

Figure S1 Summary of the route of Upland cotton introduction to Xinjiang. YRR stands for Yellow River region in China, YtRR stands for Yangtze River region in China. The integrated Figure was summarized and modified from the published data (Bowman *et al.*, 2006; Abdullaev *et al.*, 2013; Huang *et al.*, 1996; Tian *et al.*, 2016).

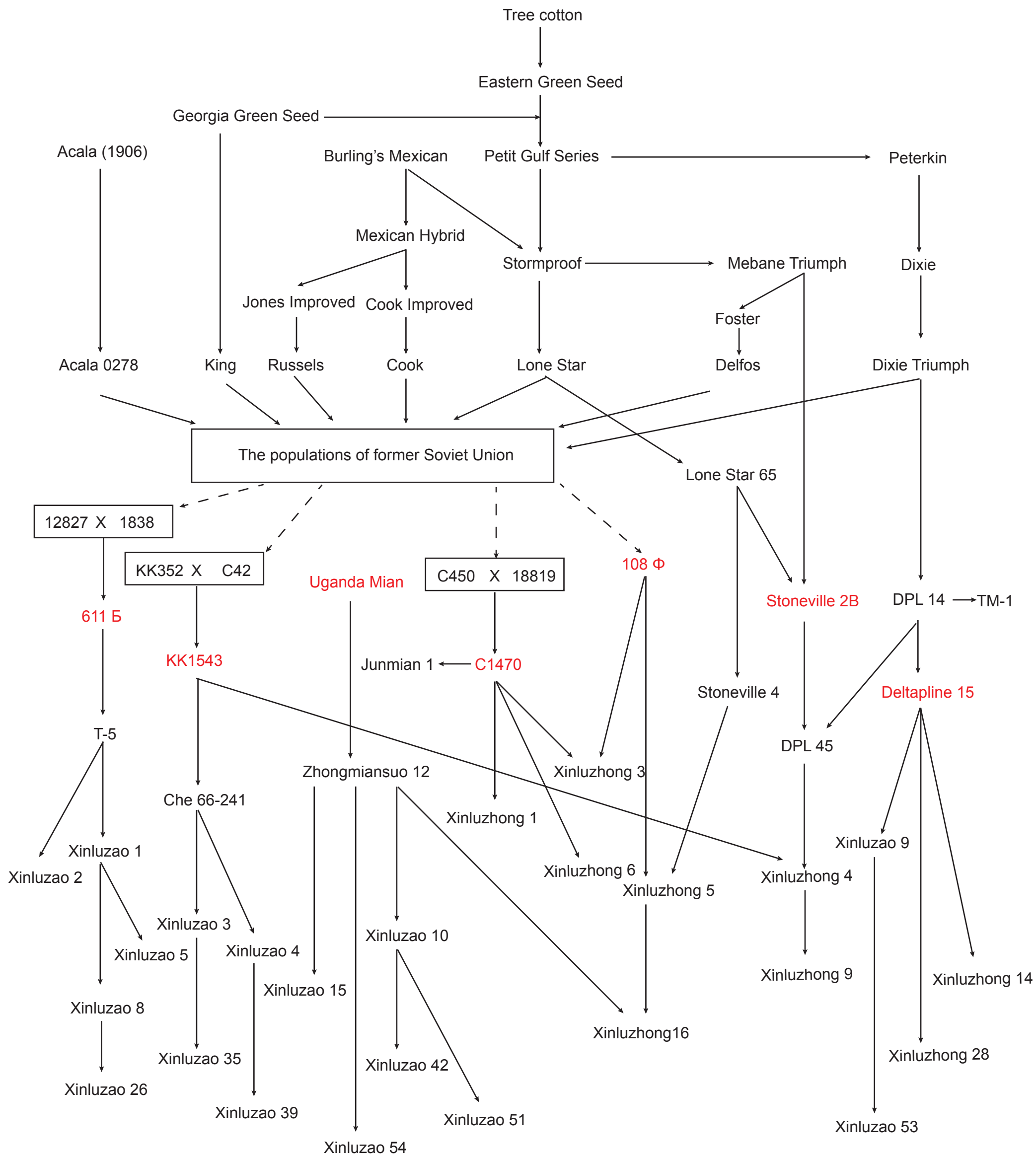


Figure S2 Partial pedigree information for Xinjiang Upland cotton breeding. The integrated figure was summarized and modified from the published data (Bowman *et al.*, 2006; Abdullaev *et al.*, 2013; Huang *et al.*, 1996; Tian *et al.*, 2016).

