

**Genome-wide association study, haplotype analysis, and genomic prediction reveal the genetic basis of yield and yield-related traits in soybean (*Glycine max* L.)**

Javaid Akhter Bhat<sup>1,2†\*</sup>, Kehinde Adewole Adeboye<sup>3†</sup>, Showkat Ahmad Ganie<sup>4</sup>,  
Rutwik Barmukh<sup>5</sup>, Dezhou Hu<sup>1</sup>, Rajeev K. Varshney<sup>5,6\*</sup>, Deyue Yu<sup>1\*</sup>

<sup>1</sup>Soybean Research Institution, National Center for Soybean Improvement, State Key Laboratory of Crop Genetics and Germplasm Enhancement, Nanjing Agricultural University, Nanjing 210095, China

<sup>2</sup>International Genome Center, Jiangsu University, Zhenjiang-212013, China

<sup>3</sup>Department of Agricultural Technology, Ekiti State Polytechnic, Isan 371106, Nigeria

<sup>4</sup>Plant Molecular Science and Centre of Systems and Synthetic Biology, Department of Biological Sciences, Royal Holloway University of London, Egham, Surrey, TW20 0EX, UK

<sup>5</sup>Center of Excellence in Genomics & Systems Biology, International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Hyderabad, 502324, India

<sup>6</sup>Murdoch's Centre for Crop & Food Innovation, State Agricultural Biotechnology Centre, Food Futures Institute, Murdoch University, Murdoch, Western Australia, Australia

**\*Correspondence:**

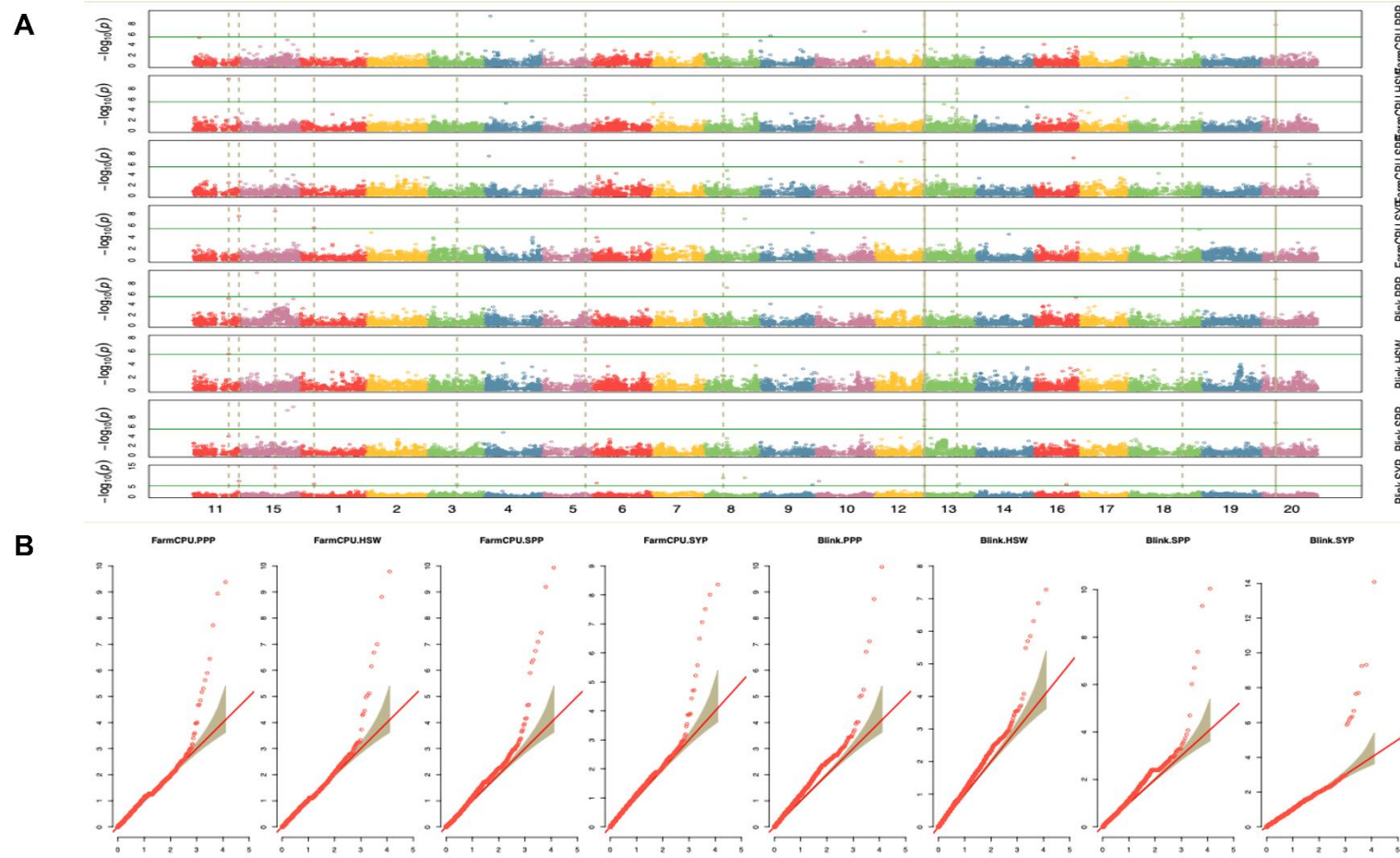
[javid.akhter69@gmail.com](mailto:javid.akhter69@gmail.com) (Javaid Akhter Bhat);

