Uluslararası Fen Araştırmalarında Yenilikçi Yaklaşımlar Dergisi
International Journal of Innovative Approaches in Science Research 2019, Vol. 3 (1), 1-8
https://doi.org/10.29329/ijiasr.2019.187.1

Original article

In Silico Analysis of *SCARECROW* Genes in Olive (*Olea europaea* L.)

Fatih Sezer *

Department of Molecular Biology and Genetics, Faculty of Science and Arts, Çanakkale Onsekiz Mart University, 17100 Çanakkale, Turkey

Abstract

The Olive (*Olea europaea* L.) is a Mediterranean crop tree with great economic importance and understanding molecular mechanisms that control developmental processes like adventitious rooting will benefit various aspects of olive tree development research. SCARECROW (*SCR)* genes are reported to be essential for asymmetric division of the cortex/endodermis progenitor cell in the root. They also have diverse roles through plant development. Thus, this study aimed to identify *SCR* genes in Olive. With this study, we investigated all GRAS family proteins in Olive and further extracted SCR proteins in Olive genome. We report *SCR* genes in olive for the first time and provide analysis of their structure with bioinformatic tools. Phylogenetic tree revealed that all 8 major sub-families of GRAS were present in the olive genome and some other sub-families like SCL18, SCL3 and SCL28 were also present. *A. thaliana SCR* (*AT3G54220*) was grouped with 2 unique olive sequences (XP_022893717.1, XP_022881217.1). Both sequences showed a 2 exon, 1 intron structure and contained the characteristic GRAS domain. Promoter regions contained several light and low temperature motifs. Auxin, gibberellin and abscisic acid response elements were present within these regions. Promoter regions of all olive SCR also contained MYB transcription factor recognition and binding sites.

Keywords: *SCARECROW, SCR, Olive, GRAS, Transcription factor.*

Received: 23 February 2019 * **Accepted:** 19 March 2019 * **DOI:** https://doi.org/10.29329/ijiasr.2019.187.1

* **Corresponding author:**

Fatih Sezer, Department of Molecular Biology and Genetics, Faculty of Science and Arts, Çanakkale Onsekiz Mart University, Çanakkale, Turkey.
Email: fsezerfatih@gmail.com

INTRODUCTION

GRAS gene family includes important transcription factors in plant development (Pysh, Wysocka-Diller, Camilleri, Bouchez, & Benfey, 1999). The family name "GRAS" comes from the first discovered members of this family, GAI, RGA and SCR (Di Laurenzio et al., 1996; Peng et al., 1997; Silverstone, 1998). Members of this family are involved in a wide area of plant development and stress response. GRAS family members were identified in several plant species including, Arabidopsis (Tian, Wan, Sun, Li, & Chen, 2004), Populus (Liu & Widmer, 2014), Chinese cabbage (Song et al., 2014), maize (Lim et al., 2005), rice (Liu & Widmer, 2014), grapevine (Grimplet, Agudelo-Romero, Teixeira, Martinez-Zapater, & Fortes, 2016), tomato (Huang, Xian, Kang, Tang, & Li, 2015). The identified members of this family are usually divided into sub-family groups like SCR, DELLA, SCL3, LISCL, SHR, LS, HAM and PAT1. These sub-families all have distinct roles in plant physiology (Tian et al., 2004). DELLA group of transcription factors include genes that repress the gibberellin signal transduction pathway (Silverstone, 1998). *SCL3* genes act as a regulator in the molecular control of DELLA proteins (Zhang et al., 2011). The PAT1, SCL5, SCL13 group proteins are involved in light signalling (Bolle, Koncz, & Chua, 2000; P. Torres-Galea, Hirtreiter, & Bolle, 2013; Patricia Torres-Galea, Huang, Chua, & Bolle, 2006). Although different groups possess different functions, all GRAS family proteins contain conserved GRAS domains in their c-terminal regions.

Another sub family of GRAS transcription factors, the *SCR* genes are essential for asymmetric division of the cortex/endodermis progenitor cell in the root. Expression of *SCR* was found in root tips, leaf primordia and while ligule and stomatal formation (Kamiya, Itoh, Morikami, Nagato, & Matsuoka, 2003). Expression profile suggests diverse roles for SCR proteins on plant development.

The Olive (*Olea europaea* L.) is a Mediterranean crop tree which belongs to the Oleaceae family, and it has great economic importance, due to its health-promoting oil content in the Mediterranean diet. Recent advances in genomic and transcriptomic data allowed us to identify SCR family proteins. Understanding the molecular control of SCR genes will benefit various aspects of olive tree development research. Thus, this study identifies SCR gene family members in olive and discusses their structures, phylogenetic relationships.

