Integrative
Plant Biotechnology
ISSN: 3007-2735 (Online), 3007-2727 (Print)

Check for updates

ROOTS PRESS

Research Article

Genome wide characterization and expression insights of Auxin Response Factor (ARF) gene family in cabbage in response to clubroot disease

Ikhlas Shafique[1], M. Ehetisham-ul-Haq[2], Muhammad Anas[1], Usman Razzaq[1], Muhammad Umar Shahbaz[2]*

[1] *Department of Plant Pathology, University of Agriculture, 38000, Faisalabad, Pakistan.*
[2] *Plant Pathology Research Institute, AARI, Faisalabad, Pakistan.*

ABSTRACT

Auxin response factors (*ARFs*) regulate genes expression in response to auxin. Auxin performs a significant role in plant development and growth such as vascular tissue differentiation, root initiation and apical dominance. Within plants, auxin promotes the growth of vascular tissue differentiation, apical dominance and at the cellular level, auxin helps cells to divide and differentiate. Members of the ARF family carry a DNA binding domain (DBD) that binds to specific DNA sequences in the promoters of target genes. Despite the crucial role of the ARF gene family, no comprehensive study has been carried out to characterize the comparative ARF in *B oleracea* and *A thaliana*. For this reason, an *in-silico* approach was adopted. During the study, we found 49 genes in *B oleracea* and 37 genes in *A thaliana*. Of these, only 17 and 22 genes were studied due to their redundancy. In the phylogenetic analysis, we included *B oleracea*, *B rapa, B napus* and *A thaliana* and divided them into four groups based on conserved domains. Chromosomal mapping identified the positions of the genes on different chromosomes. The diversity of introns in the gene structure varies from 1 to 13 and up to five protein motifs are shown on each gene. Many transcription factors were identified by cis-regulatory elements. Gene expression revealed the involvement of *BolARFs* in root gall disease. The presence of nine different conserved domains was observed, of which B3, AUX_IAA superfamily together with auxin response were dominantly repeated in maximum number of genes. The overall study provides new insights and directions for further research and analysis in the future.

Keywords: Auxin response factors; Brassica species; Phylogenetic analysis; Expression analysis; Clubroot disease.

OPEN ACCESS

***Correspondence:**
Muhammad Umar Shahbaz
umar739@yahoo.com

Article History
Received: September 25, 2023
Accepted: December 05, 2023
Published: December 15, 2023

Cite this article
Shafique, I., Shahbaz, M. U., Ehetisham-ul-Haq, M., Anas, M., & Abbas, A. (2023). Genome-wide characterization and expression insights of Auxin Response Factor (ARF) gene family in cabbage in response to clubroot disease. Integrative Plant Biotechnology, 01, 17-32.

INTRODUCTION

The genus Brassica, part of the family Brassicaceae with the tribe Brassiceae, comprises 39 species (Katche *et al.*, 2019), making it the most important genus within the family, which comprises 338 genera and 3709 species (Al-Shehbaz *et al.*, 2006; Warwick *et al.*, 2006). *Brassica oleracea* (cabbage) consists of 18 (2n) chromosomes. This species includes various vegetables such as broccoli, cauliflower, kale, cabbage, kohlrabi and Brussels sprouts. *B oleracea* varieties are rich in carotenoids (Liu et al, 2014), protein (11.67%) and fiber (3.0%) (Emebu and Anyika, 2011). They also contain glucosinolates (an important phytochemical), which help to boost plant defenses against bacterial and fungal pathogens (Halkier and Gershenzon, 2006) and when consumed by humans as food, also act as anticancer agents (Li *et al.*, 2010). Auxin is one of the earliest found plant hormones which plays a crucial role in plant growth development and regulation. It's one of the key functions in transcription factors is Auxin Response Factor (ARF) which relies on signaling pathway for its functioning (Kou *et al.*, 2022).

In ARF proteins there are three well-conserved domains i.e., DNA-binding domain (DBD), middle region (MR) and C-terminal Phox and Bem1 (PB1). The DBD domain contains the B3 domain which allows the ARFs to bind with DNA motifs known as auxin response elements (AuxREs) (Weijers and Wanger, 2016). In DBD domain multiple segments promote homo and hetero dimerization which enhance the binding preferences of ARFs to tandem repeat AuxREs. ARFs can also activate or repress transcription on chromatin binding via their MR. PB1 domain facilitates the both homo and hetro-oligomerisation at oligomerization site and also interacts with Aux/IAA proteins (Rienstra *et al.*, 2023). As a part of plant hormones, auxins are involved in multiple processes of plant development like apical dominance, root initiation and vascular differentiation (Wang *et al.*, 2007; Xing *et al.*, 2011). From early developmental stages to senescence, auxin plays key role in influencing the plant growth (Shafique *et al.*, 2023). Moreover, at molecular level, auxin triggers transcriptional changes in plant nuclei that lead to the repression or activation of gene sets and auxins also reveal the different developmental programs (Cancé *et al.*, 2022).

In *Arabidopsis thaliana*, ARF1 was the first gene belonging to auxin response factor (ARF) family to be identified using a yeast one-hybrid screen (Ulmasov *et al.*, 1997). Within the *Arabidopsis* genome, 22 ARFs are encoded, with only a limited number of loss-of-function mutants showing noticeable growth phenotypes, while double mutants have revealed redundancy between closely related genes (Cancé *et al.*, 2022). ARF expression in *Arabidopsis* is precisely and dynamically regulated in different developmental stages and tissues. (Ghelli *et al.*, 2018). Similarly, some genes play key roles in regulating plant development by controlling different and partially overlapping groups of targeting genes (Chen *et al.*, 2023).

ARF genes repress/downregulate the auxin/indoleacetic acid (Aux/IAA) pathway in *B. oleracea*. Comparatively, genes related to the IAA pathway are involved during the root hair infection stage in leaves and roots (Zhang *et al.*, 2016). Consequently, increased levels of auxin lead to the transcription of numerous genes that carry out their functions and contribute to the infection process (Oosterbeek *et al.*, 2021). Similarly, salicylic acid (SA) can indirectly influence flowering by altering auxin production and signaling pathways (Yang *et al.*, 2023). *Plasmodiophora brassicae*, the causal organism of clubroot, alters the pathways of defense and stress-related hormones, certainly SA and stops the auxins with the development of disease (Ciaghi *et al.*, 2019), which favors the abnormal growth of the plant.

However, the ARF gene family contains different numbers of gene members in different plant species, for example, 25 in rice (*Oryza sativa*) (Wang *et al.*, 2007), 20 in barley (*Hordeum vulgare*) (Tombuloglu, 2019), 31 in Chinese cabbage (*Brassica rapa* subsp. *Pekinensis*) (Mun *et al.*, 2012), 19 in grape (*Vitis vinifera*) (Wan *et al.*, 2014), 20 in pineapple (*Ananas comosus*) (Su *et al*., 2017), 17 in tomato (*Solanum lycopersicum*) (Kumar *et al*., 2011), 35 in maize (*Zea mays*) (Liu et al., 2011), 19 in sweet orange (*Citrus sinensis*) (Li *et al.*, 2015), 19 in pepper (*Capsicum annuum*) (Zhang *et al.*, 2017) and 17 in physic nut (*Jatropha curcas*) (Tang *et al.*, 2018).

The genome sequence of *B. oleracea*, known for the first time as the draft genome, was published in 2014 (Liu *et al.*, 2014). On the 9 chromosomes of *B. oleracea*, 49 ARF genes were identified. About 4 million years ago *B. oleracea* and *B. rapa* diverged on separate evolutionary paths (Parkin *et al.*, 2014). Evolutions have occurred over time. Several reference genomes of *B oleracea* have been sequenced (Parkin *et al.*, 2014; Liu *et al.*, 2014; Cheng *et al.*, 2016; Yang *et al.*, 2018; Lv et el., 2020(a) and Guo *et al.*, 2021). In 2019, Sun *et al.* (2019) published the draft genome of *B. oleracea*. Similarly, most recently, Wang *et al.*, (2022), published a database of *B. oleracea* named *Brassica oleracea* Genome Database (BoGDB) and a pan-genome. These diverse published genomic data of *B. oleracea* provided reliable experimental resources to facilitate systematic research on the *BolARF* (Luo *et al.*, 2019). Relying on these data, we performed a genome-wide identification using multiple bioinformatics tools and methods (Chen *et al.*, 2023).

MATERIALS AND METHODS

Sequences retrieval and databases search

All the protein sequences of *B. oleracea*, *B. rapa*, *B. napus* and *A. thaliana* were retrieved from Plant Transcription Factor Database (PlantTFDB) (https://planttfdb.gao-lab.org/family.php?sp=Ath&fam=ARF) (Jin *et al.*, 2016). From Phytozome (phytozome-next.jgi.doe.gov) chromosome numbers, protein sizes, gene accession numbers and genomic information of *B. oleracea* were retrieved. PlantTFDB was also used to get isoelectric points (PI) and molecular weights. By using Basic Local Alignment Search Tool (BLASTP) on National Center for Biotechnology and Information (NCBI) (https://ncbi.nlm.nih.gov/), the putative of *B. oleracea* was validated and whole genomic sequence was also downloaded from NCBI (O'Leary *et al.*, 2016).

