EXPRESSION QUANTITATIVE TRAIT LOCI ANALYSIS OF *BAN*, *F3H* AND *TT19* GENES IN *BRASSICA NAPUS*

Kun Lu, Cunmin Qu, Kai Zhang, Junxin Lu, Yourong Chai*, Jiana Li*
Chongqing Rapeseed Engineering Research Center; College of Agronomy and Biotechnology, Southwest University, Tiansheng Road 216#, Beibei, Chongqing 400716, P. R. China
*Corresponding authors: ljn1950@swu.edu.cn & chaiyourong1@163.com

Abstract

Flavonone 3-hydroxylase (F3H), anthocyanidin reductase (ANR) encoded by *BANYULS* (*BAN*) gene and glutathione S-transferase encoded by *TRANSPARENT TESTA 19* (*TT19*) gene are three key enzymes in flavonoid pathway, playing important roles in the synthesis and transport of anthocyanin and proanthocyanidin. However, little is known about the expression variations and the regulatory network in *Brassica*. In this study, we examined their expression levels in seeds of 30 days after flowering (DAF) in F_7 recombinant inbred lines (RILs) derived from a cross between *Brassica napus* cultivars Youyan2 (black-seeded) and GH06 (yellow-seeded), and characterized the locus-level regulatory network related to *B. napus BAN*, *F3H* and *TT19* genes. Expression quantitative trait loci (eQTL) mapping revealed five, seven and eight eQTLs for *BAN*, *F3H* and *TT19* genes, respectively. Marker E4M8 was associated with *qBAN-8-2* and *qTT19-8-2*, while marker H022L18-2 was linked with *qBAN-16-4* and *qF3H-16-6*, implying that there might be upstream regulatory genes in E4M8 and H022L18-2 marker flanking regions. In further analysis, we identified four *trans*-eQTLs (*qBAN-16-5*, *qF3H-16-5*, *qBAN-16-4* and *qF3H-11-2*) close to the location of the major QTL controlling yellow-seeded trait of *B. napus*. The 200-kb flanking sequences of the four eQTL on *B. rapa* chromosome A09 showed well synteny to *B. oleracea* genome sequence and partial continuous fragment of chromosome 1 of *Arabidopsis* genome, suitable for candidate gene predication of eQTLs. Based on gene finding results, seven transcription factors were suggested to be the potential upstream candidate(s) controlling expression variations of *BAN*, *F3H* and *TT19* genes. These results could provide a new approach for constructing regulatory pathways that contribute to complex traits, such as yellow-seeded trait.

Key words: *Brassica napus*; transparent testa; expression quantitative trait loci; regulatory network

Introduction

Oilseed rape (*Brassica napus*, AACC, 2n=4x=38) is one of main oil crops in the world. It is an important source to the vegetable oil, meal protein and industry material. Compared with black-seeded cultivars, yellow-seeded *B. napus* shows many good quality traits, such as lower seed coat pigment content, lower meal fiber content, higher seed oil and meal protein content, etc. Selecting of yellow-seeded oilseed rape cultivars with stably inherited yellow seed trait and good agronomic traits has become the most important breeding goals. However, it is very slow advance and difficult improvement to breeding of yellow-seeded *B. napus*, since the genetic and molecular mechanisms of yellow-seed trait were still unclear.

It had been proved that plant seed coat pigments are mainly composed of polymers of proanthocyanidin (PA), which is synthesized via the flavonoid-anthocyanin-proanthocyanidin pathway (simplified as flavonoid pathway here), a core branch pathway of phenylpropanoid pathway. F3H, BAN and TT19 are 3 key enzymes in flavonoid pathway, playing important roles in the synthesis and transport of anthocyanin and proanthocyanidin. In *Arabidopsis thaliana*, the loss of function mutants of the 3 genes all leads to the transformation of the wild-type dark-brown seed to yellow or pale yellow seed, i.e. *transparent testa* (*TT*) trait. At present, the gene families encoding the 3 key enzymes have been cloned by RACE method in our lab, which laid the foundation to further study their molecular and regulatory mechanism.

