were detected on the serine/threonine protein kinase domain. A total of five amino acid substitutions were all non-synonymous substitution identified from the amino acid sequences of Oruf_RPK1, OsI_RPK1 and OsJ_RPK1 (gi:18677097) as shown in Figure 4.8. Four amino acid substitutions were identified between Oruf_RPK1 and OsI_RPK1, and between Oruf_RPK1 and OsJ_RPK1 (gi:18677097) respectively, whilst, two amino acid substitutions were identified between OsI_RPK1 and OsJ_RPK1 (gi:18677097). The five amino acid substitutions were of either polar side chain to polar side chain or non-polar side chain to non-polar side chain.

4.2 Characterization and Identification of Full Length cDNA of Putative CLV1

4.2.1 Isolation of Full Length cDNA and Sequence Analysis of Putative CLV1

The Receptor-like Protein Kinase (RKN) used for primer design in this study is a patented sequence (WO/2000/004761) claimed to have abilities in increasing growth and yield in rice (Zhong *et al.*, 2000; Appendix B). However, the patent document did not name any species of rice. The amino acid sequences derived from this sequence showed high similarity (90 % identity) to the *Oryza sativa* ssp. *japonica* putative CLV1 (OsJ_CLV1; gi:125602183) on chromosome 8 (Figure 4.9). Primers designed from *OsJ_CLV1* (gi:125602183) successfully amplified bands of 402 bp from the cDNA of *Oryza rufipogon* and *Oryza sativa* ssp. *indica* cv. MR219 respectively (Figure 4.10). These two fragments isolated as gel extraction products were successfully cloned into pGEM-T Easy vector (Promega, USA).

	10		20	30	40	50 60
OsJ_CLV1 RKN Clustal Consensus	MPQPQCTTTA		PPPPPPTTS	MAAA	AVHVLLLLLP AVHVLLLLLP * * * * * * * * * * *	LATITSASSAPL LATITSASSAPL ***********
	70		80 r+	90 .	100	110 120
OsJ_CLV1 RKN Clustal Consensus	PLLALLSLRS PLLALLSLRS *******	S L G D P A G A L S L G D P A G A L * * * * * * * * * *	R - SWTYAAA. LRSWTYAAA. * * * * * * *	A S A G A T R S L A I A S A G A T R S L A I * * * * * * * * * * * * *	P	C D G A T G E V V G V D C D G A T G E V V G V D * * * * * * * * * * * * * * *
	130		140	150	160	170 180
OsJ_CLV1 RKN Clustal Consensus	L S R R N L S G T V L S R R N L S G T V * * * * * * * * * * * *	S P T A A R L L S S P T A A R L L S * * * * * * * * * * *		GNAFAGELPP GNGSP - ASSRI * *		LDVSHNFFNSTF LDVSHNFFNSTF **********
	190 		200	210	220	230 240 . [
OsJ_CLV1 RKN Clustal Consensus	P D G I A K L G S L P D G I A K L G G F * * * * * * * * *		FVGELPRGI FVGELPRGI		LGGSFFNGSI	PGEVGQLRRLRF PGEVGQLRRLRF *****
	250		260	270	280	290 300
OsJ_CLV1 RKN Clustal Consensus	LHLAGNALSGI LHLAGNRLSGI *****	RLPRELGEL RLPRELGEL * * * * * * * * * *	TSVEHLEIG TSVEHLEIG	YNAYDGGIPP YNAYDGGIP -	E F G KMA Q L R Y E F G KMA Q L R Y	LDIAAANVSGPL
	310		320	330	340	350 360
OsJ_CLV1 RKN Clustal Consensus	PPELGELTRLI PPELLGLTRLI ****		IGRRDPPRW		V S DNHL AGA I V S DNHL AGA I * * * * * * * * * * * *	PALG-ELTNLTT
	370		380	390	400	410 420
OsJ_CLV1 RKN Clustal Consensus	LNLMSNSLSG LNLMSNSLSG *****	T I P A A I G A L T I P A A I G A L * * * * * * * * * * *	PSLEVLQLW		SLGASRRLVR	LDVSTNSLSGPI
	430	i aa ah Basaan	440	450	460	470 480
OsJ_CLV1 RKN Clustal Consensus	PPGVCAGNRL PPGVCAGNRL	ARLILFDNR ARLILFDNR * * * * * * * * * *	FDSAIPASL. FDSAIPASL.	ADCSSLWRVRI ADCSSLWRVRI		PAGFGAIRNLTY PAGFGAIRNLTY ******
	490		500	510	520	530 540
OsJ_CLV1 RKN Clustal Consensus	MDLSSNSLTG MDLSSNSLTG *****		SPSLEYFNV		DMAWRGPKLQ DMAWRGPKLQ * * * * * * * * * * *	VFAASRCGLVGE
	550	1	560	570	580	590 600
OsJ_CLV1 RKN Clustal Consensus	L P A F G A T G C A I L P A F G A T G C A I * * * * * * * * * * * * *	NLYRLELAG NLYRLELAG *******	NALGGGIPG NALGGGIPG	DIGSCKRLVSI DIGSCKRLVSI	LRLQHNELTG LRLQHNELTG	E I PAA I AAL P S I E I PAA I AL P - S I * * * * * * * *
	610		620	630	640	650 660
OsJ_CLV1 RKN Clustal Consensus	TEVDLSWNAL TEVDLSG-TR *****	TGTVPPGFT SPAPSAGVH		V S F N H L A P A E I V S F N H L A P A E I * * * * * * * * * * * *	P S S D A G E R G S P S S D A G E R G S * * * * * * * * * * * *	PARHTAAMWVPA PRGTRRRCGCPP *
	670		680	690	700	710 720
OsJ_CLV1 RKN Clustal Consensus	VAVAFAGMVVI WRCAFAGMVVI	LAGTARWLQ LAGTARWLQ		A D A L G P G G A R I A D A R G P G G A R I * * *	H P D L V V G P W R H P D L V V G P W R * * * * * * * * * * * *	MTAFQRLSFTAD MTAFQRLSFTAD **********
	730		740	750	760 	770 780
OsJ_CLV1 RKN Clustal Consensus	DVARCVEGSDO DVPRCVEGSDO **		T V Y R A KMP N T V Y R A KMP N * * * * * * * * * *	GEVIAVKKLW		A P T E Q N Q K L R Q D A P T E Q N Q K L R Q D * * * * * * * * * * * * * * * *
	790 	n 1971 - Lasan	800 	810 .	820 .	830 840 · · · · · · · · · ·
OsJ_CLV1 RKN Clustal Consensus	SDGGGGGKRT SDGGGGGKRT *******	VAEVEVLGH VAEVEVLGH * * * * * * * * * *		L GWC T N G E S T M L GWC T N G E S T M * * * * * * * * * * * * * *	MLLYEYMPNG MLLYEYMPNG * * * * * * * * * * *	SLDELLHGAAAK SLDELLHARAKA * * * * * * * *

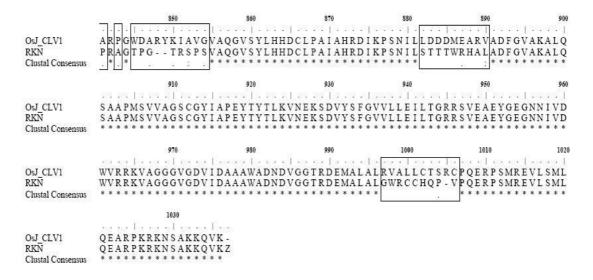


Figure 4.9: ClustalW amino acid sequence alignment of OsJ_CLV1 (gi:125602183) and RKN patent sequence (WO/2000/004761). The amino acid substitutions and deletions are boxed.

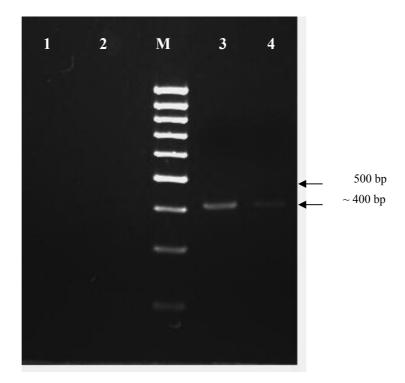


Figure 4.10: RT-PCR amplification of putative *CLV1*. The expected size of *Oruf_CLV1* and *OsI_CLV1* is around 400 bp. Lane M: 100 bp marker; Lane 1: DNase I treated total RNA of *Oryza rufipogon*; Lane 2: DNase I treated total RNA of *Oryza sativa sativa* ssp. *indica* cv. MR219; Lane 3: *Oryza rufipogon* cDNA; Lane 4: *Oryza sativa ssp. indica* cv. MR219 cDNA.

Sequencing results showed that the gene sequences of *Oryza rufipogon* putative *CLV1* (*Oruf_CLV1*) and *Oryza sativa* ssp. *indica* cv. MR219 putative *CLV1* (*OsI_CLV1*) showed 99 % and 100 % identity respectively with partial gene sequence of *OsJ_CLV1* (gi:125602183; 2,640-3,006) as shown in Figure 4.11. The partial gene sequences of *OsJ_CLV1* (gi:125602183; 2,640-3,006) and *OsI_CLV1* were identical. However, four SNPs were identified between *Oruf_CLV1* and *OsI_CLV1*, and between *Oruf_CLV1* and *OsJ_CLV1* (gi:125602183; 2,640-3,006) respectively. One of the four SNPs is a non-synonymous substitution; whereas, the remaining SNPs are synonymous substitutions.

