

Scientia Agropecuaria

Web page: http://revistas.unitru.edu.pe/index.php/scientiaagrop

Facultad de Ciencias Agropecuarias

Universidad Nacional de Trujillo

RESEARCH ARTICLE



A computational analysis revealed BES1 transcription factor and β -amylase as crosstalk elements in Upland cotton species (*Gossypium* sp.)

Ronald Oswaldo Villamar-Torres¹, Camilo Alexander Mestanza Uquillas¹, Héctor David Chévez-Vera², Marcos Raúl Heredia-Pinos¹, Christopher Robin Viot³, Seyed Mehdi Jazayeri⁴ *

¹ Facultad de Ciencias Pecuarias y Biológicas. Universidad Técnica Estatal de Quevedo. Quevedo, Los Ríos, Ecuador.

² Instituto Superior Tecnológico "Ciudad De Valencia". Puebloviejo, Los Ríos, Ecuador.

³ CIRAD, UMR AGAP Institut, INRAE, Institut Agro, Univ Montpellier, Montpellier, France.

⁴ Université d'Avignon et des Pays du Vaucluse, France.

* Corresponding author: seyed-mehdi.jazayeri@univ-avignon.fr (S. M. Jazayeri).

Received: 20 April 2024. Accepted: 19 August 2024. Published: 2 September 2024.

Abstract

Cotton is a resilient and multipurpose crop, meeting major of the world's textile needs while also yielding byproducts like edible oil and animal feed. Starch plays a crucial role in cotton fabric production. It enhances fabric strength by forming a protective film around cotton fibers, making them more resistant to wear and tear. BES1 (brassinosteroid insensitive 1) is a key regulator in brassinosteroid signaling. It controls thousands of target genes involved in development processes. Interestingly, two β -amylase proteins (BAM7 and BAM8) are part of the BES1 family, despite their primary function as β -amylases. β -Amylase (BAM) and BES1 are two gene families with functional and regulatory roles in controlling shoot growth and development by mediating brassinosteroid effects. They share similar domains and participate in various biological processes, tolerance and responses to stresses like salt and drought. In a computational analysis comparing *Arabidopsis* and *Gossypium* species, BAM and BES1 were characterized. BES1 genes were grouped into four clusters based on the comparison of the two species. Two clusters corresponded to BAM7 and BAM8, while the other two clusters were associated with BES1. The conserved nucleotide domain sequence is GCTGGATGG. Short tandem repeats include TG and TTG, which can serve as molecular markers. BES1 is specifically linked to cellulose and fiber production and holds promise as a candidate for plant selection and breeding in *Gosspipum* (cotton).

Keywords: BAM; bioinformatics; BZR1; crosstalk; gene expression; transcription factor.

DOI: https://doi.org/10.17268/sci.agropecu.2024.033

Cite this article:

Villamar-Torres, R. O., Mestanza Uquillas, C. A., Chévez-Vera, H. D., Heredia-Pinos, M. R., Viot, C. R., & Jazayeri, S. M. (2024). A computational analysis revealed BES1 transcription factor and β-amylase as crosstalk elements in Upland cotton species (*Gossypium* sp.). *Scientia Agropecuaria*, *15*(3), 449-460.

1. Introducción

Brassinosteroids (BRs) are a group of polyhydroxylated steroidal phytohormones that play crucial roles in plant development, growth, and productivity. These hormones regulate processes such as cell division, elongation, and differentiation throughout the entire plant life cycle (Manghwar et al., 2022). Interactions and crosstalk occur between BRs and other phytohormones, impacting signaling, transcription, regulation, and function (Yang et al., 2023). Notably, BRs and gibberellins exhibit crosstalk by employing downstream genes like DELLA for gibberellins and BES1 (BRI1-EMS-SUPPRESSOR 1) for BRs (Yasir & Wasaya, 2021). BES1 gain-offunction plants demonstrate involvement in jasmonic acid (JA)-related defense responses, particularly against necrotrophic pathogens. This BR-JA crosstalk has also been observed in tomatoes as a defense mechanism against insect herbivory (Campos et al., 2009).

BES1, also known as BZR1/BES1, belongs to a family of transcription factors (TFs) that play a pivotal role in mediating brassinosteroid (BR) responses in various plant biological processes (**Shi et al., 2022**). BZR1/BES1 undergoes a two-way process involving phosphorylation and dephosphorylation to regulate its activity. When phosphorylated by BIN1, BZR1 becomes a target for the 20S proteasome, leading to its degradation. Conversely, dephosphorylation of BES1 results in its nuclear accumulation, where it exerts its transcriptional functions. BES1 participates in the BES1/TPL/HDA19 repressor complex, which modulates the influence of abscisic acid (ABA) on seed germination and counteracts the inhibitory effect of BR on ABA (Ryu et al., 2014). Through BR signaling, BES1 mediates the CrRLK (receptor-like kinase protein) family, enhancing cell elongation. BES1 cooperates with other TFs to fine-tune BRregulated mechanisms. A group of WRKY TFs (including WRKY46, WRKY54, and WRKY70) interacts with BES1 in response to drought stress, growth, and development (Shi et al., 2022). Responsive to Desiccation 26 (RD26), an NAC family TF, interacts with BES1 and antagonizes BR activity, resulting in altered drought responses in plants. While BES1 is implicated in various defense and growth mechanisms, its precise roles and functions remain partially undisclosed. In summary, BES1/BZR1 stands at the crossroads of BR signaling, orchestrating intricate processes in plant development and adaptation BES1 is found in the plant kingdom with some copies in each species (https://planttfdb.gaolab.org/) but not necessarily well-distributed among the species. Among plants containing BES1, Gossypium raimondii has the highest abundance, followed by Salix purpurea and Populus trichocarpa, with 42, 38, and 32 BES1 members, respectively. However, non-fibrous plants, such as algae, do not contain BES1. Wu et al. (2016) showed that there are 14 orthologous gene pairs shared by Chinese cabbage and Arabidopsis. It functions in connecting plant signaling to BRs and controlling gene expression by a transcriptional network by which plants respond and tolerate stresses and direct growth and development (Nolan et al., 2017). Its binding motif is a G box (CCACGTGG) that interacts with BIM1 a basic helix-loop-helix protein to synergistically bind to E box (CANNTG) sequences. These sequences are seen in many BR-induced promoters (Yin et al., 2005).

β-amylase proteins (BAM) are the enzymes responsible to starch breakdown. They belong to plant kingdom and usually are found in the nucleus rather than the chloroplast. They catalyse the hydrolysis reaction that results in maltose. They have BRASSINAZOLE RESISTANT1 (BZR1)-type/BES1 DNA binding motifs. There are two β-amylase known as BAM7 and BAM8 that resemble BZR1/BES1. They possess a G box that bind a cisregulatory element in BES1/BZR1 by which they activate gene expression. Mutants of BAM7 and BAM8 (bam7 and bam8) cause leaf growth and development changes (**Reinhold et al., 2011**). However, the other BAM members also show similarity to BZR1/BES1.

BES1-BAM complex plays a role in plant response to stress and growth and development. The double-

faced complex regulates many genes that respond to BR through crosstalk signaling. There may exist metabolic signals via G box by binding a ligand in their BAM domain or regulatory cascades by their BES1 motif. However, it seems that the function of BES1-BAM can be independent and depending on plant requirement, the pertinent element starts its act. Some studies are available explaining βamylase alteration in response to plant stresses (Kaplan & Guy, 2004; Todaka et al., 2000). For example, water stress enhances activity of β -amylase and this can activate sugar signaling or regulation by interactions with other TFs and phytohormones starting from BES1 as a master key. The altered activity of β -amylase occurs in different plant life steps and organs from seeds and germination to fruit sweetening. On the other hand, similar changes in BES1 have been reported as a plant response to stresses to enhance plant tolerance (Ahammed et al., 2020; Cao et al., 2024; Gruszka et al., 2020).

Cotton plant belonging to *Gossypium* genus is one of the most important providers of cellulosic fiber and of great economic importance. Cellulosic fiber is a differentiated epidermal cell of seeds produced by cotton plant (John & Crow, 1992). This importance makes it mandatory to advance the application of biotechnological tools for cotton improvement (Rathore et al., 2015). The genome of three species including *G. raimondii*, (Wang et al., 2012), *G. hirsutum* (Li et al., 2015) and *G. arboreum* (Li et al., 2014) have been published. However, the genes those are involved in fiber and cellulose production are under more consideration.

In a previous study, researchers compared TFs and molecular markers among different Gossypium species. They found asymmetric contributions among TFs in these species (Jazayeri et al., 2020). Specifically, they assigned 11, 14, and 16 BES/BZR genes to G. arboreum, G. hirsutum, and G. raimondii, respectively. Interestingly, when comparing protein clusters, a cluster of BES1/BZR TFs revealed two copies for G. arboreum, while G. hirsutum and G. raimondii had only one copy each. Notably, G. hirsutum, a tetraploid resulting from the hybridization of diploid species G. arboreum and G. raimondii, did not double the copy numbers of BES1 family genes. These findings suggest that evolutionary events may balance the gene numbers within the BES1 family due to its multifunctionality. This article is aimed to briefly study the BES1/BZR1 family in comparatively three species of Gossypium and three species of Arabidopsis. The similarities, short sequence repeats and domains of the BES1 TF genes of these six species were disclosed.

2. Methodology

Figure 1 shows the flowchart that was used to analyze the proteins sequence data. The protein sequences of BES1 TFs were downloaded from PlantTFDB (https://planttfdb.gao-lab.org/) website for three species of Arabidopsis including *A. halleri*, A. *lyrata* and *A. thaliana*, and three species of *Gossypium* including *G. arboreum*, *G. hirsutum* and *G. raimondii*. The number of BES1 genes was 8 for *A. halleri*, 9 for *A. lyrata*, 14 for *A. thaliana*, 12 for *G. arboreum*, 24 for *G. hirsutum* and 42 for *G. raimondii*.

