Plant Omics Journal POJ

POJ 6(6):425-432 (2013) ISSN:1836-3644

Genome-wide analysis of IQ67 domain (*IQD*) gene families in *Brachypodium distachyon*

Ertugrul Filiz[1]*, Huseyin Tombuloglu[2], Ibrahim Ilker Ozyigit[3]

[1]Department of Crop and Animal Production, Cilimli Vocational School, Duzce University, 81750, Cilimli, Duzce, Turkey
[2]Fatih University, Faculty of Science and Arts, Department of Biology, 34500, Buyukcekmece, Istanbul, Turkey
[3]Marmara University, Faculty of Science and Arts, Department of Biology, 34722, Goztepe, Istanbul, Turkey

***Corresponding author: ertugrulfiliz@gmail.com**

Abstract

In plants, Ca^{2+} concentration is important for the regulation of developmental processes and responses against biotic and abiotic stress factors. The eukaryotic Ca^{2+} binding protein calmodulin (CaM: CALcium MODULating proteIN) was found in *Arabidopsis* which contains a characteristic plant-specific IQ67 (Ile, Glu) domain (IQD). In this study, a genome wide analysis was performed in *Bracyhpodium distachyon* to identify *IQD* genes. Using several bioinformatics tools, we determined 23 Bd*IQD* genes which were distributed on all chromosomes and the highest gene number was detected on chromosome 2 including 12 *IQD* genes. 22 of the predicted proteins were considered to be basic proteins. Gene duplication analysis revealed that 8 of 23 Bd*IQD* genes were involved in duplication event, either segmental or tandem. Phylogenetic analysis showed that two main groups were observed in joined tree with rice and *Arabidopsis*. Especially, monocot species (*Brachypodium* and rice) were grouped together with the highest bootstrap value (100%), whereas monocot and dicot species (*Arabidopsis*) were clustered with lower bootstrap values. Digital expression profile analysis indicated that the most of the Bd*IQD* genes were expressed in leaves (8 genes) and flowers (6 genes), respectively. In conclusion, this comparative genomics analysis contributes to understanding *IQD* genes in grass species.

Key words: *IQD* genes, IQ67 domain**,** calmodulin (CAM), *Brachypodium*, genome wide analysis.
Abbreviations: IQD_IQ (Ile, Glu) 67 domain; EST_expressed sequence tag; CAM_calmodulin; CBL_calcineurin B-like protein; CDPK_calcium-dependent protein kinase; CMLs_CaM-related proteins; NCBI_National Center for Biotechnology Information.

Introduction

In plants, intracellular Ca^{2+} plays important role as a second messenger for many biological functions including structural integrity of the cell wall and the membrane system, elongation of root hairs, pollen tube formation, and response to biotic and abiotic stress conditions (Dodd et al., 2010; Reddy et al., 2011). The cellular energy metabolism can be damaged by high Ca^{2+} concentrations and it is adjusted by removing Ca^{2+} ion from the cytosol to the other cell components (Hetherington and Brownlee, 2004). Ca^{2+} signals are being decoded by various Ca^{2+}-binding proteins that regulate many biochemical changes and most of these proteins have the EF-hand motif.

To the present, several families of Ca^{2+} sensors have been determined in higher plants including calmodulin (CaM), calcineurin B-like protein (CBL), and calcium-dependent protein kinase (CDPK) (Ranty et al., 2006). CaMs and CBLs are small proteins, which have four and three EF-hand motifs, respectively. Also, they interact with target proteins and control their activities. CAMs target proteins have been found in higher plants and contain various functions with protein kinases, metabolic enzymes, cytoskeleton-associated proteins, and others (Reddy et al., 1996; Luan et al., 2002). CDPK includes a kinase domain and Ca^{2+} -binding domain with four EF-hand motifs. Ca^{2+} activates CDPKs, thus it may be related with Ca^{2+} -mediated signaling pathways (Harper et al., 2004). Many CaM target proteins are found in both plants and animals. CaM is a protein of 148 amino acids with 4 EF hands binds to Ca^{2+} ion (Strynadka and James, 1989). Plants contain CaM isoforms and CaM-like proteins and genetically distinct CaM isoforms display nearly 90% sequence identity in plants. Expression profile data indicate that *CaM* and *CML* (CaM-related proteins) genes are actively expressed in various organs and developmental stages, including plant development, plant-microbe interaction, and abiotic stress response. The analysis of these proteins revealed that some of them may have a role in signal transduction. Also, CMLs have significant structural divergences then the typical CaM. Amino acids substitutions occurred in EF-hand loop motif in CMLs proteins suggesting that these proteins may modify Ca^{2+} affinity (Lee et al., 2000; Ranty et al., 2006). 250 EF-hand encoding genes have been determined in the *Arabidopsis* genome with six typical calmodulins and 50 calmodulin-like proteins (Day et al., 2002). In *Arabidopsis*, CDPK and CBL proteins are encoded with 34 and 10 members, respectively (Abel et al., 2005).

