J. Plant Environ. 05 (02) 2023. 109-122 DOI: 10.33687/jpe.005.02.4490

Available Online at EScience Press
Journal of Plant and Environment
ISSN: 2710-1665 (Online), 2710-1657 (Print)
https://esciencepress.net/journals/JPE

Genome-Wide Computational Analysis of Dirrigent Proteins in *Solanum Tuberosum*

Muhammad Abu Bakar Saddique[1*], Sana Abbas[2], Muhammad Dawood Amjad[3], Muhammad Faizan Khurram Maqsood[4], Zahid Hussain[5]

[1]*Department of Plant Biotechnology, National University of Sciences and Technology, Islamabad, Pakistan.*
[2]*Department of Mathematics & Statistics. Pir Mehr Ali Shah Arid Agriculture University, Rawalpindi, Pakistan.*
[3]*Department of Medical, Oral and Biotechnological Sciences, Università degli Studi 'G. d'Annunzio' Chieti – Pescara, Italy.*
[4]*Centre of Agricultural Biochemistry and Biotechnology (CABB), University of Agriculture, Faisalabad, Pakistan.*
[5]*Department of Plant Breeding and Genetics, University of Agriculture, Faisalabad, Pakistan.*

ARTICLE INFO

Article History
Received: June 29, 2023
Revised: October 03, 2023
Accepted: October 22, 2023

Keywords
Dirrigent proteins
Genome-wide study
Bioinformatics
Plant biology

ABSTRACT

Dirrigent proteins are required for the synthesis of lignans, a distinct and widely distributed class of plant-derived secondary metabolites with promising pharmacologic characteristics and a potential role in plant defense. However, no detailed data on the Dirrigent family of *Solanum tuberosum* is known. Comparatively, 33 DIR sequences were explored in *Solanum tuberosum* in this research study. There are no introns in the majority of StDIR genes. A single Dirigent domain is found in all StDIR proteins. The abundance of amino acids, sequence similarity analysis, phylogenetic analysis, genomic location, gene structure, conserved domain analysis, evolutionary analysis, and gene promoter assessment are all studied. This study lays a foundation for potential plant genetic engineering and crop improvement research by providing an in-depth and thorough explanation of the detailed molecular mechanism and structural characterization of StDIR proteins in the genome of *Solanum tuberosum*. This work will provide useful data for enabling the proper selection of dirrigent proteins in higher plants and will support future studies on the DIR gene family.

Corresponding Author: Muhammad Abu Bakar Saddique
Email: abubakarrana8899@gmail.com

INTRODUCTION

One of the biggest families in the Plantae kingdom is the Solanaceae family, with over 3000 different species (Knapp, 2002). Potato, Bell pepper, Tomato, Tobacco, and Brinjal are all members of this important family. Solanum is an example of a large genus that contains up to 1700 species. This is a widely distributed class in the Solanaceae plant family and also encompasses the major top ten species-rich phanerogam lineages (Frodin, 2004). Potatoes are a healthy, versatile food that can be used in a variety of ways. Furthermore, it adds to the fight against different inadequacies, named "hidden hunger," which might pose a serious threat to health worldwide, having an expected impact on twenty thousand lakh people, globally (Bailey *et al*., 2015). Potatoes contain a high amount of nutrients that are good for human health, such as water-soluble vitamins, synthetic resin, and supermolecules (Burlingame *et al*., 2009).

For the very first time, a preserved motif of the DIR gene was studied in the specie *Forsythia intermedia* (Davin *et al*., 1997). Dirrigent is based on the Latin word dirigere, which means "to instruct." Recently, DIR proteins have been studied in multiple taxonomic groups,

such as mosses, ferns, angiosperms, and gymnosperms. (Wu *et al*., 2009; Li *et al*., 2014). To date, no research relating to the presence of DIR proteins in early aquatic plants has been ascertained. The Dirrigent sequences arise by way of a consequence of vascular plant evolution in terrestrial systems (Shi *et al*., 2012). The scientific community has shown a growing interest in the DIR proteins due to their potential involvement in the chemical production of metabolites, particularly in response to pests and abiotic challenges. (Davin *et al*., 1997; Hosmani *et al*., 2013).

Several DIR proteins have already been structurally studied in several species, with the majority of them are linked to the fighting against diseases, insect infestation, abiotic stress resilience, or indeed combinatorial defense mechanisms. A DIR1 gene in *Pisum sativum* (Hadwiger *et al*., 1992), 6 Dirrigent genes in *Pinus* tree (Ralph *et al*. 2006), 4 homologous genes including 1 BrDIR2 in turnip (Thamil *et al*., 2013), DIR1 in *Gossypium hirsutum* (Shi *et al*., 2012), while in *Triticum aestivum* (Subramanyam *et al*., 2006, 2008; Song *et al*., 2013; Ma and Liu, 2015) Hfr1, TaDIR13 and having Dir domain in JRL genes, Another DIR1 in *Boea hygrometrica* (Wu *et al*., 2009), DIR22 in *Glycine max* (Li *et al*., 2017), DIR factor in *Tamarix androssowii* (Gao *et al*., 2010) and another DIR factor in Sugarcane (Guo *et al*., 2012).

Various Dirrigent proteins, including SHDIR11 and SHDIR16 transcription in *Saccharum officinarum* triggered by salicylic acid, Jasmonic acid, and Methyl Jasmonate, also react to the production of disease-associated hormonal substances (Damaj *et al*., 2010). It is also reported that 12 DIR proteins in *Brassica rapa* reacted to abscisic acid as well (Thamil Arasan *et al*., 2013).

Crop defense responses to biotic and abiotic threats may be better understood and improved by understanding the functions of DIR transcription factors in plant physiological and biological systems. Still, there is no detailed research on the DIR gene family in *Solanum tuberosum* has been done yet. Keeping this gap in mind, we evaluated the DIR family throughout the whole *Solanum tuberosum* genome in order to develop a clearer and more simple method for studying the DIR family.

MATERIAL AND METHODS

Retrieval of *Solanum tuberosum* Dirrigent genes

Two methodologies were employed to identify dirrigent sequences in *Solanum tuberosum*. Initially, the seed files for the Dirrigent proteins (PF03018) hidden Markov model profile from the Pfam database were obtained, (https://pfam.sanger.ac.uk/) and employed as questionary genes in HmmerWeb kind 3.3 (https://www.ebi.ac.uk/Tools/hmmer/) to study for amino-acids residues while utilizing access defaults settings. Second, TAIR (The Arabidopsis Information Resource) provided access to all Arabidopsis thalian DIR genes, which were used in BLAST database searches. Phytozome12.1 (https://phytozome.jgi.doe.gov/), with manual settings, against the *Solanum tuberosum* v4.03 proteome was also employed (Lamesch *et al*., 2012). The elimination of redundant and unnecessary sequences was carried out in a stepwise manner, and the conserved pattern of the protein was verified through automated means utilizing the CDD database (Marchler-Bauer *et al*., 2007). (https://www.ncbi.nlm.nih.gov/Structure/bwrpsb/bwrpsb.cgi).

