Available online freely at www.isisn.org

Bioscience Research

Print ISSN: 1811-9506 Online ISSN: 2218-3973

Journal by Innovative Scientific Information & Services Network

RESEARCH ARTICLE BIOSCIENCE RESEARCH, 2019 16(2): 1679-1690. **OPEN ACCESS**

In silico characterization of CDKB1 and its coding gene in some *Oryza* species

Ahmed Sayed Fouad*and Rehab Mahmoud Hafez

Botany and Microbiology Department, Faculty of Science, Cairo University, 12613 Cairo, **Egypt**.

***Correspondence:** ahmedsfouad@yahoo.com **Accepted: 04** May 2019 **Published online: 28** May. 2019

Rice (*Oryza sativa*) is one of the major staple foods for about 70% of the world population. Breeding programs for rice suffer from limited genetic diversity in cultivated rice. Estimated to be 10-20% of that in wild *Oryza* species. *In silico* analysis of genomes of wild species provide valuable information contributing effectively in development of new cultivars of better performance. CDKB1 members control M phase in cell cycle and involved in homologous recombination DNA repair that helps normal cell division under stressful conditions. In this study, a CDKB1 gene and protein were characterized in *O. nivara, O. sativa* Indica gp, *O. sativa* Japonica gp, *O. rufipogon, O. barthii, O. glumipatula, O. glaberrima, O. meridionalis, O. punctate, O. brachyantha.* Only *O. brachyantha* that exhibited a considerable variation at DNA and protein levels suggesting an important future role for its genome in enrichment of genetic variation in cultivated rice.

Keywords: ***Oryza, in silico,*** **CDKB.**

INTRODUCTION

Rice (*Oryza sativa*) is the second most cultivated cereal crop in the world with annual production of 770 million ton yielded from 167 million ha (FAO, 2017). It is one of the major staple foods for about 70% of the world population with an excellent nutritional balance of carbohydrates, proteins and lipids (Balindong et al., 2018; Szareski et al., 2018). The increasing demands of rice production in addition to stressful environmental conditions associated with global climate change necessitate production of new varieties having better agronomic performance (Kilasi et al., 2018; Szareski et al., 2018).

Among different kinases, cyclin dependent protein kinases (CDKs) play important control role for cell division through forming complexes, with cyclin, that phosphorylate proteins required for progression of cell cycle. Based on cyclin-binding domains, CDKs are divided into eight classes (CDKA – CDKG in addition to cyclin dependent kinases like) (Tank and Thaker, 2011). CDKB is plant-specific class involved in several cellular functions (De Veylder et al., 2007). It is further divided into CDKB1 and CDKB2 having PPTALRE and PPTTLRE motifs, respectively (Joubes et al., 2000). Both types are recognized in dicots while monocot grasses are deprived of CDKB2 (Tank and Thaker, 2011). CDKB1 members control M phase particularly during development of stomata (Boudolf et al., 2004; Xie et al., 2010). Recently, they were proved to be involved in homologous recombination DNA repair that helps normal cell division under stressful conditions (Weimer et al., 2016).

Only two *Oryza* species namely *O. glaberrima* from Africa and *O. sativa* from Asia are cultivated but unfortunately having limited genetic diversity, estimated to be 10-20% of that in wild *Oryza* species (Zhu et al., 2007; Palmgren et al., 2014). The growing efforts are going to establish phenotypic and DNA sequence diversity in wild species promising a wide range of inter- and intraspecific variations (Atwell et al., 2014; Li et

al., 2014; Yan et al., 2016) that provides a natural reservoir for genetic information essential for breeding programs (Rao, 2004).

Starting with the 389-Mb genome of *O. sativa* ssp. japonica cv. Nipponbare completed in 2004 (International Rice Genome Sequencing Project, 2005), efforts supported with next generation sequencing (NGS) technologies established full genome sequences for several *Oryza* species. The resulted DNA sequences are available in several free sources and constitute an ore from which genetic information can be mined. *In silico* analysis have the advantage over wet lab-based techniques of being cost and time saving where it can be completed using open-source free data and software (Murray et al., 2007). This fastens recognition of potentially important genes required for genetic improvements of present rice cultivars.

In silico analysis was employed to identify important genes in *Oryza* species including salt stress responsive genes (Bhati et al., 2016), genes encoding shikimate pathway enzymes (Yaqoob et al., 2016), regulatory elements of pathogenesis-related proteins (Kaur et al., 2017) and nucleotide binding site-leucine-rich repeats (NBS-LRR) playing an important role in the plant defense systems (Rawal et al., 2018). Therefore, the aim of this work is to characterize CDKB1 genes and proteins in cultivated and some wild *Oryza* species using *in silico* analysis.

MATERIALS AND METHODS

The amino acid sequence of cdc2 kinase (BAA19553.1) in *Oryza sativa* Japonica Group was downloaded from NCBI (http://www.ncbi.nlm.nih.gov) and targeted in three cultivated (*O. sativa* Japonica Group, *O. sativa* Indica Group and *O. glaberrima*) and seven wild (*O. nivara, O. brachyantha, O. rufipogon, O. punctata, O. glumipatula, O. meridionalis* and *O. barthii*) *Oryza* species genomes in EnsemblPlants database (http://www. http://plants.ensembl.org) using BLASTP search tool to recognize candidate genes, coding sequences and location on chromosomes.

