Research Article

Prospects for Knockout of MYB60, a Transcriptional Repressor of Anthocyanin Biosynthesis, in Brassicaceae Plants by Genome Editing

Elena Mikhaylova*, Alexander Artyukhin, Michael Shein, Khalit Musin, Anna Sukhareva, and Bulat Kuluev

[1]Institute of Biochemistry and Genetics UFRC RAS, Prospekt Oktyabrya, 71, Ufa, Russia
ORCID
Elena Mikhaylova; 0000-0001-7374-8405

Abstract. The *Brassicaceae* plant family contains many economically important crops such as *Brassica napus* L., *Brassica rapa* L., *Brassica oleracea* L., *Brassica junce*a L., *Eruca sativa* Mill., *Camelina sativa* L. and *Raphanus sativus* L. Insufficient data on the genetic regulation of agronomic traits in these species complicates the editing of their genomes. In recent years, the attention of the academic community has been drawn to anthocyanin hyperaccumulation. This trait is not only beneficial for human health, but can also increase plant resistance to stress. MYB transcription factors are the main regulators of flavonoid biosynthesis in plants. Some of them are well studied in *Arabidopsis thaliana*. The *AtMYB60* gene is a transcriptional repressor of anthocyanin biosynthesis, and it also negatively impacts plant responses to drought stress. Myb60 is one of the least studied transcription factors with similar functions in *Brassicaceae*. There is a high degree of homology between predicted *MYB60* genes of *A. thaliana* and related plant species. However, functions of these homologous genes have never been studied. Gene knockout by CRISPR/Cas technology remains the easiest way to perform genome editing in order to discover the role of individual plant genes. Disruption of genes acting as negative regulators of anthocyanin biosynthesis could result in color staining of plant tissues and an increase in stress tolerance. In the present study, we investigated the *AtMYB60* gene and its homologs in *Brassicaceae* plants and suggested universal gRNAs to knockout these genes.

Keywords: CRISPR, *Brassicaceae*, MYB60, knockout, anthocyanin

Corresponding Author: Elena Mikhaylova; email: mikhele@list.ru

Dates
Published 13 January 2022

Publishing services provided by Knowledge E

Selection and Peer-review under the responsibility of the 8th Scientific and Practical Conference Conference Committee.

1. Introduction

Crop productivity is one of the main concerns of biology and agricultural sciences. Stress factors, which are the most severe in temperate climate, decrease the yields. Despite cold tolerance of *Brassicaceae* plants, their productivity in northern regions leaves room for improvement. Canola, turnip, cabbage and false flax remain extremely susceptible to fungal diseases (powdery mildew, light leaf spot, alternaria) and pests (cabbage-stem flea beetle, cabbage leaf and flower beetle) [1]. Pesticides are widely

used to protect these crops. Genetically modified varieties, resistant to herbicides and insects, have been adopted in recent years. This strategy promotes the development of pesticide resistance in pests and weeds. The improvement of plant stress response mechanisms is a more effective and environmentally friendly method.

Genome editing is a novel technique that allows to obtain plants with desired traits. The main difference from traditional genetic engineering is that changes are being made in the particular part of the genome [2].

Hyperaccumulation of anthocyanins is associated with resistance to abiotic stress, pests, phytopathogenic fungi and bacterial diseases [3,4,5,6]. This trait could increase crop yields in temperate climate. Structural genes in the anthocyanin biosynthesis pathway are regulated by three types of transcription factors: with the MYB domain, with the helix-loop-helix domain (bHLH) and with WD repeats [7]. This complex is being actively studied in *Arabidopsis thaliana*, a popular model plant of family *Brassicaceae*. The availability of annotated genes, involved in anthocyanin biosynthesis in *A. thaliana*, opens a prospect for precise genome editing of related plant species. However, the mechanisms of regulation of anthocyanin biosynthesis can differ between them.

