Genome-wide Identification and Analysis of the *bHLH* Gene Family and Its Role in Carotenoid Biosynthesis in Wolfberry (*Lycium barbarum* L.)

Yue Yin, Xiaoya Qin, Jia Mi, Jianhua Zhao, Guoli Dai, Xinru He, Wei An, Youlong Cao*
National Wolfberry Engineering Research Center, Ningxia Academy of Agricultural and Forestry Sciences, Yinchuan, 751002, China.
*Corresponding author. Email: youlongchk@163.com

Background The basic helix-loop-helix (bHLH) transcription factor family is the second-largest family in plants, where it plays essential roles in plant growth and development, secondary metabolism, and the responses to multiple abiotic and biotic stressors [1]. Wolfberry (*Lycium barbarum* L.) is an important traditional medicinal and food supplement in China, which is a rich source of carotenoids [2]. However, the bHLH transcription factor family's function in carotenoid synthesis of wolfberry has not been studied systematically.

Methods Through a genome-wide search against the wolfberry reference genome, the genomic organization, evolution and expression of bHLH family in wolfberry were systematically analyzed.

Results In this study, we identified 148 bHLH transcription factors in wolfberry genome. The *LbabHLH* genes were grouped into 24 subgroups based on phylogenetic analysis, conserved gene structures, and motif composition. These transcription factors were named *LbabHLH1* to *LbabHLH148*, and they were distributed on all 12 chromosome of which chromosome 1 is the most distributed with 21 (Fig.1a). Gene duplication analysis revealed that DSD and WGD mainly contributed to the expansion of *LbabHLHs* and duplicated genes were subjected to strong purifying selection (Fig.1b). A synteny analysis indicated that 137, 73, 131, 92, and 120 *LbabHLH* genes were orthologous to *Solanum lycopersicum*, *Capsicum annuum*, *Solanum tuberosum*, *Solanum melongena*, and *Arabidopsis thaliana*, respectively (Fig.1c). Carotenoid contents (zeaxanthin, antheraxanthin, β-cryptoxanthin, and lutein palmitate) increased sharply as maturation progressed (Fig.2b). Furthermore, expression analysis revealed that *LbabHLHs* were widely expressed in different tissues (Fig.2c). The co-expression network was further analyzed to underpin the regulatory function of *LbabHLHs* (Fig 2e). Finally, thirteen genes were selected for qRT-PCR validation. The expression of *LbabHLH37* and *LbabHLH60* genes were consistent with carotenoid content. Therefore, we speculated that *LbabHLH37* and *LbabHLH60* participated in the carotenoid biosynthesis process. Overall, this study improves our understanding of *LbabHLH* gene family characteristics and identifies genes involved in the regulation of carotenoid biosynthesis in wolfberry.

Conclusion: This study reported the genome organization, evolutionary characteristics, and expression profile of bHLH family in wolfberry, which not only provide the targets for further functional analysis, but also facilitate better understanding of the regulatory network *bHLH* genes involved in the regulation of carotenoid biosynthesis in wolfberry.

Funding: This research was funded by the Key Research and Development projects of Ningxia Hui Autonomous Region (No.2022BBF02008).

References

[1] Bailey, P. C., Martin, C., Toledo-Ortiz, G., Quail, P. H., Huq, E., Heim, M. A., Jakoby, M., Werber, M., & Weisshaar, B. (2003). Update on the basic helix-loop-helix transcription factor gene family

in Arabidopsis thaliana. The Plant cell, 15(11), 2497–2502.

[2] Long, J. T., Fan, H. X., Zhou, Z. Q., Sun, W. Y., Li, Q. W., Wang, Y., Ma, M., Gao, H., & Zhi, H. (2020). The major zeaxanthin dipalmitate derivatives from wolfberry. Journal of Asian natural products research, 22(8), 746–753.

