Trends in Horticulture 2024, 7(2), 5483.
https://doi.org/10.24294/th.v7i2.5483

Article

Comparative analysis of the WRKY gene family reveals the gene family expansion and evolution in diverse plant species

Akshay Singh[1,2,*], Ajay Kumar Sharma[3]

[1] Division of Genomic Resources, ICAR-National Bureau of Plant Genetic Resources, Pusa Campus, New Delhi 110012, India
[2] Dr. A. P. J. Abdul Kalam Technical University, Lucknow, Uttar Pradesh 226031, India
[3] Institute of Management Studies Engineering College, Ghaziabad, Uttar Pradesh 201009, India
*** Corresponding author:** Akshay Singh, akshaybioinfo@gmail.com

CITATION

Singh A, Sharma AK. Comparative analysis of the WRKY gene family reveals the gene family expansion and evolution in diverse plant species. Trends in Horticulture. 2024; 7(2): 5483. https://doi.org/10.24294/th.v7i2.5483

ARTICLE INFO

Received: 28 March 2024
Accepted: 15 April 2024
Available online: 22 May 2024

COPYRIGHT

Copyright © 2024 by author(s). *Trends in Horticulture* is published by EnPress Publisher, LLC. This work is licensed under the Creative Commons Attribution (CC BY) license. https://creativecommons.org/licenses/by/4.0/

Abstract: The WRKY gene family plays a very diverse role in plant growth and development. These genes contained an evolutionarily conserved WRKY DNA binding domain, which shows functional diversity and extensive expansion of the gene family. In this study, we conducted a genome-wide comparative analysis to investigate the evolutionary aspects of the WRKY gene family across various plant species and revealed significant expansion and diversification ranging from aquatic green algae to terrestrial plants. Phylogeny reconstruction of WRKY genes was performed using the Maximum Likelihood (ML) method; the genes were grouped into seven different clades and further classified into algae, bryophytes, pteridophytes, dicotyledons, and monocotyledons subgroups. Furthermore, duplication analysis showed that the increase in the number of WRKY genes in higher plant species was primarily due to tandem and segmental duplication under purifying selection. In addition, the selection pressures of different subfamilies of the WRKY gene were investigated using different strategies (classical and Bayesian maximum likelihood methods (Data monkey/PAML)). The average dN/dS for each group are less than one, indicating purifying selection. Our comparative genomic analysis provides the basis for future functional analysis, understanding the role of gene duplication in gene family expansion, and selection pressure analysis.

Keywords: WRKY gene family; comparative analysis; transcription factors; gene duplication

1. Introduction

Plants need to be involved in a variety of complex mechanisms to respond to various factors and stresses in the external environment and maintain regular growth and expansion. Plants cannot move to protect themselves from adverse environmental conditions and can adapt to biotic or abiotic stress responses. Thus, plants include multiple regulatory mechanisms that allow them to amplify regulatory signals and responses to stress at the cellular, molecular, and physiological levels [1]. WRKY transcription factors (TFs) can regulate various stress responses through a composite network of genes. At the molecular level, the correlation of WRKY genes in plants could provide the most expected outcome of synchronized responses. Activation or inhibition of WRKY TFs by binding with W-box or W-box-like sequences is regulated at the transcriptional and translational levels. Due to the strong regulation associated with explicit recognition of WRKY binding to the promoter sequence, they are promising candidates for crop improvement [2].

TF is a specialized class of peptides that participates in gene regulation exclusively in plants [3]. They activate or suppress the expression of several target genes by binding to specific DNA segments in the promoter region. WRKY is one of the most important families of transcription factors for regulatory genes, primarily recognized in plants [4]. WRKY TF, an important class of the stress-responsive TF family, is actively involved in the regulation of plant growth and development, as well as in the biotic and abiotic stress responses [5]. One of the most common characteristics of all WRKYs is the presence of a highly conserved WRKY domain sequence approximately 60 amino acids in length at the N-terminus and a zinc-finger structure at the C-terminus [6]. Based on the type of WRKY domain and the pattern of zinc-finger motifs, WRKY proteins are mainly classified into groups (I, II, III) [7]. WRKY transcription factors are generally considered to be plant-specific and have been studied in several plant species such as rice (*Oryza sativa*), *Arabidopsis thaliana*, cucumber (*Cucumus sativus*), grapes (*Vitis vinifera*), poplar (*Populus trichocarpa*), and pigeon pea (*Cajanus cajan*), respectively. The members of group I contain two WRKY domains, while members of groups II and III contain only a single WRKY domain. Group II is further subdivided into subgroups (IIa-IIe) according to the presence of conserved short motifs [7]. The WRKY members of groups I and II are composed of C2H2-type zinc-finger-like sequences, while group III WRKY members include C2HC-type zinc-finger-like sequences [8].

The first WRKY *SPF1* was cloned and characterized from sweet potatoes (*Ipomoea batatas*) in early 1994 [4]. Since then, cloning and characterization of several WRKYs have been carried out in several plant species including rice [9], wheat [10], soybeans [11], tomatoes [12], and the biofuel plant *Jatropha curcas* [13]. Certain WRKY TFs participate in various developmental and physiological processes, such as seed germination [14], plant growth [15], secondary metabolism [16], seed and trichome development [17,18], panicle development [19], leaf senescence [20], floral bud differentiation [21], and hormones, such as jasmonic acid (JA) as well as salicylic acid (SA) mediated responses [22]. Several studies may suggest that WRKY genes are also involved in the abscisic acid (ABA)-mediated drought responses [23]. The WRKY TFs are also involved in a range of abiotic stress responses like heat [24], cold [25], drought and salinity [26,27], and biotic stresses including bacteria [28], fungi [29], nematodes [30], viral pathogens [31], and aphid resistance [32].

The WRKY TF gene family has been studied in many plant species for several years, but little is known about the phenomenon of expansion and evolution of the WRKY gene family in higher plants. Therefore, considering the importance of the WRKY TF gene family in plant defense mechanisms and growth, and developmental processes, an attempt was made to understand the evolution, regulation, and distribution of the WRKY gene family and its further exploration in higher plants. The main purpose of the current analysis is to gain more insight into the genomic distribution, organization, and evolution of the WRKY gene family across different crop lines. In this study, we conducted a genome-wide comparative study of the WRKY-TFs gene family in 40 plant species belonging to diverse groups such as green algae, bryophytes, pteridophytes, monocotyledons, and dicotyledons. Our analysis provides useful information about the WRKY gene family, which supports

potential functional and ecological studies of this essential gene family in higher plant species.

2. Materials and methods

2.1. Sequence retrieval

The genomic and proteomic data of *Volvox carteri*, *Ostreococcus lucimarinus*, *Chlamydomonas reinhardtii*, *Sphagnum fallax*, *Physcomitrella patens*, *Selaginella moellendorffii*, *Amborella trichopoda*, *Setaria italica*, *Sorghum bicolor*, *Zea mays*, *Oryza sativa*, *Brachypodium* distachyon, *Solanum lycopersicum*, *Solanum tuberosum*, *Phaseolus vulgaris*, *Medicago truncatula*, *Glycine max*, *Arabidopsis thaliana*, *Brassica rapa*, *Brassica oleracea*, *Linum usitatissimum*, *Manihot esculenta*, *Populus trichocarpa*, *Fragaria vesca*, *Malus domestica*, *Prunus persica*, *Theobroma cacao*, and *Gossypium raimondii* were downloaded from the Phytozome database (https://phytozome.jgi.doe.gov/pz/portal.html). *Triticum aestivum* and *Hordeum vulgare* genome and proteome data were downloaded from Ensemble Plants (https://plants.ensembl.org/). The genome sequence of *Capsicum annuum* was downloaded from the Pepper Genome Database (http://pgd.pepper.snu.ac.kr/). The *Cucumis sativus* genome was downloaded from the Cucumber Genome Database (http://cucumber.genomics.org.cn/). The *Citrullus lanatus* genome from the Cucurbit Genomics Database (http://cucurbitgenomics.org/). The *Solanum melongena* genome and proteome were downloaded from the Eggplant Genome Database (http://eggplant.kazusa.or.jp/). *Raphanus stivus* genomic data was downloaded from the Radish Genome Database (http://radish-genome.org/). *Cajanus cajan* and *Cicer arietinum* genomes and gene information were downloaded from the Legume Database (https://legumeinfo.org/). *Citrus clementina* and *Citrus sinensis* genomic data were downloaded from the Citrus Genome database (https://www.citrusgenomedb.org/), and *Gossypium hirsutum* genome was downloaded from PlantGDB (http://www.plantgdb.org/GhGDB/). The Hidden Markov Model (HMM) profile of WRKY (PF03106) is taken from the Pfam database (https://pfam.xfam.org/) and used to scan the protein dataset to identify WRKY protein candidates from each plant species using the 'hmmsearch' module of the HMMER program [33]. The predicted WRKY candidates were re-confirmed by using online search tools PfamScan (https://www.ebi.ac.uk/Tools/pfa/pfamscan/) and CDD Search (https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi).