Analysis and Structural Characterization of StDIR Genes

Expert Protein Analysis System (ExPAsy) site (https://www.expasy.org/) (Gasteiger *et al*., 2003) used to study iso-electric point, the molecular weight of each sequence, estimation of coding sequence length (CDS), Aliphatic index, Instability index, and GRAVY. Phytozome12.1 (https://phytozome.jgi.doe.gov/) (Lamesch *et al*., 2012) was employed to retrieve full length genomic and as well as CDS sequences, and the gene structures were examined employing the GSDS program (https://gsds.cbi.pku.edu.cn/) (Guo *et al*., 2007; Hu *et al.,* 2015). The subcellular sites were predicted via WoLF PSORT online tool. (Horton *et al*., 2007) (https://wolfpsort.hgc.jp/). The Multidisciplinary EM for the Motif Detection (MEME) (v.12.0; http://meme-suite.org/tools/meme) web service is employed to search for new domains in StDIRs proteins. The default settings for the motif sites are 10 and with site distribution Zero or one occurrence per sequence (Zoops). The MEME findings are displayed using the TBtool program.

Phylogeny of StDIR Genes

All *Arabidopsis thaliana*, *Oryza sativa*, and *Solanum tuberosum* DIR genes were aligned via Clustal W (http://www.ebi.ac.uk/clustalw/), (Thompson *et al*., 2003) retrieved from Phytozome (Lamesch *et al*., 2012). The Neighbor-joining (NJ) method was employed for clustering algorithms, and MEGA X (Kumar *et al*., 2018)

program was utilized to generate a phylogenetic tree with a pairwise distance model and a Bootstrap value of 1000 iterations. The remaining attributes were left at their default settings (http://www.megasoftware.net/).

Chromosomal Localization

By using the PhenoGram website tool, each StDIRs gene was allocated to the appropriate *Solanum tuberosum* chromosomal location (Wolf *et al*., 2013) (http://visualization.ritchielab.org/phenograms/plot). The TBtools program (v1.09854; http://cj-chen.github.io/tbtools/) was utilized to assign each of the StDIRs sequences to its corresponding relative distances, chromosomal length data, and chromosomal locations (Chen *et al*., 2020).

Cis-regulatory Element Analysis in the Promoters of StDIRs

The *Solanum tuberosum* Phytozome Genome Browser was used to obtain the 2000 bp upstream full-length sequences of each StDIR gene from the start site sequence (ATG) and save them in FASTA format for further usage. The web-based tool PLANTCARE (Lescot *et al*., 2002) was utilized to assess cis-acting elements on an individual sequence basis (https://bioinformatics.psb.ugent.be/webtools/plantcare/html/).

Protein-Protein association and 3D Configuration of DIR genes in *Solanum tuberosum*

By putting all *Solanum tuberosum* DIR genes to the freely downloadable software (Szklarczyk *et al*., 2017), the linkage among all DIR genes in *Solanum tuberosum* was created (https://string-db.org/). The accompanying main criteria were used to evaluate specific nodes: i) required confidence value: high (score: 0.07), ii) a highest possible number of connections five with excluded all the un-interlinked sequences. To compare the co-occurrence of all StDIR genes in *Solanum tuberosum* to those in other closely interconnected taxa, they were also visualized using String to show their convergence. According to Kelley and Sternberg (2009), the web server Phyre2 utilized to calculate the tertiary (3D) configurations and homologs of all StDIRs genes (Li *et al*., 2014).
(http://www.sbg.bio.ic.ac.uk/phyre2/html/page.cgi?id=index).

Sequence identity and transmembrane potential analysis of StDIRs

For examining the transmembrane potential mechanism involved in all StDIRs genes, all-sequences in the form of amino-acids were uploaded to the TMHMM Server, version
2.0 (http://www.cbs.dtu.dk/services/TMHMM/) (Käll L *et al*., 2007). The sequence homology evaluation was conducted using the SIAS web tool with default settings of the BLOSUM62 technique and gap penalties. (http://imed.med.ucm.es/Tools/sias.html).

Sequence similarity analysis

Circoletto, an online application that uses Circos to distinguish two sequence libraries, was used to compare DIR genes in the model plant *Arabidopsis thaliana* and subjected plant *Solanum tuberosum*. DIR genes from *Solanum tuberosum* as query were compared to DIR genes with *Arabidopsis thaliana*. From total local matches, to do genome conservation analysis, with E-value of 10 to the -40 (Strict) employed. The StDIR and AtDIR protein structures were used as searches across their individual protein databases, and the highest hits were retrieved depending on E-value.

RESULTS

Identification of DIR Genes in *Solanum tuberosum*

The Dirrigent group is unique and belongs to plants. Dirrigent genes of *Solanum tuberosum* were retrieved from the *Solanum tuberosum* genome on Phytozome. CDD and SMART techniques identified a total of 33 genes after gene manual analysis as shown in Figure 1. The names were then assigned, starting with StDIR1, and finishing with StDIR33. The physicochemical properties of the StDIR genes were studied using the ExPAsy online web. The Additional Information contains data regarding the number of amino acids, Isoelectric point, Molecular weight, GRAVY, Aliphatic Index, Instability Index, and subcellular localization of the StDIR transcripts in the *Solanum tuberosum* tree. Proteins of the StDIR family have a wide range of amino acid compositions and physicochemical properties, and the number of amino acid-forming proteins varies widely within subfamilies. The amino acid lengths of the StDIR genes varied between 147 (StDIR12) to 708 (StDIR25), having an average length of 227 base pairs. The molecular weight ranged from 16310.66 KDa to 79607.82 KDa, with an average MW of 24567.0097 kDa. The estimated average pi was 7.08, with values ranging from 4.29 (StDIR23) to 9.73 (StDIR23) (StDIR21). In 58 percent of the StDIR family proteins, the pi value is less than 7, whereas the rest have a pi value of more than 7. These results indicate that more DIR proteins have an

acidic nature than basic proteins. The Dirrigent motif in all StDIRs proteins was further validated using TBtool. Using a heatmap, the WoLF PSORT results were predicted. Figure 2 highlights the greatest likelihood values in red and a contrasting light blue.

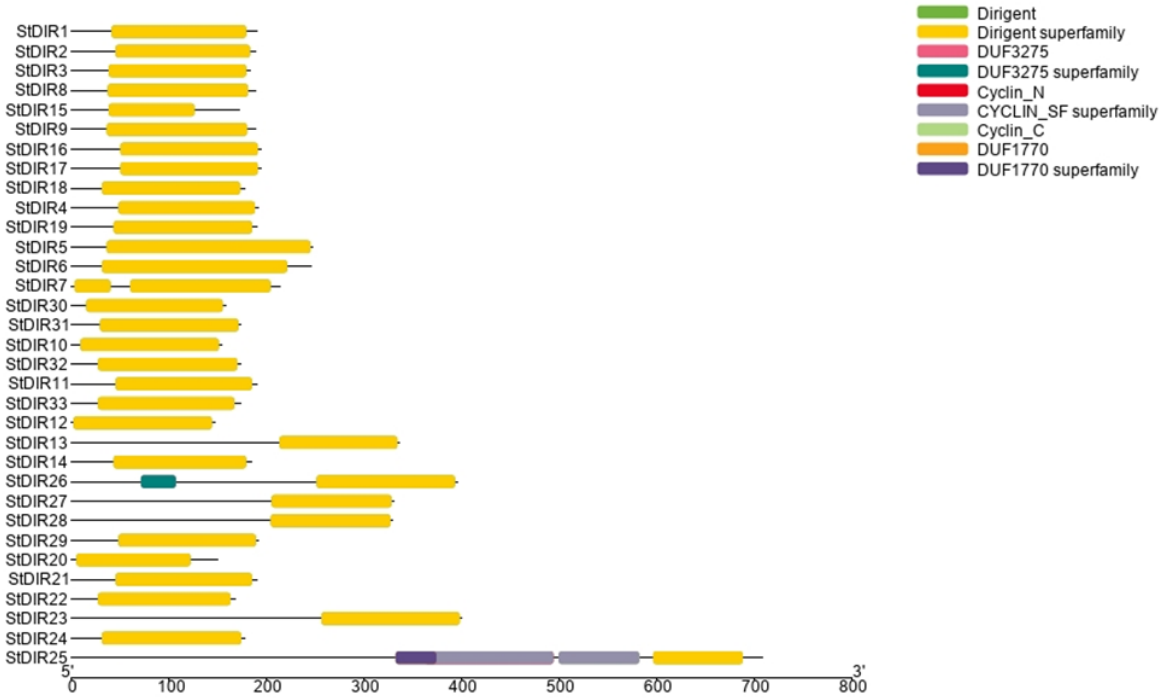

Figure 1. Confirmation of Dirrigent domains in all StDIR genes.