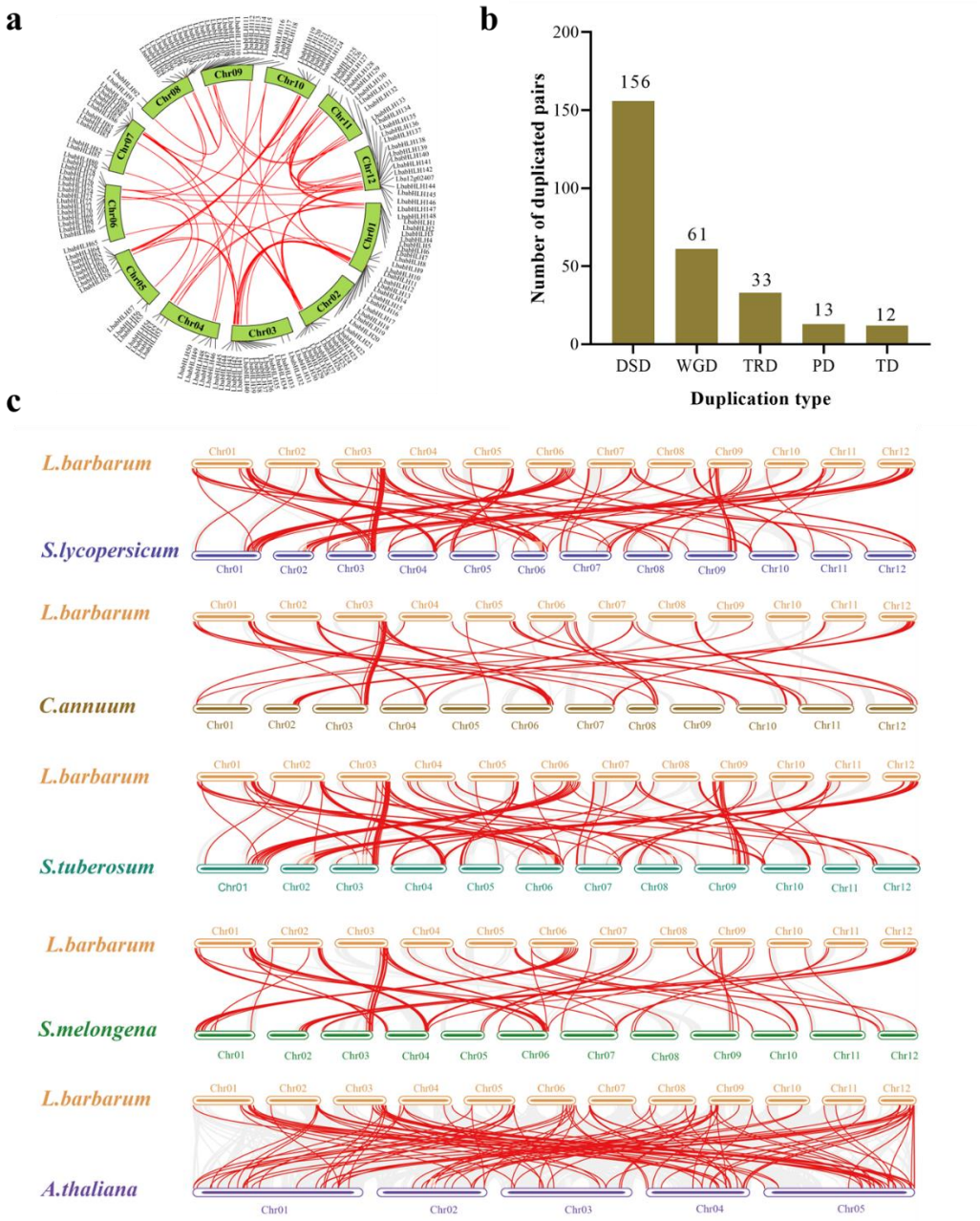

Fig.1 Gene duplication events and collinearity analysis. (a) Different models of gene duplication in *bHLH* gene family. The *x*-axis represents the duplication type. The *y*-axis represents the number of duplicated gene pairs. (b) Gene location and intraspecies collinearity analysis of *L. barbarum bHLH* genes. (c) Synteny analyses of *bHLH* genes between *L. barbarum* and five representative plant species (*Solanum lycopersicum*, *Capsicum annuum*, *Solanum tuberosum*, *Solanum melongena*, and *Arabidopsis thaliana*). Gray lines on the background indicate the collinear blocks in *L. barbarum* and other plant genomes; red lines highlight the syntenic *L. barbarum bHLH* gene pairs.