Full genomic sequences of important crops *B. napus*, *B, rapa*, *B. oleracea*, *B. juncea*, *R. sativus* and *C. sativa* are available, but most of the genes have not been annotated yet. Little is known about the functions of the genes that are homologous to those of *A. thaliana*.

Gene knockout proved to be a good technique to study gene functions. Negative regulators of flavonoid biosynthesis pathway should be chosen as targets to increase anthocyanin accumulation in this case. Several genes with required functions are already identified in *A. thaliana*. Genes AtMYB60 [8], CPC (At2g46401) [9] and AtMYBL2 [10] were reported to suppress anthocyanin accumulation. The functions of AtMYB60 gene homologs in other *Brassicaceae* plants are unknown. The study of MYB60 plant genes has significant novelty and potential for practical application.

2. The role of MYB transcription factors in anthocyanin biosynthesis

The MYB family is one of the most common groups of transcription factors described in plants. MYB proteins include the conserved R2R3-type MYB domain, comprising a DNA-binding domain and an activation/repression domain [8]. Non-MYB regions of these proteins are variable in length and amino acid content. They are also involved in different kind of interactions [11].

There are ~126 *MYB* genes in *A. thaliana* [8]. MYB transcription factors play important roles in metabolism, cell fate determination, growth and development, and stress response. For example, AtMYB59 and AtMYB77 regulate root growth, and AtMYB96 contributes to drought stress response and pathogen resistance [12]. Some MYB transcription factors regulate flavonoid biosynthesis by forming protein complexes with bHLH transcription factors and a WD40 repeat proteins which regulates the expression of structural genes for anthocyanins and proanthocyanidins.

Transcription factors PAP1 and PAP2 are well studied positive regulators of the expression of genes encoding chalcone synthase, dihydroflavonol 4-reductase and leucoanthocyanidin dioxygenase [10]. Several genes proved to be transcriptional repressors of anthocyanin synthesis in *Arabidopsis*. CPC transcription factor reduces the expression of dihydroflavonol 4-reductase (*DFR*) gene by competing with PAP1 and PAP2 to contact with bHLH proteins GL3 and EGL3 [9]. MYBL2 transcription factor negatively regulates anthocyanin biosynthesis due to the suppression of *DFR* and TRANSPARENT TESTA8 (*TT8)* genes expression [10].

Many of MYB transcription factors have multiple functions. For example, MYB60 is involved in stomatal movement, root growth and anthocyanin biosynthesis [13].

3. Role of the MYB60 genes

The *AtMYB60* gene has three exons and three splice variants, according to TAIR database (AT1G08810). Two of them are reported to have the same functions [13]. *Myb60* genes contain conserved PLN03212 model domain, which may span more than one domain and not assigned to any superfamily. It probably includes transcription repressor domain MYB5 and protein-protein interaction domain SANT [14]. Resulting proteins are localized in the nucleus.

Transgenic plants overexpressing each splice variants of the gene had no phenotypic alterations under normal conditions, but were more susceptible to drought than wild-type plants and developed greater root mass in a medium containing mannitol [13]. The knockout mutants, on the other hand, demonstrated increased resistance to drought [15].

Expression of the *AtMYB60* gene is upregulated by factors that induce stomatal opening, such as white and blue light. Factors that promote stomatal closure, such as darkness, desiccation and ABA treatment, downregulated the expression of this gene [15]. *AtMYB60* minimal promoter region is situated between -262 bp and -205 bp and contains the DOF1cis-element required for guard cell expression. Its removal resulted in

the absence of gene expression [16].This gene might be a a counterpart of transcription factor gene *MybA* for proanthocyanidins synthesis [17]

Myb60 orthologs were studied in cotton [18] and grape [19]. *VvMYB60* appeared to be the only true ortholog of *AtMYB60* in the grape genome and changed its expression level during development of grape organs and in response to ABA and drought. *MYB60* transcript levels increased in cotton at the early stages of drought stress.

4. A bioinformatic approach for identification of the MYB60 genes in Brasicaceae plants

According to OrthoDB catalog [20], there are 12 orthologs of *AtMYB60* gene in 8 *Brassicaceae* species. Most of them are represented by only one copy, however in *B. rapa* there are two copies of this gene, and in *C. sativa* there are three copies. The *MYB60* gene of *B. napus* is not annotated, however one copy can be found in A8 chomosome by sequence homology via BLAST. With regard to the facts that *B. napus* is a tetraploid, there might be several more copies. Functions of all these genes have never been studied.

Full genome, mRNA and protein sequences of *AtMYB60* gene in *Brassicaceae* species *A. thaliana*, *A. lyrata, C. sativa*, *B. rapa*, *B. oleracea*, *B. napus* and *R. sativus* were aligned using MEGA software (Figure 1).

The second copy of the gene in *B. rapa* on the A6 chromosome is clearly remote from the genes located on the eighth chromosome of other *Brassicaceae* species, and its protein product (Figure 1C) belongs to *Arabidopsis* cluster. All three gene copies in *Camelina* belong to individual cluster with high support level.

Myb60 protein structures were predicted via homology-modelling server SWISS-MODEL [21] (Figure 2).

Despite some differences in MYB60 protein structure between *Brassicaceae* species, all models contain two DNA-binding domains. In *A. thaliana* one gene has three transcript variants (Figure 2 A-C), however *C. sativa* has three copies of genes resulting in quite the same three products as those of *A. thaliana* (Figure 2 G,H). Two of them are identical smaller proteins (Figure 2 G) which may contribute to the higher importance of this isoform. Despite missing exon, shorter splice variant of *AtMYB60* mRNA has the same effect on plant fitness as the full variant [13].

The results of our investigation suggest that *MYB60*-like genes in the studied species have the same functions as *AtMYB60*.