Based on the obtained genomic and coding sequences, exon-intron structure of the mined genes was constructed utilizing Gene Structure Display Server website (http://gsds.cbi.pku.edu.cn/).Along with the CDKB1 genes in closest monocot grasses (*Hordeum vulgare*, *Triticum aestivum* and *Zea mays*) available in Gene Bank, the mined *Oryza* sequences were aligned using Clustal W. Aligned sequences were employed to construct a phylogenetic tree using Maximum Likelihood (ML) method in MEGA v. 6 (Tamura et al., 2013) according to Kimura 2-parameter model (Kimura, 1980) with gamma distribution. Bootstrap of 1000 replicate was used to assess significance of support for grouping patterns (Felsenstein, 1985**)**. Retrieved amino acid sequences in different *Oryza* species were aligned using multiple sequence alignment (multalin) (Corpet, 1988, http://multalin.toulouse.inra.fr/multalin) to determine consensus domain and characteristic motif. Physico-chemical parameters of the candidate proteins including amino acid sequence, molecular weight and isoelectric points were determined using Expasy Protparm server (Gasteiger et al., 2005**,** http://us.expasy.org/tools/protparam.html). Subcellular localization was predicted using the CELLO2GO server (Yu et al., 2014). Secondary structure data were extracted using SOPMA (Self-Optimized Prediction Method with Alignment) online tool (Combet et al., 2000, https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_sopma.html). 3-D models for predicted proteins were constructed using the Phyre2 server (Kelley et al., 2015, http://www.sbg.bio. ic.ac.uk/phyre2) and Z-score was calculated using ProSA-web server and validated using (Wiederstein and Sippl, 2007, https://prosa.services.came.sbg.ac.at/prosa.php).

RESULTS AND DISCUSSION

Data mined reflected presence of CDKB1 gene on chromosome 8 in all studied species except *O. meridionalis* where it was located on chromosome 2 (Table 1). Exon-intron structure analysis showed similar general structure of 6 exons spaced with 5 intons for CDKB1 gene in all studied *Oryza* genomes (Table 1 and Figure 1). However, Imajuku et al., (1992) recorded CDKB1 gene of 9 exons in *Arabidopsis*. On the other hand, Magwanga et al., (2018) observed only 3 exons in cotton. Such contradictory observations may reflect species-dependent structure for CDKB1 genes.

Phylogenetic analysis showed clustering of all *Oryza* species in a major clade supported with a bootstrap value of 58% (Figure 2). Regarding relationships of wild and domesticated rices, the phylogenetic tree reflected that *O. nivara* and *O. rufipogon* are the closest species to *O. sativa* while *O. barthii* is the closest species to *O. glaberrima*.

Table 1. Chromosome distribution and exons position on CDKB1 gene in some *Oryza* species.

Species	Chromo some	Gene Size (bp)	Exon 1		Exon 2		Exon 3		Exon 4		Exon 5		Exon 6	
			Start	End	Start	End	Start	End	Start	End	Start	End	Start	End
O. nivara	8	2124	1	639	789	962	1075	1168	1258	1329	1436	1572	1735	2124
***O. sativa* Indica gp**	8	2124	229	639	789	962	1075	1168	1258	1329	1436	1572	1735	1827
***O. sativa* Japonica gp**	8	2124	179	639	789	962	1075	1168	1258	1329	1436	1572	1735	2017
O. rufipogon	8	2124	139	639	789	962	1075	1168	1258	1329	1436	1572	1735	2124
O. barthii	8	2125	179	639	790	963	1076	1169	1259	1330	1437	1573	1736	2125
O. glaberrima	8	2126	229	639	791	964	1077	1170	1260	1331	1438	1574	1737	1829
O. glumipatula	8	2124	155	639	789	962	1075	1168	1258	1329	1436	1572	1735	2051
O. punctate	8	2134	159	639	808	981	1077	1170	1268	1339	1446	1582	1745	2134
O. meridionalis	2	2125	111	639	789	962	1075	1168	1259	1330	1437	1573	1736	2125
O. brachyantha	8	2072	229	639	782	955	1045	1138	1226	1297	1401	1537	1683	2072

