Protoplasma (2024) 261:397–410
https://doi.org/10.1007/s00709-023-01917-8

REVIEW

Domain of unknown function (DUF) proteins in plants: function and perspective

Chengke Luo[1] · Maryam Akhtar[2] · Weifang Min[1] · Xiaorong Bai[1] · Tianli Ma[1] · Caixia Liu[1]

Received: 20 June 2023 / Accepted: 8 December 2023 / Published online: 30 December 2023

Abstract

Domains of unknown function (DUFs), which are deposited in the protein family database (Pfam), are protein domains with conserved amino acid sequences and uncharacterized functions. Proteins with the same DUF were classified as DUF families. Although DUF families are generally not essential for the survival of plants, they play roles in plant development and adaptation. Characterizing the functions of DUFs is important for deciphering biological puzzles. DUFs were generally studied through forward and reverse genetics. Some novelty approaches, especially the determination of crystal structures and interaction partners of the DUFs, should attract more attention. This review described the identification of DUF genes by genome-wide and transcriptome-wide analyses, summarized the function of DUF-containing proteins, and addressed the prospects for future studies in DUFs in plants.

Keywords Protein family · Functional characterization · Pleiotropy · ABA signaling · Protein interactions

Introduction

The advance in sequencing technology has led to an exponential increase of data on genomics, transcriptomics, proteomics, and metabolomics in the postgenomic era. Although a majority of such data from various studies are deposited in data repositories and shared among researchers, for the most part, they are not fully explored (Nabi et al. 2020). For example, uncharacterized functional proteins (UPFs) and proteins containing the domains of unknown function (DUFs), which are collectively termed DUF families (Mistry et al. 2021), present 24% (4795/19,632) of the protein families currently stored in the Protein Family (Pfam) database (version 35.0) (Mistry et al. 2021). DUFs are protein domains with relatively conserved amino acid sequences and uncharacterized functions (Bateman et al. 2010). They were temporarily named with the prefix DUF followed by a number, such as DUF1 and DUF2. Once their functional properties are characterized, they will be appropriately renamed or merged with other well-known domains (Bateman et al. 2010). For example, DUF1 and DUF2, frequently presented in the bacterial signaling proteins, were renamed as GGDEF and EAL domains, respectively, after their functions were elucidated (Simm et al. 2004). As of 2021, the functions of 1132 DUF or UPF families were identified, and their identifiers in the Pfam database were changed (Mistry et al. 2021).

Initially, DUFs were often overlooked because they were generally not essential in standard mutant screens and were only discovered in a limited number of genomes (Jaroszewski et al. 2009). However, studies have shown that many of the most prevalent DUFs found in bacteria were also present in animals and plants (Goodacre et al. 2013). For instance, the DUF143-containing protein RsfA from bacteria can interact with ribosomal protein L14 to block the synthesis of protein. This interaction was verified using the homologs of RsfA/L14 in other bacteria, as well as the mitochondria of yeast and human and the chloroplasts of maize in vitro (Haüser et al. 2012). The homologs of RsfA, ATP25 from yeast, and DG238 from *Arabidopsis* were localized in mitochondria and chloroplasts, respectively, and function in the development of mitochondrial and chloroplast (Zeng et al. 2008; Wang et al. 2017). Therefore, eukaryotic DUF143 may be the evolutionary product of bacteria RsfA and are

Handling Editor: Pengguo Xia

Caixia Liu
liucaixia01@nxu.edu.cn

[1] School of Agriculture, Ningxia University, Yinchuan 750021, China

[2] College of Life Sciences, Northwest Normal University, Lanzhou 730070, China

functionally conservative in organelles. Evolutionary history analyses of prevalent DUFs, such as DUF143, could help to elucidate their functions and roles in plant adaptation.

Studies have shown that DUFs play essential roles in physiological processes, including the growth, development, and responses to biotic and abiotic stresses in plants (Table 1). This aspect has been addressed in a previous review (Lv et al. 2023). In the present review, we systemically summarized the identification of DUF genes and the roles of DUF-containing proteins as transferases, pleiotropic factors, ABA-responsive factors, and interaction partners and pointed out the potential future directions.

Identification and characterization of DUF genes

Phenotype-driven approaches

Phenotype-driven approaches begin with genetic screening for mutants of interest and then proceed with the identification of genes causing the observed phenotypes. Some genes encoding proteins with DUF domains were identified by this approach. For example, the *Young Leaf Chlorosis 1* (*YLC1*) gene encoding a protein with DUF3353 domain was isolated by map-based cloning from a rice mutant with a chlorosis phenotype and decreased level of chlorophyll and lutein contents at the early plant development stage. Then, the function of *YLC1* in chlorophyll and lutein biosynthesis was verified using complementary analysis and RNA inference technology. In addition, the expression levels of genes related to chlorophyll biosynthesis and photosynthesis were significantly altered in the *ylc1* mutant, implying a potential role of *YLC1* in regulating chlorophyll biosynthesis and photosynthesis (Zhou et al. 2013). Similarly, the *Young Seedling Stripe1* (*YSS1*), encoding a protein belonging to the DUF3727 superfamily, was identified through map-based cloning as the responsible gene for the striated leaves at the seedling stage of a rice mutant. Further investigation revealed abnormal chloroplast structure of the *yss1* mutant, which is completely rescued in *YSS1*-overexpressed transgenic plants (Zhou et al. 2017). In addition, a floral mutant with wrinkled petals and exposed stigma was screened in mungbean (*Vigna radiata*). The *Stigma Exposed 1* (*SE1*), a DUF1005-encoding gene, was identified through map-based cloning. Histological observation revealed irregular cell shape in petals and increased cell length in stigma of the *se1* mutant, indicating a role of *SE1* in cell shape and size regulation. This was further confirmed by the hetero-expression of *SE1* in *Arabidopsis* (Lin et al. 2020). With advances in mutant creation, sequencing technologies, bioinformatics, and phenotype screening, it will be much easier and faster to identify and characterize DUF genes by phenotype-driven approaches.

Knowledge- and data-driven approaches

The advent of next-generation sequencing and sophisticated computational algorithms enabled the fast identification and functional prediction of genes. Studying DUF genes through knowledge- and data-driven approaches generally involves some intelligent guesswork by searching for sequence and structural homologs and determining expression profiles. These predictions provide a starting point for further genetic and biochemical studies.

The function of a gene can be predicted by identifying homologous genes with known functions. For example, two DUF288-containing proteins, STELLO1 (STL1) and STELLO2 (STL2), were validated to regulate cellulose synthesis by aiding in the assembly of cellulose synthase complexes in *Arabidopsis* (Zhang et al. 2016). Their homologous genes in cotton, namely *GhSTL1-GhSTL4*, were identified by blasting the DUF288 domain in the upland cotton genome database. The role of GhSTLs in cellulose synthesis was validated by generating GhSTLs-silenced plants. The cellulose content and length of fibers were reduced in the *GhSTLs*-silenced plants compared to wild-type plants, indicating that *GhSTLs* are involved in cellulose synthesis in cotton (Guo et al. 2022). Besides sequence similarity, identifying structural homologies of a protein to other well-characterized proteins could also provide insights into its function and roles within the biological system. With structural genomics, the structures of hundreds of DUF proteins were examined, and two thirds of them may belong to functionally well-characterized protein families, which offers the first hypothesis about their functions (Bateman et al. 2010; Goodacre et al. 2013; Mudgal et al. 2015).

The expression profile of a gene under various conditions in the cell or in the whole organism can also give clues to gene function. DUF4228 genes responded to various abiotic stresses in *Arabidopsis* (Yang et al. 2020), soybean (Leng et al. 2021), and cotton (Lv et al. 2022) at the transcriptional level, indicating their functions in plant tolerance to abiotic stresses. This implication was validated later by several experiments. For example, the hetero-expression of *MsDUF4228* in tobacco negatively impacted the responses of seedlings to osmotic stress (Wang et al. 2018). Overexpression of *GmDUF4228-70* increased the resistance of soybeans to salt and drought stress (Leng et al. 2021). Silencing of GhDUF4228-67 reduced the salt tolerance in cotton (Lv et al. 2022). Some of the *OsDUF668* genes were triggered by Avr9/Cf-9 recognition and strongly expressed in response to rice blast disease and mechanical wounding, indicating that these Avr9/Cf-9-triggered *OsDUF668* genes may confer resistance to biotic stresses in rice (Zhong et al.

