

In silico Identification of MYB and bHLH Families Reveals Candidate Transcription Factors for Secondary Metabolic Pathways in *Cannabis sativa* L.

Laura Bassolino¹, Matteo Buti², Flavia Fulvio¹, Alessandro Pennesi¹, Giuseppe Mandolino¹, Justyna Milc³, Enrico Francia³ and Roberta Paris¹

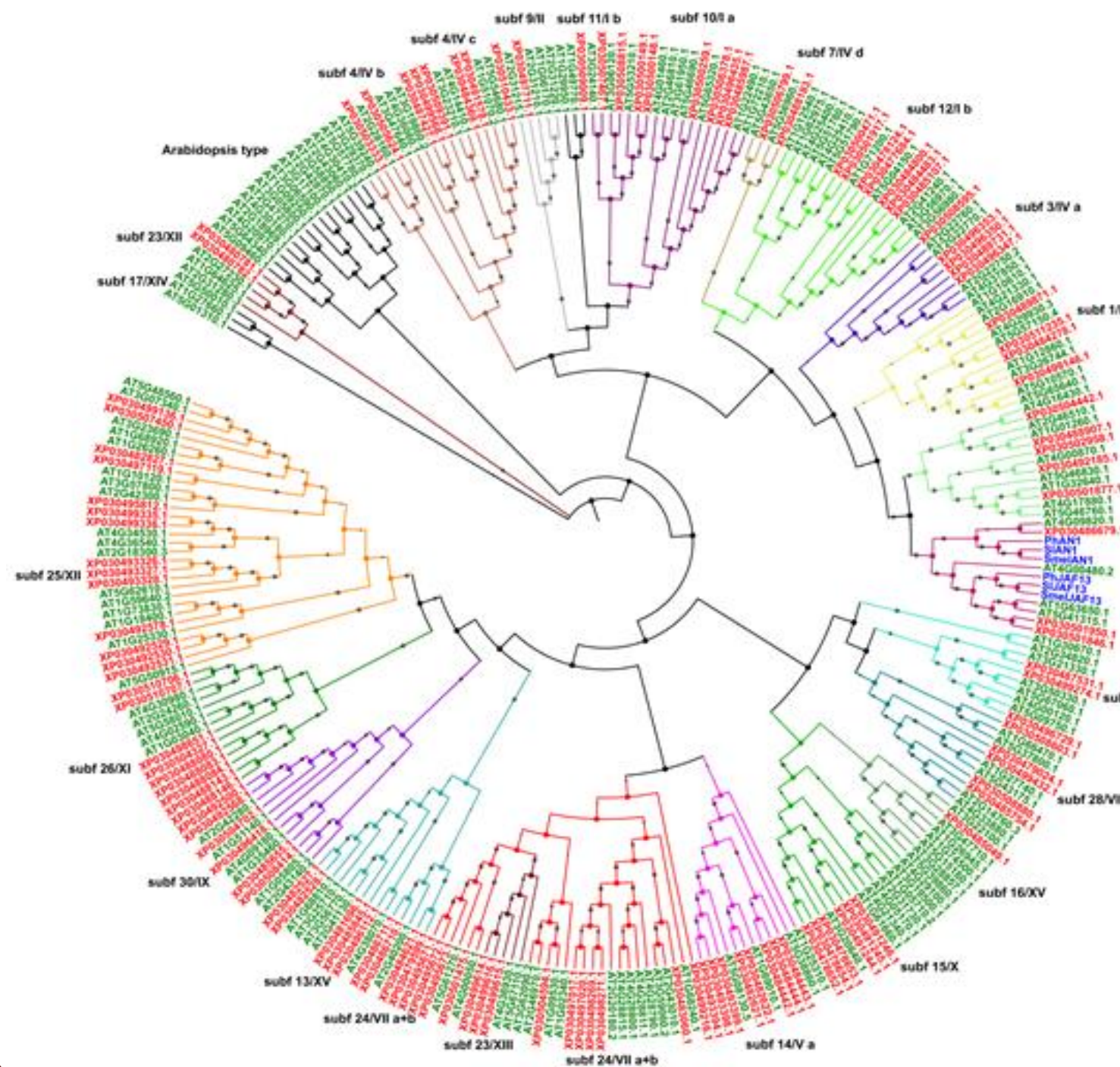
¹CREA-Research Centre for Cereal and Industrial Crops, 40128 Bologna; ²Department of Agriculture, Food, Environment and Forestry, University of Florence, 50144 Firenze; ³Department of Life Sciences, Centre BIOGEST-SITEIA, University of Modena and Reggio Emilia, 42122 Reggio Emilia, ITALY;