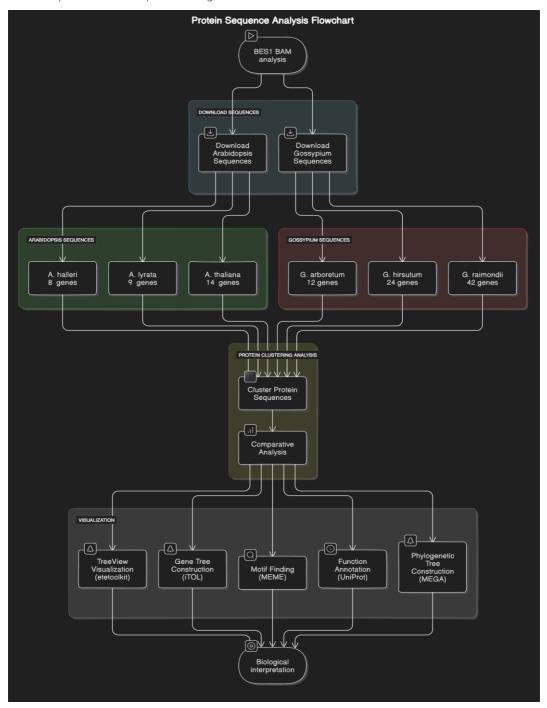


Figure 1. The protein sequence analysis workflow for analyzing BES1/BAM genes involved several steps. First, data were retrieved from public databases. Next, the retrieved sequences were subjected to protein clustering using the OrthoVenn tool. The resulting clusters were then used for further comparative analyses. To visualize the clustered BES1/BAM protein sequences, the ETETOOLKIT tool was employed. Additionally, the MEME tool revealed motifs within the BES1/BAM sequences. Finally, the iTOL tool was used to create a BES1/BAM tree, and sequence alignment and visualization of common sequence domains and repeats were performed using MEGA. The used tools in the study are written in parentheses below corresponding analysis. The flowchart was created by ERASER tool (https://app.eraser.io/).

The OrthoVenn program (Wang et al., 2015) was used to cluster the protein sequences and perform comparative analyses. It is a web-based platform that compares protein sequence datasets among different species, analyzes their annotation and generates Venn diagram and GO enrichment for the clustered proteins (http://www.bioinfogenome.net/OrthoVenn/).

The protein clusters generated by OrthoVenn were visualized by TreeView of Environment for Tree Exploration (ETE) toolkit (Huerta-Cepas et al., 2016) (http://etetoolkit.org/treeview/). It was also used to construct the gene tree for all reported plants in PlantTFDB based on BES1 TF family. The MEME Suite (Multiple Em for Motif Elicitation) was used for finding motifs for each cluster (Bailey & Elkan, 1994). The Universal Protein Resource (UniProt, https://www.uniprot.org/) was employed to reveal pertinent function of each protein cluster. UniProt serves as a comprehensive resource for protein seguence and annotation data. The UniProt databases include the UniProt Knowledgebase (UniProtKB), UniProt Reference Clusters (UniRef), and UniProt Archive (UniParc). The UniProt consortium, along with host institutions EMBL-EBI, SIB, and PIR, is dedicated to the long-term preservation of these valuable databases. The Interactive Tree Of Life (iTOL, https://itol.embl.de/) (Letunic & Bork, **2016**) was used to construct the BES1 phylogenetic tree of six studied species. iTOL is a web-based tool that allows users to display, manipulate, and annotate phylogenetic trees. The Molecular Evolutionary Genetics Analysis (MEGA7.0.21) (Kumar et al., 2016) was employed to align, find and visualize the common sequence domain and repeats among all 109 BES1 input sequences. MEGA is an integrated tool for conducting automatic and manual sequence alignment, inferring phylogenetic trees, mining web-based databases.

3. Results and discussion

The phylogenetic tree of BES1 comprises 71 leaves, representing the evolutionary relationships among these genes. Notably, there are two genus-specific branches that exclusively belong to *Gossypium* species (blue oval in **Figure 2**). Additionally, two other genera exhibit synteny pairs in different branches. As anticipated, the genes of *Arabidopsis* form clusters on branches where no *Gossypium* gene is present. However, intriguingly, certain branches within the *Gossypium* lineage feature sole *Gossypium* genes, distinct from *Arabidopsis* (red rectangles in **Figure 2**). These findings suggest that *Gossypium's* BES1 genes

evolved independently over time, separate from their Arabidopsis counterparts. *Gossypium* species possess a greater number of BES1 genes compared to *Arabidopsis* (**Table 1**). This observation aligns with the fact that BES1 is intricately involved in processes related to cellulose production and is associated with activating brassinosteroids (BRs)—key regulators of cell elongation and size. Notably, BES1 interacts with CESA (cellulose synthase) to modulate cellulose synthesis (**Xie et al., 2011**). The abundance of BES1 members in *Gossypium* underscores their potential impact on advancing cellulose production and downstream fiber-related processes.

In the context of protein clustering, here are the key findings: Proteins from each species tend to cluster together. However, there are singletons (individual proteins not grouped with others) for all species except A. thaliana (Table 1). The proteins form seven clusters in total: four intergenic clusters and three interspecies clusters. Among these, two clusters are specific to the pair of A. thaliana and A. lyrata, while one cluster is unique to G. hirsutum and G. raimondii. Interestingly, no clustered species-specific BES1 was identified. A notable observation is that A. thaliana shares all its BES1 with other studied species. In contrast, G. hirsutum possesses the most speciesspecific BES1. This suggests the potential for G. hirsutum to contribute more significantly to fiber production. By comparing different species of Arabidopsis and Gossypium, researchers can uncover valuable insights into these proteins' roles and functions (Jazayeri et al., 2019; Jazayeri et al., 2018).

Table 1

The proteins and clusters formed based on proteins. Singletons are the genes, which did not form any cluster within the species and are species-specific

Species	Proteins	Clusters	Singletons
A. halleri	8	6	1
A. lyrata	9	8	1
A. thaliana	14	8	0
G. arboreum	12	9	2
G. hirsutum	24	10	6
G. raimondii	42	10	5

All coding sequences share a common protein domain with the following sequence motif [S/G]WM*C[*/S][*/*][*/R]W[N/I]Y (Figure 3).

First position: Most genes have a serine (S) at this position, except for the Cotton_A_07763_BGI-A2_v1.0 gene which has a glycine (G).

Last position: Most genes have an asparagine (N) at this position, while *A. lyrata* 485421 and *G. hirs*utum Gh_D06G2134 genes have an isoleucine (I).

Downstream repeats: Following the conserved domain, there is a recurring pattern rich in arginine (R) and lysine (K) residues, although not perfectly consistent. Studies on α -amylase suggest that the ratio of lysine to arginine can influence protein stability in cold-adapted environments in *Pseudoalteromonas haloplanktis*.

The conserved nucleotide sequence domain in accordance with the conserved protein domain is as follow while the red nucleotides are conserved in all 109 genes.

AATGA[A/T][G/C]TT[A/C]TT[G/A][C/A][T/G]GCT[C/G]TT[G/T][C/G]T[A/T/C/G][G/T/C/A][A/T][G/A]A[A/G]GCTGGATGG[G/T/A][T/C] TGTT[C/G][T/A][T/A][C/G][C/A][T/A]GATGGA[A/T]CTA[C/T]TT[T/ A]T. However, a box as GCTGGATGG and short sequence repeats of TG in coding sequence those might bind to downstream binding sites for BES1 family was retrieved. These repeats can be used as molecular markers for screening *Gossypium* species based on BES1 to evaluate their fiber size and production. The TG repeats are shown in green. **Figure 4** shows the alignment of all 109 genes around the conserved region demonstrating TG repeats in vellow.

AATGA[A/T][G/C]TT[A/C]TT[G/A][C/A][G/T]GCT[G/C]TT[G/T][G/C]]T[A/T/C/G][G/T/C/A][A/T][G/A]A[A/G]GCTGGATGG[G/T/A][T/C] TGTT[G/C][T/A][T/A][C/G][C/A][T/A]GATGGA[A/T]CTA[C/T]TT[T/ A]T.

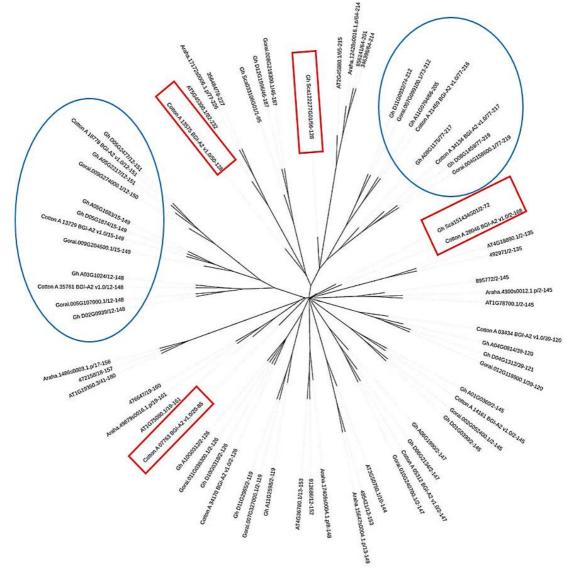


Figure 2. BES1 family tree of protein sequences for three Arabidopsis and three Gossypium. The blue circles show genus-specific gene groups of *Gossypium*, and the red rectangles show the genes of Gossypium that have separated from other genes as the unique gene of the groups they belong to.

