Supplementary Material

In silico characterisation and functional validation of chilling tolerant divergence 1 (*COLD1*) gene in monocots during abiotic stress

P. Anunathini^A, V. M. Manoj^B, T. S. Sarath Padmanabhan^B, S. Dhivya^A, J. Ashwin Narayan^B, C. Appunu^{B,C} and R. Sathishkumar^{A,C}

^APlant Genetic Engineering Laboratory, Department of Biotechnology, Bharathiar University, Coimbatore, India.

^BGenetic Transformation Lab, Division of Crop Improvement, ICAR-Sugarcane Breeding Institute, Coimbatore, India.

^cCorresponding authors. Emails: rsathish@buc.edu.in; cappunu@gmail.com

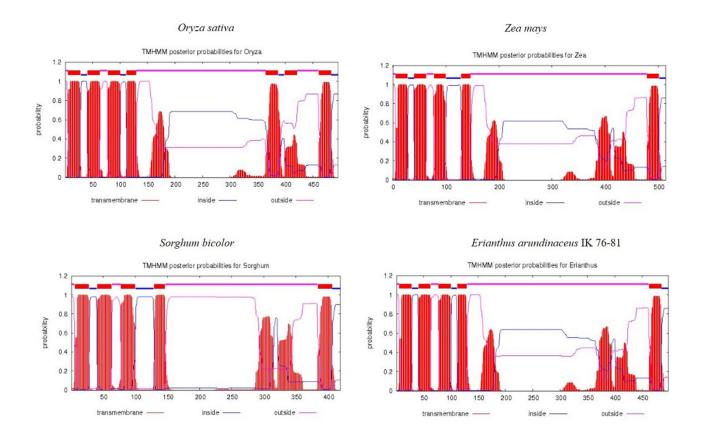


Fig. S1. TMHMM topology for the studied monocot crops.

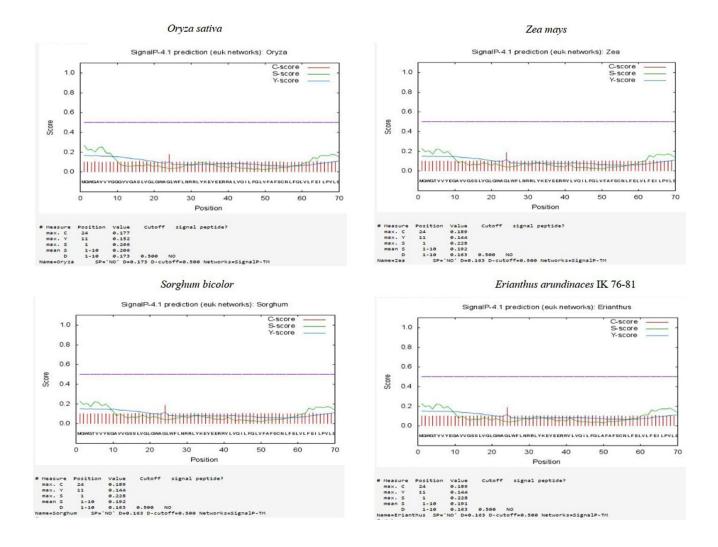


Fig. S2. Signal peptide analysis for the studied monocot crops using SignalP4.1 server. Outputs are in three different forms i.e. S, D and Y score. S-score, the average mean of the possible signal peptide; D-score (discrimination score) describes the weighted average of the mean S and the max; Y scores used to discriminate signal peptides from non-signal peptides.

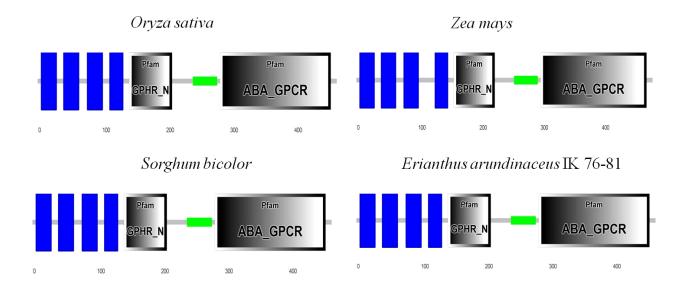


Fig. S3. Domain prediction using Pfam and SMART for the studied monocot crops.

