Genome-wide association mapping of agronomic and quality traits in foxtail millet

Yanhong Lou (yanhonglou1985@163.com)
Shandong Agricultural University https://orcid.org/0000-0002-3839-475X
Yun Chen
Hubei Normal University
Ying Chao
Taian Hi-tech industrial Zone
Fei Han
Shandong Agricultural University
Mingjie Sun
Shandong Agricultural University
Hong Pan
Shandong Agricultural University
Hui Wang
Shandong Agricultural University
Zhihao Liu
Hubei Normal University
Yuping Zhuge
Shandong Agricultural University

Research Article

Keywords: gene pyramiding, marker-assisted breeding, quantitative trait loci, Setaria italica

Posted Date: March 15th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1421666/v1

Abstract

Foxtail millet [*Setaria italica* L.] is a particularly momentous cereal and forage crop in arid and semi-arid regions, and was recognised worldwide for its potential as a health-promoting functional food. An understanding of the genomic variation and alleles underpinning its agronomic and quality traits is important for improving foxtail millet breeding. Therefore, high-quality single nucleotide polymorphisms (SNPs) to perform a genome-wide association study (GWAS) was employed to better understand the diversity of foxtail millet and facilitate the genetic dissection of its key traits. Using genotyping-by-sequencing, 107 foxtail millet accessions were sequenced; further analysis revealed 72,181 high-quality SNPs, 46 of which were significantly associated with 13 agronomic and quality traits. These SNPs were distributed across the 8 chromosomes; 36 were located in intergenic regions, whereas 1 and 5 SNPs were located in exon and intron regions, respectively. The GWAS revealed that 28 SNPs were associated with a single trait (Peak time). For some of the significant SNPs, favourable genotypes showed pyramiding effects for several traits. The 46 loci identified in present study will therefore be helpful for breeding plans aimed at foxtail millet improvement. Overall, the results provide sufficient fundamental for excellent cultivar breeding in the sustainable development of foxtail millet.

1. Introduction

Foxtail millet [*Setaria italica* L.], a perennial C4 plant belonging to the family Poaceae (subfamily Panicoideae, tribe Paniceae), is recognised as a particularly important cereal and forage crop in arid and semi-arid regions (Han et al. 2019). This ancient crop, originating from China, is grown worldwide owing to its high tolerance to drought and salinity (Werner et al. 2005). Foxtail millet is famous for both its nutritional and medicinal value, owing to the plentiful proteins, dietary fibres, phenolics, flavonoids, lysin, and minerals. It has also been recognised for its potential as a health-promoting functional food that helps reduce the risk of disease (Liu et al. 2017), reportedly by reducing blood glucose levels and controlling cholesterol in both healthy subjects and patients with diabetes (Narayanan et al. 2017). Despite its health benefits, the cultivation of foxtail millet has decreased significantly compared to other crops, due to its lower yield and reduced economic potential (Tian et al. 2018). However, increasing awareness of the importance of dietary diversification has led to increased market demand and new opportunities for this crop (Goron et al. 2015). Thus, new cultivars with improved nutritional value and higher yield, especially salt-resistant cultivars, are in demand.

Foxtail millet germplasm has great level of genetic diversity owing to their wide distribution and adaption to various agro-climatic conditions (Han et al. 2013). Agronomic traits, such as grain morphology, are the targets of breeding programme designed to improve both yield and quality, as they have important effects on crop productivity and responses to environmental stressors (Biselli et al. 2015). Understand fully the genetic basis of phenotypic variation within germplasm, especially for agronomic traits that are qualitatively important (such as salinity tolerance), has been a major focus in foxtail millet genetic studies. Notably, most agronomic traits are quantitative, making it difficult to unravel the genetic basis for phenotypes of interest, resulting to slow progress in conventional crop improvement. Literature indicated