Purified putative *CLV1* RT-PCR products from *Oryza rufipogon*, *Oryza sativa* ssp. *indica* cv. MR219, BC₂F₇ line 7 and BC₂F₇ line 23 were sequenced directly. The sequencing data is showed in Figure 4.12. Putative *CLV1* of *Oryza rufipogon* and BC₂F₇ line 23 were identical to each other; whereas, putative *CLV1* of *Oryza sativa* ssp. *indica* cv. MR219 and BC₂F₇ line 7 were identical to each other. Sequencing chromatogram of cytosine (C) and adenine (A) at position 2,710 and 2,724 respectively in *Oryza rufipogon* and BC₂F₇ line 23 were exchanged completely to thymine (T) and guanine (G) in *Oryza sativa* ssp. *indica* cv. MR219 and BC₂F₇ line 7 is homozygous for the putative *CLV1* allele from *Oryza rufipogon*; whereas, BC₂F₇ line 7 is homozygous for the putative *CLV1* allele from *Oryza sativa* ssp. *indica* cv. MR219.

	2650	2660	2670	2680	2690 2700
OsI CLV1				GAGCGATGTG	
Oruf CLV1		A	GTGAACGAGAA		TACAGCTTCGGT
OsJ_CLV1	GCACCAGAGTACACGT	ACACTCTAAAA	GTGAACGAGAA	GAGCGATGTG	TACAGCTTCGGT
Clustal Consensus			**********	*********	**********
	2710	2720	2730	2740	2750 2760
O-L CLV1		TCCTCACCCCA			
OsI_CLV1 Oruf CLV1	GTGGTACTATTGGAGA GTGGTACTACTGGAGA				
OsJ_CLV1	GTGGTACTATTGGAGA	TCCTGACGGGA	CGGCGGTCGGT	GGAGGCGGAG	TACGGGGAGGGG
Clustal Consensus	* * * * * * * * * * * * * * * *	* * * * * * * 🗌 * * *	*********	*********	* * * * * * * * * * * *
	2770	2780	2790	2800	2810 2820
0.1.07.14					
OsI_CLV1 Onuf CLV1	AACAACATCGTGGACT AACAACATCGTGGACT	GGGTGCGGCGG			
OsJ CLV1	AACAACATCGTGGACT				
Clustal Consensus	* * * * * * * * * * * * * * * *	* * * * * * * * * * *	* * * * * * * * * * *	* * * * * * * * * * *	* * * * * * * * * * * *
	2830	2840	2850	2860	2870 2880
01011					
OsI_CLV1 Oruf_CLV1	GACGCTGCGGCGTGGC				
OsJ CLV1	GACGCTGCGGCGTGGC				
Clustal Consensus	* * * * * * * * * * * * * * * *	* * * * * * *	*********	*********	* * * * * * * * * * * *
	2890	2900	2910	2920	2930 2940
575107356875532					
OsI_CLV1 Oruf CLV1	CTTAGGGTGGCGCTGC CTTAGGGTGGCGCTGC				TCGATGAGGGAG
OsJ CLV1	CTTAGGGTGGCGCTGC				
Clustal Consensus	* * * * * * * * * * * * * * * *	* * * * * * * * * * *	********	* * * * * * * * * * *	* * * * * * * * * * * *
	2950	2960	2970	2980	2990 3000
OsI_CLV1	GTGCTGTCCATGCTGC	AGGAGGCCAGG	CCGAAACGGAA		
Oruf_CLV1 OsJ_CLV1	GTGCTGTCCATGCTGC				AAGAAGCAGGTT AAGAAGCAGGTT
Clustal Consensus	* * * * * * * * * * * * * * * * *	* * * * * * * * * * *	* * * * * * * * * *	* * * * * * * * * *	* * * * * * * * * * * * *
	3010	3020	3030	3040	3050 3060
OsI_CLV1	AAGTAAAATGGGTGAT				TATATGGTGTGC
Oruf_CLV1 OsJ_CLV1	AAGTAAAATGGGTGAT AAGTAA	GGTTAAGTCTT	TAGCAGAAAGA	AGAACTAATA	TATATGGTGTGC
Clustal Consensus	*****				
	and the second second				
OsI_CLV1	TCTTCGTG				
Oruf_CLV1 OsJ_CLV1	TCTTCGTG				
Clustal Consensus	122.04.0000 EX TO TO T XE				

Figure 4.11: ClustalW gene sequence alignment of *Oruf_CLV1*, *OsI_CLV1* and *OsJ_CLV1* (gi:1256021832; 2,640-3,006). Boxes show the SNPs.

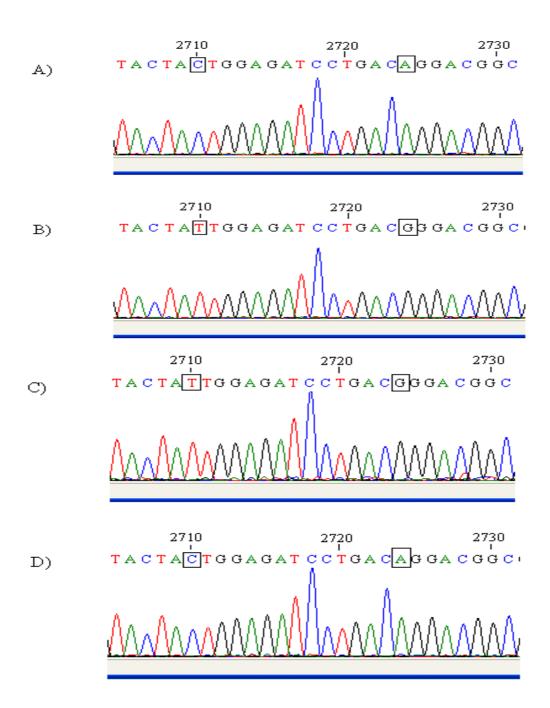
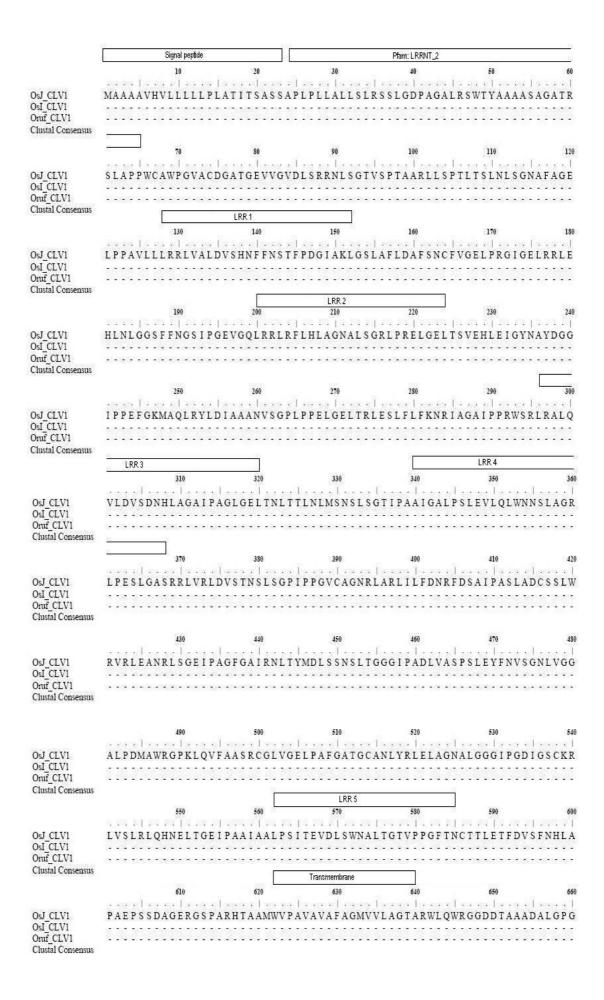


Figure 4.12: Base substitution of partial putative *CLV1* sequences revealed by sequencing from (A) *Oryza rufipogon*, (B) *Oryza sativa* ssp. *indica* cv. MR219, (C) BC_2F_7 line 7 and (D) BC_2F_7 line 23. Base substitutions are boxed. Sequencing chromatogram of *Oryza rufipogon* was identical to BC_2F_7 line 23; whereas, sequencing chromatogram of *Oryza sativa* ssp. *indica* cv. MR219 was identical to BC_2F_7 line 7.

Gene-specific outer and inner primers for 5' and 3' RLM-RACE PCR were designed and based on the 5' UTR and 3' UTR of patented sequence (WO/2000/004761) and *OsJ_CLV1* (gi:125602183). However, PCR amplifications of 5' and 3' RLM-RACE PCR of *Oruf_RPK1* and *OsI_RPK1* were unsuccessful.