Plant-specific *IQD* gene families have been analyzed in three genomes (*Arabidopsis*, rice, and tomato) with encoding about 30 predicted IQD proteins in each species. *IQD* genes were identified in *Physcomitrella* but algae have no any *IQD* genes suggests that IQD proteins are ancient family of CaM/CML-binding proteins that raised early evolution of land plants. IQD proteins have two common hallmarks: 67 conserved amino acid residues as known IQ67 domain and highly conserved exon-intron boundary with interrupts codons 16 and 17 via a phase-0 intron (Abel et al., 2005; Abel et al., 2013). 33 and 29 *IQD1*-like genes were identified with IQ67 domain in *Arabidopsis thaliana* and *Oryza sativa*. CaM target proteins (known as IQD family) include 33

members in *A. thaliana* and central domain contains 67 conserved amino acid residues (as called IQ67 domain) with CaM EF-hand motifs. Molecular mass of IQD proteins are structurally diverse (~12–87 kDa), but isoelectric points are quite uniform (pI~10.3). Also, high fractions of Arg/Lys (~17%) and Ser/Lys (~20%) residues were observed in IQDs (Abel et al., 2005). In *Arabidopsis*, *IQD1* was proposed to regulate defense metabolism by deciphering intracellular Ca^{2+} signals. According to analysis of steady-state mRNA levels, *Arabidopsis IQD*1 affects glucosinolate metabolism (Abel et al., 2005; Levy et al., 2005). In addition to *IQD*1, *IQD*22 acts as a negative regulator of the response to the plant hormone gibberellins in *Arabidopsis* (Zentella et al., 2007). A genetic analysis showed that *IQD*12/*SUN* is a major gene responsible for fruit shape variation by retrotransposon-mediated *IQD*12 duplication in tomato. Also, *IQD*12/*SUN* could affect plant developmental processes, including the shape of cotyledons, leaves, floral organs, and other morphological alterations (Xiao et al., 2008; Wu et al., 2011). In another study, 6 calmodulin (*CAM*) genes with encoding only 3 isoforms and 50 *CAM-related* (*CML*) genes were identified in *Arabidopsis*. The six *Arabidopsis CAM* genes had only one intron, while 13 of the 50 *CMLs* had two or more introns. Also, the expressed sequence tag (EST) data showed that some of the *CAM*/*CML* genes could be regulated stress or treated with hormones (McCormack and Braam, 2003). In this study, we investigated *IQD* genes in whole genome scale based on the sequence resources of *Brachypodium* genome. Also, *Brachypodium IQD* gene number, their distribution and expansion in the genome, characteristics of motifs, phylogenetic relationship, and digital expression profile were analyzed.

Results

Genome wide identification of IQD gene family

Following local database searches for *IQD* genes by using BLASTP, 23 nonredundant IQD protein sequences in *Brachypodium* genome were found. *IQD* genes were distributed on all chromosomes (Table 1). The highest chromosome numbers were identified on chromosome 2 with 12 *IQD* genes, whereas the least number was found on chromosome 4 and 5 including one *IQD* gene. Schematic structures of Bd*IQD* genes were obtained by using GSDS (Gene structure display server) program (Fig. 1). Average exon and intron numbers were found as 4.5 and 3.5, respectively. There were no intronless Bd*IQD* genes and exon numbers varied between 3 and 6. Also, all introns of most BdIQD genes are phase-0 introns. The length of the IQD proteins ranged from 340 to 583 amino acids. Except Bd*IQD*14 proteins (pI≤7, in acidic character), the other proteins (22, 95.6%) were in basic character (pI≥7). Putative subcellular localizations were predicted to be localized in the mitochondria (6), chloroplast (4), and any other locations by a TargetP analysis. PSORT analysis revealed that IQD proteins are predicted to be localized in nucleus (16), mitochondrion (2), and chloroplast (5). Also, molecular weight (kDa) and ORF length were ranged from 38.39 (Bd*IQD*18) to 63.93 kDa (Bd*IQD*9) and from 1023 (Bd*IQD*18) to 1752 bp (Bd*IQD*6), respectively (Table 1). The IQ67 domain includes 1–3 copies each of the IQ motif (IQxxxRGxxxR) or [ILV]QxxxRxxxx [R, K]), the 1-5-10 motif ([FILVW]x_3[FILV]x_4[FILVW]), and the 1-8-14 motif ([FILVW]x_6[FAILVW]x_5[FILVW]). In addition, these motifs contain some basic and hydrophobic amino acid residues (Abel et al., 2005; Levy et al., 2005). All predicted IQD proteins had a typical "IQ calmodulin-binding motif" domain (PF00612) which is a major calcium (Ca^{2+}) regulator (Fig. 2). Thus, it was observed that IQ67 domain located core region of BdIQD proteins. The most abundant amino acid residues were found as Ala (12.58%), Ser (11.50%), and Arg (10.15%), respectively with average values. Beside, average frequency of Lys was to found as 5.95% (Supplementary Table 1).

IQD gene duplications in the Brachypodium genome

According to the chromosomal distribution analysis, 23 *IQD* genes were found to be dispersed over each chromosome (Fig. 3). The number of *IQD* genes distributed to per chromosome (Chr1-5) was as, 7, 12, 2, 1, and 1, respectively. The highest *IQD* density was determined in Chr2 and the lowest in Chr4 and Chr5. In general, *IQD* genes were not found residing around the centromere regions of chromosomes. According to gene duplication prediction of *IQD* genes, two segmental and four different tandem duplications were determined. As it was indicated in Fig. 3, tandem duplications were arranged between the genes *IQD*04:*IQD*05, *IQD*20:*IQD*21, *IQD*22:*IQD*23, and *IQD*06:*IQD*10. Besides, *IQD*04:*IQD*05 and *IQD*20:*IQD*21 genes seem to be segmental duplicated in the *Brachypodium* genome.