Mapping of chromosomes, intron/exon distribution and conserved protein motifs analyses

In *B. oleracea*, the positions of ARF family genes were obtained from NCBI and their mapping on chromosomes was done by using Map Chart software (v.2.32) (Voorrips, 2002). The distribution of intron/exon of *B. oleracea* and *A. thaliana* was carried out by utilizing Gene Structure Display Server (GSDS) (https://gsds.gao-lab.org/) (Zahoor *et al.*, 2023). MEME (Multiple Em for Motif Elicitation; v5.03) was utilized to locate conserved protein motifs inside the ARF proteins of *B. oleracea* and *A. thaliana* (Bailey *et al.*, 2009). The analysis was conducted by utilizing default parameters i.e., the occurrence of a motif was limited to 0 or 1 per sequence, the optimum width of motifs ranged from 10 to 55 residues, the motifs count was set to 50 and the minimum number of sites for a motif was selected as 5.

Phylogenetic and conserved domain analyses of ARF proteins from *B. oleracea, B. napus, B. rapa* in comparison with *A. thaliana*

The putative of ARF transcription factor (TF) protein sequences from *B. oleracea*, *B. napus*, *B. rapa* and *A. thaliana* were used to construct a phylogenetic tree as shown in Figure. 2(a). All protein sequences were aligned using ClustalW (Thompson *et al.*, 2003), default options (in both pairwise and multiple alignments, gap opening penalties were the same, i.e. 10.00). However, the gap extension penalty was 0.10 in the pairwise alignment and 0.20 in the multiple alignment. The use of negative matrix was switched off and delay divergence cutoff was 30%) were used (Mun *et al.*, 2012).

The phylogenetic tree was constructed using the aligned amino acid sequences using Neighbour Joining (NJ) method (Tamura et al, 2011) with bootstrap method and the number of replications was 1000. The Poisson model was used with partial deletions with 95% site coverage cutoff in Molecular Evolutionary Genetic Analysis software MEGA (version 11) software (Tamura *et al.*, 2021). A single tree was generated with the non-redundant 17 genes of *B. oleracea*, *B. napus* with (all) 64 genes and *B. rapa* with 31 genes compared to *A. thaliana* with 22 amino acid sequences of the ARF family from PlantTFDB to understand the phylogenetic relationship between the same gene family of different species. Further modifications and visualization were done in iTOL (Letunic and Bork, 2021) (https://itol.embl.de/tree/).

The second tree (Figure 2b) was constructed to understand the evolutionary relationship between *B. oleracea* and *A. thaliana* using their amino acid sequences with the same (default) parameters mentioned above using the Neighbour Joining (NJ) method. In order to determine the existence of conserved domains in ARF transcription factors in *A. thaliana*, *B. oleracea, B. napus* and *B. rapa*, subsequent analysis phases were performed to visualize the domains. For this analysis, we used the NCBI Conserved Domain Database (CDD) (https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cg) (Marchler-Bauer *et al.*, 2015) to determine the specific domains present in *BolARF* and *AtARF* proteins.

Comparative synteny gene analysis of *B oleracea* and *A thaliana*

The synteny analysis was carried out by using Circoletto (Darzentas, 2010) web server (http://tools.bat.infspire.org/circoletto/), among the *B. oleracea* and *A. thaliana* to identify evolutionary relationship/similarities among them (Restrepo-Montoya *et al.*, 2021).

Cis-regulatory elements identification in promoter region of ARF genes

To identify the cis-regulatory elements in *B. oleracea*, promoter region of 1kb size was taken manually from phytozome for every single gene of *BolARF* (Brown *et al.*, 2015). For *in silico* analysis of promoter sequences to find cis-regulatory elements in *BolARF* genes PlantCare Database (https://bioinformatics.psb.ugent.be/webtools/plantcare/html/) was used, after retrieving the 1000bps of promoter regions (Lescot *et al.*, 2002). To generate heatmap all cis-regulatory elements data was noted for each gene. The data was visualized as a heatmap by using Tbtools (Chen *et al.*, 2020).

Expression analysis of ARF genes based on publicly available data

The transcriptomic data for root gall stress (ERP108422) was obtained from the NCBI SRA website (https://www.ncbi.nlm.nih.gov/Traces/study/?acc=ERP108422&o=acc_s%3Aa) (Leinonen *et al.*, 2010). In the mentioned bio project, Ciaghi *et al.,* (2019) observed the transcriptomic response of symptomless kohlrabi roots infected with clubroot disease and noted the up/down regulation of Salicylic Acid (SA). The reads were mapped using the reference genome (NCBI RefSeq assembly: GCF_000695525.1) of *B. oleracea* with the help of Galaxy Bowtie2, using default parameters (https://usegalaxy.org/) (De Almeida *et al.*, 2016).

To estimate the level of gene expression, FeatureCounts was used to compute FPKM values from the RNA-Seq data (Kumar *et al.*, 2021). The FPKM values from the output files were used to compare the expression of genes between the control and root gall stress conditions (Rahman *et al.*, 2023; Waqas *et al.*, 2019). Tbtools were used to create an expression profile and generated a list of resulting genes, as shown in Figure 8.

RESULTS

Identification and distribution of ARF Auxin encoding genes in B oleracea genome:

Using the complete genome assembly of *B. oleracea*, we identified 49 genes that encode putative auxin in ARF family, which is the same as the number of proteins found in *A. thaliana*. After excluding 32 ARF genes due to redundancy, we further analyzed the remaining 17 non-redundant ARF genes. These 17 ARF genes were then named BoI in ascending order, aligning with their respective positions in the chromosomal mapping (refer to Table 1 and Figure 1). The nomenclature used in this text follows the convention observed in other species of Brassica and *A. thaliana*, based on their positions on their respective chromosomes.

Table 1 provides detailed information on gene sequences, including locus IDs, transcripts, chromosome numbers, protein lengths, domain start and end positions, molecular weight and isoelectric points. However, the length of ARF gene peptides vary from 537 to 1052, while the isoelectric point ranges from 5.1049 to 8.5801. Similarly, the molecular weight ranges from 59.72 to 116.62 kDa.

Genomic Localization and Mapping of ARF genes in *B. oleracea* by using MapChart

The ARF genes from *B. oleracea* were mapped on nine different chromosomes, as shown in Figure 1. Our results indicate that *BolARF*s is present on chromosomes one, three, and four, while the maximum number of genes (three) were found on chromosomes five and eight. The remaining chromosomes (two, six, seven, and nine) each carried two genes. *BolARF01*, *BolARF09*, and *BolARF16* were located at the beginning of chromosome arms 1, 6, and 9, respectively. *BolARF06*, *BolARF07* and *BolARF08* were closely located on chromosome 5 while *BolARF13* and *BolARF*14 were found on chromosome 8.

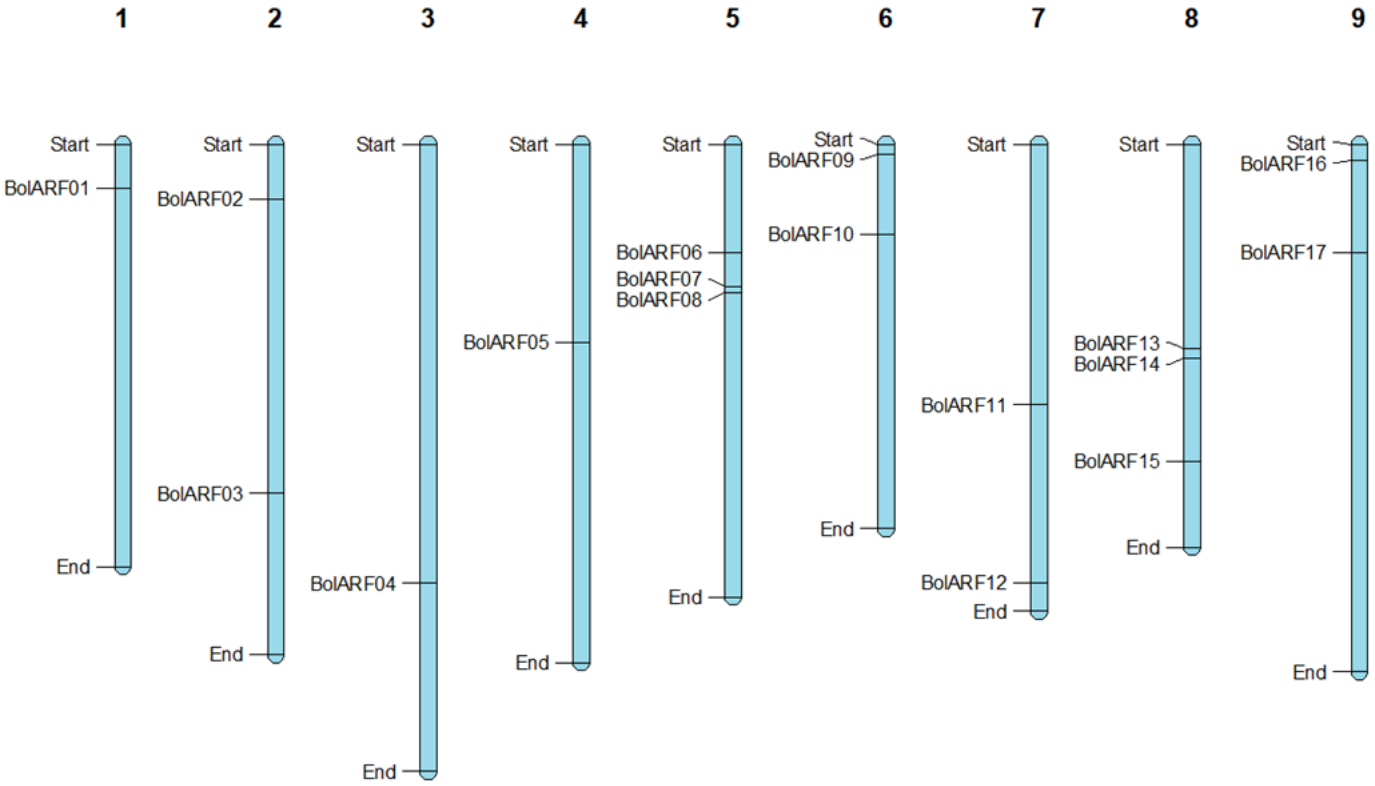

Figure 1. Chromosomal distribution of ARF genes in *Brassica oleracea*. The black lines on chromosomes represent the distributed locations of ARF genes.