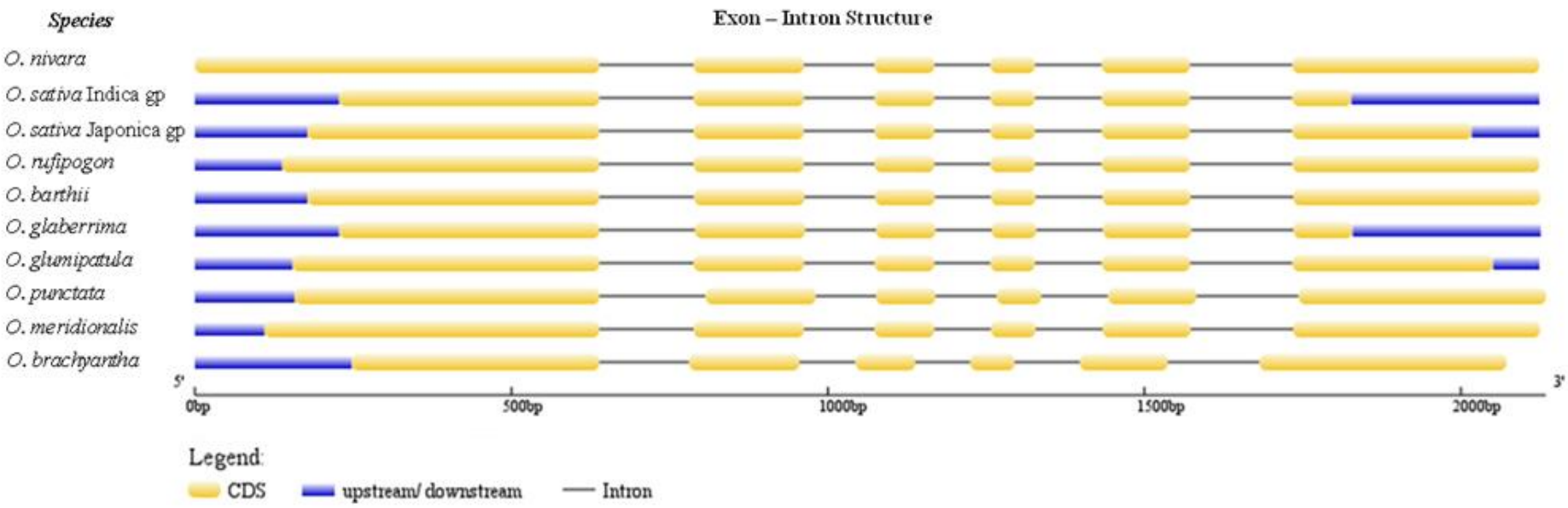

Figure 1. Exon-intron structure of CDKB1 gene in some *Oryza* species.

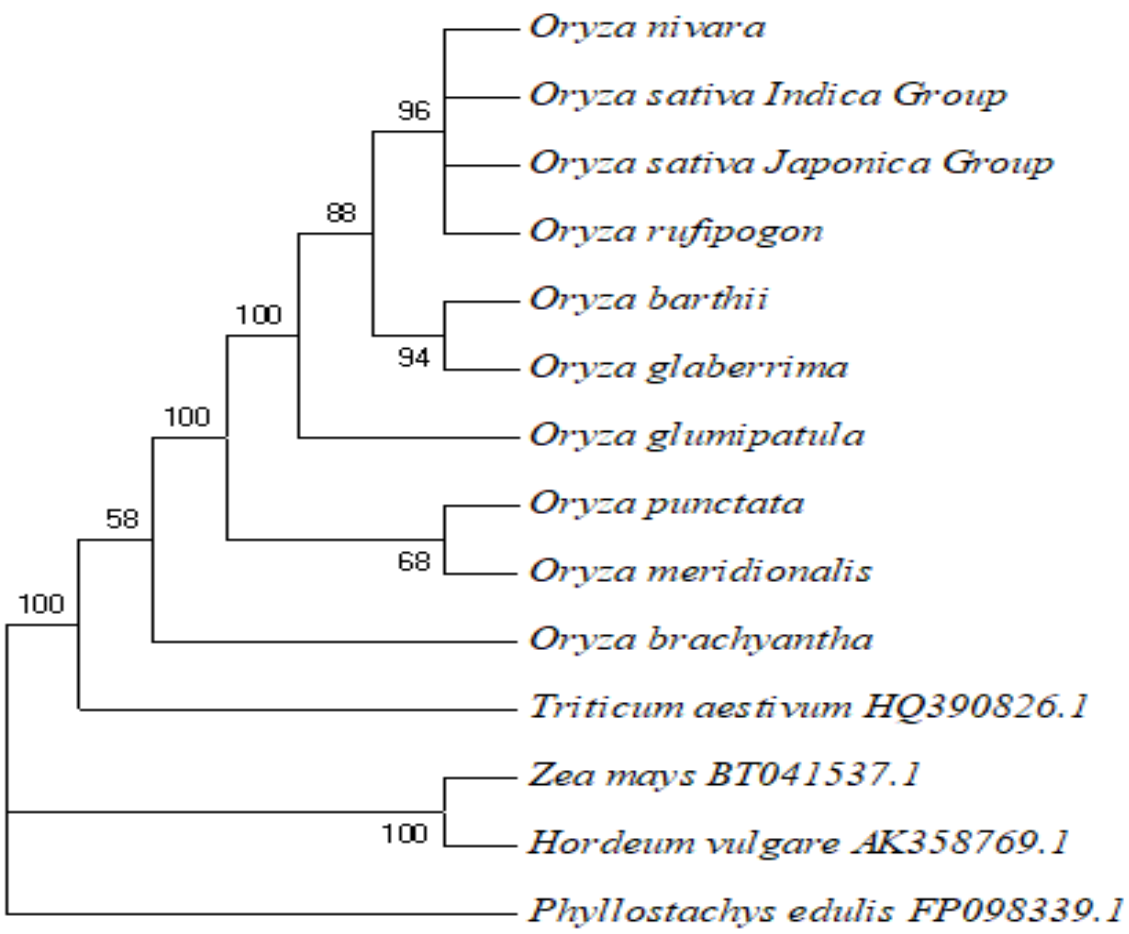

Figure 2. Phylogenetic tree of CDKB1 gene in some *Oryza* species and the closest monocot grasses (*Hordeum vulgare, Triticum aestivum* and *Zea mays*) using Maximum Likelihood method based on Kimura 2-parameter model. Based on 1000 replications, Bootstrap values (as percentages) are listed at branching points.