MATERIALS and METHODS

In-silico Identification of SCR Sequences

We obtained *A. thaliana SCR* (*AT3G54220*) sequences from TAIR and used these sequences as a reference to obtain SCR sequences of O. europaea. Blast search was performed against *O. europaea* cv. Sylvestris proteome and retrieved all sequences with e values lower than 1. All sequences were submitted to NCBI CCD search (Marchler-Bauer et al., 2015) and sequences without a GRAS domain was eliminated from the study. We obtained the SCARECROW-LIKE proteins of *A. thaliana* and

Solanum lycopersicum with the same method and used them as a reference in the phylogenetic tree. A phylogenetic tree consisting of all GRAS family proteins were computed using the Poisson correction and UPGMA method. This analysis involved 214 amino acid sequences. All ambiguous positions were removed for each sequence pair (pairwise deletion option). There were a total of 1265 positions in the final dataset. Evolutionary analyses were conducted in MEGA X (Kumar, Stecher, Li, Knyaz, & Tamura, 2018). Olive sequences that were grouped with *A. thaliana* and *S. lycopersicum* were extracted and further analyzed with bioinformatic methods.

Bioinformatic Analysis of SCR Genes in Olive

The Muscle alignment tool was used for protein and nucleotide alignments (Edgar 2004). We used Geneious R8 (Kearse et al., 2012) software to predict amino acid sequences. The molecular weight and isoelectric point of proteins were calculated with Expasy's ProtParam server (http://web.expasy.org/protparam/) (Gasteiger et al., 2005). Conserved regions were investigated with CDD: conserved domain database (Marchler-Bauer et al., 2015). Promoter regions were analyzed for regulatory motifs in PLANTCARE database (Lescot et al., 2002).

RESULTS and DISCUSSION

With this study, we investigated all GRAS family proteins in Olive and further extracted SCR proteins in Olive genome. The blast search using *A. thaliana SCR* (*AT3G54220*) protein sequence against olive proteome returned 117 protein sequences. The domain search revealed that all sequences contained the GRAS domain and 4 proteins contained the DELLA domain in addition to GRAS which is specific to the DELLA sub-family of the GRAS transcription factor family. The phylogenetic tree revealed and separated all sub families of GRAS transcription factor family (Fig 1).