Oryza	MGWGAVVYGGGVVGASLVGLGWAGLWFLNRRLYKEYEERRALVQILFGLVFAFSCNLFQL
zea	MGWGTVVYEGAVVGSSLVGLGWAGLWFLNRRLYKEYEERRVLVQILFGLVFAFSCNLFEL
Sorghum	MGWGTVVYEGAVVGSSLVGLGWAGLWFLNRRLYKEYEERRVLVQILFGLVFAFSCNLFEL
Erianthus	MGWGTVVYEGAVVGSSLVGLGWAGLWFLNRRLYKEYEERRVLVQILFGLVFAFSCNLFEL
	**** *** * *** ************************
Oryza	VLFEILPVLSKHARFLNWHLDLFCLILLLVFVLPYYHCYLLLRNSGVRRERALLVAALFL
zea	VLFEILPVLSKHARFLNWHLDLFCLILLLVFVLPYYHCYLLLRNSGMRRDRAWLVAALFL
Sorghum	VLFEILPVLSKHARFLNWHLDLFCLILLLVFVLPYYHCYLLLRNSGVRRERAWLVAALFL
Erianthus	VLFEILPVLSKHAHFLNWHLDLFCLILLLVFVLPYYHCYLLLRNSGVRRERAWLVAALFL

Oryza	LVFLYGFWRMGIHFPMPSPEKGFFTMPQLVSRIGVIGVSVMAVLSGFGAVNLPYSYLSLF
zea	LVFLYGFWRMGIHFPMPSPEKGFFTMPQLVSRIGVIGVSVMAVLSGFGAVNLPYSYLSLF
Sorghum	LVFLYGFWRMGIHFPMPSPEKGFFTMPQLVSRIGVIGVSVMAVLSGFGAVNLPYSYLSLF
Erianthus	LVFLYGFWRMGIHFPMPSPEKGFFTMPQLVSRIGMIGVSVMAVLSGFGAVNLPYSYLSLF

Oryza	IREIDEKDIKTLERQLMQSMETCIAKKKKIVLSKMEMERIQGSEEKLKARSFLKRIVGTV
zea	IREIDETDIKTLERQLMQSIETCTAKKKKIILSQMEMERIQGSEEKLKARSFLKRIVGTV
Sorghum	IREIDEADIKTLERQLMQSMETCTSKKKKIILSQMEMERIQGSEEKLKARSFLKRIVGTV
Erianthus	IREIDETDIKTLEROLMOCMETCTSKKKKIILSOMEMERIOGSEEKLKARSFLKRIVGTV
	***** ******** *** **** ** ************
Oryza	VRSVQEDQTEQDIKSLDAEVQALEELSKQLFLEIYELRQAKIAAAFSRTWRGHAQNLLGY
zea	VRSVQEDQTEQDIKNLEAEVQALEELSKQLFLEIYELRQAKIAAAYSRTWRGHLQNLLGY
Sorghum	VRSVQEDQTEQDIKNLEAEVQALEELSKQLFLEIYELRQAKIAAAYSRTWRGHLQNLLGY
Erianthus	VRSVQEDQTEQDIKNLEAEVQALEELSKQLFLEICELRQAKIAAAYSRTWRGHLQNLLGY