On the other hand *O. brachyantha* appeared as an outgroup for the remaining Oryza species.

Similar general taxonomic relations were recorded for *Oryza* species based on matk (Ge et al., 2002), *trn*L-*trn*F and ITS (Duan et al., 2007), whole chloroplast sequences (Wambugu et al., 2015), centromeres and centromere-linked genes (Liao et al., 2018) and GH3 genes (Kong et al., 2019).

Subcellular location analysis reflected that the predicted CDKB1 is a nuclear protein (Table 2). However, Boruc et al., (2010) recorded nuclear and cytoplasmic localization for CDKB1 in *Arabidopsis*. The same observations were recorded in tobacco BY2 cells by Porceddu et al., (2001). The authors explained cytoplasmic localization by nuclear envelope breakdown during cell division. Retrieved amino acid sequences showed 302 amino acid length in all *Oryza* species with PPTALRE motif (Figure 3) characteristic to CDKB1 (Mészáros et al., 2000). Not far from these results, CDKB1 proteins of 303 and 304 of amino acid length were recorded in tobacco (Sorrell et al., 2001) and Jerusalem artichoke (Freeman et al., 2003), respectively.

Physiochemical properties including molecular weight, isoelectric point (PI) and instability index for all mined CDKBs were predicted (Table 2). Secondary structure and 3-D models were constructed and validated (Table 3 and Figure 4). A narrow range of molecular weights (from 34.58 KDa in *O. brachyantha* to 34.63 kDa in *O. punctate*) and isoelectric points (8.87 in all species except *O. brachyantha* that showing PI of 9) were recorded for the predicted CDKB1s. A molecular weight of 36 KDa was predicted for CDKB1 in *Oryza sativa* (Sakaguchi et al., 2006). A similar molecular masse of 35 KDa was recorded in *Arabidopsis* (Boudolf et al., 2001) while a higher mass of 37 KDa was recorded in the green alga *Ostreococcus tauri* (Corellou et al., 2005).

Multalin-based alignment for amino acid sequences (Figure 3) showed identical sequences in *O. nivara*, *O. sativa* Indica group, *O. sativa* Japonica group, *O. rufipogon*, *O. barthii*, *O. glumipatula* and *O. glaberrima* associated with identical physiochemical properties. Species-dependent amino acids substitutions were recorded in the remaining three *Oryza* species that was more abundant in *O. brachyantha*. One substitution (threonine / asparagine at 207 position) in *O. meridionalis* and three substitutions (proline / serine, threonine / alanine and valine / leucine at 62, 165 and 187 positions, respectively) were characteristic for *O. punctate*. These substitutions were associated with variations in secondary structure parameters without corresponding alteration in PI or 3-D dimensions that may be attributed to similarity in PI of amino acids in each substitution (Table 3 and Figure 4). On the other hand, 14 amino acid substitutions were recorded in *O. brachyantha* including replacements of amino acids with others having different PI (histidine / leucine, histidine / isoleucine, arginine / glutamine and glutamine / histidine). Such replacements were accompanied with slight alteration in PI and more pronounced variations in secondary structure and 3-D dimensions.

The instability index is an estimate for the *in vitro* stability of the protein. A protein having instability index smaller than 40 is predicted to be stable (Guruprasad et al., 1990) that supports the models predicted in this study having indices of 28.82 (*O. brachyantha*) to 33.51 (*O. meridionalis*). As a measure of energy, Z-score reflected negative scores (- 6.88 in *O. brachyantha* to – 4.02 in *O. meridionalis*) indicating one of the ideal structures corresponding to the amino acid sequence (Moraes Filho et al., 2017).

CONCLUSION

In conclusion, *in silico* techniques provide valuable, fast and cost-effective information about the rapidly emerging genomes of wild relatives of strategic crops. Such information contribute effectively in breeding programs design and development of new cultivars of better performance. The used techniques revealed a very narrow range of variation at DNA and protein levels for CDKB1 in all studied *Oryza* species except *O. brachyantha.* The later exhibited a considerable variation at both levels suggesting an important future role for its genome in enrichment of genetic variation in cultivated rices.

CONFLICT OF INTEREST

The authors declared that present study was performed in absence of any conflict of interest.

AUTHOR CONTRIBUTIONS

Both authors contributed equally in finding the idea, *in silico* analyses and manuscript preparation.

Table 2. Subcellular localization and physiochemical properties of CDKB1 in some *Oryza* species.

