

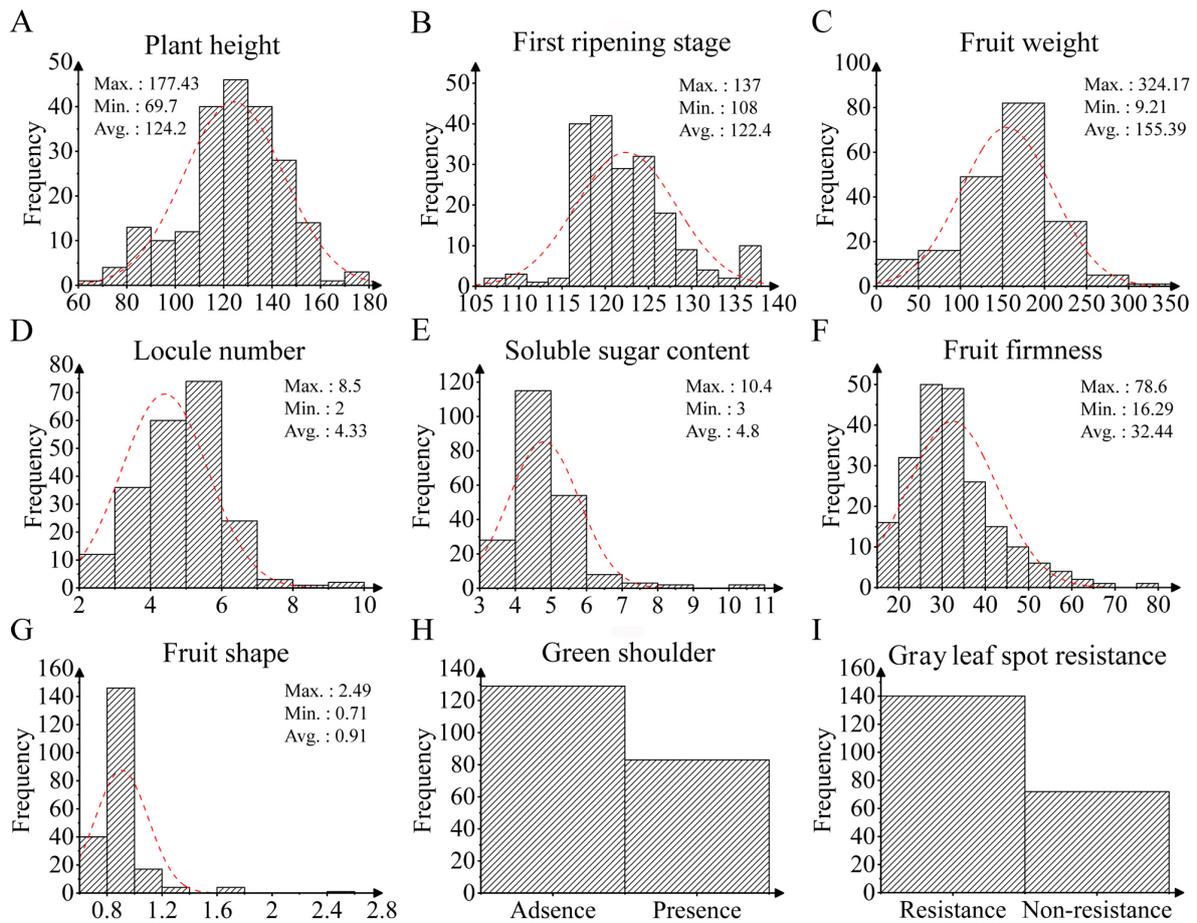
## ***Supplementary Materials***

### **GWAS Characterization of the Genetic Regions Associated with Nine Important Agronomic Traits in Tomato (*Solanum lycopersicum* L.)**