The ARF genes were primarily involved in three enrichment processes: Biological processes, Cellular Components, and Molecular Functions (Rahman *et al.*, 2023). All genes were localized within the nucleus in cellular components. Almost all genes were involved in the biological regulation of transcription, DNA-templated, and response to hormones. Furthermore, *BolARF04* was biologically involved in vegetative phase change and abaxial cell fate specification. *BolARF06* is involved in lateral root formation and leaf development in response to ethylene and auxin. *BolARF07* plays a role in flower and root development, leaf vascular tissue pattern formation and meristem development. *BolARF08* and *BolARF09* are also involved in flower development. In addition, *BolARF10* is involved in pollen wall assembly, adventitious root development, and callose deposition in cell walls. *BolARF11* is involved in pattern

specification and regulation of anthocyanin biosynthesis, as well as abscisic acid-activated and auxin-activated signaling pathways, cell division, and development of leaves, sepals and petals. Additionally, it plays a role in molecular functions such as DNA and protein binding.

Table 1. Detailed properties of identified *Brassica oleracea* ARF genes.

Locus ID	Transcript	Chr	Protein name	Start	End	Protein size	PI	M wt. (Da)	Protein
LOC106302756	Bol020966	C01	*BolARF01*	4555098	4557737 f	651	7.7272	71906.4	>XP_013594669.1
LOC106317273	Bol020075	CO2	*BolARF02*	36190281	36194013f	556	6.6944	63456.6	>XP_013610570.1
LOC106318259	Bol036113	CO2	*BolARF03*	5681886	5686418r	829	6.7988	92567.6	>XP_013611610.1
LOC106319097	Bol016113	C03	*BolARF04*	45509343.	45513501r	769	6.7165	85178.7	>XP_013612783.1
LOC106336983	Bol027304	C04	*BolARF05*	20533557	20536202r	594	7.1997	65639.3	>XP_013631443.1
LOC106295312	Bol026834	C05	*BolARF06*	14709968	14714184r	1052	6.653	116627	>XP_013586633.1
LOC106306798	Bol026901	C05	*BolARF07*	15370814	15374692	860	5.9517	95380.5	>XP_013599006.1
LOC106294993	Bol022542	C05	*BolARF08*	11180409	11184976f	900	6.2589	99001.8	>XP_013586187.1
LOC106342431	Bol012563	C06	*BolARF09*	9280028	9283836	820	6.399	90502.6	>XP_013636814.1
LOC106296377	Bol027532	C06	*BolARF10*	1036794	1039018	543	5.1049	59727.8	>XP_013587958.1
LOC106342709	Bol033635	C07	*BolARF11*	45446555	45448816	703	7.0142	77829.8	>XP_013637173.1
LOC106340796	Bol006390	C07	*BolARF12*	26973769	26975858	570	6.6944	63456.6	>XP_013635089.1
LOC106312216	Bol013157	C08	*BolARF13*	21174903	21179165r	1020	6.6471	113085	>XP_013605108.1
LOC106311742	Bol023586	C08	*BolARF14*	22166616	22170045	846	5.7474	93808.7	>XP_013604468.1
LOC106320938	Bol045672	C08	*BolARF15*	32880562	32882472	555	6.5212	62529.3	>XP_013614735.1
LOC106316816	Bol032614	C09	*BolARF16*	11212844	11215899	537	8.5801	61450.2	>XP_013610137.1
LOC106317306	Bol032178	C09	*BolARF17*	1727784	1731887	541	7.6442	62101.9	>XP_013610601.1

Comparative phylogenetic tree of *B. oleracea, B. rapa, B. napus* and *A. thaliana* and conserved domain analysis:

To predict the evolutionary lineage, we conducted a phylogenetic analysis of three different Brassica species (oleracea, napus, and rapa) and *A. thaliana* (see Figure 2a). The proteins were named according to their chromosomal positions, except for *B. napus* proteins, for which Li *et al.*, (2020) nomenclature was used. The 134 ARF proteins were classified into four groups (I, II, III, and IV) based on the presence of conserved domains (Figure 2a). Group I had 46 members, Group II had 81 members, Group III had 5 members and Group IV had only 2 members. Group I represented the presence of two conserved domains, Auxin_resp and B3, which were present in the DNA binding region. Moreover, it contained the flexible middle segment which is responsible for actvating or repressing functions and a carboxy-terminal region essentional for dimerization (Finet *et al.*, 2013). Within Group II there were three conserved domains i.e., B3, Auxin_resp and AUX_IAA superfamily.

AUX_IAA superfamily involves in the repression of activated ARF genes' expression (Luo *et al.*, 2018). Group III is containing Auxin_resp and B3 domains but the absence of Auxin_resp domain was observed in *BnARFun2a*. Aditionally, *AtARF11* showed the presence of AUX_IAA superfamily and PHA03247 superfamily domain along B3 and Auxin_resp domains. Auxin_resp and B3 domains are present in *BnARFun2b* along Zf-RVT superfamily domain while *BrARF29* shows the presence of PTZ00449 superfamily with B3, Auxin_resp and AUXIN_IAA superfamily domains. The only two members of group IV contain only the B3_DNA and Auxin_resp superfamily domains. Figure 2b shows the tree for *BolARF* and *AtARF* proteins. This tree has three groups based on conserved domains, as described in Figure 1a's phylogenetic tree. Group I have the most ARF proteins, with a total of 24 out of 39 proteins, making it the largest group. Group II has only 14 proteins, while Group III has only one protein.

Conserved domains prediction was performed on all ARFs of *BolARF*, *BnARF*, *BrARF* and *AtARF* using NCBI CDD. Tbtool was then employed for further visualization, as shown in Figure 3. Group I of the phylogenetic tree contains only the conserved domains Auxin_resp and B3, while group II genes additionally contain the AUX_IAA superfamily with Auxin-resp and B3 domains. Group III genes possess different domains. For example, *AtARF*11 has an additional PHA03247 superfamily along with Auxin_resp, B3, and AUX_IAA superfamily. *BnARFun2a* shows the presence of Auxin_resp superfamily with B3 domain. *BnARFun2b* possesses an additional zf-RVT superfamily with Auxin_resp and B3. *BrARF29* significantly carries PTZ00449 superfamily, along with the three most conserved domains. In group III, *BnARF16c* was found to have the P-loop_NTPase superfamily domain along with the Auxin_resp and B3 domains. Similarly, group IV genes also had the Auxin_resp superfamily and B3_DNA domains.