Conserved motifs and phylogenetic analysis of Bd*IQD* genes in *Brachypodium*

A total number of 23 BdIQD protein sequences were submitted to MEME suite to identify conserved domains or motifs and five different common motifs were observed (Table 2 and Fig. 4). Motif I, II, III, IV, and V were observed in IQD protein sequences as 22, 13, 15, 18, and 17 times, respectively. In other words, motif 1, 2, 3, 4, and 5 were absent as 1, 10, 8, 5, and 6 times in BdIQD proteins, respectively. The most common motif was found motif 1 with 22 times. Interestingly, motif I containing IQ67 domain was absent only in BdIQD14 protein sequence. Motif I and motif IV appeared twice in BdIQD4 and BdIQD17, respectively. Motif V was located at the first position of 17 protein sequences; in contrast motif IV was located in the end of 18 protein sequences (Fig. 4). Although, the minimum motif numbers were observed in BdIQD14 (motif II and motif III), the highest motif numbers (motif I, II, III, IV, and V) were observed in BdIQD3, BdIQD8, BdIQD15, BdIQD16, and BdIQD20. To understand the evolutionary relationships within the *IQD* gene families, we performed a combined phylogenetic analysis including *Brachypodium* (23 members), *Arabidopsis* (5 members), and rice (4 members) by using a neighbor-joining (NJ) tree (Fig. 5). The analysis revealed that phylogenetic tree was separated the IQD proteins into two main groups. Main group I had three subgroups: subgroup A (16 sequences), subgroup B (8 sequences), and subgroup C (6 sequences) while main group II had one clade (2 sequences). Intrestingly, BdIQD14 and BdIQD18 (in maingrup II) were seperated the other BdIQDs. Based on pI values, all members of main group I were basic characters, whereas BdIQD14 was acidic character in main group II with lower bootstrap value (77%). Four *Arabidopsis* sequences (AtIQD1, AtIQD2, AtIQD3, and AtIQD4) were grouped in subgroup C, while the other one was in subgroup B. Three rice sequences (OsIQD1, OsIQD2, and OsIQD4) were in subgroup A and the other one (OsIQD3) was in subgroup B. *Brachypodium* and *Arabidopsis* did not cluster in same clade, whereas *Brachypodium* and rice were grouped

Table 1. List of 23 *IQD* genes and proteins identified in *Brachypodium* with their physiochemical, structural, and sequence properties.

Gene name	Sequence ID	Chr[a]	ORF length (bp)	Exon number	Length (aa)	MW (kDa)	*pI*	PSORT Predicted Location[b]	TargetP Predicted Location[c]
Bd*IQD*1	Bradi4g01360.2	4	1308	6	435	47.58	9.95	N	?
Bd*IQD*2	Bradi2g05680.1	2	1482	3	493	52.91	10.22	N	M 0.85/2
Bd*IQD*3	Bradi2g05840.1	2	1320	5	439	48.46	10.48	M	?
Bd*IQD*4	Bradi1g06350.1	1	1248	6	415	46.36	9.60	M	M 0.88/4
Bd*IQD*5	Bradi1g13650.3	1	1260	6	419	46.29	9.67	N	?
Bd*IQD*6	Bradi3g13710.1	3	1752	5	583	63.22	10.89	N	C 0.80/2
Bd*IQD*7	Bradi1g14090.1	1	1392	3	463	49.57	10.42	C	?
Bd*IQD*8	Bradi2g18220.1	2	1308	5	435	46.94	10.81	C	C 0.75/3
Bd*IQD*9	Bradi2g18640.2	2	1731	6	576	63.93	9.78	N	?
Bd*IQD*10	Bradi5g18767.1	5	1356	4	451	48.83	10.55	C	M 0.75/4
Bd*IQD*11	Bradi2g19410.1	2	1107	3	368	40.38	10.67	N	?
Bd*IQD*12	Bradi3g26990.1	3	1431	3	476	51.52	10.11	C	?
Bd*IQD*13	Bradi2g33370.1	2	1419	3	472	50.73	10.59	N	C 0.47/5
Bd*IQD*14	Bradi1g36140.1	1	1311	6	436	46.56	6.44	N	M 0.75/5
Bd*IQD*15	Bradi2g37980.2	2	1425	5	474	51.82	10.17	N	?
Bd*IQD*16	Bradi2g47867.1	2	1281	4	426	46.43	10.61	N	?
Bd*IQD*17	Bradi2g48210.2	2	1665	6	554	60.78	9.94	N	C 0.44/5
Bd*IQD*18	Bradi1g49070.2	1	1023	5	340	38.39	10.74	N	?
Bd*IQD*19	Bradi2g49490.1	2	1173	3	390	43.41	10.55	N	M 0.82/3
Bd*IQD*20	Bradi2g54170.2	2	1479	5	492	54.97	10.25	N	?
Bd*IQD*21	Bradi2g57567.1	2	1734	6	577	63.42	11.52	C	M 0.77/4
Bd*IQD*22	Bradi1g62220.1	1	1476	5	491	54.22	10.31	N	?
Bd*IQD*23	Bradi1g74410.1	1	1233	2	410	44.50	11.24	N	?

[a] Chromosome number in which the gene resides. [b] PSORT predictions: N (nucleus), C (chloroplast), M (mitochondrion). [c] Localization of BdIQD protein supported by Target P. TargetP predictions: values indicate score (0.00 – 1.00) and reliability class (1–5; best class is 1). (M: mitochondrion, N: nucleus, and C: chloroplast)

Table 2. Different motifs commonly observed in IQD protein sequences with best possible match amino acid sequences. Red residues represents IQ67 domain.