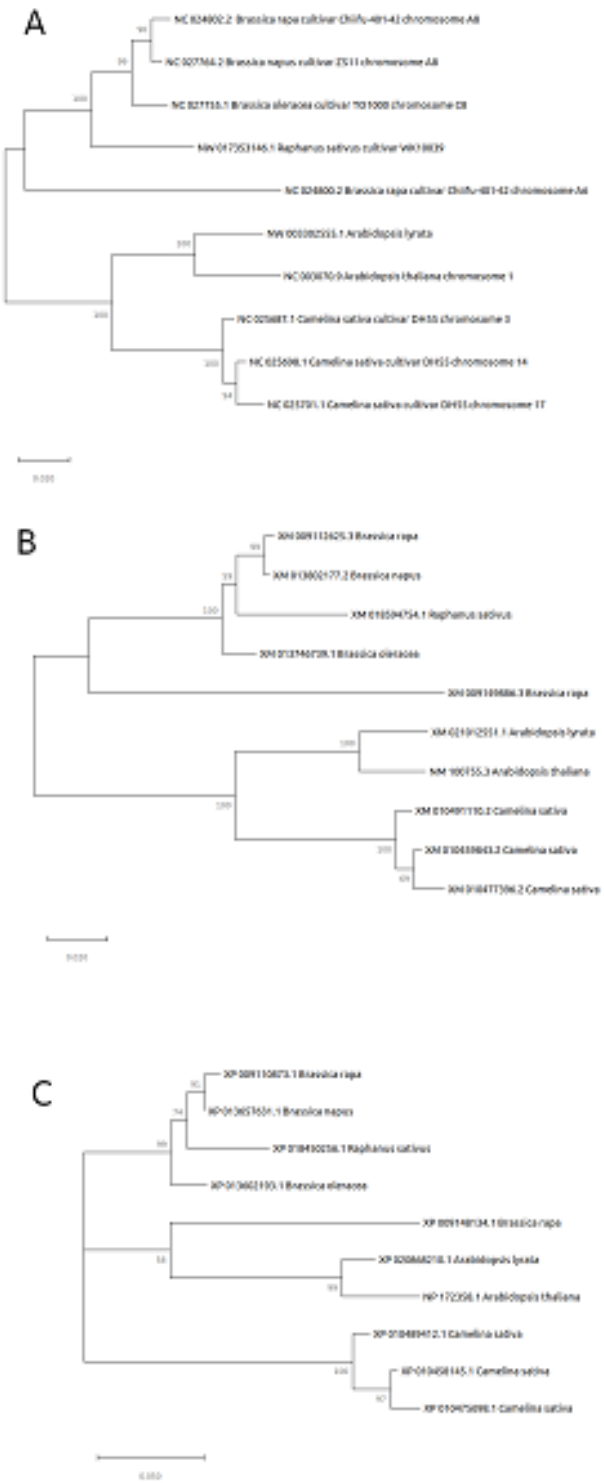

Figure 1: Unrooted phylogenetic maximum likelihood tree of *Brassicaceae* species based on the MYB60 genomic DNA (A), mRNA (B) and protein (C) sequences. Bootstrap values computed by 1000 replicates are given.

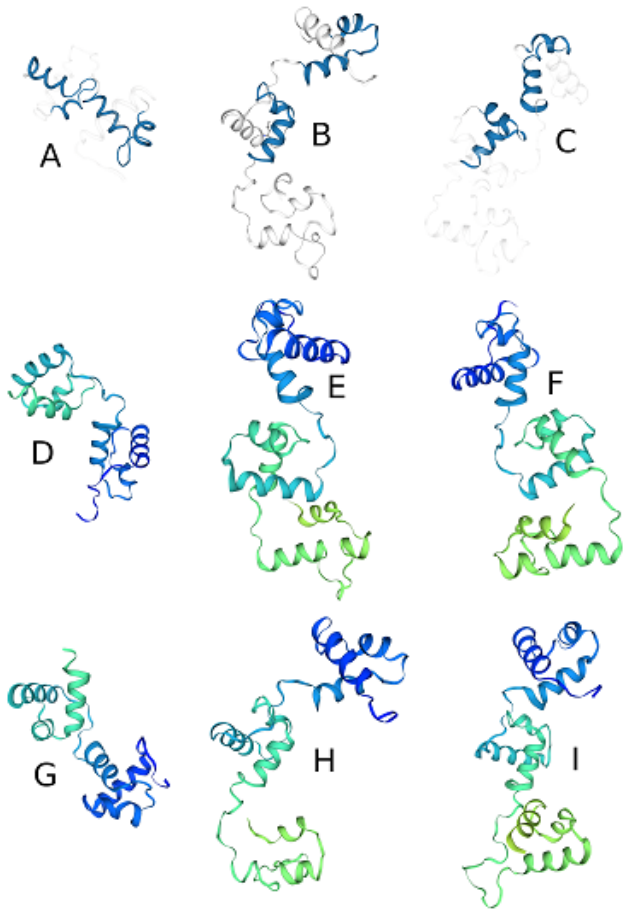

Figure 2: Predicted protein structure of *MYB60* gene product in *A. thaliana* (A-C), *B. rapa* (D-E), *B. napus* (F), *C. sativa* (G,H) and *R. sativus* (I). DNA binding sites in proteins of *A. thaliana* (A-C) are highlighted in blue.

5. The gRNA design

Silencing of the *AtPAP1* gene, which encodes positive regulator of anthocyanin biosynthesis, resulted in the inhibition of the accumulation of proanthocyanidin in the seed coat and suppression of anthocyanin accumulation in the leaves [10]. Silencing of negative regulator of anthocyanin biosynthesis *MYBL2* [10] promoted anthocyanin pigmentation of *Arabidopsis* seeds, stems and leaves due to the induction of *DFR* and *TT8* genes expression. There were no studies on the effect of the *AtMYB60* gene silencing on anthocyanin content, however knockout *Arabidopsis* mutants demonstrated an increased resistance to drought [14], and heterologous expression of *AtMYB60* in lettuce [8] repressed anthocyanin accumulation in the leaves via the inhibition of a dihydroflavonol reductase gene (*DFR*). These manipulations did not have any negative effect on the fitness of the plants. These data suggest that knockout of the genes

encoding negative regulators of anthocyanin biosynthesis, such as *MYB60*, may result in increased accumulation of anthocyanins in *Brassicaceae* plants.