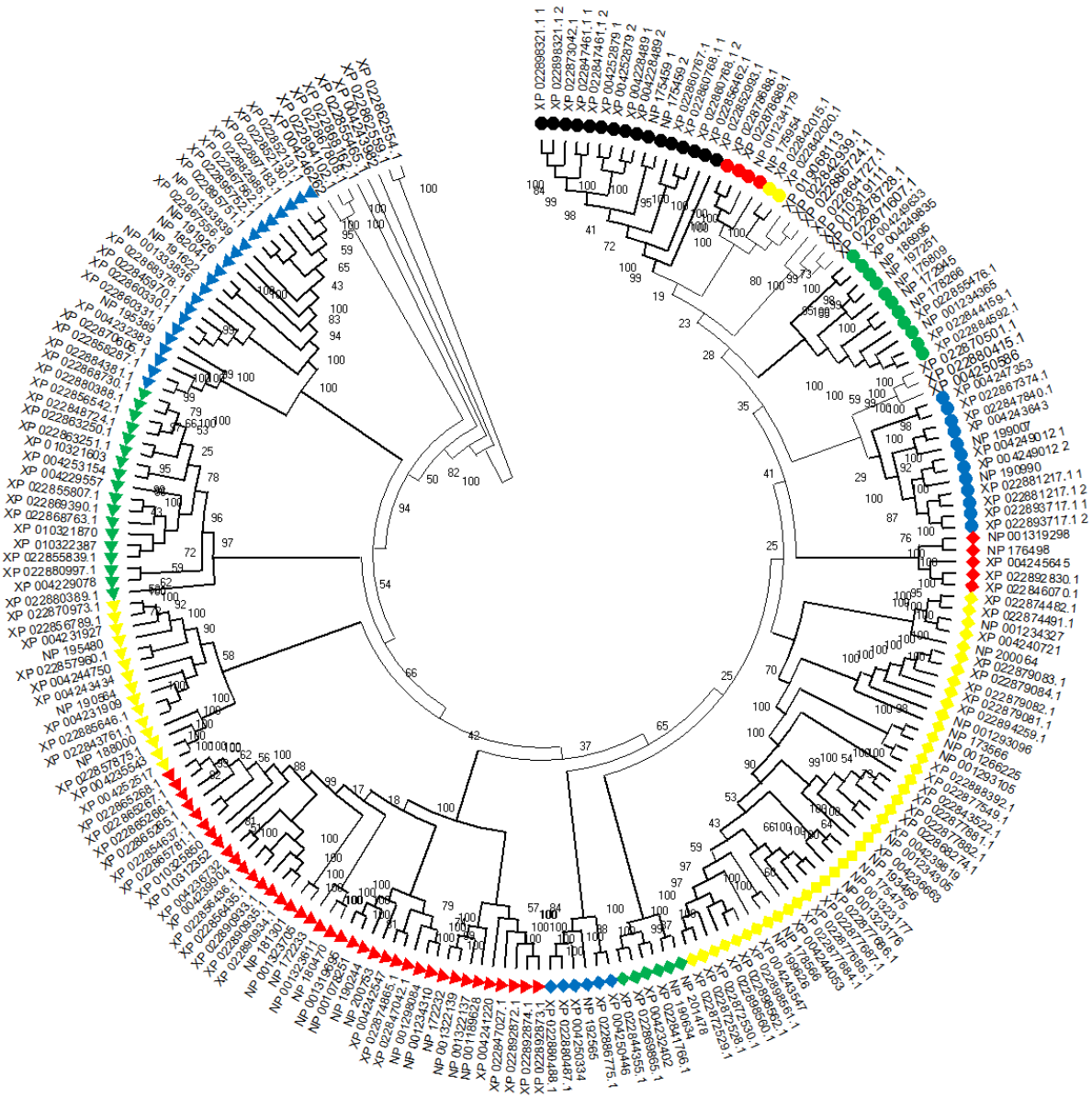

Figure 1. Phylogenetic tree of GRAS family proteins in Olive. The markers represent the GRAS sub-families: ● SCL3, ● SCL18, ● SCL23, ● DELLA-RGA, ● SCR, ♦ SCL28, ♦ PAT1, ♦ LS, ♦ SCL26, ▲ LiSC, ▲ SHR, ▲ DELLA-RGL, ▲ HAM.

The clade containing *A. thaliana SCR* (*AT3G54220*) also contained 2 unique Olive proteins (*XP_022893717.1*, *XP_022881217.1*). This clade included all SCR proteins of *A. thaliana*, *S. lycopersicum* and *O. europaea*. Thus, we named these olive sequences as *OeSCRa* (*XP_022893717.1*) and *OeSCRb* (*XP_022881217.1*). *OeSCRa* is a 806 a.a. long protein with a molecular weight of 87.2 kDa and 5.67 Theoretical pI. *OeSCRa* is 811 a.a. long with a molecular weight of 87.5 kDa and 5.83 Theoretical pI. Both proteins contain a GRAS domain which is characteristic to this gene family (Table 1).

Table 1. Protein characteristics of Olive SCR proteins

Accession	Protein Length	Molecular weight (kDa)	Theoretical pI	Conserved Domain
OeSCRa	806	87.2	5.67	GRAS
OeSCRb	811	87.5	5.83	GRAS

We performed a tBlastn search using *OeSCRa* and *OeSCRb* against *O. europaea* cv. Sylvestris transcriptome and genome. Only one transcript and genomic region showed a %100 similarity against each SCR proteins. *OeSCRa* sequences were matched against a region in the 2. chromosome and *OeSCRb* has matched against 1. chromosome. Genomic structures of Olive SCR contained 2 exons and 1 intron. The Genomic region of *OeSCRa* was 3493 bp and the mRNA was 2421 bp long. The Genomic region of *OeSCRb* was 4114 bp and the mRNA was 2436 bp long.