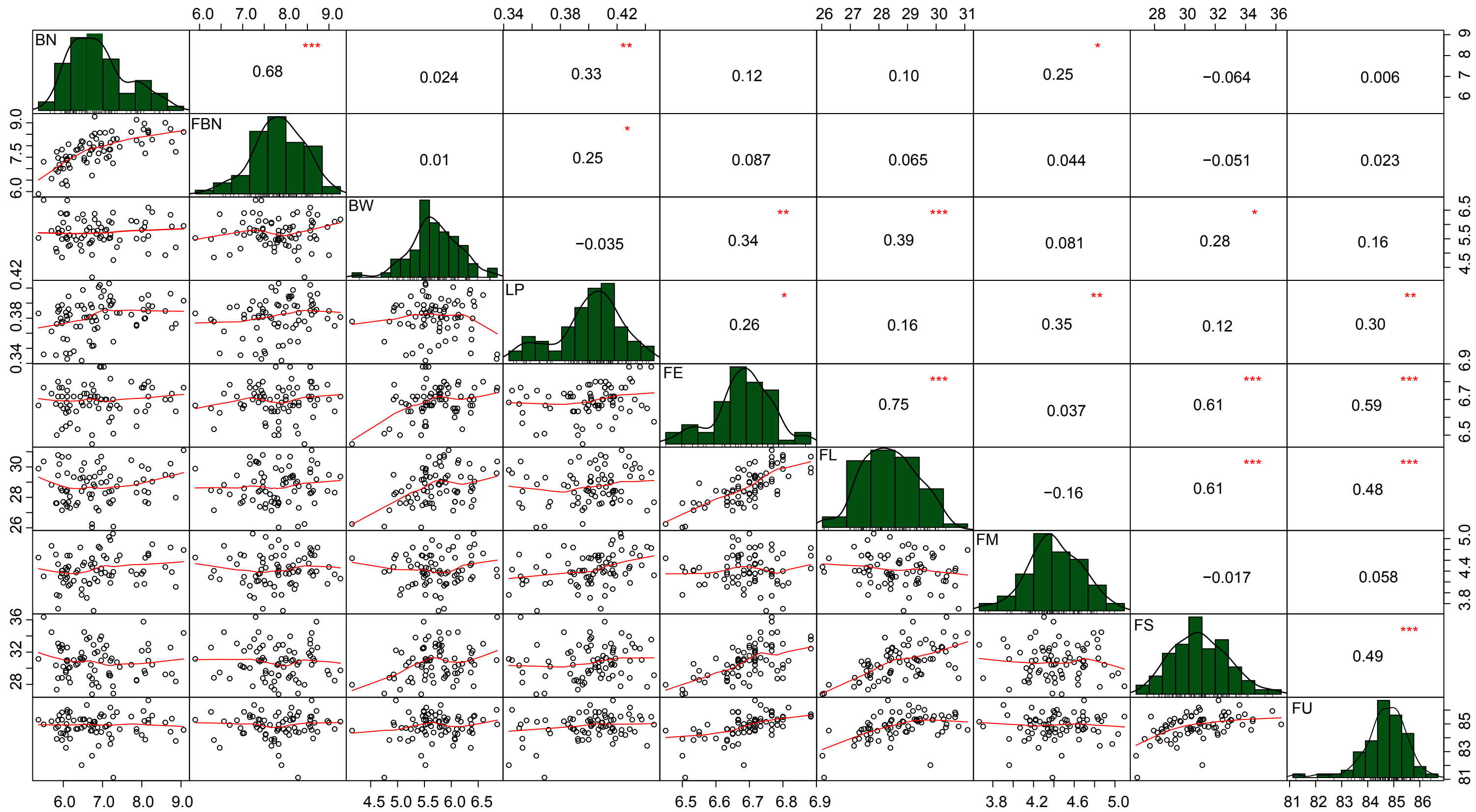


Figure S3 Frequency distribution of phenotypic variation of 9 yield and fiber quality traits and correlation coefficients among the traits in 76 accessions. BN: boll number, FBN: fruit branch number, BW: boll weight, LP: lint percentage, FE: fiber elongation, FL: fiber length, FM: fiber micronaire, FS: fiber strength, FU: fiber uniformity. ***indicates extremely significant difference at P = 0.001, ** indicates significant different at P = 0.01, * indicates different at P = 0.05.