Motif number	Width Sequence	Protein sequences
1	50	IQTAFRGYLARRALRALKGIVKLQALVRGHLVRKQAAHTLHCMQALVRVQ
2	32	SVEEIQAKIHMRQEAAIKRERAMAYAFSHQWW
3	27	YECDNNNWGWNWLERWMAARPWENRLM
4	21	PNYMANTESFKAKLRCQCAPK
5	16	MGKAGKWIKSLLGGKK

together (BdIQD22- OsIQD1 and BdIQD4- OsIQD3) in same clades with the highest bootstrap value (100%). Briefly, IQD proteins of *Brachypodium* and rice genes appear more closely related to each other than *Arabidopsis*. In *Brachypodium*, the highest bootstrap values (100%) were observed between BdIQD11-BDIQD19 and BdIQD9-BdIQD17 in subgroup A. The subgroup A had only grass species with *Brachypodium* and rice, the subgroup B contained all species (*Brachypodium*, *Arabidopsis*, and rice), and the subgroup C had *Brachypodium* and *Arabidopsis*.

Digital expression analysis

EST database can be used *in silico* analysis to study gene expression and there are about 130.000 ESTs were deposited with *Brachypodium*. To analysis the expression of the Bd*IQD* gene family, NCBI EST database was used. The coding sequences of Bd*IQD* genes were used to predict Bd*IQD* transcripts using megablast tool. According to tissue and organ types, the *Brachypodium* ESTs were divided into 4 groups (Table 3). Based on NCBI EST expression database, Bd*IQD*2, 6, 7, 8, 10, 11, 12, 13, 14, 16, 19, 21, 22, and 23 had no expression, while 9 Bd*IQD* genes (Bd*IQD*1, 3, 4, 5, 9, 15, 17, 18, and 20) were found in different organs or tissues (leaf, root, and flower). Also, Bd*IQD*3, Bd*IQD*4, and Bd*IQD*18 were observed in only one tissue or organ.

Discussion

In this study, 23 non-redundant *IQD* genes were identified in *Brachypodium* genome. Abel et al. (2005) reported that 33 and 29 *IQD*1-like genes were identified in *Arabidopsis thaliana* and *Oryza sativa*, respectively. *Brachypodium* had less *IQD* genes than *Arabidopsis* and rice. In these plants, genome sizes are significantly different such as *Brachypodium* (300Mb), *Arabidopsis* (164 Mb), and *Oryza* (441 Mb) (Opanowicz et al., 2008). At least four different large-scale duplication events were detected in *Arabidopsis* genome in 100 to 200 million years ago and 17% of all genes were shaped in tandem arrays (Vision et al., 2000; Simillion et al., 2002). These duplication events may help us to understand the highest number of *IQD* genes although the smallest genome size of *Arabidopsis*. Average physicochemical parameters of *IQD* genes and proteins were

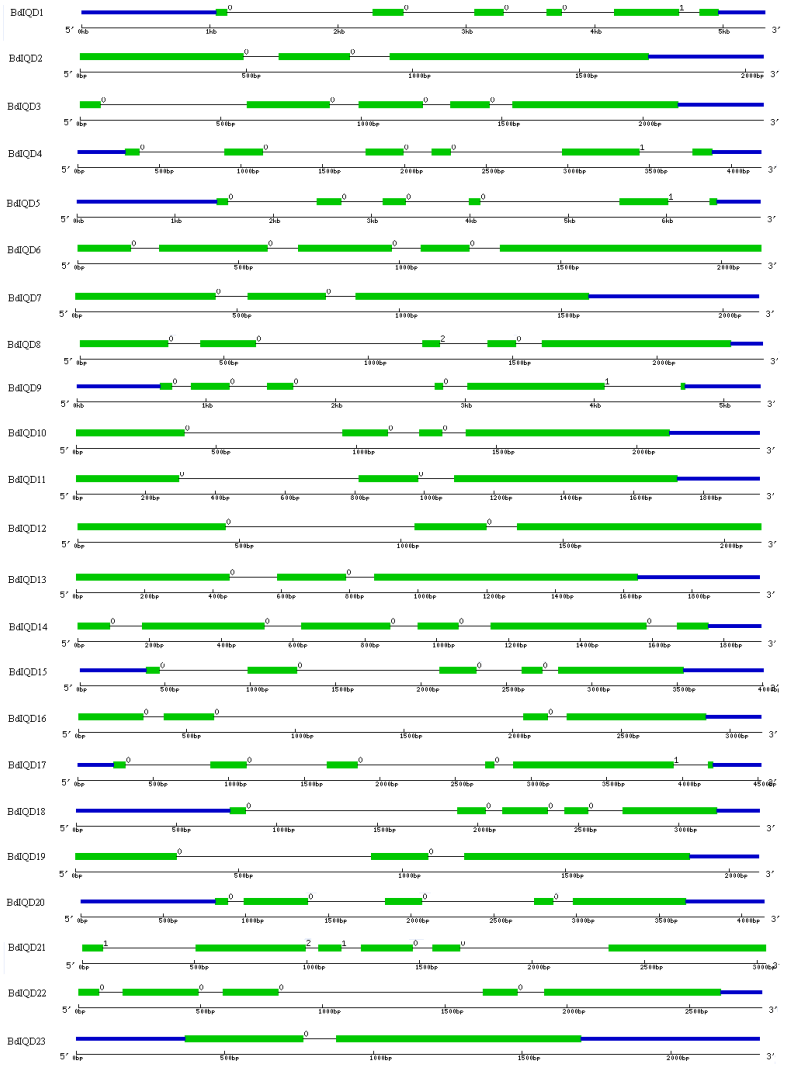

Fig 1. Gene structure of *BdIQD* genes. Exons and introns were shown by filled green boxes and single lines, respectively. UTRs were displayed by thick blue lines at both ends. Intronic phases were indicated by numbers 0, 1, and 2.