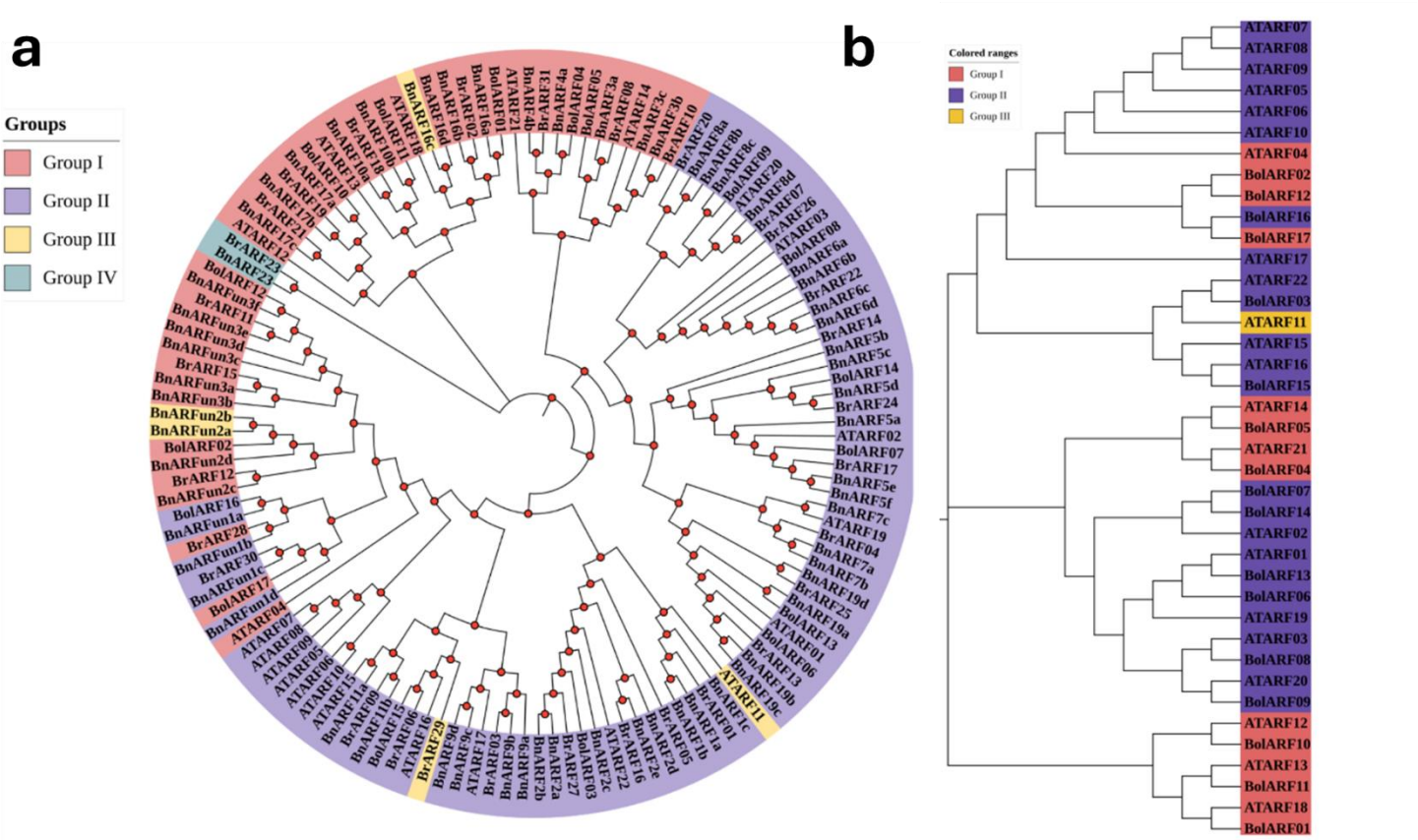

Figure 2. (a) Neighbor Joining (NJ) phylogenetic tree of 134 ARF proteins of three different species of Brassica with *A. thaliana*. The developed evolutionary relationship is grouped in 4 groups and displayed with different colors. (b) Neighbor Joining (NJ) phylogenetic tree of 39 ARF proteins of *B. oleracea* with *A thaliana*. The developed evolutionary relationship is grouped in 3 groups and displayed with different colors.

Gene structure and Conserved Motifs Investigations of *BolARF*s and *AtARF*s:

The gene structure, including the distribution of introns and exons, was analyzed to gain further insights into the evaluation of the ARF gene family in *B. oleracea* and *A. thaliana*. By scrutinizing the gene structure of 17 ARF genes in *B. oleracea* and 22 ARF genes in *A. thaliana*, we identified the positions of introns and exons within their genomic sequences (see Figure 4). The number of introns varies from one to thirteen in all *BolARF* and *AtARF* genes. Group I genes have fewer introns, while Group II and III genes have more. All groups contain genes with UTRs, represented by blue in the legend of Figure 4. All groups of motifs in *BolARF* and *AtARF* were conserved as shown in Figure 5). The most conserved motif was motif 4 (ANRQQTNIPSSVISSDSMHHGVLAAAAHAIATNSMFTVFYKPRAS) and least conserved was motif 3 (MNNKFSVGMRFRMRFETEDSSERRYMGTI).

Synteny analysis

The conservation relationship between the ARF gene family in *B. oleracea* and *A. thaliana* was found (Errum *et al.*, 2021). *A. thaliana* was used as a query against ARF family genes to obtain the ideograms of *B. oleracea* (Darzentas, 2010). Synteny was observed among *AtARF01*, *BolARF06*, and *BolARF07* as they were similar in the regulation of DNA-templated transcription process, while *BolARF08* showed synteny with *AtARF02* and *AtARF03* and is involved in the response to hormone process, as shown in Figure 6. Consequently, *BolARF09* was found to have a syntenic relation with *ATARF20*, as both enable DNA and protein binding. Similarly, *BolARF04* was found to have a relation with *AtARF21*, as both are involved in enabling DNA-binding transcription activity. *BolARF03* was found to have a relation with *AtARF22*, as both are specifically located in the nucleus and involved in the response to hormones, enabling DNA and protein binding. *BolARF06* and *BolARF07* were found to be conserved in evolution, as were *BolARF08* with *AtARF02* and *AtARF03*. Multiple replication events have occurred due to more than one pair. However,

Figure 3. Conserved domain regions of *BolARF*, *BnARF*, *BrARF* and *AtARF*s. Yellow, green and pink colors conserved domains represent the B3, Auxin_resp and AUX_IAA superfamily respectively.

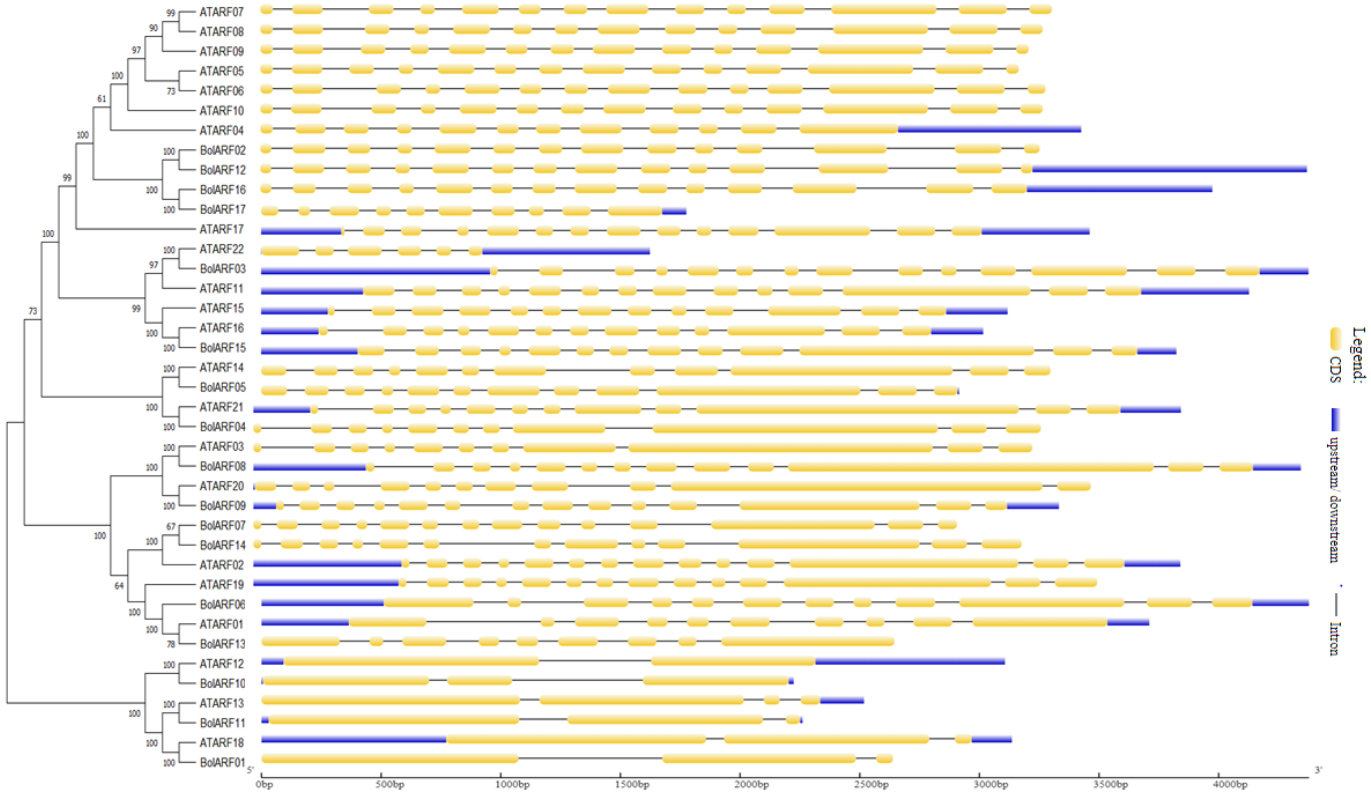

Figure 4. Phylogenetic tree-based clustering and structure of *BolARF*s and *AtARF*s genes. We created an un-rooted phylogenetic tree which was based on the full-length peptide sequences of *BolARF*s and *AtARF*s with 1000 bootstraps. The yellow, blue and grey colors are representing CDS, up/down streams and introns respectively.

Figure 5. Conserved motifs and phylogenetic analyses of *BolARF* and *AtARF*. Representative colors of motif sequences are at the basal of the figure. In length taller the motif is more significantly conserved then the shorter one.

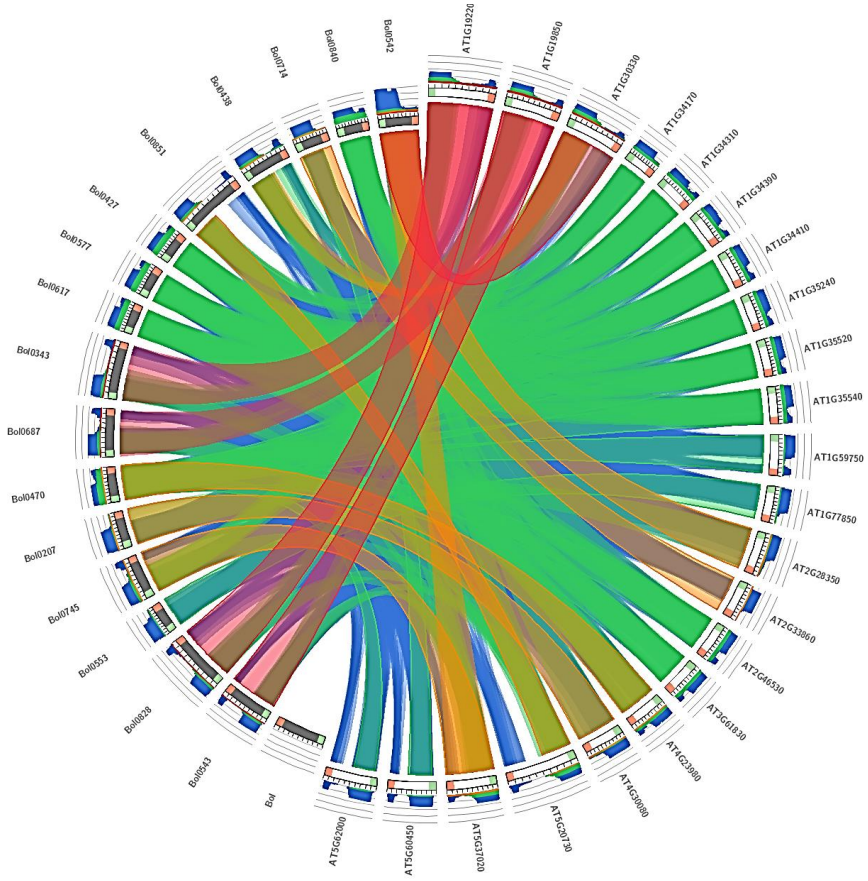

Figure 6. The syntenic regions in selected *BolARF*s proteins by taking the *AtARF*s proteins as the query sequence. Blue, green, orange and red shows 50, 75, 99.99 and 100% similarity respectively.

all the remaining pairs were less than one, suggesting significant purifying selection following replication (Chen *et al.*, 2023).