Despite the high conservation of flavonoid biosynthetic routes in model and non-model plant species, a unique report does not exist to date on the characterization of structural and regulatory genes involved in flavonoid synthesis in *Cannabis sativa* L. Further, very little is known about the regulation of cannabinoids' metabolic pathways. Thus, although performed exclusively *in silico*, this work represents the first comprehensive step in this direction.

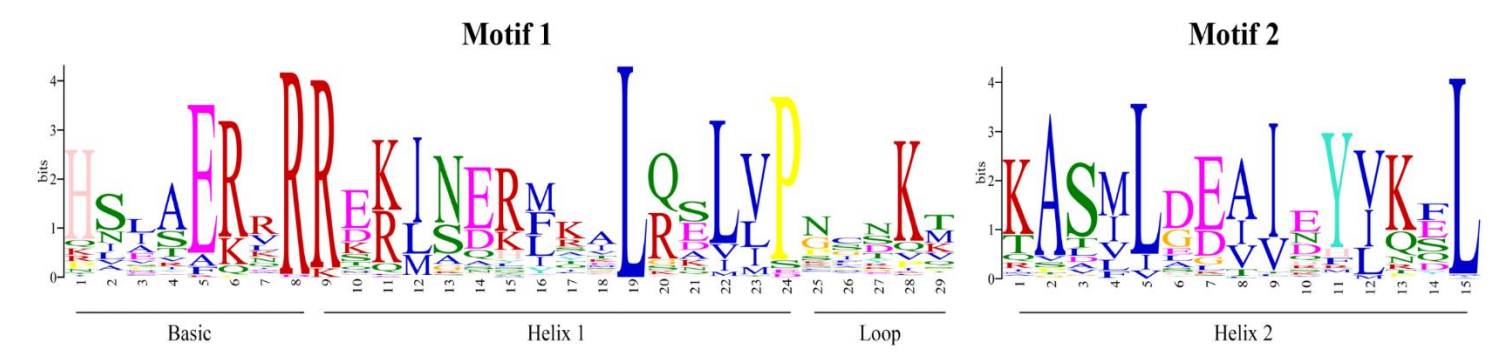


We performed a genome-wide analysis of bHLH and MYB families in *C. sativa* cs10 genome reference of cv. CDBRx. CsbHLHs and CsMYBs were classified into functional subfamilies via comparative phylogenetic analysis with *A. thaliana*. Analyses of gene structure and motif distribution confirmed that members of both superfamilies belonging to the same evolutionary clade share common features at both gene and amino acid level. Structural enzyme-coding genes for flavonoid and cannabinoid synthesis were also identified via candidate gene approach.