that twenty-three putative quantitative trait loci (QTLs) were detected for ten of fourteen traits of rice (Xu et al. 2014), twenty-three molecular markers associated with earliness-related traits of cultivated tomato (Wang et al. 2020), forty significant marker-trait associations in maize (Chen et al. 2011). Linkage mapping, which identifies quantitative trait loci that are closely related to complex traits, has been successfully applied and widely utilised in plants. A complementary approach is linkage disequilibrium (LD) mapping or association mapping. There are two basic strategies for linkage disequilibrium mapping: i) re-sequencing of candidate genes and ii), exploiting of marker polymorphisms across all chromosomes via genome-wide association analysis (Pasam et al. 2012). Genome-wide association studies (GWAS) are widely employed to detect trait-associated loci or candidate genes associated with quantitative and complex traits, including those related to growth, development, stress tolerance, and nutritional quality (Kumar et al. 2015). Owing to its high resolution and cost-effectiveness, which are beneficial for gene discovery and molecular marker identification, GWAS have successfully identified numerous loci for complex traits in various plants, including wheat (Sukumaran et al. 2015), maize (Ding et al. 2015), and alfalfa (Biazzi et al. 2017). However, analyses of foxtail millet by GWAS are limited.

To better understand the diversity of foxtail millet, facilitate the genetic dissection of its agronomic and quality traits, and accelerate its marker-assisted breeding, high-quality single nucleotide polymorphisms (SNPs) were identified in present study by genotyping-by-sequencing of 107 foxtail millet accessions and, then, used for GWAS. Some of the significant single-nucleotide polymorphisms that we identified have potential pyramiding effects for several traits. These genotypes could be helpful for improving foxtail millet breeding programs.

2. Materials And Methods

2.1 Plant materials and experiment design

This study was conducted at the Mazhuang experimental station of Shandong Agricultural University, Tai'an, China, located at 36.02°N, 117.00°E, 85 m above sea level. This site has a subtropical warm and humid continental monsoon climate, average annual temperature is 13°C, annual average precipitation is 688.3 mm, sunshine duration is 2536.2 h, and average frost-free period is 172.9 days. Basic physicochemical properties of the topsoil (0–20 cm) are as follows: pH 7.24 (soil:water = 1:2.5, m/v); soil organic matter, 11.6 g kg^{-1}; total nitrogen, 0.75 g kg^{-1}; alkali-hydrolysable nitrogen, 72.3 mg kg^{-1}; available phosphorus, 21.3 mg kg^{-1}; and available potassium, 78.6 mg kg^{-1}. Field experiments were conducted between June and September 2017, employing 107 foxtail millet accessions (collected from main cultivation areas of China or provided by the Chinese Crop Germplasm Resources Center) as plant materials.

Compound fertiliser (N:P_2O_5:K_2O = 15:15:15) was applied at 1200 kg ha^{-1} as base fertilizer, and soil was tilled before sowing. Experimental plots of 40 m^2 were randomly arranged, and seeds were sown at 7.5 kg ha^{-1} on 28 June 2017. A prophylactic programme of herbicides was applied to control weed infestation.

No significant incidences of disease, pests, or weeds affected foxtail millet throughout its growth stage. The spacing between rows was 40 cm, and three replicates were employed for each accession. All plots were thinned on 10 July 2017. Harvesting was carried out on 28 September 2017.

2.2 Measurements

Data on agronomic traits (plant height, stem diameter, leaf length, leaf width, chlorophyll SPD value, spike length, and spike diameter) were collected in the maturation period (September 27th, 2017). Ten random plants from each treatment plot were selected, and the distances from the ground to the tips of the plants were measured and averaged per replication. Stem diameters were measured with a Vernier calliper and averaged per replication. The lengths and widths of flag leaves were measured and averaged as the leaf length per replication. Ten labelled leaves in each plot were selected to assess the chlorophyll SPD value by using a portable chlorophyll meter (Konica Minolta, Tokyo, Japan); measured five times per leaf were averaged for each replication. Ten random plants from each plot were hand-harvested to determine spike length, diameter, and weight; all spikes were harvested by hand to assess crop yield. Measurements were averaged per replication for each accession.

2.3 Quality traits

At maturity, all panicles were harvested, dried naturally, and dehulled. Quality traits, including kernel length, kernel width, and kernel length/width ratio, were determined as described by China National Standards (GB/T 17891 – 1999). The foxtail millet was ground for 3 min using stainless-steel grinder, and powder was acquired for chemical analysis. Amylose content was determined as detailled by the Chinese National standard method GB7648-87, with minor modifications: 10 mg of millet powder was transferred into a 14-mL capped tube, dispersed in 0.1 mL 95% ethanol, and treated with 0.9 mL 1 M NaOH for 16 h at 25 ^{o}C. Each sample was thoroughly mixed, and 10 μL of the formed supernatant was pipetted into 96-well plates; after adding 190 μL of freshly prepared I2-KI solution (3% iodine solution diluted 100 times in 0.01 M HCl before use) to each well, the plates were incubated for 10 min. Potato amylose (Sigma-Aldrich, St. Louis, MO, USA) was employed for standard curve establishment. Amylose content was determined based on the absorbance of each sample at 620 nm, and it was normalised based on sample weight.