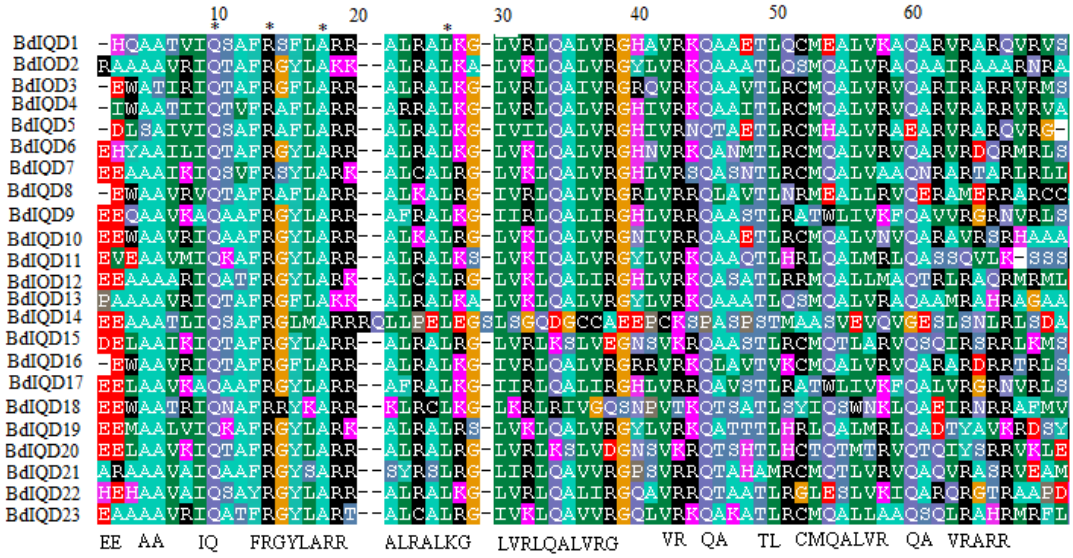

Fig 2. Amino acid sequence alignments of IQ67 domain (Abel et al., 2005). The multiple alignment results indicate the highly conserved IQD domains among putative 23 *Brachypodium* IQD protein sequences and the position of the conserved IQ calmodulin-binding motifs were shown. Symbol * shows identical residues for proteins and the consensus sequence at the bottom was consructed with greater than 50% conservation among 23 proteins.

Table 3. Digital expression analysis of Bd*IQD* genes

Gene	Tissue and organ type (NCBI)			
	Leaf	Root	Flower	Mixed
Bd*IQD*1	+		+	+
Bd*IQD*2				
Bd*IQD*3			+	
Bd*IQD*4	+			
Bd*IQD*5	+		+	+
Bd*IQD*6				
Bd*IQD*7				
Bd*IQD*8				
Bd*IQD*9	+	+	+	+
Bd*IQD*10				
Bd*IQD*11				
Bd*IQD*12				
Bd*IQD*13				
Bd*IQD*14				
Bd*IQD*15	+	+	+	+
Bd*IQD*16				
Bd*IQD*17	+		+	+
Bd*IQD*18	+			
Bd*IQD*19				
Bd*IQD*20	+	+		+
Bd*IQD*21				
Bd*IQD*22				
Bd*IQD*23				

+: Expressed; blank: not expressed.

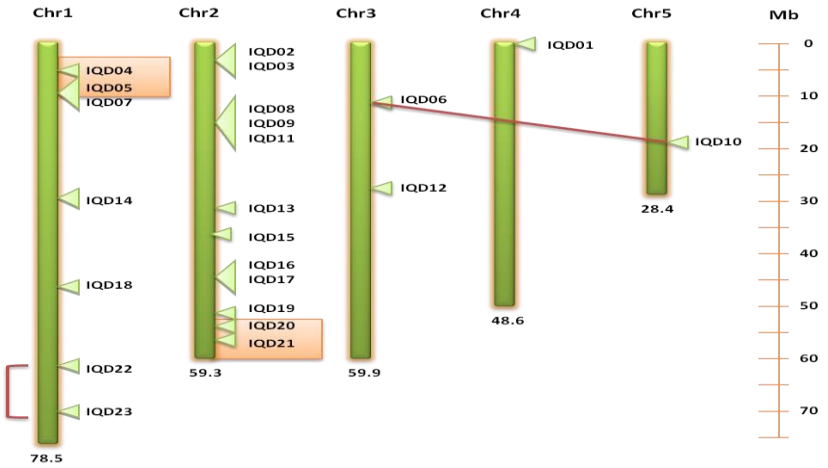

Fig 3. Location and duplication analysis of *Brachypodium IQD* genes on chromosomes 1 to 5. Colored boxes (orange) and red lines show tandem and segmental duplicated regions and genes, respectively. The ruler represented in mega bases (Mb).

reported including gene length (2.4-3 kb), exon numbers (4.5-4.4), protein length (454-471 aa), molecular mass (50.8-51.4 kD), and isoelectric point (10.3-10.4) in *A. thaliana* and *O. sativa*, respectively (Abel et al., 2005). We found these parameters as gene length (1.38), exon numbers (4.5), protein length (461), molecular mass (50.49), and isoelectric point 10.3, respectively. *IQD* gene length of *Brachypodium* was smaller than *Arabidopsis* and rice but average exon number was same with *Arabidopsis* and so similar to rice. Average frequency of Arg (9.3-10.6%), Lys (8.3-5.9%), Ser (12.2-10.2%), and Ala (8.6-12.8%) was found in *Arabidopsis* and rice, respectively (Abel et al., 2005). In this study, the average frequency of Arg, Lys, Ser, and Ala was observed at 10.15%, 5.95%, 11.50%, and 12.58% in *Brachypodium*, respectively. These data were very similar to *Arabidopsis* and rice (Abel et al., 2005). It was understood that physiochemical properties of BdIQD proteins were very similar to *Arabidopsis* and rice. Also, physiochemical data of BdIQD proteins may support that ancestral *IQD* genes were existed in plants before the monocot-dicot divergence.