Scientia Agropecuaria 15(3): 449-460 (2024)

Species/Abbrv 🛆 Grou					*	* * * * *
1. 346398	A R S S W N ? ? W R R R ? ? * ? ?					
2. 356484	WRRRR IFFWRRRRRKG??					C Y C C S C * R S W M V C * C * W N Y L * ? ? I
3. 472158	SYFYFCCCCC? - ? YGY * ? ?			DLYWT * SSRKL*	SSAL	SS*GSLF*SWMGC*RRWNYL*K?-
4. 476647 5. 485421	SYFYFCCCCCCCCCCC*??				S S ^ A L ^ ^ ^ ^	S S * G S L C * S W M G C * R R W N Y L * K ? - S S * G S L S * S W M D C S * R W I Y L * K ? -
6. 492971		SYLEGKRK**E	KRKKKK <mark>BY</mark> CC*	DECWT*NSWKE*	AS*AL****	SS*GSL**SWMDC*R*WNYL*K?-
7. 856241	ARSSWN??WRRR??*??	* T S * R K G K N * A	* R K T * K S Y Y C *	NSW	S * Y * *	CYCCSC*RSWMGCSS*WNYFSC*V
8. 895772		AYLERKRK**E	KRKKKK <mark>S</mark> YCC*		T S * A L * * * *	SS*GSL**SWMDC*T*WNYL*KG?
9. 912686	RRRR??F??W??	ISYLEGKRK**E	E R K K K K <mark>S </mark> Y Y C *	DLFWT*SSRKL*	4	S S * G S L S * S W M D C * R * W N Y L * K ? -
10. Araha.12428s0016.1.p	A R S S W N ? ? W R R R ? ? * ? ?	* T S * <mark>R K G K N * A</mark>	* R K T * K <mark>S Y Y</mark> C *	N SWRT*KTWKL*	s * s * s * <mark>Y</mark> * * i	C
11. Araha.1486s0003.1.p	SYFYFCCCC? ? YGY * ? ?	K <mark>AFLE</mark> RKRK**KI	KRKKKK <mark>S</mark> CCC*	D L Y W T * S S R K L *	S	S S * G S L F * S W M G C * R R W N Y L * K ? -
12. Araha.15647s0004.1.p				DFYWT ISRKL		SS*GSLS*SWMGCS*RWNYL*N?-
13. Araha.17172s0006.1.p 14. Araha.17172s0006.2.p	WRRRRIFFWRRRR??G?? WRRRRIFFWRRRR??G??					C Y C C S C * R S W M V C * C * W N Y L * ? ? I C Y C C S C * R S W M V C * C * W N Y L * ? ? I
15. Araha. 17406s0004.1.p				DLFWT*SSRKL*		
16. Araha.4300s0012.1.p		AYLERKRK	KRKKKKSYCC*	DFYWT * NVWKL *		SS*GSL**SWMDC*T*WNYL*KG?
17. Araha.49879s0016.1.p	SYFYFCCCCCCCCCCC*??	A F L E R K R K * * K	KRKKKK <mark>S</mark> CCC*	D L Y W T * S S R R L *	5	S S * G S L C * S W M G C * R R W N Y L * K ? -
18. AT1G19350.1	<mark>S Y F Y F </mark>		K R K K K K <mark>S</mark> C C C *		S	S
19. AT1G19350.3	S		KRKKKK <mark>S</mark> CCC*		5	S S * G S L F * SW M G C * R R W N Y L * K ? -
20. AT1G19350.4	S Y F Y F C C C C C C ? ? Y G Y * ? ? I		KRKKKK <mark>S</mark> CCC*		S	S S * G S L F * S W M G C * R R W N Y L * K ? -
21. AT1G19350.5	SYFYFCCCCCC??YGY*??		KRKKKK <mark>S</mark> CCC*		S S · A L · · · ·	S S * G S L F * S W M G C * R R W N Y L * K ? - S S * G S L F * S W M G C * R R W N Y L * K ? -
22. AT1G19350.6 23. AT1G75080.1	S Y F Y F C C C C C C ? ? Y G Y * ? ? S Y F Y F C C C C C C C C C C C C * ? ?		KRKKKKSCCC*		S S ^ A L ^ ^ ^ ^	S S * G S L F * S W M G C * R R W N Y L * K ? - S S * G S L C * S W M G C * R R W N Y L * K ? -
24. AT1G75080.2	SYFYFCCCCCCCCCCC*??				5 5 * A L * * * *	S S * G S L C * S W M G C * R R W N Y L * K ? -
25. AT1G78700.1		AYLERKRK**E			T S * A L * * * *	SS*GSL**SWMDC*T*WNYL*KG?
26. AT2G45880.1		* T S * R K G K N * A		N SWRT*KTWKL*		C
27. AT3G50750.1	R ? ? Y C C ? ? W ? ?	AYLEGKRK**E	ERKKKKSYCC*		4 S * A L * * * *	S S * G S L S * SW M D C S * R W N Y L * K ? -
28. AT4G18890.1	? <mark>D F W N</mark> * I	I S <mark>Y L E G</mark> K R K * * E I	ĸ <mark>ĸĸĸĸ<mark>s</mark> y</mark> c c *	DFCWT*NSWKF*	4 <mark>5</mark> * <mark>A L</mark> * * * *	S S * G S L * * S W M D C * R * W N Y L * K ? -
29. AT4G36780.1		 S 				SS*GSLS*SWMDC*R*WNYL*K?-
30. AT5G45300.1	WRRRR IFWRRRRRRK G??				SSC · S · Y E · I	UYUCSC*RSWMVC*S*WNYL*??I
31. AT5G45300.2	WRRRR IFWRRRRRRKG??		* R K T * K S Y Y F *		S	C Y C C S C * R S W M V C * S * W N Y L * ? ? I
32. Cotton_A_03434_BGI-A2_	TLGCS*NY*GR??KSY*I*	SY*R*KT*E*A		DFCWT*GTWKL*	S * E S * F * *	SS*GCL*RSWMAC*S*WNYL*K?-
33. Cotton_A_05312_BGI-A2_			KKKKKKSYCC*		AS*AL**** SS*AL	S S * G S L * * S W M D C * R R W N Y L * E ? - S S * G S L C * G W M G C * R * W N Y L * K * G
34. Cotton_A_07763_BGI-A2_ 35. Cotton_A_13575_BGI-A2_	CWSFYCFFW? 7 G ? ?	RKRKCKCVM	* RKT * KRVVC *			
35. Cotton_A_13575_BGI-A2_ 36. Cotton_A_13729_BGI-A2_				DLYWT*SSRKL*		
37. Cotton_A_14161_BGI-A2_	?DFWN*	SYLEGKRK**E	KRKKKK <mark>S</mark> YCC*	DFCWTINVWKLI	AS AL · · · ·	5 5 * 6 5 L * * 5 W M D C * T * W N Y L * K ? -
38. Cotton_A_18778_BGI-A2_	SYFYACC? ? ? * ? ?	(KRKKKK <mark>S</mark> YCC*	D L Y W T * S S W K L * I	3	5 5 * 6 5 L C * 5 W M 6 C * R R W N Y L * K ? -
39. Cotton_A_21459_BGI-A2_	SRTSFY??WRR??*??	* T S * R K G K N * A	* R K T * K S S Y C *	N S C W T * K A W K L *	3 * S * S * <mark>Y</mark> * * (C Y C C S C * R S W M G C S S * W N Y F S F * I
40. Cotton_A_28046_BGI-A2_	? <mark>D</mark> F W N * I	A Y L E G K R K * * E I	KRKKKK <mark>S</mark> YCC*	DFCWT*NVWKL*	A S * A L * * * *	5 5 * 6 5 L * * 5 W M D C * T * W N Y L * K E M
41. Cotton_A_34134_BGI-A2_	SRTSF*SRRRR??*??			NISCWT * KTWKL *	S * <mark>S</mark> * <mark>S</mark> * <mark>Y</mark> * * (C
42. Cotton_A_34170_BGI-A2_					A S * A L * * * * * S S * A L * * * * *	SS*GSL**SWMDC*RRWNYL*K?-
43. Cotton_A_35761_BGI-A2_	SYFYSCS????		KRKKKKSYCC*	DLYWT*SSRKL*		SS*GSLF*SWMGC*R*WNYL*K?-
44. Gh_A01G0360 45. Gh_A03G1024	SVEVSCS2	<mark>S Y L E G</mark> K R K * * <mark>E</mark> I (A F L <mark>E</mark> R K R K * * K I	KRKKKKKKKCC	D F C W T * N V W K L * A D L Y W T * S S R K L *		
46. Gh_A04G0814	TLGCS*NY*GR??KSY*I*	SY*R*KT*E*A	KRKKEKSCCS		S * E S * F * *	
47. Gh_A05G1683	SYFYSCSC? ? ? * ? ?	AFLERKRK**KI	KRKKKK <mark>SY</mark> CC*	DLYWT'SSRKL'	3 5 * A L * * * * *	5 5 * 6 5 L C * 5 W M 6 C * R * W N Y L * K 7 -
48. Gh_A05G2217	SYFYESC? ? ? * ? ?	K <mark>A F S * G</mark> K R K * * K I	K R K K K K <mark>S Y</mark> C C *	D L Y W T * S S W K L * :	3	S
49. Gh_A06G1990	? D F W N * I	AYLEGKRK**EI	KRKKKK <mark>SY</mark> CC*	DFCWTINVWKLI	4 S * A L * * * *	S S * G S L * * SWMDC * R R WN Y L * E ? -
50. Gh_A08G1175		* T S * R K G K N * A		N S C W T * K T W K L * I	8 * 8 * 8 * <mark>Y</mark> * * (C
51. Gh_A10G0312	? <mark>D F W N</mark> *	AYLEGKRK**E	KRKKKK <mark>S</mark> YCC*	D F C W T * N V W K L * /	A S * A L * * * * *	S S * G S L * * SWMD C * R R WN Y L * K ? -
52. Gh_A11G0794	SRTSFY??WRR??*??		* RKT * KSSYC *	DFCWT*NVWKL*	3 * <mark>3 * 5</mark> * <mark>Y</mark> * * (CYCCSC*RSWMGCSS*WNYFSF*I SS*GSI**SWMDC*T*WNYI*K?-
53. Gh_A11G2598 54. Gh_D01G0390	2 D F W N *		KRKKKKSTCC*	DFCWTNVWKL	S AL	SS*GSL**SWMDC*T*WNYL*K?- SS*GSL**SWMDC*T*WNYL*K?-
55. Gh_D02G0939		AFLERKRK**K			S * A L * * * *	SS*GSLF*SWMGC*R*WNYL*K?-
56. Gh_D04G1312	TLGCS*NY*GR??KSY*I*	SY*R*KT*E*A	K R K K <mark>E</mark> K S C C S *	DFCWT*GTWKL*	S * <mark>E</mark> S * F * *	SS*GCL*RSWMAC*S*WNYL*K?-
