doi 10.18699/bgrs2024-06-53

The transcription factor WRKY multigene family in grapevine (*Vitis vinifera*): genome-wide identification and evolution

Vodiasova E.[1]*, Sinchenko A.[1], Khvatkov P.[1], Dolgov S.[1, 2]

[1] *The Nikitsky Botanical Gardens – National Scientific Center of the RAS, Nikita, Yalta, Republic of the Crimea, Russia*

[2] *Branch of Shemyakin and Ovchinnikov Institute of Bioorganic Chemistry, Puschino, Russia*

* *eavodiasova@gmail.com*

Key words: WRKY transcription factor; *Vitis vinifera*; genome-wide analyses: phylogeny; grape cultivars

Motivation and Aim: WRKY is one of the largest multigenic families of transcription factors in higher plants. These proteins are key regulators involved in the response to a variety of biotic and abiotic factors and play a pivotal role in plant growth and development [1]. Many theories have been proposed about the evolution of WRKY genes from algae to multicellular plants, but there is no consensus [2]. *Vitis vinifera* L. (grape) is one of the most widely cultivated plants for agricultural purposes. To date, several genomic analyses of WRKY genes in grape have been performed [3], but the number of genes encoding WRKY varied and seemed to depend on the version of the genomic assembly analysed. This study focuses on the identification of WRKYs by genomic analysis from different grapevine assemblies and attempts to shed light on their evolution.

Methods and Algorithms: The genome-wide analyses were performed using six grape genome assemblies. The identification of WRKY proteins was based on the detection of the WRKY domain signature using InterProScan v. 5.63-95.0. The exon-intron structure of each WRKY transcription factor-encoding gene was described according to the annotation of *V. vinifera* genomes from the NCBI and GrapeGenomics databases. A motif analysis was performed using the MEME online tool in classic mode. Phylogenetic analysis was performed using IQ-TREE v. 2.3.0. The data set included all identified *V. vinifera* WRKY protein sequences from six genome assemblies. Multiple alignment of 965 protein sequences was performed using the MAFFT v. 7.0 online service.

Results: The domain analysis identified WRKY DNA-binding domain proteins in each grape assembly. The number of proteins identified differed between the grape cultivars. A total of 115 proteins (including isoforms) were found for Cabernet Franc, 106 for Cabernet Sauvignon, 129 for Pinot Noir clone FPS123, 89 for the 12X assembly (GCA_000003745.2), 87 for the RefSeq reference assembly (GCA_030704535), 65 for the GenBank reference assembly (GCA_030704535) and 84 for the assembly from the GrapeGenomics database. The phylogenetic analysis identified 62 gene clusters with 100 bootstrap supports characterising different WRKY classes. Each WRKY class was numbered in accordance with the position of the encoding gene in the majority of the assemblies analysed. It was observed that some WRKYs were not present in all grape cultivars. The number of exons ranged from 2 to 8 and the length of VvWRKY from 151 to 746 aa. The mean genetic distances between varieties ranged from 0 to 0.113. Variability of the conserved DNA-binding heptapeptide WRKYGQK was detected. Some additional domains were identified in several grape VvWRKYs: a Zn cluster

domain (IPR018872); a Frigida-like domain (IPRO12474); the COILS structural motif and the signal peptide in the N-terminal and non-cytoplasmic domains. Four novel chimeric WRKY TFs were thus revealed. The LxLxLx repressor and LxxLL co-activator motifs were also identified in eight and twelve VvWRKY proteins, respectively. An association between the clustering of WRKY groups with the functional family of the protein, WRKY signature sequences and the presence of other domains in the protein with the formation of chimeric WRKY TFs was revealed. Based on the phylogenetic analysis performed and the probability of evolutionary events, we suggest that there were two dynamic phases of complexity and simplification in the evolution of WRKY. The two clades C and D evolved from a common ancestor as a result of the complexity of the gene structure, then divergence occurred and a characteristic motif (NTWD or Coil) emerged in each clade. During the simplification phase, loss of one of the motifs resulted in a decrease in exon number and protein length, consistent with a bimodal distribution of exon length and number. The proposed evolutionary model is consistent with the fact that macroevolutionary patterns are characterised by a periodicity of dynamics in opposite directions. Quantitatively, which is also consistent with the proposed theory, genome evolution is dominated by reduction and simplification, followed by episodes of increasing complexity.

Conclusion: In *V. vinifera*, sixty-two *WRKY* transcription factor genes have been identified. The structure of each of the *VvWRKY* genes has been studied and its chromosomal location has been determined. The genes were re-numbered by chromosome location in the reference genome of *V. vinifera* cv. PN40024, v.5. was proposed. Inter-varietal amino acid variability was revealed, reaching 5 % for some genes. In addition, chimeric *VvWRKY* genes have been found, which may have specific regulatory functions. Phylogenetic analysis shows that the evolution of *WRKY* genes went through a phase of complication (gaining introns and forming genes with two domains) and a phase of simplification (losing introns and reducing protein length). Consistent with the wide range of processes in which WRKY transcription factors are involved, the data obtained indicate a high functional flexibility of this family.

Funding: The study is supported by the Russian Science Foundation Grant (No. 23-76-10013).

References

1. Wani S.H., Anand S., Singh B., Bohra A., Joshi R. WRKY Transcription Factors and Plant Defense Responses: Latest Discoveries and Future Prospects. *Plant Cell Rep.* 2021;40(7):1071-1085. doi 10.1007/s00299-021-02691-8
2. Goyal P., Devi R., Verma B., Hussain S., Arora P., Tabassum R., Gupta S. WRKY Transcription Factors: Evolution, Regulation, and Functional Diversity in Plants. *Protoplasma*. 2023;260(2):331-348. doi 10.1007/s00709-022-01794-7
3. Wu W., Fu P., Lu J. Grapevine WRKY Transcription Factors. *Fruit Res.* 2022;2(1):1-8. doi 10.48130/frures-2022-0010