Molecular Bioinformatics and Expression Analysis of the COBRA Gene Family in *Huperzia serrata* Postprint

Authors: Huang Yumei, Teng Jianbei, Tu Dongping, Liang Liuguan

Date: 2023-12-24T00:00:00+00:00

Abstract

To elucidate the molecular bioinformatics characteristics and tissue expression patterns of COBRA gene family members in Huperzia serrata, thereby providing a foundation for further research on COBRA genes. Based on full-length transcriptome data of Huperzia serrata, bioinformatics techniques were utilized to analyze the physicochemical properties, domains, conserved motifs, cis-acting elements, and gene expression levels of these family members (HsCOBRAs). The results demonstrated that: (1) A total of 24 HsCOBRAs family members were identified from the full-length transcriptome of Huperzia serrata, among which there were 9 acidic proteins, 11 stable proteins, 5 hydrophobic proteins, 7 proteins with transmembrane structures, and 3 proteins with signal peptides. (2) Subcellular localization occurred in the cell wall, chloroplast, nucleus, and cell membrane. (3) Structural analysis revealed that HsCOBRAs possess 7 types of domains and 6 types of conserved motifs, with some members containing a highly conserved CCVS structure. (4) HsCOBRAs harbor 46 types of cis-acting elements, including CAAT-box and TATA-box. (5) HsCOBRA2 exhibited the highest expression levels in leaf, spore, stem, and gemma. The molecular bioinformatics and expression characteristics of the COBRA gene family in Huperzia serrata can provide a theoretical basis for further investigation of HsCOBRAs and validation of their biological functions.

Full Text

Molecular Bioinformatics and Expression Analysis of the COBRA Gene Family in *Huperzia serrata*

HUANG Yumei, TENG Jianbei, TU Dongping*, LIANG Liuguan
College of Pharmacy, Guangxi University of Traditional Chinese Medicine, Nanning 530200, China

Abstract

This study elucidates the molecular bioinformatics characteristics and tissue expression patterns of COBRA gene family members in *Huperzia serrata* to provide a foundation for further research on COBRA genes. Based on full-length transcriptome data of *H. serrata*, we analyzed the physicochemical properties, structural domains, conserved motifs, cis-acting elements, and gene expression levels of family members (HsCOBRAs) using bioinformatics approaches. The results revealed: (1) A total of 24 HsCOBRA family members were identified in the full-length transcriptome of *H. serrata*, including 9 acidic proteins, 11 stable proteins, 5 hydrophobic proteins, 7 proteins with transmembrane structures, and 3 proteins with signal peptides. (2) Subcellular localization occurred in the cell wall, chloroplast, nucleus, and cell membrane. (3) Structural analysis showed that HsCOBRAs possess 7 domains and 6 conserved motifs, with some members containing a highly conserved CCVS structure. (4) HsCOBRAs contain 46 types of cis-acting elements such as CAAT-box and TATA-box. (5) HsCOBRA2 exhibited the highest expression levels in leaves, spores, stems, and gemma. The molecular bioinformatics and expression characteristics of the COBRA gene family in *H. serrata* can provide a theoretical basis for further research and functional validation of HsCOBRAs.

Keywords: *Huperzia serrata*, COBRA gene family, bioinformatics, full-length transcriptome, expression analysis

---

The COBRA gene was first discovered in *Arabidopsis* root cells (Borner et al., 2003) and represents a class of cellulose synthase genes widely present in higher plants and algae (Ren Angyan et al., 2021). COBRA encodes glycosyl-phosphatidyl inositol (GPI)-anchored proteins containing an N-terminal signal peptide, a carbohydrate-binding domain, a cysteine-rich CCVS domain, and a C-terminal ω-site for GPI protein anchoring (Roudier et al., 2002). To date, COBRA genes have been identified in various plants including cotton, Chinese cabbage, tea, tomato, and maize. Studies demonstrate that COBRA plays a central role in primary and secondary cell wall cellulose biosynthesis by regulating cellulose microfibril synthesis and deposition (Li et al., 2019; Ren Angyan et al., 2021), and participates extensively in root, stem, leaf, flower, and fruit development, thereby influencing plant biomass (Yuan Zhicheng et al., 2020; Zaheer et al., 2022; Li et al., 2022). For instance, overexpression of cotton *GhCOBL9A* increases biomass (Niu et al., 2018), while tea *CsCOBRA* genes affect tea yield by regulating leaf cell wall composition and mechanical strength (Ai Antao et al., 2021).