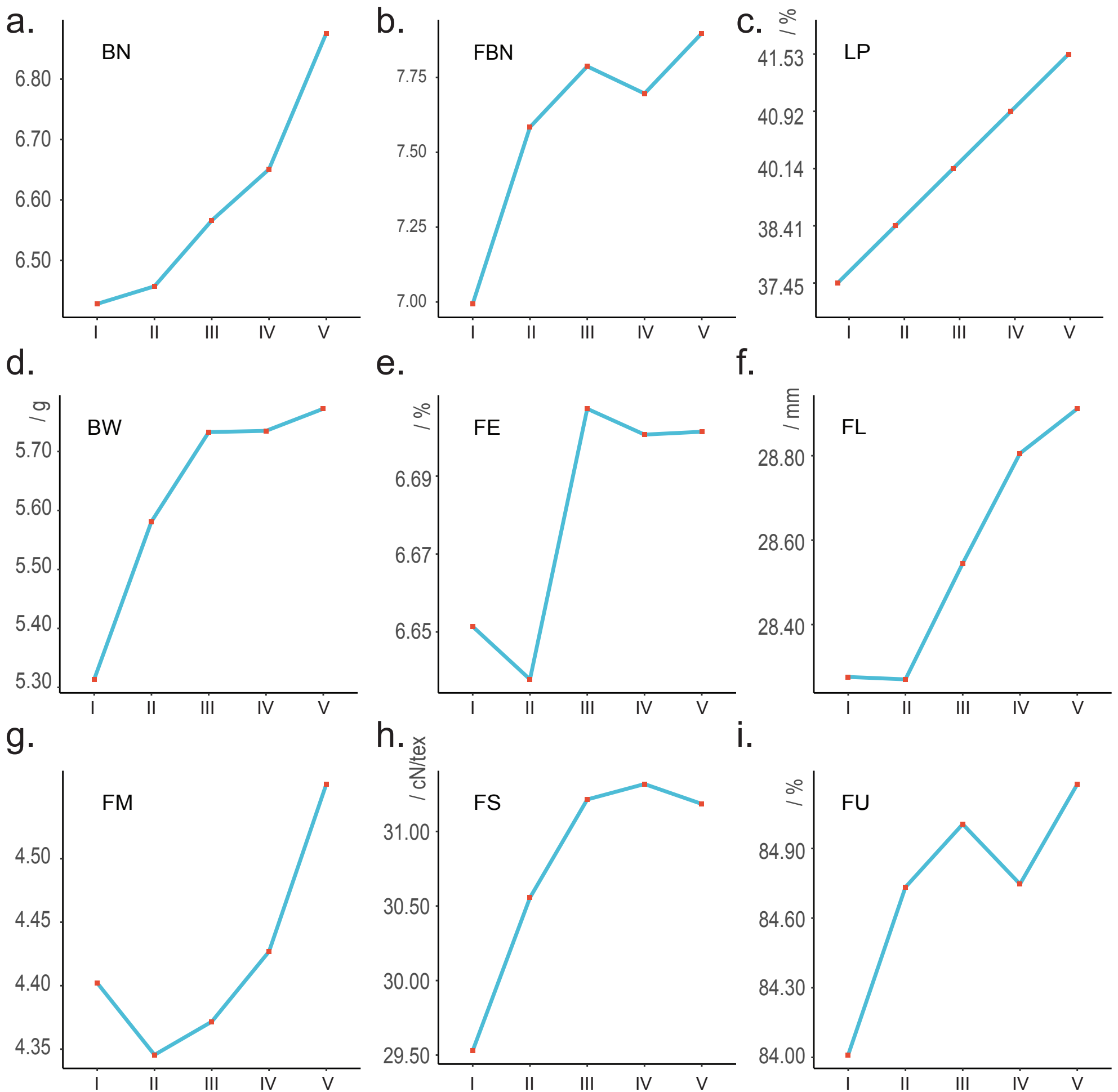


Figure S4 Trait improvement over the breeding process in Xinjiang. Improvement of nine traits comprising Boll number (a), fruit branch number (b), lint percentage (c), boll weight (d), fiber elongation (e), length (f), micronaire (g), strength (h) and uniformity (i) according to 5 breeding periods (I:1947-1956, II:1969-2000, III:2000-2005, IV:2006-2010, and V:2010-2013). The x-axis represents breeding periods from I to V, the y-axis the corresponding mean value of each trait.

