BioTechnologia vol. 101 (4) • pp. 309–322 • 2020
Journal of Biotechnology, Computational Biology and Bionanotechnology RESEARCH PAPERS
http://doi.org/10.5114/bta.2020.100423

Comparative transcriptomic analyses of four *Phalaenopsis* species to identify and characterize the *WUSCHEL*-related homeobox (*WOX*) gene family

MADHVI KANCHAN, JASPREET K. SEMBI *
Department of Botany, Panjab University, Chandigarh, India

Abstract

In the present study, we identified and characterized the plant-specific *WUSCHEL*-related homeobox (*WOX*) gene family that plays a major role in the determination of cell fate, early embryonic pattern formation, stem cell maintenance, organogenesis, flower development and somatic embryogenesis. For this purpose, the transcriptomes of four species of *Phalaenopsis*, namely *P. aphrodite*, *P. modesta*, *P. bellina and P. lueddemanniana*, were analysed, and nine *PaWOX*, ten *PmWOX*, eight *PbWOX* and nine *PlWOX* transcripts were identified. The duplication event analysis showed the presence of two duplication events in *P. lueddemanniana*, one each in *P. bellina* and *P. modesta* and no duplication event in *P. aphrodite*. During the evolutionary analysis, all the WOX proteins were clustered with those of *Arabidopsis thaliana* (AtWOX), *Phalaenopsis equestris* (PeWOX), *Apostasia shenzhenica* and *Dendrobium catenatum* (DcWOX). The expression analysis of the *WOX* genes suggested their critical role in floral development and in other developmental processes. The secondary and tertiary structural analysis of seven selected WOX proteins was then performed, with each protein representing its respective clade. The results provide a valuable resource for further studies of the molecular mechanisms of floral and vegetative developments in *Phalaenopsis* species.

Key words: *WUSCHEL*, WOX, expression analysis, *Phalaenopsis*, orchids

Introduction

The *WUSCHEL*-related homeobox (*WOX*) gene family is a subgroup of the homeobox transcription factor superfamily containing the conserved domain of 60–66 amino acid helix-loop-helix-turn-helix structure (Gehring et al., 1994). It controls cell fate determination, early embryonic pattern formation, stem cell maintenance, organogenesis, flower development, somatic embryogenesis and stress tolerance in plants (Deveaux et al., 2008; Costanzo et al., 2014; Jha et al., 2020). The *WOX* gene family was first identified in *Arabidopsis thaliana*, where the *WUSCHEL* (*WUS*) gene of modern clade was isolated and functionally characterized (Laux et al., 1996; Haecker et al., 2004). In *Arabidopsis*, the *WOX* gene family consists of 15 members and is divided into three distinct clades: WUS (*WUS* and *WOX 1–7*), intermediate (*WOX8*, *WOX9*, *WOX11* and *WOX12*) and ancient (*WOX10*, *WOX13* and *WOX14*) (van der Graaff et al. 2009). The WUS clade is a modern clade restricted to angiosperms, the intermediate clade belongs to the intermediate plant groups such as pteridophytes and gymnosperms, and the ancient clade is present among all plant groups from green algae to angiosperms (van der Graaff et al., 2009; Lian et al., 2014). The *WUS* gene plays a role in the maintenance of shoot apical meristem in *Arabidopsis* (Zuo et al., 2002), somatic embryogenesis in *Gossypium hirsutum* (Bouchabké-Coussa et al., 2013) and organogenesis in *Coffea canephora* (Arroyo-Herrera et al., 2008). *WUS* also regulates floral development by acting as an activator of the *AGAMOUS* gene in flowers (Lohmann et al., 2001). The other members of the WUS clade contribute to different stages of plant development; *AtWOX5* performs the same function as *AtWUS* in root and shoot apical meristems (Oshchep-

* Corresponding author: Department of Botany, Panjab University, Chandigarh, India; e-mail: jaspreet.sembi@pu.ac.in

kova et al., 2017). The *WOX3* (*PRESSED FLOWER* or *PRS*) gene of *A. thaliana* regulates the development of lateral sepals, stamens and stipules (Matsumoto and Okada, 2001), whereas its ortholog in *Zea mays*, *NARROW SHEATH* (*NS*), enhances the growth of the leaf sheath and the proximal blade region (Scanlon et al., 1996; Nardmann et al., 2004). The *WOX4* gene regulates vascular cell differentiation through the auxin-dependent pathway for the lateral growth of plants (Suer et al., 2011). *AtWOX6 (PRETTY FEW SEEDS2* or *PFS2*) is required for ovule development and pattern formation (Park et al., 2005). *AtWOX7* plays the central integrating role in lateral root development and inhibits lateral root development in response to plant sugar status (Kong et al., 2016). *WOX2* was also found to be involved in somatic embryogenesis in *Larix decidua* (Rupps et al., 2016). The *WOX2*, *WOX8* and *WOX9* genes of *Arabidopsis* play a critical role in early embryo development where they act as an essential cell fate regulator and show the highest functional redundancy (Breuninger et al., 2008; Ueda et al., 2011). *WOX8* and *WOX9* are known as *STIMPY* (*WOX9*) and *STIMPY-LIKE* (*WOX8*) genes in *A. thaliana* and *EVERGREEN* (*EVG*) and *SISTER OF EVERGREEN* (*SOE*) in *Petunia*; they are essential for inflorescence development (Rebocho et al., 2008; Costanzo et al., 2014). *WOX11* and its homolog *WOX12* play an essential role in *de novo* root organogenesis (Liu et al., 2014). *WOX13* of ancient clade shows zestful expression during various developmental processes such as primary lateral root initiation, gynoecium and embryo development, floral transitions and vegetative fruit development (Deveaux et al., 2008; Romera-Branchat et al., 2013).

The current study was performed to characterize the *WOX* genes in *Phalaenopsis* orchids. Orchids are economically important plants renowned for their spectacular flowers with high longevity. *Phalaenopsis* is a tropical epiphytic orchid accounting for nearly 80% of the orchid trade in floriculture industry (Wu et al., 2012). This genus comprises nearly 70 species, of which *P. aphrodite*, *P. modesta*, *P. bellina* and *P. lueddemanniana* were selected for the present study. This study was conducted to identify and characterize the *WOX* genes in different *Phalaenopsis* species with a particular focus on the evolutionary analysis and expression profiling.

Materials and methods

Identification of the WOX family protein

To identify the WOX proteins of *P. aphrodite* (Su et al., 2011), *P. bellina*, *P. modesta*, and *P. lueddemanniana*, the TBLASTN search was performed in Orchidstra 2.0 database (http://orchidstra2.abrc.sinica.edu.tw/orchidstra2/index.php; Chao et al., 2017). The WOX protein sequences of *A. thaliana* (AtWOX) (van der Graaff et al., 2009), *Phalaenopsis equestris* (PeWOX) and *Dendrobium catenatum* (DcWOX) (Ramkumar et al., 2018) were used as queries against four species of *Phalaenopsis*. The presence of a homeobox domain (pfam00046) was verified using SMART (http://smart.embl-heidelberg.de/; Schultz et al., 2000) tools.

Conserved domain and motif analysis

The conserved domain (homeobox-domain) was identified using Expasy – Prosite (https://prosite. expasy.org/; Sigrist et al., 2012). MULTALIN (http://multalin.toulouse.inra.fr/multalin/; Corpet 1988) was used to identify the location of the DNA binding helix-loop-helix-turn-helix region. The MEME suite online server (http://meme-suite.org/tools/meme; Bailey et al., 2009) was used to identify the conserved motifs, with pre-set parameters (maximum number of motifs: 05, number of repetitions: any, optimal motif width: ≥ 6 and ≤ 150).

Physicochemical characterization

The molecular weight, aliphatic index, instability index, pI and grand average of hydropathicity (GRAVY) were calculated using the Expasy Protparam tool (https://web.expasy.org/protparam/; Gasteiger et al., 2005). The online tools CELLO v.2.5 (http://cello.life.nctu.edu.tw/; Yu et al. 2006) and WoLF PSORT (https://www.genscript.com/wolf-psort.html; Horton et al., 2007) were used to predict the subcellular location of the WOX proteins. The presence of a signal peptide and the transmembrane helix was predicted using the online tools SignalP.4.0 (http://www.cbs.dtu.dk/services/signalp/; Petersen et al., 2011) and TMHMM v.2.0 (http://www.cbs.dtu.dk/services/TMHMM/; Krogh et al., 2001).