Viscosity profiles were analysed by a Rapid Visco Analyser (Super 3; Newport Scientific, Warriewood, Australia) following the protocal of the American Association of Cereal Chemists. Data were recorded using the RVA-3D model Thermocline Windows Control 1.2 software (New Port Scientific, Sydney, Australia), and peak viscosity, trough viscosity, peak time, gelatinisation temperature, breakdown value, setback value, and consistency value were measured and expressed in centipoises (cp).

2.4 DNA isolation and sequencing

Fresh leaves were selected and obtained for genomic DNA extraction followed the protocol of cetyltrimethylammonium bromide method. Then, 500 ng of DNA per sample were used for the construction of double-digest restriction-associated DNA libraries, with slightly modifications. Briefly,

genomic DNA was digested with the restriction enzymes Hind*III* and Bfa*I* at 37°C for 5 h, followed by a ligation step in which each sample was assigned 1 of 24 unique adaptors. Pooled digests of 24 individuals were run on agarose gel; fragments ranging in size from 220 bp to 450 bp were manually excised and purified using a Zymoclean Gel DNA recovery kit (Zymo Research, Irvine, CA, USA). Each pool was amplified using 14 PCR cycles in 25-μL reactions containing 5 μL 5× reaction buffer, 5 μL 5× high GC enhancer, 0.25 μL Q5 polymerase, 4 μL library DNA, and a unique indexing primer for each pool that corresponds with the standard Illumina (Illumina Inc, San Diego, CA, USA) multiplexed sequencing protocol. The PCRs were carried out in a Veriti 96-well thermal cycler (Life Technologies, Carlsbad, CA, USA) using the following profile: initial denaturation at 98°C for 30 s, 14 cycles at 98°C for 15 s, 65°C for 30 s, and 72°C for 30 s, followed by a final extension step at 72°C for 5 min. The DNA libraries were quantified using a high-sensitivity DNA analysis kit (No. 5067 − 4626; Agilent Technologies, Santa Clara, CA, USA) in a 2100 Bioanalyzer (Agilent Technologies). Pools were combined in equimolar concentrations to form a single genomic library and were sequenced in one lane of a HiSeq 2500 Illumina sequencer (pair-end, 2×150 bp).

2.5 Sequencing quality check and filtering

Adapter sequences were removed using AdapterRemoval 2. Reads with Phred scores < 20 (average on sliding window), with incorrect restriction sites, and reads of < 50 bp were removed.

2.6 Sequence alignment and SNPs detection

Illumina paired-end reads were aligned to the foxtail millet reference sequence using BWA-MEM, with default parameters, and SNPs were called using GATK (UnifiedGenotyper, stand_call_conf 30, stand_emit_conf 10) with subsequent filtering based on read map quality score (≥ 20), base quality score (≥ 5), and read depth (≥ 3).

2.7 Annotation of genetic variants

Variants were annotated using ANNOVAR 2016-02-01 with gene-based annotation to assess whether SNPs or indels caused protein-coding changes and to identify which amino acids were affected.

2.8 Population structure and linkage disequilibrium analyses

We conducted both PCA and ancestry analysis to evaluate genetic structure using Plink 1.90 beta and ADMIXTURE 1.3 (http://www.genetics.ucla.edu /software/admixture/), respectively. An individual-based NJ tree was constructed based on p-distance using TreeBest 1.92 (http://treesoft.sourceforge.net/treebest.shtml).

2.9 Genome-wide association study

In total, 107 accessions were used in GWAS for different traits. Association analysis was conducted with the genome-wide efficient mixed-model association (GEMMA) software package. For mixed-linear-model analysis, the following equation was used:

$$y = Xa + S\beta + K\mu + e; \quad (1)$$

where, y represents phenotypes; a and β are fixed effects representing marker effects and non-marker effects, respectively; and μ represents unknown random effects. X, S, and K are the incidence matrices for a, β, and μ, respectively.