Table 1 The DUFs involved in diverse biological processes in plants

Biological process	DUFs	DUF-containing gene	Genetic approach	Specific function	Species	Reference
Synthesis and modification of cell wall	DUF231	*AXY4L, BdXyBAT1, MOAT1-MOAT4, TBL3, TBL10, TBL28-TBL35*	Reverse genetics	Polysaccharide (mannan, pectin, xylan, xyloglucan) O-acetyltransferases activity	*Arabidopsis*, *Brachypodium distachyon*, rice, tomato	(Gille et al. 2011; Xiong et al. 2013; Yuan et al. 2013; Urbanowicz et al. 2014; Liu et al. 2016; Yuan et al. 2016a, b, c; Zhong et al. 2017, 2018, 2020; Stranne et al. 2018)
	DUF266	*BC10, PdDUF266A*	Forward genetics (*BC10*), reverse genetics (*PdDUF266A*)	Glycosyltransferase activity	Poplar, rice	(Zhou et al. 2009; Yang et al. 2017)
	DUF579	*AGM1, AGM2, GXM1-GXM3,*	Reverse genetics	Glucuronoxylan methyltransferase activity	*Arabidopsis*	(Lee et al. 2012; Temple et al. 2019)
	DUF246	*PAGR*	Reverse genetics	Pectin synthesis	*Arabidopsis*	(Stonebloom et al. 2016)
	DUF288	*GhSTL1-GhSTL4, STL1, STL2,*	Reverse genetics	Cellulose synthesis	*Arabidopsis*, cotton	(Zhang et al. 2016; Guo et al. 2022)
	DUF1218	*MWL-1, MWL-2*	Reverse genetics	Lignin synthesis	*Arabidopsis*	(Mewalal et al. 2016)
Seedling growth and development	DUF239, DUF506, DUF584, DUF793, DUF966, DUF1005, DUF1645	*AtAuxRP3, AtS40.4, B1L, CiDUF1005, LOTR1, OsSGL, RXR1*	Forward genetics (*LOTR1*), reverse genetics (The others)	Root development	*Arabidopsis*, *Caragana intermedia*, rice	(Cui et al. 2016; Li et al. 2017; Shen et al. 2019; Shi et al. 2021; Kolbeck et al. 2022; Tian et al. 2022; Yang et al. 2022; Ying et al. 2022)
	DUF266, DUF538, DUF584, DUF630, DUF632, DUF966	*AtAuxRP3, AtS40.4, GL6, OsPLS3, REL2, SVB*	Forward genetics (*OsPLS3*, *REL2*), reverse genetics (*AtAuxRP3*, *AtS40.4*, *GL6*, *SVB*)	Leaf development	*Arabidopsis*, maize, rice	(Yang et al. 2016; Li et al. 2018; Li et al. 2019; Shen et al. 2019; Yu et al. 2020; Shi et al. 2021; Yu et al. 2021)
	DUF295, DUF584, DUF640, DUF4228	*MsDUF, LSH1, LSH2, AtS40.4, FDB2*	Reverse genetics	Seed germination	Alfalfa, *Arabidopsis*	(Wang et al. 2018; Lee et al. 2020; Shi et al. 2021; Gong et al. 2022)
	DUF538, DUF3353, DUF3727	*CcDUF538, YLC1, YSS1*	Forward genetics (*YLC1*, *YSS1*), reverse genetics (*CcDUF538*)	Chloroplast development	*Celosia cristata*, rice	(Zhou et al. 2013; Gholizadeh 2016; Zhou et al. 2017; Gholizadeh 2018)

Table 1 (continued)

Biological process	DUFs	DUF-containing gene	Genetic approach	Specific function	Species	Reference
Reproductive organ development	DUF1005, DUF1313	*EFL1, EFL3, SE1*	Forward genetics (*SE1*), reverse genetics (*EFL1, EFL3*)	Flower development	*Arabidopsis*, mungbean	(Lin et al. 2019, 2020)
	DUF239, DUF1644, DUF1645	*AtDUF239, OsGSL, OsSIDP301, SbSGL*	Reverse genetics	Seed development	*Arabidopsis*, rice, *sorghum*	(Cui et al. 2016; Zhang et al. 2018; Ge et al. 2022; Liu et al. 2022; Verges et al. 2022)
	DUF246, DUF647, DUF726, DUF1618, DUF1620, DUF1668	*PAGR, RFC3, RUS4, S22B, Sc, STS1*	Forward genetics (*S22B, Sc, STS1*), reverse genetics (*PAGR, RFC3, RUS4*)	Pollen fertilization	*Arabidopsis*, rice	(Qin et al. 2016; Stonebloom et al. 2016; Shen et al. 2017; Zhao et al. 2019; Chen et al. 2020b; Sakata et al. 2020; Yuan et al. 2022)
Biotic stress resistance	DUF26, DUF668, DUF740, DUF966, DUF4228	*AFP1, AFP2, AtPADRE, Gnk2, OsDUF668, OsSRDP, TaCRK3, ZmAuxRP1*	Forward genetics (*ZmAuxRP1*), reverse genetics (The others)	Response to fungi	*Arabidopsis*, *Gingko biloba*, maize, rice, wheat,	(Miyakawa et al. 2014; Ma et al. 2018; Ye et al. 2019; Zhong et al. 2019; Didelon et al. 2020; Guo et al. 2021; Jayaraman et al. 2023)
	DUF569, DUF1117	*AtDUF569, GhRDUF4D*	Reverse genetics	Response to bacteria	*Arabidopsis*, cotton,	(Nabi et al. 2020; Zhao et al. 2021)
Abiotic stress tolerance	DUF151, DUF221, DUF538, DUF584, DUF617, DUF740, DUF962, DUF966, DUF1117, DUF1618, DUF1644, DUF1645, DUF2358, DUF4228,	*AtAuxRP3, AtBBD1, AtMHP1, AtRDUF1, AtRDUF2, AtS40-1, GhDUF4228-67, GL6, GmDUF4228-70, LOC_Os01g68450, MIZ1, OsDSR2, OsSGL, OsSIDP301, OsSIDP361, OsSIDP366, OsSRDP, PtoMPO1, Sc, TaDUF966-9B, ZmOSCA2.4*	Forward genetics (*Sc*), reverse genetics (The others)	Response to drought and salt stresses	*Arabidopsis*, cotton, maize, poplar, rice, soybean, wheat,	(Kim et al. 2012; Luo et al. 2014; Wang et al. 2014; Cui et al. 2016; Guo et al. 2016; Li et al. 2016; Li et al. 2019; Shen et al. 2019; Cao et al. 2020; Kaur et al. 2020; Zhou et al. 2020; Huque et al. 2021; Leng et al. 2021; Zheng et al. 2021; Ge et al. 2022; Lv et al. 2022; Punchkhon et al. 2022; Wang et al. 2022; Zheng et al. 2022; Jayaraman et al. 2023; Luo et al. 2023)
	DUF740, DUF793	*B1L, OsSRDP*	Reverse genetics	Response to cold stress	*Arabidopsis*, rice	(Chen et al. 2019, 2020a; Jayaraman et al. 2023)

Table 1 (continued)

Biological process	DUFs	DUF-containing gene	Genetic approach	Specific function	Species	Reference
ABA signaling pathway	DUF151, DUF295, DUF584, DUF962, DUF966, DUF1117, DUF4228	*AtBBD1, AtRDUF1, AtRDUF2, AtS40-1, AtS40.4, AtMHP1, FDB2, MsDUF, OsDSR2, OsSIDP301, PtoMPO1*	Reverse genetics	ABA signaling	*Arabidopsis*, rice, *Medicago sativa*, poplar	(Kim et al. 2012; Luo et al. 2014, 2023; Wang et al. 2018, 2022; Huque et al. 2021; Shi et al. 2021; Zheng et al. 2021, 2022; Ge et al. 2022; Gong et al. 2022)

2019). Similarly, a total of 28 DUF966 genes were identified in wheat, and some of them were strongly induced by salt stresses, implying a role of *TaDUF966* in salt tolerance. This was experimentally confirmed by the fact that *TaDUF966-9B* knockout plants showed increased sensitivity to salt stress (Zhou et al. 2020). The expression of *GhRDUF4D*, a DUF1117 gene, was increased upon *Verticillium dahliae* infection in upland cotton, indicating its role against *V. dahliae* infection in cotton (Zhao et al. 2021). Consistent with the suggestion from expression data, the *V. dahliae* resistance was significantly enhanced in *GhRDUF4D* transgenic *Arabidopsis* and weakened in *GhRDUF4D* silenced cotton plants (Zhao et al. 2021). A total of 12 DUF221 domain-containing proteins (DDP) genes were identified in the genome of tomato. Among them, *SlDDP6, SlDDP11*, and *SlDDP12* were downregulated under salt stress, whereas *SlDDP1*, *SlDDP2*, *SlDDP3*, *SlDDP4*, *SlDDP7*, *SlDDP8*, and *SlDDP10* were upregulated, suggesting the involvement of these *DUF221* genes in responses to abiotic stress (Waseem et al. 2021), although further investigations are required.

Instead of monitoring a single gene, RNA-seq can investigate the expressions of tens of thousands of genes at once. By identifying differently expressed genes under various conditions, researchers were able to associate genes with physiological processes and provide insight into gene functions. For example, from the available RNA-Seq datasets of *Arabidopsis* leaves exposed to S-nitroso-L-cysteine, 231 upregulated and 206 downregulated DUF genes were identified, indicating the roles of these genes in nitro-oxidative stress responses (Nabi et al. 2020). Among these genes, the involvement of *AtDUF569* in nitro-oxidative stress responses was validated experimentally (Nabi et al. 2020). Clustering genes by expression pattern can provide an additional layer of information for gene function prediction. For example, an *Oryza sativa* stress-responsive DUF740 protein (*OsSRDP*) was characterized for its role in rice resistance to abiotic and biotic stresses (Jayaraman et al. 2023). Further investigation of the expression data showed that two genes (*LOC_Os05g09640* and *LOC_Os06g50370*) were co-expressed with *OsSRDP* under abiotic and biotic stresses, indicating that these two genes are candidate interaction partners of *OsSRDP* and may be coordinately regulated (Jayaraman et al. 2023).

DUF-containing proteins function as transferases in plant cell wall formation

The cell wall is an intricate and crucial component of plant cells (Fig. 1a). It needs to be strong to resist internal and external pressures while remaining flexible to allow cell growth (Temple et al. 2022). Polysaccharides are fundamental for the proper structure and function of cell walls

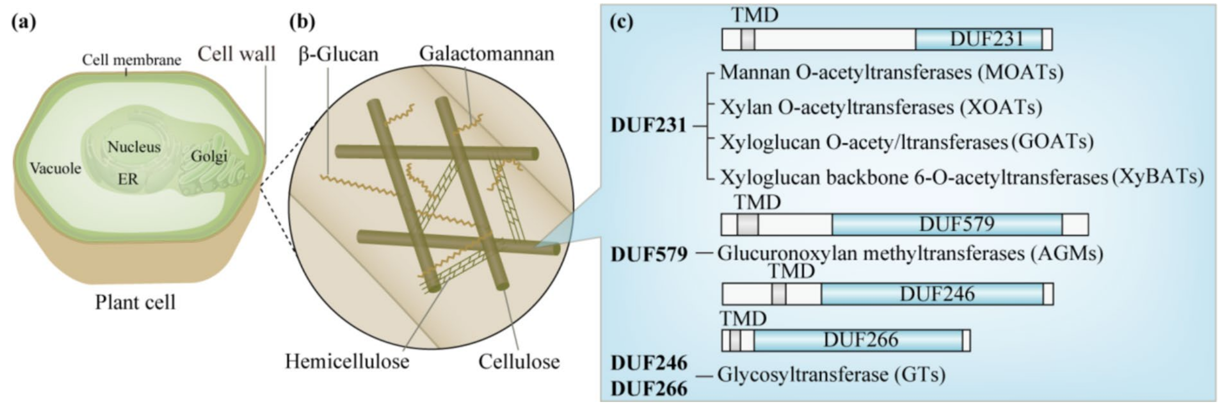

Fig. 1 The roles of DUF-containing proteins in regulating the synthesis and modification of plant cell wall components. **a** Model of the plant cell. The cell wall is shown as the closed brown circle. **b** The main polysaccharides in the plant cell wall are cellulose and hemicellulose, which are polymerized forms of β-glucan and galactomannan. **c** Various DUF-containing proteins function as transferases in the synthesis and modification of polysaccharides. The DUF domains are indicated as blue boxes, and the transmembrane domains (TMD) are indicated as the gray boxes

(Fig. 1b). Many DUF genes have been shown to participate in the synthesis and modification of polysaccharides as transferases.