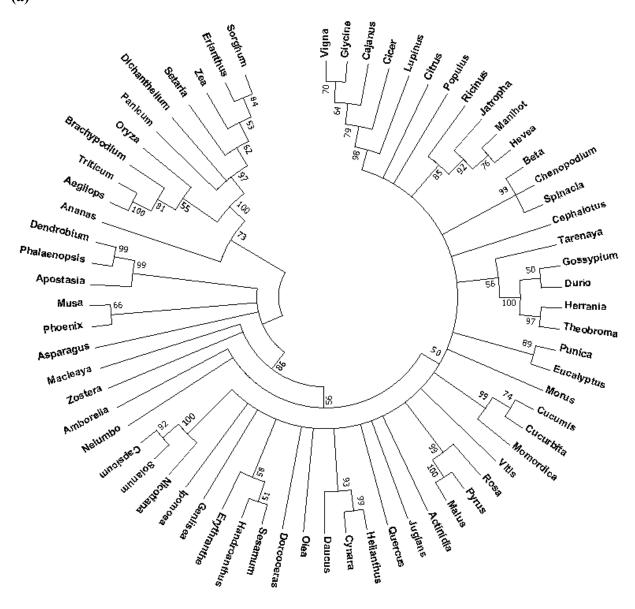
Oryza	ALSVYCVYKMLKSLQSVVFKEAGSVDPVTMTITIFLRHFDIGIDVTLLSQYISLIFIGML
zea	ALSVYCVYKMLKSLQSVVFKESGSVDPVTMSITIFLRHFDIGIDVALLSQYISLMFIGML
Sorghum	ALSVYCVYKMLKSLQSVVFKESGSVDPVTMTITIFLRHFDIGIDVALLSQYISLMFIGML
Erianthus	ALSVYCVYKMLKSLQSVVFKESGSVDPVTMTITIFLRHFDIGIDVALLSQYISLMFIGML

Oryza	VVISVRGFLANVMKFFFAVSRVGSGSTTNVVLFLSEIMGMYFISSILLIRKSLANEYRV
zea	VVISVRGFLANVMKFFFAVSRVGSGSTTNVVLFLSEIMGMYFISSILLIRKSLANEYRV
Sorghum	VVISVRGFLANVMKFFFAVSRVGSGSTTNVVLFLSEIMGMYFISSILLIRKSLANEYRV
Erianthus	VVISVRGFLANVMKFFFAVSRVGSGSTTNVVLFLSEIMGMYFISSILLIRKSLANEYRV

Oryza	ITDVLGGDIQFDFYHRWFDAIFVASAFLSLLLISAQYTSRQTDKHPID*
zea	ITDVLGGDIQFDFYHRWFDAIFVASAFLSLLLISAQYTTRQTDKHPID-
Sorghum	ITDVLGGDIQFDFYHRWFDAIFVASAFLSLLLISAQYTTRQTDKHPID-
Erianthus	ITDVLGGDIQFDFYHRWFDAIFVASAFLSLLLISAQYTTRQTDKHPID-

Fig. S4. Multiple sequence alignment of studied COLD1 proteins using ClustalO for the studied monocot crop species.

- * Represent the similarity : Represent the conserved substitutions
- . Represent the semi-conserved substitutions
- Gap represents no match.



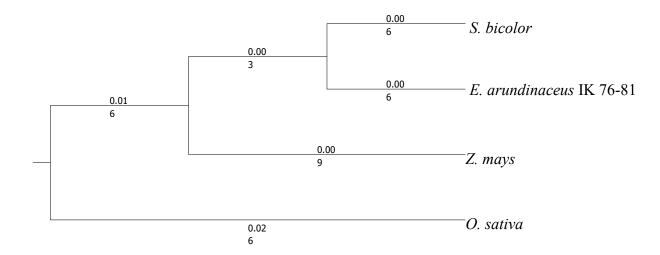


Fig. S5. (a) Phylogenetic analysis of COLD1 proteins using BLAST sequence (67 sequences); (b) Phylogenetic analysis of COLD1 proteins using MEGA7 software for the studied monocot crop species.