4.2.2 Structure Analysis of Putative CLV1

Although the amplification of Oruf CLV1 and OsI_CLV1 full length cDNA were unsuccessful, the partial predicted 112 amino acids for Oruf CLV1 and OsI CLV1 respectively were compared with the predicted 1,001 amino acids of OsJ CLV1 (gi:125602183). Partial amino acid sequences of OsI CLV1 showed 100 % identity with OsJ CLV1 (gi:125602183), whilst partial amino acid sequences of Oruf CLV1 showed 99 % identity with OsJ CLV1 (gi:125602183) as shown in Figure 4.13. SMART analysis showed that the OsJ CLV1 (gi:125602183) contained extracellular receptor domains and cytoplasmic protein kinase domains (Figure 4.13). Eight conserved domains were identified, namely signal peptide, LRRNT 2, LRR motif, and serine/threonine protein kinase. The LRRNT 2 and LRR motif belong to extracellular receptor domains; whereas, serine/threonine protein kinase belongs to cytoplasmic protein kinase domains (Figure 4.13). As mentioned in section 4.1.2 previously, the SMART able to cross-references to other domain databases. The signal peptide was predicted using SignalP 3.0 Server (http://www.cbs.dtu.dk/services /SignalP; Nielsen et al., 1997; Bendtsen et al., 2004; Emanuelsson et al., 2007. The LRRNT 2 was predicted using protein domain families (PFAM) database (Bateman



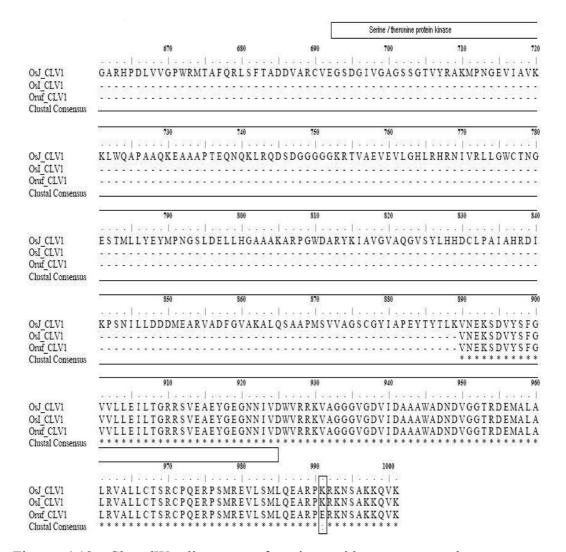


Figure 4.13: ClustalW alignment of amino acid sequence and gene structure prediction of Oruf_CLV1, OsI_CLV1 and OsJ_CLV1 (gi:125602183). Functional domains identified by SMART analysis are indicated above the sequence. The amino acid substitutions are boxed.

et al., 2004). The transmembrane helix was predicted using Transmembrane Prediction Server (http://www.ch.embnet.org/software/TMPRED_form.html; Hofmann and Stoffel, 1993). The remaining conserved domains were identified using Basic Local Alignment Search Tool (BLAST) database (Karlin and Altschul, 1993). There was only one different amino acid identified between Oruf_CLV1 and OsI_CLV1, and this also occurs between Oruf_CLV1 and OsJ_CLV1 (gi:125602183). This substitution of lysine to glutamine acid converses a polar side chain.

4.3 Characterization of Kinase of Putative RPK1 and Putative CLV1

A number of known RLK kinases were aligned with the predicted amino acid sequences of the kinase domains of Oruf_RPK1 (382-643), OsI_RPK1 (382-643), OsJ_RPK1 (gi:18677097; 382-643) and OsJ_CLV1 (gi:125602183; 699-952). *Arabidopsis thaliana* CLAVATA1 At_CLV1 (sp:Q9SYQ8.3; 692-968), rice FLORAL NUMBER1 OsJ_FON1 (gi:56790017; 708-915) and *Arabidopsis thaliana* ERECTA At_ERECTA (gi:1345132: 653-910) are known to belong to the RD kinase class (Clark *et al.*, 1997; Dardick and Ronald, 2006; Afzal *et al.*, 2008). Meanwhile, rice bacterial blight resistance protein OsI_Xa21 (gi:1122443; 708-922) and *Arabidopsis thaliana* flagellin perception At_FLS2 (sp:Q9FL28.1; 882-1152) are known to belong to the non-RD kinase class (Song *et al.*, 1995, Gomez-Gomez and Boller, 2000; Dardick and Ronald, 2006; Afzal *et al.*, 2008). The *Solanum lycopersicum* receptor-like protein kinase SI_PEPRK1 (gi:3015488; 372-638) and *Glycine max* receptor-like kinase protein Gm_RHG1 (gi:206584433; 578-834) were

identified to belong to the kinase minus class (Muschietti *et al.*, 1998; Dardick and Ronald, 2006; Afzal and Lightfoot, 2007; Afzal *et al.*, 2008). Based on the absence of a conserved arginine (R) and aspartic acid (D) in the kinase subdomain VI, Oruf_RPK1 (382-643), OsI_RPK1 and OsJ_RPK1 (gi:18677097; 382-643) most probably belong to the RD minus kinase class (shown in box B; Figure 4.14); whereas, OsJ_CLV1 (gi:125602183; 699-952) most probably belong to the RD kinase class based on the presence of conserved arginine (R) and aspartic acid (D) residues in the kinase subdomain VIb (shown in box B; Figure 4.14).