Synthesis of plant cell wall components

Cellulose is the major component in both the primary and secondary cell walls. Two DUF266-containing proteins participated in cellulose biosynthesis, probably as glycosyltransferases (GTs) (Fig. 1c). The loss-of-function mutation in a DUF266 gene, *brittle culm 10 (BC10)*, reduced the cellulose content in rice (Zhou et al. 2009). Overexpression of *PdDUF266A* significantly increased cellulose content in *Populus* (Yang et al. 2017).

Pectin is a major constituent of the primary cell wall and along with hemicelluloses form a matrix into which cellulose is embedded. Rhamnogalacturonan-I (RG-I), a pectic polysaccharide, is likely present in the primary cell walls of all vascular plants (Cankar et al. 2014). The Pectic ArabinoGalactan synthesis-Related Protein (PAGR), containing a highly conserved DUF246 domain, may function in the biosynthesis of pectic RG-I arabinogalactans as glycosyltransferases in *Arabidopsis* (Stonebloom et al. 2016) (Fig. 1c).

Modification of plant cell wall components

Xylans, as a major component of hemicellulose, can be *O*-acetylated at multiple positions of its backbone residuals. Some members in the DUF231 family were proven to modify xylan as *O*-acetyltransferases (Fig. 1c). For example, the *eskimo1/trichome birefringence-like* 29 (*esk1/tbl29*) mutant had significantly decreased xylan acetyltransferase activity and xylan 2-*O*- and 3-*O*-monoacetylation (Xiong et al. 2013; Yuan et al. 2013). Eight closely related genes to *TBL29* from the DUF231 family, namely *TBL3*, *TBL28*, and *TBL30-TBL35*, were investigated for their functions in xylan acetylation. Double mutants (*tbl3 tbl31*, *tbl32 tbl33*, *tbl34 tbl35*) had a significantly lower level of xylan acetylation than the *tbl29* mutant (Yuan et al. 2016a, b, c). Moreover, triple mutants (*tbl29 tbl3 tbl31*, *tbl29 tbl32 tbl33*, *tbl29 tbl34 tbl35*) showed a much more dramatic decrease in xylan acetylation than double mutants, indicating functional redundancy of these genes in xylan acetylation (Yuan et al. 2016a, b, c). The xylan acetyltransferase activity of TBL29 and its homologs (TBL3, TBL28, and TBL30-TBL35) has been biochemically demonstrated in *Arabidopsis* (Urbanowicz et al. 2014; Zhong et al. 2017).

Besides xylan acetylation, several members in the DUF231 family were also shown to catalyze the acetylation of xyloglucan (Fig. 1c). *Arabidopsis* ALTERED XYLOGLUCAN 4 (AXY4) acetylated xyloglucan in tissues except seeds, and its paralog AXY4L acetylated xyloglucan in seeds (Gille et al. 2011). Another Xyloglucan Backbone 6-*O*-Acetyltransferase 1 (XyBAT1) can acetylate xyloglucan in rice, tomato, and *Brachypodium distachyon* (Liu et al. 2016; Zhong et al. 2020). Specifically, AXY4 and AXY4L may acetylate the side-chain of xyloglucan, while BdXyBAT1 acetylated the glucan backbone of xyloglucan (Liu et al. 2016). In addition, mannan and pectin are generally acetylated by acetyltransferase, including mannan *O*-acetyltransferase1 (MOAT1), MOAT2, MOAT3, MOAT4, and TBL10 (Stranne et al. 2018; Zhong et al. 2018). The above results indicate the conserved function of the DUF231 family in the acetylation of polysaccharides.

Besides acetylation, cell wall polysaccharides can be methylated, which may be critical for cell expansion (Levesque-Tremblay et al. 2015) and signal responses (Mizukami et al. 2016). Members in the DUF579 family may methylate polysaccharides as methyltransferases. Five proteins of the DUF579 family, GLUCURONOXYLAN METHYLTRANSFERASE1 (GXM1), GXM2, GXM3 (Lee et al. 2012), ArabinoGalactan Methyltransferase1 (AGM1), and AGM2 (Temple et al. 2019), aided the 4-*O*-methylation of xylan in *Arabidopsis* (Fig. 1c).

It would be interesting to investigate the functions of uncharacterized proteins in the DUF231 and DUF579 families in the acetylation and methylation of polysaccharides, considering the conserved functions of characterized proteins in these families.

The pleiotropy of DUF genes as regulatory factors

Some DUF genes have pleiotropic effects on the growth, development, and stress tolerance of plants. A DUF538 gene, *SMALLER TRICHOMES with VARIABLE BRANCHES* (*SVB*), was significantly highly expressed in an *Arabidopsis* mutant with defective trichomes, compared to the wild-type plant. T-DNA insertion in *SVB* resulted in smaller trichomes with variable branches (Marks et al. 2009). Besides trichome development, *SVB* also modulates plant growth with *SVB-like* (*SVBL*) through the transcriptional regulation of *GLABRA1*, a hub gene for trichome development (Yu et al. 2021). In addition, SVB may also involve in the endoplasmic reticulum (ER) stress tolerance through signal transduction as a putative phosphoinositide-binding protein (Yu and Kanehara 2020). BYPASS1-LIKE (B1L) protein from the DUF793 family is also a multifunctional regulator. B1L interacts with 14-3-3λ and TRANSTHYRETIN-LIKE (TTL) to co-regulate seedling growth and freezing tolerance by the C-REPEAT BINDING FACTOR (CBF) pathway (Chen et al. 2019, 2020a). Moreover, B1L modulates lateral root initiation via exocytic vesicular trafficking-mediated PIN-FORMED (PIN) recycling in *Arabidopsis* (Yang et al. 2022) (Fig. 2b). The Stress Induced DUF1644 Protein 301 (OsSIDP301) is a negative regulator for both salt stress and grain size in rice (Ge et al. 2022). OsSIDP301 regulates the tolerance of rice to salt stress through the abscisic acid (ABA) signaling pathway and affects grain size by influencing cell expansion in spikelet hulls (Ge et al. 2022). The *Sc* from the DUF1618 family functions as a pollen-essential factor. In *Sc-j*/*Sc-i* (*japonica* allele/*indica* allele) hybrids, the high expression of *Sc-i* in sporophytic cells suppressed the expression of *Sc-j* in pollen, leading to transmission ratio distortion. Knocking out one or two of the three *Sc-i* copies by CRISPR/Cas9 rescues *Sc-j* expression and pollen fertility (Shen et al. 2017) (Fig. 2d). In addition, *Sc* (*#51*, *Os03g0247300*) also responded to salt, drought, and cold stress in rice, suggesting a role of *Sc* in abiotic stress tolerance (Wang et al. 2014).

The importance of Stress tolerance and Grain Length (OsSGL), a DUF1645-containing protein, has been revealed by several studies. The *OsSGL* was upregulated by a wide spectrum of abiotic stress. The overexpression and heteroexpression of *OsSGL* increased the expressions of antioxidative and stress-responsive genes and enhanced the drought tolerance in rice and *Arabidopsis* (Cui et al. 2016). Moreover, the overexpression of *OsGSL* also altered an array of other traits, including increased grain length, grain weight, grain number per panicle, and extensive root systems (Fig. 2c), probably via a cytokinin signaling pathway (Cui et al. 2016; Wang et al. 2016). In addition, overexpression or silencing of *OsGSL* both reduced the starch content of grain (Liu et al. 2022). OsSGL was shown to inhibit the expression of starch-biosynthesis-related genes as a regulatory suppressor. Further, the interaction of OsSUS1 with OsSGL alleviated the transcriptional repression of OsSGL. Therefore, OsGSL functions as a regulator in controlling grain yield and quality (Liu et al. 2022).