In order to determine the upstream regulatory network of the 3 gene families, analysis of transcript abundance was carried out on RNA from seeds of 30 DAF in RIL population. Regarding as quantitative traits, the transcript levels of the 3 gene families were examined by QTL mapping method for eQTL detection. Using this method, it was possible to construct regulatory pathways that contribute to complex traits, demonstrating that the principles of eQTL mapping could be applied to *B. napus*.

Materials and Methods

Plant materials and total RNA extraction

The RILs were developed through successive selfing up to six generations from a cross between yellow-seeded female parent 'GH06' and black-seeded male parent 'Youyan2' by single seed descent (SSD). Parental lines and RILs were sown in field trials at plant breeding station of Chongqing Rapeseed Technology Research Center (CRTRC) in 2009 as previously described (Fu et al. 2007). Seeds of 30 DAF in 180 F_7 recombinant inbred lines were harvested and used for total RNA isolation. Total RNA was extracted using the Plant RNA Mini Kit (Watson Biotechnologies, Inc., China). To remove contaminated genomic DNA, the total RNA was treated with RNase-free DNase I (TaKaRa). Quality and concentration of total RNA samples were assessed by agarose gel electrophoresis and spectrophotometry.

RT-PCR detection and expression quantification

One microgram of total RNA extracted from each sample was used to make first-strand cDNA using the RNA PCR Kit (AMV) Ver. 3.0 (TaKaRa) in a 10-μl reaction with Oligo dT-Adaptor Primer. *Brassica napus 26S rRNA* gene was used as the internal control to monitor sample uniformity of initial RNA input and RT efficiency. The specific primers were F26S (5′-CACAATGATAGGAAGAGCCGAC-3′) and R26S (5′-CAAGGGAACGGGCTTGGCAGAATC-3′). Amplification protocol for *Bn26S* was 94°C for 5 min, followed by 21 cycles of amplification (94°C for 1 min, 60°C for 1 min and 72°C for 1 min), then 72°C for 10 min. The transcription levels of *BAN*, *F3H* and *TT19* genes in the RILs were detected by RT-PCR analyses. RT-PCR products were examined with agarose gel electrophoresis and quantified with Gel-Pro Analyzer 4.0 for optical density assessment. In order to control experimental error, all RT-PCRs were done and analyzed with 3 replicates.

eQTL analysis of BAN, F3H and TT19 genes

The genetic linkage map was constructed by Fu et al (2007) in 2007. A total of 444 loci (158 for SSR, 94 for RAPD, 229 for SRAP and 6 for TRAP) were mapped on 26 linkage groups, covering 1375 cM of *B. napus* genome and the average distance between two adjacent markers was 3.09 cM. Composite interval mapping (CIM) (Zeng, 1994) was used for eQTL analysis of expression of *BAN*, *F3H* and *TT19* genes in *B. napus*, using Windows QTL Cartographer 2.5 (Wang et al, 2007). Model 3 of the Zmapqtl module of QTL Cartographer was used to identify eQTL with a scanning interval of 1 cM between markers and putative eQTL along with 1,000 permutations to set an experiment wide significance LOD threshold at 0.01. Default LOD threshold values of 2.0 was used to declare the presence of an eQTL. The eQTL was named according to the suggestions of McCouch et al (1997), starting with '*q*', followed by an abbreviation of the trait name, the name of the linkage group, and the number of eQTLs affecting the trait on the linkage group, such as '*qBAN-4-1*' is the first eQTL for expression of *BAN* gene in the fourth linkage group.

Analysis of flanking sequences of trans-eQTL

To determine the location of *BAN*, *F3H* and *TT19* genes on *B. rapa* chromosome and the type of eQTL, the cDNA sequences of *Arabidopsis* orthologous genes were used as query for BLASTN search against *B. rapa* scaffold database in the *Brassica* Database (BRAD, http://brassicadb.org/brad/). Sequences of eQTL makers were used for BLASTN search in the BRAD and our local *B. oleracea* genome sequence database. The 200-kb of flanking sequences of each maker was extracted from abovementioned two databases. Dot matrix analysis of *B. rapa* and *B. oleracea* genomic sequences in eQTL region were performed using the bl2seq program on the NCBI website. *Ab initio* gene identification in the flanking sequences was obtained using Fgenesh program with a set of trained *Arabidopsis* gene model. Function of predicted genes was annotated based on matches against the GenBank protein database. A *cis*-eQTL coincides with the location of the underlying gene. In case of *trans*-eQTL, the observed location of eQTL does not coincide with the location of the gene, implying that the observed eQTL represents the position of a locus that controls the expression variation of target gene.