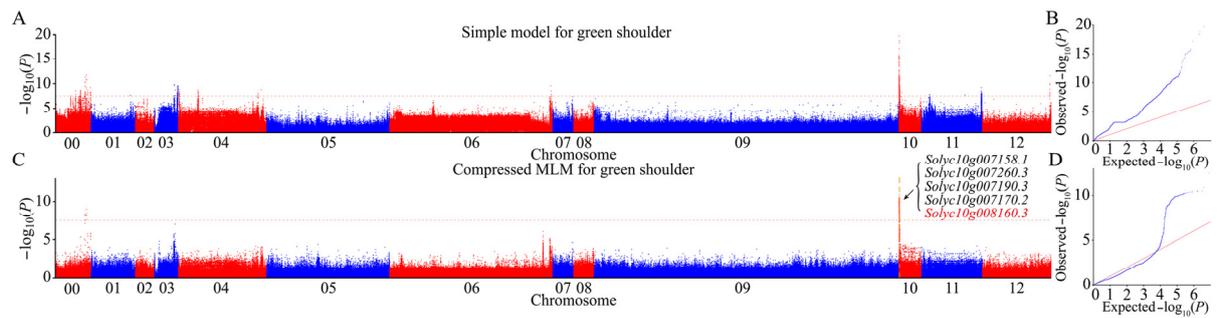
*Zengbing Liu*<sup>1</sup>, *Jingbin Jiang*<sup>1</sup> and *Jingfu Li*<sup>1\*</sup>

College of Horticulture and Landscape Architecture, Northeast Agricultural University, Harbin 150030, China.

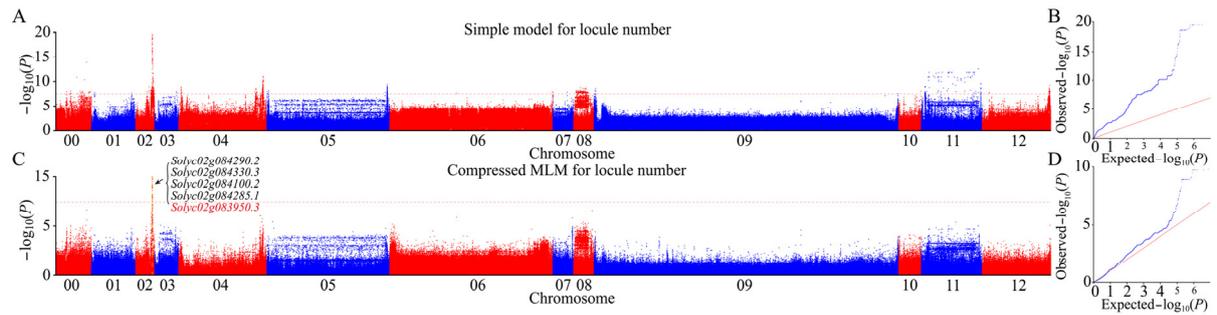
\* Correspondence: Jingfu Li, [LiJfNeau@gmail.com](mailto:LiJfNeau@gmail.com); Tel/Fax: +86-451-55190748



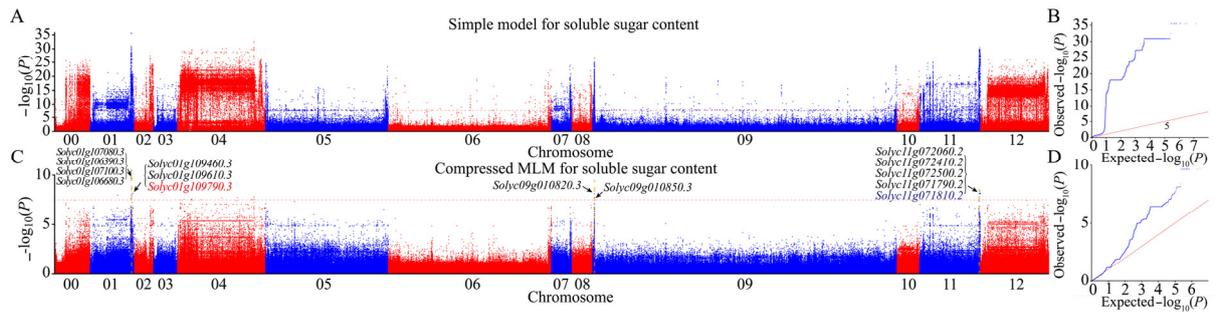
**Figure S1** Frequency distribution of variation of nice traits in 212 cultivated tomato accessions. The red dotted line represents the normal distribution fitted curve.