2.2. Multiple sequence alignment and phylogenetic analysis

Multiple sequence alignments of all conserved WRKY protein sequences from 14 plant species representing major plant groups, such as monocot, dicot, and lower plant species were performed using ClustalX version 2.0 [34] with default parameters. The alignment files generated from the ClustalX were converted into the MEGA format after that MEGA6 software [35] was used to build the phylogenetic tree. The phylogenetic trees were generated by using the neighbor-joining (NJ) method under the Jones-Thornton-Taylor (JTT) method based on the following parameters, including 1000 bootstrap replicates, amino acid type of substitution,

Poisson model, 95% of site coverage, and partial deletion of gap/missing data. Based on the multiple sequence alignment, phylogenetic analyses as well as previously reported classification of *AtWRKY* and *OsWRKY* genes, WRKY genes were assigned to different groups and subgroups.

2.3. Gene duplication and evolutionary analysis

Gene duplication analysis was done using the MCScanX software package [36]. To perform the duplication analysis, homologs/paralogs among all predicted WRKY protein sequences were found by deploying all vs all BLASTP programs with the parameters including V = 10, B = 100, filter = seg, *e*-value $<1 \times 10^{-10}$ and -m 8 (for tabular format). All vs. all BLASTP outputs were integrated into the MCScanX program, along with the predicted chromosomal coordinates of all protein-coding genes. MCScanX can classify the origin of duplicate genes within a gene family into various categories, including proximal, dispersed, segmental/WGD, and tandem duplicates based on the copy number and genomic distribution. The duplication events were predicted to compare the contribution of duplication events toward the expansion of the WRKY gene family across selected plant species. Orthologous gene locations and segmental duplication were visualized by using Circos v0.68 (http://circos.ca/).

Synonymous (Ks: which do not alter amino acids) and non-synonymous (Ka: alters the amino acids) substitutions of each orthologous gene pair were used to calculate the selection pressure for all orthologous gene pairs of selected species WRKY family [37]. As a result, the PAL2NAL program was used to convert protein alignments and their corresponding DNA (or mRNA) sequences into codon alignments [38]. The PAL2NAL program automatically determines the Ka and Ks values by CODEML script in the PAML program. The synonymous (*K*s) values above 5.0 were excluded from the analysis due to saturated substitutions at synonymous sites [39].

3. Results

3.1. Identification of WRKY genes across 40 diverse plant species

In our study, we thoroughly analyzed the WRKY gene family in 40 plant species, starting from marine algae to flowering plants (**Figure 1**). A total of 3234 WRKY homologs were identified from the selected plant genomes (**Figure 2**). The presence of the WRKY gene in unicellular aquatic algae indicates their ancient origin and their functional conservation. In addition, the presence of greater numbers of WRKY genes in higher plants indicates a broader degree of expansion as compared to marine algae. Initially, only two homologs were identified in all three marine algae *V. carteri*, *O. lucimarinus*, and *C. reinhardtii* respectively, which was extremely lower than those present in land plants. Besides, in *P. patens* a most primitive land plant which is diverged from marine plants [40], and *Amborella trichopoda*, a single living species of the most basal lineage of the clade angiosperms [41], thirty-two and twenty-nine WRKYgenes were recognized, respectively. While, in *S. moellendorffii* a non-seeded vascular plant [42], thirty-five WRKYs were

present. In seven monocotyledonous plant species, 45 (in *H. vulgare*) to 172 (*T. aestivum*) WRKY genes were identified, and in the other twenty-six eudicots, 44 (*S. melongena*) to 188 (in *G. max*) WRKY genes have been observed, which indicating the widespread gene expansion and duplication events.

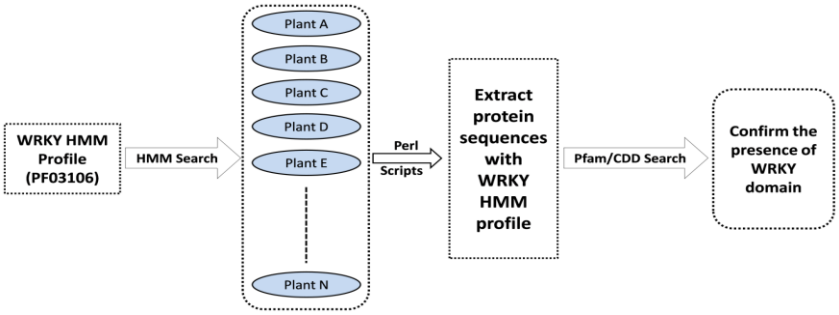

Figure 1. Schematic representation of the methodology used for the identification of the WRKY gene family across selected plant genomes.

		Species	Group							Total
			I	IIa	IIb	IIc	IId	IIe	III	
Monocotyledons	Algae	*Ostreococcus lucimarinus*	0	0	0	0	0	0	2	2
		Volvox carteri	1	0	0	0	0	1	0	2
		Chlamydomonas reinhardtii	2	0	0	0	0	0	0	2
	Bryophyte	*Physcomitrella patens*	3	0	4	16	5	0	4	32
		Sphagnum fallax	6	0	2	8	6	7	10	39
	Pteridophyte	*Selaginella moellendorffii*	8	0	6	7	8	2	4	35
	Amborellales	*Amborella trichopoda*	6	2	4	4	3	4	6	29
	Poaceae	*Setaria italica*	10	5	8	20	9	12	41	105
		Zea mays	27	7	11	29	14	16	31	135
		Sorghum bicolor	20	6	11	26	10	19	41	133
		Oryza sativa	19	5	6	18	8	11	40	107
		Brachypodium distachyon	17	3	6	22	6	10	24	88
		Triticum aestivum	30	11	8	50	17	10	46	172
		Hordeum vulgare	18	7	9	23	5	6	18	86
Dicotyledons	Solanaceae	*Capsicum annuum*	12	4	6	14	5	7	11	59
		Solanum lycopersicum	13	5	6	18	7	16	14	79
		Solanum tuberosum	16	5	8	16	6	17	14	82
		Solanum melongena	12	6	4	14	8	9	7	60
	Cucurbitaceae	*Cucumis sativus*	12	5	4	15	8	7	11	62
		Citrullus lanatus	11	2	5	16	7	6	10	57
	Fabaceae	*Cicer arietinum*	12	4	11	15	6	10	7	65
		Medicago truncatula	16	5	11	18	7	16	25	98
		Phaseolus vulgaris	16	5	14	12	7	11	15	90
		Glycine max	32	14	33	42	21	20	26	188
		Cajanus cajan	16	5	17	24	8	10	14	94
	Brassicaceae	*Arabidopsis thaliana*	13	4	8	18	7	9	14	73
		Raphanus sativus	31	6	15	27	14	12	21	126
		Brassica rapa	31	7	17	38	14	13	25	145
		Brassica oleracea	36	6	17	36	14	15	24	148
	Malpighiales	*Linum usitatissimum*	24	4	13	16	11	12	17	97
		Manihot esculenta	17	5	14	20	8	9	12	85
		Populus trichocarpa	48	5	9	13	13	4	10	102
	Rosaceae	*Fragaria vesca*	10	3	8	13	6	7	15	62
		Prunus persica	10	3	8	15	7	7	8	58
		Malus domestica	24	8	12	33	15	13	23	127
	Rutaceae	*Citrus clementina*	9	4	8	12	5	5	8	51
		Citrus sinensis	9	2	11	12	5	6	7	52
	Malvales	*Theobroma cacao*	10	3	8	17	6	6	11	61
		Gossypium raimondii	20	7	16	35	15	13	12	120
		Gossypium hirsutum	34	23	28	47	19	17	29	197

Figure 2. NCBI taxonomy trees of 40 plant species showing the no. of predicted WRKY genes in each crop. The species from different taxonomic groups were marked with a specific color.