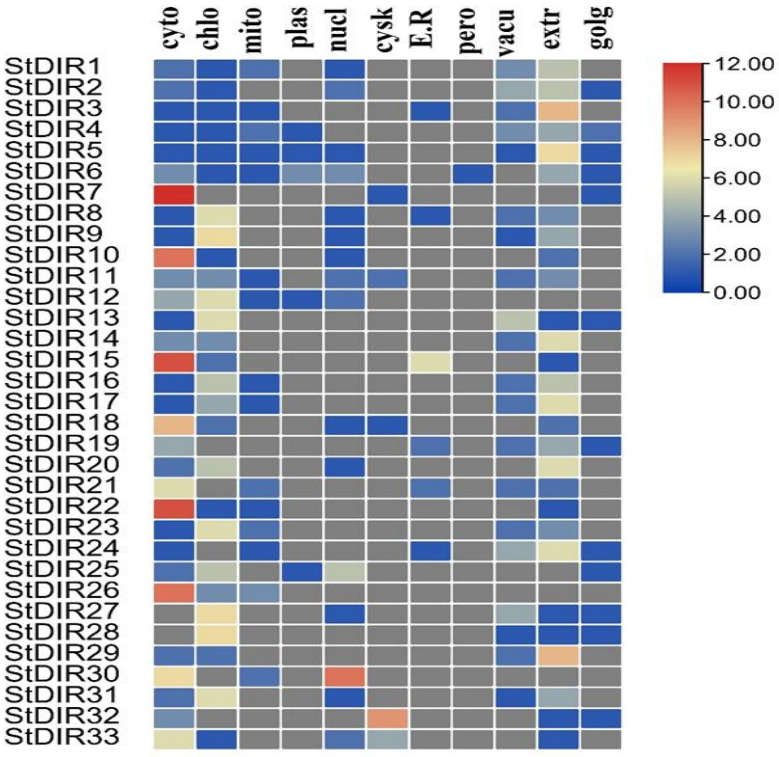

Figure 2. Heatmap interpretation of Subcellular localization of StDIR genes.

Phylogeny of the StDIR gene family

ClustalW was used for comparison of the amino-acids sequences of 33 *Solanum tuberosum* DIR genes, 52 of *Oryza sativa* DIR genes, and 26 from model plant Arabidopsis DIR genes in order to study the evolutionary perspectives, as well as to analyses the characteristics of the *Solanum tuberosum* DIR genes exclusively. The phylogenetic trees were created via MEGA X, minimum evolution, and neighbor-joining (NJ) approaches. The DIR genes of *Solanum tuberosum*, AtDIR, and Oryza sativa were found to be grouped together indicating that the StDIR genes that can be classified with the model plant Arabidopsis and rice subgroup sequences belong to the same subgroup. All of these sequences may be grouped into A-K major subclasses by how closely they resemble *Arabidopsis thaliana* DIR sequences., as in Figure 3.

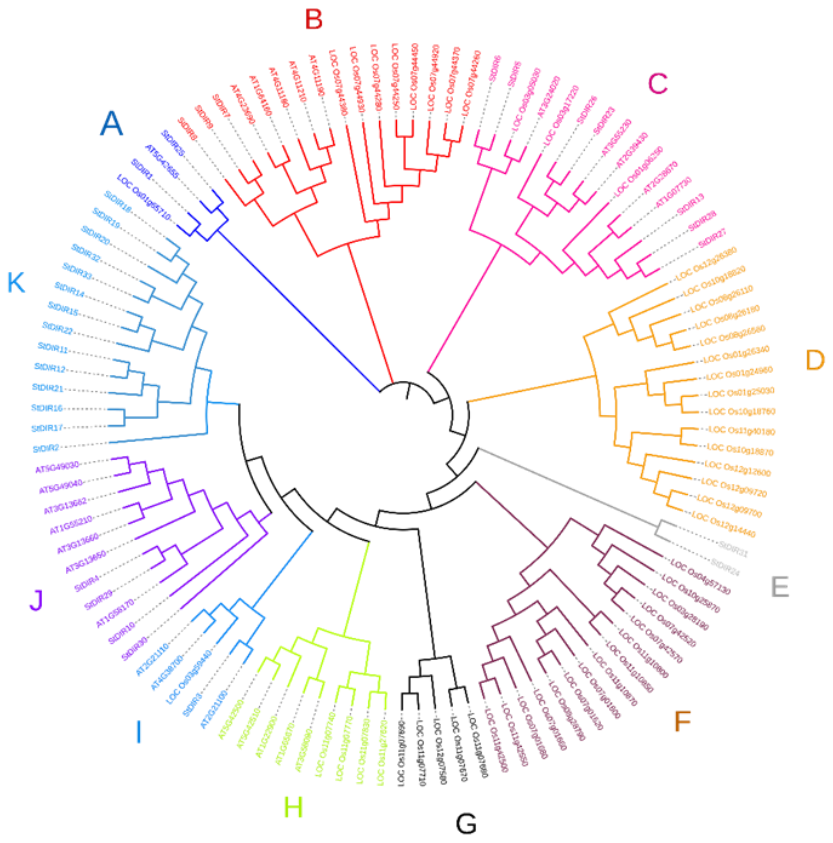

Figure 3. Phylogenetic relationship of *Solanum tuberosum*, *Oryza sativa*, and *Arabidopsis thaliana* DIR genes.

Gene structural characterization and conserved motifs analysis of StDIRs

We studied the DNA sequences of the StDIR genes and the structure of their intronic and exonic regions to get a better understanding of how they are structured. The GSDS 2.0 tool was used to examine the intron-exon development of the *Solanum tuberosum* DIR transcription factors (http://gsds.gao-lab.org/). According to with data, only 7 of the 33 StDIR genes (21 %) contained one intron. It's worth mentioning that 26 of the 33 proteins are intronless. According to this analysis, the single StDIR25 gene contains the most introns to exons, with a 7:8 ratio. StDIR15 is the only gene that has a 3:4 (intron: exon) ratio.

In terms of motif assessment, the group I differs from the others. The number of motifs varies per sequence, with the most motifs being found in StDIR32, StDIR33, StDIR19, StDIR18, StDIR14, StDIR22, StDIR21, StDIR2, StDIR17, and StDIR16. The remaining sequences have between four and five motifs. The members of Group II

have a total of 5 motifs that have been conserved. Group III, on the other hand, shows the greatest range of variances from low to high. It has a total of 2 (StDIR1) to 6 (StDIR27, StDIR28, StDIR13) motifs. Groups IV and V have the same number of motifs, which is four.