Duplication events and ortholog prediction

The NCBI BLASTp search was performed to predict ortholog sequences of all four species of *Phalaenopsis* against closely related orchids, *P. equestris* (PeWOX),

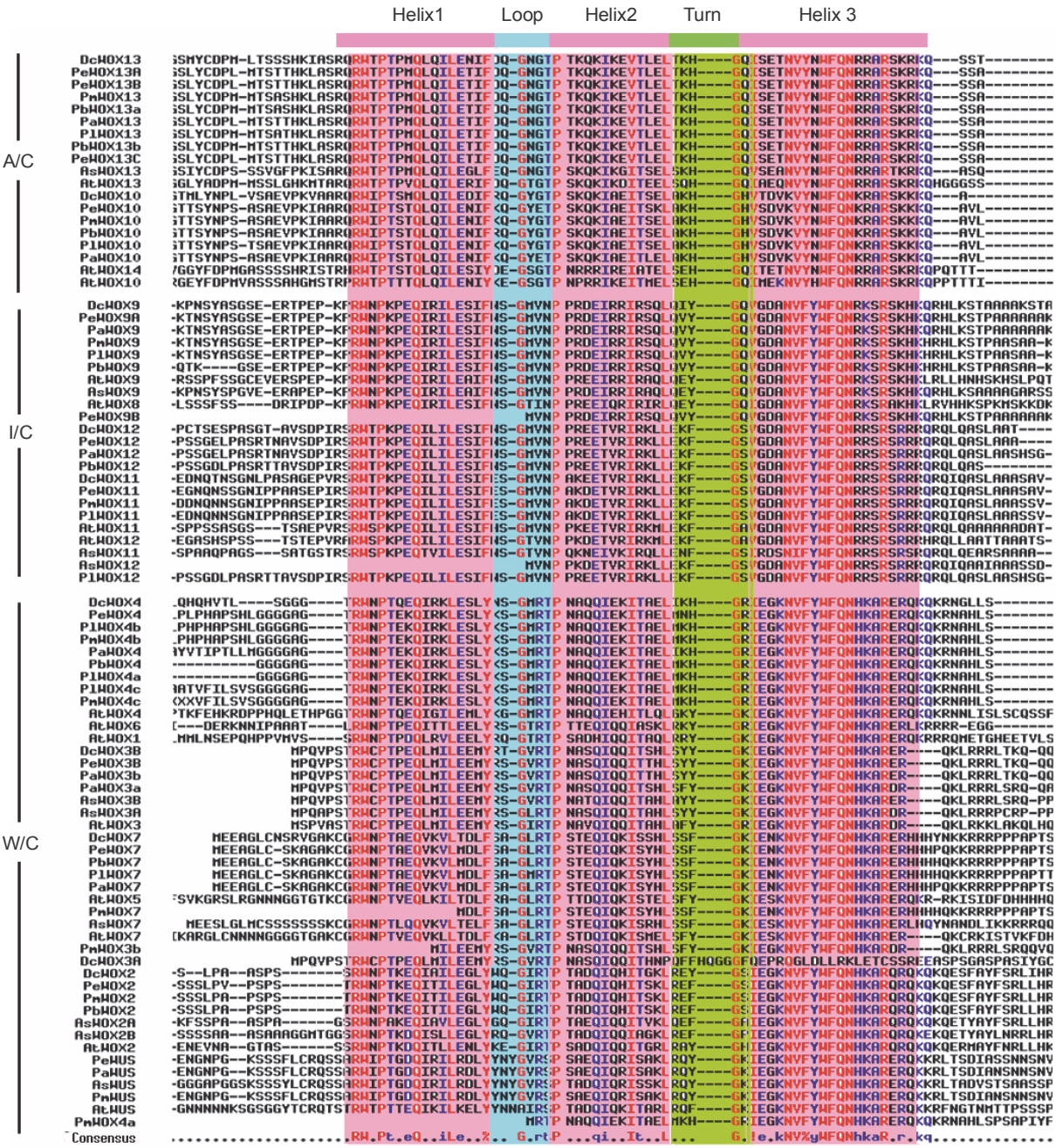

Fig. 1. Multiple sequence alignment of WOX protein sequences from *P. bellina, P. modesta, P. aphrodite* and *P. lueddemanniana* with *P. equestris, D. catenatum, A. shenzhenica* and *A. thaliana* shows homeodomain with the helix-loop-helix-turn-helix region, which is a characteristic of the WOX protein family (highly conserved amino acid sequence regions are shown in red and the less conserved ones are indicated in blue); the helix regions are marked in pink, the loop region in sky blue and the turn region in green

D. catenatum (DcWOX), *A. shenzhenica* (AsWOX) and the model plant *A. thaliana* (AtWOX), where each WOX protein query was independently blasted against WOX protein sequences of the target species. The percentage identity matrix of *WOX* gene sequences was analysed on the basis of the alignment of CDS sequences using MUSCLE (https://www.ebi.ac.uk/Tools/msa/muscle/; Edgar, 2004) server. More than 80% identity was shared by duplicated genes at the nucleotide level.

Phylogenetic analysis

The full-length and conserved homeodomain sequences of the putative PaWOX, PmWOX, PlWOX and PbWOX family members were aligned with the MUSCLE

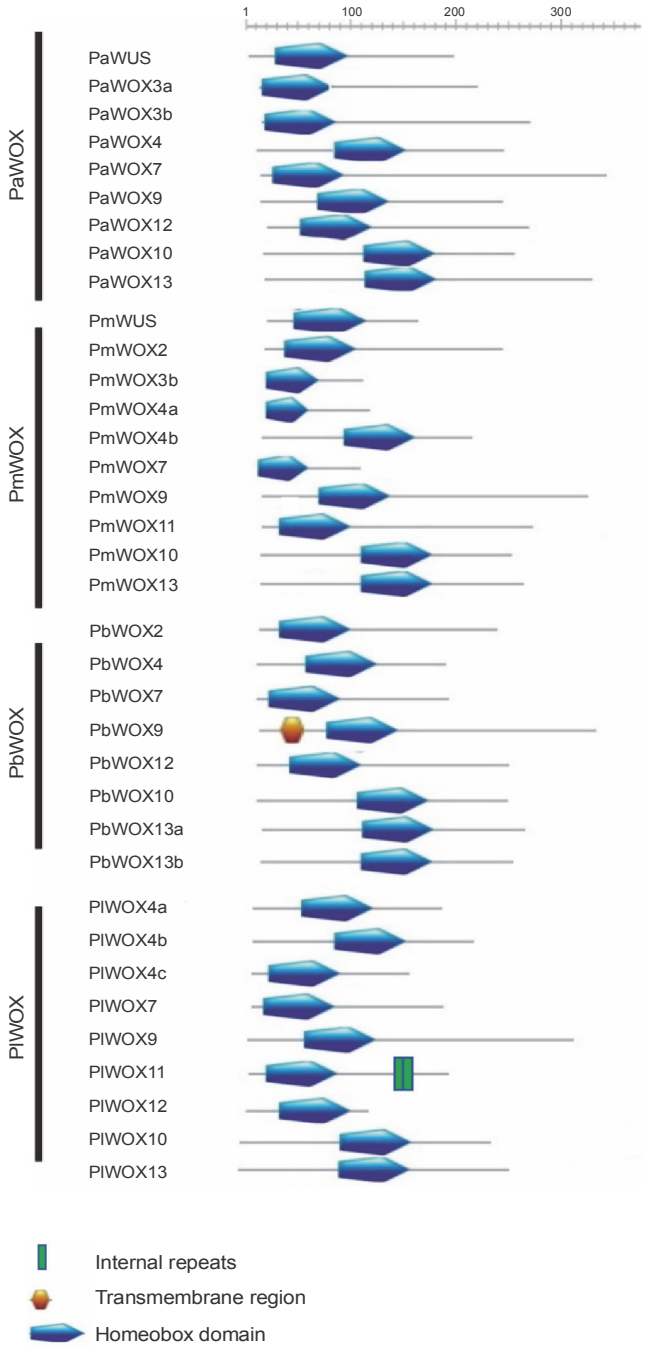

Fig. 2. The domain architecture analysis shows a homeobox domain for PaWOX, PmWOX, PbWOX and PlWOX protein sequences

program. A phylogenetic tree of the WOX protein sequences from six plant species (*A. thaliana, P. equestris*, *P. bellina*, *P. modesta*, *P. aphrodite*, *P. lueddemanniana*, *A. shenzhenica* and *D. catenatum*) was constructed based on the maximum likelihood method using the MEGA7 tool (http://www.megasoftware.net/; Kumar et al., 2016) with 1000 iterations for the bootstrap values.