Species	Subcellular location	Molecular weight	Formula:	No. of Amino acids	No. of negatively charged residues	No. of positively charged residues	PI	Instability index
O. nivara	**Nucleus**	34604	$C_{1571}H_{2488}N_{420}O_{437}S_{11}$	302	36	41	8.87	33.01
***O. sativa* Indica gp**	**Nucleus**	34604	$C_{1571}H_{2488}N_{420}O_{437}S_{11}$	302	36	41	8.87	33.01
***O. sativa* Japonica gp**	**Nucleus**	34604	$C_{1571}H_{2488}N_{420}O_{437}S_{11}$	302	36	41	8.87	33.01
O. rufipogon	**Nucleus**	34604	$C_{1571}H_{2488}N_{420}O_{437}S_{11}$	302	36	41	8.87	33.01
O. barthii	**Nucleus**	34604	$C_{1571}H_{2488}N_{420}O_{437}S_{11}$	302	36	41	8.87	33.01
O. glumipatula	**Nucleus**	34604	$C_{1571}H_{2488}N_{420}O_{437}S_{11}$	302	36	41	8.87	33.01
O. glaberrima	**Nucleus**	34604	$C_{1571}H_{2488}N_{420}O_{437}S_{11}$	302	36	41	8.87	33.01
O. meridionalis	**Nucleus**	34591	$C_{1571}H_{2489}N_{419}O_{437}S_{11}$	302	36	41	8.87	33.51
O. punctate	**Nucleus**	34630	$C_{1573}H_{2490}N_{420}O_{437}S_{11}$	302	36	41	8.87	33.48
O. brachyantha	**Nucleus**	34580	$C_{1564}H_{2476}N_{424}O_{436}S_{12}$	302	36	42	9.00	28.82

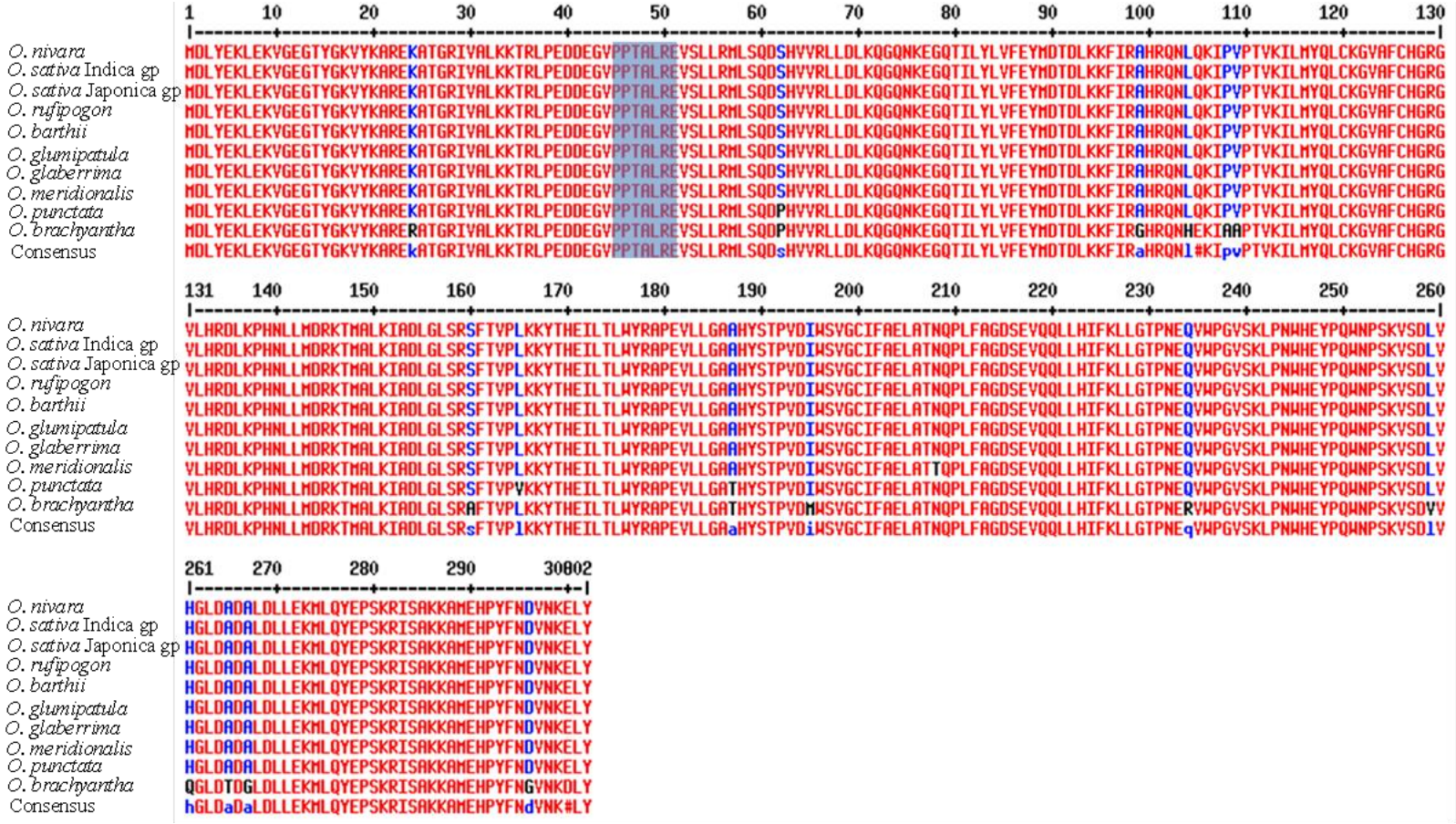

Figure 3. Multalin-based amino acid sequence alignment of CDKB1 in some *Oryza* species showing PPTALRE motif.