Cis-regulatory elements identification in *BolARF*s

All *BolARF* promoters contain TGTCTC auxin response elements (ARE), in which conjunction with Aux/IAA repressors, form dimers with ARF activators in a manner regulated by auxin (Guilfoyle and Hagen, 2007). The heatmap in Figure 7 visualizes only 11 of the 43 cis-regulatory elements found. The ABRE cis-acting element was significantly observed in the *BolARF01* gene, indicating its involvement in abscisic acid responsiveness. *BolARF02* actively showed the presence of the GT-1 motif, making it a light responsive element, more so than *BolARF17* and 05. In Figure 7, Box4 was found to be repeatedly present in multiple genes, including *BolARF17*, 08, 16, and 07, which constitute as a conserved DNA module part, involved in light responsiveness. Additionally, MBS was present as a MYB binding site involved in drought-inducibility in *BolARF03*, 04, and 09. Table 2 contains all the sites with their sequences and functions.

***BolARF* expression in *B. oleracea* roots infected with root galls**

Plasmodiophora brassicae causes clubroot disease, which reduces the quality and yield of Brassicaceae crops worldwide. The formation of galls in the roots disrupts the uptake of water and nutrients, resulting in stunted growth and wilting of the host plant's aboveground parts (Ning *et al.*, 2019). However, *P. brassicae's* effector BA/SA-methyltransferase can reduce salicylic acid (SA) levels and promote pathogen colonization (Djavaheri *et al.*, 2019). We analyzed gene expression data in *B. oleracea* using New Generation Sequencing (NGS). The bio project PRJEB26435 had a total of 10 runs, and 16.76 GB of experimental data was uploaded to the Sequence Read Archive (SRA) with accession number ERP108422. Experimental data was used to generate FPKM values through the Galaxy server-based tool. FPKM values were calculated for all *BolARF* genes and a heatmap was created using Tbtools (Waqas *et al.*, 2020).

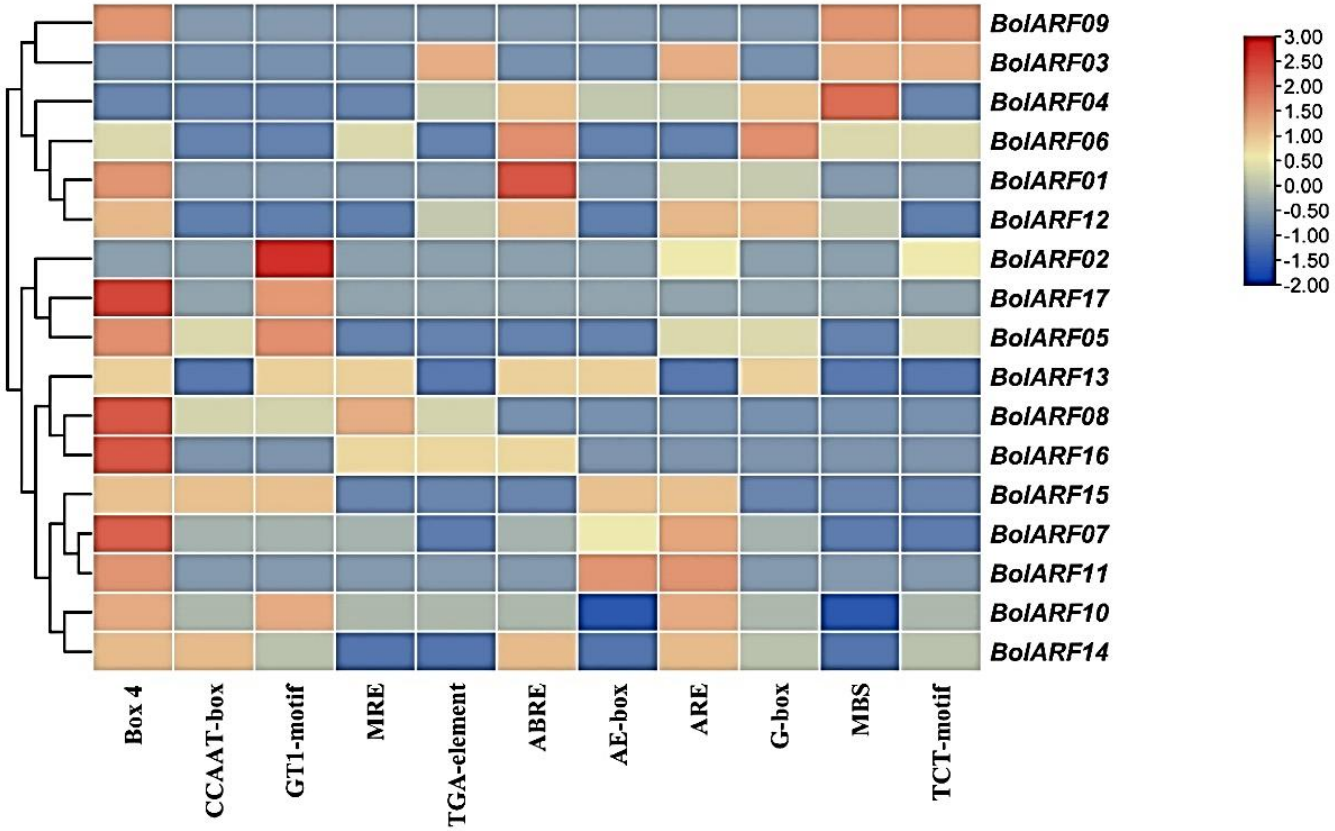

Figure 7. The heatmap showing the presence or absence of cis-regulatory elements on upstream promotor region of a gene.

Table 2. All the sites of cis-elements with their sequences and functions.

Cis-elements	Sequences	Functions
Box 4	ATTAAT	A conserved DNA module part involved in light responsiveness
CCAAT-box	CAACGG	MYBHv1 binding site
GT1-motif	GGTTAA	Element responds to light
MRE	AACCTAA	MYB binding site responds to light
TGA-element	AACGAC	Auxin-responsive element
ABRE	ACGTG	Cis-acting element involved in the abscisic acid responsiveness
AE-box	AGAAACAA	Part of a module for light response
ARE	AAACCA	Cis-acting regulatory element essential for the anaerobic induction
G-box	CACGTT	Cis-acting regulatory element responds to light
MBS	CAACTG	MYB binding site involved in drought-inducibility
TCT-motif	TCTTAC	Part of a light responsive element

It is noteworthy that *BolARF09* (Figure 8) gene upregulated the SA in young root gall from group II and showed no response in older root gall and symptomless root. In older root gall, several genes from Group I (*BolARF01*, 04, 08, and 10) showed significant downregulation of SA, while same genes in symptomless roots showed upregulation of SA, except for *BolARF08*, which did not show significant up/down regulation. In young root gall, the same genes did not show any particular response to SA, except for minor downregulation in *BolARF01*, 03, and 10. *BolARF02*, 12, and 17 from Group I and *BolARF16* from Group II did not respond to SA in any experiment. Interestingly, there was a lack of control observed. The remaining genes, *BolARF05* and 11 from Group I and *BolARF06* from Group II, upregulated the response of SA in older root gall, while in symptomless root, they were all downregulated. In young root gall, *BolARF11* did not show any specific regulation, while *BolARF05* and 06 were slightly upregulated. The factors are discussed in comparison to the control parameters.