The results indicate that while individuals belonging to identical subfamilies within the StDIR family generally exhibit similar quantities and types of motifs, there exist differences in the manner in which motifs are organized among subfamily representatives. The correctness of the phylogenetic tree is enhanced through the identification of analogous gene architectures and conserved patterns within a given subcategory. Conversely, the existence of structural variations among subfamilies suggests the presence of functional diversity within the StDIR gene family in *Solanum tuberosum*.

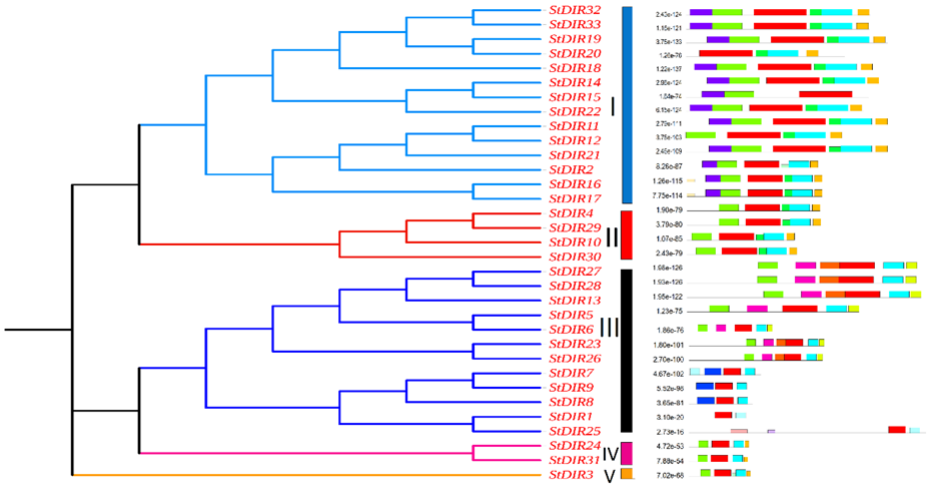

Figure 4. Phylogenetic aspects, Conserved Motif analysis in StDIR genes (A) The phylogeny of StDIR genes was analyzed using MEGA X software with the Neighbor-Joining methodology. (B) Distinct colors represent the diverse conserved motifs domains of StDIR genes.

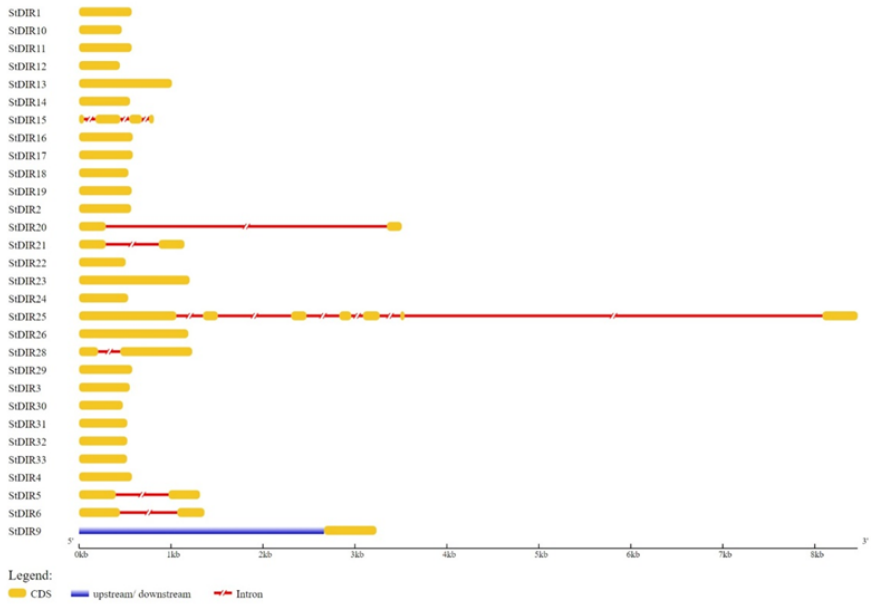

Figure 5. Intron-Exon structure of StDIR genes, yellow boxes showing Exon part and red represents the intronic region with blue color upstream/downstream region.

Promotor Analysis of DIRs in *Solanum tuberosum*

In order to study the potential biochemical processes of StDIR during modulation, maturation, and resilience to abiotic and biotic stress responses, the nucleotide 2000 bp upstream from every transcription start site of the StDIRs was analyzed using PlantCARE to identify cis-acting regions. The study conducted a comprehensive analysis of the responsive components of each gene and observed that they exhibit diverse activities related to biotic and abiotic stress, including plant development, growth, and regulation (Lescot *et al*. 2002).