Table 3. Details of secondary structures and 3-D model of CDKB in some *Oryza* species.

Species	Secondary structure			3-D Model			
	Alpha helix (%)	Extended Strand (%)	Beta turn (%)	dimensions (Å)			Z-Score
				X	Y	Z	
O. nivara	43.38	12.58	5.63	46.8	59.4	61.2	- 4.08
***O. sativa* Indica gp**	43.38	12.58	5.63	46.8	59.4	61.2	- 4.08
***O. sativa* Japonica gp**	43.38	12.58	5.63	46.8	59.4	61.2	- 4.08
O. rufipogon	43.38	12.58	5.63	46.8	59.4	61.2	- 4.08
O. barthii	43.38	12.58	5.63	46.8	59.4	61.2	- 4.08
O. glumipatula	43.38	12.58	5.63	46.8	59.4	61.2	- 4.08
O. glaberrima	43.38	12.58	5.63	46.8	59.4	61.2	- 4.08
O. meridionalis	42.38	14.57	5.96	46.8	59.4	61.2	- 4.02
O. punctate	41.06	12.91	6.62	46.8	59.4	61.2	- 4.19
O. brachyantha	43.05	13.58	6.29	64.4	56.2	50.0	- 6.88

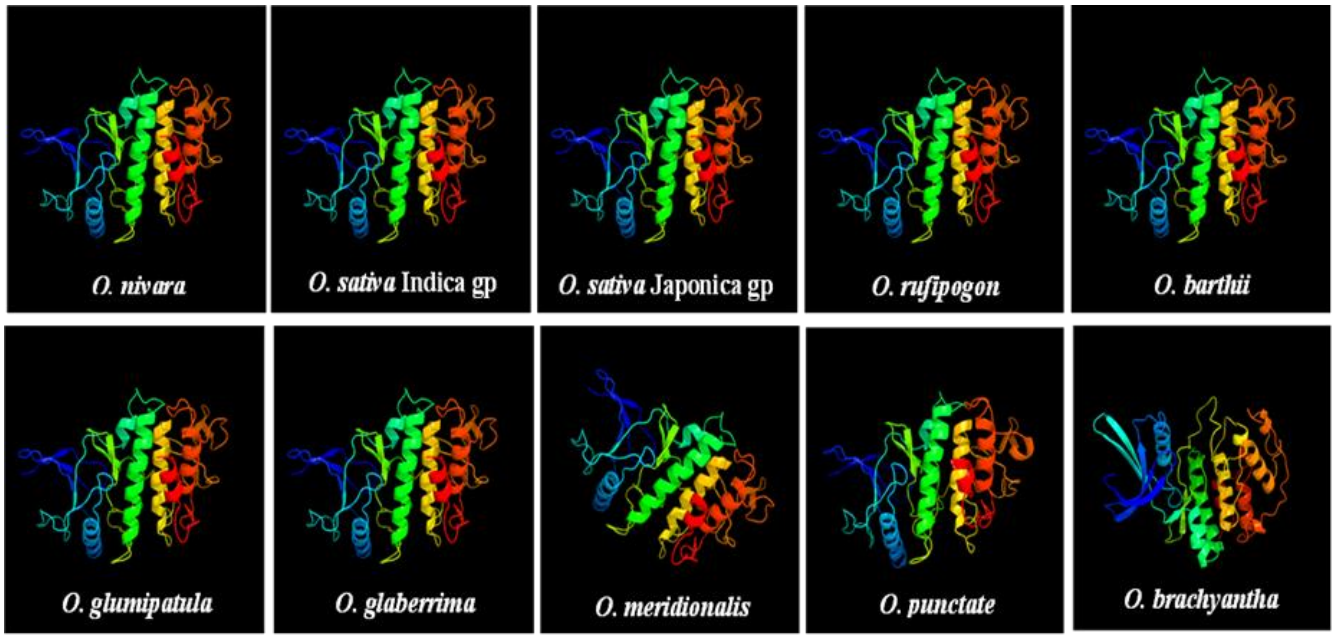

Figure 4. Predicted 3-D models for CDKB1 in some *Oryza* species.

REFERENCES

Atwell BJ, Wang H, Scafaro AP, 2014. Could abiotic stress tolerance in wild relatives of rice be used to improve *Oryza sativa*?. Plant Sci 215: 48-58.

Balindong JL, Ward RM, Liu L, Rose TJ, Pallas LA, Ovenden BW, Snell PJ, Waters DL, 2018. Rice grain protein composition influences instrumental measures of rice cooking and eating quality. J Cereal Sci 1(79): 35-42.

Bhati J, Chaduvula KP, Rai A, Gaikwad K, Soma Marla S, 2016. *In-silico* prediction and functional analysis of salt stress responsive genes in rice (*Oryza sativa*). J Rice Res 4(164): 2.

Boruc J, Mylle E, Duda M, De Clercq R, Rombauts S, Geelen D, Hilson P, Inzé D, Van Damme D, Russinova E, 2010. Systematic localization of the *Arabidopsis* core cell cycle proteins reveals novel cell division complexes. Plant physiol 152(2): 553-565.