The bHLH subfamily structure in *Cannabis sativa* cv CDBRx

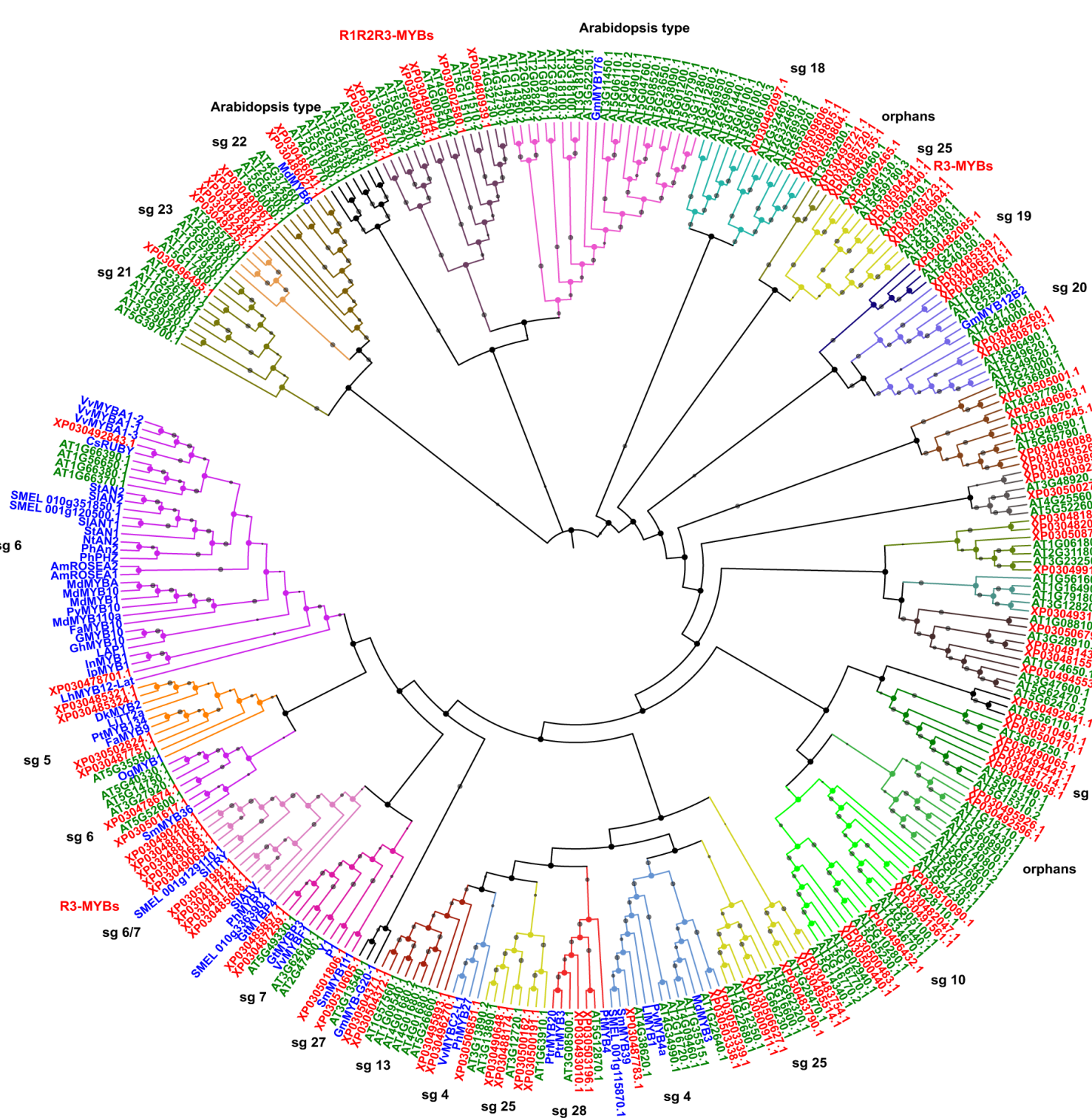


- 89 genomic loci encoding for 121 *Cannabis* bHLH
- Comparative phylogenetic analysis with AtbHLH, classified *Cannabis* bHLHs into distinct 25 functional subfamilies with Subf25 being the largest group with 14 members
- CsbHLH111-121 share the conserved PF14215.6 signature, thus these proteins can be involved in the regulation of phenylpropanoids
- Cannabis* orthologs of JAF13 (CsbHLH113-114) and AN1 (CsbHLH112), known anthocyanin regulatory bHLH, were found in two distinct evolutionary clades of Subf11f thus confirming a common pattern to other species
- In Subf10 we found CsFMA, CsSPCH, CsMUTE clustering with corresponding Arabidopsis sequences suggesting a conserved role in stomatal control in *Cannabis*
- Candidate CsbHLH for brassinosteroids, response to biotic and abiotic stimuli, hormones, light signaling and perception were also identified