The maximum number of the WRKY genes was identified in the leguminous plant, *G.max* (188) and in *T. aestivum* (172) respectively. The expansion of the WRKYgene family in wheat is primarily due to gene duplication events, and compared with tandem duplication, segmental duplication might play a more critical role [43]. Although differences in copy number among the plant species appear very complex, our analysis showed that the number of WRKY genes in each plant species is positively correlated with the number of genes present in that species. In addition, the WRKY gene was unevenly distributed within the species or among the plant species. For example, no WRKY genes were present in 10 out of the 27 *P. patens* chromosomes, while about 51% of the total *Oryza nivara* WRKY genes were present only on chromosomes 1, 5, and 11 out of the total 12 chromosomes. Similarly, in *G. max* and *G. raimondii*, both the gene copy numbers and distribution of WRKY genes were different.

3.2. Comparative phylogenetic analysis

Comparative analysis of the WRKY genes was performed in 40 plant genomes consisting of monocotyledonous, dicotyledonous, and lower eukaryotes but the main focus was on the pigeonpea WRKY gene family. The WRKY gene plays an important role in the regulation of gene networks which are associated with several important developmental processes and defense responses in plants. Genome-wide analysis of the WRKY gene family showed that the unicellular green algae have lesser no. of WRKY genes (1–4), followed by *P. patens* (non-vascular plant) and *S. moellendorffii* (vascular plant) of the lower plant group having 32 and 35 WRKY genes, respectively.

To understand the evolutionary relationships between WRKY homologs of different plant species, we constructed a phylogenetic tree using 1020 WRKY genes from 14 different plant species representing a large group of plants. These 14 plant species represent; Algae (*O. lucimarinus*, *V. carteri*, *C. reinhardtii*), Bryophytes (*P. patens*, S. *moellendorf*), Monocots (*Z. mays, O. sativa, B. distachyon*), and Dicots (*C. cajan*, *G. max*, *M. truncatula*, *P. tricocarpa*, *A. thaliana*, *F. vesca*). The phylogenetic clustering between diverse plant lineages reveals the evolutionary relationship among WRKY proteins.

According to phylogenetic analysis plant WRKYs can be divided into seven clades and named I, II, III, IV, V, VI, and VII, respectively (**Figure 3**), as reported earlier in Arabidopsis, rice, and other plant species. Monocots and dicots were distributed between all clades (shown by light green and light blue colors). The algal species (indicated by orange color) are observed in clades I and VI, and bryophyte (*P. patens*) are predominantly observed in clades number VII, VI, and V (indicated by dark blue color), while pteridophyte (*S. moellendorffii*) was observed in all clades except clade IV (represented by red color). Clade IV contained only monocotyledonous and dicotyledonous species.

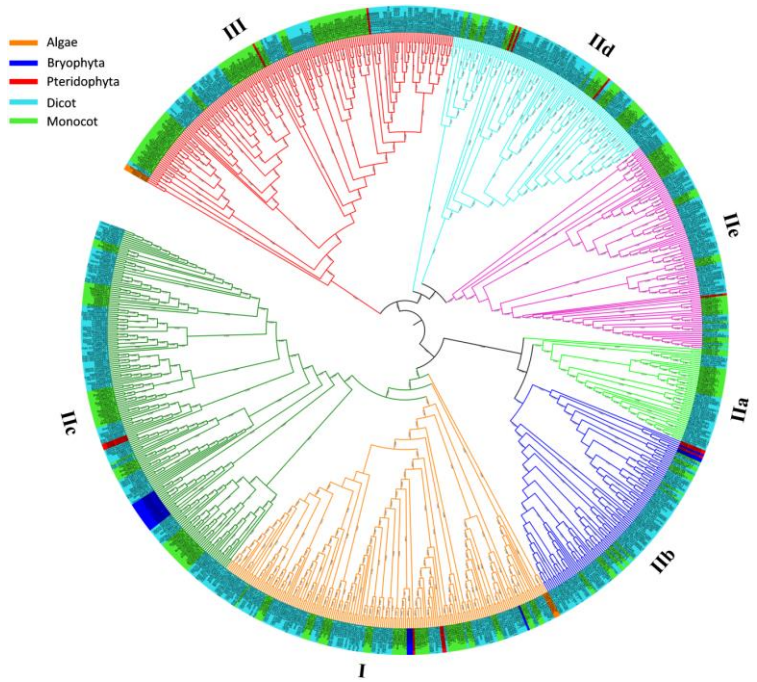

Figure 3. Phylogenetic relationships of WRKY proteins from 14 different plant species. The tree was built using the neighbor-joining method by MEGA6.0. The different colors of the inner circle represent WRKY groups, and the colors of the outer circle represent the crop category; (orange-algae, blue-bryophytes, red-pteridophytes, light blue-dicots, and green color represents monocts).

3.3. Expansion of WRKY gene family among plant genomes

Gene expansion or duplication, arising from polyploidy or through tandem and segmental duplication associated with replication, is a major factor of gene family expansion. Various types of gene duplications such as segmental duplications (SD) or whole-genome duplication (WGD), and single gene duplication (including proximal, tandem duplications as well as dispersed duplication), have been identified. Out of these gene duplication patterns, tandem and segmental type of duplication are the two major causes of gene family expansion in plants. Therefore, assessment of the duplication type in WRKY genes family for the selected plant genomes was performed using MCscanX program, including algae *O. lucimarinus*, *V. carteri*, *C. reinhardtii,* the basal land plant species *S. moellendorffii, P. patens* as well as angiosperms (monocotyledonous and dicotyledonous). The results suggested that the number of WRKY genes maintained by different type of gene duplication events. The dispersed type of duplication mode preferentially detected in all selected plant species. While, WGD or segmental duplication pattern of WRKYgenes were detected in most of the higher plant species, excluding algae and mosses, which might be correlated to the fact that all flowering plants go through one or more whole-genome duplication events. For example, there are 16 *GmWRKY* gene pairs (Chr03&Chr19: Glyma03G176600-Glyma19G177400, Glyma03G220100-Glyma19G217000, Glyma03G220800-Glyma19G217800, Glyma03G224700-Glyma19G221700, Glyma03G256700-Glyma19G254800), (Chr09&Chr15: Glyma09G005700-Glyma15G110300, Glyma09G029800-Glyma15G135600, Glyma09G034300-Glyma15G139000, Glyma09G061900-Glyma15G168200,

Glyma09G080000-Glyma15G186300), and (Chr09&Chr18: Glyma09G240000-Glyma18G256500, Glyma09G250500-Glyma18G242000, Glyma09G254400-Glyma18G238600, Glyma09G254800-Glyma18G238200, Glyma09G274000-Glyma18G213200, Glyma09G280200-Glyma18G208800) were identified as segmental duplicated genes (**Figure 4**), suggesting that the expansion of *GmWRKY* genes were possibly occurs due to gene segmental duplication.