Boudolf V, Barrôco R, de Almeida Engler J, Verkest A, Beeckman T, Naudts M, Inzé D, De Veylder L, 2004. B1-type cyclin-dependent kinases are essential for the formation of stomatal complexes in *Arabidopsis thaliana*. Plant Cell 16(4): 945-955.

Boudolf V, Rombauts S, Naudts M, Inzé D, De Veylder L, 2001. Identification of novel cyclin-dependent kinases interacting with the CKS1 protein of *Arabidopsis*. J Exp Bot 52(359):1381-1382.

Combet C, Blanchet C, Geourjon C, Deleage G, 2000. NPS@: network protein sequence analysis. Trends Biochem Sci 25(3): 147-150.

Corellou F, Camasses A, Ligat L, Peaucellier G, Bouget FY, 2005. A typical regulation of a green lineage-specific B-type cyclin-dependent kinase. Plant Physiol 138(3):1627-1636.

Corpet F, 1988. Multiple sequence alignment with hierarchical clustering. Nucleic Acids Res 16(22):10881-10890.

De Veylder L, Beeckman T, Inzé D, 2007. The ins and outs of the plant cell cycle. Nat Rev Mol Cell Bio 8(8): 655-665.

Duan S, Lu B, Li Z, Tong J, Kong J, Yao W, Li S, Zhu Y, 2007. Phylogenetic analysis of AA-genome *Oryza* species (Poaceae) based on chloroplast, mitochondrial, and nuclear DNA sequences. Biochem Genet 45(1-2): 113-129.

FAO, 2017. http://www.fao.org/faostat/en/#data/QC

Felsenstein J, 1985. Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39: 783-791.

Freeman D, Riou-Khamlichi C, Oakenfull EA, Murray JA, 2003. Isolation, characterization and expression of cyclin and cyclin-dependent kinase genes in Jerusalem artichoke (*Helianthus tuberosus* L.). J Exp Bot 54(381): 303-308.

Ge S, Li A, Lu BR, Zhang SZ, Hong DY, 2002. A phylogeny of the rice tribe Oryzeae (Poaceae) based on matK sequence data. Am J Bot 89(12): 1967-1972.

Guruprasad K, Reddy BB, Pandit MW, 1990. Correlation between stability of a protein and its dipeptide composition: a novel approach for predicting *in vivo* stability of a protein from its primary sequence. Protein Eng Des Sel 4(2): 155 - 161.

Imajuku Y, Hirayama T, Endoh H, Oka A, 1992. Exon—intron organization of the *Arabidopsis thaliana* protein kinase genes CDC2a and CDC2b. FEBS Lett 304(1): 73-77.

International Rice Genome Sequencing Project, 2005. The map-based sequence of the rice genome. Nature 436(7052): 793–800.

Jain E, Bairoch A, Duvaud S, Phan I, Redaschi N, Suzek BE, Martin MJ, McGarvey P, Gasteiger E, 2009. Infrastructure for the life sciences: design and implementation of the UniProt website. BMC Bioinformatics 10(1): 136.

Kaur A, Pati PK, Pati AM, Nagpal AK, 2017. *In-silico* analysis of cis-acting regulatory elements of pathogenesis-related proteins of *Arabidopsis thaliana* and *Oryza sativa*. PloS One. 12(9): e0184523.

Kelley LA, Mezulis S, Yates CM, Wass MN,

Sternberg MJ, 2015. The Phyre2 web portal for protein modeling, prediction and analysis. Nat Protoc 10(6): 845.

Kilasi NL, Singh J, Vallejos CE, Ye C, Jagadish SK, Kusolwa P, Rathinasabapathi B, 2018. Heat stress tolerance in rice (*Oryza sativa* L.): Identification of quantitative trait loci and candidate genes for seedling growth under heat stress. Front Plant Sci 9: 1578.

Kimura M, 1980. A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. J Mol Evol 16: 111-120.

Kong W, Zhong H, Deng X, Gautam M, Gong Z, Zhang Y, Zhao G, Liu C, Li Y, 2019. Evolutionary analysis of GH3 genes in six *Oryza* species/subspecies and their expression under salinity stress in *Oryza sativa* ssp. japonica. Plants 8(2): 30.

Li JY, Liu J, Dong D, Jia X, McCouch SR, Kochian LV, 2014. Natural variation underlies alterations in Nramp aluminum transporter (NRAT1) expression and function that play a key role in rice aluminum tolerance. P Natl Acad Sci 111(17): 6503-6508.

Liao Y, Zhang X, Li B, Liu T, Chen J, Bai Z, Wang M, Shi J, Walling JG, Wing RA, Jiang J, 2018. Comparison of *Oryza sativa* and *Oryza brachyantha* genomes reveals selection-driven gene escape from the centromeric regions. Plant Cell 30(8):1729-1744.

Magwanga R, Lu P, Kirungu J, Cai X, Zhou Z, Wang X, Diouf L, Xu Y, Hou Y, Hu Y, Dong Q, 2018. Whole genome analysis of cyclin dependent kinase (CDK) gene family in cotton and functional evaluation of the role of CDKF4 Gene in drought and salt stress tolerance in plants. Int J Mol Sci 19(9): 2625.