57. Gh_D05G1874	SYFYSCSC? ? ? * ? ?	(5 5 * <mark>A L</mark> * * * * *	5
58. Gh_D05G2477	SYFYESC? ? ? * ? ?		KRKKKK <mark>S</mark> YCC*		S	S
59. Gh_D06G2134	? D F W N * I	AYLEGKRK**E		DFCWT * NVWKL * /	A S * A L * * * * *	SS*GSL**SWMDC*RRWIYL*E?-
60. Gh_D08G1459	SRTSF*SIWRRR? ? * ? ?	* T S * R K G K N * A	* <mark>R K T</mark> * <mark>K S Y Y</mark> C *	NSCWT * KTWKL *	8 * S * S * Y * * (CYCCSC*RSWMGCSS*WNYFSF*I
61. Gh_D10G0318	? D F W N * I	AYLEGKRK**E		DFCWT*NVWKL*/	S *AL****	5 5 * <mark>6 5 L * * 5 W M D</mark> C * <mark>R R W N Y L * K</mark> ? -
62. Gh_D11G0932	SRTSFY??WRR??L??			NSCWT * KAWKL * :		CCSC*RSWMGCSF*WNYFSF*I
63. Gh_D11G2965		A Y L E G K R K * * E I R K R K G K G K N * A			S * A L * * * *	S ^ G S L ^ ^ S W M D C ^ I ^ W N Y L ^ K ? - C S C C S C * R S W M D C * T * W N Y L * ? ? F
64. Gh_D12G1996 65. Gh_Sca033106G01	CWaaterrw / / G / /		* R K T * K S Y Y F *			SCCSC*RSWMDC*T*WNYL*??F
66. Gh_Sca122277G01	SRTSF*SIWRRR? ? * ? ?	* T S * <mark>R K G K N</mark> * A	* R K T * K S Y Y C *	NSCWT* <mark>KTWKL</mark> *		YCCSC*RSWMGCSS*WNYFSF*I
67. Gh_Sca151434G01	? D F WN *	AYLEGKRK**E		D F C W T * N V W K L * /		8 5 * 6 5 L * * 5 W M D C * T * W N Y L * K ? -
68. Gorai.002G052400.1	? <mark>D F W N</mark> * 1	S	KRKKKK <mark>S</mark> YCC*	DFCWT*NVWKL*/	S*AL****	S
69. Gorai.002G052400.2	? <mark>D</mark> F W N * 1			D F C W T * N V W K L * /	S * A L * * * *	8 5 * <mark>6 5 L * * 5 W M D</mark> C * T * W N <mark>Y</mark> L * K ? -
70. Gorai.002G052400.3	• • • • • • • • • • • • • • • ? <mark>D F W N</mark> *			DFCWT*NVWKL*/	S AL A A	S * G S L * * SWMDC * T * WN Y L * K ?
71. Goral.002G052400.4 72. Goral.004G158600.1			KRKKKKSYCC* *RKT*KSYYC*			SIA GOOGLE SWMDC T WNYL KC?
72. Gorai.004G158600.1 73. Gorai.004G158600.10	SRTSF*SIWRRR??*??		* R K T * K S Y Y C *			CYCCSC*RSWMGCSS*WNYFSF*I
74. Goral.004G158600.10	SRTSF*SIWRRR??*??		RKTKSYYC		* S * S * Y * * (YCCSC*RSWMGCSS*WNYFSF*I
75. Gorai.004G158600.12	SRTSF*SIWRRR??*??	* T S * R K G K N * A	* R K T * K S Y Y C *	NSCWT*KTWKL*	3 * 5 * 5 * <mark>Y</mark> * * (YCCSC*RSWMGCSS*WNYFSF*I
76. Gorai.004G158600.13	S R T S F * S I W R R R ? ? * ? ?	* T S * R K G K N * A		N S C W T * <mark>K</mark> T W K <mark>L</mark> * !		C
77. Gorai.004G158600.14	S R T S F * S I W R R R ? ? * ? ?			N S C W T * <mark>K</mark> T W K <mark>L</mark> * :		CYCCSC*RSWMGCSS*WNYFSF*I
78. Gorai.004G158600.2	SRTSF*SIWRRR??*??	* T S * R K G K N * A	* R K T * K S Y Y C *	NSCWT*KTWKL*		CYCCSC*RSWMGCSS*WNYFSF*I
79. Goral.004G158600.3 80. Gorai.004G158600.4	SRTSF*SIWRRR??*??		* R K T * K S Y Y C * * R K T * K S Y Y C *	N S C W T * K T W K L * : N S C W T * K T W K L * :		C Y C C S C * R SWM G C S S * WN Y F S F * I C Y C C S C * R SWM G C S S * WN Y F S F * I
80. Gorai.004G158600.4 81. Gorai.004G158600.5	SRTSF*SIWRRR??*?? SRTSF*SIWRRR??*??		* R K T * K S Y Y C *	NSCWT*KTWKL* NSCWT*KTWKL*	3 × 5 × 5 × V × × /	CYCCSC*RSWMGCSS*WNYFSF*I
82. Gorai.004G158600.6		TSTRKGKNTA		NSCWT*KTWKL*	5 * 5 * 5 * <mark>7</mark> * * (YCCSC*RSWMGCSS*WNYFSF*I
83. Gorai.004G158600.7	SRTSF*SIWRRR??*??	* T S * R K G K N * A	* R K T * K S Y Y C *	N S C W T * <mark>K</mark> T W K L * :	8 * S * S * <mark>Y</mark> * * (YCCSC*RSWMGCSS*WNYFSF*I
84. Gorai.004G158600.8	SRTSF*SIWRRR??*??	* T S * R K G K N * A		N S C W T * <mark>K</mark> T W <mark>K</mark> L * !		
85. Gorai.004G158600.9		* T S * R K G K N * A		NSCWT*KTWKL*	3 * <mark>5 * 5</mark> * <mark>Y</mark> * * (C
86. Gorai.005G107000.1		AFLERKRK**KI				8 8 * 6 8 L F * 8 W M 6 C * R * W N Y L * K ? -
87. Gorai.007G099100.1				NSCWT * KAWKL * !		YCCSC · RSWMGCSS · WN YFSF · I
88. Gorai.007G099100.2 89. Gorai.007G099100.3		* T S * R K G K N * A * * T S * R K G K N * A *	BKT KSYYC*	N S C W T * K A W K L * : N S C W T * K <mark>A W K L</mark> * :	3 × 5 × 5 × V × × /	C
90. Gorai.007G099100.3	SRTSFT??WRR??*??			NSCWT*KAWKL*		YCCSC*RSWMGCSS*WNTFSF*I
91. Gorai.007G099100.5		TSTRKGKNTA	RKTKSYYC	NSCWT * KAWKL*	s·s·s·y··	CYCCSC*RSWMGCSS*WNYFSF*I
92. Gorai.007G099100.6	SRTSFY??WRR??*??	* T S * R K G K N * A	* R K T * K S Y Y C *	NSCWT*KAWKL*	3 *	CYCCSC*RSWMGCSS*WNYFSF*I
	SRTSFY??WRR??*??	* T S * R K G K N * A	* R K T * K S Y Y C *	N S C W T * K A W K L *	S * S * S * <mark>Y</mark> * *	C Y C C S C * R S W M G C S S * W N Y F S F * I
93. Gorai.007G099100.7	SRTSFY??WRR??*??	* T S * R K G K N * A	* R K T * K S Y Y C *	N S C W T * K A W K L *	3 * 5 * 5 * <mark>7</mark> * * 1	C
		AYLEGKRK**E	KRKKKK <mark>S</mark> YCC*	DFCWT*NVWKL*	A S * A L * * * *	SS*GSL**SWMDC*T*WNYL*K?-
93. Gorai.007G099100.7 94. Gorai.007G099100.8 95. Gorai.007G327000.1	?DFWN*		A TO REAL AND A VALUE A	NACWENTSKES	SSC'S'YE'	C S C C S C * R S W M D C * T * W N Y L * ? ? F
93. Goral.007G099100.7 94. Gorai.007G099100.8 95. Gorai.007G327000.1 96. Gorai.008G218300.1		R K R K G K G K N * A				
93. Goral.007G099100.7 94. Gorai.007G099100.8 95. Goral.007G327000.1 96. Goral.008G218300.1 97. Gorai.008G218300.2	CWSFYCFFW? ? G?? CWSFYCFFW? ? G?? CWSFYCFFW? ? G??	R	* R K T * K S Y Y F *	N A C W T * T I W K F S		C S C C S C * R S W M D C * T * W N Y L * ? ? F
93. Goral.007G099100.7 94. Goral.007G099100.8 95. Goral.007G327000.1 96. Goral.008G218300.1 97. Goral.008G218300.2 98. Goral.008G218300.3	CWSFYCFFW?	R K R K G K G K N * A R K R K G K G K N * A R K R K G K G K N * A	* <mark>R K T * K S Y Y F</mark> * * <mark>R K T * K S Y Y F</mark> *	N	SSC * S * <mark>Y E</mark> * 1	C
93. Goral.007G099100.7 94. Goral.007G099100.8 95. Goral.007G327000.1 96. Goral.008G218300.1 97. Goral.008G218300.2 98. Goral.008G218300.3 99. Goral.008G218300.4	CWSFYCFFW? ? G?? CWSFYCFFW? ? G?? CWSFYCFFW? ? G?? CWSFYCFFW? ? G??	R K R K G K G K N * A R K R K G K G K N * A R K R K G K G K N * A R K R K G K G K N * A	* R K T * K <mark>S Y Y F</mark> * * R K T * K S Y Y F * * R K T * K <mark>S Y Y F</mark> *	N A C W T * T I W K F S N A C W T * T I W K F S N A C W T * T I W K F S	S S C * S * <mark>Y E</mark> * (S S C * S * <mark>Y E</mark> * (C S C C S C * R S W M D C * T * W N Y L * ? ? F C S C C S C * R S W M D C * T * W N Y L * ? ? F C S C C S C * R S W M D C * T * W N Y L * ? ? F