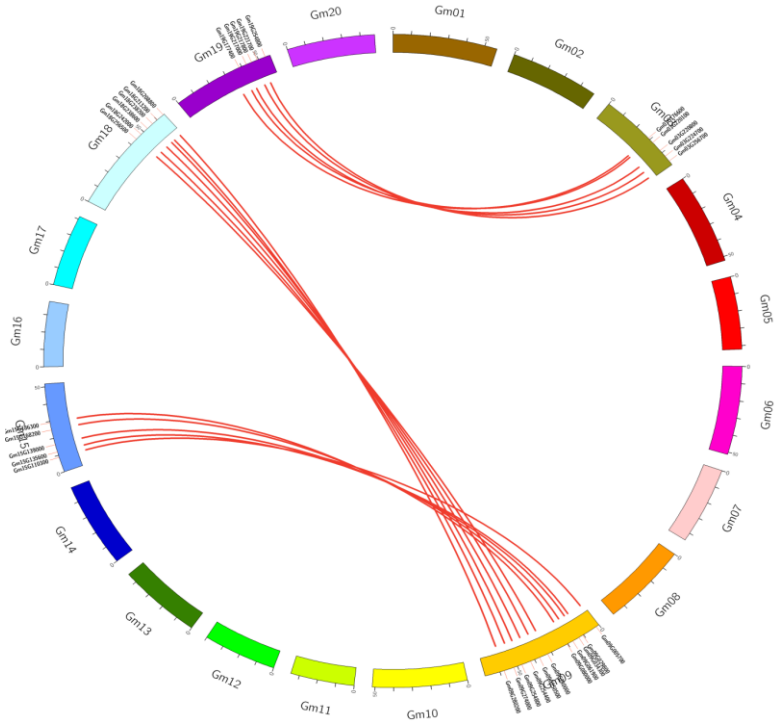

Figure 4. Segmental duplications of *GmWRKY* genes in *Glycine max* genome.
Note: Red color lines denote segmental duplicated *GmWRKY* gene pairs. Each chromosome in the circle is represented with a different color.

Specifically, the WRKY gene accounted for 37.1%, 32.4%, 17.4%, and 4.5%, of the duplication, retained from WGD/segmental duplication, dispersed, tandem, and proximal duplication respectively. Due to the segmental and tandem duplication events, a significant increase in the number of WRKY genes was observed in higher plants compared to the basal land plants. A species-specific duplication model was detected and the proportion of the segmental and tandem duplication in all plant species was not equal. For example, in the case of monocots, similar type of trends of tandem, as well as WGD/segmental duplication was observed in all selected monocotyledons except *T. aestivum* and *H. vulgare* in which WGD/segmental type of duplication is absent. Conversely, in dicotyledonous plants, WGD/segmental duplication mainly enriched the expansion of *G. hirsutum*, *B. oleracea*, *G. max* and *P. tricocarpa* WRKY genes and the tandem duplication event confers to the expansion of WRKY genes in *M. truncatula* and *S. tuberossum* plant species. In *G. raimondii*, more than 74% of WRKY genes, were derived through WGD/segmental duplication and 32.0% of genes are WGD/segmental duplicated in case of *B. oleracea* and *Z. mays*, while 33.0% of genes in *S. tuberossum*, 29% in *M. truncatula* and24 % in *Z. mays* were derived through tandem duplication event, which were much greater than

other plant species (**Table 1**).

Table 1. Distribution of the type of duplication in all selected 40 plant species.

Crop species	Total No. of genes	Dispersed	Proximal	Tandem	WGD/ Segmental
Ostreococcus lucimarinus	2	0	0	0	0
Volvox carteri	2	0	0	0	0
Chlamydomonas reinhardtii	2	0	0	0	0
Physcomitrella patens	32	26	2	4	0
Sphagnum fallax	39	33	0	6	0
Selaginella moellendorffii	35	29	0	6	0
Amborella trichopoda	29	22	0	7	0
Setaria italica	105	29	34	26	16
Zea mays	135	36	22	33	44
Sorghum bicolor	133	38	31	28	36
Oryza sativa	107	47	22	24	14
Brachypodium distachyon	88	30	23	19	16
Triticum aestivum	172	168	0	4	0
Hordeum vulgare	86	39	29	18	0
Capsicum annuum	59	40	13	4	0
Solanum lycopersicum	79	45	16	18	0
Solanum tuberosum	82	42	13	27	0
Solanum melongena	60	48	0	12	0
Cucumis sativus	62	56	2	4	0
Citrullus lanatus	57	29	22	6	0
Cicer arietinum	65	39	20	6	0
Medicago truncatula	98	41	29	28	0
Phaseolus vulgaris	90	44	28	18	0
Glycine max	188	115	22	19	32
Cajanus cajan	94	74	3	17	0
Arabidopsis thaliana	73	27	31	15	0
Raphans sativus	126	102	9	15	0
Brassica rapa	145	128	4	13	0
Brassica oleracea	148	51	30	19	48
Linum usitatissimum	97	92	0	5	0
Manihot esculenta	85	55	10	4	16
Populus trichocarpa	102	45	18	9	30
Fragaria vesca	62	14	14	22	12
Prunus persica	58	20	31	7	0
Malus domestica	127	127	0	0	0
Citrus clementina	51	25	18	8	0
Citrus sinensis	52	43	4	5	0
Theobroma cacao	61	26	21	14	0
Gossypium raimondii	120	62	26	10	22
Gossypium hirsutum	197	37	10	4	146

3.4. Identification of orthologous gene pairs and evolutionary analysis

For the selection pressure analysis, we identified 42 orthologous gene pairs between *C.cajan* and *G.max*, 60 orthologous pairs between *G.max* and adzuki bean, and 36 orthologous gene pairs between *C.cajan* and adzuki bean respectively, based on the phylogenetic as well as sequence homology (**Table 2**). The higher and lower levels of protein sequence identity between *C.cajan* and *G.max* were observed in the pairs CcWRKY86-Glyma02G297400 (93.42%) and CcWRKY71-Glyma02G141000 (81.14%) with an average sequences identity of 85.24%. The higher and lower level of amino acid identity between the *G.max* and adzuki bean gene pairs were Glyma05G127600-Vang0333s00130 (96.63%) and Glyma09G254400-Vang04g03920 (81.14%) with an average of 85.40%. The higher and lower levels of protein sequence identity between *C.cajan* and adzuki bean were the gene pairs CcWRKY75-Vang0333s00130 (96.21%) and CcWRKY13-Vang01g02180 (81.14%) with an average sequences identity of 85.48%.

The chromosomal distribution and syntenic relationship between *C. cajan*-*G. max*, *G. max*-adzuki bean and *C. cajan*-adzuki bean orthologous gene pairs were shown in **Figure 5**. The physical mapping of WRKY genes revealed that (56%–64%) of WRKY genes in most of the leguminous species such as pigeonpea, adzuki bean, common bean and mung bean were not located in the corresponding chromosomes, suggesting the occurrence of substantial chromosomal rearrangement in the leguminous genomes. In addition, the level of variation between the synonymous substitution (*dS*) and non-synonymous substitution (*dN*) values was used to infer the direction and magnitude of natural selection acting on WRKY orthologous gene pairs in pigeonpea, soybean and adzuki bean. The *dN/dS* distribution showed that the WRKY orthologous gene pairs in leguminous species subject to stronger purifying selection pressure during evolution (**Table 2**).

Table 2. Identification of synonymous (*dS*) and non- synonymous (*dN*) substitution rates for WRKY orthologous gene pairs between pigeonpea-soybean, soybean-adzuki bean and pigeonpea-adzuki bean.

Orthologous gene pairs		**Protein identity (%)**	***S***	***N***	***dS***	***dN***	***dN/dS***
CcWRKY86	Glyma02G297400	93.42	386.2	1248.8	0.1994	0.0296	0.1487
CcWRKY54	Glyma18G081200	90.68	430.0	1298.0	0.1743	0.0471	0.2704
CcWRKY12	Glyma14G200200	92.14	452.9	1254.1	0.2111	0.0326	0.1544
CcWRKY61	Glyma18G208800	85.27	396.6	1205.4	0.1935	0.0699	0.3611
CcWRKY32	Glyma15G110300	88.24	440.1	1176.9	0.3081	0.0525	0.1705
CcWRKY21	Glyma19G177400	87.41	313	1085	0.3007	0.0586	0.1948
CcWRKY01	Glyma18G056600	81.17	380.8	1164.2	0.2535	0.0879	0.3466
CcWRKY90	Glyma14G006800	82.09	353.7	1137.3	0.2227	0.0692	0.3106
CcWRKY29	Glyma18G263400	85.32	347.7	957.3	0.2432	0.0344	0.1412
CcWRKY27	Glyma01G053800	81.72	355.5	1000.5	0.3914	0.0819	0.2092
CcWRKY55	Glyma18G081200	91.23	303.1	1049.9	0.3186	0.1115	0.3498
CcWRKY35	Glyma05G215900	88.43	292.4	790.6	0.2199	0.0536	0.2437
CcWRKY28	Glyma09G250500	83.81	344.8	1026.2	0.3451	0.0651	0.1885
CcWRKY33	Glyma04G238300	82.63	259.8	799.2	0.3409	0.0655	0.1921

Table 2. (*Continued*).