Expression analysis

The CDS sequences of the WOX genes of all *Phalaenopsis* species were used as a query during the BLASTn search against the high-throughput RNA-seq data generated (NCBI-SRA database) from different developmental stages such as roots (SRX2439763), leaf (SRX2439762), fully opened flower (SRX2439761), large flower bud (SRX2196321) and small flower bud (SRX2439759) of *P. aphrodite*; small bud (SRX2210819), large bud (SRX2210818) and fully opened flower (SRX2210817) of *P. modesta*; root (SRX2210822), leaf (SRX2210823), large flower bud (SRX2210821) and fully opened flower (SRX2210820) of *P. bellina*; and root (SRX2210813), leaf (SRX2210814), large bud (SRX2210811), small bud (SRX2210812) and fully opened flower (SRX2210810) of *P. lueddemanniana* (https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastn&PAGE_TYPE=BlastSearch&BLAST_SPEC=SRA&LINK_LOC=blasttab), and the hit counts were noted. The RPKM (reads per kilobase per million) values were calculated using the formula [RPKM = (C × 109) / (N × L), where N is the total mapped reads in the RNA-seq experiment concerned, L is the base-pair length of the gene, and C is the number of hits for the candidate gene)] (Mortazavi et al., 2008). The heat maps were generated using the Hierarchical Clustering Explorer 3.5 (http://www.cs.umd.edu/hcil/hce/; Seo et al. 2006).

Secondary structures of WOX proteins

Secondary structures (alpha-helices, random coils, beta turns and extended strands) for the selected WOX protein sequences were predicted using the SOPMA (self-optimized prediction method with alignment), a secondary structure prediction tool (https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_sopma.html; Sapay et al., 2006). The tool includes the homologue method that takes information from an alignment of sequences belonging to the same protein family.

3D structure prediction and analysis of WOX proteins

The molecular modelling tool I-TASSER (https://zhanglab.ccmb.med.umich.edu/I-TASSER/; Yang et al., 2015) was used to predict the tertiary structure using top 10 homologues available in the protein data bank (PDB) with specific parameters [c-scores (confidence score for estimating the quality of the models predicted by I-TASSER), BS-score, TM-scores and IDEN coverage

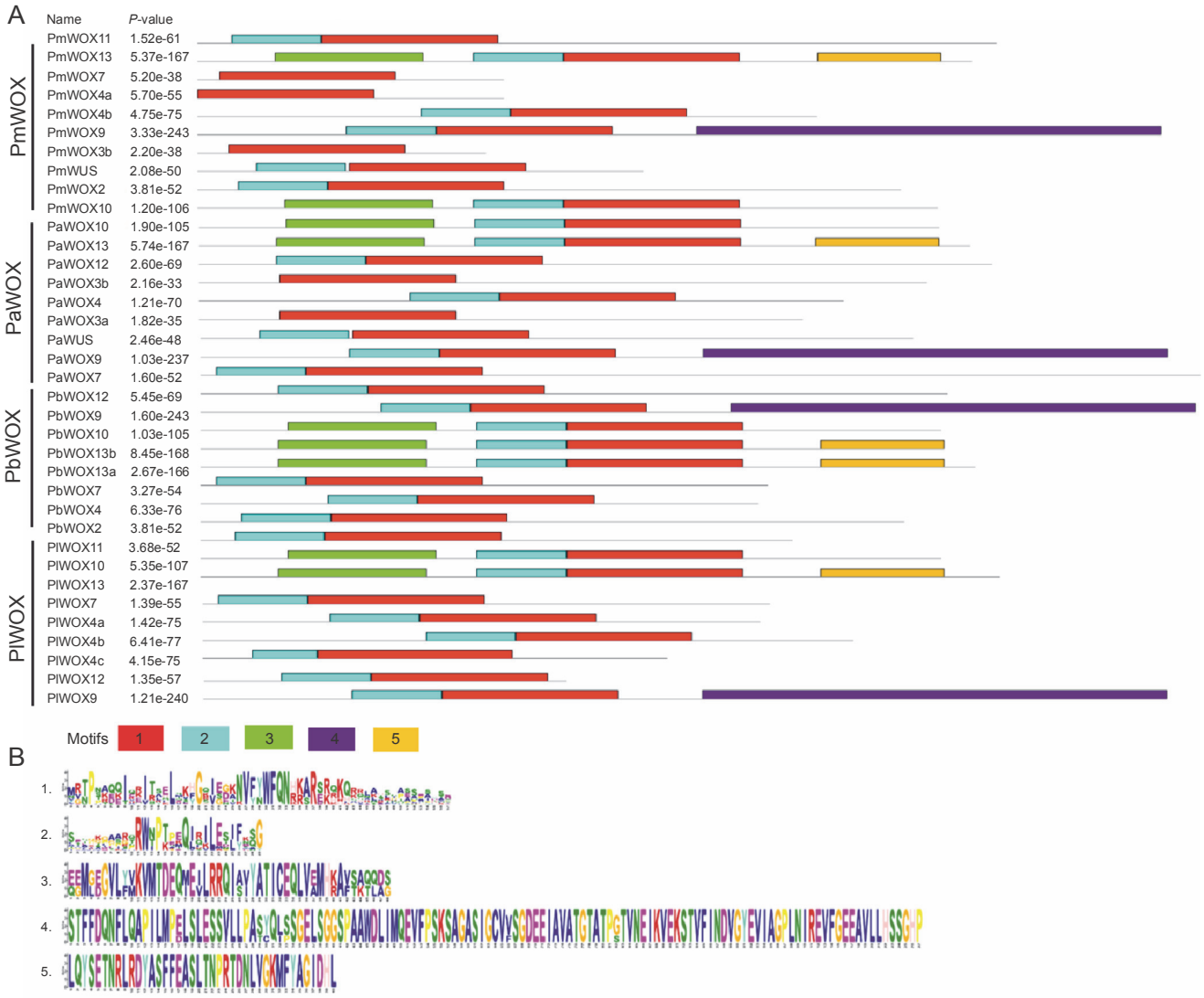

Fig. 3. The motif analysis for PaWOX, PmWOX, PbWOX and PlWOX: A) conserved motifs in sequences are marked in coloured boxes, B) sequence logo of WOX protein motifs

of the alignment by TM-align]. The binding site prediction with the BS-score value of >0.5 was considered as highly confident. The molecular viewer PyMOL (https://pymol.org/; DeLano 2002) was used to visualize the 3D protein structure.

Results

Protein identification and domain analysis

In the present study, we identified and characterized 36 *WOX* genes from *P. bellina, P. modesta, P. aphrodite* and *P. lueddemanniana*. The nomenclature of proteins and their respective genes was performed following their closest phylogenetic homologs in *A. thaliana, P. equestris* and *D. catenatum* (van der Graa et al., 2009; Ramkumar et al., 2018). All the protein sequences identified contained a *WUSCHEL*-related homeobox domain. The multiple sequence alignment showed the occurrence of the conserved helix-loop-helix-turn-helix region in homeodomain-encoded WOX proteins (Fig. 1). One transmembrane region was found in PbWOX9, and two internal repeats were found in PlWOX11 (Fig. 2). The maximum number of motifs was found in the ancient clade, whereas the least number of motifs was identified in the WUS clade (Fig. 3). Motifs 1 and 2 were present in all WOX proteins and were predicted to encode the homeodomain. Members of the modern and intermediate clades had similar motif compositions,

Table 1. Physiochemical characterization of WOX protein family

Protein name	Length [aa]	MW [kDa]	IP	INS	AI	GRAVY	Localization	Sp	TMD
Phalaenopsis aphrodite									
PaWUS	230	25832.44	5.99	57.57	56.04	−0.834	nuclear	no	0
PaWOX3a	195	21906.56	8.63	73.86	58.05	−0.756	nuclear	no	0
PaWOX3b	236	26956.41	6.52	81.90	61.14	−0.639	nuclear	no	0
PaWOX4	208	23454.73	10.01	67.30	72.26	−0.766	nuclear	no	0
PaWOX7	329	37804.14	6.44	48.52	71.43	−0.592	cytoplasm	no	0
PaWOX9	312	33887.09	6.8	51.26	77.24	−0.325	nuclear	no	0
PaWOX12	256	28069.26	5.61	69.40	72.38	−0.304	nuclear	no	0
PaWOX10	239	27208.62	5.72	53.71	66.19	−0.644	nuclear	no	0
PaWOX13	249	28559.01	5.23	59.81	64.26	−0.781	nuclear	no	0
Phalaenopsis modesta									
PmWUS	144	16681.78	9.63	57.43	62.36	−0.750	nuclear	no	0
PmWOX2	227	25638.89	6.71	89.12	63.66	−0.705	nuclear	no	0
PmWOX3b	93	11020.54	10.21	78.64	60.75	−0.823	nuclear	no	0
PmWOX4a	99	11713.19	6.84	53.56	63.13	−1.186	nuclear	no	0
PmWOX4b	200	22624.71	9.75	62.84	66.40	−0.800	nuclear	no	0
PmWOX7	99	11677.29	10.11	77.61	59.09	−0.979	nuclear	no	0
PmWOX9	311	33931.17	6.83	51.15	77.78	−0.358	nuclear	no	0
PmWOX11	258	27220.26	5.64	77.19	68.49	−0.234	nuclear	no	0
PmWOX10	239	26979.48	5.85	54.68	69.41	−0.544	nuclear	no	0
PmWOX13	250	28646.06	5.16	64.12	61.68	−0.814	nuclear	no	0
Phalaenopsis bellina									
PbWOX2	227	25621.9	6.71	86.85	63.66	−0.696	nuclear	no	0
PbWOX4	180	20529.26	9.76	60.51	68.33	−0.917	nuclear	no	0
PbWOX7	183	20937.99	9.18	59.00	71.42	−0.650	nuclear	no	0
PbWOX9	321	35235.05	7.1	48.60	88.38	−0.188	plasma membrane	no	1
PbWOX12	241	26531.48	5.44	73.83	69.59	−0.412	nuclear	no	0
PbWOX10	239	27120.55	5.86	56.38	69.00	−0.597	nuclear	no	0
PbWOX13a	240	27438.72	5.44	64.21	62.21	−0.807	nuclear	no	0
PbWOX13b	250	28706.16	5.16	64.61	61.68	−0.799	nuclear	no	0
Phalaenopsis lueddemanniana									
PlWOX4a	180	20486.19	9.76	63.91	66.17	−0.932	nuclear	no	0
PlWOX4b	210	23751.89	9.82	66.51	66.00	−0.843	nuclear	no	0
PlWOX4c	150	17134.28	9.45	60.99	65.07	−0.988	nuclear	no	0
PlWOX7	183	21031.13	9.30	56.07	67.16	−0.683	nuclear	no	0
PlWOX9	311	33935.16	6.83	51.43	78.10	−0.348	nuclear	no	0
PlWOX11	191	19883.87	6.59	81.24	60.84	−0.355	nuclear	no	0
PlWOX12	117	13123.66	11.05	73.83	66.75	−0.779	nuclear	no	0
PlWOX10	239	27191.68	6.02	56.81	68.58	−0.621	nuclear	no	0
PlWOX13	258	29538.14	5.34	65.51	63.18	−0.776	nuclear	no	0