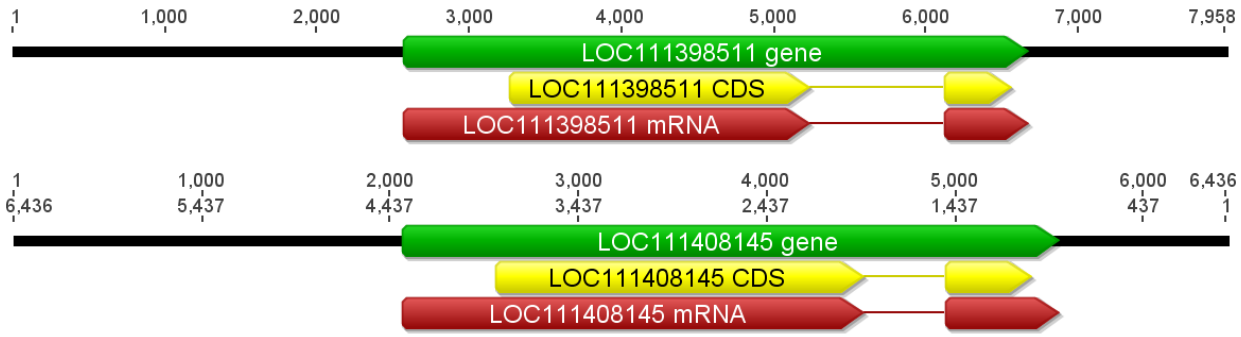

Figure 2. Genomic structure of SCR genes in Olive.

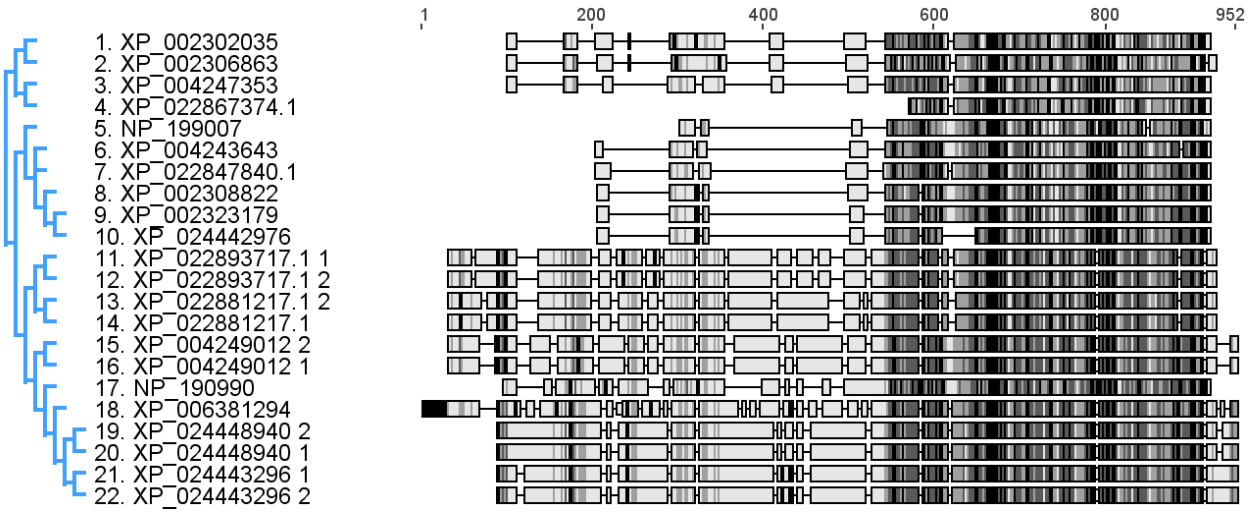

Figure 3. Alignment of SCR LIKE Proteins of Olive

Alignment of the protein sequences that grouped closely with *A. thaliana SCR* (*AT3G54220*) which includes *SCR* and SCR-LIKE 23 groups showed a more conserved region in carboxy terminal

region. This region also contains the GRAS domain. The amino terminal region, however, showed higher polymorphic regions (Fig 3).