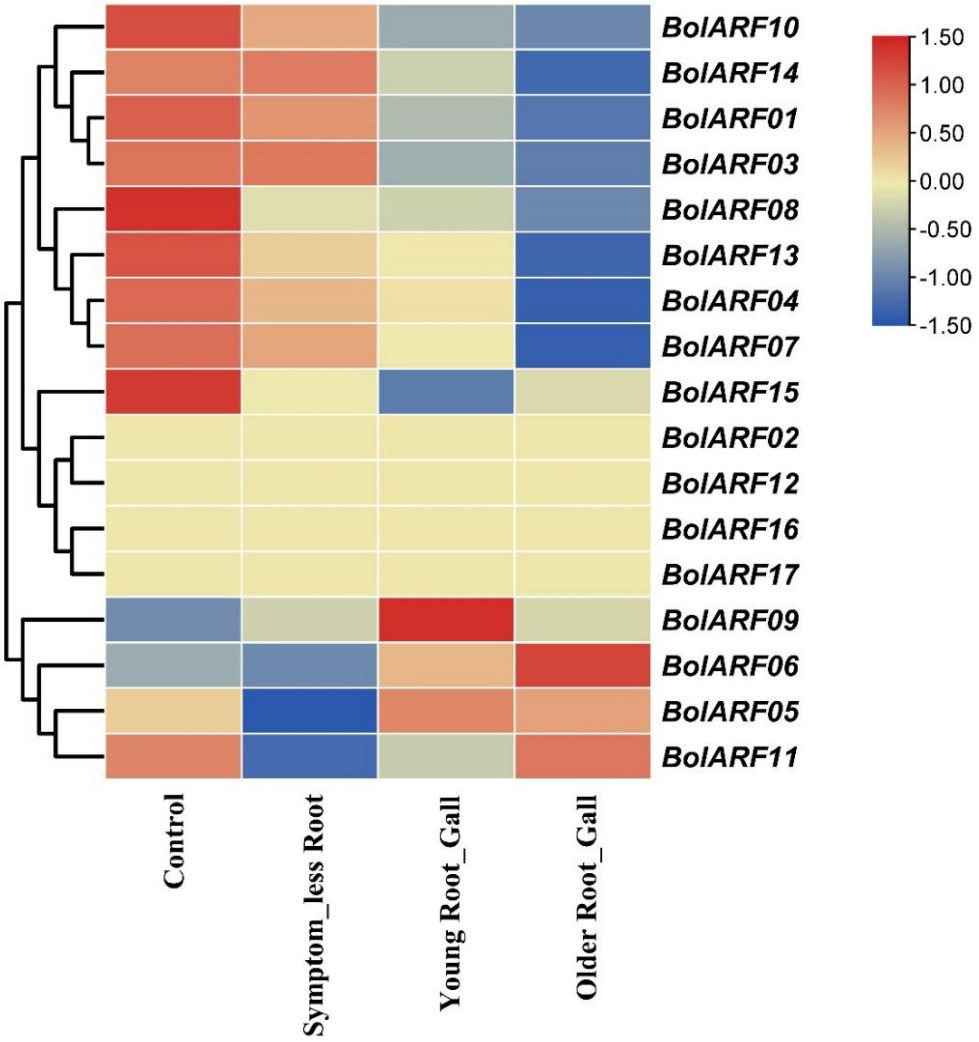

Figure 8. ARF genes expression profiling against clubroot disease. Every single row is representing the genes while each column is representing an experiment (either stress or control) in the above heatmap. The heatmap was generated using Tbtools software, with log2-transformed values representing the relative expression levels of *BolARF*s.

DISCUSSION

The presence of multiple members in a specific gene family arises from prolonged periods of evolution experienced by a particular microorganism (Wang *et al.*, 2007; Danilevskaya *et al.*, 2008). The Brassicaceae family contains highly diverse species which includes nutrients rich vegetables such as cauliflower, cabbage, kale, kohlrabi and brussels sprouts (Yu *et al.*, 2013). This research layouts the strong foundation for further investigations on the regulatory functions of *BolARF* proteins against root gall disease. It outlines the makeup of the ARF gene family in *B. oleracea* genome (Song *et al.*, 2014).

This research identified the 17 ARF genes in *B. oleracea* and 22 in *A. thaliana*, alongside 31 in *B. rapa* and 64 in *B. napus*. These genes were characterized, and phylogenetic analysis revealed four groups i.e., Group I, II, III and IV. Group II contained the most members (81), trailed by Group I with 46 members and Group III comprised only 5 members. However, Group IV contained only 2 genes. The presence of duplicated *BolARF* and *AtARF* genes raised questions about their functional redundancy (Xing *et al.*, 2011). We identified overlapping genes and examined their chromosomal distribution. On the generated map, closely related loci such as *BolARF06*, *BolARF07*, and *BolARF08* are closely linked on chromosome 5. Similarly, *BolARF13* and *BolARF14* are closely placed on chromosome 8, but *BolARF15* is located at a greater distance on the same chromosome. The gene structure, based on phylogenetics, includes introns, conserved domains, and up/downstream regions. We found 1 to 13 introns present in every *BolARF* and *AtARF* genes. However, up/downstream regions were absent in *AtARF03*, 05-10, 14, and 20. In *BolARF* genes,

the absence of upstream regions was identified in *BolARF01*, 02, 04, 05, 07, 13, and 14. Conserved motif regions were identified in every *BolARF* and *AtARF* protein. Motif 4 was the most conserved, while motif 3 was least conserved.

Studying *B. oleracea* has the potential to provide new insights into the morphology evolution, complementing and expanding existing knowledge from *Arabidopsis* (Bowman, 2006; Wang *et al.*, 2011). Cis-regulatory sequences, such as enhancers and promoters, regulate gene expression, influencing development and physiology. Recent research aims to uncover the genetic and molecular mechanisms behind regulatory divergence, which has been linked to phenotypic evolution in numerous case studies (Wittkopp and Kalay, 2012). The TATA box (TATATAA), a core promoter element in transcription initiation, is present in all genes. Box 4 (ATTAAT) is expressed in the maximum number of genes, including *BolARF01*, 05, 07, 08, 09, 11, 16, and 17, compared to any other regulatory elements. The cis-elements GT1 motif (GGTTAAT), Box 4 (ATTAAT), MRE (AACCTAA), AE-box (AGAAACAA), G-box (TACGTG), and TCT-motif (TCTTAC) are all parts of a conserved DNA module involved in light responsiveness. Other cis-elements, such as the ABRE (ACGTG) involved in abscisic acid responsiveness, TGA-element (AACGAC) auxin-responsive element, ARE (AAACCA) essential for anaerobic induction, and MBS (CAACTG) involved in drought-inducibility, act for different stresses or responses. The cis elements were differentially up/down regulated in various genes.

One of the main objectives in comparative genomics is to identify syntenic regions (Sinha and Meller, 2007). Synteny provides a framework for recognizing the preservation of similar genes and their sequential arrangements across the genomes of different species (Liu and Tsai, 2018). Our analysis revealed distinct results in terms of syntenic relations and similarities. The most similar regions generated were *AtARF01* with *BolARF06*, *07*, and *08*, and *BolARF13* also showed a syntenic relation with *AtARF*01 and 02. We performed a conserved domain analysis among all *BolARF*, *BrARF*, *BnARF*, and *AtARF* genes by using protein sequences. We identified nine different conserved responses, from which Auxin_resp and B3 were highly conserved in all proteins of group I. Group II contains an additional third domain, AUX_IAA superfamily, with Auxin_resp and B3 domains. Group III contains all domains except B3_DNA, which includes Auxin_resp, B3, AUX_IAA superfamily, PHA03247 superfamily, zf-RVT superfamily, PTZ00449 superfamily, and P-loop_NTPase superfamily. Group IV possessed only two domains: Auxin_resp superfamily and B3_DNA on both genes.

B. oleracea is susceptible to various diseases, including white rust, downy mildew, *Alternaria* leaf spot, powdery mildew, and clubroot, with clubroot, downy mildew and white rust being the most destructive (Lv *et al.*, 2020). Clubroot is characterized by roots swelling resulting from changes in auxin, salicylic acid (SA) metabolism, and cytokinin metabolism, while downy mildew and white rust also cause significant damage (Mourou *et al.*, 2023). *P brassicae* disrupts the levels of auxin and cytokinin (Ludwig-Muller *et al.*, 2009) and salicylic acid (Manoharan *et al.*, 2016), leading to clubroot formation through the transport of IAA (Paesold *et al.*, 2010; Devos, 2005). Expression analysis of ARF transcription factors in roots revealed three categories: Symptomless roots (SL), Young root gall (YG) and older root gall (OG). Notably, *BolARF09* was upregulated in YG and *BolARF06* was upregulated in OG. As a result of increased auxin levels, gall formation occurs in roots (Yuan *et al.*, 2021). *BolARF05*, 06, and 11 were downregulated in SL, while *BolARF02*, 12, 16, and 17 did not respond to any disease factors. *BolARF01*, *03, 04, 07, 08, 10, 13,* and *14* were all downregulated in OG. *BolARF03* and 14 showed slight upregulation. The time-dependent regulation of these clusters highlights the importance of ARFs in coordinating the plant's response to biotic stress at specific stages (Rahman *et al.*, 2023). These findings offer valuable insights for improving strategies and advancing research in crops development.

ACKNOWLEDGEMENTS

The authors are thankful to HEC and the department of Plant Pathology, University of Agriculture Faisalabad for support during this study.

AUTHOR CONTRIBUTIONS

IS mainly conducted the research, ME and UR assisted in data recording and helped in data analysis, MA reviewed the manuscript, MUS supervised the study.

COMPETING OF INTEREST

The authors declare no competing interests.

REFERENCES

Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., & Walter, P. (2002). The structure and function of DNA. In Molecular Biology of the Cell (4th ed.). Garland Science.

Al-Shehbaz, I. A., Beilstein, M. A., & Kellogg, E. A. (2006). Systematics and phylogeny of the Brassicaceae (Cruciferae): An overview. Plant Systematics and Evolution, 259, 89-120.

Bailey, T. L., Boden, M., Buske, F. A., Frith, M., Grant, C. E., Clementi, L., & Noble, W. S. (2009). MEME SUITE: tools for motif discovery and searching. Nucleic acids research, 37(suppl_2), W202-W208.

Bowman, J. L. (2006). Molecules and morphology: Comparative developmental genetics of the Brassicaceae. Plant Systematics and Evolution, 259, 199-215.