Huperzia serrata, the source plant of the traditional Chinese medicine Qian-CengTa, possesses therapeutic effects including blood circulation promotion, hemostasis, wound healing, swelling reduction, detoxification, and fever relief, and is used to treat traumatic injuries, hemoptysis, hemorrhoids, lung abscesses, and burns (Huang Yumei et al., 2023). Its primary component, huperzine A,

demonstrates efficacy against Alzheimer's disease (AD) with minimal peripheral side effects (Ji Shengguo, 2007). With global AD cases projected to reach 139 million by 2050 (Jia et al., 2018) and Chinese elderly cases expected to hit 30.03 million (Wang Yingquan et al., 2019), demand for *H. serrata* continues to grow, indicating strong market prospects. Although artificial cultivation has progressed, large-scale commercial production remains unachieved due to slow growth and difficult propagation, leaving supply far below demand. Developing high-yield, efficient modern production systems for *H. serrata* has become urgent, and manipulating cellulose synthase genes such as COBRA may offer a novel approach to increasing biomass production.

This study utilized *H. serrata* as experimental material to identify COBRA genes from full-length transcriptome sequencing data. Through bioinformatics analysis of gene characteristics and tissue expression patterns, we investigated: (1) the bioinformatics features of the *H. serrata* COBRA gene family; and (2) the potential roles of COBRA genes in *H. serrata* growth and development, laying the groundwork for future functional identification and regulation of *H. serrata* COBRA genes to provide a basis for molecular breeding and yield improvement.

Materials and Methods

1.1 Plant Material

Huperzia serrata plants were collected from Daming Mountain, Nanning, Guangxi in August 2019 and identified as *H. serrata* (Huperziaceae) by Senior Engineer Feng Shixin from Guangxi Medicinal Botanical Garden. Full-length transcriptome data were provided by the PacBio Sequel platform of BGI Genomics Co., Ltd., Shenzhen.

1.2.1 Identification of *H. serrata* COBRA Gene Family Members

Based on full-length transcriptome sequencing data and annotation results, we used *Arabidopsis* COBRA gene family DNA sequences as seed sequences to BLAST search against the *H. serrata* whole-genome sequence database with a screening threshold of $E \leq 10^{-5}$ to identify COBRA gene family members in *H. serrata*.

1.2.2 Analysis of Physicochemical Properties, Signal Peptides, Subcellular Localization, and Domains

Original nucleotide sequences of *H. serrata* COBRA gene family members were processed through ORF Finder (http://www.ncbi.nlm.nih.gov/orffinder) to obtain potential protein-coding fragments. ExPAsy (ProtParam tool) was used to predict physicochemical properties of COBRA-encoded protein sequences, SignalP-6.0 (https://services.healthtech.dtu.dk/service.php?SignalP) for signal peptide prediction, Plant-mPLoc (http://www.csbio.sjtu.edu.cn/bioinf/plant-multi/) for subcellular localization prediction, Batch CD-Search (https://www.ncbi.nlm.nih.gov/Structure/bwr

for domain prediction, and TBtools1.112 for visualization.

1.2.3 Analysis of Secondary Structure, Tertiary Structure, and Cis-Acting Elements

Secondary and tertiary structures were predicted using Prabi (https://npsa-prabi.ibcp.fr/cgi-bin/npsa_{automat}.pl?page=npsa_{sopma}.html) and phyre2 (http://www.sbg.bio.ic.ac.uk/phyre2/html/page.cgi?id=index), respectively. The 2,000 bp upstream sequence of *H. serrata* COBRA family genes was used as the promoter region for cis-acting element prediction using PlantCARE (https://bioinformatics.psb.ugent.be/webtools/plantcare/html/). Functionally annotated cis-acting elements were selected and visualized using TBtools1.112.