Identical/similar DNA sequences empower genome redundancy with decreasing its complexity. Also, coding and non-coding DNA sequences were affected by redundancy. Therefore, gene families were shaped by gene duplications (Grassi et al., 2008). In our study, we detected one segmental and four tandem duplications and 8 of 23 Bd*IQD* genes were found to be involved in duplication event. Duplications and the increasing copy number of *IQD* genes in *Brachypodium* could make contribution to plant genetic rearrangement. Since the gene duplications might be one of the major evolutionary forces for new protein functions (Kondrashov et al., 2002), *IQD* gene duplications can contribute to plant adaptation to variable environmental conditions. Also, orthology and paralogy are the key concepts in the field of

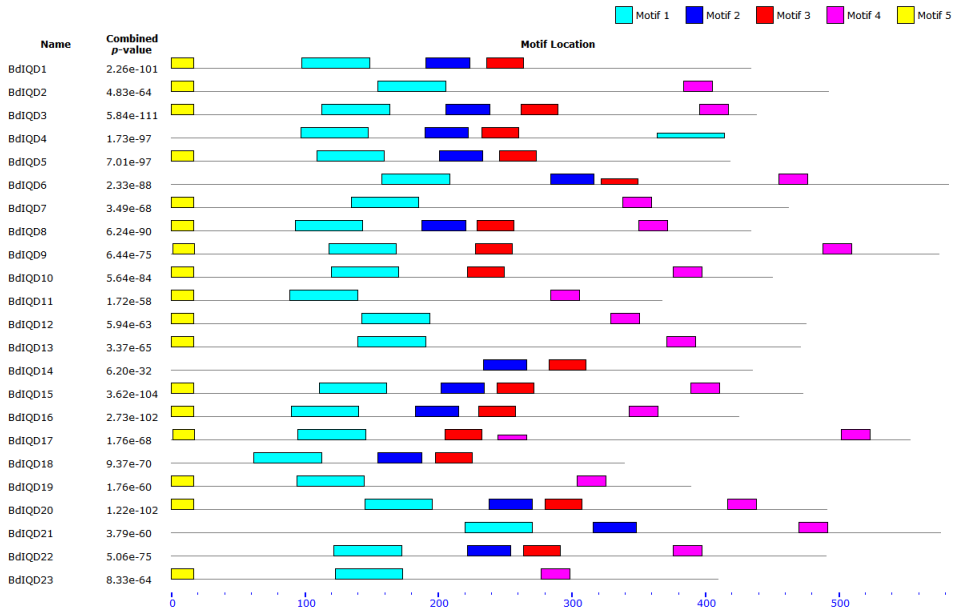

Fig 4. Predicted motif distribution in *Brachypodium* IQD proteins by using MEME server. Each motif was represented in boxes with different colors: Motif 1, cyan; Motif 2, blue; Motif 3, red; Motif 4, pink; and Motif 5, yellow.

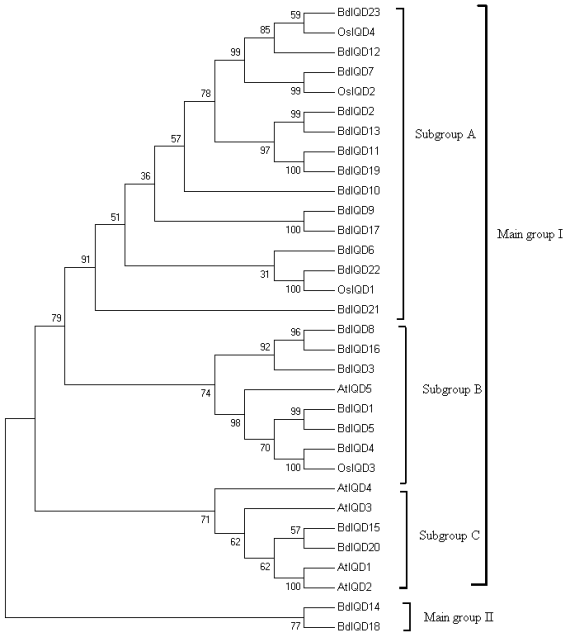

Fig 5. Phylogenetic relationship of IQD proteins among *Brachypodium*, *Arabidopsis*, and rice. The analysis revealed that the phylogenetic tree was divided into two main maingoups (Main group I and II). Except BdIQD14 and BdIQD18, others were composed of the main group I which was sub-divided into 3 subgroups (Subgroup A, B, and C).

protein and proteome evolution (Makarova et al., 1999). Considering to their sequence similarity, the genes involved in the duplication event seem to be paralogues. 19 of 23 *IQD* genes (82.6%) were distributed on chromosome 1 and 2; it was observed non-uniform distribution of *IQD* genes. Chromosome number of *B. distachyon* was mostly found in 10, 20 or 30 (Draper et al., 2001) and polyploidy nature of *B. distachyon* was discussed widely (Jaroszewicz et al., 2012). This *IQD* gene distribution of *Brachypodium* may be affected by ancient polyploidy or aneuploidy events than whole-genome duplication event during evolutionary history. Our hypothesis was agreement with *IQD* genes in rice genome (Abel et al., 2005). Predicted *IQD*-like gene in *Physcomitrella patens* showed that the *IQD* gene family evolved during the early evolution of land plants 450–700 Myr ago (Hedges, 2002). *Brachypodium* genus diverged from the *Pooideae* subfamily in family *Poaceae* approximately 35–40 Mya (Catalan et al., 1997). The genus *Brachypodium* belong to the *Pooideae* subfamily diverged from the subfamily *Ehrhartoideae* which contains rice approximately 50 million years ago (Kellogg, 2001). The phylogenetic trees of the 33 *Arabidopsis* and 28 rice *IQD* genes revealed closely clustered nodes and a high degree of sequence divergence (Abel et al., 2005). In our combined phylogenetic tree, *Brachypodium* and rice (monocots) were clustered particularly with high bootstrap value (100%) for internal nodes; by contrast, the outer nodes received lower bootstrap support. This finding supports systematic data that