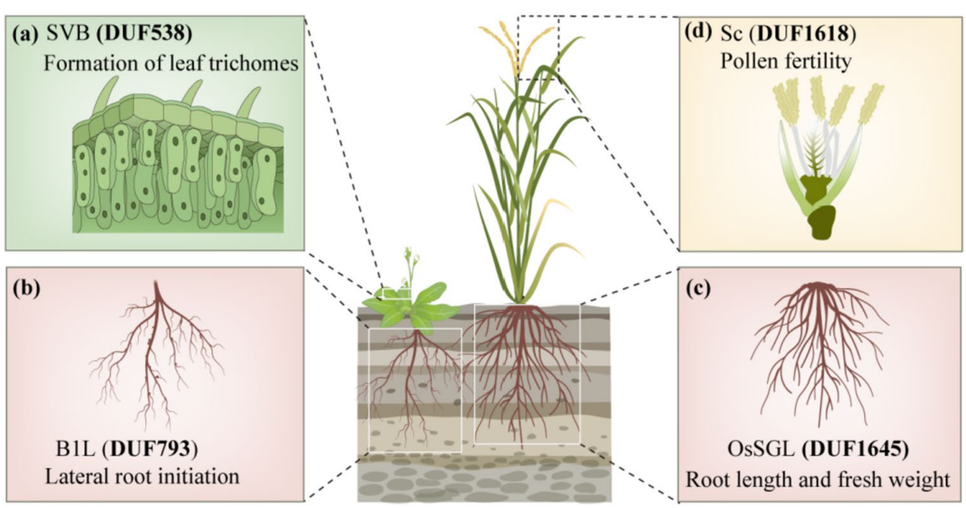

Fig. 2 The roles of several DUF-containing proteins in controlling plant growth and development. **a** SVB is involved in regulating the formation of leaf trichomes in *Arabidopsis*. **b** B1L and **c** OsSGL are involved in regulating roots development, respectively, in *Arabidopsis* and rice. **d** Sc is involved in regulating pollen fertility in rice. The SVB, B1L, OsSGL, and Sc proteins contain conserved DUF538, DUF793, DUF1645, and DUF1618 domains, respectively

The responses of plants to biotic and abiotic stresses are deployed at the expense of growth (Huot et al. 2014). Therefore, balancing the tradeoff between defense and growth is crucial. Some DUF genes represent potential roles in the trade-off by reallocating resources through phytohormonal crosstalk (Ning et al. 2017). For example, *Zea mays Auxin-Regulated Protein 1* (*ZmAuxRP1*) from the DUF966 family can increase the biosynthesis of auxin (IAA) and inhibit the biosynthesis of benzoxazinoid, which is a potent secondary metabolite that contributes to allelopathy and defense (Ye et al. 2019). Upon pathogen infection, the expression of the resistant *ZmAuxRP1* allele was transiently decreased so that plants were able to allocate more resources for benzoxazinoid biosynthesis and less for IAA biosynthesis. This led to arrested root growth but enhanced the resistance of maize to pathogens. When the pathogen attacks were averted, the expression of *ZmAuxRP1* increased back to the level that is required for normal growth (Ye et al. 2019). Similarly, overexpression of *AtAuxRP3,* another DUF966 gene, increased the endogenous IAA level in *Arabidopsis* and reduced the NaCl and osmotic stress tolerance. Transcriptomic analysis showed increased expressions of a variety of development-related genes and decreased expressions of many stress-responsive genes in the *AtAuxRP3* overexpression lines (Shen et al. 2019).

The functions of DUF genes mediated by the ABA signaling pathway

ABA is recognized as a "stress hormone" because of its critical role in mediating adaptive responses to various stresses, especially drought and salinity (Nambara and Marion-Poll 2005). Many DUF genes contribute to stress resistance through ABA-dependent pathways. For example, two RING-DUF1117 E3 ubiquitin ligase genes, *AtRDUF1* and *AtRDUF2*, were upregulated by both ABA and drought stresses. Single- and double-deletion mutants of these two genes exhibited reduced ABA-dependent drought resistance in *Arabidopsis* (Kim et al. 2012). *Arabidopsis thaliana mpo1 homolog in plants* (*AtMHP1*) encodes a protein from the DUF962 family. The *mhp1* mutant was hypersensitive to salt stress and ABA, indicating that MHP1 may be involved in the ABA-dependent salt stress response pathway in *Arabidopsis* (Zheng et al. 2021). A homolog of *AtMHP1* in Poplar, *Metabolism of PHS to Odd-numbered FA 1* (*PtoMPO1*), also encodes a protein containing the DUF962 domain. The hetero-expression of *PtoMPO1* in *Arabidopsis* greatly reduced the sensitivity of the *mhp1* mutant to salt stress and ABA (Zheng et al. 2022). The *Bifunctional nucleases in Basal Defense response 1* (*AtBBD1*) from DUF151 family and a *cytoplasmic S40 protein (AtS40-1)* from DUF584 family were upregulated by ABA and respectively enhanced the drought and salt tolerance by ABA signaling in *Arabidopsis* (Huque et al. 2021; Wang et al. 2022). The *DUF966-stress repressive gene 2* (*OsDSR2*) negatively regulated ABA-dependent salt and simulated drought stresses by down-regulating the expression of ABA- and stress-responsive genes (Luo et al. 2014, 2023). Similarly, *OsSIDP301* from the DUF1644 family negatively regulated the tolerance of rice to salt stress through the ABA signaling pathway (Ge et al. 2022).

ABA can suppress seed germination by preventing cell wall loosening of the embryo and inhibiting water uptake (Xi et al. 2010). Overexpression of *MsDUF*, a DUF4228 gene in *Medicago sativa*, induced the expression levels of ABA synthesis genes, increased the accumulation of ABA, and caused a reduction in seed germination (Wang et al. 2018). Similarly, *AtS40.4* encoding a DUF584-containing protein negatively regulated seed germination and seedling growth by the ABA signaling pathway in *Arabidopsis* (Shi et al. 2021). *F-box/DUF295 Brassiceae specific 2 (FDB2)*, belonging to the DUF295 family, conferred ABA insensitivity during seed germination and post-germination growth in *Arabidopsis* (Gong et al. 2022). In the presence of ABA, overexpression of *FDB2* increased seed germination and seedling growth, while *fdb2* mutants showed opposite phenotypes (Gong et al. 2022).

Interactions of DUF-containing proteins with other proteins

Instead of being static, proteins often change their configuration over time and frequently interact with other molecules to perform biological roles. Some DUFs perform their biological roles through protein interactions (Table 2). For instance, the DUF593 domain at the C-terminus of *Zea mays* floury1 (ZmFL1) binds to the 22-kD α-zein to facilitate its localization, which is critical for the formation of vitreous endosperm (Holding et al. 2007). The DUF724 domain at the C-terminal regions of DUF724-containing proteins (AtDuf3, AtDuf5, AtDuf7) engages with microtubules or actin filaments, which may play a role in RNA transport in the plant cell (Cao et al. 2010). The DUF827 domains in two coiled-coil proteins, weak chloroplast movement under blue light 1 (WEB1) and plastid movement impaired 2 (PMI2), provide a protein-protein interaction surface for forming the WEB1–PMI2 complex, which is essential for chloroplast movement response (Kodama et al. 2011). The DUF581 domain interacts with SUCROSE-NON-FERMENTING1-RELATED KINASE 1 (SnRK1) and serves as a bridge between SnRK1 and the DUF581-containing proteins while regulating stress signaling in *Arabidopsis* (Nietzsche et al. 2014). The DUF1620 domain and the WD40 repeat motif at the C-terminus and the N-terminus of the restoration of

Table 2 DUF-containing proteins and their interacting proteins in plants

DUF domain	DUF-containing protein	Its interacting protein	Species	Reference
DUF581	DUF581-1, DUF581-9	SnRK1	*Arabidopsis*	(Nietzsche et al. 2014)
DUF593	ZmFL1	α-zein	Maize	(Holding et al. 2007)
DUF724	AtDuf3, AtDuf5, AtDuf7	Interact with one another	*Arabidopsis*	(Cao et al. 2010)
DUF827	WEB1, PMI2	Interact with one another	*Arabidopsis*	(Kodama et al. 2011)
DUF1620	RFC3	RF5	Rice	(Qin et al. 2016)
DUF1645	OsSGL	OsSUS1	Rice	(Wang et al. 2016; Liu et al. 2022)
DUF3755	PtrDRIF1	PtrWOX13c, PtrKNAT7	*Populus trichocarpa*	(Petzold et al. 2018)

fertility complex 3 (RFC3) can interact with RF5 and Glycine-Rich Protein 162 (GRP162), respectively, to form the RFC complex, which is required for the restoration of fertility in Hong-Lian (HL) rice (Qin et al. 2016). The C-terminal DUF3755 domain of DIVARICATA AND RADIALIS INTERACTING FACTOR 1 (PtrDRIF1) is necessary and sufficient for the interactions of PtrDRIF1 with the homeodomain (HD) proteins (PtrWOX13c and PtrKNAT7). The formed heterotrimer may affect wood formation by mediating vascular cambium cell division and lignocellulose deposition in *Populus trichocarpa* (Petzold et al. 2018).

Conclusions and perspectives

Over the past 20 years, the number of DUF genes discovered in the plant kingdom has considerably increased. This number will further increase with the sequencing of new species and resequencing of other species. Exploring this treasure trove will provide a wealth of data for solving biological puzzles.

Model plants *Arabidopsis* and rice have been the focus of most studies on the biological characterization of numerous DUF genes. Since orthologs in the same DUF family often have biological functions that are identical among species (Leng et al. 2021; Lv et al. 2022), findings from rice and *Arabidopsis* may be transferable to other species. The phylogenetic analysis of well-characterized genes in model plants among species may provide insights into their biological functions in non-model plants.

Previous genome-wide identification of DUF genes showed that each DUF family normally has multiple duplicated gene members in one genome. Single mutants of any of these DUF genes exhibited no dramatic variation in the phenotype, whereas double, triple, or even quadruple mutants of these DUF genes exhibited distinct phenotypic variation compared to the wild-type plants. This suggested that genes containing the same DUF domain may be functional redundancy (Cao et al. 2010; Kim et al. 2012; Mewalal et al. 2016). On the one hand, this makes it particularly difficult to study the function of DUF genes. On the other hand, characterizing the function of one DUF gene can give clues to the functions of other genes in the same DUF family.

DUF genes can perform various molecular functions, such as transferases, ABA-responsive factors, and regulatory factors, in plant growth and stress responses. The architecture of cell wall is important in plant morphogenesis and the tolerance of plant to external stresses. For example, a DUF-246 family glycosyltransferase-like gene, *PAGR*, involves in biosynthesis of pectic arabinogalactans and affects male fertility in *Arabidopsis* (Stonebloom et al. 2016). In addition, *AhDGR2*, a DUF642 gene, affects the structure and composition of cell walls and causes salt and ABA hyper-sensibility in *Arabidopsis*. Similar to *PAGR* and *AhDGR2*, whether other cell wall formation-related DUF genes affect plant morphogenesis and stress resistance will also be interesting to explore. Studies have shown several DUF genes may function as ABA-responsive factors in abiotic stress tolerances (Zheng et al. 2021; Wang et al. 2022). *ZmAuxRP1*, a stalk rot disease resistance related gene from the DUF966 family, regulates plant growth by influencing IAA biosynthesis (Ye et al. 2019), indicating it may involve in IAA signaling pathway. Because of the importance of signaling pathway in plant growth and stress responses, whether DUF genes encoding transferases and regulatory factors (e.g., *OsSGL*, *SVB*) may act through signaling (ABA, IAA, JA, SA etc.) pathways or hormone crosstalk (Fig. 3) deserves further investigation.