1.2.4 Analysis of Conserved Motifs, Expression Levels, and Phylogeny

Conserved motifs were predicted using MEME Suite 5.5.1 (https://meme-suite.org/meme/tools/meme) and visualized with TBtools1.112. Expression levels of COBRA gene family members in *H. serrata* leaves, spores, stems, and gemma were analyzed by calculating mean fragments per kilobase of exon model per million mapped fragments (FPKM). Expression heatmaps were generated and refined using TBtools1.112. Protein sequences were multiply aligned using MEGA11, and a neighbor-joining phylogenetic tree was constructed (bootstrap = 1,000) and beautified using iTOL (http://itol.embl.de).

Results

2.1 Physicochemical Properties of *H. serrata* COBRA Gene Family

A total of 24 COBRA gene family nucleotide sequences were screened from *H. serrata* full-length transcriptome data and designated HsCOBRA1–HsCOBRA24, with base sequences available through [Figure 1: see original paper]A. Amino acid sequences were obtained based on *H. serrata* transcriptome sequencing and ORFfinder ([Figure 1: see original paper]B). Physicochemical property prediction of HsCOBRA-encoded proteins using ExPAsy online tools yielded the results shown in . Among the 24 HsCOBRA protein sequences, amino acid numbers ranged from 73 (HsCOBRA19) to 663 (HsCOBRA14), with molecular weights between 8,494.91 (HsCOBRA15) and 73,084.23 (HsCOBRA14) Da (average: 26,012.36 Da). Nine proteins (HsCOBRA1, HsCOBRA5, HsCOBRA6, HsCOBRA7, HsCOBRA9, HsCOBRA14, HsCOBRA16, HsCOBRA17, HsCOBRA18) had isoelectric points between 4.53–6.97 (acidic proteins), while the remaining 15 had isoelectric points between 8.17–10.13 (basic proteins). Eleven proteins (HsCOBRA1, HsCOBRA2, HsCOBRA3, HsCOBRA4, HsCOBRA6, HsCOBRA7, HsCOBRA11, HsCOBRA12, HsCOBRA13, HsCOBRA14, HsCOBRA23) were stable, while the other 13 were unstable. The aliphatic index of HsCOBRA family members ranged from 61.70 to 92.37.

2.2 Subcellular Localization and Signal Peptide Analysis of *H. serrata* COBRA Gene Family

Transmembrane structure prediction revealed that 7 proteins (HsCOBRA2, HsCOBRA3, HsCOBRA4, HsCOBRA9, HsCOBRA11, HsCOBRA12, HsCOBRA14) each possessed one transmembrane structure, while the remainder had none. Signal peptide prediction identified typical GPI-anchored protein N-terminal signal peptides in HsCOBRA2, HsCOBRA4, and HsCOBRA14, with signal peptide indices of 0.9758, 0.9758, and 0.9998, respectively. Cleavage sites were predicted between amino acid residues 41–42, 41–42, and 28–29 at the N-terminus; all other proteins had signal peptide indices below 0.5. Subcellular localization analysis showed widespread distribution of HsCOBRA family members: 1 (HsCOBRA9) localized to the cell wall, 4 (HsCOBRA6, HsCOBRA7, HsCOBRA10, HsCOBRA23) to chloroplasts, 10 (HsCOBRA1, HsCOBRA5, HsCOBRA7, HsCOBRA16, HsCOBRA17, HsCOBRA19, HsCOBRA20, HsCOBRA21, HsCOBRA22, HsCOBRA24) to the nucleus, and the remaining 10 to the cell membrane. Notably, HsCOBRA7 localized to both chloroplasts and nucleus, while HsCOBRA10 localized to both chloroplasts and peroxisomes ().