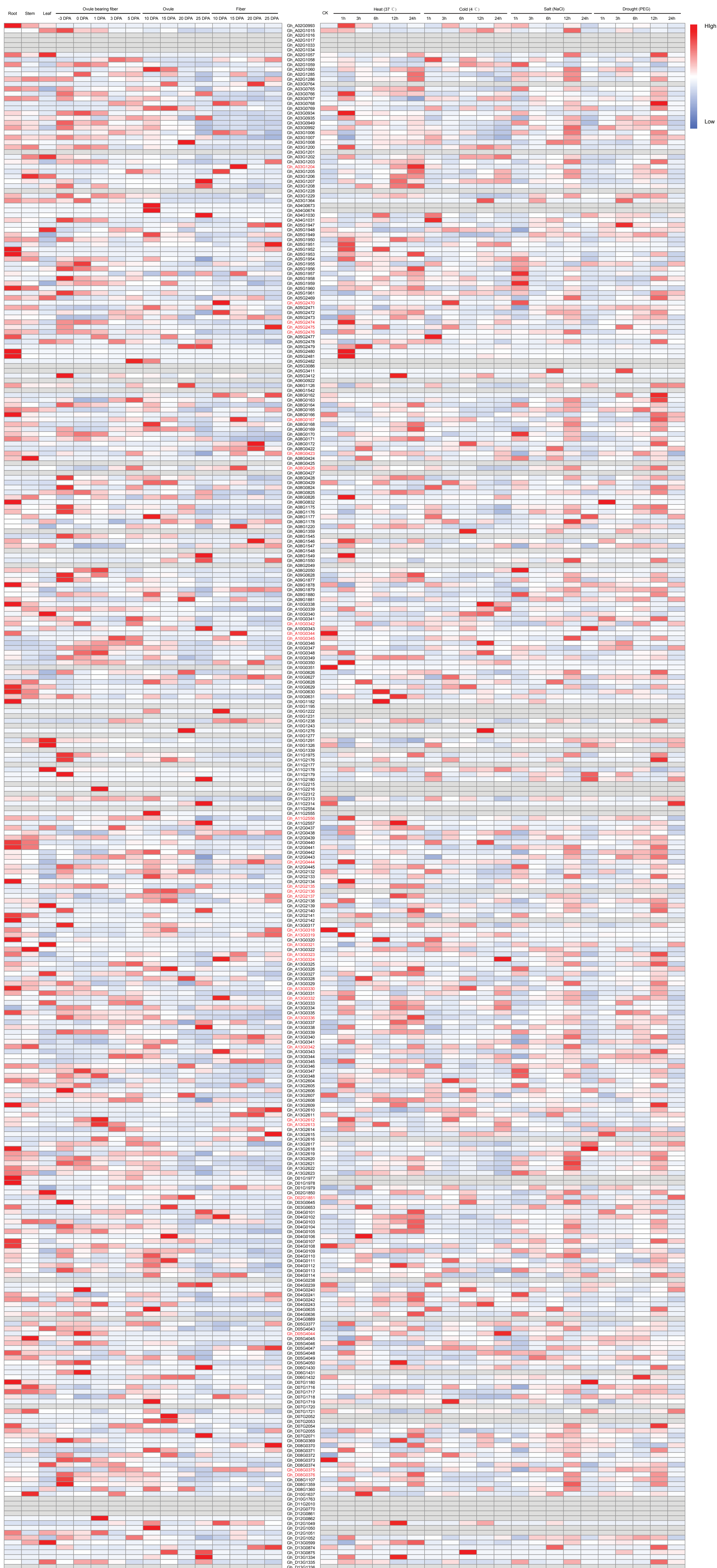


Figure S6 Heat map of transcriptomic patterns of 318 candidate genes in selection sweeps. Transcriptomic patterns of 318 genes under artificial selection in distinct tissues following stress treatment, based on the Z-score of fragments per kilobase of the exon model per million mapped reads (FPKM) in a single experiment, including root, stem, and leaf tissue during ovule- and fiber-development stages (left), and 1h, 3h, 6h, 12h and 24h post treatment by heat (37 °C), cold (4 °C), salt and drought. 31 genes with nonsynonymous SNPs are marked in red font. Blue to red represents the increase in expression.