* IP – isoelectric point, MW – protein molecular weight in kDa, INS – instability index, AI – aliphatic index, GRAVY – grand average of hydropathy, SP – signal peptide, TMD – transmembrane domain

Table 2. Clade-wise distribution of WOX proteins of *P. aphrodite, P. modesta, P. bellina* and *P. lueddemanniana*

CLADES	*P. bellina*	*P. modesta*	*P. lueddemanniana*	*P. aphrodite*
Ancient	PbWOX10 PbWOX13a PbWOX13b	PmWOX10, PmWOX13	PlWOX10 PlWOX13	PaWOX10 PaWOX13
Intermediate	PbWOX12 PbWOX9	PmWOX11 PmWOX9	PlWOX11 PlWOX12 PlWOX9	PaWOX12 PaWOX9
WUS	PbWOX7 PbWOX4 PbWOX2	PmWOX4a PmWOX4b PmWOX3b PmWUS PmWOX2 PmWOX7	PlWOX4a PlWOX4b PlWOX4c PlWOX7	PaWOX4 PaWOX3a PaWOX3b PaWUS PaWOX7

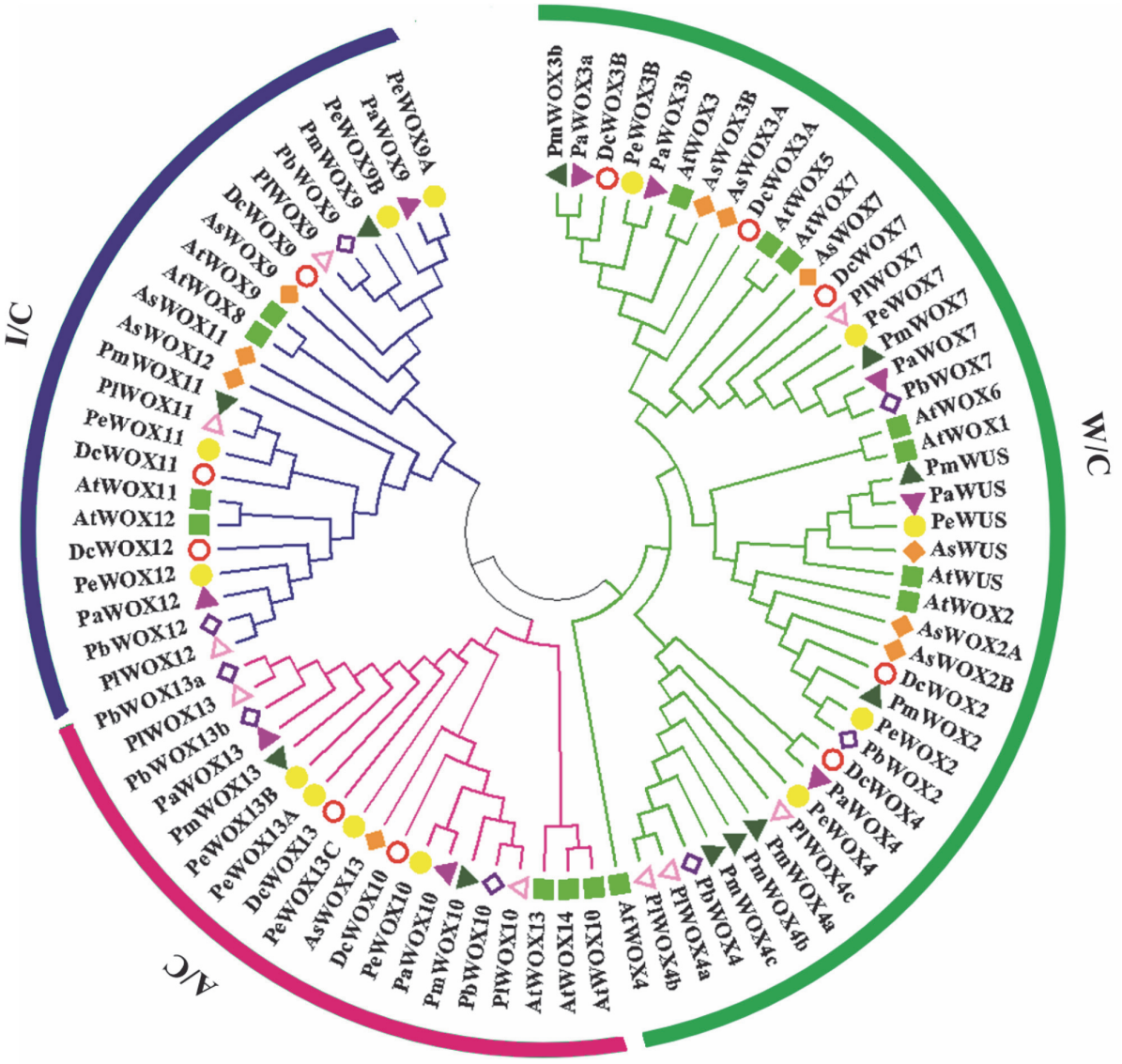

Fig. 4. The phylogenetic analysis of PaWOX, PmWOX, PbWOX and PlWOX proteins with AtWOX, PeWOX, AsWOX and DcWOX sequences shows clustering into ancient (A/C), intermediate (I/C) and WUS (W/C) clades (marked in pink, blue and green, respectively)

except for motif 4 which was unique to the intermediate clade. On the other hand, motifs 3 and 5 were unique to the ancient clade (Fig. 3).

Physicochemical analysis and topology study

The average size of PaWOX, PbWOX, PlWOX and PmWOX was 250 aa (amino acid), 236 aa, 204 aa and 232 aa, respectively. The average molecular mass of PaWOX, PbWOX, PlWOX and PmWOX was 27.6 kDa, 26.4 kDa, 23.24 kDa and 20.6 kDa, respectively. The isoelectric point for all the WOX proteins ranged from 5.16 to 11.01. The average values of the aliphatic index were 70.75, 69.29, 64.79 and 65.34 for PaWOX, PbWOX, PlWOX and PmWOX, respectively. A negative GRAVY value in all the sequences indicates that all WOX proteins were hydrophilic. All the proteins were found to be located in the nucleus, except for PbWOX9, which was predicted to be located in the plasma membrane, and PaWOX7, which was found in cytoplasm (Table 1).

Evolutionary analysis

The protein sequences of PaWOX, PmWOX, PbWOX and PlWOX were analysed along with 15 AtWOX, 14 PeWOX, 10 AsWOX (*A. shenzhenica* WOX) and 10 DcWOX sequences to predict their evolutionary relationship. A total of 36 protein sequences were clustered into three clades, namely ancient, intermediate and WUS, along with the respective members of AtWOX, PeWOX, AsWOX and DcWOX (Table 2, Fig. 4).

