

木荷 WRKY 基因家族的全基因组鉴定与分析

陈昌娅, 蔡燕灵, 何波祥, 张谦、梁东成, 汪迎利, 尧俊*

(广东省林业科学研究院, 广东 广州 510520)

摘要:【目的】基于木荷基因组数据, 利用生物信息学手段, 鉴定木荷 *WRKY* 基因家族所有成员, 并对其理化性质、系统进化、顺式作用元件、基因结构、基序分析、保守结构域、染色体定位以及共线性进行分析, 为后续探究 *SsWRKY* 基因功能作用提供参考。【方法】采用生物信息学方法, 基于木荷基因组数据对木荷 *WRKY* 基因家族进行鉴定, 通过 TBtool 软件对该家族进行理化性质、基因结构、保守结构域、染色体定位、共线性分析。利用 MEGA 软件构建系统发育树, PlantCARE 网站对顺式作用元件分析。【结果】在木荷基因组数据中共鉴定出 82 个 *WRKY* 基因家族成员, 并根据相似性命名为 *SsWRKY1*—*SsWRKY75*, 不均匀分布在 18 条染色体上, 编码氨基酸个数在 153—770, 分子量 17701—783888.06 Da, 有 30 个 *SsWRKY* 为碱性蛋白 ($pI > 7$), 酸性蛋白 52 个 ($pI < 7$), 且有 80 个为不稳定蛋白。基于系统进化关系将其划分为 3 大类群(I-III), 第II类群又被分成 5 个亚群(IIa-e), II组成员最多。*SsWRKY* 基因的外显子数量为 2~7 个, 内含子数量为 1~6 个, 同一类家族中的结构域及保守基序相似, 所有成员都含有 WRKYGQK 七肽保守结构域和锌指结构。木荷 *WRKY* 基因家族中片段重复在基因家族扩展中发挥的作用大于串联重复, 且木荷与拟南芥的共线性关系远高于水稻。分析启动子发现, 存在多种植物生长发育、激素、干旱、低温等响应元件。【结论】在木荷中共鉴定出 82 个 *WRKY* 家族基因, 分为 7 个亚族, 不同亚族间的基因结构和蛋白保守结构域有所不同, 片段重复可能是进化驱动力, 以上结果为开展木荷 *WRKY* 基因功能进一步研究其生物学功能提供了有价值的线索。

Genome-Wide Identification and Analysis of the *WRKY* Gene Family in *Schima superba*

Changya Chen¹, Yanling Cai¹, Boxiang He¹, Qian Zhang¹, Dongcheng Liang¹, Yingli Wang¹, Jun Yao^{1,*}

(1. Guangdong Provincial Key Laboratory of Silviculture Protection and Utilization, Guangdong Academy of Forestry, Guangzhou 510520, China)

Abstract: 【Objective】Based on the genome data of *Schima superba*, bioinformatics methods were used to identify all members of the *WRKY* gene family in *Schima superba*. Analyses were conducted on their physicochemical properties, phylogeny, cis-acting elements, gene structure, motif analysis, conserved domains, chromosomal localization, and collinearity. This provides a reference for subsequent exploration of the functional roles of *SsWRKY* genes. 【Method】Using bioinformatics methods, the *WRKY* gene family in *Schima superba* was identified based on its genome data. The physicochemical properties, gene structure, conserved domains, chromosomal localization, and collinearity of this family were analyzed using TBtools software. Phylogenetic trees were constructed using MEGA software, and cis-acting elements were analyzed using the PlantCARE website.

【Result】A total of 82 members of the *WRKY* gene family were identified in the *Schima superba* genome data and named *SsWRKY1* to *SsWRKY75* based on their similarity. These genes are unevenly distributed across 18 chromosomes, encoding amino acids ranging from 153 to 770, with molecular weights between 17,701 and

783,888.06 Da. Among them, 30 SsWRKY proteins are basic ($pI > 7$), 52 are acidic ($pI < 7$), and 80 are unstable proteins. Based on phylogenetic relationships, they were divided into three major groups (I-III), with group II further subdivided into five subgroups (IIa-e), and group II having the most members. The number of exons in *SsWRKY* genes ranges from 2 to 7, and the number of introns ranges from 1 to 6. The domains and conserved motifs within the same family are similar, and all members contain the WRKYGQK heptapeptide conserved domain and zinc finger structure. In the *SsWRKY* gene family, segmental duplication plays a more significant role in gene family expansion than tandem duplication, and the collinearity relationship between *Schima superba* and *Arabidopsis thaliana* is much higher than that with *Oryza sativa*. Promoter analysis revealed the presence of various response elements related to plant growth and development, hormones, drought, and low temperature. **【Conclusion】** We identified a total of 82 WRKY family genes were identified in *Schima superba*, which are divided into 7 subfamilies. The gene structures and conserved protein domains differ among the various subfamilies. Segmental duplication may be a driving force in their evolution. These results provide valuable clues for further research into the biological functions of *WRKY* genes in *Schima superba*.