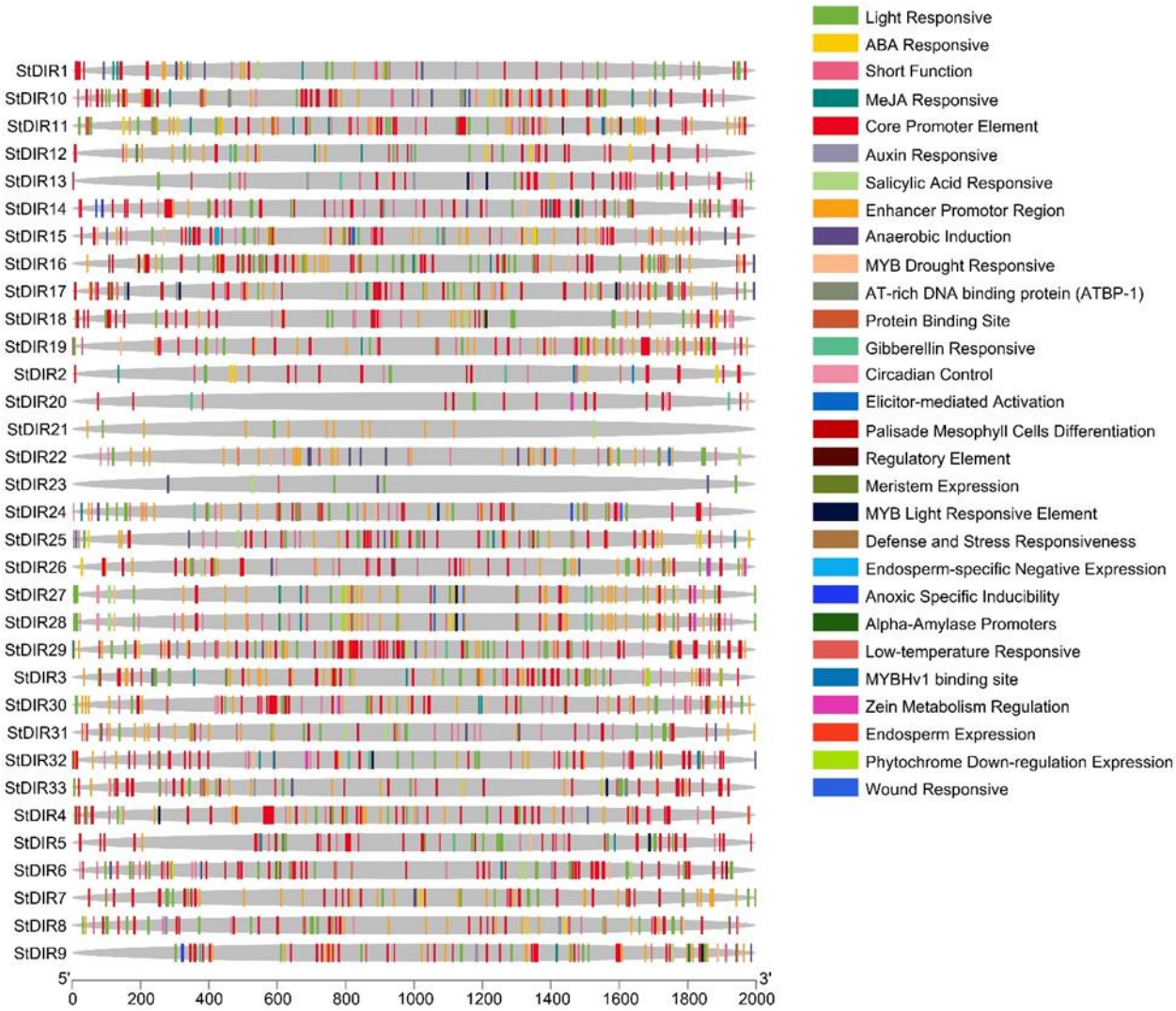

Figure 6. A comprehensive investigation of the promoter regions of all StDIR genes about stress responses and developmental processes.

Protein-Protein linkage association, 3D structures and co-expression of StDIRs

We examined data from an associated protein and co-expression research with interconnected taxa to gain more about our targeted *Solanum tuberosum* StDIR genes. The image indicates that there is a substantial link between genes at various levels revealed by String Browser as shown in Figure 7. As illustrated by all the colorful nodes, the initial shell of interaction can be observed across the junction. It should be emphasized that none of the proteins have a three-dimensional structure. The divergence, conservation, and co-expression of StDIR genes in a set of related species are depicted in Figure 8. The 3D images of all StDIRs proteins were formed as shown in Figure 9.

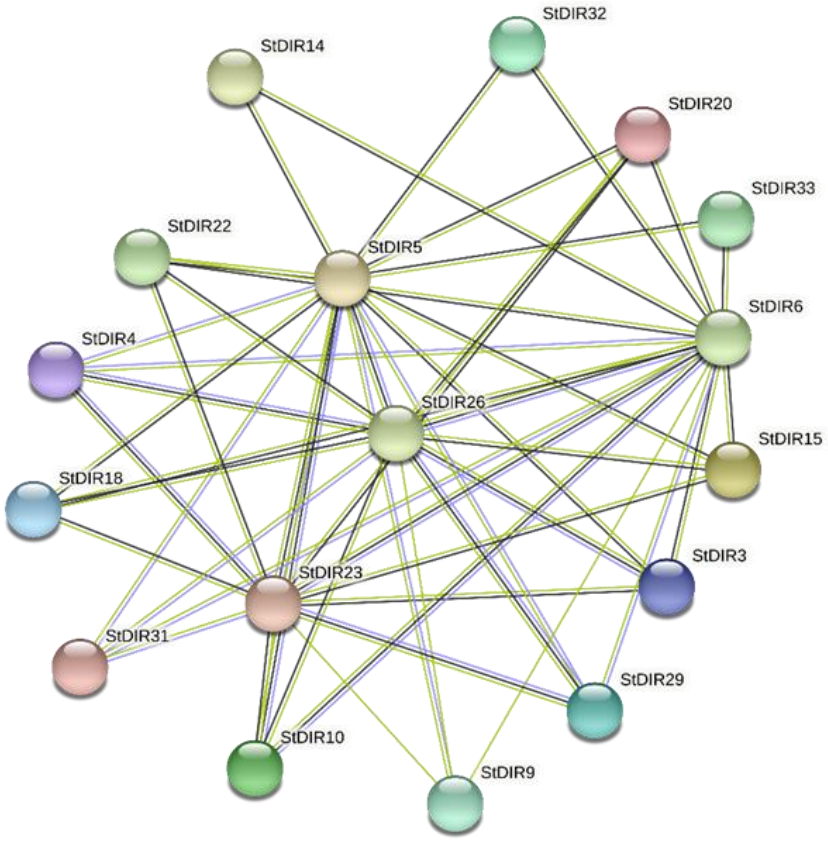

Figure 7. String database projection of StDIR genes.

Figure 8. String database predicted StDIR gene co-expression in related taxa. Compared to bright colors, black has the greatest expression level in the Solanum tuberosum genome.