93. Goral.007G099100.7 94. Gorai.007G099100.8 95. Gorai.008G218300.1 97. Gorai.008G218300.1 97. Gorai.008G218300.2 98. Gorai.008G218300.3 99. Gorai.008G218300.4 100. Gorai.009G204600.1	C W S F Y C F F W ? ? G ? ? C W S F Y C F F W ? ? G ? ? C W S F Y C F F W ? ? G ? ? C W S F Y C F F W ? ? G ? ? C W S F Y C F F W ? ? G ? ?	R K R K G K G K N * A R K R K G K G K N * A R K R K G K G K N * A R K R K G K G K N * A K A F L E R K R K * * K	* R K T * K S Y Y F * * R K T * K S Y Y F * * R K T * K S Y Y F * K R K K K K S Y C C *	N A C W T * T I W K F S N A C W T * T I W K F S N A C W T * T I W K F S D L Y W T * S S R K L *	S S C * S * <mark>Y E</mark> * 1 S S C * S * <mark>Y E</mark> * 1 S S * A L * * * *	C S C C S C * R S W M D C * T * W N Y L * ? ? F C S C C S C * R S W M D C * T * W N Y L * ? ? F C S C C S C * R S W M D C * T * W N Y L * ? ? F S S * C S C * C * R S W M D C * T * W N Y L * ? ? F
93. Goral.007G099100.7 94. Goral.007G099100.8 95. Goral.007G327000.1 96. Goral.008G218300.1 97. Goral.008G218300.2 98. Goral.008G218300.3 99. Goral.008G218300.4	CW S F Y C F W A P Q P W A	R K R K G K G K N * A R K R K G K G K N * A R K R K G K G K N * A R K R K G K G K N * A A F L E R K R K * * K A F L E R K R K * * K	* R K T * K S Y Y F * * R K T * K S Y Y F * * R K T * K S Y Y F * K R K K K K S Y C C * K R K K K K S Y C C *	N A C W T * T I W K F S N A C W T * T I W K F S N A C W T * T I W K F S D L Y W T * S S R K L * D L Y W T * S S R K L *	S S C * S * <mark>Y E</mark> * S S C * S * <mark>Y E</mark> * S S * A L * * * * S S * A L * * * *	C S C C S C * R S W M D C * T * W N Y L * ? ? F C S C C S C * R S W M D C * T * W N Y L * ? ? F C S C C S C * R S W M D C * T * W N Y L * ? ? F S * C S L C * S W M S C * R * W N Y L * K ? - S S * C S L C * S W M S C * R * W N Y L * K ? -
93. Gorai.007G09100.7 94. Gorai.007G327000.1 95. Gorai.007G327000.1 96. Gorai.008G218300.1 97. Gorai.008G218300.2 98. Gorai.008G218300.3 99. Gorai.008G218300.4 100. Gorai.008G204600.1 101. Gorai.009G204600.2	C M S F Y C F F M 7 ? P M N * C M S F Y C F M 7	R R K N A R K K K N A R K K K K N A R K K K K K K K R K K K K K K K R K K K K K K K R K K K K K K K K C A F L C K K K C A F L C K K K C A F S G K K K	* R K Y Y F * * R K Y Y F * * R K Y Y F * * R K K Y Y F * * R K K S Y Y F * * R K K K S Y C * K R K K K S C C * K R K K K S C C * K R K K K S C C *	N A C WT * T I WK F S N A C WT * T I WK F S D L YWT * S S K L * D L YWT * S S K L * D L YWT * S S K L *	S S C • S • Y E • S S C • S • Y E • S S • A L • • • • S S • A L • • • • S S * A L • • • •	C C C C C S C * R S M M D C * B W M Y L * ? F C C S C * R S M M D C * B W M Y L * ? F C C S C * R S M M D C * B W M Y L * ? F S C C S C * R S M M D C * B W M Y L * ? F S S C S S C * R S M D C * R S M N Y L * ? F S S C S S C * R S M D C * R S M N Y L * ? F S S C S S C * R S M S C * R S M N Y L * ? F S S S S S L C * S M S C * R S M N Y L * ? F S S S S L C * S M S C * R W N Y L * ? F S S * S L C * S M S C * R W N Y L * ? F
93. Garai.007G099100.7 94. Garai.07G03910.8 95. Garai.007G327000.1 96. Garai.008G218300.1 97. Garai.008G218300.2 98. Garai.008G218300.4 100. Garai.008G204600.1 110. Garai.005G204600.2 112. Garai.009G274000.1 113. Garai.01G240700.2	C M S F Y C F M P C M S P C F M P C M P C M P C M P C M P C M P C M P C M P C M P C M P C M P C M P C M P C M P C M P C M P C M P C M P M C M P M	R K	* R K T * K S Y Y F * R K T * K S Y Y F * R K T * K S Y Y F * R K K K K S Y C C K R K K K K S Y C C K R K K K K S Y C C K R K K K K S Y C C K R K K K K S Y C C K R K K K K S Y C C K R K K K K S Y C C K R K K K K S Y C C	N A C T T K F S N A C W T T K F S N A C W T T K F S Q L Y T S S R L * Q L Y T S S R K L Q L Y T S S R K L Q L Y T S S R K L Q L Y T S S K L L Q F C T N Y K L	S S C · S · Y E · S · S C · S · Y E · S · S · Y E · S · S · S · S · S · S · S · S · S ·	C C C T W Y
93 Garai (007G099100 7 94 Gorai (007G099100 8 95 Gorai (007G092100 1 96 Gorai (006218300 1 97 Garai (006218300 2 99 Gorai (0062018300 4 100 Garai (006201800 2 100 Garai (006201800 2 102 Garai (006201800 1 103 Garai (10622100 1 104 Gorai (10628070 1 105 Garai (11603830 1	C M S F Y C F F M 7 7 0 7 0 7 0 7 0 7 0 7 0 7 0 7 0 7	R K K K K R K K K K R K K K K R K K K K R K K K K R K K K K R K K K K R K K K K R K K K K R K K K K R K K K K R K K K K R K K K K K K K K K K K K K K K K K K K K K K K K K K K K K	R T K Y Y F R K T K S Y Y F R K T K S Y Y F R K T K S Y Y F R K T K S Y Y F K K K S Y C * K K K S Y C * K K K S Y C * K K K K S Y C * K K K K S Y C * K K K K S Y C * K K K K K S Y C *	N A C W T * T I W K F S N A C W T * T I W K F S O L Y W T * S S R K L * D L Y W T * S S R K L * D L Y W T * S S R K L * D F C W T * N V W K L * D F C W T * N V W K L *	S C S Y E S C S Y E S C S Y E S C S Y E S C A Y E S A L Y E S A L Y E S A L Y E S A L Y E S A L Y E A S A L Y	C C T W Y
93. Garai.007G099100.7 94. Garai.07G03910.8 95. Garai.007G327000.1 95. Garai.008G218300.1 97. Garai.008G218300.2 98. Garai.008G218300.4 100. Garai.008G218300.4 100. Garai.008G24600.2 112. Garai.008G274000.1 103. Garai.010G240700.2 105. Garai.010G240700.2 105. Garai.010G24070.2	C M B F Y C F M C M S Y C F M C M S Y C F M C M S Y C N T C N T C N T C N T C N T C N T C N T C N T C N T C N T C N N C N N C N N N C N	R K	R T K S Y Y F K S Y Y F K K S Y Y F K K S Y Y F K K S Y Y F K K K S Y Y F K K K S Y Y F K	N A C W T * T I W K F S N A C W T * T I W K F S D L Y W T * S S R K L * D L Y W T * S S R K L * D L Y W T * S S R K L * D F C W T * N V W K L * D F C W T * N V W K L * D F C W T * N V W K L *	S C S Y E Z Y E Y Z Y Z Y Z Y Z Y Z Y	C C C C T
93 Garai (007G099100 7 94 Gorai (007G099100 8 95 Gorai (007G092100 1 96 Gorai (006218300 1 97 Garai (006218300 2 99 Gorai (0062018300 4 100 Garai (006201800 2 100 Garai (006201800 2 102 Garai (006201800 1 103 Garai (10622100 1 104 Gorai (10628070 1 105 Garai (11603830 1	C M S F Y C F F M ? ? 2 F M N * C M S F Y C F F M ? . ? 2 F M N * C M S F Y C F F M ? . ? 2 7 ? C M S F Y C F M ? . ? 2 7 ? C M S F Y C F M ? . ? 2 7 ? S Y F Y S C G ? ? ? ? ? S Y F Y S C G ? ? ? ? ? ? S Y F Y S C G ? ? ? ? ? ? ? ? S Y F Y S C G ? ?		R K T K S Y Y F R K T K S Y Y F R K T K S Y Y F R K T K S Y Y F K R K K K S Y C C K C C K R K K K S Y C C K C C K R K K K S Y C C K C C K R K K K S Y C C K C C K R K K K S Y C C K C C K R K K K S Y C C K C C K R K K K S Y C C K C C K R K K K S Y C C K C C K R K K K S Y C C K C C	N A C W T * T I W K F S N A C W T * T I W K F S O L Y W T * S S R K L * D L Y W T * S S R K L * D L Y W T * S S R K L * D F C W T * N V W K L * D F C W T * N V W K L *	B C S Y E Y Z Y Y E Y Z Y Y Z Y Y Z Y Z Y Z Y Z Y Z Y Z Y Z Y Z Z	C C C T W Y