Currently, at the genetic level, the functions of DUF genes in controlling plant growth, development, and stress responses were mainly investigated by phylogenetic analysis, chromosomal locations, gene structures, motif compositions, gene duplications, *cis*-elements prediction, and expression profiling analysis, as well as through phenotypic, physiological, and transcriptional differences between wild-type and transgenic plants. At the protein level, it is commonly thought that the functional units of proteins are represented by domains, which often have unique structures and roles. It is also the respective DUF domain that defines the DUF protein family. Therefore, the functions of specific DUF domains deserve more investigation. Such studies could begin with mining the core motifs of the DUF domains

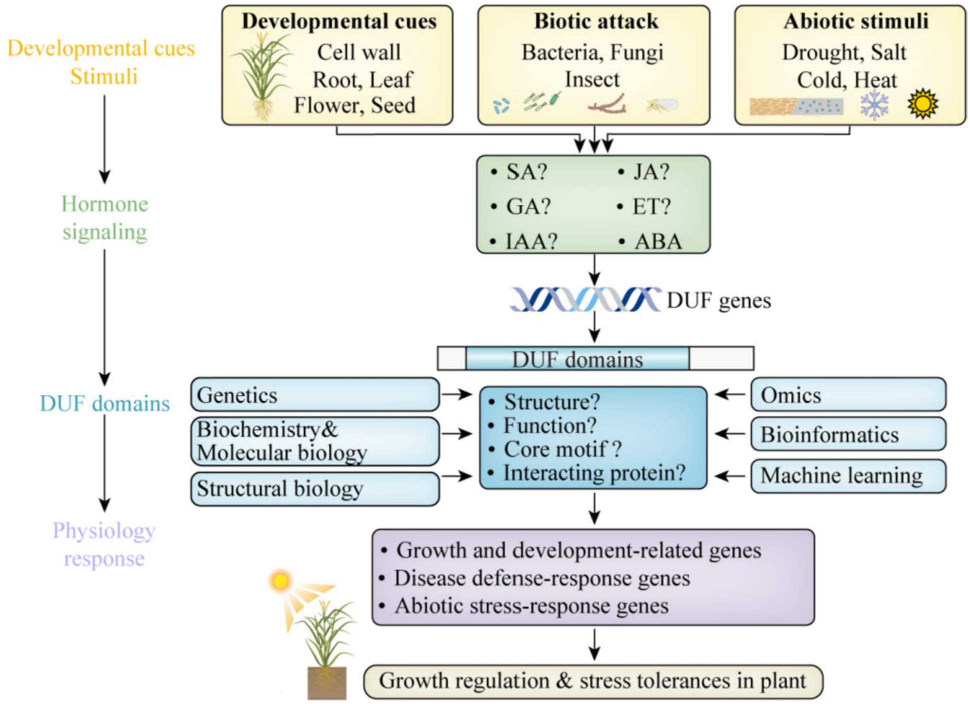

Fig. 3 Proposed model for DUF genes involved in growth regulation and stress tolerances in plants. The developmental cues or various stimuli (biotic and abiotic stresses) trigger the expression of DUF genes through an ABA-dependent pathway in plants. Other signaling molecules, besides ABA, that mediated the function of DUF genes should be elucidated in future studies. The DUF families contain at least one conserved DUF. The structures and functions of DUFs should be elucidated using various biotechnologies (genetics, biochemistry and molecular biology, structural biology, omics, and bioinformatics) and information techniques (machine learning). The DUF genes regulate the growth and stress tolerances of plants by influencing the expressions of developmental, defense, and stress-responsive genes

through amino acid substitutions and deletions (Wang et al. 2017) and then dissecting their functions using gene editing (Fig. 3). It was also suggested that individual DUF domains may have unique functions determined by their structure (Bateman et al. 2010). To the best of our knowledge, the crystal structure was solved for only a few DUF domains (PatG-DUFsp, DUF1110) in plants (Mann et al. 2014; Harada et al. 2016). In the future, the functions and structures of specific DUF domains need to be solved to fuel the characterization of proteins with these domains. In addition, most biological processes are mediated through protein interactions. Therefore, to systemically study the function of DUFs as well as their involved biological networks, diverse approaches at the protein level should be applied (Harada et al. 2016; Liu et al. 2022).

The continuous innovation in algorithms and computational frameworks will enable the integration of likely relevant multi-omics data, including phenomics, genomics, epigenomics, transcriptomics, proteomics, proteogenomics, interactomics, ionomics, metabolomics, et al., across experiments to boost the insights, broaden the horizon, and generate new hypotheses for DUFs studies. The computationally generated hypotheses can then be tested efficiently with advanced technologies in functional genomics. The high-quality experimental data will in turn improve algorithms to generate new hypotheses. Though the future is bright, there is a gap between computing and biology. To build the bridge and harness the power of biotechnologies and information techniques (e.g., machine learning) in DUFs studies, multidisciplinary research collaboration among geneticists, biologists, data analysts, computer scientists, as well as researchers of other related disciplines is required (Fig. 3). With more and more DUFs being deciphered, our understanding of the intricate mechanisms underlying the biological processes in the plant kingdom will be enhanced.

Acknowledgements We thank the reviewers for their constructive comments on this review. We thank the members in the Laboratory of Crop Genetics and Breeding, Ningxia University for their assistance in this study.

Author contribution Chengke Luo and Caixia Liu wrote the manuscript. Chengke Luo, Caixia Liu and Maryam Akhtar revised the manuscript. Weifang Min, Xiaorong Bai, and Tianli Ma helped to make tables and handle figures.

Funding This work was funded by the National Natural Science Foundation of China (Grant number 32060425), the Ningxia Natural Science

Foundation (Grant number 2022AAC03008), and the Ningxia Key R&D Project for Talent Introduction (Grant number 2021BEB04081).

References

Bateman A, Coggill P, Finn RD (2010) DUFs: families in search of function. Acta Crystallogr F: Struct Biol Cryst Commun 66:1148–1152. https://doi.org/10.1107/S1744309110001685

Cankar K, Kortstee A, Toonen MA, Wolters-Arts M, Houbein R, Mariani C, Ulvskov P, Jorgensen B, Schols HA, Visser RG, Trindade LM (2014) Pectic arabinan side chains are essential for pollen cell wall integrity during pollen development. Plant Biotechnol J 12(4):492–502. https://doi.org/10.1111/pbi.12156

Cao X, Yang KZ, Xia C, Zhang XQ, Chen LQ, Ye D (2010) Characterization of DUF724 gene family in *Arabidopsis thaliana*. Plant Mol Biol 72:61–73. https://doi.org/10.1007/s11103-009-9551-5

Cao L, Zhang P, Lu X, Wang G, Wang Z, Zhang Q, Zhang X, Wei X, Mei F, Wei L et al (2020) Systematic analysis of the maize *OSCA* genes revealing *ZmOSCA* family members involved in osmotic stress and *ZmOSCA2.4* confers enhanced drought tolerance in transgenic Arabidopsis. Int J Mol Sci 21. https://doi.org/10.3390/ijms21010351

Chen T, Chen JH, Zhang W, Yang G, Yu LJ, Li DM, Li B, Sheng HM, Zhang H, An LZ (2019) BYPASS1-LIKE, A DUF793 family protein, participates in freezing tolerance via the CBF pathway in *Arabidopsis*. Front Plant Sci 10:807. https://doi.org/10.3389/fpls.2019.00807

Chen T, Zhang W, Yang G, Chen JH, Chen BX, Sun R, Zhang H, An LZ (2020a) TRANSTHYRETIN-LIKE and BYPASS1-LIKE co-regulate growth and cold tolerance in *Arabidopsis*. BMC Plant Biol 20:1–11. https://doi.org/10.1186/s12870-020-02534-w

Chen YJ, Yang XX, Li WC, Zhao SQ (2020b) Knockdown of the DUF647 family member RUS4 impairs stamen development and pollen maturation in *Arabidopsis*. Plant Sci 301:110645. https://doi.org/10.1016/j.plantsci.2020.110645

Cui Y, Wang M, Zhou H, Li M, Huang L, Yin X, Zhao G, Lin F, Xia X, Xu G (2016) OsSGL, a novel DUF1645 domain-containing protein, confers enhanced drought tolerance in transgenic rice and *Arabidopsis*. Front Plant Sci 7:2001. https://doi.org/10.3389/fpls.2016.02001

Didelon M, Khafif M, Godiard L, Barbacci A, Raffaele S (2020) Patterns of sequence and expression diversification associate members of the PADRE gene family with response to fungal pathogens. Front Genet 11. https://doi.org/10.3389/fgene.2020.00491

Ge L, Guo H, Li X, Tang M, Guo C, Bao H, Huang L, Yi Y, Cui Y, Chen L (2022) OsSIDP301, a member of the DUF1644 family, negatively regulates salt stress and grain size in rice. Front Plant Sci 13:863233. https://doi.org/10.3389/fpls.2022.863233

Gholizadeh A (2016) DUF538 protein superfamily is predicted to be chlorophyll hydrolyzing enzymes in plants. Physiol Mol Biol Plants 22:77–85. https://doi.org/10.1007/s12298-015-0331-1

Gholizadeh A (2018) Chlorophyll binding ability of non-chloroplastic DUF538 protein superfamily in plants. Proc Natl Acad Sci India Sect B: Biol Sci 88:967–976. https://doi.org/10.1007/s40011-016-0834-8

Gille S, de Souza A, Xiong G, Benz M, Cheng K, Schultink A, Reca IB, Pauly M (2011) O-acetylation of *Arabidopsis* hemicellulose xyloglucan requires AXY4 or AXY4L, proteins with a TBL and DUF231 domain. Plant Cell 23:4041–4053. https://doi.org/10.1105/tpc.111.091728