For each significant SNP in the GWAS results, accessions with phenotypic value were classified into three groups according to their genotypes (reference homozygosis, heterozygosis, and altered homozygosis). Pairwise comparisons of phenotypic values between genotype groups were performed using *t*-tests, and significant SNPs were filtered according to the *t*-test results (at least one *p*-value ≤ 0.05).

3. Results

3.1 Double-digest restriction-associated DNA sequencing and variation detection

The double-digest restriction-associated DNA (ddRAD) sequencing carried out to genotype the 107 millet accessions generated 187 Gb of data with reads of 150 bp in length, on average. After quality control, 169 Gb of high-quality sequences were obtained and, then, mapped to the reference genome sequence of foxtail millet (NCBI; GCF_000263155.2_v2.0). The mapping rates of the 107 accessions varied between 78.75% and 100.00% (Supplementary file S1), and, then, these data were used for SNP calling by GATK. Initially, 506,788 SNPs were called for the 107 accessions, and 72,181 high-quality SNPs [coverage depth ≥ 3, mapping quality ≥ 20, missing ratio of samples within population ≤ 20%, and multiple allele frequency (MAF) ≥ 0.05] were retained for subsequent analyses. Of these, 71,947 SNPs were spread across the nine chromosomes, with NC_028457.1 containing the most SNPs (13,010) and NC_028450.1 containing the least SNPs (4,854) (Fig. 1), while the remaining 234 SNPs were scattered across 31 scaffolds.

3.2 Population structure and linkage disequilibrium

To correct the GWAS model for population structure, we performed three analyses to determine the relationships between the 107 accessions. These were classified into four groups (G1, G2, G3, and G4) according to the results of the principal components analysis (PCA), neighbour-joining (NJ)-tree analysis, and ancestry structure analysis (Fig. 2). Groups G1, G2, G3, and G4 included 36, 14, 27, and 30 accessions, respectively. Whereas G1 primarily included accessions from northern China, G2 and G3 mainly contained accessions from northern and eastern China, respectively, and G4 included accessions distributed from western to northeast China (Fig. 2a). To determine the mapping resolution of the GWAS, the LD of the 107 accessions was analysed (Supplementary file S1), and the LD decay rate was estimated as 100 kb (r^2 = 0.2).

3.3 Genome-wide association analysis

Based on the 72,181 high-quality SNPs obtained, we performed an association analysis for 14 traits, most of which presented unimodal distributions (Fig. 3 and Supplementary file S2). In total, 68 SNPs were significantly associated with 13 traits (none SNP were significantly associated with Leaf width) after Bonferroni correction (Fig. 4), and 46 significant SNPs were retained after filtering (Supplementary file S3). These SNPs were distributed across eight chromosomes (no significant SNP was detected in NC_028454.1). Of these SNPs, 36 were located in intergenic regions, whereas one and five SNPs were located in exon and intron regions, respectively. Up to 28 SNPs were associated with a single trait (Peak time). The 318 nonredundant genes in the LD decay regions around each significant SNP were associated with the 13 traits (Supplementary file S4). The number of genes associated with the different traits ranged from 2 to 47.

A favourable genotype is defined as one in which a significant SNP leads to an increase in phenotypic value (Supplementary file S5). Nine traits were associated with more than one significant SNP. To assess the potential pyramiding effects between favourable genotypes for traits, the mean phenotypic values of the accessions that contained multiple favourable genotypes were analysed. As shown in Fig. 5, the correlation between the number of favourable genotypes and phenotypic value was > 0.4 for three traits (amylose content, spike length, and trough viscosit). These findings suggested a certain degree of pyramiding effect between the favourable genotypes and these three traits.

3.4 Frequencies of favourable genotypes in accession groups

The frequencies of favourable genotypes at different significant SNPs were calculated for G1, G2, G3, and G4. The favourable genotypes of SNP locus NC_028452.1::13,868,887 (the 13,868,887th base of chromosome NC_028452.1) were all distributed in G2 (Fig. 6 and Supplementary file S6). Favourable genotypes involving 13, 8, and 25 SNP loci were obtained from two, three, and four different accession groups, respectively. These results suggested that accessions in the different groups might have had different evolutionary or domestication directions.