Brachypodium and rice belong to *Poaceae* family and it may reflect common origin for *IQD* genes in monocots. IQD proteins of *Arabidopsis* were clustered with high bootstrap value (100%) in subgroup C for internal node, whereas outer nodes had lower bootstap value (62% and 71%) in subgroup C. It can be thought that monocot-dicot divergence of *IQD* genes may be responsible for occurring different bootstrap values in combined tree among dicot (*Arabidopsis*) and monocot (*Brachypodium* and rice) species. Expression profile of Bd*IQD* genes revealed that 14 of 23 *IQD* genes had no expression data. There were many reasons for explaining this situation. Possible explanation of this situation could be that some Bd*IQD* genes exhibit temporal and spatial expression pattern. For example, *IQD*1 gene regulates defense metabolism in biotic challenge (Levy et al., 2005) and *IQD*22 is a negative regulator in response to the plant hormone gibberellins (Zentella et al., 2007). These data proved our hypothesis that some *IQD* genes are related to special physiological and biochemical conditions. The results presented here support to understand *IQD* genes in plants and to contribute to studies of genome wide analysis in future.

Materials and Methods

Isolation of IQD genes in Brachypodium

The sequences of 31 *Arabidopsis* and 10 rice IQD protein sequences were downloaded the TAIR database (http://www.arabidopsis.org) and TIGR database (http://rice.plantbiology.msu.edu). For the identification of *Brachypodium IQD* gene family, *Arabidopsis* and rice IQD protein sequences were firstly used as query sequences to search against *Brachypodium distachyon* genome (http://www.phytozome.net/search.php) at the Joint Genome Institute (JGI) (http://www.phytozome.net) using BLASTP program. The sequences were selected as predicted proteins if their E value satisfied E was $\leq e^{-10}$ and redundant sequences were removed. Also, all candidate sequences were analyzed in the Pfam (PF00612) (Sonnhammer et al., 1997) and SMART (Letunic et al., 2009) database for detecting IQ domain.

Analysis of protein sequences and motifs

Physiochemical data were generated from the Expasy's ProtParam server (Gasteiger, 2005) including sequence length, molecular weight, and theoretical isoelectric point (pI) values (Table 1). Conserved motif of IQD proteins were identified by using MEME suite (http://meme.sdsc.edu/meme4_4_0/intro.html) (Timothy et al., 2009). The following parameters were adopted: (1) the optimum motif width was set to ≥6 and ≤50; (2) the maximum number was set to identify 5 motifs. The subcellular distribution of the IQD proteins was predicted by using TargetP 1.1 (http://www.cbs.dtu.dk/services/TargetP/) server. Exons and introns structures of Bd*IQD* genes were determined by using GSDS (Gene structure display server) (http://gsds.cbi.pku.edu.cn/) (Guo et al., 2007).

Phylogenetic analysis

Amino acid sequences of the predicted BdIQD proteins were aligned using Clustal W (Thompson et al. 1994). Phylogenetic analysis was performed with MEGA 5.1 (Tamura et al. 2011) by a neighbor-joining (NJ) tree based on the multiple sequence alignment of all predicted IQD protein sequences including the following parameters: Poisson correction, pairwise deletion, and bootstrap analysis with 1000 replicates. Also, *Arabidopsis* (5 members) and rice (4 members) protein sequences were used as out-groups in phylogenetic tree.

Digital expression profiles of the Brachypodium IQD genes

Transcript levels of *Brachypodium IQD* genes were analyzed in NCBI. EST mining was performed in the NCBI EST database (http://www.ncbi.nlm.nih.gov/dbEST/) using megablast tool. Parameters of searching were as followings: maximum identity ≥95%, length ≥200 bp and E value $\leq 10^{-10}$ (Wang et al., 2010). Furthermore, EST mining was performed in the PlantGDB database (http://www.plantgdb.org/) with default parameters.

Conclusions

Calmodulin (CaM) and CaM-related (CML) proteins are related with Ca^{+2} and generated by plants in various biotic and abiotic changes. We have performed genome wide analysis in *Brachypodium* genome and 23 predicted calmodulin target proteins were identified. Based on phylogenetic analysis, two main groups were observed and main group I was organized into 3 subgroups (A, B, and C). Especially, *Brachypodium* and rice were clustered together with the highest bootstrap value (100%). Our analysis revealed that segmental and tandem duplications were contributed to *IQD* gene expansion in *Brachypodium.* Therefore, gene duplication events may have caused gene diversification in expansion of the *IQD* gene family. 14 of 22 *IQD* genes had no expresion data. This finding may be related with various expression patterns of *IQD* genes by stress conditions in different plant tissues. Our study results will contribute to genome wide analysis of *IQD* gene families in different plant species. Also, new experimental studies are needed to understand *IQD* gene functions in plants.

Acknowledgements

We would like to thank Miss. Berçem Yeman for the manuscript proofreading.