Brown, G. R., Hem, V., Katz, K. S., Ovetsky, M., Wallin, C., Ermolaeva, O., & Murphy, T. D. (2015). Gene: A gene-centered information resource at NCBI. Nucleic Acids Research, 43(D1), D36-D42.

Cancé, C., Martin-Arevalillo, R., Boubekeur, K., & Dumas, R. (2022). Auxin response factors are keys to the many auxin doors. New Phytologist, 235, 402-419.

Chen, C., Chen, H., Zhang, Y., Thomas, H. R., Frank, M. H., He, Y., & Xia, R. (2020). TBtools: An integrative toolkit developed for interactive analyses of big biological data. Molecular Plant, 13, 1194-1202.

Chen, F., Zhang, J., Ha, X., & Ma, H. (2023). Genome-wide identification and expression analysis of the Auxin-Response factor (ARF) gene family in *Medicago sativa* under abiotic stress. BMC Genomics, 24, 498.

Cheng, F., Wu, J., Cai, C., Fu, L., Liang, J., Borm, T., & Wang, X. (2016). Genome resequencing and comparative variome analysis in a *Brassica rapa* and *Brassica oleracea* collection. Scientific Data, 3, 1-9.

Ciaghi, S., Schwelm, A., & Neuhauser, S. (2019). Transcriptomic response in symptomless roots of clubroot infected kohlrabi (*Brassica oleracea* var. *gongylodes*) mirrors resistant plants. BMC Plant Biology, 19, 1-14.

Danilevskaya, O. N., Meng, X., Hou, Z., Ananiev, E. V., & Simmons, C. R. (2008). A genomic and expression compendium of the expanded PEBP gene family from maize. Plant Physiology, 146, 250-264.

Darzentas, N. (2010). Circoletto: Visualizing sequence similarity with Circos. Bioinformatics, 26(20).

De Almeida, Á. O., Piattella, O. F., & Rodrigues, D. C. (2016). A method for evaluating models that use galaxy rotation curves to derive the density profiles. Monthly Notices of the Royal Astronomical Society, 462, 2706-2714.

Devos, S. (2005). Plant hormones: A key in clubroot development. Universiteit Antwerpen (Belgium).

Djavaheri, M., Ma, L., Klessig, D. F., Mithöfer, A., Gropp, G., & Borhan, H. (2019). Mimicking the host regulation of salicylic acid: A virulence strategy by the clubroot pathogen *Plasmodiophora brassicae*. Molecular Plant-Microbe Interactions, 32, 296-305.

Emebu, P. K., & Anyika, J. U. (2011). Proximate and mineral composition of kale (*Brassica oleracea*) grown in Delta State, Nigeria. Pakistan Journal of Nutrition.

Errum, A., Rehman, N., Khan, M. R., & Ali, G. M. (2021). Genome-wide characterization and expression analysis of pseudo-response regulator gene family in wheat. Molecular Biology Reports, 48, 2411-2427.

Finet, C., Berne-Dedieu, A., Scutt, C. P., & Marlétaz, F. (2013). Evolution of the ARF gene family in land plants: Old domains, new tricks. Molecular Biology and Evolution, 30, 45-56.

Ghelli, R., Brunetti, P., Napoli, N., De Paolis, A., Cecchetti, V., Tsuge, T., & Cardarelli, M. (2018). A newly identified flower-specific splice variant of AUXIN RESPONSE FACTOR8 regulates stamen elongation and endothecium lignification in *Arabidopsis*. The Plant Cell, 30, 620-637.

Guilfoyle, T. J., & Hagen, G. (2007). Auxin response factors. Current Opinion in Plant Biology, 10, 453-460.

Guo, N., Wang, S., Gao, L., Liu, Y., Wang, X., Lai, E., & Liu, F. (2021). Genome sequencing sheds light on the contribution of structural variants to *Brassica oleracea* diversification. BMC Biology, 19, 1-15.

Halkier, B. A., & Gershenzon, J. (2006). Biology and biochemistry of glucosinolates. Annual Review of Plant Biology, 57, 303-333.

Jin, J., Tian, F., Yang, D. C., Meng, Y. Q., Kong, L., Luo, J., & Gao, G. (2016). PlantTFDB 4.0: Toward a central hub for transcription factors and regulatory interactions in plants. Nucleic Acids Research, gkw982.

Katche, E., Quezada-Martinez, D., Katche, E. I., Vasquez-Teuber, P., & Mason, A. S. (2019). Interspecific hybridization for Brassica crop improvement. Crop Breeding, Genetics and Genomics, 1.

Kou, X., Zhao, X., Wu, B., Wang, C., Wu, C., Yang, S., & Xue, Z. (2022). Auxin response factors are ubiquitous in plant growth and development and involved in crosstalk between plant hormones: A review. Applied Sciences, 12(3), 1360.

Kumar, A., Rasche, H., Grüning, B., & Backofen, R. (2021). Tool recommender system in Galaxy using deep learning. GigaScience, 10, giaa152.

Kumar, R., Tyagi, A. K., & Sharma, A. K. (2011). Genome-wide analysis of auxin response factor (ARF) gene family from tomato and analysis of their role in flower and fruit development. Molecular Genetics and Genomics, 285, 245-260.

Leinonen, R., Sugawara, H., Shumway, M., & International Nucleotide Sequence Database Collaboration. (2010). The sequence read archive. Nucleic Acids Research, 39, 19-21.

Lescot, M., Déhais, P., Thijs, G., Marchal, K., Moreau, Y., Van de Peer, Y., & Rombauts, S. (2002). PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. Nucleic Acids Research, 30, 325-327.

Letunic, I., & Bork, P. (2021). Interactive Tree of Life (iTOL) v5: An online tool for phylogenetic tree display and annotation. Nucleic Acids Research, 49, 293-296.

Lewin, B. (Ed.). (2007). Cells. Jones & Bartlett Learning.

Li, S. B., OuYang, W. Z., Hou, X. J., Xie, L. L., Hu, C. G., & Zhang, J. Z. (2015). Genome-wide identification, isolation and expression analysis of auxin response factor (ARF) gene family in sweet orange (*Citrus sinensis*). Frontiers in Plant Science, 6, 119.

Li, W., Chen, F., Wang, Y., Zheng, H., Yi, Q., Ren, Y., & Gao, J. (2020). Genome-wide identification and functional analysis of ARF transcription factors in *Brassica juncea* var. *tumida*. PLoS ONE, 15, e0232039.

Li, Y., Zhang, T., Korkaya, H., Liu, S., Lee, H. F., Newman, B., & Sun, D. (2010). Sulforaphane, a dietary component of broccoli/broccoli sprouts, inhibits breast cancer stem cells. Clinical Cancer Research, 16, 2580-2590.

Liu, D., Hunt, M., & Tsai, I. J. (2018). Inferring synteny between genome assemblies: A systematic evaluation. BMC Bioinformatics, 19, 1-13.

Liu, S., Liu, Y., Yang, X., Tong, C., Edwards, D., Parkin, I. A., & Paterson, A. H. (2014). The *Brassica oleracea* genome reveals the asymmetrical evolution of polyploid genomes. Nature Communications, 5, 3930.

Liu, Y., Jiang, H., Chen, W., Qian, Y., Ma, Q., Cheng, B., & Zhu, S. (2011). Genome-wide analysis of the auxin response factor (ARF) gene family in maize (*Zea mays*). Plant Growth Regulation, 63, 225-234.

Ludwig-Müller, J., Prinsen, E., Rolfe, S. A., & Scholes, J. D. (2009). Metabolism and plant hormone action during clubroot disease. Journal of Plant Growth Regulation, 28, 229-244.

Luo, D., Zhou, Q., Wu, Y., Chai, X., Liu, W., Wang, Y., & Liu, Z. (2019). Full-length transcript sequencing and comparative transcriptomic analysis to evaluate the contribution of osmotic and ionic stress components towards salinity tolerance in the roots of cultivated alfalfa (*Medicago sativa L.*). BMC Plant Biology, 19, 1-20.

Luo, J., Zhou, J. J., & Zhang, J. Z. (2018). Aux/IAA gene family in plants: Molecular structure, regulation, and function. International Journal of Molecular Sciences, 19, 259.

Lv, H., Fang, Z., Yang, L., Zhang, Y., & Wang, Y. (2020). An update on the arsenal: Mining resistance genes for disease management of Brassica crops in the genomic era. Horticulture Research, 7.

Lv, H., Wang, Y., Han, F., Ji, J., Fang, Z., Zhuang, M., & Yang, L. (2020a). A high-quality reference genome for cabbage obtained with SMRT reveals novel genomic features and evolutionary characteristics. Scientific Reports, 10, 12394.

Manoharan, R. K., Shanmugam, A., Hwang, I., Park, J. I., & Nou, I. S. (2016). Expression of salicylic acid-related genes in *Brassica oleracea* var. *capitata* during *Plasmodiophora brassicae* infection. Genome, 59, 379-391.

Marchler-Bauer, A., Derbyshire, M. K., Gonzales, N. R., Lu, S., Chitsaz, F., Geer, L. Y., & Bryant, S. H. (2015). CDD: N'BI's conserved domain database. Nucleic Acids Research, 43, 222-226.

Mourou, M., Raimondo, M. L., Lops, F., & Carlucci, A. (2023). Brassicaceae Fungi and Chromista Diseases: Molecular Detection and Host–Plant Interaction. *Plants*, *12*(5), 1033.

Mun, J. H., Yu, H. J., Shin, J. Y., Oh, M., Hwang, H. J., & Chung, H. (2012). Auxin response factor gene family in *Brassica rapa*: Genomic organization, divergence, expression, and evolution. Molecular Genetics and Genomics, 287, 765-784.