	10 20 _30 40 50 60	
At CVL1 OsJ_FON1 OsJ_CLV1 At ERECTA OsI_Xa21 At FLS2	LKEENIIGKGGAGIVYRGSMPNNVDVAIKRLVGR NIIGKGGAGIVYHGVTRG-AELAIKRLVGR GGGEH GAGSSGTVYRAKMPNGEVIAVKKLWQAPAAQKEAAAPTEQNQKLRQDSDGGGG IGHGASSTVYKCVLKNCKPVAIKRLYSHN FAPTNLLGSGSFGSVYKGKLNIQDHVAVKVLKLEN STVYKGQLEDGTVIAVKVLNLKE STVYKGQLEDGTVIAVKVLNLKE STVYKGQLEDGTVIAVKVLNLKE STVYKGQLEDGTVIAVKVLNLKE	
OsJ_RPK1 OsI_RPK1 Oruf_RPK1 Gm_RHG1 S1_PEPRK1 Clustal Consensus	EVLGKGSVGTSYKAVLEEGTTVVVKRLKD EVLGKGSVGTSYKAVLEEGTTVVVKRLKD EVLGKGSVGTSYKAVLEEGTTVVVKRLKD 	
	70 80 90 100 110 120	
At CVL1 OsJ FON1 OsJ CLV1 At ERECTA OsJ Xa21 At FLS2 OsJ RPK1 OsJ RPK1 OsJ RPK1 OmJ RPK1 Gm RHG1 S1 PEPRK1 GT THCC	DHGFTAEIQTLGRIRHRHIVRLLGVVANK DTNLLLYEYMPNGSLGELLHG S DRGFSAEVTTLGRIRHRNIVRLLGFVSNR ETNLLLYEYMPNGSLGEMLHG G GKRTVAEVEVLGHLRHRNIVRLLGWCTNG ESTMLLYEYMPNGSLDELLHG A MKQFETELEMLSSIKHRNLVSLQAYSLSH LGSLLFYDYLENGSLWDLHHG P S FTAECEALRNMRHRNLVKIVTICSSIDNRGNDFKAIVYDFMPNGSLEDWIHPETNDQ DKWFYTEAKTLSQLKHRNLVKIUTICSSIDNRGNDFKAIVYDFMPNGSLEDWIHPETNDQ DKWFYTEAKTLSQLKHRNVLPVRAYYFS KDEKLLVFDYLPNGSLSAMLHGS - RG RREFDAHMDALGKVEHRNVLPVRAYYFS KDEKLLVFDYLPNGSLSAMLHGS - RG RREFDAHMDALGKVEHRNVLPVRAYYFS KDEKLLVFDYLPNGSLSAMLHGS - RG RREFDAHMDALGKVEHRNVLPVRAYYFS KDEKLLVFDYLPNGSLSAMLHGS - RG RKEFDAHMDALGKVEHRNVLPVRAYYFS KDEKLLVFDYLPNGSLSAMLHGS - RG RKEFDAHMDALGKVEHRNVLPVRAYYFS KGEKLLVFDYLPNGSLSAMLHGS - RG KEDFHEHMRRIGRLSHKNLLPVVAYYR KEEKLLVSEYVNNVSLAVHLHGN - KS	
Clustal Consensus	130 140 150 B 160 170 C 180	í
At_CVL1 OsJ_FON1 OsJ_CLV1 At_ERECTA OsJ_Xa21 At_FLS2 OsJ_RPK1 OsJ_RPK1 Ond_RPK1 Gm_RHG1 S1_PEPRK1	KG - GHLQWE TRHRVAVEAAKGLCYLHHDCSP - LILHRDVKSNNILLD - SDFEAHVADFGL KG - GHLGWEARARVAAEAACGLCYLHHDCAP - RIIHRDVKSNNILLD - SAFEAHVADFGL AAKARPGWDARYKIAVGVAQGVSYLHHDCLP - AIAHRDIKPSNILLD - DDMEARVADFGV TKKKTLDWDTRLKIAYGAAQGLAYLHHDCSP - RIIHRDVKSSNILLD - KDLEARLTDFGI ADQRHLNLHRRVTILLDVACALDYLHRHGPE - PVVHCDIKSSNVLLD - SDMVAHVGDFGL G SLLEKIDLCVHIASGIDYLHSGYGF - PIVHCDLKPANILLD - SDRVAHVSDFGT SGKTPLDWDARMRSALSAARGLAHLHTVH SLVHGNVKSSNVLLRPDADAAALSDFCL SGKTPLDWDAQMRSALSAARGLAHLHTVH SLVHGNVKSSNVLLRPDADAAALSDFCL SGKTPLDWDARMRSALSAARGLAHLHTVH SLVHGNVKSSNVLLRPDADAAALSDFCL GTETFIDWPTRMKIAQDLARGLFCLHSQE - NIIHGNLTSSNVLLRPDADAAALSDFCL	
Clustal Consensus	· * · · * · · * · · · · · · · · · · · ·	
At CVL1 OsJ_FON1 OsJ_CLV1 At ERECTA OsJ_RPK1 OsJ_RPK1 OsJ_RPK1 Omf_RPK1 Gm_RHG1 S1_PEPRK1 Clustal Consensus	AKFLVDGAASECMSSIAGSYGYIAPEYAYTLKVDEKSDVYSFGVVLLELIAGKK AKFLGG-ATSECMSAIAGSYGYIAPEYAYTLKVDEKSDVYSFGVVLLELITGRR AKALQSAAPMSVVAGSCGYIAPEYAYTLKVDEKSDVYSFGVVLLELITGRRS- AKSLCVS-KSHTSTYVMGTIGYIDPEYARTSRLTEKSDVYSYGIVLLELLTGKRP- ARILVDGTSLIQQSTSSMGFIGTIGYAAPEYGVGLIASTHGDIYSYGILVLELLTRKA- ARILVDGTSLIQQSTSSMGFIGTIGYLAPEFAYMRKVTTKADVFSFGIIMMELMTKQRPT HPIFAPSSARPGAGGYRAPEVVDTRRPTYKADVYSLGVLLLELLTGKSPT HPIFAPSSARPGAGGYRAPEVVDTRRPTYKADVYSLGVLLLELLTGKSPT HPIFAPSSARPGAGGYRAPEVVDTRRPTYKADVYSLGVLLLELLTGKSPT SRLMSTAANSNVIATAGALGYRAPELSKLKKANTKTDIYSLGVILLELLTRKSPG LPVVNLEHAQEHMIAYKSPEFKHNGRITRKNDVWTLGILILEMLTGKFPS	
At CVL1 OsJ FON1 OsJ CLV1 At ERECTA OsJ Xa21 At FLS2 OsJ RPK1 OsJ RPK1 OsJ RPK1 Om RHG1 S1_PEPRK1	- PVGEFGEGVDIVRWVRNTEEEITQPSDAAIVVAIVDPRLTGYPLTSVIHVFKIAMMCV - PVGGFGDGVDIVH - VEAEYGEGNNIVDWVR - RKVAGGGVGDVIDAAAWADNDVGGTRDEMALALRVALLCT - VDESNLHHLIMSKTG - NNEVMEMADPDITSTCKDLGVVKKVFQLALLCT - SLNDEDSQDMTLRQLVEKSIGNGRKGMVRVLDMELGDSIVSLKQEEAIEDFLKLCLFCT - HASLEGDGTLDLPRWVQSVVREEWTAEVFDVELVRLGASA - EEEMVALLQVAMACV - HASLEGDGTLDLPRWVQSVVREEWTAEVFDVELVRLGASA - EEEMVALLQVAMACV - HASLEGDGTLDLPRWVQSVVREEWTAEVFDVELVRLGASA - EEEMVALLQVAMACV - MASLEGDGTLDLPRWVQSVVREEWTAEVFDVELVRLGASA - EEEMVALLQVAMACV - MASLEGDGTLDLPRWVQSVVREEWTAEVFDVELVRLGASA - EEEMVALLQVAMACV - VSMNG LDLPQWVASVVREEWTAEVFDADLMRDASTV - GDELLNTLKLALHCV NFLQQGKGSDTDLATWVRSVVNEDMTEVDVFEKEMRGTTNS - EGEMMKLLKIALGCC	
Clustal Consensus		
Clustal Consensus	310	
Clustal Consensus At_CVL1 OsJ_FON1 OsJ_CLV1 At_ERECTA OsJ_R21 At_FLS2 OsJ_RPK1 OsI_RPK1 Ond_RPK1 Omf_RPK1 Gm_RHG1	310 EEEAAARPTMREVVHMLT SRCPQERPSMREVLSML- KRQPNDRPTMHQV SSRPEDRPDMNEILTH ATVPDARPDAPDVV ATVPDARPDAPDVV DPSPSARPEVHQVLQ	

Figure 4.14: ClustalW amino acid alignment of kinase domains with known RLKs. Characterization of kinase was based on presence or absence conserved lysine (K) kinase subdomain II (shown in box A), conserved arginine (R) and aspartic acid (D) in kinase subdomain VIb (shown in box B), and conserved aspartic acid (D) in kinase subdomain VII (shown in box C; Dardick and Ronald, 2006; Afzal *et al.*, 2008). At_CLV1 (sp:Q9SYQ8.3; 692-968), At_ERECTA (gi:1345132: 653-910), At_FLS2 (sp:Q9FL28.1; 882-1152) from *Arabidopsis thaliana*; Gm_RHG1 (gi:206584433; 578-834) from *Glycine max*; OsI_RPK1 (382-643), OsI_Xa21 (gi:1122443; 708-922), OsJ_CLV1 (gi:125602183; 699-952), OsJ_FON1 (gi:56790017; 708-915), OsJ_RPK1 (gi:18677097; 382-643) from *Oryza sativa*; Oruf_RPK1 (382-643) from *Oryza rufipogon*; SI_PEPRK1 (gi:3015488; 372-638) from *Solanum lycopersicum*.

4.4 Phylogenetic Analysis

A total of 14 orthologous RPK1 amino acid sequences and 16 orthologous CLV1 amino acid sequences were selected from the GenBank protein database and the OrthoMCL database (Chen *et al.*, 2006). The relatively conserved regions of orthologous RPK1 amino acid sequences and orthologous CLV1 amino acid sequences were screened by Gblocks 0.91 (Appendix C; Appendix D; Castresana, 2000; Dereeper *et al.*, 2008). The Gblocks 0.91 screened sequences were used to construct phylogenetic trees (Appendix E; Appendix F). Figure 4.15 shows phylogenetic tree putative RPK1 amino acid sequences. The Oruf_RPK1, OsI_RPK1 and OsJ_RPK1 (gi:18677097) are grouped together with At_TMK (At5g58300), At_RLK (At4g23740), OsJ_RLL1 (gi:7573610), OsJ_RLL2 (gi:15128407) and OsJ_

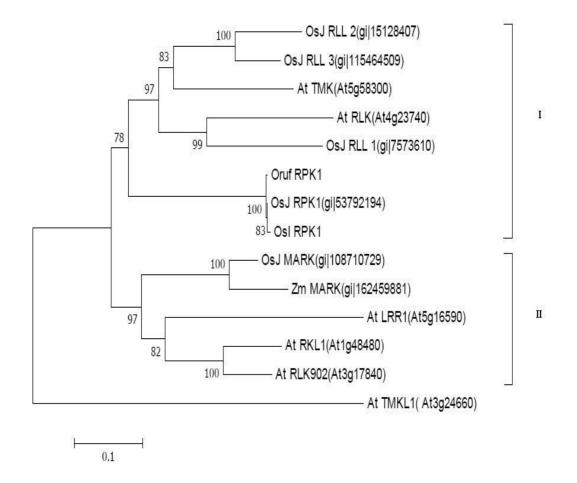
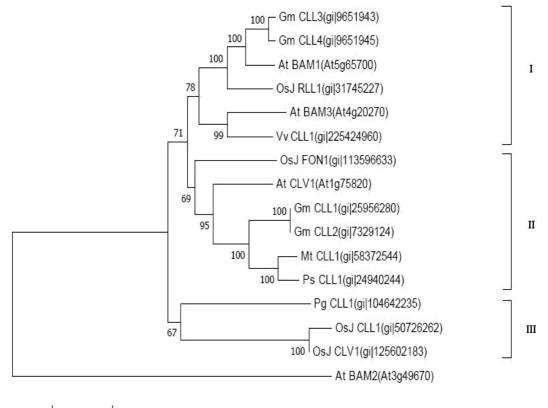


Figure 4.15: Phylogenetic analysis of orthologous RPK1 amino acid sequences from *Arabidopsis thaliana*, *Oryza sativa*, *Oryza rufipogon* and *Zea mays*. The tree was formatted with MEGA 4.0.2 program. The bootstrap value greater than 50 % was selected from the neighbor-joining method with Poisson correction. At_LRR1 (At5g16590), At_RKL1 (At1g48480), At_RLK (At4g23740), At_RLK902 (At3g17840), At_TMK (At5g58300) and At_TMKL1 (At3g24660) from *Arabidopsis thaliana*; OsJ_MARK (gi:108710729), OsI_RPK1, OsJ_RPK1 (gi:18677097) OsJ_RLL1 (gi:7573610), OsJ_RLL2 (gi:15128407) and OsJ_RLL3 (gi:115464509) from *Oryza sativa*; Oruf_RPK1 from *Oryza rufipogon*; Zm_MARK (gi:226498594) from *Zea mays*.