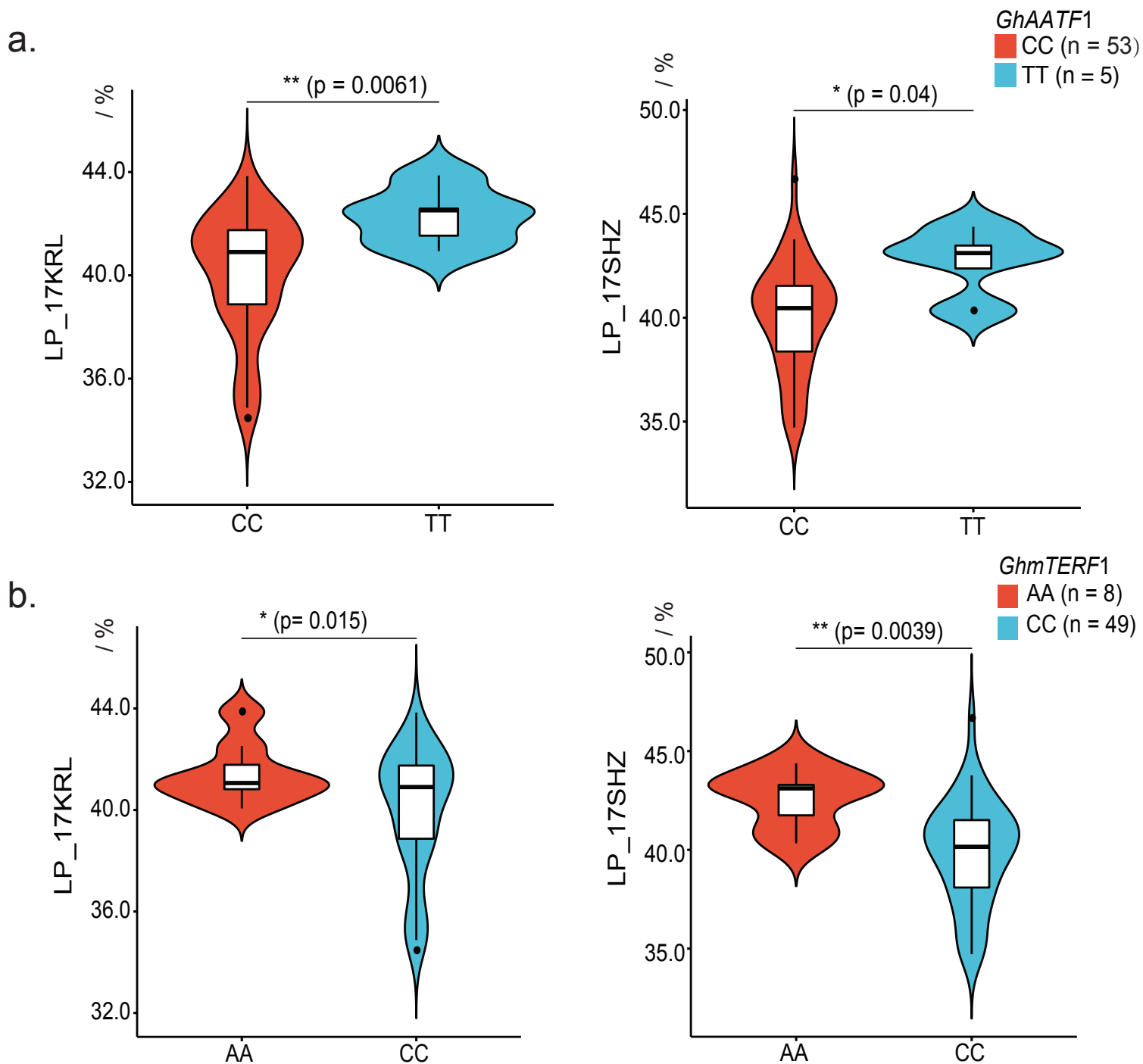


Figure S7 Box plot of *GhAATF1* and *GhmTERF1* related to lint percentages in two environments. Lint percentage analyses of *GhAATF1* accessions with CC and TT genotypes and *GhmTERF1* accessions with AA and CC genotypes in two environments. a. Box plot of lint percentage in 17KRL and 17SHZ of *GhAATF1*. b. Box plot of lint percentage in 17KRL and 17SHZ of *GhmTERF1*. Centre line, median; box limits, upper and lower quartiles; whiskers, 1.5x the interquartile range (* $P < 0.05$, ** $P < 0.01$, two-sided t-test).

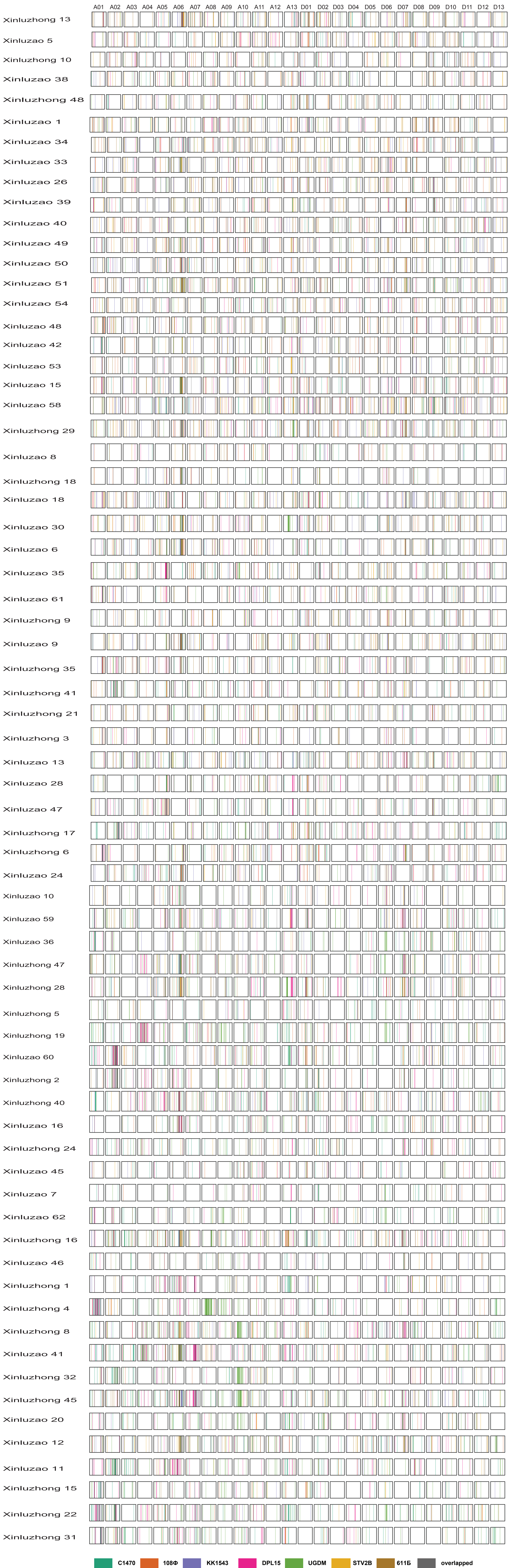


Figure S8 Identification of IBD segments inherited from foundation parents. Different colors represent different parents, and the corresponding color of each parent is noted under the images.