Ning, Y., Wang, Y., Fang, Z., Zhuang, M., Zhang, Y., Lv, H., & Yang, L. (2019). Comparative transcriptome analysis of cabbage (*Brassica oleracea* var. *capitata*) infected by *Plasmodiophora brassicae* reveals drastic defense response at secondary infection stage. Plant and Soil, 443, 167-183.

O'Leary, N. A., Wright, M. W., Brister, J. R., Ciufo, S., Haddad, D., McVeigh, R., & Pruitt, K. D. (2016). Reference sequence (RefSeq) database at NCBI: Current status, taxonomic expansion, and functional annotation. Nucleic Acids Research, 44, 733-745.

Oosterbeek, M., Lozano-Torres, J. L., Bakker, J., & Goverse, A. (2021). Sedentary plant-parasitic nematodes alter auxin homeostasis via multiple strategies. Frontiers in Plant Science, 12, 668548.

Paesold, S., Siegel, I. N. A., Seidel, C., & Ludwig-Müller, J. (2010). Flavonoid accumulation in *Arabidopsis thaliana* root galls caused by the obligate biotrophic pathogen *Plasmodiophora brassicae*. Molecular Plant Pathology, 11, 545-562.

Parkin, I. A., Koh, C., Tang, H., Robinson, S. J., Kagale, S., Clarke, W. E., & Sharpe, A. G. (2014). Transcriptome and methylome profiling reveals relics of genome dominance in the mesopolyploid *Brassica oleracea*. *Genome biology*, *15*, 1-18.

Rahman, S., Rehman, A., Waqas, M., Mubarik, M. S., Alwutayd, K., AbdElgawad, H., & Rizwan, M. (2023). Genome-wide exploration of bZIP transcription factors and their contribution to alkali stress response in *Helianthus annuus*. Plant Stress, 10, 100204.

Restrepo-Montoya, D., McClean, P. E., & Osorno, J. M. (2021). Orthology and synteny analysis of receptor-like kinases "RLK" and receptor-like proteins "RLP" in legumes. BMC Genomics, 22, 1-17.

Rienstra, J., Hernández-García, J., & Weijers, D. (2023). To bind or not to bind: How Auxin Response Factors select their target genes. Journal of Experimental Botany, 74, 6922-6932.

Shafique, I., Shahbaz, U., ul Haq, E., Abbas, A., Wahid, M. A., Anas, M., & Nasir, S. (2023). In silico genome-wide comparative analysis of Auxin Response Factors (ARFs) from *Brassica oleracea* and *Arabidopsis thaliana*. In International Congresses of Turkish Science and Technology (pp. 377-377).

Sinha, A. U., & Meller, J. (2007). Cinteny: Flexible analysis and visualization of synteny and genome rearrangements in multiple organisms. BMC Bioinformatics, 8, 1-9.

Song, X., Liu, G., Duan, W., Liu, T., Huang, Z., Ren, J., & Hou, X. (2014). Genome-wide identification, classification and expression analysis of the heat shock transcription factor family in Chinese cabbage. Molecular Genetics and Genomics, 289, 541-551.

Su, Z., Wang, L., Li, W., Zhao, L., Huang, X., Azam, S. M., & Qin, Y. (2017). Genome-wide identification of auxin response factor (ARF) genes family and its tissue-specific prominent expression in pineapple (*Ananas comosus*). Tropical Plant Biology, 10, 86-96.

Sun, D., Wang, C., Zhang, X., Zhang, W., Jiang, H., Yao, X., & Shan, X. (2019). Draft genome sequence of cauliflower (*Brassica oleracea* L. var. *botrytis*) provides new insights into the C genome in Brassica species. Horticulture Research, 6.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., & Kumar, S. (2011). MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular biology and evolution*, *28*(10), 2731-2739.

Tamura, K., Stecher, G., & Kumar, S. (2021). MEGA11: Molecular evolutionary genetics analysis version 11. Molecular Biology and Evolution, 38, 3022-3027.

Tang, Y., Bao, X., Liu, K., Wang, J., Zhang, J., Feng, Y., & Li, C. (2018). Genome-wide identification and expression profiling of the auxin response factor (ARF) gene family in physic nut. *PLoS One*, *13*(8), e0201024.

Thompson, J. D., Gibson, T. J., & Higgins, D. G. (2003). Multiple sequence alignment using ClustalW and ClustalX. *Current protocols in bioinformatics*, (1), 2-3.

Tombuloglu, H. (2019). Genome-wide analysis of the auxin response factors (ARF) gene family in barley (*Hordeum vulgare* L.). Journal of Plant Biochemistry and Biotechnology, 28, 14-24.

Ulmasov, T., Hagen, G., & Guilfoyle, T. J. (1997). ARF1, a transcription factor that binds to auxin response elements. Science, 276, 1865-1868.

Voorrips, R. (2002). MapChart: software for the graphical presentation of linkage maps and QTLs. Journal of heredity, 93(1), 77-78.

Wan, S., Li, W., Zhu, Y., Liu, Z., Huang, W., & Zhan, J. (2014). Genome-wide identification, characterization and expression analysis of the auxin response factor gene family in *Vitis vinifera*. Plant Cell Reports, 33, 1365-1375.

Wang, D., Pei, K., Fu, Y., Sun, Z., Li, S., Liu, H., & Tao, Y. (2007). Genome-wide analysis of the auxin response factors (ARF) gene family in rice (*Oryza sativa*). Gene, 394, 13-24.

Wang, X., Torres, M. J., Pierce, G., Lemke, C., Nelson, L. K., Yuksel, B., & Paterson, A. H. (2011). A physical map of *Brassica oleracea* shows complexity of chromosomal changes following recursive paleopolyploidizations. BMC Genomics, 12, 1-15.

Wang, Y., Ji, J., Fang, Z., Yang, L., Zhuang, M., Zhang, Y., & Lv, H. (2022). BoGDB: An integrative genomic database for *Brassica oleracea* L. Frontiers in Plant Science, 13, 852291.

Waqas, M., Azhar, M. T., Rana, I. A., Azeem, F., Ali, M. A., Nawaz, M. A., & Atif, R. M. (2019). Genome-wide identification and expression analyses of WRKY transcription factor family members from chickpea (*Cicer arietinum* L.) reveal their role in abiotic stress-responses. Genes & Genomics, 41, 467-481.

Waqas, M., Shahid, L., Shoukat, K., Aslam, U., Azeem, F., & Atif, R. M. (2020). Role of DNA-binding with one finger (Dof) transcription factors for abiotic stress tolerance in plants. In Transcription factors for abiotic stress tolerance in plants (pp. 1-14). Academic Press.

Warwick, S. I., Francis, A., & Al-Shehbaz, I. A. (2006). Brassicaceae: Species checklist and database on CD-Rom. Plant Systematics and Evolution, 259, 249-258.

Weijers, D., & Wagner, D. (2016). Transcriptional responses to the auxin hormone. Annual Review of Plant Biology, 67, 539-574.

Wittkopp, P. J., & Kalay, G. (2012). Cis-regulatory elements: Molecular mechanisms and evolutionary processes underlying divergence. Nature Reviews Genetics, 13, 59-69.

Xing, H., Pudake, R. N., Guo, G., Xing, G., Hu, Z., Zhang, Y., & Ni, Z. (2011). Genome-wide identification and expression profiling of auxin response factor (ARF) gene family in maize. BMC Genomics, 12, 1-13.

Yang, H., Fang, R., Luo, L., Yang, C., & Hui, W. (2023). Uncovering the mechanisms of salicylic acid-mediated abiotic stress tolerance in horticultural crops. Frontiers in Plant Science, 14, 1226041.

Yang, K., Nath, U. K., Biswas, M. K., Kayum, M. A., Yi, G. E., Lee, J., & Nou, I. S. (2018). Whole-genome sequencing of *Brassica oleracea* var. *capitata* reveals new diversity of the mitogenome. PloS One, 13, e0194356.

Yu, J., Zhao, M., Wang, X., Tong, C., Huang, S., Tehrim, S., & Liu, S. (2013). Bolbase: A comprehensive genomics database for *Brassica oleracea*. BMC Genomics, 14, 1-7.

Yuan, Y., Qin, L., Su, H., Yang, S., Wei, X., Wang, Z., & Zhang, X. (2021). Transcriptome and coexpression network analyses reveal hub genes in Chinese cabbage (*Brassica rapa* L. ssp. *pekinensis*) during different stages of *Plasmodiophora brassicae* infection. Frontiers in Plant Science, 12, 650252.

Zahoor, A., Niaz, Z., Seo, H., Jabran, M., Azeem, F., Joyia, F. A., & Ali, M. A. (2023). Computational insights into evolutionary and functional biology of plant cell wall degrading pectate lyases from plant parasitic nematodes of the order Tylenchida. Pakistan Journal of Agricultural Sciences, 60.

Zhang, H., Ning, C. A. O., Chunjuan, D. O. N. G., & Shang, Q. (2017). Genome-wide identification and expression of ARF gene family during adventitious root development in hot pepper (*Capsicum annuum*). Horticultural Plant Journal, *3*(4), 151-164.

Zhang, X., Liu, Y., Fang, Z., Li, Z., Yang, L., Zhuang, M., & Lv, H. (2016). Comparative transcriptome analysis between broccoli (*Brassica oleracea* var. *italica*) and wild cabbage (*Brassica macrocarpa* Guss.) in response to *Plasmodiophora brassicae* during different infection stages. Frontiers in Plant Science, 7, 1929.