RLL3 (gi:115464509) in group I. The phylogenetic tree of orthologous CLV1 amino acid sequences as shown in Figure 4.16. Three well-supported groups (designated as group I, group II and group III) were formed from 16 orthologous putative CLV1 amino acid sequences. The OsJ_CLV1 (gi:125602183) are grouped together with Pg_CLL1 (gi:104642235) and OsJ_CLL1 (gi:50726262) in group III.

Most of the orthologous RPK1 and CLV1 amino acid sequences are known to belong to LRR-III and LRR-XI subfamilies in the plant RLK/*Pelle* family respectively (Shiu *et al.*, 2002; Shiu *et al.*, 2004). The phylogenetic analysis of these sequences (Figure 4.15 and 4.16) is in agreement with the suggestion that Oruf_RPK1, OsI_RPK1 and OsJ_RPK1 (gi:18677097) could belong to the LRR-III subfamily in the plant RLK/*Pelle* family; whereas, OsJ_CLV1 (gi:125602183) may be grouped in the LRR-XI subfamily in the plant RLK/*Pelle* family. The partial 402 bp long gene sequences of *Oruf_CLV1* and *OsI_CLV1* were not selected for the construction of phylogenetic tree because only full length amino acid sequence was selected in this study. It is suggested that Oruf_CLV1 and OsI_CLV1 belong to the LRR XI subfamily in the plant RLK/*Pelle* family too.



0.2

Figure 4.16: Phylogenetic analysis of orthologous CLV1 amino acid sequences from *Arabidopsis thaliana*, *Oryza sativa*, *Glycine max*, *Medicago truncatula*, *Picea glauca*, *Pisum sativum* and *Vitis vinifera*. The tree was formatted with MEGA 4.0.2 program. The bootstrap value greater than 50 % was selected from the neighbor-joining method with Poisson correction. At_BAM1 (At5g65700), At_BAM2 (At3g49670), At_BAM3 (At4g20270) and At_CLV1 (At1g75820) from *Arabidopsis thaliana*; Gm_CLL1 (gi:25956280), Gm_CLL2 (gi:7329124), Gm_CLL3 (gi:9651943) and Gm_CLL4 (gi:9651945) from *Glycine max*; Mt_CLV1 (gi:58372544) from *Medicago truncatula*; OsJ_CLL1 (gi:50726262), OsJ_CLV1 (gi:125602183), OsJ_FON1 (gi:113596633), OsJ_LRK1 (gi:255677496) and OsJ_RLL1 (gi:31745227) from *Oryza sativa*; Pg_CLL1 (gi:104642235) from *Picea glauca*; Ps_CLL1 (gi:24940244) from *Pisum sativum*; Vv_CLL1 (gi:225424960) from *Vitis vinifera*.

4.5 Gene Expression Study of Putative RPK1 and Putative CLV1

4.5.1 Validation of Comparative C_T Method Real Time qRT-PCR

The total RNA of *Oryza rufipogon* (accession number IRGC105491), *Oryza sativa* ssp. *indica* cv. MR219, BC₂F₇ line 7 and BC₂F₇ line 23 was successfully extracted from the leaves and the whole plant at the seedling stage, the panicles at the booting, heading and flowering stages, and grains at the milk grain stage (Figure 4.17A). After that, the total RNA from different tissues at different developmental stages was treated with DNase I to remove DNA (Figure 4.17B). Next, all treated total RNAs were reverse transcribed to synthesize cDNA. The cDNA of different tissues at the synthesize cDNA. The cDNA of different tissues at different tissues at different tissues at tissues at the total stages in rice were used during the real time qRT-PCR experiments.

Housekeeping genes are used to normalize target genes in a real time qRT-PCR experiment. The normalization of real time qRT-PCR data against housekeeping genes is to obtain an accurate and reliable gene expression profile. *Actin, eEF-1a* and *UBQ5* were selected to examine gene expression stability at different developmental stages in rice using geNORM v3.4 software (Primer-Design, UK) in this study. Measure M value of *Actin* was 0.360, whereas measure M value of *eEF-1a* and *UBQ5* were 0.18 (Figure 4.18). Thus, *eEF-1a* and *UBQ5* were selected as pair-wise housekeeping genes to normalize the gene expression of putative *RPK1* and putative *CLV1*. The *eEF-1a* and *UBQ5* were found to be the most stable housekeeping genes

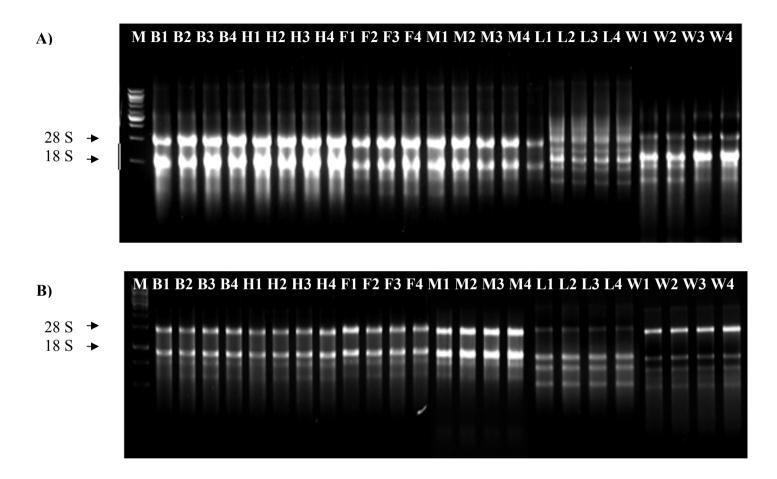


Figure 4.17: Agarose gel electrophoresis of (A) total RNA and (B) DNaseI treated total RNA of *Oryza rufipogon* (1), *Oryza sativa* ssp. *indica* cv. MR219 (2), BC₂F₇ line 7 (3) and BC₂F₇ line 23 (4) at different stages. M: 1 kb DNA marker; B: Booting stage; H: Heading stage; F: Flowering stage; M: Milk grain stage; L: Leaf of 8th seedling day; W: Whole plant of 8th seedling day.

(A)	Change Data	Actin	eEF-1α	UBQ5	Normalisation Factor
	Booting	1.87E+01	1.80E+01	1.82E+01	0.9364
	Heading	1.86E+01	1.82E+01	1.82E+01	0.9385
	Flowering	2.07E+01	1.74E+01	1.75E+01	0.9451
	Milk grain	3.33E+01	1.67E+01	1.70E+01	1.0811
	Leaf of Seedling	2.32E+01	1.97E+01	2.01E+01	1.0726
	Whole plant	2.13E+01	1.94E+01	2.02E+01	1.0384
	-				
	M < 1.5	0.360	0.189	0.189	

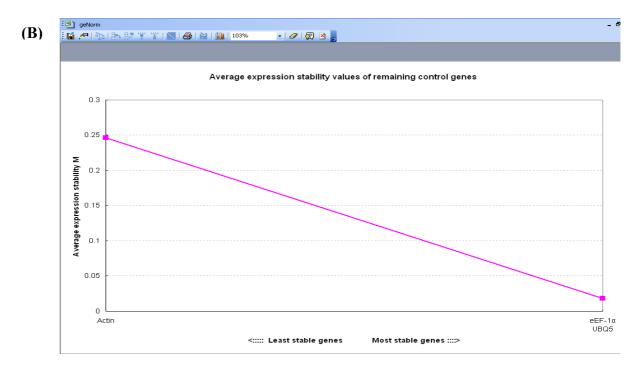
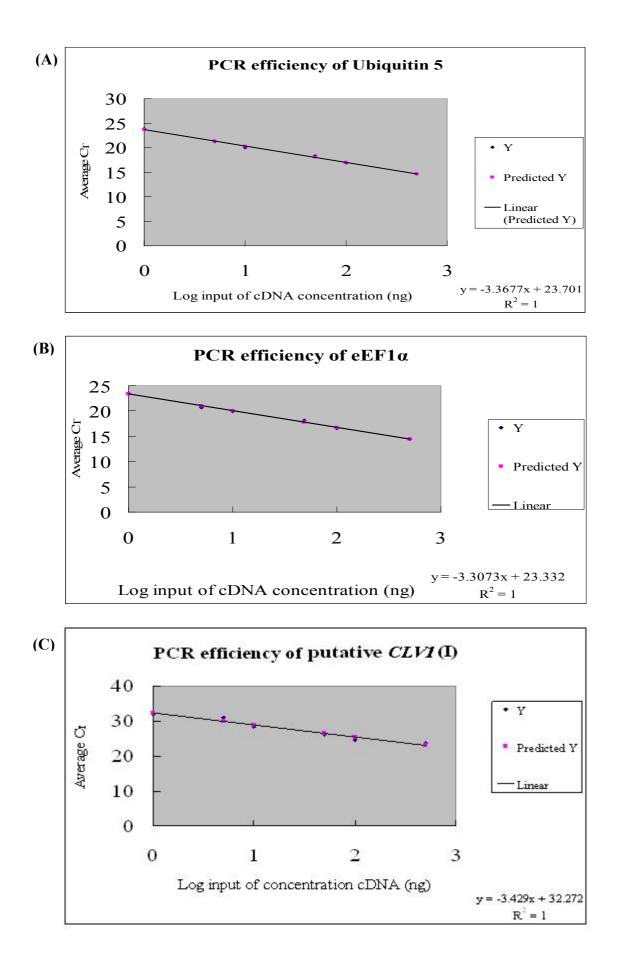


Figure 4.18: Selection of housekeeping genes by geNORM v3.4. (A) Calculation of gene stability measure (M) of *Actin*, *eEF-1* α and *UBQ5* at different developmental stages in rice. (B) Average expression stability values of *Actin*, *eEF-1* α and *UBQ5*.

across developmental stages, showing the lowest gene stability measure M values (Figure 4.18).