4. Discussion

Well understanding the genetic variation of foxtail millet germplasm is pre-requisite for conservation and utilization, and progress in crop breeding relys on the extent of genetic variability. Indeed, genetic diversity is recognized as the essential necessity that provide new genes for yield, acclimatization, stress resistance, and CV usually employed as an effective indicator for the interest of phenotypic variants selection for breeding purpose. For the core germplasm of 107 foxtail millet accessions, greatest CV (50.54%) was detected in breakdown value, and was much higher than barely (with 16.8%) (Cozzolino et al. 2013), and was similar to rice (46.0%) (Pang et al. 2016). To our best known, the extent of breakdown is refelected the stability of the paste during cooking and is characteristics of the fragility of the starch

granules (Han et al. 2001). Meanwhile, the CV of chlorophyll SPD value was 29.02%, and the CVs of plant height, spike length, peak viscosity, and trough viscosity were exceeded 15.0%. However, low CVs of peak time, kernel length, amylose content, and leaf length were investigated that below 10%. These results were consisted with wheat, which the Cvs of kernel length and kernel width were 1.78% and 2.30% (Alemu et al. 2020), indicted the kernel size was low efficiency to access genetic variability. Overall, the difference accessions variation was attributed to the complex heredity of morphological and quality traits or their diverse geographical origins (Lou et al. 2015).

Genome-wide association analysis is an effective method for identifying trait-associated loci in natural populations, which have become more and more mature, and widely used in excellent rice (Yang et al. 2019), and maize (Liang et al. 2020) gene mining gradually. To date, only a few SNP-based association analyses in foxtail millet have been reported, with one SNP-based association analysis performed using low-depth (0.8×) re-sequencing (Jia et al. 2013). In the current study, we sequenced 107 millet accessions through ddRAD sequencing, with an average coverage of 3.58×, and identified 46 SNP loci associated with 13 traits in the subsequent GWAS. These results provided the key theoretical foundations for new cultivar breeding of faoxtail millet. Meanwhile, the accessions used in this study were collected primarily from China, with two accessions from the United States. Population structure analysis indicated that these cultivars could be classified into four groups, although the obtained division did not correlate with the geographic origin of the accessions. Similar findings have been reported for other plant species, e.g. *Brassica rapa* L. (Tanhuanpää et al. 2016) and *Matricaria chamomilla* L. (Pirkhezri et al. 2010). This discrepancy might be explained by the transport routes and exchange of plant materials between regions.

Moderate LD decay is important for association analysis, ranging over several hundred kilobases between plants such as rice, soybean (Gupta et al. 2005), and cotton (Ma et al. 2018). The LD decay in the present study was estimated as 100 kb when r^2 decreased to 0.2, which is consistent with previous findings (Zhang et al. 2014). This LD decay will be useful for identifying unknown genes that are linked to significant SNPs. We employed a mixed linear model in the present association analysis; where, the remarkable threshold was evaluated as $P = 10^{-4.86}$ after the Bonferroni correction (1/72,181). The *t*-tests used to analyse differences in phenotypic values of different genotypes at each significant SNP locus identified 46 loci associated with 13 traits and established favourable genotypes for each locus. Marker-based gene pyramiding strategies have been demonstrated in several studies (Zhang et al. 2014; Sacco et al. 2013), and the favourable genotypes obtained in the present study also showed dosage pyramiding effects for several traits. The favourable genotypes identified here have substantial potential for future breeding programs, whereas the different frequencies of favourable genotypes in the different accession groups may imply that these accessions have had different evolutionary or domestication directions.

5. Conclusion

We identified a substantial number of SNP markers in foxtail millet and performed a GWAS to identify trait-associated loci. For a portion of the significant SNPs, favourable genotypes showed pyramiding

effects for several traits. These favourable genotypes are expected to be useful for breeding programs aimed at foxtail millet improvement.

Declarations

Acknowledgments: We thanks the Chinese Crop Germplasm Resources Centre of the Chinese Academy of Agricultural Sciences for providing foxtail millet accessions. This study was financially supported by the Modern Agriculture Industrial Technology Systems Project of Shandong Province (SDAIT-15-04), the Postdoctoral Innovation Project of Shandong Province (201603052).

Authors' contributions: Y.ZG**.** and Z.L. designed the experiment; Y.L. wrote the manuscript; Y.C. revised the manuscript; Z.L. provided cDNA libraries and generated Illumina libraries. H.W. and Y.C. performed statistical analysis. F.H., H.P., and M.S. conducted the experiment.