Orthologs and duplication event analysis

Orthologs counterparts for PaWOX, PmWOX, PbWOX and PlWOX protein sequences against AtWOX, PeWOX, DcWOX and AsWOX protein sequences were identified (Table S1 in the supplementary materials). Duplication events were predicted using *PaWOX, PmWOX, PbWOX* and *PlWOX* CDS sequences. Two duplication events were predicted in *P. lueddemanniana,* and only one duplication event was predicted in *P. bellina* and *P. modesta* each, where *PbWOX13a* and *PbWOX13b* shared 98.06% identity and *PmWOX4b* and *PmWOX4a* shared 86.19% identity, respectively. In *P. lueddemanniana*, *PlWOX4c* shared 96.64% identity with *PlWOX4a* and 96.84% identity with *PlWOX4b*, and *PlWOX4a* and *PlWOX4b* shared 99.63% identity. No duplication events were found in *P. aphrodite* (Table S2 in the supplementary materials).

Expression profiling of the WOX genes

The expression profiling of the *WOX* genes was performed in tissues at various developmental stages (leaf, root, flower, small bud and large buds) in all the species of *Phalaenopsis* investigated (*P. bellina, P. modesta, P. aphrodite* and *P. lueddemanniana*). The expression profile was determined by calculating the RPKM values (Table S3 in the supplementary materials) from hit counts in the NCBI-SRA database, and heat maps were then generated (Fig. 5). The expression of the *WOX* genes was found to be highly tissue-specific in all the four orchids. Most of the *WOX* genes were expressed in roots; however, this high expression was not replicated in other vegetative tissues, i.e. leaves. The expression of *PbWOX2, PbWOX4, PlWOX4a, PlWOX4b, PlWOX4c, PmWOX4a, PmWOX4b, PmWUS, PmWOX2, PmWOX3b, PmWOX7* and *PaWUS* (members of the WUS clade) was elevated in reproductive tissues. The intermediate clade genes, namely *PbWOX9, PmWOX9* and *PaWOX9*, also showed predominant expression in floral buds, except for *PlWOX9* which did not show high expression in any of the reproductive tissues. Similarly, the ancient clade genes also showed enhanced expression in the reproductive tissues. *PmWOX13* and *PaWOX13* showed high expression in floral buds, while *PbWOX13* and *PlWOX13* showed low expression in the reproductive tissues and high expression in roots. Similarly, *PbWOX10* was abundantly expressed in root tissues.

Homology modelling of WOX proteins

Homology modelling revolutionized the deciphering of protein structure and the mechanism of protein function. Seven WOX proteins (PaWOX9, PaWUS, PbWOX2, PbWOX4, PmWOX13, PmWOX11 and PmWOX10) were selected from *P. bellina*, *P. modesta* and *P. aphrodite* for the secondary and tertiary structure prediction on the basis of their significant expression level in the reproductive tissues and a clade-wise distribution (Table 3, Fig. 6). Alpha-helix and random coils were found to be dominant in the secondary structures (Fig. 6.1A–G). The 3D structure prediction showed the presence of almost an equal number of alpha-helices in PaWUS, PbWOX2, PbWOX4, PmWOX13, PmWOX11 and PmWOX10 (Fig. 6.2 I–N). The presence of beta-strand was also predicted in a single protein, i.e. PaWOX9 (Fig. 6.2H). The ligand binding analysis showed that PbWOX2, PmWOX10 and PmWOX13 protein sequences have

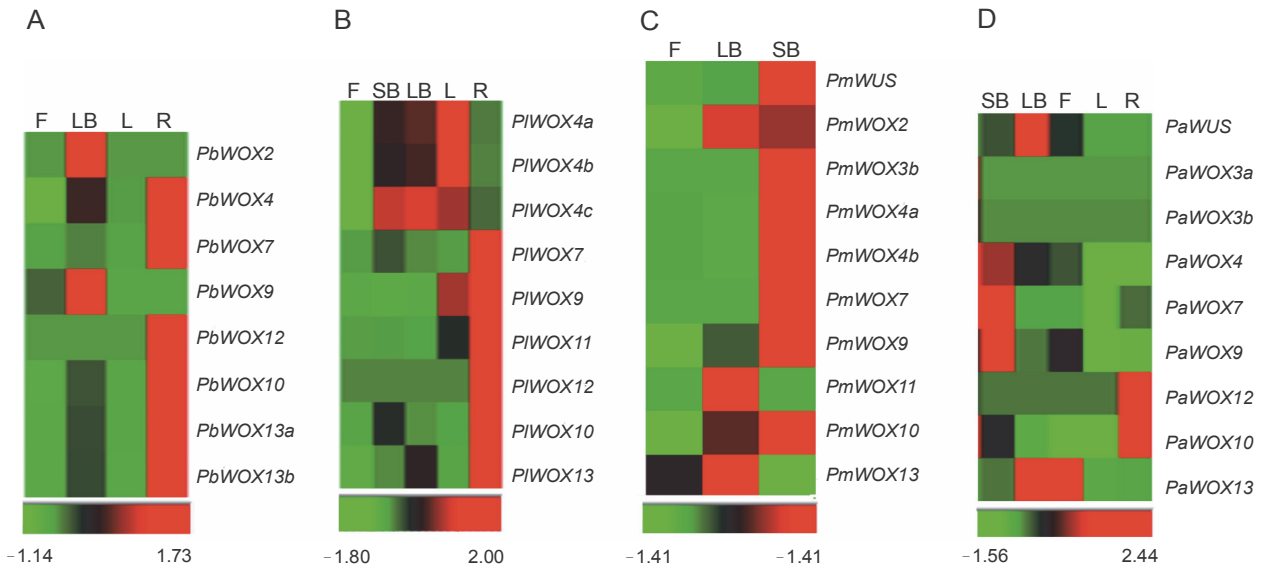

Fig. 5. The expression analysis of the WOX genes in various tissues: A) *PbWOX*, B) *PlWOX*, C) *PmWOX* and D) *PaWOX* genes (fully opened flower – F, leaf – L, root – R, small bud – SB, large bud – LB, small floral bud – SB, large floral bud – LB)

Table 3. Prediction of secondary structure and ligand binding sites in selected WOX proteins

Name	AH	ES	BT	RC	LI	BS
PaWUS	16.96% (39)	6.52% (15)	3.48% (8)	73.04% (168)	Zn2+	83, 87
PaWOX9	20.83% (65)	13.46% (42)	4.81% (15)	60.90% (190)	Mn2+	71, 97, 104, 106
PbWOX2	24.67% (56)	9.69% (22)	5.29% (12)	60.35% (137)	NU	21, 22, 24, 45, 48, 70, 74, 77, 81
PbWOX4	33.33% (60)	14.44% (26)	4.44% (8)	47.78% (86)	phosphoric acid	65, 66, 70, 80, 81, 82
PmWOX11	28.68% (74)	16.28% (42)	8.14% (21)	46.90% (121)	Mg2+	26, 29
PmWOX10	52.30% (125)	5.44% (13)	6.69% (16)	35.56% (85)	NU	100, 101, 103, 144, 144, 147, 148, 150, 151, 155
PmWOX13	45.60% (114)	6.80% (17)	4.80% (12)	42.80% (107)	NU	97, 98, 99, 100, 101, 103, 144, 147, 148, 151, 155

* AH – alpha helix, RC – random coil, ES – extended strand, BT – beta turn, LI – ligand, BS – binding sites, NU – nucleic acid, Zn^{2+} – zinc, Mn^{2+} – manganese^{2+}, Mg^{2+} – magnesium

specific binding sites for the nucleic acid ligand (Fig. 6.2 J, L, N). PmWOX11, PbWOX4, PaWUS and PaWOX9 contained magnesium (Mg^{2+}), phosphoric acid, zinc (Zn^{2+}) and manganese (Mn^{2+}) binding sites, respectively (Fig. 6.2 M, K, I, H).