[rajeev.varshney@murdoch.edu.au](mailto:rajeev.varshney@murdoch.edu.au) (Rajeev K. Varshney);

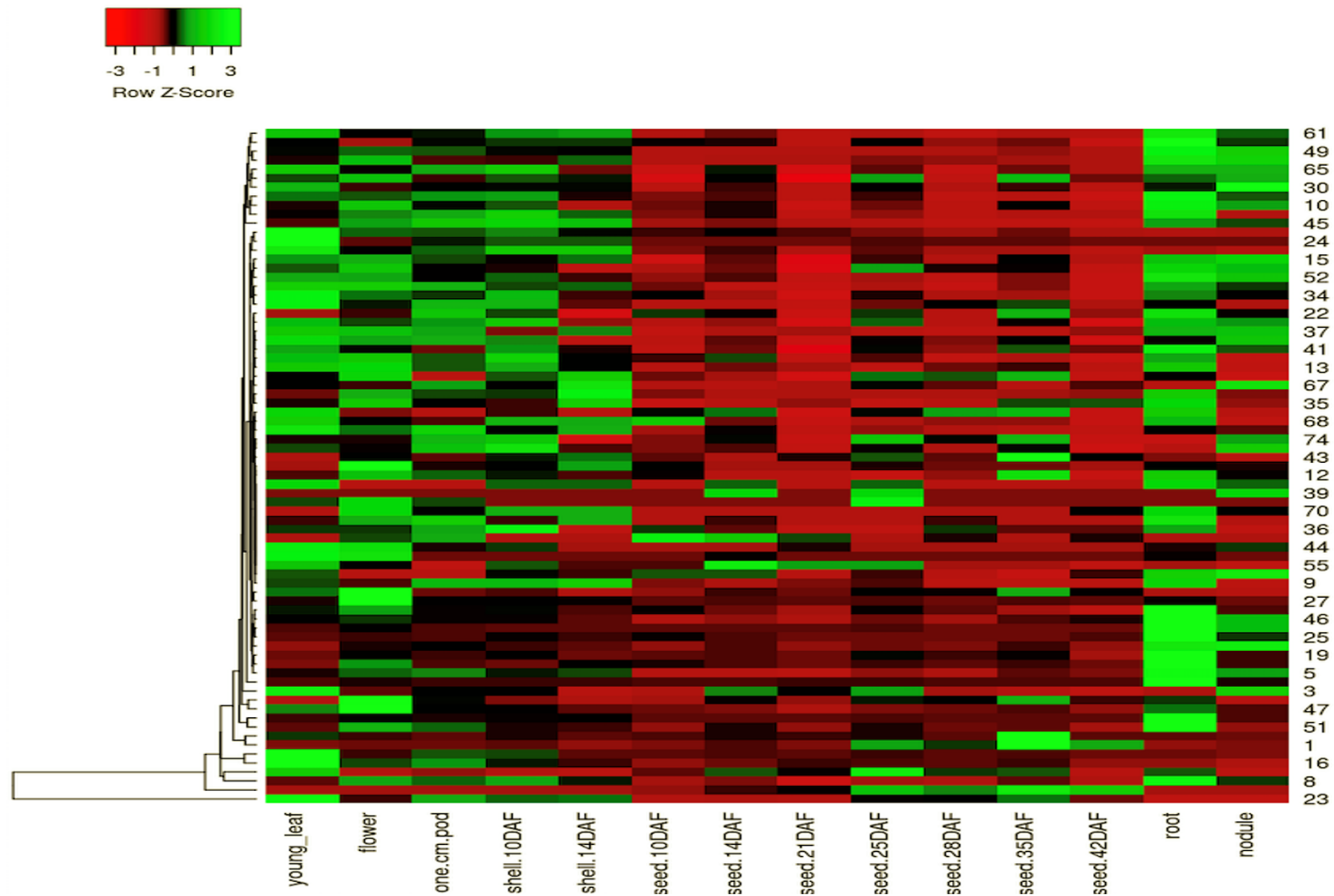
[dyyu@njau.edu.cn](mailto:dyyu@njau.edu.cn) (Deyue Yu)

<sup>†</sup>Equal first author

## Supplementary Figures



**Supplementary Figure S1:** Manhattan plot and Q-Q plot for traits with significant SNPs. (A) Manhattan plot for four traits viz., PPP, SPP, HSW, and SYP, identified using different GWAS models (FarmCPU and Blink). The green line indicates the significant threshold while the blue line shows the suggestive threshold. (B) Q-Q plot for GWAS results of four traits viz., PPP, SPP, HSW, and SYP, identified using different GWAS models (FarmCPU and Blink).



**Supplementary Figure S2:** Heat map showing the expression profiles of 75 candidate genes underlying six QTL regions (*qPPP-SPP4*, *qHSW5*, *qHSW-SPP11*, *qHSW-SPP13*, and *qHSW-PPP18*, and *qPPP-SPP20*), among different soybean tissues and development stages. Heat map was generated using the RNA-seq data downloaded from SoyBase database. Youngleaf—young leaf, One-cm.pod—1 cm of pod, PS—pod shell, DAF—days after flowering, and S—seed.