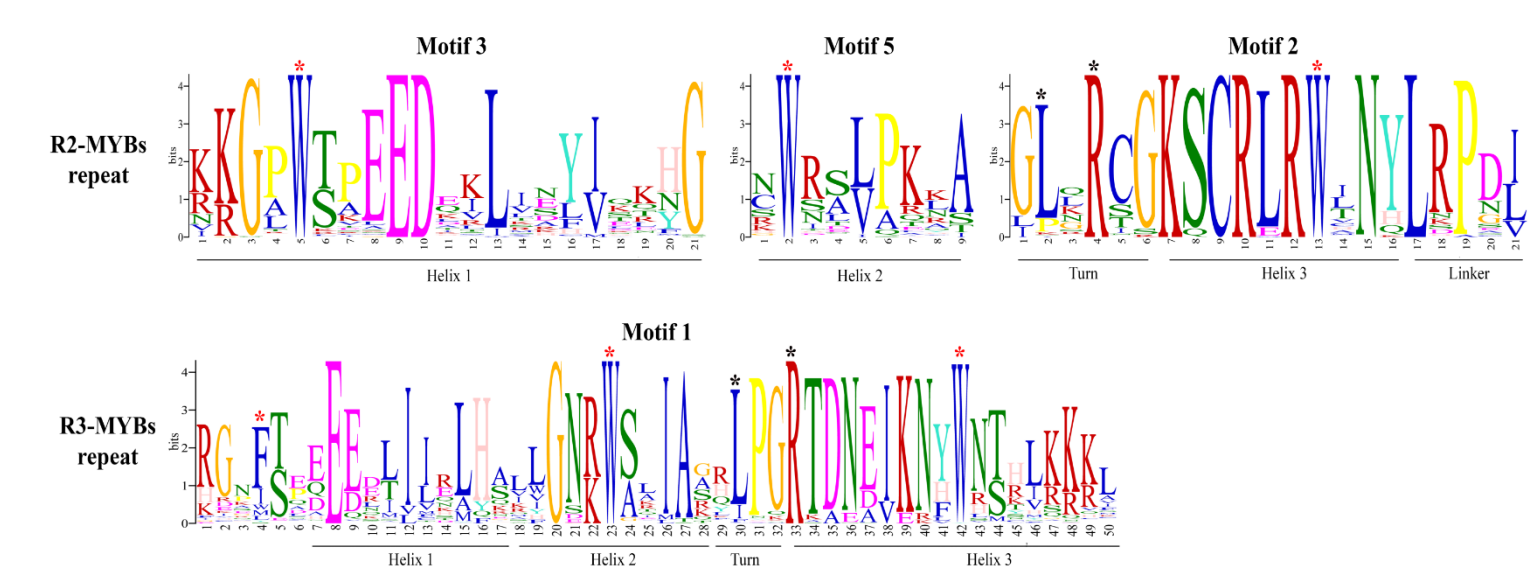


- The logo consensus of Cannabis bHLHs. Sequences of motif 1 and 2 represent the bHLH domain in *Cannabis* which is conserved respect to other species.
- bHLHs falling into the same or adjacent subfamilies share a similar pattern of conserved motifs and intron-exon structure