Results and Discussion

Expression levels of *B. napus BAN*, *F3H* and *TT19* genes in the RILs were assayed by RT-PCR and then digitally treated. The results showed that the ranges of optical density value were 3.12~42818, 61.03~39986 and 4206.5~39460. Expression of the 3 genes displayed continuous normal distribution, belongs to typical quantitative traits, and could be analyzed by QTL mapping method for eQTL analysis.

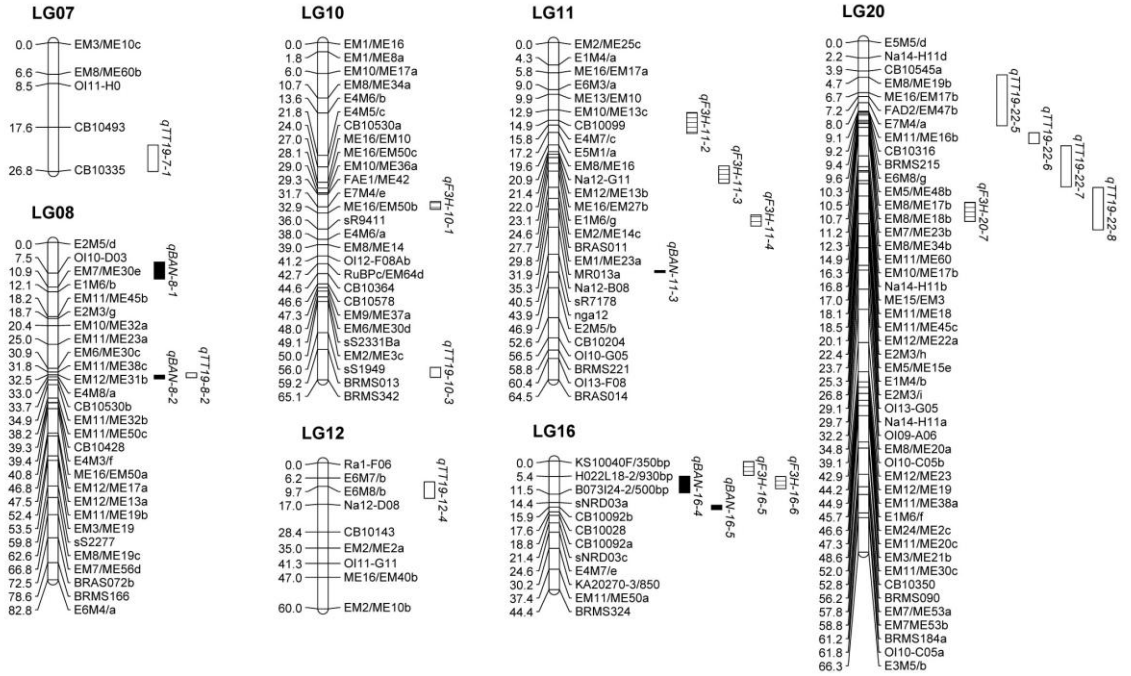

Figure 1 Putative eQTL location for *BAN*, *F3H* and *TT19* genes in genetic linkage map