Design of a gRNA is essential for the knockout of target gene and relies on availability of high□quality genome sequence and gene annotations. Cross-pollination and interspecific hybridization are common in *Brassicaceae* plants, therefore sequences available in the databases can not be extrapolated to many existing varieties of the same species. Targeting the most conserved and the most essential gene regions should increase the chances of the knockout.

CRISPOR software [22] was used to design gRNAs using *AtMYB60* gene as a template. However 43 possible guide sequences were predicted by the program, only few of them fell within the most conserved regions and could be used as universal gRNAs for multiple plant species. *AtMYB60* minimal promoter region [16], which would be a perfect target to prevent gene expression, contains only two possible gRNA spacers (TGACTACGCAAAATAGTTAA TGG and CGCAAAATAGTTAATGGAAA AGG). Both are specific for *A. thaliana*. In other plant species this region contains multiple single nucleotide polymorphisms (SNPs).

However GACAAGATAGGGATCAAGAA AGG (Figure 3A) seems to be the perfect spacer, suitable for the knockout of *MYB60* gene in five plant species, it does not interfere with DNA binding site. One of the predicted spacers (TTGTGGTAGGTATGAAGCTA TGG) fall within the desired region, however it has one nucleotide mismatch with sequences of *MYB60* gene of six species. With minimal alternations, it could also be used as a universal guide sequence (Fig. 3B). Most of the other results generated by CRISPOR were specific only for *A. thaliana* (Figure 3C) or *Arabidopsis* sp. and *C. sativa* (Figure 3D). All discussed guide sequences have high specificity. None of them had off-targets with 0-2 mismatches.

6. Conclusion

In the present study we have investigated the *AtMYB60* gene and its homologs in *Brassicaceae* plants. The results of the alignment of genomic DNA sequence, mRNA sequence and protein sequence, as well as the prediction of protein structure, suggest that all studied genes have similar functions.

This research has proposed guide sequences for the knockout of *MYB60* gene in multiple plant species. Obtained results can be used to generate plants with increased anthocyanin pigmentation and stress tolerance via genome editing.

Figure 3: Predicted possible guide sequences for *MYB60* gene (A-D) highlighted in yellow.

7. Funding notice

The reported study was funded by Russian Science Foundation according to the research project №20-74-10053.

8. Ethics polices

Authors declare no conflict of interest. This article does not contain any studies with human participants or animals performed by any of the authors.

References

[1] Warwick SI. (2011). Genetics and genomics of the *Brassicaceae*. New York: Springer; 2011. *Brassicaceae* in Agriculture.

[2] Kuluev, Bulat & Gumerova, Gulnar & Mikhaylova, Elena et al. Delivery of CRISPR/Cas components into higher plant cells for genome editing. Russian Journal of Plant Physiology. 2019;66(5):694-706.

[3] Shoeva OY, Khlestkina EK. Anthocyanins participate in the protection of wheat seedlings against cadmium stress. Cereal Research Communications.

2018;46(2):242-252.

[4] Mitsunami T, Nishihara M, Galis I, et al. Overexpression of the PAP1 transcription factor reveals a complex regulation of flavonoid and phenylpropanoid metabolism in *Nicotiana tabacum* plants attacked by *Spodoptera litura*. PLoS One. 2014;9(9). 108849-108849

[5] Gandikota, Madhuri & De Kochko, Alexandre & Chen, Lili & Ithal, Nagabhushana & Fauquet, Claude & Reddy, Arjula. Development of transgenic rice plants expressing maize anthocyanin genes and increased blast resistance. Molecular Breeding. 2001;7(1):73-83.

[6] Kangatharalingam, N. & Pierce, Margaret & Bayles, Melanie & Essenberg, Margaret Epidermal anthocyanin production as an indicator of bacterial blight resistance in cotton. Physiological and Molecular Plant Pathology. 2002;61(3):189-195.