2.3 Domain Analysis of *H. serrata* COBRA Gene Family

Protein domain analysis revealed seven domain types in the HsCOBRA family ([Figure 2: see original paper]): COBRA, COBRA super family, PI31_{Prot}N, PI31_{Prot}C, PI31_{Prot}N super family, RING_{Ubox} super family, and TilS super family. COBRA and COBRA super family domains, both annotated as glycosylphosphatidylinositol, were distributed across four protein sequences each. Except for HsCOBRA1, which contained two domains, 13 proteins (HsCOBRA2, HsCOBRA3, HsCOBRA4, HsCOBRA5, HsCOBRA6, HsCOBRA7, HsCOBRA8, HsCOBRA12, HsCOBRA13, HsCOBRA14, HsCOBRA16, HsCOBRA17, HsCOBRA18) possessed only one domain, while the remaining 10 proteins had no identifiable domains.

2.4 Secondary Structure Analysis of *H. serrata* COBRA Gene Family

Secondary structure prediction using Prabi software indicated that HsCOBRA proteins comprise random coils, α-helices, extended strands, and β-turns, with average proportions of 46.02%, 29.34%, 18.95%, and 5.70%, respectively (). In six proteins (HsCOBRA10, HsCOBRA19, HsCOBRA20, HsCOBRA21, HsCOBRA22, HsCOBRA23), α-helices predominated (38.27%–58.25%), followed by random coils. In all other HsCOBRA proteins, random coils were most abundant (38.67%–64.29%). β-turns showed the lowest proportion across all proteins (1.55%–14.81%).

2.5 Tertiary Structure Analysis of *H. serrata* COBRA Gene Family

Tertiary structure prediction using phyre2 revealed 19 distinct structural models among the 24 proteins (, [Figure 3: see original paper]). HsCOBRA7 showed the highest amino acid coverage (81%). HsCOBRA13, HsCOBRA16, HsCOBRA19, and HsCOBRA24 possessed hydrolase structures, while HsCOBRA1, HsCOBRA6, and HsCOBRA7 contained hydrolase inhibitor structures, and HsCOBRA2, HsCOBRA3, HsCOBRA4, and HsCOBRA14 contained carbohydrate structures. HsCOBRA6, HsCOBRA7, and HsCOBRA17 shared 99.9% sequence homology, with HsCOBRA6 and HsCOBRA7 showing similar tertiary structures that may indicate analogous biological functions. HsCOBRA1 and HsCOBRA5 exhibited 100% sequence homology.

2.6 Cis-Acting Element Analysis of *H. serrata* COBRA Gene Family

Cis-acting element analysis identified 46 functionally annotated elements in HsCOBRAs ([Figure 4: see original paper]). Twenty-three elements were involved in light responsiveness, including MRE, Box 4, G-box, ACE, circadian, AE-box, GT1-motif, Sp1, 3-AF1 binding site, CAG-motif, GA-motif, Gap-box, TCT-motif, chs-CMA2a, chs-CMA2c, GATA-motif, LAMP-element, Box II, I-box, AT1-motif, ATCT-motif, GTGGC-motif, and TCCC-motif. Some proteins contained elements associated with meristem expression (CAT-box), endosperm expression (GCN4_{motif}), and hormone responses including abscisic acid (ABRE), auxin (TGA-element), gibberellin (P-box, TATC-box), and salicylic acid (TCA-element). Additionally, environmental response elements for anaerobic conditions (ARE), hypoxia (GC-motif), drought (MBS), and low temperature (LTR) were present. CAAT-box elements were most abundant (307 copies), present in all proteins as common cis-acting elements in promoter and enhancer regions, followed by TATA-box (206 copies) as core promoter elements and MBS (72 copies) as MYB-binding sites for drought induction. HsCOBRA18 contained the most cis-acting elements (78), followed by HsCOBRA14 (75) and HsCOBRA24 (70), while HsCOBRA7 had the fewest (10). Some elements were unique to single proteins, such as TCCC-motif (HsCOBRA5), 3-AF1 binding site (HsCOBRA12), GTGGC-motif (HsCOBRA17), ACA-motif and Box II-like sequence (HsCOBRA18), Box II and chs-CMA2c (HsCOBRA19), ATCT-motif, AT1-motif, and A-box (HsCOBRA20), GA-motif and ACE (HsCOBRA23), and TC-rich repeats (HsCOBRA24).