While, PCR amplification efficiency was used to measure the overall performance of real time qRT-PCR experiment. The PCR amplification efficiency of the housekeeping gene (*eEF-1* α and *UBQ5*) and gene specific primer target set (putative *RPK1* and putative *CLV1*) were within 90-110 % with the standard curve slope between -3.587 to -3.103 (Figure 4.19). The PCR amplification efficiency results suggest that they are suitable for investigation of gene expression of putative *RPK1* and putative *CLV1*. Two different sets of forward primers were designed for putative *CLV1* (see section 3.4). The lengths of the two different forward primers of putative *CLV1* (I) and putative *CLV1* (II) were 24 bases. The 16th base of putative *CLV1* (I) forward primer was substituted from thymine (T) to cytosine (C) for the putative *CLV1* (II) primer. Putative *CLV1* (I) forward primer was used for *Oryza rufipogon* and BC₂F₇ line 23 samples, whereas, putative *CLV1* (II) forward primer was used for *Oryza sativa* ssp. *indica* cv. MR219 and BC₂F₇ line 7 samples.



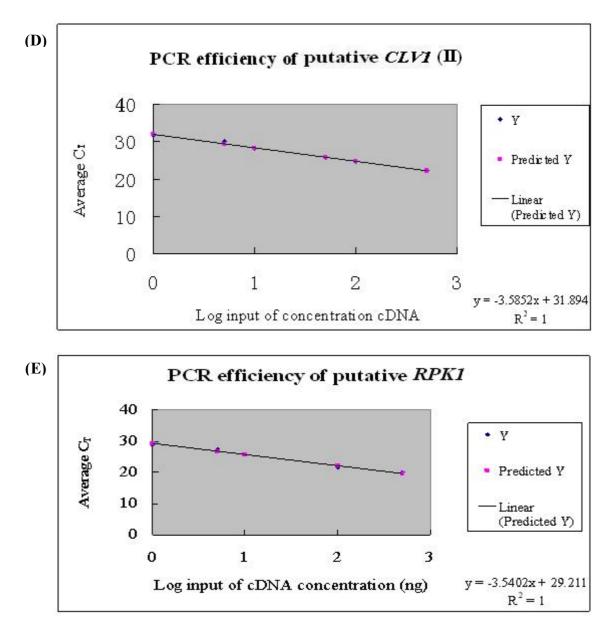


Figure 4.19: PCR efficiency of targets and housekeeping genes. The PCR efficiencies of (A) UBQ5, (B) $eEF-1\alpha$, (C) putative CLV1 (I), (D) putative CLV1 (II) and (E) putative RPK1 were plotted.

4.5.2 Expression Profiles of Putative RPK1

Figure 4.20 shows real time qRT-PCR gene expression profiling of putative RPK1 of Oryza rufipogon, Oryza sativa ssp. indica cv. MR219, BC₂F₇ line 7 and BC₂F₇ line 23 at different developmental stages. Leaf of 8th day seedling stage was selected as calibrator for expression profiling of putative RPK1. This is because putative RPK1 transcript leaf of 8th day seedling stage had the lowest expression compared with other stages. Gene expression of putative RPK1 was the highest in the panicle of booting stage, except in BC₂F₇ line 23. After the booting stage, gene expression decreased gradually until the flowering stage. According to sequencing data as shown in Figure 4.4, BC₂F₇ line 23 is homozygous for the putative RPK1 allele from *Oryza rufipogon*; whereas, BC_2F_7 line 7 is homozygous for the putative RPK1 allele from Oryza sativa ssp. indica cv. MR219. However, the putative RPK1 expression profile did not observe correlation between Oryza rufipogon and BC₂F₇ line 23, and between Oryza sativa ssp. indica cv. MR219 and BC₂F₇ line 7 as shown in Figure 4.20. Statistical analysis using two-way ANOVA with rescaled normalized expression level from the putative RPK1 of Oryza rufipogon, Oryza sativa ssp. indica cv. MR219, BC₂F₇ line 7 and BC₂F₇ line 23 collected at different developmental stages (Table 4.2) revealed that the lines, stages, and correlation between lines and stages were significant (P<0.01; Table 4.3).

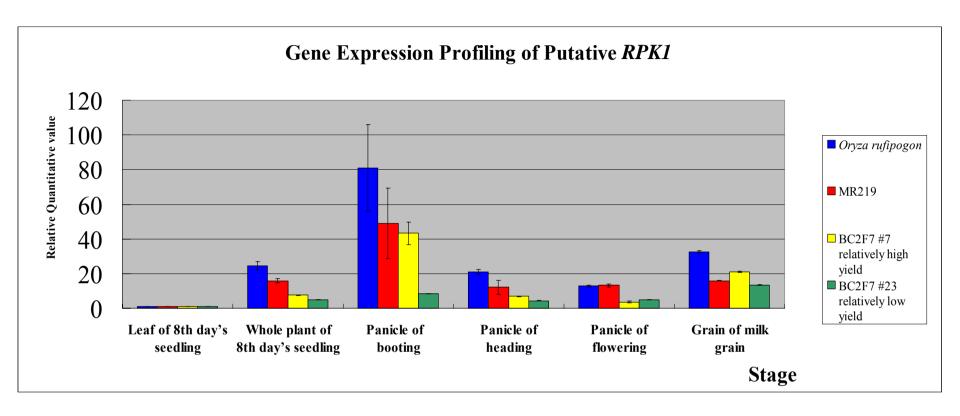


Figure 4.20: Comparative gene expression of putative *RPK1* between vegetative, reproductive and ripening phase. *Oryza rufipogon*, *Oryza sativa* ssp. *indica* cv. MR219, BC₂F₇ line 7 and BC₂F₇ line 23 were used on leaves and the whole plant at the seedling stage, the panicles at the booting, heading and flowering stages, and grains at the milk grain stage. Error bars indicate standard error of the means.

Sample and stage	Normalized Expression level of target (putative <i>RPK1</i>)	Standard error of normalized Expression level of target (putative <i>RPK1</i>)	Rescaled normalized expression level of target (putative <i>RPK1</i>)	Standard error of rescaled normalized expression level of target (putative <i>RPK1</i>)
OR-L	0.0300	0.0010	1	0.0010
OR-W	0.7351	0.1040	24.5033	2.5483
OR-B	2.4295	0.3088	80.9833	25.0076
OR-H	0.6303	0.0701	21.0100	1.4728
OR-F	0.3870	0.0392	12.9000	0.5057
OR-M	0.9767	0.0173	32.5567	0.5632
MR-L	0.0519	0.0022	1	0.0022
MR-W	0.8218	0.1152	15.8343	1.8241
MR-B	2.5419	0.4116	48.9769	20.1589
MR-H	0.6270	0.0578	12.0809	0.6983
MR-F	0.5345	0.0493	10.2987	0.5077
MR-M	0.8180	0.0147	15.7611	0.2317
BC ₂ F ₇ #7 -L	0.0331	0.0018	1	0.0018
BC ₂ F ₇ #7-W	0.2521	0.0324	7.6163	0.2468
BC ₂ F ₇ #7-B	1.4306	0.1469	43.2205	6.3491
BC ₂ F ₇ #7-H	0.2292	0.0305	6.9245	0.2112
BC ₂ F ₇ #7-F	0.1187	0.0184	3.6012	0.4275
BC ₂ F ₇ #7-M	0.6979	0.0139	21.0846	0.2931
BC ₂ F ₇ #23 -L	0.0749	0.0051	1	0.0051
BC ₂ F ₇ #23-W	0.3712	0.0537	4.9560	0.0598
BC ₂ F ₇ #23-B	0.6292	0.0392	8.4000	0.0697
BC ₂ F ₇ #23-H	0.3275	0.0116	4.3725	0.0107
BC ₂ F ₇ #23-F	0.3628	0.0544	4.8438	0.0501
BC ₂ F ₇ #23-M	1	0.0640	13.3500	0.1808

Table 4.2: Data of rescaled normalized expression level of putative *RPK1* and standard error of rescaled normalized expression level of putative *RPK1* of *Oryza rufipogon* (OR), *Oryza sativa* ssp. *indica* cv. MR219 (MR), BC₂F₇ line 7 and BC₂F₇ line 23 at different developmental stages. L: Leaf of 8th seedling day; W: Whole plant of 8th seedling day; B: Booting stage; H: Heading stage; F: Flowering stage; M: Milk grain stage.

Item	Sum of Squares	Degrees of Freedom	Mean Squares	Variances	P value summary	Probability
Between lines	6392	3	2131	14.01	**	< 1 %
Between stages	19302	5	3860	42.29	***	< 1 %
Interaction	6912	15	460.8	15.15	***	< 1 %
Residual (error)	13030	72.0	181.0			
Total	45637	95.0				

Table 4.3: Statistical analysis using two-way ANOVA with putative RPK1 profile of Oryza rufipogon, Oryza sativa ssp. indica cv. MR219,

BC₂F₇ line 7 and BC₂F₇ line 23 collected at different developmental stages. * indicates that the degrees of statistical significance.