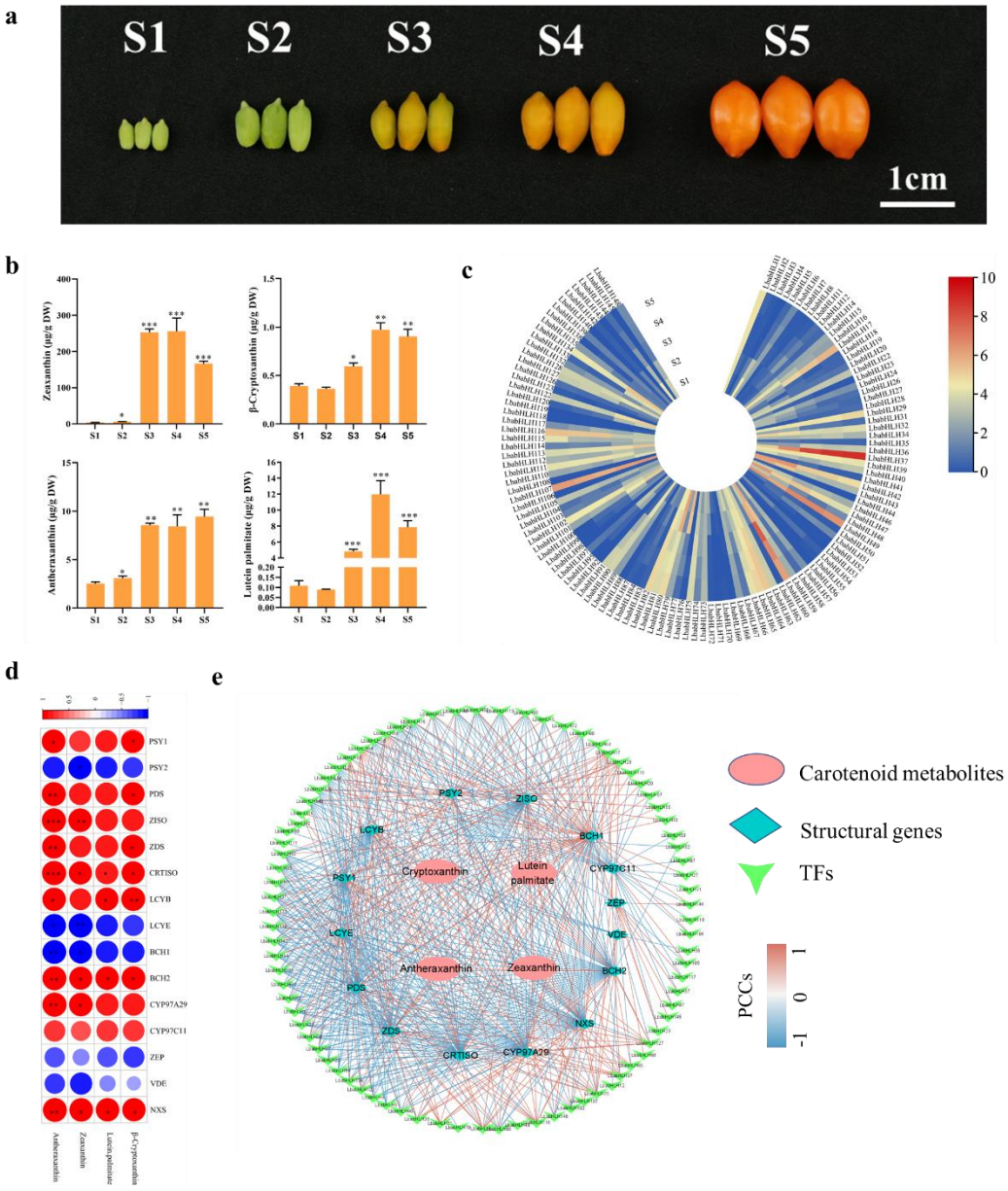

Fig.2 Identification bHLH transcription factor involved in the regulation of carotenoid biosynthesis in wolfberry. (a) Fruits of *Lycium barbarum* var. auranticarpum at different stages of development. S1, S2, S3, S4, and S5 period represent 12, 19, 25, 30, and 37 days after full bloom (DAF), respectively. Scale bars represent 1 cm. (b) Trends in carotenoids (zeaxanthin, antheraxanthin, β–cryptoxanthin and lutein palmitate) at five developmental stages. The data contains the averages and standard deviations of three individual replicates. Asterisks indicate a significant difference (*$p<0.05$, **$p<0.01$, ***$p<0.001$) compared with S1 at the different time points during development. (c) (d) Correlation analysis was constructed using the expression levels of structure genes and carotenoids content in five different developmental stages. The blue color means negative correlation, the red color means positive correlation. Asterisks indicate a significant difference (*$p<0.05$, **$p<0.01$, ***$p<0.001$). (e)