

***In silico* identification of MYB and bHLH families reveals candidate transcription factors for secondary metabolic pathways in *Cannabis sativa* L.**

Laura Bassolino *, Matteo Buti, Flavia Fulvio, Alessandro Pennesi, Giuseppe Mandolino, Justyna Milc, Enrico Francia and Roberta Paris

Abstract: Plant secondary metabolic pathways are finely regulated by the activity of transcription factors, among which members of the bHLH and MYB Subfamilies play a main role. *Cannabis sativa* L. is a unique officinal plant species with over 600 synthesized phytochemicals having diverse scale-up industrial and pharmaceutical usage. Despite a comprehensive knowledge on cannabinoids metabolic pathway, very little is known about their regulation, while literature on flavonoids metabolic pathway is still scarce. In this study we provide the first genome-wide analysis of bHLH and MYB families in *C. sativa* reference cultivar CBDRx and identify 93 and 89 genes coding for these transcription factors, respectively. CsbHLHs and CsMYBs were then classified into functional Subfamilies through comparative phylogenetic analysis with *A. thaliana* transcription factors. Analyses of gene structure and motif distribution confirmed that CsbHLHs and CsMYBs belonging to the same evolutionary clade share common features at both gene and amino acidic level. Candidate regulatory genes for key metabolic pathways leading to flavonoids and cannabinoids synthesis in Cannabis were also retrieved. Furthermore, a candidate gene-approach was used to identify structural enzymes-coding genes for flavonoids and cannabinoids synthesis. Taken as a whole, this work represents a valuable resource of candidate genes for further investigating of the *C. sativa* cannabinoids and flavonoids metabolic pathways for genomic studies and breeding programs.

Keywords: Cannabis sativa; Genome-wide analysis; MYBs; bHLH; Flavonoids