

Genomic regions associated with fiber-related traits in flax

Bankin M.^{1*}, Kanapin A.², Samsonova A.², Rozhmina T.³, Samsonova M.¹

¹ *Mathematical Biology & Bioinformatics Laboratory, Peter the Great St. Petersburg Polytechnic University, St. Petersburg, Russia*

² *Centre for Computational Biology, Peter the Great St. Petersburg Polytechnic University, St. Petersburg, Russia*

³ *Laboratory of Breeding Technologies, Federal Research Center for Bast Fiber Crops, Torzhok, Russia*

* *mikle.p.bankin@gmail.com*

Key words: GWAS, flax, fiber-related traits

Motivation and Aim: For thousands of years mankind have cultivated flax for its fiber [1]. Flax fiber was used as a textile raw material for production of cords, weaving yarn and more later fashionable garment and high-quality fabric upholstery. Since 1930's a unique mechanical properties of flax fibers paved the way to their application as reinforcement in composite materials [2]. The current information on genomic regions associated with fiber is not limited but is also quite inaccurate as the reduced representation sequencing methods were employed instead of whole genome sequencing and in most cases the biparental crosses were used. Several studies identified QTLs associated with fiber yield and fiber content, fiber technical length, straw weight and cell wall [3–5] analyzing RIL populations from crosses between linseed and fiber flax parents. Here we expand the knowledge on genetic control of fiber traits by mining 306 flax accessions maintained at Federal Research Centre of the Bast Fiber Crops. An in-depth characterization of genetic regions controlling fiber-related traits could provide foundation for further genomic selection and breeding efforts aimed on development of cultivars with superior fiber quality.

Methods and Algorithms: 306 flax specimens containing 182 fiber flax accessions, 120 linseed flax accessions and 4 accessions of unknown morphotype were selected from the collection of the Federal Research Center for Bast Fiber Crops. The dataset included landraces, elite cultivars and breeding lines. All plants have been grown in the experimental field of Federal Research Center for Bast Fiber Crops in Torzhok, Tver region, Russia. A detailed description of plant growing and phenotyping was published in [6]. DNA was extracted from collected leaves with DNeasy Plant Mini Kit (Qiagen). DNA sequencing was performed at the BGI using Illumina protocol generating paired-end reads 150bp in length. 1143.850625 Gbytes of raw data comprising 7.626 billion reads with an average of 9.3x coverage or 3.7 Gbp per sample were generated. The genome-wide association analyses were performed with GAPIT3 R package (models: GLM, MLM, CMLM, FarmCPU, SUPER, Blink) [7] and GEMMA model [8].

Results: Univariate GWAS analysis identified 206 significantly associated QTNs, of which 21 were associated with several traits. Most QTNs were associated with technical length (69 QTNs), elementary fibre length (49 QTNs) and fiber content (33 QTNs). Analysis of genes around QTNs revealed 7 potential candidate genes. GWAS results using multivariate model implemented in GEMMA package identified 354 QTNs. Analysis of genes around QTNs, taking into account the LD value for each chromosome, revealed 13 potential candidate genes.

Conclusion: The predicted candidate genes were subdivided into five functional groups: (1) genes related to the biosynthesis and subsequent modification of polysaccharides that make up the secondary and tertiary cell walls of flax fiber; (2) genes associated with vesicular transport and maintenance of intracellular trafficking of the cell wall components; (3) genes associated with the processes accounted for by acid-growth hypothesis (4) genes associated with hormonal regulation, and (5) genes associated with resistance to plant diseases. Our results provide novel insights into the genetic architecture of flax fiber-related traits and pinpoint potential candidate genes for further in-depth studies.

Acknowledgements: This work is supported by Peter the Great St. Petersburg Polytechnic University in the framework of Russian Federation's Priority Strategic Academic Leadership Programme (Agreement 75-15-1038-2021-1333).

References

1. Allaby R.G. et al. Evidence of the domestication history of flax (*Linum usitatissimum* L.) from genetic diversity of the sad2 locus. *Theor Appl Genet.* 2005;112:58-65.
2. Fogorasi M.S., Barbu I. The potential of natural fibers for the automotive sector – review. *IOP Conf Ser: Mater Sci Eng.* 2017;252:12-44.
3. Zhang J. et al. Consensus genetic linkage map construction and QTL mapping for plant height-related traits in linseed flax (*Linum usitatissimum* L.). *BMC Plant Biol.* 2018;18:160.
4. Kumar S. et al. QTL for fatty acid composition and yield in linseed (*Linum usitatissimum* L.). *Theor Apple Genet.* 2015;128:965-84.
5. Xie D. et al. Genomic variations and association study of agronomic traits in flax. *BMC Genomics.* 2018;19:512.
6. Rozhmina T. et al. A comprehensive dataset of flax (*Linum usitatissimum* L.) phenotypes. *Data Brief.* 2021;37:107224.
7. Wang J., Zhang Z. GAPIT Version 3: Boosting power and accuracy for the genomic association and prediction. *Genom. Proteom. Bioinform.* 2021;19(4):629-640.
8. Zhou X., Stephens M. Genome-wide efficient mixed-model analysis for association studies. *Nat Genet.* 2012;44:821-824.