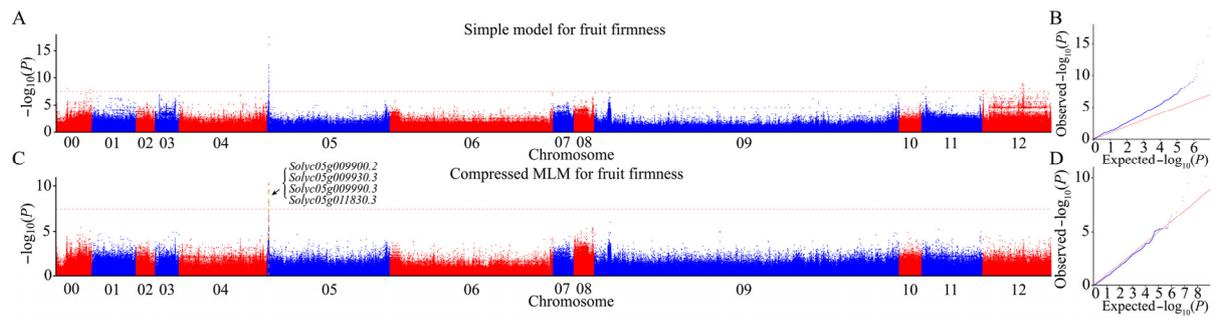
**Figure S2** Genome wide association studies for green shoulder. **(A)** Manhattan plots of the simple model. Chr00 represents unanchored scaffolds. Red horizontal dashed line indicates the genome-wide significance threshold. **(B)**  $Q$ - $Q$  plot of the simple model. **(C)** Manhattan plots of compressed MLM, as in A. **(D)**  $Q$ - $Q$  plot of compressed MLM. Gene number in red represent known gene in the peak region, while gene number in black represent top four unknown genes with a large phenotypic interpretation rate detected in the region. The yellow dotted line indicates the peak region where the labeled gene was located.



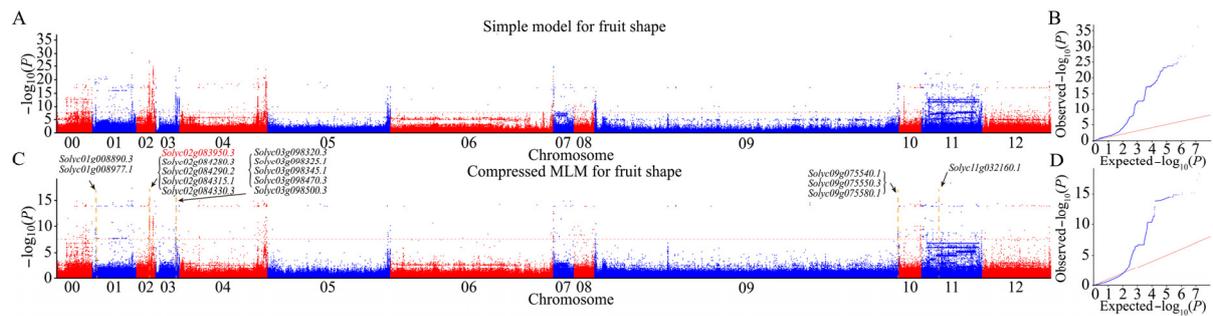
**Figure S3** Genome wide association studies for locule number. **(A)** Manhattan plots of the simple model. Chr00 represents unanchored scaffolds. Red horizontal dashed line indicates the genome-wide significance threshold. **(B)**  $Q$ - $Q$  plot of the simple model. **(C)** Manhattan plots of compressed MLM, as in A. **(D)**  $Q$ - $Q$  plot of compressed MLM. Gene number in red represent known genes in the peak region, while gene number in black represent top four unknown genes with a large phenotypic interpretation rate detected in the region. The yellow dotted line indicates the peak region where the labeled gene was located.



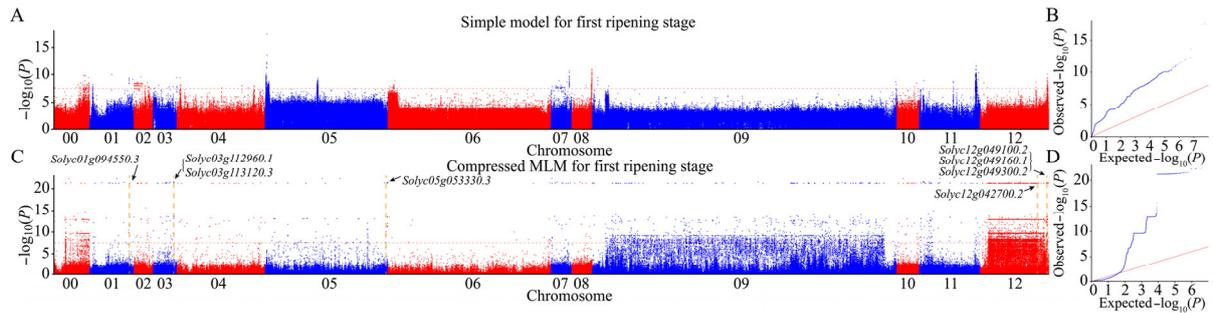
**Figure S4** Genome wide association studies for soluble sugar content. **(A)** Manhattan plots of the simple model. Chr00 represents unanchored scaffolds. Red horizontal dashed line indicates the genome-wide significance threshold. **(B)**  $Q-Q$  plot of the simple model. **(C)** Manhattan plots of compressed MLM, as in A. **(D)**  $Q-Q$  plot of compressed MLM. Gene numbers in red represent known genes in the peak region, while gene number in black represent top four unknown gene with a large phenotypic interpretation rate detected in the region. Gene number in blue font indicate gene known to have other functions. The yellow dotted line indicates the peak region where the labeled gene was located.