Discussion

The *in silico* characterization of genes and gene families has become an important research tool in molecular biology to understand various biological pathways involved in growth and development. The *WOX* gene family plays a versatile role in embryo pattern formation, organogenesis and florigenesis (van der Graaff et al., 2009; Costanzo et al., 2014). The present study involved protein characterization, topology study and phylogenetic analysis of WOX proteins in four commercially important orchids, namely *P. aphrodite, P. modesta, P. bellina* and *P. lueddemanniana.* The transcriptomic analysis of *P. aphrodite, P. modesta, P. bellina* and *P. lueddemanniana* showed that they encode nine, ten, nine and eight full-length WOX proteins, respectively. This finding was comparable to that reported for *Phalaenopsis* (14), *Dendrobium* (10) and *Apostasia* (10)

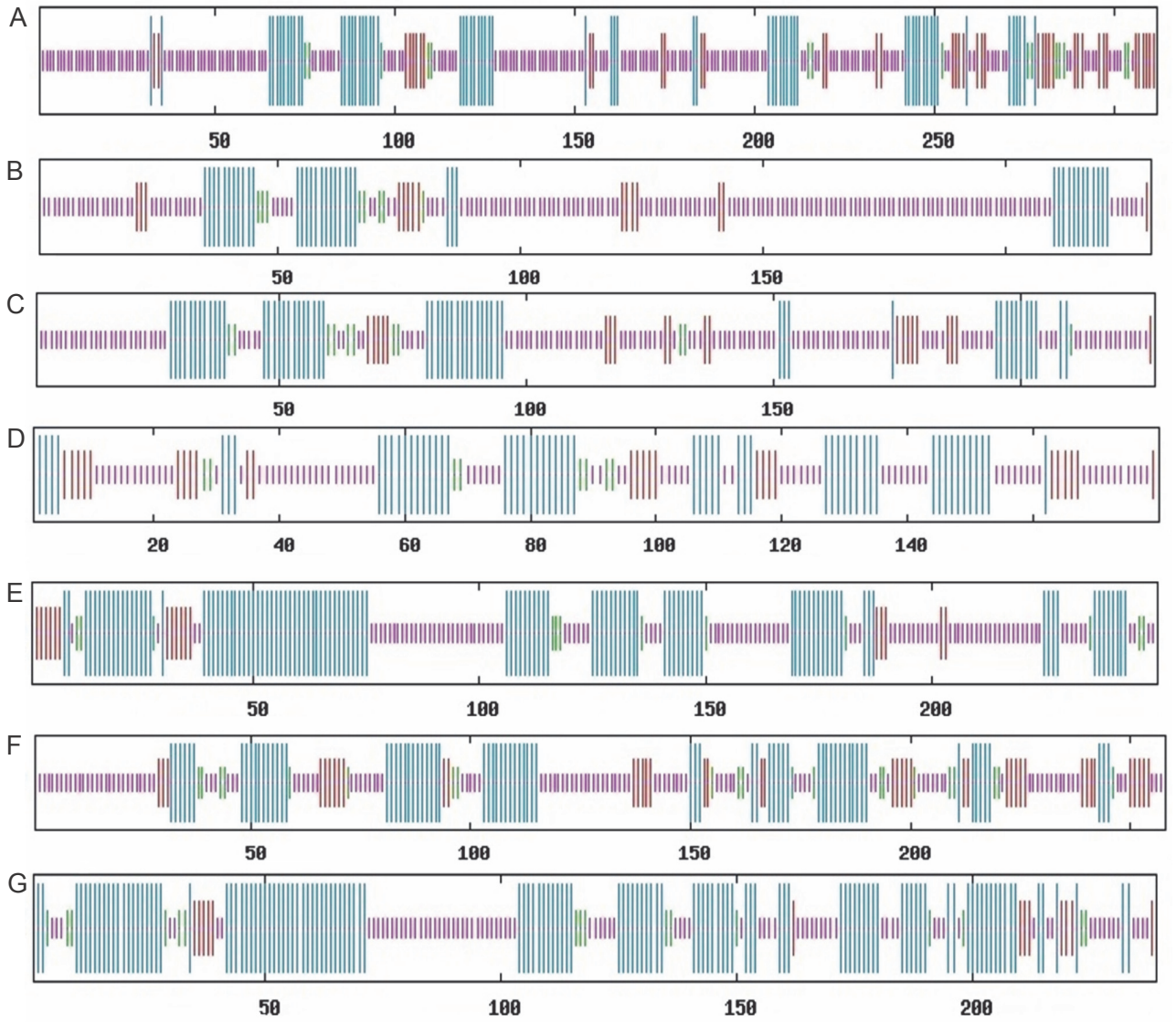

Fig. 6.1. Secondary structure analysis: A) PaWOX9, B) PaWUS, C) PbWOX2, D) PbWOX4, E) PmWOX10, F) PmWOX11 and G) PmWOX13

(Ramkumar et al., 2018, 2020). The multiple sequence alignment showed that all the predicted sequences have a *WUSCHEL*-related homeobox structure with a helix-loop-helix-turn-helix region, which is the major characteristic feature of this gene family. Evolutionary studies showed that the *WOX* genes of *P. aphrodite, P. modesta, P. bellina* and *P. lueddemanniana* can be sub-grouped into three clades (Ancient, Intermediate and WUS) in a ratio of 2 : 2 : 5, 2 : 2 : 6, 3 : 2 : 3 and 2 : 3 : 4, respectively. The tight clustering formed during the construction of the phylogenetic tree along with *P. equestris* and *D. catenatum* showed the common origin of these plants, which is substantiated by their inclusion in the common sub-family *Epidendroideae* (Freudenstein et al., 2015). The predicted average size and molecular mass of all the PaWOX, PbWOX, PlWOX and PmWOX protein sequences were 250 aa, 236 aa, 204 aa and 232 aa and 27.6 kDa, 26.4 kDa, 23.24 kDa and 20.6 kDa, respectively; these values are in sync with those reported for *P. equestris*, *D. catenatum* and *A. shenzhenica* (Ramkumar et al., 2018, 2020). All the predicted proteins were hydrophilic in nature and located in the nucleus, which is in conformity with previous studies (Sakakibara et al., 2014; Ramkumar et al., 2018, 2020). The duplication event prediction showed the presence of two duplication events in the WUS clade of *P. lueddemanniana*

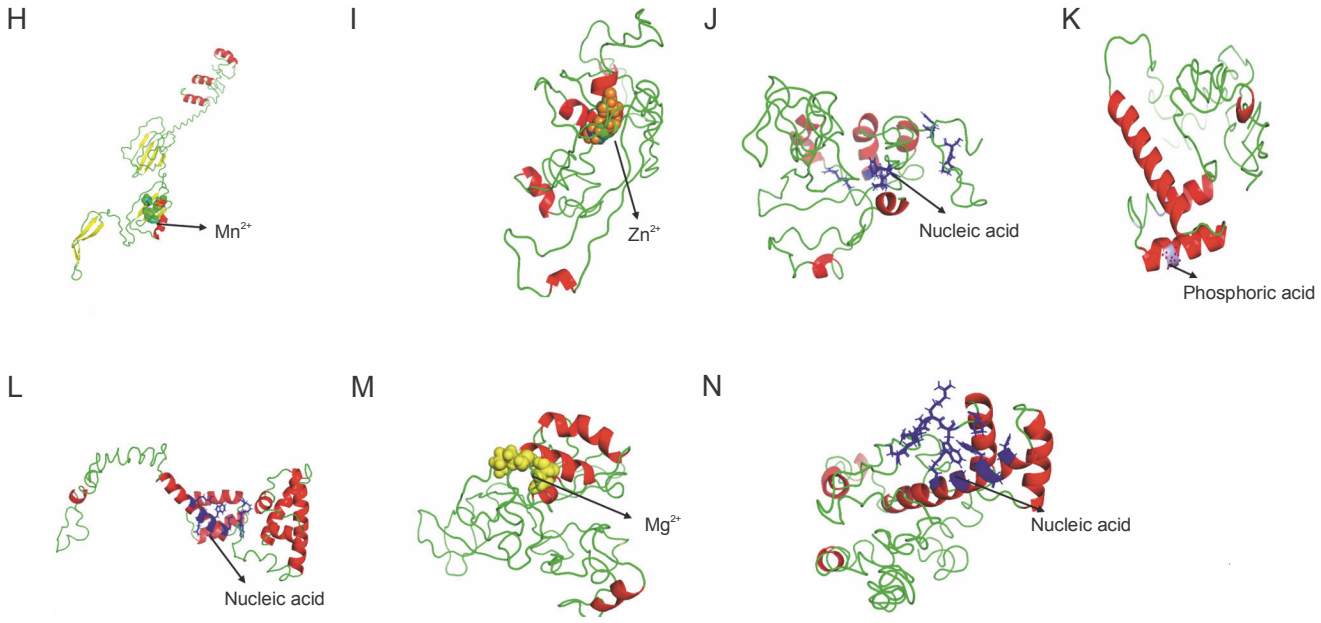

Fig. 6.2. Simulated 3D structures with ligand binding sites: H) PaWOX9, I) PaWUS, J) PbWOX2, K) PbWOX4, L) PmWOX10, M) PmWOX11 and N) PmWOX13

(PlWOX4a, PlWOX4b and PlWOX4c) and one duplication event in the ancient clade of *P. bellina* (PbWOX13a and PbWOX13b). PbWOX13a and PbWOX13b clustered with PeWOX13A/B/C of *P. equestris*, which also has been reported to have two duplication events (PeWOX13A, PeWOX13B and PeWOX13C) (Ramkumar et al., 2018).