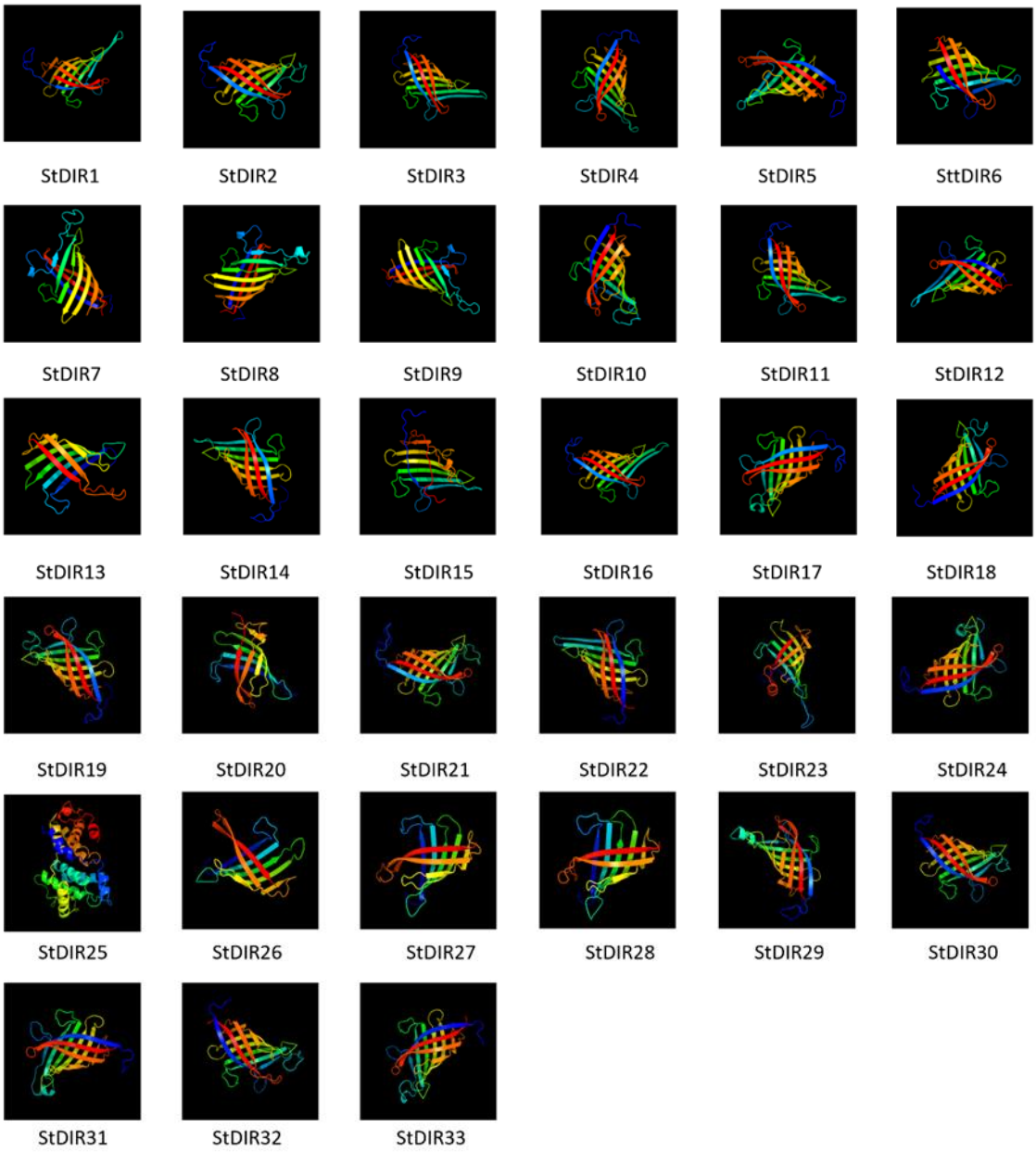

Figure 9. Tertiary Prediction of all StDIRs proteins.

Transmembrane potential and Sequence Identity and Similarity Analysis of StDIRs genes

In order to determine the universality of the active transmembrane potential in the proteins under inquiry, all StDIR genes were subjected to analysis using TMHMM, a freely available web-based program. As indicated in the analysis, 13 of the 33 StDIR were found to be engaged in the cell membrane's integral tasks. The SIAS analysis was conducted to govern the uniqueness, homology, and global resemblance of all the sequences. The table displays standard findings in tabular format. We may conclude from the findings that StDIR indicated a higher value for their similarity and identity analysis than Global similarity. A supplemental file with particular results for each gene is provided.

Table 1. Tabular representation of SIAS tool Results.

	Minimum Value	Maximum Value	Mean Value	Standard Deviation
Identity Analysis	1.78	97.42	8.01	6.94
Similarity Analysis	6.32	100	17.02	16.22
Global Similarity (Blosum62)	-0.5	0.97	-0.15	0.1139

Chromosomal Location and Synteny Analysis of StDIR genes

The localization of the *Solanum tuberosum* DIR genes from across the genome was studied using PhenoGram. Using annotated information from the *Solanum tuberosum* genome obtained from Phytozome, the DIR genes of *Solanum tuberosum* were assigned to their perspective chromosomes, over 11 of the 12 chromosomes, according to the Figure 10. Five StDIR genes are found on chromosome I, three on chromosome two, two on chromosome four, and one on chromosomes five, nine, eleven, and twelve. While Chromosome ten, which possesses 11, has the greatest number of genes, it is Chromosome 11 that has the least. While chromosomes 6 and 8 have three StDIR genes each, chromosome 7 only contains two. While chromosome 3 has no DIR gene. This study suggests an unbalance in the DIR gene distribution in *Solanum tuberosum* chromosomes, which could be connected to chromosomal size and variations.

For synteny analysis, a total of 236 ribbons were created via local alignments by coloring ribbons on bit score.

The maximum bit score was observed at 342 and a minimum of 130 in the study. The color pattern was chosen 7-colour rainbow, i.e. (in order) red, orange, yellow, green, blue, indigo, violet as shown in Figure 11. These findings revealed relationships among DIR genes in potato and model plant *Arabidopsis thaliana*. StDIR 25 from potato and AT5G49030 from *Arabidopsis thaliana* show the relationship among each other, while other show a diverse range of relationship with each other from high to low. Arabidopsis thaliana AT5G49040 DIR transcription factor establish collinearity wit *Solanum tuberosum* DIR genes like StDIR10, StDIR11, StDIR16, StDIR17, StDIR18, StDIR19, StDIR21, StDIR22, StDIR29, StDIR30, StDIR32, and StDIR33. In the same way Arabidopsis thaliana DIR factors AT5G42655, At5G42655, AT5G42510, AT5G42500 resembles with StDIR2, StDIR4, StDIR10, StDIR16, StDIR17, StDIR21, StDIR22, StDIR29, StDIR30, and StDIR32. The Arabidopsis thaliana DIR factors like AT4G38700, AT4G23690, AT4G11210, AT4G11190, AT4G11180 resembles with StDIR3, StDIR4, StDIR7, StDIR8, StDIR9, and StDIR11. In the same way Arabidopsis thaliana DIR factors AT3G58090, AT3G55230, AT3G24020, AT3G13662, resembles with StDIR5, StDIR6, StDIR9, StDIR8, StDIR10, StDIR14, StDIR11, StDIR20, StDIR21, StDIR22, StDIR23, StDIR26, StDIR27, StDIR28, StDIR29, StDIR30, StDIR32, and StDIR33. This group shows the highest level of resemblance. In the same way Arabidopsis thaliana DIR factors AT3G13660, AT3G13650, AT2G39430, AT2G28670, resembles with StDIR4, StDIR5, StDIR9, StDIR10, StDIR13 StDIR14, StDIR19, StDIR21, StDIR22, StDIR23, StDIR26, StDIR27, and StDIR28. Hence in such a case, all the sequences form a relationship with each other.

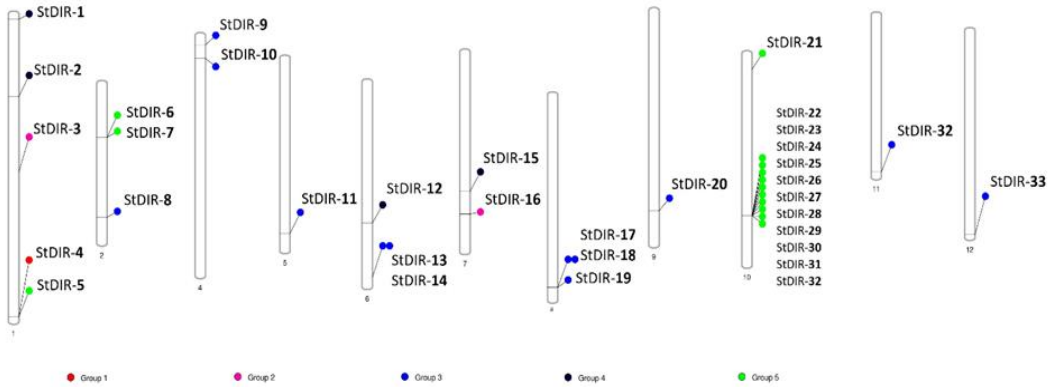

Figure 10. The distribution of DIR genes within the genome of *Solanum tuberosum*.