Gong C, Yin X, Ye T, Liu X, Yu M, Dong T, Wu Y (2022) The F-Box/DUF295 Brassiceae specific 2 is involved in ABA-inhibited seed germination and seedling growth in *Arabidopsis*. Plant Sci 323:111369. https://doi.org/10.1016/j.plantsci.2022.111369

Goodacre NF, Gerloff DL, Uetz P (2013) Protein domains of unknown function are essential in bacteria. mBio 5:e713-e744. https://doi.org/10.1128/mBio.00744-13

Guo C, Luo C, Guo L, Li M, Guo X, Zhang Y, Wang L, Chen L (2016) *OsSIDP366*, a DUF1644 gene, positively regulates responses to drought and salt stresses in rice. J Integr Plant Biol 58:492–502. https://doi.org/10.1111/jipb.12376

Guo F, Wu T, Shen F, Xu G, Qi H, Zhang Z (2021) The cysteine-rich receptor-like kinase TaCRK3 contributes to defense against *Rhizoctonia cerealis* in wheat. J Exp Bot 72:6904–6919. https://doi.org/10.1093/jxb/erab328

Guo Y, Chen F, Luo J, Qiao M, Zeng W, Li J, Xu W (2022) The DUF288 domain containing proteins GhSTLs participate in cotton fiber cellulose synthesis and impact on fiber elongation. Plant Sci 316:111168. https://doi.org/10.1016/j.plantsci.2021.111168

Harada K, Yamashita E, Inoue K, Yamaguchi K, Fujiwara T, Nakagawa A, Kawasaki T, Kojima C (2016) Plant-specific DUF1110 protein from *Oryza sativa*: expression, purification and crystallization. Acta Crystallogr F Struct Biol Commun 72:480–484. https://doi.org/10.1107/S2053230X16007573

Haüser R, Pech M, Kijek J, Yamamoto H, Titz B, Naeve F, Tovchigrechko A, Yamamoto K, Szaflarski W, Takeuchi N et al (2012) RsfA (YbeB) proteins are conserved ribosomal silencing factors. PLoS Genet 8:e1002815. https://doi.org/10.1371/journal.pgen.1002815

Holding DR, Otegui MS, Li B, Meeley RB, Dam T, Hunter BG, Jung R, Larkins BA (2007) The maize floury1 gene encodes a novel endoplasmic reticulum protein involved in zein protein body formation. Plant Cell 19:2569–2582. https://doi.org/10.1105/tpc.107.053538

Huot B, Yao J, Montgomery BL, He SY (2014) Growth-defense tradeoffs in plants: a balancing act to optimize fitness. Mol Plant 7(8):1267–1287. https://doi.org/10.1093/mp/ssu049

Huque A, So W, Noh M, You MK, Shin JS (2021) Overexpression of AtBBD1, *Arabidopsis* bifunctional nuclease, confers drought tolerance by enhancing the expression of regulatory genes in ABA-mediated drought stress signaling. Int J Mol Sci 22. https://doi.org/10.3390/ijms22062936

Jaroszewski L, Li Z, Krishna SS, Bakolitsa C, Wooley J, Deacon AM, Wilson IA, Godzik A (2009) Exploration of uncharted regions of the protein universe. PLoS Biol 7:e1000205. https://doi.org/10.1371/journal.pbio.1000205

Jayaraman K, Sevanthi AM, Raman KV, Jiwani G, Solanke AU, Mandal PK, Mohapatra T (2023) Overexpression of a DUF740 family gene (LOC_Os04g59420) imparts enhanced climate resilience through multiple stress tolerance in rice. Front Plant Sci 13:947312. https://doi.org/10.3389/fpls.2022.947312

Kaur V, Yadav SK, Wankhede DP, Pulivendula P, Kumar A, Chinnusamy V (2020) Cloning and characterization of a gene encoding MIZ1, a domain of unknown function protein and its role in salt and drought stress in rice. Protoplasma 257:475–487. https://doi.org/10.1007/s00709-019-01452-5

Kim SJ, Ryu MY, Kim WT (2012) Suppression of *Arabidopsis* RING-DUF1117 E3 ubiquitin ligases, AtRDUF1 and AtRDUF2, reduces tolerance to ABA-mediated drought stress. Biochem Biophys Res Commun 420:141–147. https://doi.org/10.1016/j.bbrc.2012.02.131

Kodama Y, Suetsugu N, Wada M (2011) Novel protein-protein interaction family proteins involved in chloroplast movement response. Plant Signal Behav 6:483–490. https://doi.org/10.4161/psb.6.4.14784

Kolbeck A, Marhavy P, De Bellis D, Li B, Kamiya T, Fujiwara T, Kalmbach L, Geldner N (2022) CASP microdomain formation requires cross cell wall stabilization of domains and non-cell autonomous action of LOTR1. eLife 11. https://doi.org/10.7554/eLife.69602

Lee C, Teng Q, Zhong R, Yuan Y, Haghighat M, Ye ZH (2012) Three *Arabidopsis* DUF579 domain-containing GXM proteins are methyltransferases catalyzing 4-o-methylation of glucuronic acid on xylan. Plant Cell Physiol 53(11):1934–1949. https://doi.org/10.1093/pcp/pcs138

Lee M, Dong X, Song H, Yang JY, Kim S, Hur Y (2020) Molecular characterization of *Arabidopsis thaliana LSH1* and *LSH2* genes. Genes Genomics 42:1151–1162. https://doi.org/10.1007/s13258-020-00985-x

Leng ZX, Liu Y, Chen ZY, Guo J, Chen J, Zhou YB, Chen M, Ma YZ, Xu ZS, Cui XY (2021) Genome-wide analysis of the DUF4228 family in soybean and functional identification of GmDUF4228-70 in response to drought and salt stresses. Front Plant Sci 12:628299. https://doi.org/10.3389/fpls.2021.628299

Levesque-Tremblay G, Pelloux J, Braybrook SA, Müller K (2015) Tuning of pectin methylesterification: consequences for cell wall biomechanics and development. Planta 242(4):791–811. https://doi.org/10.1007/s00425-015-2358-5

Li M, Gou L, Gou C, Wang L, Chen L (2016) Over-expression of a DUF1644 protein gene, *SIDP361*, enhances tolerance to salt stress in transgenic rice. J Plant Biol 59:62–73. https://doi.org/10.1007/s12374-016-0180-7

Li B, Kamiya T, Kalmbach L, Yamagami M, Yamaguchi K, Shigenobu S, Sawa S, Danku JM, Salt DE, Geldner N et al (2017) Role of LOTR1 in nutrient transport through organization of spatial distribution of root endodermal barriers. Curr Biol 27:758–765. https://doi.org/10.1016/j.cub.2017.01.030

Li K, Chen Y, Luo Y, Huang F, Zhao C, Cheng F, Xiang X, Pan G (2018) A 22-bp deletion in *OsPLS3* gene encoding a DUF266-containing protein is implicated in rice leaf senescence. Plant Mol Biol 98:19–32. https://doi.org/10.1007/s11103-018-0758-1

Li L, Du Y, He C, Dietrich CR, Li J, Ma X, Wang R, Liu Q, Liu S, Wang G et al (2019) Maize *glossy6* is involved in cuticular wax deposition and drought tolerance. J Exp Bot 70:3089–3099. https://doi.org/10.1093/jxb/erz131

Lin K, Zhao H, Gan S, Li G (2019) *Arabidopsis* ELF4-like proteins EFL1 and EFL3 influence flowering time. Gene 700:131–138. https://doi.org/10.1016/j.gene.2019.03.047

Lin Y, Laosatit K, Chen J, Yuan X, Wu R, Amkul K, Chen X, Somta P (2020) Mapping and functional characterization of stigma exposed 1, a DUF1005 gene controlling petal and stigma cells in mungbean (*Vigna radiata*). Front Plant Sci 11:575922. https://doi.org/10.3389/fpls.2020.575922

Liu L, Hsia MM, Dama M, Vogel J, Pauly M (2016) A xyloglucan backbone 6-O-acetyltransferase from *Brachypodium distachyon* modulates xyloglucan xylosylation. Mol Plant 9(4):615–617. https://doi.org/10.1016/j.molp.2015.11.004

Liu Z, Jiang S, Jiang L, Li W, Tang Y, He W, Wang M, Xing J, Cui Y, Lin Q et al (2022) Transcription factor OsSGL is a regulator of starch synthesis and grain quality in rice. J Exp Bot 73:3417–3430. https://doi.org/10.1093/jxb/erac068

Luo C, Guo C, Wang W, Wang L, Chen L (2014) Overexpression of a new stress-repressive gene *OsDSR2* encoding a protein with a DUF966 domain increases salt and simulated drought stress sensitivities and reduces ABA sensitivity in rice. Plant Cell Rep 33:323–336. https://doi.org/10.1007/s00299-013-1532-0

Luo C, Akhtar M, Min W, Alam Y, Ma T, Shi Y, She Y, Lu X (2023) The suppressed expression of a stress responsive gene 'OsDSR2' enhances rice tolerance in drought and salt stress. J Plant Physiol 282:153927. https://doi.org/10.1016/j.jplph.2023.153927

Lv P, Wan J, Zhang C, Hina A, Al Amin GM, Begum N, Zhao T (2023) Unraveling the diverse roles of neglected genes containing domains of unknown function (DUFs): progress and perspective. Int J Mol Sci 24(4):4187. https://doi.org/10.3390/ijms24044187

Lv X, Wei F, Lian B, Yin G, Sun M, Chen P, An L, Jian H, Wang H, Fu X et al (2022) A comprehensive analysis of the DUF4228 gene family in *Gossypium* reveals the role of GhDUF4228–67 in salt tolerance. Int J Mol Sci 23. https://doi.org/10.3390/ijms232113542

Ma LS, Wang L, Trippel C, Mendoza-Mendoza A, Ullmann S, Moretti M, Carsten A, Kahnt J, Reissmann S, Zechmann B et al (2018) The Ustilago maydis repetitive effector Rsp3 blocks the antifungal activity of mannose-binding maize proteins. Nat Commun 9:1711. https://doi.org/10.1038/s41467-018-04149-0