Orthologous gene pairs		Protein identity (%)	*S*	*N*	*dS*	*dN*	*dN/dS*
CcWRKY50	Glyma03G042700	83.29	252.3	836.7	0.2682	0.0714	0.2664
CcWRKY23	Glyma19G254800	82.38	255.2	797.8	0.2802	0.0748	0.2669
CcWRKY71	Glyma02G141000	81.14	210.7	707.3	0.241	0.0312	0.1295
CcWRKY77	Glyma01G189100	85.27	240.5	686.5	0.5871	0.0532	0.0907
CcWRKY42	Glyma02G010900	90.49	232.9	706.1	0.2237	0.0407	0.1818
CcWRKY34	Glyma05G211900	84.61	196.4	664.6	0.282	0.0592	0.2097
CcWRKY38	Glyma07G023300	85.23	204.7	626.3	0.2677	0.0228	0.0851
CcWRKY62	Glyma18G213200	83.51	226.5	661.5	0.2898	0.1124	0.3877
CcWRKY82	Glyma02G115200	84.14	220.6	625.4	0.255	0.0549	0.2155
CcWRKY83	Glyma05G123000	87.28	200.2	822.8	0.4936	0.1444	0.2925
CcWRKY87	Glyma08G118200	82.19	180.3	599.7	0.5144	0.082	0.1594
CcWRKY64	Glyma19G217000	85.24	199.4	544.6	0.2288	0.0511	0.2234
CcWRKY80	Glyma15G003300	81.25	218.9	693.1	0.2498	0.0621	0.2487
CcWRKY24	Glyma08G143400	90.73	188.4	510.6	0.1273	0.047	0.3692
CcWRKY79	Glyma12G212300	87.45	187.9	577.1	0.293	0.0613	0.2093
CcWRKY18	Glyma04G218400	84.38	166.1	523.9	0.2038	0.0576	0.2827
CcWRKY60	Glyma07G116300	89.72	155.8	465.2	0.2855	0.0591	0.207
CcWRKY81	Glyma04G061400	82.21	154.2	415.8	0.3265	0.0826	0.2529
CcWRKY04	Glyma16G031400	85.63	133.3	442.7	0.2945	0.0631	0.2142
CcWRKY67	Glyma09G254800	85.81	145.4	418.6	0.3433	0.0397	0.1155
CcWRKY70	Glyma17G224800	83.25	102.6	374.4	0.2429	0.0647	0.2665
CcWRKY25	Glyma20G030500	82.31	139.5	346.5	0.1868	0.0742	0.3971
CcWRKY76	Glyma16G054400	83.57	88.6	292.4	0.308	0.0361	0.1173
CcWRKY89	Glyma08G011300	81.62	76.4	328.6	0.3152	0.0723	0.2293
CcWRKY19	Glyma06G168400	84.51	103.0	296	0.157	0.1209	0.7698
CcWRKY75	Glyma01G056800	82.26	213.3	677.7	51.9354	0.7458	0.0144
CcWRKY88	Glyma04G218400	84.53	148.6	523.4	6.7035	0.5915	0.0882
CcWRKY07	Glyma06G242200	82.42	135.7	497.3	19.1289	0.9472	0.0495
Glyma01G053800	Vang0005s00450	84.65	261.3	710.7	0.4622	0.0865	0.1870
Glyma01G056800	Vang04g17060	82.81	166	569	12.3005	0.6527	0.0531
Glyma01G128100	Vang0103s00010	82.23	346.9	1126.1	0.301	0.0792	0.263
Glyma02G112100	Vang0005s00450	83.14	266.3	705.7	0.5046	0.0925	0.1833
Glyma02G115200	Vang0005s00190	81.22	236.8	627.2	0.2959	0.0672	0.227
Glyma02G141000	Vang06g17510	96.27	189.5	623.5	0.2819	0.0363	0.1286
Glyma02G232600	Vang01g03800	89.43	333.7	1103.3	0.3742	0.0637	0.1701
Glyma02G285900	Vang0333s00130	87.21	86.6	336.4	3.7293	0.2899	0.0777
Glyma02G297400	Vang01g01490	84.29	404.8	1290.2	0.2578	0.0687	0.2663
Glyma02G306300	Vang01g00540	82.51	338.4	1131.6	0.1911	0.0934	0.489
Glyma03G042700	Vang0103s00010	81.25	358.3	1114.7	0.3003	0.0812	0.2705
Glyma03G109100	Vang08g06450	83.26	163.2	523.8	0.5319	0.0625	0.1175
Glyma03G159700	Vang01g12760	81.73	169.5	538.5	51.5733	0.854	0.0166
Glyma03G176600	Vang04g14650	82.11	289	1052	0.324	0.0649	0.2004

Table 2. (*Continued*).

Orthologous gene pairs		Protein identity (%)	*S*	*N*	*dS*	*dN*	*dN/dS*
Glyma04G076200	Vang01g12760	89.41	144.3	503.7	1.3079	0.2744	0.2098
Glyma04G115500	Vang0039ss00390	86.28	541.6	1684.4	0.3717	0.0606	0.1631
Glyma05G127600	Vang0333s00130	96.63	78.9	344.1	0.6192	0.1705	0.2754
Glyma05G211900	Vang10g06430	82.61	212.5	645.5	0.3774	0.06	0.1591
Glyma05G215900	Vang10g06010	87.62	280.5	802.5	0.2095	0.0604	0.2883
Glyma06G077400	Vang01g12760	87.59	157.1	562.9	1.3158	0.2676	0.2034
Glyma06G147100	Vang0605s00070	82.53	59.7	192.3	0.3559	0.1045	0.2936
Glyma06G242200	Vang01g01490	81.26	130.8	496.2	0.3609	0.1738	0.4815
Glyma06G320700	Vang0039ss00390	85.36	556	1667	0.3403	0.0588	0.1727
Glyma08G011300	Vang10g07150	88.18	75.7	365.3	0.7634	0.2505	0.3282
Glyma08G018300	Vang10g06430	81.24	215.1	651.9	0.483	0.0792	0.164
Glyma08G021900	Vang10g06010	85.41	281.4	789.6	0.2131	0.0715	0.3357
Glyma08G082400	Vang0333s00130	96.54	95.9	327.1	0.4903	0.2024	0.4129
Glyma08G118200	Vang0013ss00970	87.21	184.3	589.7	51.5741	0.9462	0.0183
Glyma08G143400	Vang10g04840	83.13	164.5	513.5	0.2383	0.0717	0.3008
Glyma08G240800	Vang04g00270	89.42	453.5	1097.5	0.3012	0.04	0.1328
Glyma09G029800	Vang03g08430	81.46	267.8	812.2	0.3241	0.1003	0.3094
Glyma09G061900	Vang09g06970	89.22	134.8	408.2	0.5129	0.0517	0.1007
Glyma09G254400	Vang04g03920	81.14	124	416	0.3516	0.0555	0.1579
Glyma09G280200	Vang04g08110	85.14	348.7	1073.3	0.3528	0.0879	0.249
Glyma10G032900	Vang06g17510	95.23	205.8	715.2	0.4278	0.1217	0.2845
Glyma11G053100	Vang07g02340	83.81	188.8	576.2	0.5302	0.0479	0.0904
Glyma12G212300	Vang0304s00120	83.46	202.1	562.9	0.2891	0.0725	0.2508
Glyma13G102000	Vang11g13040	82.42	252.1	4.9	0.4874	0.0828	0.1698
Glyma13G289400	Vang0304s00120	83.19	192.7	572.3	0.2598	0.0756	0.2909
Glyma13G310100	Vang09g01900	82.09	395.6	1146.4	0.3403	0.0727	0.2137
Glyma13G370100	Vang0027ss00340	84.47	230.5	675.5	0.3449	0.0732	0.2123
Glyma14G006800	Vang01g00540	83.26	345.4	1127.6	0.2229	0.0838	0.376
Glyma14G200200	Vang01g03800	90.14	344.3	1092.7	0.3251	0.0621	0.1909
Glyma15G003300	Vang0027ss00340	85.25	215.2	699.8	0.3865	0.075	0.1941
Glyma15G135600	Vang03g08430	86.43	266.2	813.8	0.2691	0.048	0.1782
Glyma15G168200	Vang09g06970	89.16	130.1	421.9	0.5157	0.1211	0.2348
Glyma16G031400	Vang08g00900	93.17	97.1	304.9	0.2921	0.0542	0.1856
Glyma16G031900	Vang08g01000	81.59	156	402	0.2689	0.0819	0.3045
Glyma16G054400	Vang04g12730	83.07	106.9	463.1	0.6239	0.071	0.1137
Glyma16G176700	Vang06g08640	83.57	205.5	616.5	23.5547	1.597	0.0678
Glyma16G219800	Vang1880s00010	87.43	43.2	181.8	0.5683	0.2187	0.3848
Glyma17G057100	Vang11g13040	81.31	257.4	690.6	0.4482	0.0735	0.1639
Glyma18G183100	Vang0051s00140	82.17	217.1	658.9	0.3314	0.1144	0.3452
Glyma18G208800	Vang04g08110	86.62	330.1	1094.9	0.3436	0.0796	0.2318
Glyma18G238600	Vang04g03920	82.39	123.1	416.9	0.3456	0.0527	0.1524
Glyma18G242000	Vang04g04330	83.42	379.7	1321.3	0.4418	0.1227	0.2778