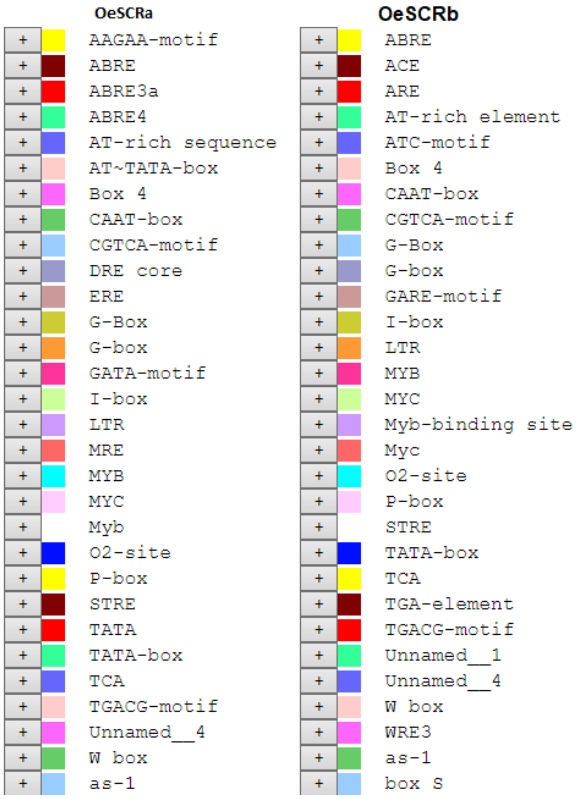

Figure 4. Regulatory motifs found in promoter regions of olive SCR proteins

We also investigated regulatory motifs found in promoter regions of olive *SCR* genes. We analyzed 1000 bp upstream of the first codon (ATG) for all three genes. We found several regulatory elements and core promoter regions like TATA and CAAT-box within all promoters. Among the regulatory motifs, MYB transcription factor recognition and binding sites were common for all SCR genes. MYB transcription factors include several genes that control plant development, differentiation, stress and defence (Ambawat, Sharma, Yadav, & Yadav, 2013). The MYB motifs that are present in olive SCR promoters suggests a role in their regulation with an MYB transcription factor. Promoter regions also contained several light and low temperature motifs. Auxin, gibberellin and abscisic acid response elements were also present within these regions (Fig 4).

Conclusion

Recent advances in olive genomic and transcriptomic research enabled the discovery of several gene and gene families related to development and fruit quality. With this study, we report *SCR* genes in olive for the first time and provide analysis of their structure with bioinformatic tools. *SCR* genes are a part of a large family of transcription factors called GRAS (Pysh et al., 1999). While this study is focused on the *A. thaliana SCR* (*AT3G54220*) genes of olive, the phylogenetic tree includes all GRAS

family transcription factors in olive. Phylogenetic tree revealed that all 8 major sub-families of GRAS were present in the olive genome and some other sub-families like SCL18, SCL3 and SCL28 were also present. *A. thaliana SCR* (*AT3G54220*) was grouped with 2 unique olive sequences (XP_022893717.1, XP_022881217.1). Both sequences showed a 2-exon structure and contained the characteristic GRAS domain. Olive *SCR* genes were more conserved in their carboxy terminal regions. Promoter regions of all olive SCR contained MYB transcription factor recognition and binding sites. MYB transcription factors include several genes that control plant development, differentiation, stress and defence (Ambawat et al., 2013). The MYB motifs that are present in olive SCR promoters suggests a role in their regulation with an MYB transcription factor. Further functional studies will enlighten molecular control of GRAS family and will benefit various aspects of olive tree development research.

REFERENCES

Ambawat, S., Sharma, P., Yadav, N. R., & Yadav, R. C. (2013). MYB transcription factor genes as regulators for plant responses: An overview. Physiology and Molecular Biology of Plants, 19(3), 307–321. https://doi.org/10.1007/s12298-013-0179-1

Bolle, C., Koncz, C., & Chua, N. H. (2000). PAT1, a new member of the GRAS family, is involved in phytochrome A signal transduction. Genes and Development, 14(10), 1269–1278. https://doi.org/10.1016/j.enpol.2015.10.044

Di Laurenzio, L., Wysocka-Diller, J., Malamy, J. E., Pysh, L., Helariutta, Y., Freshour, G., … Benfey, P. N. (1996). The SCARECROW gene regulates an asymmetric cell division that is essential for generating the radial organization of the Arabidopsis root. Cell, 86(3), 423–433. https://doi.org/10.1016/S0092-8674(00)80115-4

Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M. R., Appel, R. D., & Bairoch, A. (2005). Protein Identification and Analysis Tools on the ExPASy Server. In J. Walker (Ed.), The Proteomics Protocols Handbook (pp. 571–607). Humana Press. https://doi.org/10.1385/1-59259-890-0:571

Grimplet, J., Agudelo-Romero, P., Teixeira, R. T., Martinez-Zapater, J. M., & Fortes, A. M. (2016). Structural and Functional Analysis of the GRAS Gene Family in Grapevine Indicates a Role of GRAS Proteins in the Control of Development and Stress Responses. Frontiers in Plant Science, 7(March). https://doi.org/10.3389/fpls.2016.00353

Huang, W., Xian, Z., Kang, X., Tang, N., & Li, Z. (2015). Genome-wide identification, phylogeny and expression analysis of GRAS gene family in tomato. BMC Plant Biology, 15(1), 209. https://doi.org/10.1186/s12870-015-0590-6