Mann G, Koehnke J, Bent AF, Graham R, Houssen W, Jaspars M, Schwarz-Linek U, Naismith JH (2014) The structure of the cyanobactin domain of unknown function from PatG in the patellamide gene cluster. Acta Crystallogr F Struct Biol Commun 70:1597–1603. https://doi.org/10.1107/S2053230X1402425X

Marks MD, Wenger JP, Gilding E, Jilk R, Dixon RA (2009) Transcriptome analysis of *Arabidopsis* wild-type and gl3-sst sim trichomes identifies four additional genes required for trichome development. Mol Plant 2:803–822. https://doi.org/10.1093/mp/ssp037

Mewalal R, Mizrachi E, Coetzee B, Mansfield SD, Myburg AA (2016) The *Arabidopsis* domain of unknown function 1218 (DUF1218) containing proteins, MODIFYING WALL LIGNIN-1 and 2 (At1g31720/MWL-1 and At4g19370/MWL-2) function redundantly to alter secondary cell wall lignin content. PLoS ONE 11:e150254. https://doi.org/10.1371/journal.pone.0150254

Mistry J, Chuguransky S, Williams L, Qureshi M, Salazar GA, Sonnhammer E, Tosatto S, Paladin L, Raj S, Richardson LJ et al (2021) Pfam: the protein families database in 2021. Nucleic Acids Res 49:D412–D419. https://doi.org/10.1093/nar/gkaa913

Miyakawa T, Hatano K, Miyauchi Y, Suwa Y, Sawano Y, Tanokura M (2014) A secreted protein with plant-specific cysteine-rich motif functions as a mannose-binding lectin that exhibits antifungal activity. Plant Physiol 166:766–778. https://doi.org/10.1104/pp.114.242636

Mizukami AG, Inatsugi R, Jiao J, Kotake T, Kuwata K, Ootani K, Okuda S, Sankaranarayanan S, Sato Y, Maruyama D, Iwai H, Garénaux E, Sato C, Kitajima K, Tsumuraya Y, Mori H, Yamaguchi J, Itami K, Sasaki N, Higashiyama T (2016) The AMOR arabinogalactan sugar chain induces pollen-tube competency to respond to ovular guidance. Curr Biol 26(8):1091–1097. https://doi.org/10.1016/j.cub.2016.02.040

Mudgal R, Sandhya S, Chandra N, Srinivasan N (2015) De-DUFing the DUFs: Deciphering distant evolutionary relationships of Domains of Unknown Function using sensitive homology detection methods. Biol Direct 10:38. https://doi.org/10.1186/s13062-015-0069-2

Nabi R, Tayade R, Imran QM, Hussain A, Shahid M, Yun BW (2020) Functional insight of nitric-oxide induced DUF genes in *Arabidopsis thaliana*. Front Plant Sci 11:1041. https://doi.org/10.3389/fpls.2020.01041

Nambara E, Marion-Poll A (2005) Abscisic acid biosynthesis and catabolism. Annu Rev Plant Biol 56:165–185. https://doi.org/10.1146/annurev.arplant.56.032604.144046

Nietzsche M, Schiessl I, Bornke F (2014) The complex becomes more complex: protein-protein interactions of SnRK1 with DUF581 family proteins provide a framework for cell- and stimulus type-specific SnRK1 signaling in plants. Front Plant Sci 5:54. https://doi.org/10.3389/fpls.2014.00054

Ning Y, Liu W, Wang GL (2017) Balancing immunity and yield in crop plants. Trends Plant Sci 22(12):1069–1079. https://doi.org/10.1016/j.tplants.2017.09.010

Petzold HE, Chanda B, Zhao C, Rigoulot SB, Beers EP, Brunner AM (2018) DIVARICATA AND RADIALIS INTERACTING FACTOR (DRIF) also interacts with WOX and KNOX proteins associated with wood formation in *Populus trichocarpa*. Plant J 93:1076–1087. https://doi.org/10.1111/tpj.13831

Punchkhon C, Chutimanukul P, Chokwiwatkul R, Saputro TB, Grennan AK, Diego N, Spichal L, Chadchawan S (2022) Role of

LOC_Os01g68450, containing DUF2358, in salt tolerance is mediated via adaptation of absorbed light energy dissipation. Plants 11. https://doi.org/10.3390/plants11091233 (Basel)

Qin X, Huang Q, Xiao H, Zhang Q, Ni C, Xu Y, Liu G, Yang D, Zhu Y, Hu J (2016) The rice DUF1620-containing and WD40-like repeat protein is required for the assembly of the restoration of fertility complex. New Phytol 210:934–945. https://doi.org/10.1111/nph.13824

Sakata M, Takano-Kai N, Miyazaki Y, Kanamori H, Wu J, Matsumoto T, Doi K, Yasui H, Yoshimura A, Yamagata Y (2020) Domain Unknown Function DUF1668-containing genes in multiple lineages are responsible for F(1) pollen sterility in rice. Front Plant Sci 11:632420. https://doi.org/10.3389/fpls.2020.632420

Shen R, Wang L, Liu X, Wu J, Jin W, Zhao X, Xie X, Zhu Q, Tang H, Li Q, Chen L, Liu Y (2017) Genomic structural variation-mediated allelic suppression causes hybrid male sterility in rice. Nat Commun 8(1):1310. https://doi.org/10.1038/s41467-017-01400-y

Shen L, Zhong T, Wang L, Zhang Q, Jin H, Xu M, Ye J (2019) Characterization the role of a UFC homolog, *AtAuxRP3*, in the regulation of *Arabidopsis* seedling growth and stress response. J Plant Physiol 240:152990. https://doi.org/10.1016/j.jplph.2019.152990

Shi XP, Ren JJ, Qi HD, Lin Y, Wang YY, Li DF, Kong LJ, Wang XL (2021) Plant-specific AtS40.4 acts as a negative regulator in abscisic acid signaling during seed germination and seedling growth in Arabidopsis. Front. Plant Sci 12:622201. https://doi.org/10.3389/fpls.2021.622201

Simm R, Morr M, Kader A, Nimtz M, Romling U (2004) GGDEF and EAL domains inversely regulate cyclic di-GMP levels and transition from sessility to motility. Mol Microbiol 53:1123–1134. https://doi.org/10.1111/j.1365-2958.2004.04206.x

Stonebloom S, Ebert B, Xiong G, Pattathil S, Birdseye D, Lao J, Pauly M, Hahn MG, Heazlewood JL, Scheller HV (2016) A DUF-246 family glycosyltransferase-like gene affects male fertility and the biosynthesis of pectic arabinogalactans. BMC Plant Biol 16:90. https://doi.org/10.1186/s12870-016-0780-x

Stranne M, Ren Y, Fimognari L, Birdseye D, Yan J, Bardor M, Mollet JC, Komatsu T, Kikuchi J, Scheller HV et al (2018) TBL10 is required for O-acetylation of pectic rhamnogalacturonan-I in *Arabidopsis thaliana*. Plant J 96:772–785. https://doi.org/10.1111/tpj.14067

Temple H, Mortimer JC, Tryfona T, Yu X, Lopez-Hernandez F, Sorieul M, Anders N, Dupree P (2019) Two members of the DUF579 family are responsible for arabinogalactan methylation in *Arabidopsis*. Plant Direct 3:e117. https://doi.org/10.1002/pld3.117

Temple H, Phyo P, Yang W, Lyczakowski JJ, Echevarría-Poza A, Yakunin I, Parra-Rojas JP, Terrett OM, Saez-Aguayo S, Dupree R, Orellana A, Hong M, Dupree P (2022) Golgi-localized putative S-adenosyl methionine transporters required for plant cell wall polysaccharide methylation. Nat Plants 8(6):656–669. https://doi.org/10.1038/s41477-022-01156-4

Tian X, Niu X, Chang Z, Zhang X, Wang R, Yang Q, Li G (2022) DUF1005 family identification, evolution analysis in plants, and primary root elongation regulation of *CiDUF1005* from *Caragana intermedia*. Front Genet 13:807293. https://doi.org/10.3389/fgene.2022.807293

Urbanowicz BR, Pena MJ, Moniz HA, Moremen KW, York WS (2014) Two *Arabidopsis* proteins synthesize acetylated xylan in vitro. Plant J 80:197–206. https://doi.org/10.1111/tpj.12643

Verges V, Bellenger L, Pichon O, Giglioli-Guivarc'H N, Dutilleul C, Ducos E (2022) The *Arabidopsis* DUF239 gene family encodes Neprosin-like proteins that are widely expressed in seed endosperm. Plant Genome e20290. https://doi.org/10.1002/tpg2.20290

Wang L, Shen R, Chen LT, Liu YG (2014) Characterization of a novel DUF1618 gene family in rice. J Integr Plant Biol 56:151–158. https://doi.org/10.1111/jipb.12130

Wang M, Lu X, Xu G, Yin X, Cui Y, Huang L, Rocha PSCF, Xia X (2016) *OsSGL*, a novel pleiotropic stress-related gene enhances grain length and yield in rice. Sci Rep 6:38157. https://doi.org/10.1038/srep38157

Wang M, Jiang L, Da Q, Liu J, Feng D, Wang J, Wang H, Jin H (2017) Domain of Unknown Function 143 is required for the functioning of PEP-associated protein DG238 in the chloroplast. J Plant Biol 60:604–611. https://doi.org/10.1007/s12374-017-0294-6

Wang Y, Zhang Z, Liu H, An Y, Han B, Wu Y, Chang L, Hu T, Yang P (2018) Overexpression of an alfalfa (*Medicago sativa*) gene, *MsDUF*, negatively impacted seed germination and response to osmotic stress in transgenic tobacco. Plant Cell Tissue Organ Cult (PCTOC) 132:525–534. https://doi.org/10.1007/s11240-017-1348-7