Mészáros T, Miskolczi P, Ayaydin F, Pettkó-Szandtner A, Peres A, Magyar Z, Horváth GV, Bakó L, Fehér A, Dudits D, 2000. Multiple cyclin-dependent kinase complexes and phosphatases control G 2/M progression in alfalfa cells. Plant Mol Biol 43: 595-605.

Moraes Filho RM, Menezes AF, Martins LS, 2017. *In silico* modeling and characterization of phytoparasitic nematodes translationally-controlled tumor proteins. Genet Mol Res 16(3).

Murray D, Doran P, MacMathuna P, Moss AC, 2007. *In silico* gene expression analysis–an overview. Mol Cancer 6(1): 50.

Palmgren MG, Edenbrandt AK, Vedel SE, Andersen MM, Landes X, Østerberg JT, Falhof J, Olsen LI, Christensen SB, Sandøe P, Gamborg C, 2015. Are we ready for back-to-nature crop breeding?. Trends in Plant Sci 20(3):155-164.

Porceddu A, Stals H, Reichheld JP, Segers G, De Veylder L, de Pinho Barrôco R, Casteels P, Van Montagu M, Inzé D, Mironov V, 2001. A plant-specific cyclin-dependent kinase is involved in the control of G2/M progression in plants. Journal Biol Chem 276(39): 36354-36360.

Rao NK, 2004. Plant genetic resources: Advancing conservation and use through biotechnology. Afr J Biotechnol 3(2): 136-145.

Rawal HC, Mithra SA, Arora K, Kumar V, Goel N, Mishra DC, Chaturvedi KK, Rai A, Devi SV, Sharma TR, Solanke AU, 2018. Genome-wide analysis in wild and cultivated *Oryza* species reveals abundance of NBS genes in progenitors of cultivated rice. Plant Mol Biol Rep 36(3): 373-386.

Sakaguchi N, Furukawa T, Shimada H, Hashimoto J, Sakaguchi K, Umeda M, 2006. Isolation and characterization of a rice cDNA encoding B1-type cyclin-dependent kinase. Plant Biotechnol 23(2): 211-214.

Sorrell DA, Menges M, Healy JS, Deveaux Y, Amano C, Su Y, Nakagami H, Shinmyo A, Doonan JH, Sekine M, Murray JA, 2001. Cell cycle regulation of cyclin-dependent kinases in tobacco cultivar Bright Yellow-2 cells. Plant Physiol 126(3): 1214-23.

Szareski VJ, Carvalho IR, da Rosa TC, Dellagostin SM, de Pelegrin AJ, Barbosa MH, dos Santos OP, Muraro DS, de Souza VQ, Pedó T, Aumonde TZ, 2018. *Oryza* wild species: an alternative for rice breeding under abiotic stress conditions. Am J Plant Sci 9(6): 1093.

Tamura K, Stecher G, Peterson D, Filipski A, Kumar S, 2013. MEGA6: Molecular evolutionary genetics analysis version 6.0. Mol Biol Evol 30: 2725-2729.

Tank JG, Thaker VS, 2011. Cyclin dependent kinases and their role in regulation of plant cell cycle. Biol Plantarum 55(2): 201-212.

Wambugu PW, Brozynska M, Furtado A, Waters DL, Henry RJ, 2015. Relationships of wild and domesticated rices (*Oryza* AA genome species) based upon whole chloroplast genome sequences. Sci Rep 5: 13957.

Weimer AK, Biedermann S, Harashima H, Roodbarkelari F, Takahashi N, Foreman J, Guan Y, Pochon G, Heese M, Van Damme D, Sugimoto K, 2016. The plant-specific

CDKB1-CYCB1 complex mediates homologous recombination repair in *Arabidopsis*. EMBO J 35(19): 2068-2086.

Wiederstein M, Sippl MJ, 2007. ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. Nucleic Acids Res 35(suppl_2): 407-410.

Xie Z, Lee E, Lucas JR, Morohashi K, Li D, Murray JA, Sack FD, Grotewold E, 2010. Regulation of cell proliferation in the stomatal lineage by the *Arabidopsis* MYB FOUR LIPS via direct targeting of core cell cycle genes. Plant Cell 22(7): 2306-2321.

Yan J, Wang P, Wang P, Yang M, Lian X, Tang Z, Huang CF, Salt DE, Zhao FJ, 2016. A loss-of-function allele of OsHMA3 associated with high cadmium accumulation in shoots and grain of Japonica rice cultivars. Plant Cell Environ 39(9): 1941-1954.

Yaqoob U, Kaul T, Pandey S, Nawchoo IA, 2016. *In-silico* characterization, structural modelling, docking studies and phylogenetic analysis of 5-enolpyruvylshikimate-3-phosphate synthase gene of *Oryza sativa* L. Med Aromat Plants (Los Angel) 5(274): 2167-0412.

Yu CS, Cheng CW, Su WC, Chang KC, Huang SW, Hwang JK, Lu CH, 2014. CELLO2GO: a web server for protein subCELlular LOcalization prediction with functional gene ontology annotation. Plos One 9(6): e99368.

Zhu Q, Zheng X, Luo J, Gaut BS, Ge S, 2007. Multilocus analysis of nucleotide variation of *Oryza sativa* and its wild relatives: severe bottle neck during domestication of rice. Mol Biol Evol 24: 875 - 888.