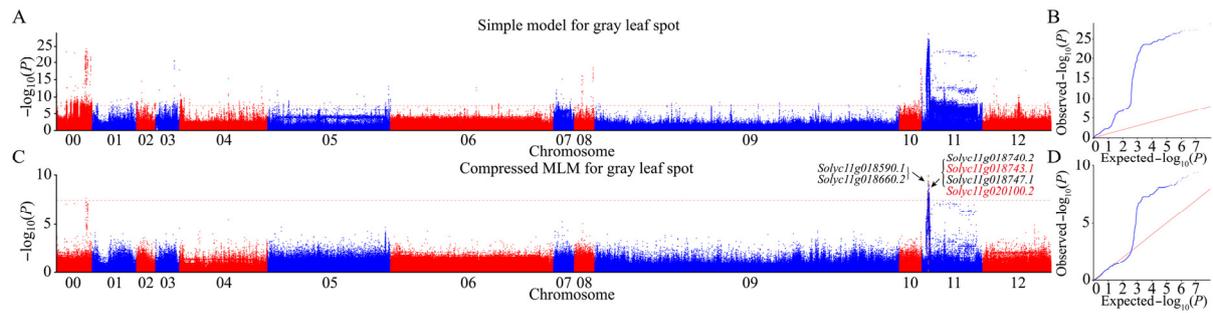
**Figure S5** Genome wide association studies for fruit firmness. **(A)** Manhattan plots of the simple model. Chr00 represents unanchored scaffolds. Red horizontal dashed line indicates the genome-wide significance threshold. **(B)**  $Q-Q$  plot of the simple model. **(C)** Manhattan plots of compressed MLM, as in A. **(D)**  $Q-Q$  plot of compressed MLM. Gene number in red represent known gene in the peak region, while gene number in black represent top four unknown genes with a large phenotypic interpretation rate detected in the region. The yellow dotted line indicates the peak region where the labeled gene was located.



**Figure S6** Genome wide association studies for fruit shape. **(A)** Manhattan plots of the simple model. Chr00 represents unanchored scaffolds. Red horizontal dashed line indicates the genome-wide significance threshold. **(B)**  $Q$ - $Q$  plot of the simple model. **(C)** Manhattan plots of compressed MLM, as in A. **(D)**  $Q$ - $Q$  plot of compressed MLM. Gene number in red represent known genes in the peak region, while gene number in black represent top four unknown genes with a large phenotypic interpretation rate detected in the region. The yellow dotted line indicates the peak region where the labeled gene was located.



**Figure S7** Genome wide association studies for first ripening stage. **(A)** Manhattan plots of the simple model. Chr00 represents unanchored scaffolds. Red horizontal dashed line indicates the genome-wide significance threshold. **(B)**  $Q$ - $Q$  plot of the simple model. **(C)** Manhattan plots of compressed MLM, as in A. **(D)**  $Q$ - $Q$  plot of compressed MLM. Gene number in red represent known gene in the peak region, while gene number in black represent top four unknown genes with a large phenotypic interpretation rate detected in the region. The yellow dotted line indicates the peak region where the labeled gene was located.



**Figure S8** Genome wide association studies for gray leaf spot. **(A)** Manhattan plots of the simple model. Chr00 represents unanchored scaffolds. Red horizontal dashed line indicates the genome-wide significance threshold. **(B)**  $Q-Q$  plot of the simple model. **(C)** Manhattan plots of compressed MLM, as in A. **(D)**  $Q-Q$  plot of compressed MLM. Gene number in red represent known gene in the peak region, while gene number in black represent top four unknown genes with a large phenotypic interpretation rate detected in the region. The yellow dotted line indicates the peak region where the labeled gene was located.