Table 2. (*Continued*).

Orthologous gene pairs		Protein identity (%)	*S*	*N*	*dS*	*dN*	*dN/dS*
Glyma18G263400	Vang04g00270	90.61	448.6	1102.4	0.2823	0.0397	0.1405
Glyma19G094100	Vang04g12730	83.26	114.6	446.4	0.4768	0.0657	0.1379
Glyma19G177400	Vang04g14650	86.21	313.8	1096.2	0.3611	0.0689	0.1907
Glyma20G030500	Vang1037s00010	81.22	137.3	351.7	0.3824	0.0911	0.2383
CcWRKY86	Vang01g01490	82.53	370	1253	0.3222	0.1015	0.3149
CcWRKY85	Vang0032ss02430	81.24	451.9	1288.1	0.2809	0.0814	0.2897
CcWRKY61	Vang04g08110	87.13	346.1	1075.9	0.2853	0.0768	0.2693
CcWRKY12	Vang01g03800	89.43	341.8	1086.2	0.3647	0.0542	0.1486
CcWRKY21	Vang04g14650	86.19	324.4	1073.6	0.3865	0.0683	0.1767
CcWRKY94	Vang09g01900	81.54	381.7	1106.3	0.3608	0.0605	0.1676
CcWRKY29	Vang04g00270	83.73	348.9	956.1	0.386	0.0368	0.0954
CcWRKY35	Vang10g06010	86.51	291	792	0.2846	0.0709	0.2489
CcWRKY28	Vang04g04330	85.59	314.2	1080.8	0.4815	0.1217	0.2528
CcWRKY50	Vang0103s00010	82.54	239.6	840.4	0.3191	0.0857	0.2685
CcWRKY14	Vang01g02960	84.21	254.9	726.1	0.4479	0.0681	0.1521
CcWRKY34	Vang10g06430	82.23	210.2	686.8	0.4149	0.0805	0.194
CcWRKY13	Vang01g02180	81.14	213.2	725.8	0.4342	0.0574	0.1322
CcWRKY64	Vang04g16950	82.71	186.5	548.5	0.3382	0.0717	0.2121
CcWRKY77	Vang07g02340	85.47	195	573	0.5342	0.053	0.0991
CcWRKY24	Vang10g04840	84.64	170.3	504.7	0.2328	0.0507	0.2179
CcWRKY79	Vang0304s00120	83.53	182.3	567.7	0.3348	0.0616	0.1839
CcWRKY60	Vang08g06450	84.61	156.4	452.6	0.5111	0.0711	0.1391
CcWRKY27	Vang0005s00450	88.46	283.7	700.3	0.5151	0.0958	0.186
CcWRKY36	Vang09g06970	87.13	114.7	431.3	0.8636	0.1512	0.1751
CcWRKY83	Vang0322s00110	91.26	140	583	0.7537	0.2035	0.27
CcWRKY76	Vang04g12730	87.58	80.1	306.9	0.4974	0.0385	0.0774
CcWRKY75	Vang0333s00130	96.21	74	349	0.6218	0.1466	0.2358
CcWRKY04	Vang08g00900	91.42	92.8	309.2	0.2894	0.0675	0.2332
CcWRKY06	Vang05g08140	82.18	135.3	518.7	2.6702	0.185	0.0693
CcWRKY68	Vang08g00900	93.35	78.1	287.9	0.8689	0.2299	0.2646
CcWRKY89	Vang10g07150	86.12	72.3	338.7	0.9673	0.2583	0.267
CcWRKY25	Vang0322s00110	86.56	260.9	891.1	54.697	0.7937	0.0145
CcWRKY82	Vang04g17060	82.43	165.7	569.3	43.0101	0.6861	0.016
CcWRKY88	Vang0942s00010	81.64	135.7	497.3	58.4138	0.6047	0.0104
CcWRKY51	Vang0322s00110	82.22	230.1	780.9	54.659	0.7217	0.0132
CcWRKY52	Vang01g12760	83.17	173.7	525.3	50.2306	0.6929	0.0138
CcWRKY39	Vang01g12760	81.24	170.9	528.1	50.6758	0.8076	0.0159
CcWRKY71	Vang01g12760	83.76	170.1	528.9	51.1333	0.7335	0.0143
CcWRKY67	Vang0304s00120	82.48	139.5	415.5	49.6073	0.7088	0.0143
CcWRKY03	Vang0013ss00970	95.26	211.6	619.4	45.349	1.9509	0.043

* *S*—Number of synonymous sites, *N*—Number of non-synonymous sites, *dS*—Synonymous substitution rate, *dN*—Non-synonymous substitution rate.

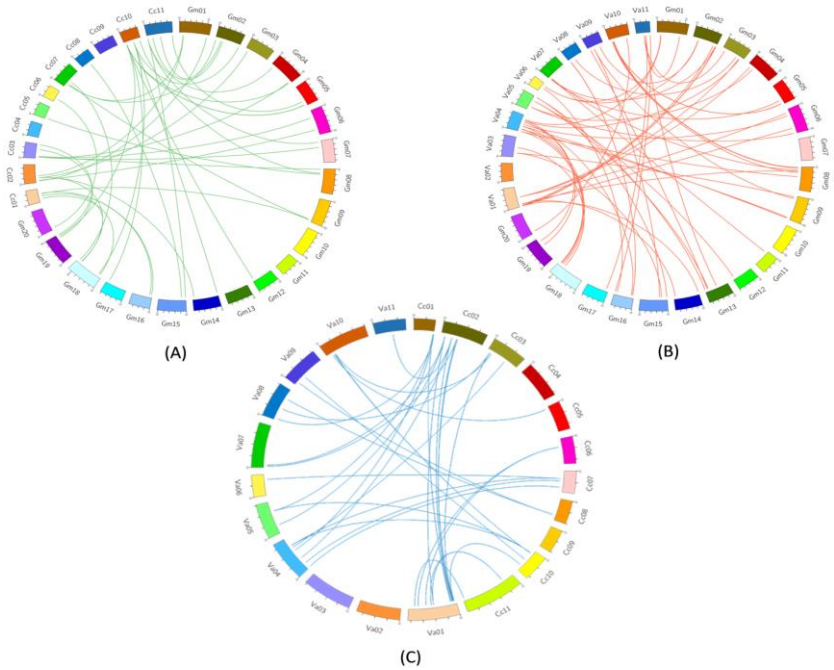

Figure 5. Orthologous WRKY genes between the *Cajanus cajan* and *Glycine max and Vigna angularis*. Cc, Gm, and Va indicates the chromosomes of *Cajanus cajan, Glycine max* as well as *Vigna angularis* and each chromosome represented with different color. **(A)** Green color lines indicate the homologous genes pairs between the pigeonpea and soybean chromosomes; **(B)** red color lines represents the homologous genes pairs between soybean and adzuki bean chromosomes; and **(C)** blue color lines indicates the homologous gene pairs between pigeonpea and adzuki bean.