[7] Li S. Transcriptional control of flavonoid biosynthesis: Fine-tuning of the MYB-bHLH-WD40 (MBW) complex. Plant Signaling & Behavior. 2014;9(1) 27522- 27522

[8] Park JS, Kim JB, Cho KJ, et al. (2008). *Arabidopsis* R2R3-MYB transcription factor AtMYB60 functions as a transcriptional repressor of anthocyanin biosynthesis in lettuce (*Lactuca sativa*). Plant Cell Reports. 2008;27(6):985-994.

[9] Zhu HF, Fitzsimmons K, Khandelwal A, Kranz RG CPC, a single-repeat R3 MYB, is a negative regulator of anthocyanin biosynthesis in *Arabidopsis*. Molecular Plant. 2009;2(4):790-802,

[10] Matsui K, Umemura Y, Ohme□Takagi M. AtMYBL2, a protein with a single MYB domain, acts as a negative regulator of anthocyanin biosynthesis in *Arabidopsis*. The Plant Journal. 2008;55(6):954-967.

[11] Millard PS, Kragelund BB, Burow M. R2R3 MYB transcription factors–functions outside the DNA-binding domain. Trends in Plant Science. 2019;24(10):934-946.

[12] Wang B, Luo Q, Li Y, et al. Structural insights into target DNA recognition by R2R3-MYB transcription factors. Nucleic Acids Research. 2020;48(1):460-471.

[13] Oh JE, Kwon Y, Kim JH et al. A dual role for MYB60 in stomatal regulation and root growth of *Arabidopsis thaliana* under drought stress. Plant Molecular Biology. 2011;77(1-2):91-103.

[14] Lu S, Wang J, Chitsaz F, et al. CDD/SPARCLE: The conserved domain database in 2020. Nucleic Acids Research. 2020;48(D1):265-268.

[15] Cominelli E, Galbiati M, Vavasseur A et al. A guard-cell-specific MYB transcription factor regulates stomatal movements and plant drought tolerance. Current Biology. 2005;15(13):1196-1200.

[16] Cominelli E, Galbiati M, Albertini A et al. DOF-binding sites additively contribute to guard cell-specificity of AtMYB60promoter. BMC Plant Biology. 2011;11(1):162-167.

[17] Falginella L, Di Gaspero G, Castellarin SD. Expression of flavonoid genes in the red grape berry of 'Alicante Bouschet' varies with the histological distribution of anthocyanins and their chemical composition. Planta. 2012;236(4):1037-1051.

[18] Batista, V G L, Pedro Dantas Fernandes, R.C. dos Santos, Péricles A. Melo Filho and Liziane Maria de Lima. (2019) Expression profile of MYB60 and GUSP1 genes during early growth of cotton genotypes submitted to water stress. Genetics and Molecular Research. 2019;18(4). 1-18

[19] Galbiati M, Matus JT, Francia P, et al. The grapevine guard cell-related VvMYB60 transcription factor is involved in the regulation of stomatal activity and is differentially expressed in response to ABA and osmotic stress. BMC Plant Biology. 2011;11(1):142-142.

[20] Waterhouse A, Bertoni M, Bienert S, et al. SWISS-MODEL: Homology modelling of protein structures and complexes. Nucleic Acids Research. 2018;46(W1):296-303.

[21] Kuznetsov, Dmitry & Tegenfeldt, Fredrik & Manni, Mosè & Dias, Renata & Simão, Felipe & Zdobnov, Evgeny. OrthoDB v10: Sampling the diversity of animal, plant, fungal, protist, bacterial and viral genomes for evolutionary and functional annotations of orthologs. Nucleic Acids Research. 2019;47(D1):807-D811.

[22] Haeussler M, Schönig K, Eckert H, et al. (2016). Evaluation of off-target and on-target scoring algorithms and integration into the guide RNA selection tool CRISPOR. Genome Biology. 2016;17(1):148-148.