2.7 Conserved Motif Analysis of *H. serrata* COBRA Gene Family

Conserved motif analysis revealed six motif types in the HsCOBRA family, designated motif 1–motif 6 ([Figure 5: see original paper], [Figure 6: see original paper]). Only 13 proteins contained motifs, with three proteins (HsCOBRA2, HsCOBRA3, HsCOBRA4) possessing all six motif types. Seven proteins (HsCOBRA1, HsCOBRA5, HsCOBRA7, HsCOBRA9, HsCOBRA10, HsCOBRA12, HsCOBRA24) contained only one motif type. Motif 6 showed the broadest distribution, followed by motif 1, motif 3, and motif 4. Motif 3 contained the highly

conserved CCVS structure present in seven proteins. Specific motif patterns were shared among certain groups: HsCOBRA1, HsCOBRA6, and HsCOBRA7 (motif 4); HsCOBRA2, HsCOBRA3, and HsCOBRA4 (motif 1–motif 6); and HsCOBRA9 and HsCOBRA10 (motif 6).

2.8 Tissue Expression Pattern Analysis of *H. serrata* COBRA Gene Family

Expression patterns were analyzed using transcriptome data from *H. serrata* leaves, spores, stems, and gemma ([Figure 7: see original paper]). HsCOBRA2 showed the highest expression across all four tissues, while HsCOBRA4 ranked second highest in leaves, spores, and stems. HsCOBRA7, HsCOBRA15, and HsCOBRA20 were barely expressed in any tissue. Additionally, three genes (HsCOBRA6, HsCOBRA12, HsCOBRA23) showed negligible expression in leaves; one gene (HsCOBRA21) in spores; four genes (HsCOBRA6, HsCOBRA12, HsCOBRA13, HsCOBRA23) in stems; and four genes (HsCOBRA4, HsCOBRA6, HsCOBRA21, HsCOBRA23) in gemma. Overall COBRA gene family expression in *H. serrata* followed the pattern: leaves > spores > stems > gemma. HsCOBRA2 and HsCOBRA16 showed high expression in gemma and spores, suggesting potential involvement in reproductive development.

2.9 Phylogenetic Analysis of *H. serrata* COBRA Gene Family

To understand the evolutionary relationship between *H. serrata* and fern COBRA gene families, we constructed a neighbor-joining phylogenetic tree of COBRA family members from *H. serrata* and *Selaginella moellendorffii* ([Figure 8: see original paper]). Sequence alignment identified CCVS structures in HsCOBRA2, HsCOBRA3, HsCOBRA4, HsCOBRA11, HsCOBRA13, and HsCOBRA14 at positions 238/241, 268/271, 238/241, 20/23, 109/112, and 431/434, respectively, indicating high protein conservation. The phylogenetic tree showed homologous relationships between HsCOBRA1 and HsCOBRA7, HsCOBRA2 and HsCOBRA4, and HsCOBRA16 and HsCOBRA17, demonstrating evolutionary relationships between *H. serrata* and *S. moellendorffii* COBRA gene family members.