Figure 3. The conserved protein sequence among 109 genes. The stars on the top mean the conserved amino acid in all proteins.

Scientia Agropecuaria 15(3): 449-460 (2024)

	6. 492971	CTAAGCATTGTGATAATAATGAAGTTCTTAAGGCTCTTTGTAATGAAGCTGGATGGA
	7.856241	A C C T G A T A T T A A T G A T G T T A T T G C T G C T C T T G C T A G A A G C T G G A T G G G T T G T T C T T C C T G A T G G A A C T A C T T T C C T G
	8. 895772	
	15. Araha.17406s0004.1.p	
	22. AT1G19350.6	
	23. AT1G75080.1	CTAAGCATTGTGATAATAATGAAGTTCTTAAGGCTCTTTGTGTTGAAGCTGGATGGGTTGTTGAAGAAGATGGAACTACTACTATAGA
	24. AT1G75080.2	
	29. AT4G36780.1	
	30. AT5G45300.1	
	31. AT5G45300.2	CT C C T A G A G C T S A T A T S A T S A T S A T A T T G C T C C T C C T C C T A G A A G C T G G T C C T C A T S G A A C T A C T A T A G A C T A
	32. Cotton_A_03434_BGI-A2_	CTAASAGAGCTCATTCTAATGATCTTCTTAAGGCTGTTTGTGAAGAAGCTGGATGGCATGGTGAAGCTGATGGAACTATTTATAGA
	33. Cotton_A_05312_BGI-A2_	
	34. Cotton_A_07763_BGI-A2_	
	39. Cotton_A_21459_BGI-A2_	GAGTTAGAGCTGATATTAATGATGTTATTGCTGCTCTTGCTAGAGAAGCTGGATGGGTTGTTCTTCCTGATGGAACTACTTTTCCT
	40. Cotton_A_28046_BGI-A2_	CTAAGCATTGTGATAATAATGAAGTTCTTAAGGCTCTTTGTAATGAAGCTGGATGGA
	41. Cotton_A_34134_BGI-A2_	GAGTTAGAGCTGATATTAATGATGATGTTATTGCTGCTGCTGCTGGATGGGTTGTTCTTCCTGATGGAACTACTATTTCCT
	47. Gh_A05G1683	
	48. Gh_A05G2217	
	49. Gh_A06G1990	
	54. Gh_D01G0390	
	55. Gh_D02G0939	CTAACCATIGIGATAATAATGAAGTTCTTAAGGCTCTTIGTTCTGAAGCTGGGTGGTTGTTGAAGATGATGGAACTACTTATAGA
	56. Gh_D04G1312	CTARGERGETGATTCTAATGATCTTCTTARGECTGTTGTGAAGCTGGATGGCATGTTGAAGCTGGATGGCAGCTATTTATAGA
	64. Gh_D12G1996	C T C C T A C A C T C A T A T C A A T C A T C T T C C T C C T C C T A C A C
	65. Gh_Sca033106G01	
	70. Gorai.002G052400.3	
	71. Gorai.002G052400.4	
	72. Gorai.004G158600.1	ONGTINGNOCICNININICNICNICNICNICNICCICCICCICCICCIC
	73. Gorai.004G158600.10	
	77. Gorai.004G158600.14	GAGTTAGA CCTGATATTAATGATGTTATTGCTGCTCCTGCTAGAAGCTGGATGCGGTTGTTCTTCCTGATGGAACTACTTTTCCT
	78. Gorai.004G158600.2	GAGTTAGAGCTGATATTAATGATGTTATTGCTGCTCTTGCTAGAGAAGCTGGATGGGTTGTTCTTCCTGATGGAACTACTTTTCCT
		GAGTTAGAGCTGATATTAATGATGTTATTGCTGCTCTTGCTAGAGAAGCTGGATGGGTTGTTCTTCCTGATGGAACTACTTTTCCT
	80. Gorai.004G158600.4	
		NON TRANSFER OF CONTRACTOR CONTROL OF CONTRO
	82. Gorai.004G158600.6 83. Gorai.004G158600.7	
	84. Gorai.004G158600.8	
	85. Gorai.004G158600.9	GAGTTAGAGCTGATATTAATGATGATGATGCTGCTGCTGCTGCAGAGAGCTGGGTTGTTCTTCCTGCTGCAGACAACTACTATTTCCT
	86. Gorai.005G107000.1	
	87. Gorai.007G099100.1	GAGITAGAGCTGATAITATGATGATGTTGTTGCTGCTCGTGCTGGAGCTGGACTACTTGCTGCTGCTGCTGGACTACTTTCCT
		NAME AND A DESCRIPTION OF A DESCRIPTIONO
	91. Gorai.007G099100.5	GAGTTAGARCTGATATTATTGATGGTTGCTTGCTTGCTTGCTTGCTTGC
	92. Gorai.007G099100.6	GAGTTAGAGCTGATATTAATGATGTTATTGCTGCTCTTGCTAGAGAAGCTGGATGGGTTGTTCTTCCTGATGGAACTACTTTTCCT
	93. Gorai.007G099100.7	
	94. Gorai.007G099100.8	
		C T S C T A S A S C T S A T A T S A T C A T S T T C T T S C T A C A S A A S C T S C T A C A C T S A A C T S A
	98. Gorai.008G218300.2	C T G C T A G A G C T G A T G A A T G A T G A T G C T T G C T G C T G C T G G A G G A G G A G C T G G A G G A G C T G A T G G A A G C T A C T A C A G A G A G C T
	99. Gorai.008G218300.4	CTGCTAGACCTGATATGATGATGTTCTTGCTGCTCCTCCTAGAAGCTGGATGGA
	100. Gorai.009G204600.1	CTAAGCATTGTGATAATAATGAAGTTCTTAAGGCTCTTTTGTGTTGAAGCTGGATGGGTTGAAGATGGAAGATGGAACTACTTATAGA
		C T A A G C A T T G T G A A T A A T G A A G T T C T T A A G G C T C T T T G T G T G G A A G C T G G T T G T T G A A G A T G A T G G A A C T A C T T A T A G A G A G A G A T G A G A T G A A G A T A G A T A G A T A G A T A G A T A G A T A G A T A G A T A G A T A A T A A T A A G A T A A T A A T A A T A A T A A T A A A T A A A T A A A T A A T A A G A T A G A T A G A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A G A T A
	102. Gorai.009G274000.1	CTARSCATTSTGATAATAATGAAGTTCTTTAAGGCTCTTTGTGCTGAAGCTGGATGGGTTGTTGAAGATGGAACTACTACTTATAGA
105. Goral 0116035300.1 CTARS CATIGINATATATGAGETICTEAGCCTCTTIGINATGAGCTGGTGGTGGTGGTGGTGGGCGTGTGGGGGGGGGG		
106. Geral 0110633300.2 CTARGCATTGTGATATATATGAGTTCTTAGGCTCTTTGTAAGCTGGATGGA		
107. Goral 0116036300.3 CTAAGCATTGTGATGATAATGAAGTTCTTAAGGCTCTTTGTAAGCGTGGATGGA		C TA A G C A TT G T G A TA A T G A A G TT C TT T A A G G C T C TT T G T A A T C A A G C T G G A T G G A C T G TT G A A G A A G A T G G A A C T A C TT A T A G A
108. Goral 011G036300.4 CTAACCATTGTGATAATGAAGTCTTAAGGCTCTTTGTAATGAACCTGGATGGA		CTARGCATIGIGATAATAATGAAGTTCTTAAGGCTCTTTGTAATGAAGCTGGATGGA
109. Goral 0125118900.1 CTAAGAGAGAGCTGATTCTAATGATCTTAAGGCTGTTTCTGAAGAAGCTGGCATGGCATGGCTGTTGTGGAAGCTGGATGGCATGGCATGGCATGGAAGCTGATGGAAGCTGATGGAAGCTGATGGAAGCTGATGGAAGCTGATGGAAGCTGATGGAAGCTGATGGAAGCTGATGGAAGCTGATGGAAGCTG	108. Gorai.011G036300.4	C TAAGCATIGIGATAATAATGAAGTICITAAGGCTCITIGIAATGAAGCIGGATGGACTGITGAAGAAGAIGAACTACITATAGA
	109. Gorai.012G118900.1	CINA GAGAGECIGANICIANICANCINCINALGECIGINIGIGAAGEIGGAIGECAIGUAGEIGGAIGEAGEIGAAGEIGIGAAGEIGAAGEIGAAGEIGAAGEIGAAGE

Figure 4. The alignment of 109 BES1 genes among the 3 Arabidopsis and 3 Gossypium species. The stars on the top of columns means conserved nucleotide. TG motif is highlighted in yellow.