The expression pattern analysis of the *WOX* genes at various developmental stages of plants showed that the *WOX* gene family has diverse roles in several developmental processes. In our studies, maximum genes were predicted to have good expression in floral organs, which indicates that they might play an active role during floral development as previously reported in many plants (Deyhle et al., 2007; Ikeda et al., 2009; Romera- Branchat et al., 2013; Costanzo et al., 2014). The expression profiling revealed that *PaWUS* and *PmWUS* showed high expression in floral buds, which is in accordance with earlier studies on the *WUS* gene of *A. thaliana* in which a high expression was related to the regulation of floral patterning by meristem maintenance in floral buds (Ikeda et al., 2009). The predominant expression of the *PbWOX13a/b, PlWOX13* and *PaWOX13* genes was noted in floral buds and flowers; additionally, abundant expression of *PbWOX13a/b* and *PlWOX13* was also noted in roots. Importantly, it was reported that *WOX13* plays an important role in the primary and lateral root development, floral transition and replum development in *A. thaliana* (Romera-Branchat et al., 2013); thus, *PmWOX2* and *PbWOX2* with similar expression levels might be playing similar roles. The *PlWOX11* and *PlWOX12* genes and their orthologs *PbWOX12* and *PaWOX12* had preferential expression in all the tested tissues, but abundant expression was noted in roots; this indicates their involvement in root development, as already reported in *A. thaliana* and rice, where these genes function in growth promotion (Hu et al., 2016; Jiang et al., 2017). Protein homology modelling is necessary to understand protein structure, ligand binding and functional mechanisms. In the present work, we selected seven WOX proteins clade wise for the protein homology analysis on the basis of their expression pattern. The predicted 3D structures and ligand binding according to their alignment with top 10 homologous PDB templates of other homeodomain proteins revealed the conserved nature of WOX proteins. The protein-ligand interaction analysis of PaWOX9, PaWUS, PbWOX2, PbWOX4, PmWOX13, PmWOX11 and PmWOX10 showed the presence of five different ligand binding sites, i.e. nucleic acid, Mg^{2+}, Mn^{2+}, phosphoric acid and zinc. PbWOX2, PmWOX10 and PmWOX13 have nucleic acid ligand binding sites, which supports the finding that these proteins have a DNA binding domain

(Gehring et al., 1994). However, the binding sites of phosphoric acid in PbWOX4, Mn^{2+} in PaWOX9, Zn^{2+} in PaWUS and Mg^{2+} in PmWOX11 were unexpected (Fig. 6.2H–N). The present study is insufficient to confirm these metal-binding properties of WOX proteins, and an in depth *in vitro* study is required to provide more supporting evidence.

Conclusions

As the molecular mechanisms of growth and development are poorly studied in orchids, the present study provides an insight into the WOX gene family in *P. aphrodite, P. modesta, P. bellina* and *P. lueddemanniana*. This study illustrates that WOX members are conserved in nature at both the sequence and structural levels. They are clustered in their respective clades and show increased expression in floral organs. The present study can help in the functional elucidation of candidate genes for understanding the growth and development in these commercially important orchid species.

Acknowledgements

MK is thankful to the Department of Science and Technology (DST) for INSPIRE Fellowship for Research Students (File Number: DST/INSPIRE/03/2017/002346). JKS and MK are grateful to the Department of Science and Technology (DST), Government of India for the partial financial support under the Promotion of University Research and Scientific Excellence (PURSE) grant scheme.

Authors' contributions

JKS designed the work. MK executed the experiments. MK and JKS prepared the manuscript. All authors have read and approved the final manuscript.

Competing interests

The authors declare that they have no competing interests.

References

Arroyo-Herrera A., Ku Gonzalez A., Canche Moo R., Quiroz-Figueroa F.R., Loyola-Vargas V.M., Rodriguez-Zapata L.C., Burgeff D Hondt C., Suárez-Solís V.M., Castaño E. (2008) *Expression of WUSCHEL in Coffea canephora causes ectopic morphogenesis and increases somatic embryogenesis*. Plant Cell Tissue Organ Cult. 94: 171–180. doi.org/10.1007/s11240-008-9401-1

Bailey T.L., Boden M., Buske F.A., Frith M., Grant C.E., Clementi L., Ren J., Li W.W., Noble W.S. (2009) *MEME SUITE: tools for motif discovery and searching*. Nucl. Acids Res. 37: W202–W208. doi.org/10.1093/ nar/gkp335

Bouchabké-Coussa O., Obellianne M., Linderme D., Montes E., Maia-Grondard A., Vilaine F., Pannetier C. (2013) *WUSCHEL overexpression promotes somatic embryogenesis and induces organogenesis in cotton (Gossypium hirsutum L.) tissues cultured in vitro*. Plant Cell Rep. 32: 675–686. doi.org/10.1007/s00299-013-1402-9

Breuninger H., Rikirsch E., Hermann M., Ueda M., Laux T. (2008) *Differential expression of WOX genes mediates apical-basal axis formation in the Arabidopsis embryo*. Develop. Cell 14: 867–876. doi.org/10.1016/j.devcel.2008.03.008

Chao Y.T., Yen S.H., Yeh J.H., Chen W.C., Shih M.C. (2017) *Orchidstra 2.0-a transcriptomics resource for the orchid family*. Plant Cell Physiol. 58: e9. doi.org/10.1093/pcp/pcw220

Corpet F. (1988) *Multiple sequence alignment with hierarchical clustering*. Nucl. Acids Res. 16: 10881–10890. doi.org/10.1093/nar/16.22.10881

Costanzo E., Trehin C., Vandenbussche M. (2014) *The role of WOX genes in flower development*. Ann. Bot. 114: 1545–1553. doi.org/10.1093/aob/mcu123

DeLano W.L. (2002) *Pymol: an open-source molecular graphics tool*. CCP4 Newsletter Protein Crystallogr. 40: 82–92.

Deveaux Y., Toffano-Nioche C., Claisse G., Thareau V., Morin H., Laufs P., Moreau H., Kreis M., Lecharny A. (2008) *Genes of the most conserved WOX clade in plants affect root and flower development in Arabidopsis*. BMC Evolut. Biol. 8: 291. doi.org/10.1186/1471-2148-8-291

Deyhle F., Sarkar A.K., Tucker E.J., Laux T. (2007) *WUSCHEL regulates cell differentiation during anther development*. Develop. Biol. 302: 154–159. doi.org/10.1016/j.ydbio.2006.09.013

Edgar R.C. (2004) *MUSCLE: multiple sequence alignment with high accuracy and high throughput*. Nucl. Acids Res. 32: 1792–1797. doi.org/10.1093/nar/gkh340

Freudenstein J.V., Chase M.W. (2015) *Phylogenetic relationships in Epidendroideae (Orchidaceae), one of the great flowering plant radiations: progressive specialization and diversification*. Ann. Bot. 115: 665–681. doi.org/ 10.1093/aob/mcu253

Gasteiger E., Hoogland C., Gattiker A., Duvaud S., Wilkins M.R., Appel R.D., Bairoch A. (2005) *Protein identification and analysis tools on the ExPASy server*. Springer.

Gehring W.J., Affolter M., Burglin T. (1994) *Homeodomain proteins*. Ann. Rev. Biochem. 63: 487–526. doi.org/10.1146/annurev.bi.63.070194.002415

Haecker A., Groß-Hardt R., Geiges B., Sarkar A., Breuninger H., Herrmann M., Laux T. (2004) *Expression dynamics of WOX genes mark cell fate decisions during early embryonic patterning in Arabidopsis thaliana*. Development 131: 657–668. doi:10.1242/dev.00963

Horton P., Park K.J., Obayashi T., Fujita N., Harada H., Adams-Collier C.J., Nakai K. (2007) *WoLF PSORT: protein localization predictor*. Nucl. Acids Res. 35: 585–587. doi.org/10.1093/nar/gkm259

Hu X., Xu L. (2016) *Transcription factors WOX11/12 directly activate WOX5/7 to promote root primordia initiation and organogenesis*. Plant Physiol. 172: 2363–2373. doi.org/10.1104/pp.16.01067

Ikeda M., Mitsuda N., Ohme-Takagi M. (2009) *Arabidopsis WUSCHEL is a bifunctional transcription factor that acts as a repressor in stem cell regulation and as an activator in floral patterning*. Plant Cell 21: 3493–3505. doi.org/10.1105/tpc.109.069997

Jha P., Ochatt S.J., Kumar V. (2020) *WUSCHEL: a master regulator in plant growth signaling*. Plant Cell Rep. 1: 1–4. doi.org/10.1007/s00299-020-02511-5

Jiang W., Zhou S., Zhang Q., Song H., Zhou D.X., Zhao Y. (2017) *Transcriptional regulatory network of WOX11 is involved in the control of crown root development, cytokinin signals, and redox in rice*. J. Exp. Bot. 68: 2787–2798. doi.org/10.1093/jxb/erx153

Kong D., Hao Y., Cui H. (2016) *The WUSCHEL related homeobox protein WOX7 regulates the sugar response of lateral root development in Arabidopsis thaliana*. Mol. Plant 9: 261–270. doi.org/10.1016/j.molp.2015. 11.006

Krogh A., Larsson B., von Heijne G., Sonnhammer E.L. (2001) *Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes*. J. Mol. Biol. 305: 567–580. doi.org/10.1006/jmbi.2000. 4315