Discussion and Conclusion

This study identified 24 COBRA gene family members in *H. serrata*, predominantly basic proteins whose secondary structures mainly comprise random coils, α-helices, and extended strands. Most members contain light-responsive cis-acting elements, similar to COBRA gene families in Chinese cabbage (Lian Ruiting et al., 2022), flax (Qi Yanni et al., 2019), sorghum (Yuan Zhicheng et al., 2020), tomato (Cao et al., 2012), maize (Pan Yitian, Huang Min, 2022), and tea (Ai Antao et al., 2021), suggesting conserved family characteristics across species. However, subcellular localization patterns of HsCOBRAs differed from those reported in maize, flax, and millet (Ren Angyan et al., 2021),

indicating potential functional divergence during evolution. Six HsCOBRA proteins (HsCOBRA15, HsCOBRA19, HsCOBRA20, HsCOBRA21, HsCOBRA22, HsCOBRA23) lacked identifiable domains and conserved motifs and showed negligible expression across all four tissues, suggesting they may not belong to the *H. serrata* COBRA family.

Huperzine A from *H. serrata* demonstrates superior efficacy and safety in AD treatment (Chen Sisi et al., 2021), making the plant's growth, reproduction, and sustainable production critically important. COBRA genes primarily function in plant cell expansion (Lian Ruiting et al., 2022) and influence biomass and yield by regulating plant growth and development. HsCOBRAs contain cis-acting elements involved in meristem and endosperm expression, and 16 HsCOBRAs showed varying expression levels across leaves, spores, stems, and gemma, with highest expression in leaves and spores, indicating involvement in tissue development (Yang Lanfeng et al., 2023). Previous studies revealed abundant cellulose in large leaves, old leaves, and sporangial walls of *H. serrata*, suggesting HsCOBRAs may regulate cellulose and cell wall growth, mechanical strength, and developmental processes to enhance biomass. As spores represent a major reproductive organ, HsCOBRAs may influence sporangial wall development, affecting spore maturation and germination. The gene family also contains cis-acting elements responding to abiotic stresses including drought, hypoxia, low temperature, and salinity. ACA elements respond to cold, heat, salt, and UV stress (Chen Taofei et al., 2023); ABRE elements bind specifically to BpbZIP1 to enhance salt tolerance by scavenging reactive oxygen species (Guo Yiping et al., 2020); and LTR retrotransposons respond to salt, ABA, H_2O_2, and drought stress (Zhou Binhan et al., 2023). The abundant presence of ACA, ABRE, and LTR retrotransposons in HsCOBRA promoter regions suggests the *H. serrata* COBRA gene family plays important roles in abiotic stress responses.

The *H. serrata* COBRA gene family exhibits bioinformatics features similar to those in other species while potentially participating in growth, development, reproduction, and stress responses. Characterizing the molecular bioinformatics and tissue expression patterns of this gene family will help elucidate molecular regulatory mechanisms and promote growth, providing new avenues for biomass increase and a theoretical foundation for future molecular breeding and functional validation studies.

References

1. AI AT, LI YL, ZHANG WY, et al., 2021. Genome-wide identification and expressions of COBRA family in *Camellia sinensis*. Acta Tea Sin, 62(3): 95-104.

2. CAO Y, TANG XF, LIU YS, et al., 2012. Cloning, expression pattern and bioinformation analyses of COBRA gene in Tomato(*Solanum lycopersicum*). Bull Bot Res, 32(3): 304-310.

3. CHEN SS, ZHANG MH, WANG JX, et al., 2021. Original plant and re-

search progress of the medicinal plant *Huperzia javanica*. Guihaia, 41(11): 1794-1809.