4.5.3 Expression Profiles of Putative CLV1

Gene expression profiling of putative CLV1 of Oryza rufipogon, Oryza sativa ssp. indica cv. MR219, BC₂F₇ line 7 and BC₂F₇ line 23 at different developmental stages shown in Figure 4.21. Putative CLV1 transcript leaf of 8th day seedling stage had lowest expression compared with other stages, and was selected as calibrator for expression profiles. Gene expression of putative CLV1 of BC₂F₇ line 7 was highly up-regulated at booting, heading and flowering stages compared to Oryza rufipogon, Oryza sativa ssp. indica cv. MR219 and BC₂F₇ line 23. According to sequencing data as shown in Figure 4.12, putative CLV1 of Oryza rufipogon and BC₂F₇ line 23 were identical to each other; whereas, putative CLV1 of Oryza sativa ssp. indica cv. MR219 and BC₂F₇ line 7 were identical to each other. However, the correlation of putative CLV1 expression between Oryza rufipogon and BC₂F₇ line 23, and between Oryza sativa ssp. indica cv. MR219 and BC₂F₇ line 7 did not observe (Figure 4.21). Statistical analysis using two-way ANOVA with rescaled normalized expression level from the putative CLV1 of Orvza rufipogon, Orvza sativa ssp. indica cv. MR219, BC_2F_7 line 7 and BC_2F_7 line 23 collected at different developmental stages (Table 4.4) revealed that the lines, stages, and correlation between lines and stages was extremely significant (P<0.0001; Table 4.5).

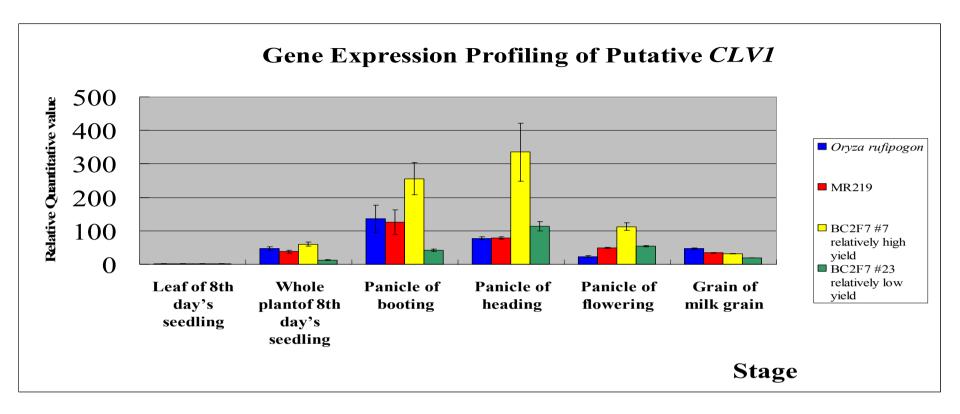


Figure 4.21: Comparative gene expression of putative *CLV1* between vegetative, reproductive and ripening phase. *Oryza rufipogon, Oryza sativa* ssp. *indica* cv. MR219, BC₂F₇ line 7 and BC₂F₇ line 23 were used on leaves and the whole plant at the seedling stage, the panicles at the booting, heading and flowering stages, and grains at the milk grain stage. Error bars indicate standard error of the means.

Sample and stage	Normalized expression level of target	Standard error of normalized expression	Rescaled normalized expression	Standard error of rescaled normalized	
	(putative	level of target	level of target	expression level	
	CLV1)	(putative	(putative	of target	
		CLV1)	CLV1)	(putative CLV1)	
OR-L	0.0083	0.0008	1	0.0008	
OR-W	0.3744	0.0647	45.1084	2.9185	
OR-B	2.4295	0.2682	292.7161	78.5065	
OR-H	1.0828	0.0561	130.4578	7.3187	
OR-F	0.8993	0.0132	108.3493	1.4302	
OR-M	0.2172	0.0057	26.1686	0.1492	
MR-L	0.0209	0.0033	1	0.0033	
MR-W	0.7864	0.1020	37.6268	3.8379	
MR-B	2.6300	0.2869	125.8373	36.1027	
MR-H	1.6578	0.0438	79.3206	3.4742	
MR-F	1.0281	0.0411	49.1914	2.0218	
MR-M	0.7176	0.0231	34.3349	0.7931	
BC ₂ F ₇ #7 -L	0.0025	0.00004	1	0.00004	
BC ₂ F ₇ #7-W	0.1269	0.0393	50.7600	1.9949	
BC ₂ F ₇ #7-B	1.6802	0.3626	672.0800	134.9162	
BC ₂ F ₇ #7-H	1.8671	0.0946	746.8400	70.6511	
BC ₂ F ₇ #7-F	0.8405	0.0563	336.2000	18.9281	
BC ₂ F ₇ #7-M	0.4762	0.0178	190.4800	3.3905	
BC ₂ F ₇ #23 -L	0.0256	0.0070	1	0.0070	
BC ₂ F ₇ #23-W	0.3192	0.0865	12.4688	1.0786	
BC ₂ F ₇ #23-B	1.0490	0.1491	40.9766	6.1096	
BC ₂ F ₇ #23-H	1.5913	0.0763	62.1602	4.5904	
BC ₂ F ₇ #23-F	0.8509	0.0574	33.2383	1.9079	
BC ₂ F ₇ #23-M	0.6455	0.0195	25.2148	0.4917	

Table 4.4: Data of rescaled normalized expression level of putative *CLV1* and standard error of rescaled normalized expression level of putative *CLV1* of *Oryza rufipogon* (OR), *Oryza sativa* ssp. *indica* cv. MR219 (MR), BC₂F₇ line 7 and BC₂F₇ line 23 at different developmental stages. L: Leaf of 8th seedling day; W: Whole plant of 8th seedling day; B: Booting stage; H: Heading stage; F: Flowering stage; M: Milk grain stage.

Item	Sum of Squares	Degrees of Freedom	Mean Squares	Variances	P value summary	Probability
Between lines	1.389e+006	3	463072	34.55	***	< 1 %
Between stages	1.088e+006	5	217690	27.07	***	< 1 %
Interaction	1.169e+006	15	77960		***	< 1 %
Residual (error)	374322	72.0	5199			
Total	4.021e+006	95.0				

Table 4.5: Statistical analysis using two-way ANOVA with and putative CLV1 profile of Oryza rufipogon, Oryza sativa ssp. indica cv. MR219,

BC₂F₇ line 7 and BC₂F₇ line 23 collected at different developmental stages. * indicates that the degrees of statistical significance.

4.6 Southern Hybridization Analysis

Genomic DNA isolated from young leaves of *Oryza rufipogon* and *Oryza sativa* ssp. *indica* cv. MR219 was single digested with *Eco*RI, with *Bam*HI and with *Hind*III. The putative *RPK1* and putative *CLV1* cDNA fragments from *Oryza rufipogon* and *Oryza sativa* ssp. *indica* cv. MR219 respectively were used as probes for Southern hybridization. The length of putative *RPK1* probe was 320 bp, while the length of putative *CLV1* probe was 402 bp.

4.6.1 Gene Structure of Putative RPK1

Southern hybridization analysis showed the presence of one to three hybridized bands differing in size and with weak intensities (Figure 4.22). The putative *RPK1* probe did not contain *Eco*RI, *Bam*HI or *Hind*III restriction sites. The putative *RPK1* is probably present at one to three copies in the genomes of *Oryza rufipogon* and *Oryza sativa* ssp. *indica* cv. MR219.

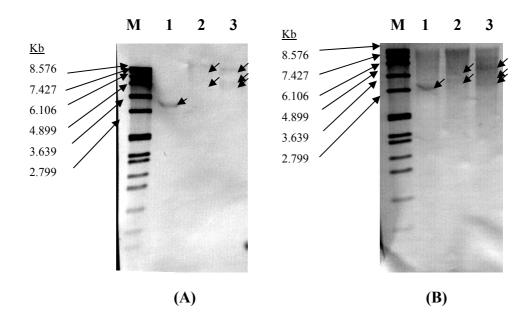


Figure 4.22: Southern hybridization analysis of putative *RPK1*. Genomic DNA of (A) *Oryza rufipogon* and (B) *Oryza sativa* ssp. *indica* cv. MR219 were digested with *Eco*RI (lane 1), *Bam*HI (lane 2) and *Hind*III (lane 3). Lane M: DIG-labeled DNA Molecular Weight Marker VII (Roche, Germany); Lane 1: Genomic DNA digested with *Eco*RI; Lane 2: Genomic DNA digested with *Bam*HI; Lane 3: Genomic DNA digested with *Hind*III.

4.6.2 Gene Structure of Putative *CLV1*

Southern hybridization analysis showed the presence of three to six hybridized bands differing in size and with strong and weak intensities (Figure 4.23). The putative *CLV1* probe did not contain *Eco*RI, *Bam*HI or *Hind*III restriction sites. The putative *CLV1* may be present at three to five copies in the genomes of *Oryza rufipogon* and *Oryza sativa* ssp. *indica* cv. MR219.