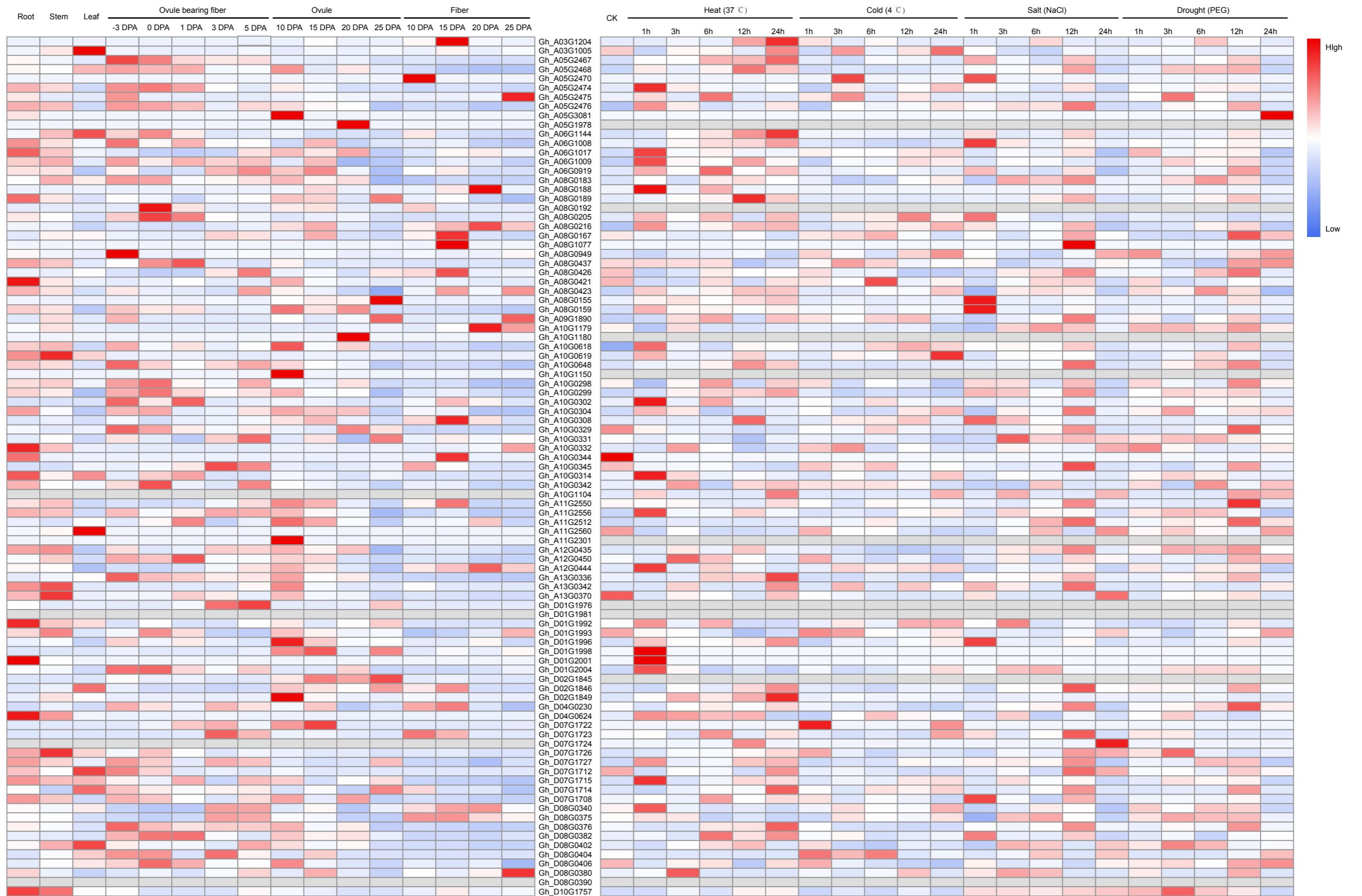


Figure S9 Heat map of transcriptomic patterns of 93 candidate nonsynonymous genes in IBD segments. Transcriptomic patterns of 93 genes in IBD segments with nonsynonymous SNPs in distinct tissues and following stress treatment, based on the Z-score of fragments per kilobase of the exon model per million mapped reads (FPKM) in a single experiment, including root, stem, and leaf tissues during ovule- and fiber-development stages (left), and 1h, 3h, 6h, 12h, 24h post treatment by heat (37 °C), cold (4 °C), salt and drought. Blue to red represents an increase in expression.