In an intriguing discovery, the four clusters of common BES1 genes shared between Arabidopsis and Gossypium were categorized into two distinct groups: β-amylase and BES1. Specifically, there were two clusters associated with each of these gene families (as indicated in Table 2). Notably, BMY7 and BMY8 exhibit attributes of BES1 TFs, consistent with previous reports by (Reinhold et al., 2011). When overexpressed in plants, BMY7 and BMY8 enhance tolerance to drought, heat, and Conversely, osmotic stress. plants with overexpressed BES1 genes also demonstrate resilience to these environmental challenges. The fascinating aspect lies in the polyvalent nature of BES1 and BMY genes, driven by their shared functional and regulatory motifs. This dual functionality allows them to respond differentially in various contexts, depending on the plant's requirements. BMY genes, acting as β-amylases, help balance starch content and produce maltose as a signaling sugar. Meanwhile, BES1 TFs regulate

gene expression and facilitate crosstalk among phytohormones and other TFs (**Song et al., 2018**).

In Figure 5, G. arboreum genes exhibit greater proximity to the Arabidopsis clade compared to the two other Gossypium species. Meanwhile, G. hirsutum genes are partially situated within their own clade. Notably, G. raimondii genes form a cohesive cluster, except for the outlier Gorai.008G218300.2. The Arabidopsis species follow a similar pattern: A. halleri and A. lyrata share a branch, while A. thaliana occupies a separate branch. Interestingly, A. halleri and A. lyrata are more closely related to each other than to A. thaliana, suggesting distinct evolutionary patterns in A. thaliana compared to the other two species. Furthermore, among the 28 BAM7 genes, three G. hirsutum genes exhibit varying relationships: one closely resembles G. arboreum, another is similar to G. raimondii, and a third stands alone in a distant clade. Additionally, Cluster 2 encompasses 11 BAM8 genes (Figure 6).

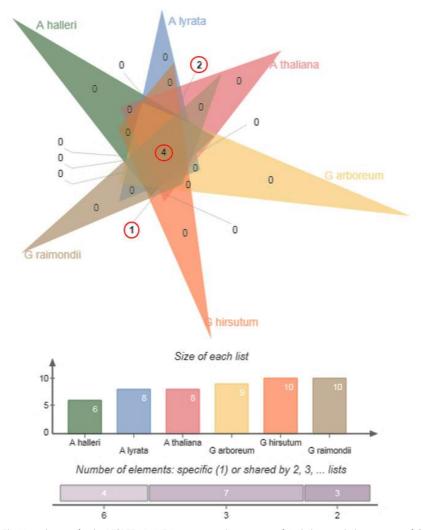


Figure 5. The Venn diagram for the BES1 TF / BAM7/8 comparing three species of Arabidopsis with three species of Gossypium.

Table 2

The four common clusters among all 6 species. GO terms are presented for each cluster. In total 39 genes clustered as BAM and 18 genes as BES1 while the inputs were known as BES1

ID	No. of proteins	Swiss-Prot Hit	GO Annotation
Cluster 1	28	B-amylase 7 O80831	GO:0005737; C:cytoplasm; GO:0005634; C:nucleus; GO:0016161; F: β -amylase activity; GO:0003700; F:sequence-specific DNA binding transcription factor activity; GO:0000272; P:polysaccharide catabolic process; GO:0048831; P:regulation of shoot system development; GO:0006355; P:regulation of transcription, DNA-templated
Cluster 2	11	B-amylase 8 Q9FH80	GO:0005737; C:cytoplasm; GO:0005634; C:nucleus; GO:0016161; F:β-amylase activity; GO:0003700; F:sequence-specific DNA binding transcription factor activity; GO:0000272; P:polysaccharide catabolic process; GO:0048831; P:regulation of shoot system development; GO:0006355; P:regulation of transcription, DNA-templated;
Cluster 3	9	Protein BRASSINAZOLE- RESISTANT 1 Q8S307	GO:0005829; C:cytosol; GO:0005634; C:nucleus; GO:0003677; F:DNA binding; GO:0003700; F:sequence-specific DNA binding transcription factor activity; GO:0009742; P:brassinosteroid mediated signaling pathway; GO:0045892; P:negative regulation of transcription, DNA-templated; GO:0048481; P:ovule development; GO:0040008; P:regulation of growth; GO:0006355; P:regulation of transcription, DNA-templated; GO:0048316; P:seed development; GO:0006351; P:transcription, DNA-templated;
Cluster 4	9	BES1/BZR1 homolog protein 4 Q9ZV88	GO:0003677; F:DNA binding; GO:0006355; P:regulation of transcription, DNA-templated; TAS:TAIR; GO:0006351; P:transcription, DNA-templated

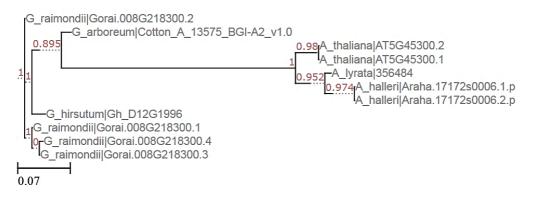
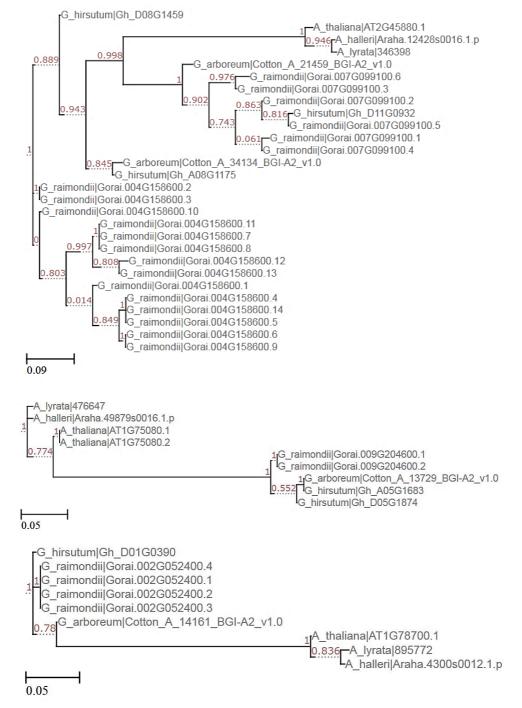


Figure 6. Gene tree for cluster 1 including 28 genes of BAM7 (above) and cluster2 with 11 genes of BAM8 (below).

In the context of BES1 clusters, as depicted in **Figure** 7, a consistent pattern emerges among *Arabidopsis* species: *A. lyrata* and *A. halleri* share a branch, while *A. thaliana* occupies a distinct branch. Notably, all *G. raimondii* genes cluster together within a single clade. In contrast, *G. hirsutum* genes form a separate clade, except for Gh_A05G1683, which aligns with the *G. arboreum* clade. These findings align with those observed for BAM genes, suggesting that these specific genes may serve as representative markers for evolutionary events within the studied species.

Among the motif within Cluster 1, which comprises 28 genes (as illustrated in **Figure 8**), recurring occurrences of the amino acid glycine (G) in the peptide sequences was observed. Glycine stands out due to its unique properties: it lacks a chiral carbon and possesses a hydrogen side chain (unlike

other amino acids, which feature carbon-based side chains). Consequently, glycine exhibits greater conformational flexibility. This intriguing observation leads us to hypothesize that glycine may play a role in the functionality of amylases. Notably, glycine is highly abundant among the 39 genes of both BAM7 and BAM8. Additionally, another recurring amino acid residue: glutamic acid (E) was identified. Glutamic acid is essential for α -amylase activity, as previously reported by (MacGregor et al., 2001). It is speculated that a similar pattern may exist for βamylase. Furthermore, both L-glycine and Lglutamic acid are known to be effective chelating agents in proteins, aligning with existing literature. These findings highlight the importance of further investigation to ascertain whether these amino acid residues are involved in the dual functionalities of BES1 and BAM.



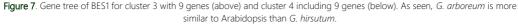




Figure 8. The most conserved motif of cluster1. It is rich of E and G.

4. Conclusions

The BES1 and BAM7/8 gene families share common domains, which makes them crosstalk elements capable of regulating gene expression and the function of downstream genes. Initially, their similar stress response patterns in plants suggested that they belonged to the same group due to their common binding domain and conserved motifs. However, recent findings indicate that they belong to two distinct gene groups. Despite this separation, their shared domains allow them to collaborate with each other. This study disclosed that these genes are known as BES1 but they share highly similar domains with BAM7 and BAM8. Therefore, motifs and BES1/BAM-based family are the same suggesting these genes are multifunctional.

Interestingly, BES1 and BAM7/8 directly generate maltose, a sugar that may function as a regulatory signal and play a role in plant responses to stress, growth, and development processes. In the context of Arabidopsis and Gossypium (cotton), a specific domain—G box GCTGGATGG—and repeats of TG have been identified in all 109 BES1 genes. This domain can be used for marker-assisted selection in Gossypium species. Notably, G. hirsutum exhibits less kinship, while G. arboreum and G. raimondii are closer species. This research proposes a similar structure and function for the RK repeats in plant βamylase, potentially impacting enzyme activity. However, this connection needs further investigation by functional genomics and molecular biology to disclose their functionality.