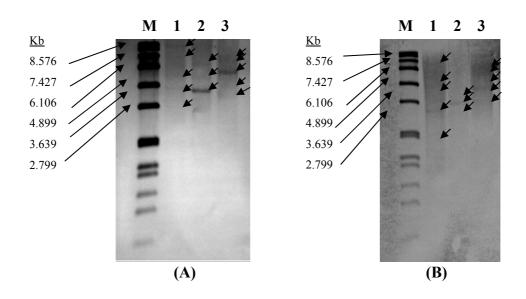


Figure 4.23: Southern hybridization analysis of putative *CLV1*. Genomic DNA of (A) *Oryza rufipogon* and (B) *Oryza sativa* ssp. *indica* cv. MR219 were digested with *Eco*RI (lane 1), *Bam*HI (lane 2) and *Hind*III (lane 3). Lane M: DIG-labeled DNA Molecular Weight Marker VII (Roche, Germany); Lane 1: Genomic DNA digested with *Eco*RI; Lane 2: Genomic DNA digested with *Bam*HI; Lane 3: Genomic DNA digested with *Hind*III.

4.7 Construction of RNAi Vectors for Knockdown of Gene Expression

4.7.1 Cloning of Putative *RPK1* and Putative *CLV1* Sequences into pANDA Vector

The putative *RPK1* and putative *CLV1* were successfully amplified from the cDNA of *Oryza rufipogon* (Figure 4.24). RT-PCR products of 105 bp for putative *RPK1* and 104 bp putative *CLV1* were amplified; each including the four bases "CACC" at the 5' end from each forward primer (see section 3.4) to allow for orientation based cloning into pENTR/D-TOPO (Invitrogen, California). After transformation in *E. coli*, putative positive colonies of pENTR/D-TOPO vector containing the sequences were confirmed by colony PCR (Figure 4.25) and DNA sequencing (Appendix G). The length of inserts in the putative positive clones of pENTR/D-TOPO vector containing the sequences of interest were 247 bp, including M13 forward (-20) priming site and *attL1* region.

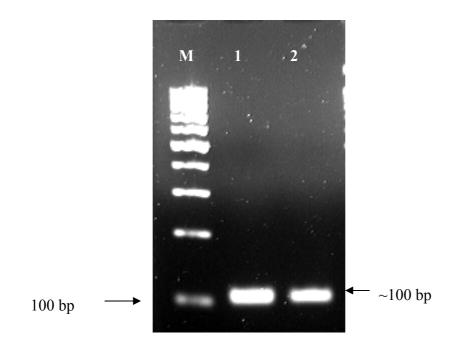


Figure 4.24: RT-PCR amplification of genes of interest. Lane M: 100 bp marker; Lane 1: RT-PCR amplified putative *RPK1* sequences; Lane 2: RT-PCR amplified putative *CLV1* sequences.

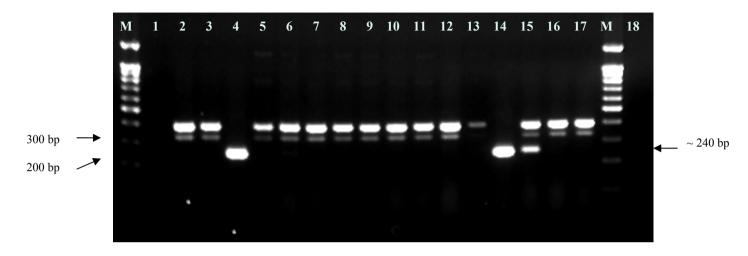


Figure 4.25: PCR confirmation of pENTR/D-TOPO cloning vector. Lane M: 100 bp marker; Lane 1: H₂O template; Lane 1-9: Putative subclone of putative subclone of putative *RPK1* sequences; Lane 10-18: putative subclone of putative *CLV1* sequences; Lane 14: Positive subclone of putative *CLV1* sequences; Lane 18: H₂O template.

Next, the pENTR/D-TOPO cloning vectors containing the sequences of interest were successfully mobilized into pANDA vector by an LR Clonase enzyme reaction. After transformation, putative positive colonies were confirmed by colony PCR (Figure 4.26) and DNA sequencing (Appendix G). Colony PCR was performed by using Gus linker primers and gene of interest primers: Gus linker forward and a gene of interest reverse primer were used to confirm the presence of the sense orientation, while a gene of interest reverse and Gus linker reverse primer were used to confirm the presence of the antisense orientation for each construct (Figure 4.26). The length of the PCR product of pANDA containing the gene sequences of interest was 984 bp, including 829 bp of Gus linker fragment and 50 bp *attB* sequences; whereas, the length of the PCR product antisense orientation of pANDA containing the gene sequences of interest was 887 bp, including 732 bp of Gus linker fragment and 50 bp *attB* sequences.

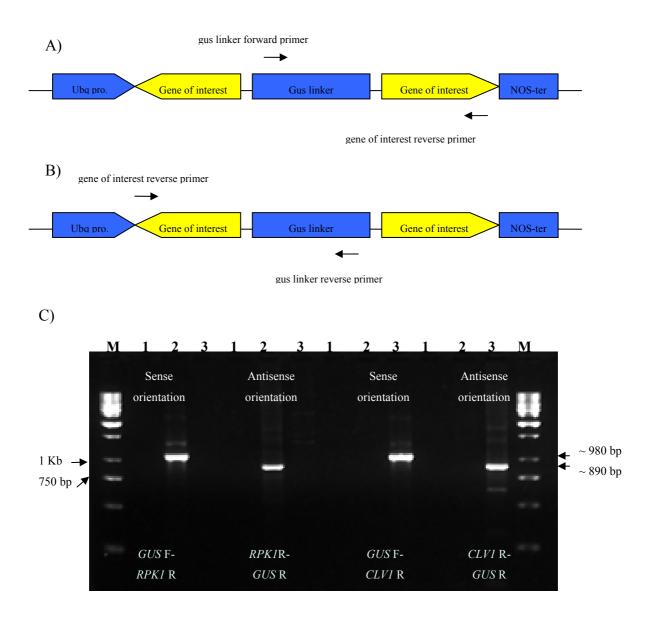


Figure 4.26: The sense orientation and antisense orientation of genes of interest for pANDA construct. (A) Gus linker forward primer and gene of interest reverse primer were used to check the sense orientation. (B) Gene of interest reverse primer and Gus linker reverse primer were used to check the antisense orientation. (C) PCR confirmation of vector construction orientation. Lane M: 1 kb marker; Lane 1: H₂O template; Lane 2: positive clone of pANDA containing putative *RPK1* sequences; Lane 3: positive clone of pANDA containing putative *CLV1* sequences.

4.7.2 Plant Transformation

pANDA vectors containing the sequences of putative *RPK1* (*pANDA_RPK1*) and putative *CLV1* (*pANDA_CLV1*) were transformed into *Agrobacterium tumefaciens* strain EHA105 through a freeze and thaw method (Jyothishwaran *et al.*, 2007). Next, positive *Agrobacterium* colonies were confirmed by PCR with a Gus linker primer (Figure 4.27). A 636 bp long fragment was successfully amplified from positive *Agrobacterium* colonies. Following this, *Agrobacterium tumefaciens* strain EHA105 containing *pANDA_CLV1* used to transform callus of *Oryza rufipogon* IRGC105491 (Sivakumar, unpublished data). A 636 bp long Gus linker fragment was successfully amplified from leaves of hygromycin resistant T₀ plants of *Oryza rufipogon* IRGC105491 by PCR with Gus linker primer (Figure 4.28).

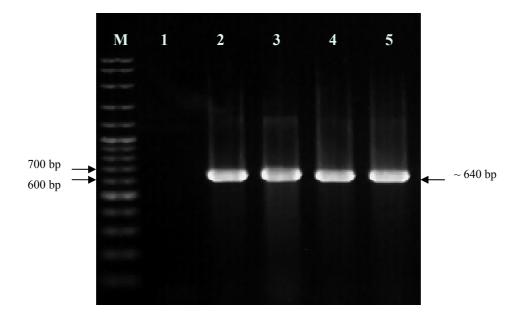


Figure 4.27: PCR screening of *Agrobacterium* colonies with Gus linker primer. Lane M: 100 bp DNA marker; Lane 1: H₂O template; Lane 2 and 3: positive clone of *pANDA_RPK1*; Lane 4 and 5: positive clone of *pANDA_CLV1*.

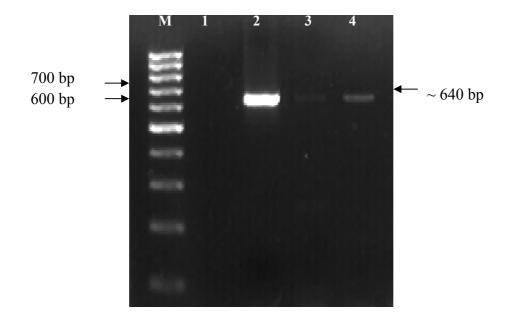


Figure 4.28: PCR screening of leaves hygromycin resistant T_0 plants at flowering stage with Gus linker primers. Lane M: 100 bp DNA marker; Lane 1: H₂O template; Lanes 2, 3 and 4: DNA from leaves of hygromycin resistant plants.