Wang Y, Li Y, Tian H, Wang W, Wang X, Hussain S, Yuan Y, Lin R, Hussain H, Wang T et al (2022) AtS40-1, a group I DUF584 protein positively regulates ABA response and salt tolerance in *Arabidopsis*. Gene 846:146846. https://doi.org/10.1016/j.gene.2022.146846

Waseem M, Aslam MM, Shaheen I (2021) The DUF221 domain-containing (DDP) genes identification and expression analysis in tomato under abiotic and phytohormone stress. GM Crops Food 12:586–599. https://doi.org/10.1080/21645698.2021.1962207

Xi W, Liu C, Hou X, Yu H (2010) MOTHER OF FT AND TFL1 regulates seed germination through a negative feedback loop modulating ABA signaling in *Arabidopsis*. Plant Cell 22(6):1733–1748. https://doi.org/10.1105/tpc.109.073072

Xiong G, Cheng K, Pauly M (2013) Xylan O-acetylation impacts xylem development and enzymatic recalcitrance as indicated by the *Arabidopsis* mutant tbl29. Mol Plant 6:1373–1375. https://doi.org/10.1093/mp/sst014

Yang SQ, Li WQ, Miao H, Gan PF, Qiao L, Chang YL, Shi CH, Chen KM (2016) *REL2*, a gene encoding an unknown function protein which contains DUF630 and DUF632 domains controls leaf rolling in rice. Rice 9:37. https://doi.org/10.1186/s12284-016-0105-6. (**N Y**)

Yang Y, Yoo CG, Guo HB, Rottmann W, Winkeler KA, Collins CM, Gunter LE, Jawdy SS, Yang X, Guo H et al (2017) Overexpression of a Domain of Unknown Function 266-containing protein results in high cellulose content, reduced recalcitrance, and enhanced plant growth in the bioenergy crop *Populus*. Biotechnol Biofuels 10:74. https://doi.org/10.1186/s13068-017-0760-x

Yang Q, Niu X, Tian X, Zhang X, Cong J, Wang R, Zhang G, Li G (2020) Comprehensive genomic analysis of the DUF4228 gene family in land plants and expression profiling of ATDUF4228 under abiotic stresses. BMC Genomics 21:12. https://doi.org/10.1186/s12864-019-6389-3

Yang G, Chen B, Chen T, Chen J, Lin X, Yue X, An L, Zhang H (2022) BYPASS1-LIKE regulates lateral root initiation via exocytic vesicular trafficking-mediated PIN recycling in *Arabidopsis*. J Integr Plant Biol 64(5):965–978. https://doi.org/10.1111/jipb.13243

Ye J, Zhong T, Zhang D, Ma C, Wang L, Yao L, Zhang Q, Zhu M, Xu M (2019) The auxin-regulated protein ZmAuxRP1 coordinates the balance between root growth and stalk rot disease resistance in maize. Mol Plant 12:360–373. https://doi.org/10.1016/j.molp.2018.10.005

Ying S, Blancaflor EB, Liao F, Scheible WR (2022) A phosphorus-limitation induced, functionally conserved DUF506 protein is a repressor of root hair elongation in plants. New Phytol 233:1153–1171. https://doi.org/10.1111/nph.17862

Yu CY, Kanehara K (2020) The unfolded protein response modulates a phosphoinositide-binding protein through the IRE1-bZIP60 pathway. Plant Physiol 183:221–235. https://doi.org/10.1104/pp.19.01488

Yu CY, Sharma O, Nguyen P, Hartono CD, Kanehara K (2021) A pair of DUF538 domain-containing proteins modulates plant growth and trichome development through the transcriptional regulation of GLABRA1 in *Arabidopsis thaliana*. Plant J 108:992–1004. https://doi.org/10.1111/tpj.15487

Yuan Y, Teng Q, Zhong R, Ye ZH (2013) The *Arabidopsis* DUF231 domain-containing protein ESK1 mediates 2-O- and 3-O-acetylation of xylosyl residues in xylan. Plant Cell Physiol 54:1186–1199. https://doi.org/10.1093/pcp/pct070

Yuan Y, Teng Q, Zhong R, Haghighat M, Richardson EA, Ye ZH (2016a) Mutations of *Arabidopsis* TBL32 and TBL33 affect xylan acetylation and secondary wall deposition. PLoS ONE 11:e146460. https://doi.org/10.1371/journal.pone.0146460

Yuan Y, Teng Q, Zhong R, Ye ZH (2016b) Roles of *Arabidopsis* TBL34 and TBL35 in xylan acetylation and plant growth. Plant Sci 243:120–130. https://doi.org/10.1016/j.plantsci.2015.12.007

Yuan Y, Teng Q, Zhong R, Ye ZH (2016c) TBL3 and TBL31, Two *Arabidopsis* DUF231 domain proteins, are required for 3-O-monoacetylation of xylan. Plant Cell Physiol 57:35–45. https://doi.org/10.1093/pcp/pcv172

Yuan G, Zou T, He Z, Xiao Q, Li G, Liu S, Xiong P, Chen H, Peng K, Zhang X et al (2022) SWOLLEN TAPETUM AND STERILITY 1 is required for tapetum degeneration and pollen wall formation in rice. Plant Physiol 190:352–370. https://doi.org/10.1093/plphys/kiac307

Zeng X, Barros MH, Shulman T, Tzagoloff A (2008) ATP25, a new nuclear gene of Saccharomyces cerevisiae required for expression and assembly of the Atp9p subunit of mitochondrial ATPase. Mol Biol Cell 19:1366–1377. https://doi.org/10.1091/mbc.e07-08-0746

Zhang Y, Nikolovski N, Sorieul M, Vellosillo T, McFarlane HE, Dupree R, Kesten C, Schneider R, Driemeier C, Lathe R, Lampugnani E, Yu X, Ivakov A, Doblin MS, Mortimer JC, Brown SP, Persson S, Dupree P (2016) Golgi-localized STELLO proteins regulate the assembly and trafficking of cellulose synthase complexes in *Arabidopsis*. Nat Commun 7:11656. https://doi.org/10.1038/ncomms11656

Zhang B, Zhang X, Xu G, Li M, Cui Y, Yin X, Yu Y, Xia X, Wang M (2018) Expression of sorghum gene *SbSGL* enhances grain length and weight in rice. Mol Breeding 38:40. https://doi.org/10.1007/s11032-018-0799-9

Zhao SQ, Li WC, Zhang Y, Tidy AC, Wilson ZA (2019) Knockdown of *Arabidopsis* ROOT UVB SENSITIVE4 disrupts anther dehiscence by suppressing secondary thickening in the endothecium. Plant Cell Physiol 60:2293–2306. https://doi.org/10.1093/pcp/pcz127

Zhao YP, Shen JL, Li WJ, Wu N, Chen C, Hou YX (2021) Evolutionary and characteristic analysis of RING-DUF1117 E3 ubiquitin ligase genes in *Gossypium* discerning the role of *GhRDUF4D* in *Verticillium dahliae* resistance. Biomolecules 11. https://doi.org/10.3390/biom11081145

Zheng M, Peng T, Yang T, Yan J, Yang K, Meng D, Hsu YF (2021) *Arabidopsis* MHP1, a homologue of yeast Mpo1, is involved in ABA signaling. Plant Sci 304:110732. https://doi.org/10.1016/j.plantsci.2020.110732

Zheng M, Wang Q, Lei S, Yang D, Liu Y, Feng D, Huang X, Yang K, Qian J, Hsu YF (2022) PtoMPO1, a negative mediator, functions in poplar drought tolerance. Plant Physiol Biochem 190:156–163. https://doi.org/10.1016/j.plaphy.2022.09.004

Zhong R, Cui D, Ye ZH (2017) Regiospecific acetylation of xylan is mediated by a group of DUF231-containing O-acetyltransferases. Plant Cell Physiol 58:2126–2138. https://doi.org/10.1093/pcp/pcx147

Zhong R, Cui D, Ye ZH (2018) Members of the DUF231 family are O-acetyltransferases catalyzing 2-O- and 3-O-acetylation of mannan. Plant Cell Physiol 59:2339–2349. https://doi.org/10.1093/pcp/pcy159

Zhong R, Cui D, Phillips DR, Richardson EA, Ye ZH (2020) A group of O-acetyltransferases catalyze xyloglucan backbone acetylation and can alter xyloglucan xylosylation pattern and plant growth when expressed in *Arabidopsis*. Plant Cell Physiol 61:1064–1079. https://doi.org/10.1093/pcp/pcaa031

Zhong H, Zhang H, Guo R, Wang Q, Huang X, Liao J, Li Y, Huang Y, Wang Z (2019) Characterization and functional divergence of a novel DUF668 gene family in rice based on comprehensive expression patterns. Genes 10. https://doi.org/10.3390/genes10120980 (Basel)

Zhou Y, Li S, Qian Q, Zeng D, Zhang M, Guo L, Liu X, Zhang B, Deng L, Liu X et al (2009) BC10, a DUF266-containing and Golgi-located type II membrane protein, is required or cell-wall biosynthesis in rice (*Oryza sativa* L.). Plant J 57:446–462. https://doi.org/10.1111/j.1365-313X.2008.03703.x

Zhou K, Ren Y, Lv J, Wang Y, Liu F, Zhou F, Zhao S, Chen S, Peng C, Zhang X et al (2013) Young Leaf Chlorosis 1, a chloroplast-localized gene required for chlorophyll and lutein accumulation during early leaf development in rice. Planta 237:279–292. https://doi.org/10.1007/s00425-012-1756-1

Zhou K, Ren Y, Zhou F, Wang Y, Zhang L, Lyu J, Wang Y, Zhao S, Ma W, Zhang H et al (2017) Young Seedling Stripe1 encodes a chloroplast nucleoid-associated protein required for chloroplast development in rice seedlings. Planta 245:45–60. https://doi.org/10.1007/s00425-016-2590-7

Zhou X, Zhu X, Shao W, Song J, Jiang W, He Y, Yin J, Ma D, Qiao Y (2020) Genome-wide mining of wheat DUF966 gene family provides new insights into salt stress responses. Front Plant Sci 11:569838. https://doi.org/10.3389/fpls.2020.569838