Kumar S., Stecher G., Tamura K. (2016) *MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets*. Mol. Biol. Evol. 33: 1870–1874. doi.org/10.1093/molbev/msw054

Laux T., Mayer K.F., Berger J., Jurgens G. (1996) *The WUSCHEL gene is required for shoot and floral meristem integrity in Arabidopsis*. Development 122: 87–96. doi.org/10.1371/journal.pone.0038161

Lian G., Ding Z., Wang Q., Zhang D., Xu J. (2014) *Origins and evolution of WUSCHEL-related homeobox protein family in plant Kingdom*. Sci. World J. 2014: 1–12. doi.org/10.1155/2014/534140

Liu J., Sheng L., Xu Y., Li J., Yang Z., Huang H., Xu L. (2014) *WOX11 and 12 are involved in the first-step cell fate transition during de novo root organogenesis in Arabidopsis*. Plant Cell 26: 1081–1093. doi.org/10.1105/tpc.114.122887

Lohmann J.U., Hong R.L., Hobe M., Busch M. A., Parcy F., Simon R., Weigel D. (2001) *A molecular link between stem cell regulation and floral patterning in Arabidopsis*. Cell 105: 793–803. doi.org/10.1016/S0092-8674(01) 00384-1

Matsumoto N., Okada K. (2001) *A homeobox gene, PRESSED FLOWER, regulates lateral axis-dependent development of Arabidopsis flowers*. Genes Develop. 15: 3355–3364. www.genesdev.org/cgi/doi/10.1101/gad.931001.

Mayer K.F., Schoof H., Haecker A., Lenhard M., Jurgens G., Laux T. (1998) *Role of WUSCHEL in regulating stem cell fate in the Arabidopsis shoot meristem*. Cell 95: 805–815. doi.org/10.1016/S0092-8674(00)81703-1

Mortazavi A., Williams B.A., McCue K., Schaeffer L., Wold B. (2008) *Mapping and quantifying mammalian transcriptomes by RNA-Seq*. Nature Meth. 5: 621–628. doi.org/10.1038/nmeth.1226

Nardmann J., Ji J., Werr W., Scanlon M.J. (2004) *The maize duplicate genes narrow sheath1 and narrow sheath2 encode a conserved homeobox gene function in a lateral domain of shoot apical meristems*. Development 131: 2827–2839. doi: 10.1242/dev.01164

Oshchepkova E.A., Omelyanchuk N.A., Savina M.S., Pasternak T., Kolchanov N.A., Zemlyanskaya E.V. (2017) *Systems biology analysis of the WOX5 gene and its functions in the root stem cell niche*. Russ. J. Gen. Appl. Res. 7: 404–420. doi.org/10.1134/S2079059717040086

Park S.O., Zheng Z., Oppenheimer D.G., Hauser B.A. (2005) *The PRETTY FEW SEEDS2 gene encodes an Arabidopsis homeodomain protein that regulates ovule development*. Development 132: 841–849. doi: 10.1242/dev. 01654

Petersen T.N., Brunak S., von Heijne G., Nielsen H. (2011) *SignalP 4.0: discriminating signal peptides from transmembrane regions*. Nature Meth. 8: 785–786. doi.org/10.1038/nmeth.1701

Ramkumar T.R., Kanchan M., Upadhyay S.K., Sembi J.K. (2018) *Identification and characterization of WUSCHEL-related homeobox (WOX) gene family in economically important orchid species Phalaenopsis equestris and Dendrobium catenatum*. Plant Gene 14: 37–45. doi.org/ 10.1016/j.plgene.2018.04.004

Ramkumar T.R., Kanchan M., Sembi J.K. (2020) *Genome wide characterization of WUSCHEL-related homeobox (WOX) gene family in Apostasia shenzhenica, a primeval orchid*. Plant Sci. Today 7: 164–171. doi.org/10.14719/pst.2020.7.2.703

Rebocho B.A., Bliek M., Kusters E., Castel R., Procissi A., Roobeek I., Souer E., Koes R. (2008) *Role of EVERGREEN in the development of the Cymose petunia inflorescence*. Develop. Cell 15: 437–447. doi.org/10.1016/j.devcel.2008.08.007

Romera-Branchat M., Ripoll J.J., Yanofsky M.F., Pelaz S. (2013) *The WOX13 homeobox gene promotes replum formation in the Arabidopsis thaliana fruit*. Plant J. 73: 37–49. doi.org/10.1111/tpj.12010

Rupps A., Raschke J., Rummler M., Linke B., Zoglauer K. (2016) *Identi cation of putative homologs of Larix decidua to babyboom (BBM), leafy cotyledon1 (LEC1), WUSCHEL-related homeobox2 (WOX2) and somatic embryogenesis receptor-like kinase (SERK) during somatic embryogenesis*. Planta 243: 473–488. doi.org/10.1016/j.gene.2018.12.049

Sakakibara K., Reisewitz P., Aoyama T., Friedrich T., Ando S., Sato Y., Tamada Y., Nishiyama T., Hiwatashi Y., Kurata T., Ishikawa M. (2014) *WOX13-like genes are required for reprogramming of leaf and protoplast cells into stem cells in the moss Physcomitrella patens*. Development 141: 1660–1670. https://doi.org/10.1242/dev.097444

Sapay N., Guermeur Y., Deléage G. (2006) *Prediction of amphipathic in-plane membrane anchors in monotopic proteins using a SVM classifier*. BMC Bioinformatics 7: 255. doi.org/10.1093/bioinformatics/btk052

Scanlon M.J., Schneeberger R.G., Freeling M. (1996) *The maize mutant narrow sheath fails to establish leaf margin identity in a meristematic domain*. Development 122: 1683–1691.

Schultz J., Copley R.R., Doerks T., Ponting C.P., Bork P. (2000) *SMART: a web-based tool for the study of genetically mobile domains*. Nucl. Acids Res. 28: 231–234. doi.org/10.1093/nar/28.1.231

Seo J., Gordish-Dressman H., Hoffman E.P. (2006) *An interactive power analysis tool for microarray hypothesis testing and generation*. Bioinformatics 22: 808–814. doi.org/10.1093/bioinformatics/btk052

Sigrist C.J.A., de Castro E., Cerutti L., Cuche B.A., Hulo N., Bridge A., Bougueleret L., Xenarios I. (2012) *New and continuing developments at PROSITE*. Nucl. Acids Res. 41: 344–347. doi.org/10.1093/nar/gks1067

Su C.L., Chao Y.T., Alex Chang Y.C., Chen W.C., Chen C.Y., Lee A.Y., Shih M.C. (2011) *De novo assembly of expressed transcripts and global analysis of the Phalaenopsis aphrodite transcriptome*. Plant Cell Physiol. 52: 1501–1514. doi.org/10.1093/pcp/pcr097

Suer S., Agusti J., Sanchez P., Schwarz M., Greb T. (2011) *WOX4 imparts auxin responsiveness to cambium cells in Arabidopsis*. Plant Cell 23: 3247–3259. doi.org/10.1105/tpc.111.087874

Ueda M., Zhang Z., Laux T. (2011) *Transcriptional activation of Arabidopsis axis patterning genes WOX8/9 links yygote polarity to embryo development*. Develop. Cell 20: 264–270. doi.org/10.1016/j.devcel.2011.01.009

van der Graaff E., Laux T., Rensing S.A. (2009) *The WUS homeobox-containing (WOX) protein family*. Genome Biol. 10: 248. doi.org/10.1186/gb-2009-10-12-248

Wu P.H., Chang C.N.D. (2012) *Cytokinin treatment and flower quality in Phalaenopsis orchids: Comparing N-6-benzyladenine, kinetin and 2-isopentenyl adenine*. Afr. J. Biotech. 11: 1592–1596. doi.org/10.5897/AJB11.2472

Yang J., Yan R., Roy A., Xu D., Poisson J., Zhang Y. (2015) *The I-TASSER Suite: protein structure and function prediction*. Nature Meth. 12: 7–8. doi.org/10.1038/nmeth. 3213

Yu C.S., Chen Y.C., Lu C.H., Hwang J.K. (2006) *Prediction of protein subcellular localization*. Prot. Struct. Func. Bioinform. 64: 643–651. doi.org/10.1002/prot.21018

Zhang X., Zong J., Liu J., Yin J., Zhang D. (2010) *Genome-wide analysis of WOX gene family in rice, sorghum, maize, Arabidopsis and poplar*. J. Integr. Plant Biol. 52: 1016–1026. dx.doi.org/10.1111/j.1744-7909.2010. 00982.x

Zuo J., Niu Q.W., Frugis G., Chua N.H. (2002) *The WUSCHEL gene promotes vegetative-to-embryonic transition in Arabidopsis*. Plant J. 30: 349–359. doi.org/10.1046/ j.1365-313X.2002.01289.x