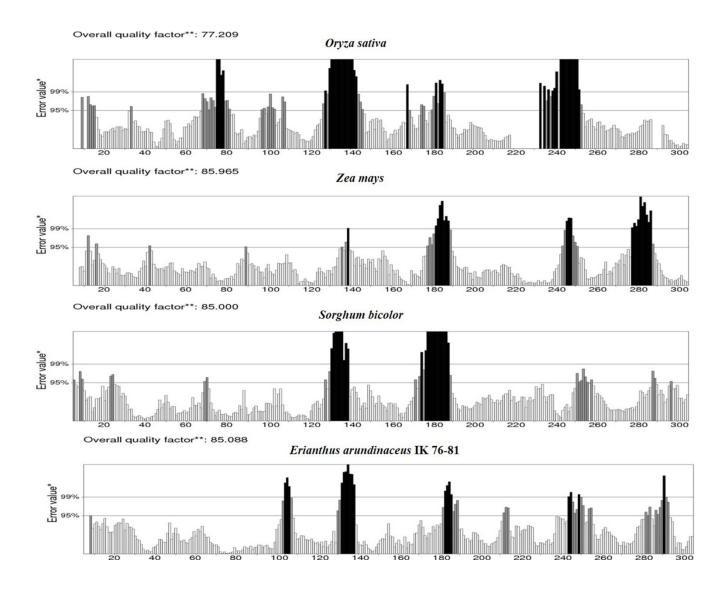


Fig. S6. ERRAT analysis of studied for the studied monocot crop species. Error value is represented on X axis, two lines drawn to indicate the confidence and possibility to reject regions that exceed error value. Overall quality factor is expressed as the percentage of the protein calculated. 95% or higher provide good high-resolution structures. For lower resolutions (2.5 to 3A) the average overall quality factor is around 91%.

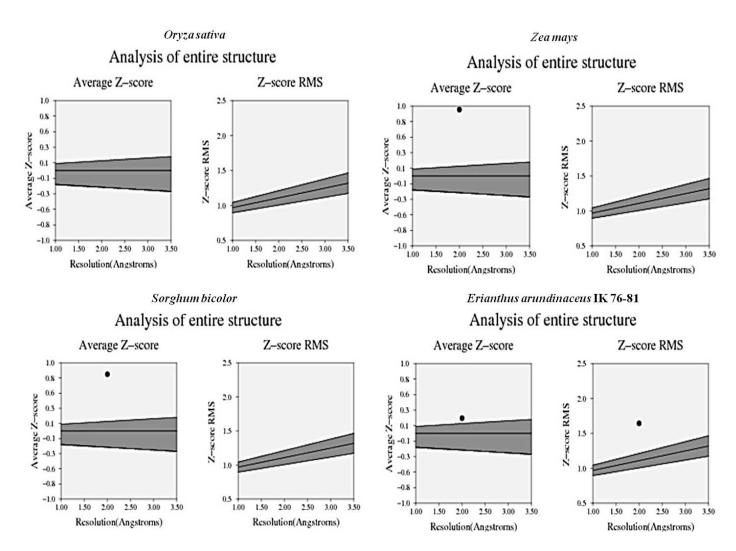


Fig. S7. PROVE analysis of studied COLD1 proteins. Output shows the level of Z score. 2.0 Å or better is highly resolved structures for the studied monocot crop species.

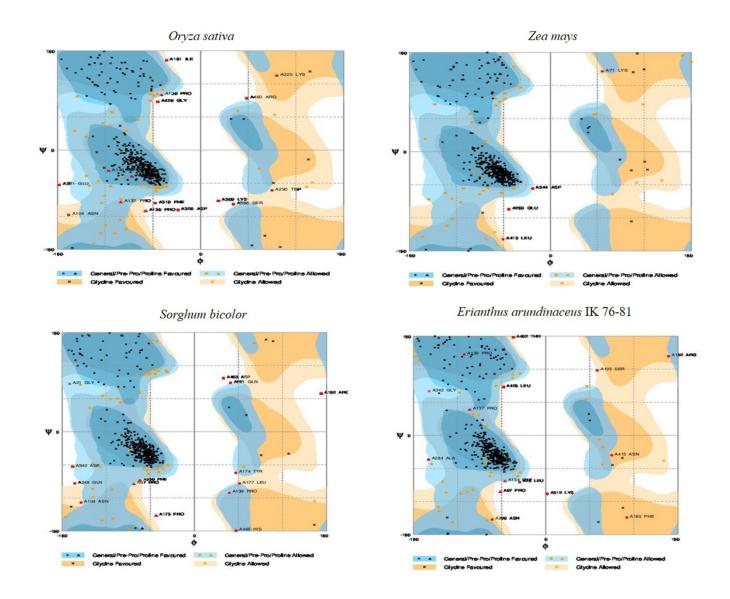


Fig. S8. Ramachandran plot using RAMPAGE for the studied COLD1 proteins.