4. CHEN TF, HUANG M, LIU R, et al., 2023. Bioinformatics analysis of maize ACA gene family. Mol Plant Breed, 21(13): 4188-4201.
5. GUO YP, SHI JJ, ZHOU MQ, et al., 2020. Drought and salt tolerance analysis of BpbZIP1 gene in birch and ABRE element binding identification. For Res, 33(5): 68-76.
6. HUANG YM, XIE CL, ZHOU LM, et al., 2023. Literature research on the original and efficacy of *Huperzia Serrata* (Thunb.) Trev. Asia-Pacific Trad Med, 19(3): 183-189.
7. JI SG, 2007. Research on molecular biology of Huperziaceae medicinal plants. Shanghai: Fudan University: 1-140.
8. LIAN RT, REN H, ZHAI W, et al., 2022. Identification and expression analysis under low temperature stress of COBRA gene family in *Brassica rapa* L. J Yunnan Agric Univ(Nat Sci), 37(6): 1031-1039.
9. PAN YT, HUANG M, 2022. Genome-wide identification and expression pattern analysis of COBRA family members in maize. Gansu Agric Sci Technol, 53(2): 67-72.
10. QI YN, LI WJ, WANG LM, et al., 2019. Identification and bioinformatics analysis of the COBRA gene family in Flax. Gansu Agric Sci Technol, (9): 33-38.
11. REN AY, CHEN HY, WANG GP, et al., 2021. Genome-wide identification and expression files of COBRA family in foxtail millet. Mol Plant Breed: 1-10[2023-08-05]. http://kns.cnki.net/kcms/detail/46.1068.S.20211015.0104.002.html.
12. WANG YQ, LIANG JH, JIA RX, et al., 2019. Alzheimer disease in China (2015-2050) estimated using the 1% population sampling survey in 2015. Chin J Alzheimer's Dis and Relat Disord, 2(1): 289-298.
13. YUAN ZC, KE YS, WU FJ, et al., 2020. Genome wide identification and expression of COBRA gene family in *Sorghum bicolor*. J Henan Agric Sci, 49(10): 33-41.
14. YANG LF, ZHU XD, ZHOU BH, et al., 2023. Genome-wide identification and abiotic stress response analysis of TCP transcription factor gene family in tartary buckwheat (*Fagopyrum tataricum*). Guihaia: 1-15[2023-10-23]. http://kns.cnki.net/kcms/detail/45.1134.q.20230824.1928.013.html.
15. ZHOU BH, YANG Z, WANG SP, et al., 2023. Screening of active LTR retrotransposons in wheat(*Triticum aestivum* L.) seedlings and analysis of their responses to abiotic stresses. Acta Agron Sin, 49(4): 966-977.
16. BORNER GH, LILLEY KS, STEVENS TJ, et al., 2003. Identification

of glycosylphosphatidylinositol-anchored proteins in *Arabidopsis*. A proteomic and genomic analysis. Plant Physiol, 132(2): 568-577.

17. JIA J, WEI C, CHEN S, et al., 2018. The cost of Alzheimer's disease in China and re-estimation of costs worldwide. Alzheimerors Demen, 14(4): 483-491.
18. KASIRAJAN L, HOANG NV, FURTADO A, et al., 2018. Transcriptome analysis highlights key differentially expressed genes involved in cellulose and lignin biosynthesis of sugarcane genotypes varying in fiber content. Sci Rep, 8(1): 11612.
19. LI P, LIU Y, TAN W, et al., 2019. Brittle culm 1 encodes a COBRA-like protein involved in secondary cell wall cellulose biosynthesis in Sorghum. Plant Cell Physiol, 60(4): 788-801.
20. LI Z, ZHOU T, SUN P, et al., 2022. COBL9 and COBL7 synergistically regulate root hair tip growth via controlling apical cellulose deposition. Biochem Biophys Res Commun, 596: 6-13.
21. NIU E, FANG S, SHANG X, et al., 2018. Ectopic expression of GhCOBL9A, a cotton glycosyl-phosphatidyl inositol-anchored protein encoding gene, promotes cell elongation, thickening and increased plant biomass in transgenic *Arabidopsis*. Mol Genet Genomics, 293(5): 1191-1204.
22. ROUDIER F, SCHINDELMAN G, DESALLE R, et al., 2002. The COBRA family of putative GPI-anchored proteins in *Arabidopsis*. A new fellowship in expansion. Plant Physiol, 130(2): 918-930.
23. ZAHEER M, REHMAN SU, KHAN SH, et al., 2022. Characterization of new COBRA like (COBL) genes in wheat (*Triticum aestivum*) and their expression analysis under drought stress. Mol Biol Rep, 49(2): 1379-1387.

Note: Figure translations are in progress. See original paper for figures.