4. Discussion

Plants are constantly being challenged by invaders that influence their growth and development processes. To cope with invaders plants,use their sentries and WRKY plays an important role. Therefore, in this analysis, an attempt was made to identify the distribution of WRKY throughout the plant lineages. Publicly available genomes of 40 plant species whose genomes were completely sequenced were used. Interestingly, some plant genomes such as *S. italica* (105), *Z. mays* (135), *S. bicolor* (133), *O. sativa* (107), *T. aestivum* (172), *R. sativus* (126), *B. rapa* (145), *B. oleracea* (148), *P. trichocarpa* (102), *M. domestica* (127), *G. raimondii* (120), and *G. hirsutum* (197) contained more than hundreds of WRKYs (**Figure 3**). In our current study, WRKY homologs have been investigated in 40 plant species, ranging from unicellular algae to higher plants, revealing their functional importance and ancient origin. Only one or two WRKY homologs are predicted for the three aquatic algae, and 29 to 197 homologs are predicted in terrestrial plants, indicating rapid gene expansion of the WRKYgene family in higher plants (especially angiosperms).

Phylogenetic analysis of selected plant species revealed that the WRKY-TF is present in monocotyledons, dicotyledons, and lower eukaryotes. This means that the origin of most WRKY-TFs in plants precedes the divergence of these species. No species-specific groups, subgroups, or clades have been observed in the phylogenetic tree. This indicates that the WRKY gene family is more conserved during evolution.

In addition, WRKY domains from similar ancestors tended to cluster together in the phylogenetic tree, which was not observed in this study. This suggests that they accomplished duplication after divergence. The WRKY genes that clustered together are orthologous ones and also evolutionarily closer than others. The phylogenetic relationship found in this study showed that WRKY TFs may evolve conservatively. Only a few WRKY genes were identified in lower eukaryotes, including *O. lucimarinus*, *V. carteri*, and *C. reinhardtii*, while higher plants had a larger number of WRKY genes. This suggests that the most primitive evolutionary origin of the genes containing the WRKY-TF was from aquatic algae. This suggests that the WRKY protein evolved before plants transitioned from an aquatic to a terrestrial habitat. With the uninterrupted evolution of plant species, terrestrial plants have evolved a series of highly sophisticated signaling processes that help them to adjust to the ever-changing environmental circumstances, and hence, the number of WRKY-TFs increased in different plant species.

Gene family expansion is generally due to gene duplication events commonly occurring in plants and is considered to be an important evolutionary force in the expansion and evolution of gene families that are the source of new biological functions. Estimation of the recent gene duplication events, including segmental, tandem, proximal and dispersed duplication was conducted that led to the expansion of the WRKYgene family in flowering plants. Dispersed type of duplication models mainly contributes to genetic novelty and adaptation to new environments. The results indicate that distinct duplication patterns were observed, that implies a range of functional divergence between different plant species or taxa. The tandem type duplication was observed in most of the plant species except *Malus domestica* and aquatic algae. Furthermore, only 12 segmental duplicated *FvWRKY* have been identified in the *Fragaria vesca* genome, suggesting that tandem and segmental duplication events occur at low levels in the *FvWRKY* gene family. Therefore, the number of WRKY gene members may be lesser due to the absence of segmental and tandem duplication events in several plant genomes. Based on our results, we suggested that the ancestral core WRKY genes possibly mainly dispersed duplications. Consequently, segmental and tandem duplications mainly contributed to the expansion of the WRKY gene family in angiosperms. The duplication and expansion of genes followed by functional diversification, and diversification of gene function can play a significant role in providing novel genes for adaptation to new environments. The expansion of the WRKY gene family, and their distinct roles in various processes, evidently indicate their functional variations and evolutionary history.

The orthologous WRKY genes between different plants are usually supposed to maintain related properties and provide other important functions. Therefore, comparative analysis of the WRKY orthologous genes between legume plants might be helpful in predicting their hereditary relationship and possible functions of the WRKY protein in pigeon pea, soybean, and adzuki bean. Our study identified 42 orthologous gene pairs between pigeon pea and soybean, 60 orthologous gene pairs between soybean and adzuki bean, and 36 orthologous gene pairs between pigeon pea and adzuki bean, which gives a very important key for further functional prediction of WRKY genes in pigeon pea, soybean, and adzuki bean.

5. Conclusion

Overall, our comparative genomic analysis of the WRKY gene family provides a comprehensive framework for understanding the evolutionary dynamics, duplication patterns, and selection pressures shaping this important gene family across diverse plant lineages. This study not only enhances our knowledge of plant evolution but also paves the way for targeted approaches to harness the potential of WRKY genes in crop improvement and environmental resilience.

Author contributions: Conceptualization, AS; methodology, AS; software, AS; formal analysis, AS; investigation, AS; data curation, AS; writing—original draft preparation, AS; writing—review and editing, AKS; supervision, AKS. All authors have read and agreed to the published version of the manuscript.

Conflict of interest: The authors declare no conflict of interest.

References

1. Rejeb I, Pastor V, Mauch-Mani B. Plant Responses to Simultaneous Biotic and Abiotic Stress: Molecular Mechanisms. Plants. 2014; 3(4): 458-475. doi: 10.3390/plants3040458
2. Phukan UJ, Jeena GS, Shukla RK. WRKY Transcription Factors: Molecular Regulation and Stress Responses in Plants. Frontiers in Plant Science. 2016; 7. doi: 10.3389/fpls.2016.00760
3. Bakshi M, Oelmüller R. WRKY transcription factors. Plant Signaling & Behavior. 2014; 9(2): e27700. doi: 10.4161/psb.27700
4. Ishiguro S, Nakamura K. Characterization of a cDNA encoding a novel DNA-binding protein, SPF1, that recognizes SP8 sequences in the 5′ upstream regions of genes coding for sporamin and β-amylase from sweet potato. Molecular and General Genetics MGG. 1994; 244(6): 563-571. doi: 10.1007/bf00282746
5. Rushton PJ, Macdonald H, Huttly AK, et al. Members of a new family of DNA-binding proteins bind to a conserved cis-element in the promoters of ?-Amy2 genes. Plant Molecular Biology. 1995; 29(4): 691-702. doi: 10.1007/bf00041160
6. Rushton PJ, Somssich IE, Ringler P, et al. WRKY transcription factors. Trends in Plant Science. 2010; 15(5): 247-258. doi: 10.1016/j.tplants.2010.02.006
7. Eulgem T, Rushton PJ, Robatzek S, et al. The WRKY superfamily of plant transcription factors. Trends in Plant Science 2000; 5(5): 199-206. doi: 10.1016/s1360-1385(00)01600-9
8. Jiang J, Ma S, Ye N, et al. WRKY transcription factors in plant responses to stresses. Journal of Integrative Plant Biology. 2017; 59(2): 86-101. doi: 10.1111/jipb.12513
9. QIU Y. Cloning and analysis of expression profile of 13 WRKY genes in rice. Chinese Science Bulletin. 2004; 49(20): 2159. doi: 10.1360/982004-183
10. Wu H, Ni Z, Yao Y, et al. Cloning and expression profiles of 15 genes encoding WRKY transcription factor in wheat (Triticum aestivem L.). Progress in Natural Science. 2008; 18(6): 697-705. doi: 10.1016/j.pnsc.2007.12.006
11. Zhou Q, Tian A, Zou H, et al. Soybean WRKY-type transcription factor genes,GmWRKY13, GmWRKY21, andGmWRKY54, confer differential tolerance to abiotic stresses in transgenicArabidopsisplants. Plant Biotechnology Journal. 2008; 6(5): 486-503. doi: 10.1111/j.1467-7652.2008.00336.x
12. Hofmann MG, Sinha AK, Proels RK, et al. Cloning and characterization of a novel LpWRKY1 transcription factor in tomato. Plant Physiology and Biochemistry. 2008; 46(5-6): 533-540. doi: 10.1016/j.plaphy.2008.02.009
13. Agarwal P, Dabi M, Agarwal PK. Molecular Cloning and Characterization of a Group II WRKY Transcription Factor fromJatropha curcas, an Important Biofuel Crop. DNA and Cell Biology. 2014; 33(8): 503-513. doi: 10.1089/dna.2014.2349
14. Jiang W, Yu D. Arabidopsis WRKY2 transcription factor mediates seed germination and postgermination arrest of development by abscisic acid. BMC Plant Biology. 2009; 9(1): 96. doi: 10.1186/1471-2229-9-96