References

Abel S, Savchenko T, Levy M (2005) Genome-wide comparative analysis of the *IQD* gene families in *Arabidopsis thaliana* and *Oryza sativa*. BMC Evol Biol 5:72

Abel S, Bürstenbinder K, Müller J (2013) The emerging function of IQD proteins as scaffolds in cellular signaling and trafficking. Plant Signal Behav. 26:8(6)

Catalan P, Kellogg EA, Olmstead RG (1997) Phylogeny of Poaceae subfamily Pooideae based on chloroplast *ndhF* gene sequences. Mol Phylogenet. Evol. 8:150–166

Day IS, Reddy VS, Shad Ali G, Reddy AS (2002) Analysis of EF-hand-containing proteins in *Arabidopsis*. Genome Biol 3:RESEARCH0056

Dodd AN, Kudla J, Sanders D (2010) The language of calcium signaling. Annu Rev Plant Biol. 61:593–620

Gasteiger E (2005) Protein Identification and Analysis Tools on the ExPASy Server. In: John M. Walker ed, The Proteomics Protocols Handbook, Humana Press:571-607

Grassi AD, Lanave C, Saccone C (2008) Genome duplication and gene-family evolution: The case of three *OXPHOS* gene families. Gene 421:1–6

Guo AY, Zhu QH, Chen X, Luo JC (2007) GSDS: a gene structure display server. Yi Chuan 29(8):1023–1026.

Harper JF, Breton G, Harmon A (2004) Decoding Ca^{2+} signals through plant protein kinases. Annu Rev Plant Biol 55:263-88

Hedges SB (2002) The origin and evolution of model organisms. Nat Rev Genet 3(11):838-849

Hetherington AM, Brownlee C (2004) The generation of Ca^{2+} signals in plants. Annu Rev Plant Biol 55:401-27

Jaroszewicz AM, Kosina R, Stankiewicz PR (2012) RAPD, karyology and selected morphological variation in a model grass, *Brachypodium distachyon*. Weed Res. 52: 204–216

Kellogg EA (2001) Evolutionary history of the grasses. Plant Physiol 125:1198–1205

Lee SH, Johnson JD, Walsh MP, Van Lierop JE, Sutherland C et al. (2000) Differential regula-tion of Ca^{2+}/calmodulin-dependent enzymes by plant calmodulin isoforms and free Ca^{2+} concentration. Biochem J 350:299-306

Letunic I, Doerks T, Bork P (2009) SMART 6: recent updates and new developments. Nucleic Acids Res. 37: 229-232

Levy M, Wang Q, Kaspi R, Parrella MP, Abel S (2005) *Arabidopsis* IQD1, a novel calmodulin-binding nuclear protein, stimulates glucosinolate accumulation and plant defense. Plant J. 43:79–96

Luan S, Kudla J, Rodriguez-Concepcion M, Yalovsky S, Gruissem W (2002) Calmodulins and calcineurin B-like proteins: calcium sensors for specific signal response coupling in plants. Plant Cell 14 Suppl: 389-400

McCormack E, Braam J (2003) Calmodulins and related potential calcium sensors of *Arabidopsis*. New Phytologist. 159:585–598

Opanowicz M, Vain P, Draper J, Parker D, Doonan JH (2008) *Brachypodium distachyon*: making hay with a wild grass. Trends Plant Sci. 13:172-177

Ranty B, Aldon D, Galaud JP (2006) Plant calmodulins and calmodulin-related proteins. Plant Signal Behavior. 1:96-104

Reddy AS, Narasimhulu SB, Safadi F, Golovkin M (1996) A plant kinesin heavy chain-like protein is a calmodulin-binding protein. Plant J. 10:9–21

Reddy AS, Ali GS, Celesnik H, Day IS (2011) Coping with stresses. Roles of calcium- and calcium/calmodulin-regulated gene expression. Plant Cell 23:2010–2032

Simillion C, Vandepoele K, Van Montagu MC, Zabeau M, Van de Peer Y (2002) The hidden duplication past of *Arabidopsis thaliana*. Proc Natl Acad Sci USA 99(21):13627-13632

Sonnhammer EL, Eddy SR, Durbin R (1997) Pfam: a comprehensive database of protein domain families based on seed alignments. Proteins 28:405–420

Strynadka NCJ, James MNG (1989) Crystal structure of the helix-loop-helix calcium-binding proteins. Annu Rev Biochem. 58:951–998

Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 22:4673–4680

Timothy L, Mikael Bodén B, Buske FA, Frith M, Grant CE, Clementi L, Ren J, Li WW, Noble WS (2009) MEME SUITE: tools for motif discovery and searching. Nucleic Acids Res 37:202-208

Vision TJ, Brown DG, Tanksley SD (2000) The Origins of Genomic Duplications in *Arabidopsis*. Science. 290:2114-2117

Wang Y, Deng D, Bian Y, Lv Y, Xie Q (2010) Genome-wide analysis of primary auxin-responsive *Aux/IAA* gene family in maize (*Zea mays* L.). Mol Biol Rep 37:3991–4001

Wu S, Xiao H, Cabrera A, Meulia T, Van derKnaap E (2011) SUN regulates vegetative and reproductive organ shape by changing cell division patterns. Plant Physiol. 157:1175–1186

Xiao H, Jiang N, Schaffner E, Stockinger EJ, Van der Knaap E (2008) A retrotransposon-mediated gene duplication underlies morphological variation of tomato fruit. Science 319:1527–1530

Zentella R, Zhang ZL, Park M, Thomas SG, Endo A, Murase K, Fleet CM, Jikumaru Y, Nambara E, Kamiya Y, Sun TP (2007) Global analysis of della direct targets in early gibberellin signaling in *Arabidopsis*. Plant Cell 19:3037–3057