Kamiya, N., Itoh, J. I., Morikami, A., Nagato, Y., & Matsuoka, M. (2003). The SCARECROW gene's role in asymmetric cell divisions in rice plants. Plant Journal, 36(1), 45–54. https://doi.org/10.1046/j.1365-313X.2003.01856.x

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., … Drummond, A. (2012). Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics, 28(12), 1647–1649. https://doi.org/10.1093/bioinformatics/bts199

Kumar, S., Stecher, G., Li, M., Knyaz, C., & Tamura, K. (2018). MEGA X: Molecular evolutionary genetics analysis across computing platforms. Molecular Biology and Evolution, 35(6), 1547–1549. https://doi.org/10.1093/molbev/msy096

Lescot, M., Déhais, P., Thijs, G., Marchal, K., Moreau, Y., Van de Peer, Y., … Rombauts, S. (2002). PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. Nucleic Acids Research, 30(1), 325–327. https://doi.org/10.1093/nar/30.1.325

Lim, J., Jung, J. W., Chae, E. L., Lee, M. H., Bong, J. K., Kim, M., … Benfey, P. N. (2005). Conservation and diversification of SCARECROW in maize. Plant Molecular Biology, 59(4), 619–630. https://doi.org/10.1007/s11103-005-0578-y

Liu, X., & Widmer, A. (2014). Genome-wide Comparative Analysis of the GRAS Gene Family in Populus, Arabidopsis and Rice. Plant Molecular Biology Reporter, 32(6), 1129–1145. https://doi.org/10.1007/s11105-014-0721-5

Marchler-Bauer, A., Derbyshire, M. K., Gonzales, N. R., Lu, S., Chitsaz, F., Geer, L. Y., … Bryant, S. H. (2015). CDD: NCBI's conserved domain database. Nucleic Acids Research, 43(D1), D222–D226. https://doi.org/10.1093/nar/gku1221

Peng, J., Carol, P., Richards, D. E., King, K. E., Cowling, R. J., Murphy, G. P., & Harberd, N. P. (1997). The Arabidopsis GAI gene defines a signaling pathway that negatively regulates gibberellin responses. Genes and Development, 11(23), 3194–3205. https://doi.org/10.1101/gad.11.23.3194

Pysh, L. D., Wysocka-Diller, J. W., Camilleri, C., Bouchez, D., & Benfey, P. N. (1999). The GRAS gene family in Arabidopsis: Sequence characterization and basic expression analysis of the SCARECROW-LIKE genes. Plant Journal, 18(1), 111–119. https://doi.org/10.1046/j.1365-313X.1999.00431.x

Silverstone, A. L. (1998). The Arabidopsis RGA Gene Encodes a Transcriptional Regulator Repressing the Gibberellin Signal Transduction Pathway. The Plant Cell Online, 10(2), 155–170. https://doi.org/10.1105/tpc.10.2.155

Song, X. M., Liu, T. K., Duan, W. K., Ma, Q. H., Ren, J., Wang, Z., … Hou, X. L. (2014). Genome-wide analysis of the GRAS gene family in chinese cabbage (brassica rapa ssp. pekinensis). Genomics, 103(1), 135–146. https://doi.org/10.1016/j.ygeno.2013.12.004

Tian, C., Wan, P., Sun, S., Li, J., & Chen, M. (2004). Genome-wide analysis of the GRAS gene family in rice and Arabidopsis. Plant Molecular Biology, 54(4), 519–532. https://doi.org/10.1023/B:PLAN.0000038256.89809.57

Torres-Galea, P., Hirtreiter, B., & Bolle, C. (2013). Two GRAS Proteins, SCARECROW-LIKE21 and PHYTOCHROME A SIGNAL TRANSDUCTION1, Function Cooperatively in Phytochrome A Signal Transduction. Plant Physiology, 161(1), 291–304. https://doi.org/10.1104/pp.112.206607

Torres-Galea, P., Huang, L. F., Chua, N. H., & Bolle, C. (2006). The GRAS protein SCL13 is a positive regulator of phytochrome-dependent red light signaling, but can also modulate phytochrome A responses. Molecular Genetics and Genomics, 276(1), 13–30. https://doi.org/10.1007/s00438-006-0123-y

Zhang, Z.-L., Ogawa, M., Fleet, C. M., Zentella, R., Hu, J., Heo, J.-O., … Sun, T. (2011). SCARECROW-LIKE 3 promotes gibberellin signaling by antagonizing master growth repressor DELLA in Arabidopsis. Proceedings of the National Academy of Sciences, 108(5), 2160–2165. https://doi.org/10.1073/pnas.1012232108