The MYB subfamily organization in *Cannabis sativa* cv CDBRx



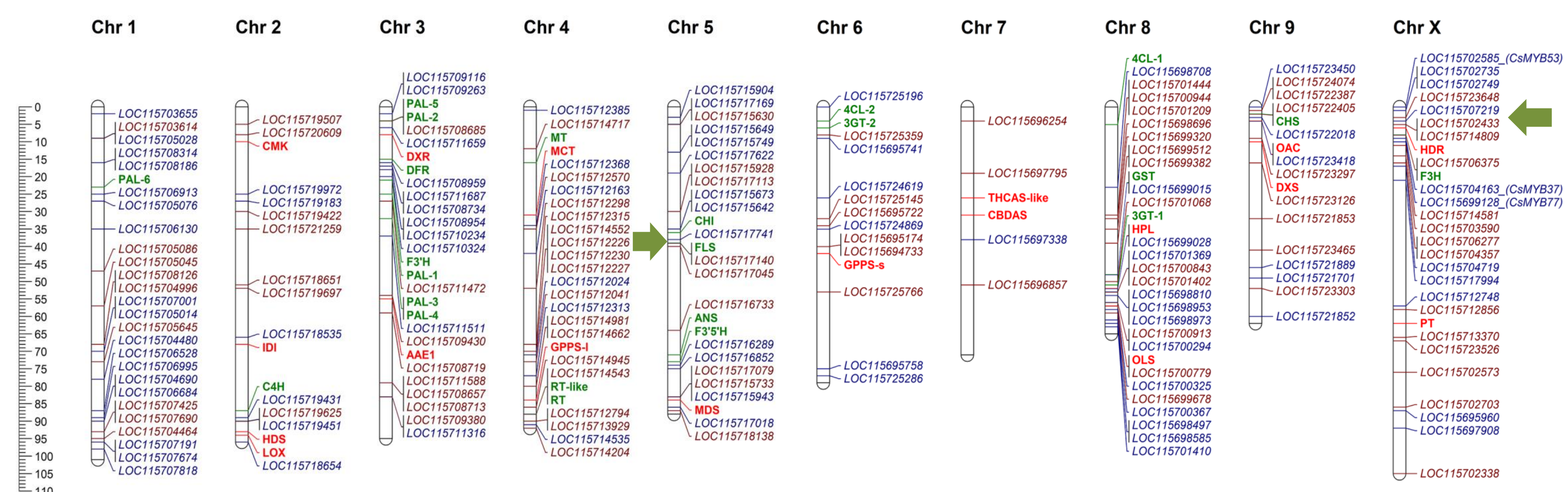
- 94 genomic loci encoding for 104 *Cannabis* MYB
- Based on the number of R repeats in the N-terminal domain, MYB proteins were classified into three subclasses: 10 belonging to R3-MYBs, 6 to R1R2R3-MYBs, and 88 to R2R3-MYBs
- Phylogenetic analysis clustered CsMYBs into 27 distinct subgroups
- Candidate inducers of anthocyanin pigmentation in *Cannabis* were found in subg6, like MYB82 and MYB87 that are orthologs of CsRUBY and InMYB1, respectively
- Candidate R2R3-MYBs repressors of anthocyanins were also found like CsMYB59 which clustered in subg4 with PhMYB27 and VvMYBC2-L1
- Further, CsR3-MYBs spanned from subg6 and 7 with known R3-type repressors for anthocyanins suggesting a similar role in *Cannabis*.
- The *Cannabis* orthologs of AtMYB78 and AtMYB12 were found in subg20 and 7, respectively. These MYBs are both putatively involved in cannabinoids regulation.
- Orthologs of known monolignol and phenolic acid related MYBs were identified (subg 4,13)
- The CsR1R2R3-MYBs clustered with Atr3MYBs which are involved in cell division thus, these proteins might exert a similar role in regulation of trichomes differentiation in *Cannabis*



- Logo consensus sequences of Cannabis MYBs. The sequences of motifs 3, 5 and 2 together constitute the R2 repeat, motif 1 correspond to R3 repeat of MYBs domain, conserved tryptophan residues are marked with red stars;
- Members of the same class of R type MYBs clustering in the same phylogenetic clade share a common exon-intron structure and motifs distribution

Chromosomal Distribution of Genomic Loci for CsMYBs, CsbHLHs, and biosynthetic enzymes for flavonoids and cannabinoids

Chromosomal location of *Cannabis* MYBs, bHLHs was displayed with 20 genomic loci encoding for flavonoid structural enzymes and 18 sequences of cannabinoids biosynthetic genes, identified by direct protein-to-protein search via the BLAST tool. We found interesting regions on Chr X and 5 with both structural and candidate regulatory genes for both flavonoids and cannabinoids. Indeed, the flavonoids-related biosynthetic genes *FLS*, *CHI*, *ANS*, *F3'5'H* co-map with *bHLH112-113* (subf 11f), orthologs of *PhAN1* and *PhJAF13*, respectively which are known anthocyanin regulators. Further, we noticed that adjacent genes on the same chromosome, clustered in the same phylogenetic clade thus suggesting a similar functional role such as the case of bHLH86-88 belonging to subf28 and involved in root hair formation.



This work represents a valuable resource of candidate genes for further investigation of the *C. sativa* cannabinoid and flavonoid metabolic pathways for genomic studies and breeding programs.

The full story is out: Bassolino, L.; Buti, M.; Fulvio, F.; Pennesi, A.; Mandolino, G.; Milc, J.; Francia, E.; Paris, R. *In Silico Identification of MYB and bHLH Families Reveals Candidate Transcription Factors for Secondary Metabolic Pathways in Cannabis sativa* L. *Plants* 2020, 9, 1540.



UNIONE EUROPEA
Fondo Europeo di Sviluppo Regionale



Ministero dell'Università e della Ricerca



PON
Ricerca e Innovazione



UNIHEMP
CANNABIS PROJECTO



crea
Consiglio per la ricerca in agricoltura e l'analisi dell'economia agraria