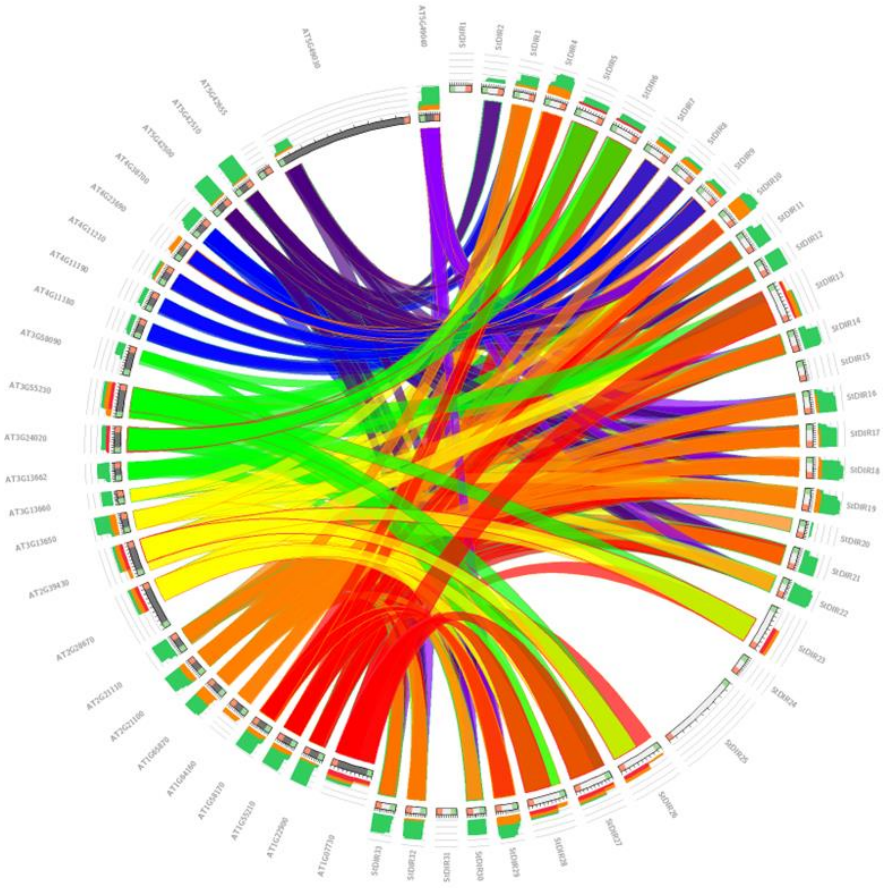

Figure 11. A Synteny map among all identified sequences of StDIR and AtDIR genes.

DISCUSSION

Researching the phenomena of transcription is now a debated issue in biological science. They influence a range of activities such as metabolic processes, physiology, embryonic functions, and so on through controlling the expression channels of genes associated. In recent years, many transcriptional components that regulate dehydration, excessive salt, and other climatic impacts have been identified. Cold temperature, hormone levels, and pathogenic reactions were studied using plants. A multitude of factors can improve a plant's resilience and adaptability to diverse stresses.

DIR genes have a crucial role in the metabolism of lignan and lignin, as well as plant defense mechanisms. In this work, the genome of *Solanum tuberosum* revealed 33 DIR genes. StDIR proteins have a basic domain structure, with just one Dirrigent conserved domain found in all of them. Similar to the previously identified Dirrigent genes of Arabidopsis thaliana, Populus, and Piper nigrum, the DIR gene structures of *Solanum tuberosum* are simple, mostly without introns (Khan *et al*., 2018). Still, one-third of the rice genome has 1–5 introns. (Liao *et al*., 2017). This implies that following differentiation, *Oryza sativa*, dicotyledonous *Solanum tuberosum*, poplar, and the model plant *Arabidopsis thaliana* may well have distinct developmental paths.

Our findings, on the evaluation of cis-elements, reveal that cis-acting elements that were discovered in the upstream area, indicating stress, defense, and stress, sunlight sensitivity, cell cycle development, protein association sites, promotor enhancer parts, gibberellin sensitivity, central promoter element, anaerobic induction, auxin sensitive, cold-responsive, ABA sensitive, dehydration, MeJA sensitive, zein Metabolism development, salicylic acid sensitive, AT-rich DNA Binding Protein (ATBP-1), Wound sensitive, Meristem regulation, Sequence conserved in Alpha-Amylase Promotors, MYBHv1 Binding site, Root Specific Element,

Endosperm regulation, Anoxic Specific Inducibility, Flavonoid Biosynthetic development, and light may play a regulatory role in StDIRs. The results from the study indicate that the mechanisms of StDIRs hormone responses are quite intricate. It is believed that unique DIR genes serve varying purposes in a heterogeneous environment during specific time periods.

The physical characteristics (amino acid, pi, MW, etc.) of the StDIR family genes in *Solanum tuberosum*, and also gene structure (introns, exons), preserved motifs assessment, phylogeny, evolved analysis, mapping of genes, and promoter enrich cis-acting elements, were studied in this study. The StDIR proteins family was identified to have a moderate amount of acidic and basic amino acids, and members with the same subclass were found to be fairly similar, implying that they may play key positions.

Furthermore, proteins with the same evolutionary connections had structural similarities. The StDIR gene family members were uneven among subclasses in both *Solanum tuberosum* and the model plant, according to analyses of the phylogenetic trees of the *Solanum tuberosum* and Arabidopsis trees. Given that the subfamilies of the StDIR gene family had a single parent and may have shared roles, this implies that the StDIR gene family originated before the two species diverged. Protein-protein interaction analysis and SIAS both showed a clear correlation between them.

The genomes of tree species can be examined through the utilization of data mining and phylogenetic methodologies. The bulk of the *Solanum tuberosum* tree StDIR genes were not misplaced owing to environmental selection and instead exhibited great preservation throughout evolution, however, they still need to be investigated more.

CONFLICT OF INTEREST

The authors declare that they have no conflicts of interest.

REFERENCES

Bailey, R.L., K.P.JR. West and R.E. Black. 2015. The epidemiology of global micronutrient deficiencies. Annals of Nutrition and Metabolism. 66(Suppl 2): 22–33.

Burlingame, B., B. Mouillé and U.R. Charrondiére. 2009. Review: nutrients, bioactive non-nutrients, and anti-nutrients in potatoes. Journal of Food Composition and Analysis, 22: 494–502.

Chen, C., H. Chen, Y. Zhang, H.R. Thomas, M.H. Frank, Y. He and R. Xia. 2020. TBtools: an integrative toolkit developed for interactive analyses of big biological data. Molecular plant. 3;13(8): 1194-202.

Davin, L.B., H.B. Wang, A.L. Crowell, D.L. Bedgar, D.M. Martin, S. Sarkanen and N.G. Lewis. 1997. Stereoselective bimolecular phenoxy radical coupling by anauxiliary (120irigent) protein without an active center. Science, 275:362–366.

Frodin, D.G. 2004. History and concepts of big plant genera. Taxon, 53: 753–776.

Gao, C.Q., G.F. Liu, Y.C. Wang, J. Jiang and C.P. Yang. 2010. Cloning and analysis of dirigent-like protein in gene from *Tamarix androssowii*. Bulletin of Botanical Research, 30(1): 81–86.

Guo, A.Y., Q.H. Zhu, X. Chen and J.C. Luo. GSDS: a gene structure display server. Yi chuan= Hereditas, 2007, Aug 1;29(8): 1023-6.

Guo, J.L., L.P. Xu, Y.C. Su, H.Y. Fu, Y.X. Que and J.S. Xu. 2012. A novel dirigent protein gene with highly stemspecific expression from sugarcane, response to drought, salt and oxidative stresses. Plant Cell Reports, 31(10): 1801–181.

Hadwiger, L.A., C.C. Chiang and D. Horovitz. 1992. Expression of disease resistance response genes in nearisogenic pea cultivars following challenge by Fusarium oxysporum race 1. Physiological and Molecular Plant Pathology, 40: 259-269.

Horton, P, Park KJ, Obayashi T, Fujita N, Harada H, Adams-Collier CJ, Nakai K. WoLF PSORT: protein localization predictor. Nucleic acids research. 2007, Jul 1;35(suppl_2): W585-7.