15. Chen J, Nolan T, Ye H, et al. Arabidopsis WRKY46, WRKY54 and WRKY70 Transcription Factors Are Involved in Brassinosteroid-Regulated Plant Growth and Drought Response. The Plant Cell. Published online June 2, 2017: tpc.00364.2017. doi: 10.1105/tpc.17.00364
16. Wang H, Avci U, Nakashima J, et al. Mutation of WRKY transcription factors initiates pith secondary wall formation and increases stem biomass in dicotyledonous plants. Proceedings of the National Academy of Sciences. 2010; 107(51): 22338-22343. doi: 10.1073/pnas.1016436107
17. Johnson CS, Kolevski B, Smyth DR. Transparent testa glabra2, a Trichome and Seed Coat Development Gene of Arabidopsis, Encodes a WRKY Transcription Factor. The Plant Cell. 2002; 14(6): 1359-1375. doi: 10.1105/tpc.001404
18. Zhang CQ, Xu Y, Lu Y, et al. The WRKY transcription factor OsWRKY78 regulates stem elongation and seed development in rice. Planta. 2011; 234(3): 541-554. doi: 10.1007/s00425-011-1423-y
19. Xiang J, Tang S, Zhi H, et al. Loose Panicle1 encoding a novel WRKY transcription factor, regulates panicle development, stem elongation, and seed size in foxtail millet [Setaria italica (L.) P. Beauv.]. PLOS ONE. 2017; 12(6): e0178730. doi: 10.1371/journal.pone.0178730
20. Miao Y, Laun T, Zimmermann P, et al. Targets of the WRKY53 transcription factor and its role during leaf senescence in Arabidopsis. Plant Molecular Biology. 2004; 55(6): 853-867. doi: 10.1007/s11103-005-2142-1
21. Liu H, Sun M, Du D, et al. Whole-Transcriptome Analysis of Differentially Expressed Genes in the Vegetative Buds, Floral Buds and Buds of Chrysanthemum morifolium. PLOS ONE. 2015; 10(5): e0128009. doi: 10.1371/journal.pone.0128009
22. Chen L, Zhang L, Li D, et al. WRKY8 transcription factor functions in the TMV-cg defense response by mediating both abscisic acid and ethylene signaling in Arabidopsis. Proceedings of the National Academy of Sciences. 2013; 110(21). doi: 10.1073/pnas.1221347110
23. Ren X, Chen Z, Liu Y, et al. ABO3, a WRKY transcription factor, mediates plant responses to abscisic acid and drought tolerance in Arabidopsis. The Plant Journal. 2010; 63(3): 417-429. doi: 10.1111/j.1365-313x.2010.04248.x
24. Wu ZJ, Li XH, Liu ZW, et al. Transcriptome-wide identification of Camellia sinensis WRKY transcription factors in response to temperature stress. Molecular Genetics and Genomics. 2015; 291(1): 255-269. doi: 10.1007/s00438-015-1107-6
25. Chen L, Yang Y, Liu C, et al. Characterization of WRKY transcription factors in Solanum lycopersicum reveals collinearity and their expression patterns under cold treatment. Biochemical and Biophysical Research Communications. 2015; 464(3): 962-968. doi: 10.1016/j.bbrc.2015.07.085
26. Wu J, Chen J, Wang L, et al. Genome-Wide Investigation of WRKY Transcription Factors Involved in Terminal Drought Stress Response in Common Bean. Frontiers in Plant Science. 2017; 8. doi: 10.3389/fpls.2017.00380
27. Singh A, Singh PK, Sharma AK, et al. Understanding the Role of the WRKY Gene Family under Stress Conditions in Pigeonpea (Cajanus Cajan L.). Plants. 2019; 8(7): 214. doi: 10.3390/plants8070214
28. Deslandes L, Olivier J, Theulières F, et al. Resistance to Ralstonia solanacearum in Arabidopsis thaliana is conferred by the recessive RRS1-R gene, a member of a novel family of resistance genes. Proceedings of the National Academy of Sciences. 2002; 99(4): 2404-2409. doi: 10.1073/pnas.032485099
29. Zheng Z, Qamar SA, Chen Z, et al. Arabidopsis WRKY33 transcription factor is required for resistance to necrotrophic fungal pathogens. The Plant Journal. 2006; 48(4): 592-605. doi: 10.1111/j.1365-313x.2006.02901.x
30. Yang Y, Zhou Y, Chi Y, et al. Characterization of Soybean WRKY Gene Family and Identification of Soybean WRKY Genes that Promote Resistance to Soybean Cyst Nematode. Scientific Reports. 2017; 7(1). doi: 10.1038/s41598-017-18235-8
31. Huh SU, Choi LM, Lee GJ, et al. Capsicum annuum WRKY transcription factor d (CaWRKYd) regulates hypersensitive response and defense response upon Tobacco mosaic virus infection. Plant Science. 2012; 197: 50-58. doi: 10.1016/j.plantsci.2012.08.013
32. Li P, Song A, Gao C, et al. The over-expression of a chrysanthemum WRKY transcription factor enhances aphid resistance. Plant Physiology and Biochemistry. 2015; 95: 26-34. doi: 10.1016/j.plaphy.2015.07.002
33. Eddy SR. Profile hidden Markov models. Bioinformatics. 1998; 14(9): 755-763. doi: 10.1093/bioinformatics/14.9.755
34. Larkin MA, Blackshields G, Brown NP, et al. Clustal W and Clustal X version 2.0. Bioinformatics. 2007; 23(21): 2947-2948. doi: 10.1093/bioinformatics/btm404
35. Tamura K, Stecher G, Peterson D, et al. MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. Molecular Biology and Evolution. 2013; 30(12): 2725-2729. doi: 10.1093/molbev/mst197
36. Wang Y, Tang H, DeBarry JD, et al. MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. Nucleic Acids Research. 2012; 40(7): e49-e49. doi: 10.1093/nar/gkr1293

37. Patil G, Valliyodan B, Deshmukh R, et al. Soybean (Glycine max) SWEET gene family: insights through comparative genomics, transcriptome profiling and whole genome re-sequence analysis. BMC Genomics. 2015; 16(1). doi: 10.1186/s12864-015-1730-y
38. Suyama M, Torrents D, Bork P. PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. Nucleic Acids Research. 2006; 34(Web Server): W609-W612. doi: 10.1093/nar/gkl315
39. Vanneste K, Van de Peer Y, Maere S. Inference of Genome Duplications from Age Distributions Revisited. Molecular Biology and Evolution. 2012; 30(1): 177-190. doi: 10.1093/molbev/mss214
40. Rensing SA, Ick J, Fawcett JA, et al. An ancient genome duplication contributed to the abundance of metabolic genes in the moss Physcomitrella patens. BMC Evolutionary Biology. 2007; 7(1): 130. doi: 10.1186/1471-2148-7-130
41. Albert VA, Barbazuk WB, et al. The Amborella Genome and the Evolution of Flowering Plants. Science. 2013; 342(6165). doi: 10.1126/science.1241089
42. Banks JA, Nishiyama T, Hasebe M, et al. The Selaginella Genome Identifies Genetic Changes Associated with the Evolution of Vascular Plants. Science. 2011; 332(6032): 960-963. doi: 10.1126/science.1203810
43. Ning P, Liu C, Kang J, et al. Genome-wide analysis of WRKY transcription factors in wheat (*Triticum aestivum* L.) and differential expression under water deficit condition. PeerJ. 2017; 5: e3232. doi: 10.7717/peerj.3232