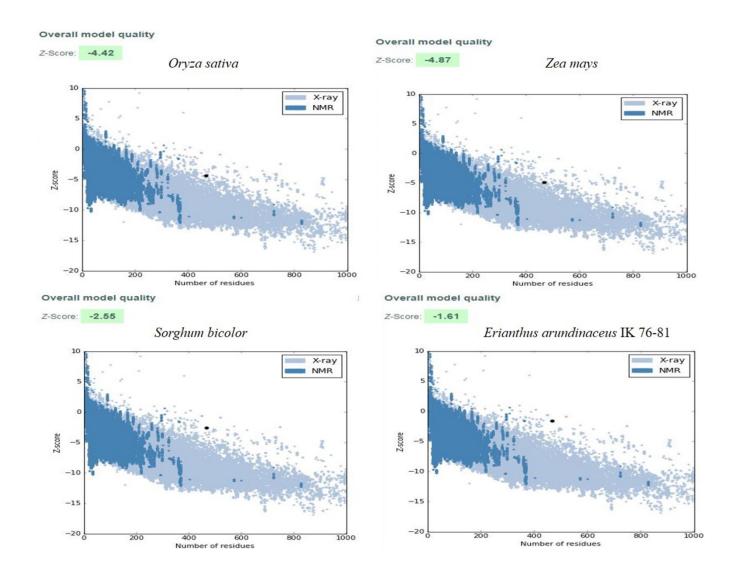


Fig. S9. ProSA analysis for the studied monocot crop species. Output display Z-scores, which indicate the overall model quality represented by large black dot. All protein chains in PDB are determined from different sources by X-ray crystallography and NMR spectroscopy with respect to their length.

Table S1. qRT-PCR primers designed for the studied monocot crop species using IDT software

S.NO	PLANT	Primers	Internal gene reference	
1	Oryza sativa	Cold1 Fw - TGGCGTGATTGGAGTAAGTG	Jain et al (2006)	
		Cold1 Rv - CATGAGCTGCCTTTCCAATG	Sakuraba et al (2017)	
		UBQ5 Fw - ACCACTTCGACCGCCACTACT	(2017)	
		UBQ5 Rv - ACGCCTAAGCCTGCTGGTT		
2	Zea mays	Cold1 Fw - TTCACTTCCCCATGCCTTC	Galli et al (2013)	
		Cold1 Rv - CTCCAATCACCCCAATCCTAC	Lin et al (2014)	
		Tubulin Fw - AGAACTGCGACTGCCTCCAAAG		
		Tubulin Rv - AGATGAGCAGGGTGCCCATTC		
3	Sorghum bicolour	Cold1 Fw - CCGCTGGTTTGATGCTATATTTG	Reddy et al (2016)	
		Cold1 Rv - GATGCTTGTCTGTTTGCCTG		
		PP2A Fw - AACCCGCAAAACCCCAGACTA		
		PP2A Rv - TACAGGTCGGGCTCATGGAAC		
4	Erianthus arundinaceus	Cold1 Fw - CGCAGTTGGTCAGTAGGATT	Iskandar et al 2012	
	IK 76-81	Cold1 Rv - CCTGATGAAGAGCGACAGATAA	Guo et al 2014	
		GAPDH Fw - CACGGCCACTGGAAGCA	Ling et al 2014	
		GAPDH Rv - TCCTCAGGGTTCCTGATGCC	Silva et al 2014	
			Andrade et al 2017	

Table S2. Similarity percentage of *Erianthus arundinaceus* IK 76-81 protein sequence with *Oryza sativa, Zea mays* and *Sorghum bicolor* obtained using Emboss needle tool

Erianthus	Oryza sativa (%)		Zea mays (%)		Sorghum bicolour (%)	
arundinaceus IK 76-81	Similarity	Identity	Similarity	Identity	Similarity	Identity
	97.6	95.3	96.3	95.2	76.1	76.1