Conflict of interest: The author(s) declare that they have no competing interests.

Code of availability: Not applicable.

References

Alemu A, Feyissa T, Tuberosa R, Maccaferri M, Sciara G, Letta T, Abeyo B (2020) Genome-wide association mapping for grain shape and color traits in Ethiopian durum wheat (*Triticum turgidum* ssp. *durum*). Crop J https://doi.org/10.1016/j.cj.2020.01.001.

Biazzi E, Nazzicari N, Pecetti L, Brummer EC, Palmonari A, Tava A, Annicchiarico P (2017) Genome-wide association mapping and genomic selection for alfalfa (*Medicago sativa*) forage quality traits. PLoS One 12: e0169234.

Biselli C, Bagnaresi P, Cavalluzzo D, Urso S, Desiderio F, Orasen G, Gianinetti A, Righettini F, Gennaro M, Perrini R, Ben-Hassen M, Sacchi GA, Cattivelli L, Valè G (2015) Deep sequencing transcriptional fingerprinting of rice kernels for dissecting grain quality traits. BMC Genomics 16: 1091.

Chen JT, Hu LZ, Zhu LY, Guo JJ, Zhao YF, Huang YQ (2011) Diversity, structure, and marker-trait association analysis of the maize recombinant inbred line population. Agr Sci China 10: 975-986.

Cozzolino D, Roumeliotis S, Eglinton J (2013) Relationships between starch pasting properties, free fatty acids and amylose content in barely. Food Res Int 51: 444-449.

Ding JQ, Ali F, Chen GS, Li HH, Mahuku G, Yang N, Narro L, Magorokosho C, Makumbi D, Yan J (2015) Genome-wide association mapping reveals novel sources of resistance to northern corn leaf blight in maize. BMC Plant Biol 15: 206.

Goron TL, Raizada MN (2015) Genetic diversity and genomic resources available for the small millet crops to accelerate a New Green Revolution. Front Plant Sci 24: 157.

Gupta PK, Rustqi S, Kulwal PL (2005) Linkage disequilibrium and association studies in higher plants: present status and future prospects. Plant Mol Biol 57: 461-485.

Han B, Huang XH (2013) Sequencing-based genome-wide association study in rice. Curr Opin Plant Bio 16: 133-138.

Han F, Sun MJ, He W, Cui XM, Pan H, Wang H, Song FP, Lou YH, Zhuge YP (2019) Ameliorating effects of exogenous Ca^{2+} on foxtail millet seedlings under salt stress. Funct Plant Biol 46: 407-416.

Han XZ, Hamaker BR (2001) Amylopectin fine structure and rice starch paste breakdown. J Cereal Sci 34: 279-284.

Jia GQ, Huang XH, Zhi H, Zhao Y, Zhao Q, Li WJ, Chai Y, Yang LF, Liu KY, Lu HY, Zhu CR, Lu YQ, Zhou CC, Fan DL, Weng QJ, Guo YL, Huang T, Zhang L, Lu TT, Feng Q, Hao HF, Liu HK, Lu P, Zhang N, Li YH, Guo EH, Wang SJ, Wang SY, Liu JR, Zhang WF, Chen GQ, Zhang BJ, Li W, Wang YF, Li HQ, Zhao BH, Li JY, Diao XM, Han B (2013) A haplotype map of genomic variations and genome-wide association studies of agronomic traits in foxtail millet (*Setaria italica*). Nat Genet 45: 957-961.

Kumar V, Singh A, Mithra SVA, Krishnamurthy L, Parida SW, Jain S, Tiwari KK, Kumar P, Rao AR, Sharma SK, Khurana JP, Singh NK, Mohapatra T (2015) Genome-wide association mapping of salinity tolerance in rice (*Oryza sativa*). DNA Res 22: 133-145.

Liang ZK, Qiu YM, Schnablem JC (2020) Genome-phenome wide association in maize and Arabidopsis identifies a common molecular and evolutionary signature. Mol Plant doi: https://doi.org/10.1016/j.molp.2020.03.003.

Liu KG, Qi SH, Li D, Jin CY, Gaom CH, Duan SW, Feng B, Chen M (2017) TRANSPARENT TESTA GLABRA 1 ubiquitously regulates plant growth and development from *Arabidopsis* to foxtail millet (*Setaria italica*). Plant Sci 254: 60-69.