Given their regulatory functionality in cellulose and fiber production, BES1/BAM members are promising candidate genes for further analysis in *Gossypium*. Additionally, in *Arabidopsis* species, BES1 and BAM7/8 relationships reveal that *A. halleri* and *A. lyrata* are more closely related to each other than to *A. thaliana*. The precise bifunctionality of the BES1/BAM complex remains to be fully uncovered.

Declaration of conflict of interest

The authors have no conflicts of interest to declare.

Author Statement

R. O. Villamar-Torres: Conceptualization, Data curation, Methodology, Visualization, Writing – original draft, Writing – review & editing, C. A. Mestanza Uquillas: , H. D. Chévez-Vera: Investigation, Validation, Visualization, M. R. Heredia-Pinosi: Investigation, Validation, Visualization, C. R. Viot. Conceptualization, Data curation, Investigation, Methodology, Project administration, Resources, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing, S. M. Jazyeri: Conceptualization, Resources, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing.

ORCID

- R. O. Villamar-Torres https://orcid.org/0000-0003-2511-1789
- C. A. Mestanza Uquillas D https://orcid.org/0000-0001-9299-170X
- H. D. Chévez-Vera https://orcid.org/0009-0002-5620-7628
- M. R. Heredia-Pinos D https://orcid.org/0000-0003-1662-7337 C. R. Viot D https://orcid.org/0000-0001-6321-4741
- S. M. Jazayeri D https://orcid.org/0000-0003-2852-5063

References

- Ahammed, G. J., Li, X., Liu, A., & Chen, S. (2020). Brassinosteroids in plant tolerance to abiotic stress. *Journal of Plant Growth Regulation*, 39 (4), 1451–1464. https://doi.org/10.1007/S00344-020-10098-0
- Bailey, T. L., & Elkan, C. (1994). Fitting a mixture model by expectation maximization to discover motifs in biopolymers. *Proceedings. International Conference on Intelligent Systems* for Molecular Biology, 2, 28–36.
- Campos, M. L., de Almeida, M., Rossi, M. L., Martinelli, A. P., Litholdo Junior, C. G., et al. (2009). Brassinosteroids interact negatively with jasmonates in the formation of anti-herbivory traits in tomato. *Journal of Experimental Botany*, 60(15), 4347– 4361. https://doi.org/10.1093/jxb/erp270
- Cao, X., Wei, Y., Shen, B., Liu, L., & Mao, J. (2024). Interaction of the transcription factors BES1/BZR1 in plant growth and stress response. *International Journal of Molecular Sciences*, 25(13), 6836. https://doi.org/10.3390/JJMS25136836
- Gruszka, D., Srivastava, A. K., Yin, Y., & Kono, A. (2020). Updates on BES1/BZR1 regulatory networks coordinating plant growth and stress responses. *Front. Plant Sci*, *11*, 617162. https://doi.org/10.3389/fpls.2020.617162
- Huerta-Cepas, J., Serra, F., & Bork, P. (2016). ETE 3: reconstruction, analysis, and visualization of phylogenomic data. *Molecular Biology and Evolution*, 33(6), 1635–1638. https://doi.org/10.1093/molbev/msw046
- Jazayeri, S. M., Pooralinaghi, M., Villamar-Torres, R. O., & Cruzatty, L. C. G. (2018). An in silico comparative genomic report on transcription factors in three Arabidopsis species. *Ciencia y Tecnología*, 11(1), 1–9. https://doi.org/10.18779/cyt.v11i1.185
- Jazayeri, S. M., Villamar-Torres, R. O., Zambrano-Vega, C., Cruzatty, L. C. G., Oviedo-Bayas, B., et al. (2020). Transcription factors and molecular markers revealed asymmetric contributions between allotetraploid Upland cotton and its two diploid ancestors. *Bragantia*, 79(1), 30-46. https://doi.org/10.1590/1678-4499.20190161
- Jazayeri, S.M., García Cruzatty, L. C., Villamar-Torres, R., Cruzatty, L. C. G., & Villamar-Torres, R. (2019). Genomic comparison among three arabidopsis species revealed heavy metal responsive genes. *Journal of Animal and Plant Sciences*, 29(2).
- John, M. E., & Crow, L. J. (1992). Gene expression in cotton (Gossypium hirsutum L.) fiber: cloning of the mRNAs. Proceedings of the National Academy of Sciences of the United States of America, 89(13), 5769–73.
- Kaplan, F., & Guy, C. L. (2004). beta-Amylase induction and the protective role of maltose during temperature shock. *Plant Physiology*, 135(3), 1674–84. https://doi.org/10.1104/pp.104.040808
- Kumar, S., Stecher, G., & Tamura, K. (2016). MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Molecular Biology and Evolution*, 33(7), 1870–1874. https://doi.org/10.1093/molbev/msw054
- Letunic, I., & Bork, P. (2016). Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Research*, *44*(W1), W242–W245. https://doi.org/10.1093/nar/gkw290
- Li, F., Fan, G., Lu, C., Xiao, G., Zou, C., et al. (2015). Genome sequence of cultivated Upland cotton (*Gossypium hirsutum* TM-1) provides insights into genome evolution. *Nature Biotechnology*, 33(5), 524–530. https://doi.org/10.1038/nbt.3208

- Li, F., Fan, G., Wang, K., Sun, F., Yuan, Y., et al. (2014). Genome sequence of the cultivated cotton *Gossypium arboreum*. *Nature Genetics*, 46(6), 567–572. https://doi.org/10.1038/ng.2987
- MacGregor, E. A., Janecek, S., & Svensson, B. (2001). Relationship of sequence and structure to specificity in the alpha-amylase family of enzymes. *Biochimica et Biophysica Acta*, 1546(1), 1–20.
- Manghwar, H., Hussain, A., Ali, Q., & Liu, F. (2022). Brassinosteroids (BRs) role in plant development and coping with different stresses. *International Journal of Molecular Sciences*, 23(3), 1012. https://doi.org/10.3390/JJMS23031012
- Nolan, T., Chen, J., & Yin, Y. (2017). Cross-talk of Brassinosteroid signaling in controlling growth and stress responses. *The Biochemical Journal*, 474(16), 2641–2661. https://doi.org/10.1042/BCJ20160633
- Rathore, K. S., Campbell, L. M., Sherwood, S., & Nunes, E. (2015). Cotton (Gossypium hirsutum L.). Methods in molecular biology (Clifton, N.J.), 1224, 11–23. https://doi.org/10.1007/978-1-4939-1658-0_2
- Reinhold, H., Soyk, S., Simková, K., Hostettler, C., Marafino, J., et al. (2011). β-amylase-like proteins function as transcription factors in Arabidopsis, controlling shoot growth and development. *The Plant Cell*, *23*(4), 1391–403. https://doi.org/10.1105/tpc.110.081950
- Ryu, H., Cho, H., Bae, W., & Hwang, I. (2014). Control of early seedling development by BES1/TPL/HDA19-mediated epigenetic regulation of ABI3. *Nature Communications*, 5, 4138. https://doi.org/10.1038/ncomms5138
- Shi, H., Li, X., Lv, M., & Li, J. (2022). BES1/BZR1 family transcription factors regulate plant development via brassinosteroiddependent and independent pathways. *International Journal of Molecular Sciences*, *23*(17), 10149. https://doi.org/10.3390/JJMS231710149/S1
- Song, X., Ma, X., Li, C., Hu, J., Yang, Q., et al. (2018). Comprehensive analyses of the BES1 gene family in Brassica napus and examination of their evolutionary pattern in representative

species. BMC Genomics, 19(1), 1–15. https://doi.org/10.1186/S12864-018-4744-4

- Todaka, D., Matsushima, H., & Morohashi, Y. (2000). Water stress enhances beta-amylase activity in cucumber cotyledons. *Journal of Experimental Botany*, 51(345), 739–745.
- Wang, K., Wang, Z., Li, F., Ye, W., Wang, J., et al. (2012). The draft genome of a diploid cotton *Gossypium raimondii*. *Nature Genetics*, 44(10), 1098–1103. https://doi.org/10.1038/ng.2371
- Wang, Y., Coleman-Derr, D., Chen, G., & Gu, Y. Q. (2015). OrthoVenn: a web server for genome wide comparison and annotation of orthologous clusters across multiple species. *Nucleic Acids Research*, 43(W1), W78-W84. https://doi.org/10.1093/nar/gkv487
- Wu, P., Song, X., Wang, Z., Duan, W., Hu, R., et al. (2016). Genomewide analysis of the BES1 transcription factor family in Chinese cabbage (*Brassica rapa* ssp. pekinensis). *Plant Growth Regulation*, 80(3), 291–301. https://doi.org/10.1007/s10725-016-0166-y
- Xie, L., Yang, C., & Wang, X. (2011). Brassinosteroids can regulate cellulose biosynthesis by controlling the expression of CESA genes in Arabidopsis. *Journal of Experimental Botany*, 62(13), 4495–4506. https://doi.org/10.1093/jxb/err164
- Yang, T., Li, H., Li, L., Wei, W., Huang, Y., Xiong, F., & Wei, M. (2023). Genome-wide characterization and expression analysis of αamylase and β-amylase genes underlying drought tolerance in cassava. BMC Genomics, 24(1), 1–14. https://doi.org/10.1186/S12864-023-09282-9
- Yasir, T. A., & Wasaya, A. (2021). Brassinosteroids Signaling pathways in plant defense and adaptation to stress. *Plant Growth Regulators: Signalling under Stress Conditions*, 197–206. https://doi.org/10.1007/978-3-030-61153-8_9
- Yin, Y., Vafeados, D., Tao, Y., Yoshida, S., Asami, T., & Chory, J. (2005). A new class of transcription factors mediates brassinosteroidregulated gene expression in Arabidopsis. *Cell*, 120(2), 249–259. https://doi.org/10.1016/j.cell.2004.11.044