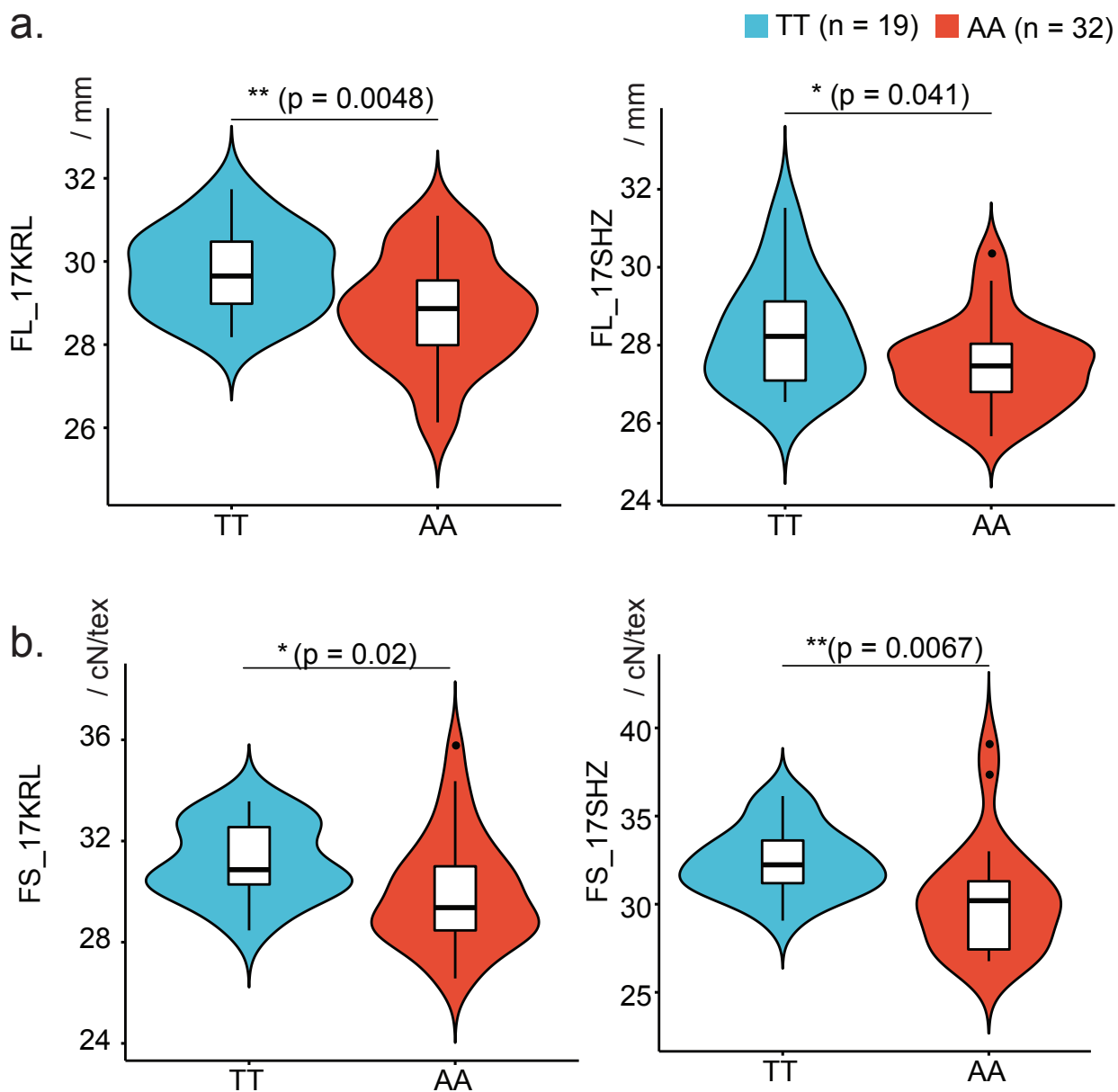


Figure S10 Box plot of fiber length and strength in two environments. Fiber length and fiber strength analyses of *GhARF3* accessions with AA and TT genotypes of in two environments. a. Box plot of fiber length in 17KRL and 17SHZ. b. Box plot of fiber strength in 17KRL and 17SHZ. Centre line, median; box limits, upper and lower quartiles; whiskers, 1.5x the interquartile range (*P < 0.05, **P < 0.01, two-sided t-test).