Composite interval mapping method was employed in eQTL analysis of expression of *BAN*, *F3H* and *TT19* genes. Based on our published genetic map, a total of twenty eQTLs were mapped on eight linkage groups (LG7, LG8, LG10, LG11, LG12, LG16, LG20 and LG22) with LOD≥2.0 (Fig. 1). Phenotypic variation explained by the single putative eQTL varied from 4.09% to 8.95%, and the LOD score ranged from 2.04 to 4.29. Besides, two eQTLs markers E4M8/a and H022L18-2 explaining variation of expression *TT19* and *BAN* genes were located on LG8 and LG16, respectively. Five eQTLs locating on linkage groups LG8, LG11 and LG16 were detected to be responsible for expression of *BAN* with the single contribution to phenotypic variance ranged from4.09% to 8.95%, explained 30.01% total phenotypic variation. Among the 5 eQTLs, only additive effect of *qBAN-8-2* inherited from allelic gene of male parent, the rest additive effect all came from female parent. Two eQTL locating on linkage groups LG8 and LG16 accounted for 14.86% and 11.06% of the phenotypic variation, respectively. Seven eQTLs mapping on linkage groups LG10, LG11, LG16 and LG20 were identified to account for expression of *F3H* with the single contribution to phenotypic variance ranged from 5.28% to 7.79%, explained 45.72% total phenotypic variation. Only additive effect of *qF3H-20-7* was inherited from allelic gene of male parent, the rest additive effect all came from allelic gene of female parent among the 7 eQTLs. And three eQTL locating on linkage groups LG11 explained 19.34% of the phenotypic variation. As for expression of *TT19*, eight eQTLs were mapped on L7, L8, L10, L12 and L22 linkage groups, explaining 49.48% of total phenotypic variation, with the single contribution to phenotypic variance ranged from 4.38% to 8.44%. Additive effect of *qTT19-12-4* was inherited from allelic gene of female parent, while the rest came from those of male parent. Four eQTLs mapping on LG22 could explain 26.04% of the phenotypic variation.

According to BLASTN analysis, two *BAN* genes were identified, and both of them could be located on chromosome A01 of *B. rapa* genome. Both *F3H* and *TT19* have two copies in *B. rapa*. Two *F3H* genes were located on chromosome A03 and A09, while two *TT19* genes were located on chromosome A02 and A09 in *B. rapa*. Among the 20 markers linked to eQTLs, only 8 SSR markers could be located on chromosome A01, A03, A08 and A09 of *B. rapa* genome precisely. Location of these SSR markers was different from those of *BAN*, *F3H* and *TT19* genes in *B. rapa* genome, indicating the 8 eQTLs might be *trans*-eQTL. The 4 eQTL markers on chromosome A09 of *B. rapa* genome linked to 2 and 3 eQTLs controlling expression variation of *BAN* and *F3H* genes. The marker H022L18-2 simultaneously linked to eQTL of *BAN* and *F3H* genes, implying there must be an upstream regulatory gene in H022L18-2 marker flanking region. Several previous studies have proved that there might be a major QTL for yellow seed coat color on linkage group A09/N09 of the A-genome of *B. napus* (Fu et al., 2007; Xiao et al., 2007; Rahman et al., 2009; Snowdon et al., 2010). So, the 200-kb flanking sequences of *trans*-eQTL markers were extracted and analyzed from *B. rapa* and *B .oleracea* genomes.

The synteny between *B. rapa* and *B. oleracea* genome sequences in the 200-kb of flanking sequences of 4 eQTL markers on LG16 (linkage group A09 of *B. napus*) were conserved. The

sequences also showed well synteny to partial continuous fragment of chromosome 1 and 5 of *Arabidopsis* genome, which laid the good foundation for candidate gene predication of *trans*-eQTL. Gene finding and annotation results showed that a zinc finger protein (AT1G08290) and two embryo defective binding proteins (AT1G08840 and AT1G08910) in the H022L18-2 marker region was the primary upstream candidate genes controlling expression variation of *BAN* and *F3H* genes. In flanking regions of another three eQTL markers CB10092b, KS10040F and MR013a, several important MYB, bHLH and zinc finger transcription factors (such as *MYB68*, *bHLH125* and *AZF1*) might also involved in simultaneously regulating expression variation of our target genes. By combining traditional QTL and eQTL approaches, discovery of candidate genes in the major QTL of complex traits would be accelerated. In the future, combination of several kinds of QTL would be more popular in identification of key regulatory regions or genes, and be widely used in molecular plant breeding.

References

Omitted

Acknowledgements

This work was supported by Fundamental Research Funds for the Central Universities (XDJK2009C183), Doctor Foundation of Southwest University (SWU110015) and Specialized Research Fund for the Doctoral Program of Higher Education (20100182120016).