Hosmani, P.S., T. Kamiya, J. Danku, S. Naseer, N. Geldner and M.L. Guerinot. 2013. Dirigent domain-containing protein is part of the machinery required for formation of the lignin-based Casparian strip in the root. Proceedings of the National Academy of Sciences of the United States of America, 110: 14498–14503.

Hu, B., J. Jin, A.Y. Guo, H. Zhang, J. Luo and G. Gao. GSDS 2.0: an upgraded gene features visualization server. Bioinformatics. 2015, Apr 15;31(8): 1296-7.

Käll, L., A. Krogh and E.L. Sonnhammer. 2007. Advantages of combined transmembrane topology and signal peptide prediction—the Phobius web server. Nucleic Acids Research, 1;35(suppl_2): W429-32.

Kelley, L.A. and M.J.E. Sternberg. 2009. Protein structure prediction on the Web: a case study using the Phyre

server. Nature Protocols, 4:3,4(3), 363–371. https://doi.org/10.1038/nprot.2009.2.

Khan, A., R.J. Li, J.T. Sun, F. Ma, H.X. Zhang, J.H. Jin, M. Ali, S. Huq, J.E. Wang and Z.H. Gong. 2018. Genome-wide analysis of dirigent gene family in pepper (*Capsicum annuum* L.) and characterization of CaDIR7 in biotic and abiotic stresses. Scientific Reports, 8:5500.

Knapp, S. 2002. Tobacco to tomatoes: a phylogenetic perspective on fruit diversity in the Solanaceae. Journal of Experimental Botany, 53(377): 2001–2022. doi:10.1093/jxb/erf068. PMID:12324525.

Kumar, S., G. Stecher, M. Li, C. Knyaz and K. Tamura. 2018. MEGA X: molecular evolutionary genetics analysis across computing platforms. Molecular Biology and Evolution, 35(6):1547.

Lamesch, P., T.Z. Berardini, D. Li, D. Swarbreck, C. Wilks, R. Sasidharan, R. Muller, K. Dreher, D.L. Alexander, M. Gracia-Hernandez and A.S. Karthikeyan. 2012. The Arabidopsis information resource (tair): improved gene annotation and new tools. Nucleic Acids Research, 40: 1202–1210.

Li, N., M. Zhao, T. Liu, L. Dong, Q. Cheng, J. Wu, L. Wang, X. Chen, C. Zhang, W. Lu and P. Xu. 2017. A novel soybean dirigent gene GmDIR22 contributes to promotion of lignan biosynthesis and enhances resistance to *Phytophthora sojae*. Frontiers in Plant Science, 8: 1185.

Li, Q., J. Chen, Y. Xiao, P. Di, L. Zhang, and W. Chen. 2014. The dirigent multigene family in *Isatis indigotica*: gene discovery and differential transcript abundance. BMC Genome, 15: 388.

Li, Q., J. Chen, Y. Xiao, P. Di, L. Zhang and W. Chen. 2014. The dirigent multigene family in Isatis indigotica: gene discovery and differential transcript abundance. BMC Genomics, 15:1, 15(1), 1–13. https://doi.org/10.1186/1471-2164-15-388.

Liao, Y., S. Liu, Y. Jiang, C. Hu, X. Zhang, X. Cao, Z. Xu, X. Goa, L. Li, J. Zhu and R. Chen. 2017. Genome-wide analysis and environmental response profiling of dirigent family genes in rice (*Oryza sativa*). Genes Genom, 39(1): 47–62.

Ma, Q.H. and Y.C. Liu. 2015. Tadir13, a dirigent protein from wheat, promotes lignan biosynthesis and enhances pathogen resistance. Plant Molecular Biology Reporter, 33(1): 143–152.

Lescot, M., P. Dhais, G. Thijs, K. Marchal, Y. Moreau, Y. Van de Peer, P. Rouz and S. Rombauts. 2002. Nucleic Acids Research, 1;30(1): 325-327.

Ralph, S.G., J.Y. Park, J. Bohlmann and S.D. Mansfield. 2006. Dirigent proteins in conifer defense: gene discovery, phylogeny, and differential wound- and insect-induced expression of a family of DIR and DIRlike genes in spruce (*Picea* spp.). Plant Molecular Biology, 60: 21–40.

Shi, H., Z. Liu, L. Zhu, C. Zhang, Y. Chen, Y. Zhou, F. Li and X. Li. 2012. Overexpression of cotton (*Gossypium hirsutum*) dirigent1 gene enhances lignification that blocks the spread of *Verticillium dahliae*. Acta Biochimica et Biophysica Sinica, 44: 555–564.

Song, M., W. Xu, Y. Xiang, H. Jia, L. Zhang and Z. Ma. 2013. Association of jacalin-related lectins with wheat responses to stresses revealed by transcriptional profiling. Plant Molecular Biology, 84: 95–110.

Subramanyam, S., N. Sardesai, D.P. Puthoff, J.M. Meyer, J.A. Nemacheck, M. Gonzalo and C.E. Williams. 2006. Expression of two wheat defense-response genes, Hfr-1 and Wci-1, under biotic and abiotic stresses. Plant Science, 170:90–103.

Subramanyam, S., D.F. Smith, J.C. Clemens, M.A. Webb, N. Sardesai and C.E. Williams. 2008. Functional characterization of HFR1, a high-mannose N-glycan-specific wheat lectin induced by Hessian fly larvae. Plant Physiology, 147(3): 1412–1426.

Szklarczyk, D., J.H. Morris, H. Cook, M. Kuhn, S. Wyder, M. Simonovic, A. Santos, N.T. Doncheva, A. Roth, P. Bork and L.J. Jensen. 2017. The STRING database in 2017: quality-controlled protein–protein association networks, made broadly accessible. Nucleic Acids Research, 18: gkw937.

Szklarczyk, G.E., A. Gattiker, C. Hoogland, I. Ivanyi, R.D. Appel and A. Bairoch. 2003. ExPASy: the proteomics server for in-depth protein knowledge and analysis. Nucleic acids research, 1;31(13): 3784-8.

Thamil, A.S.K., J.I. Park, N.U. Ahmed, H.J. Jung, Y. Hur, K.K. Kang, Y.P. Lim and I.S. Nou. 2013. Characterization and expression analysis of dirigent family genes related to stresses in brassica. Plant physiology and biochemistry, 67(3): 144–153.

Thompson, J.D., T.J. Gibson and D.G. Higgins. 2003. Multiple sequence alignment using ClustalW and ClustalX. Current Protocols in Bioinformatics. (1): 2-3.

Wolfe, D., S. Dudek, M.D. Ritchie and S.A. Pendergrass. 2013. Visualizing genomic information across chromosomes with PhenoGram. BioData Mining, 6(1):1-2.

Wu, R., L. Wang, Z. Wang, H. Shang, X. Liu, Y. Zhu, D. Qi, and X. Deng. 2009. Cloning and expression analysis of a dirigent protein gene from the resurrection plant *Boea hygrometrica*. Progress in Natural Science, 19: 347–352.

Wu, R., L. Wang, Z. Wang, H. Shang, X. Liu, Y. Zhu, D. Qi and X. Deng. 2009. Cloning and expression analysis of a dirigent protein gene from the resurrection plant *Boea hygrometrica*. Progress in Natural Science, 19: 347–352.