Lou YH, Hu LX, Chen L, Sun XY, Liu HM, Xu QG (2015) Association analysis of simple sequence repeat (SSR) markers with agronomic traits in Tall Fescue (*Festuca arundinacea* Schreb.). Plos One 10: e0133054.

Ma ZY, He SP, Wang XF, Sun JL, Zhang Y, Zhang GY, Wu LQ, Li ZK, Liu ZH, Sun GF, Yan YY, Jia YH, Yang J, Pan ZE, Gu QS, Li XY, Sun ZW, Dai PH, Liu ZW, Gong WF, Wu JH, Wang M, Liu HW, Feng KW, Ke HF, Wang JD, Lan HY, Wang GN, Peng J, Wang N, Wang LR, Pang BY, Peng Z, Li RQ, Tian SL, Du XM (2018) Resequencing a core collection of upland cotton identifies genomic variation and loci influencing fiber quality and yield. Nat Genet 50: 803-813.

Narayanan J, Sanjeevi V, Rohini U, Trueman P, Viswanathan V (2017) Postprandial glycaemic response of foxtail millet dosa in comparison to a rice dosa in patients with type 2 diabetes. Indian J Med Res 144: 712-717.

Pang YL, Ali JH, Wang XQ, Franje NJ, Revilleza EJ, Xu JL, Li ZK (2016) Relationship of rice grain amylose, gelatinization temperature and pasting properties for breeding better eating and cooking quality of rice varieties. PLoS ONE 11: e0168483.

Pasam RK, Sharma R, Malosetti M, van Eeuwijk FA, Gaseneyer G, Kilian B, Graner A (2012) Genome-wide association studies for agronomical traits in a worldwide spring barley collection. BMC Plant Biol 12: 16.

Pirkhezri M, Hassani ME, Hadian J (2010) Genetic diversity in different populations of *Matricaria chamomilla* L. growing in southwest of Iran, based on morphological and RAPD markers. J Med Plants Res 4: 1-13.

Sacco A, Di Matteo A, Lombardi N, Trotta N, Punzo B, Mari A, Barone A (2013) Quantitative trait loci pyramiding for fruit quality traits in tomato. Mol Breed 31: 217-222.

Sukumaran S, Dreisigacker S, Lopes M, Chavez P, Reynolds MP (2015) Genome-wide association study for grain yield and related traits in an elite spring wheat population grown in temperate irrigated environments. Theor Appl Genet 128: 353-363.

Tanhuanpää P, Erkkilä M, Tenhola-Roininen T, Tanskanen JA, Manninen O (2016) SNP diversity within and among Brassica rapa accessions reveals no geographic differentiation. Genome 59: 1-21.

Tian BH, Luan SR, Zhang LX, Liu YL, Zhang L, Li HJ (2018) Penalties in yield and yield associated traits caused by stem lodging at different developmental stages in summer and spring foxtail millet cultivars. Field Crops Res 217: 104-112.

Wang T, Zhang ZJ, Zhu H, Zhang YM, Gao W, Wang XF, Piao ZY, Zou QD (2020) Phenotypic diversity and genome-wide association mapping of earliness-related traits in cultivated tomato (*Solanum lycopersicum* L.). Scientia Horticulturae 164: 109194.

Werner K, Friedt W, Ordon F (2005) Strategies for pyramiding resistance genes against the barley yellow mosaic virus complex (BaMMV, BaYMV, BaYMV-2). Mol Breed 16: 45-55.

Xu FF, Tang FF, Shao YF, Chen YL, Tong C, Bao JS (2014) Genotype × environment interactions for agronomic traits of rice revealed by association mapping. Rice Sci 21: 133-141.

Yang L, Wang YY, Noushin J, Hu HT, Chen P, Shang LG, Lin HY, Dong GJ, Hu J, Gao ZY, Qian Q, Zhang Y, Guo LB (2019) Genome-wide association analysis and allelic mining of grain shape-related traits in rice. Rice Science 26: 384-392.

Zhang B, Li WY, Chang XP, Li RZ, Jing RL (2014) Effects of favorable alleles for water-soluble carbohydrates at grain filling on grain weight under drought and heat stresses in wheat. PLoS One 18: e102917.

Figures

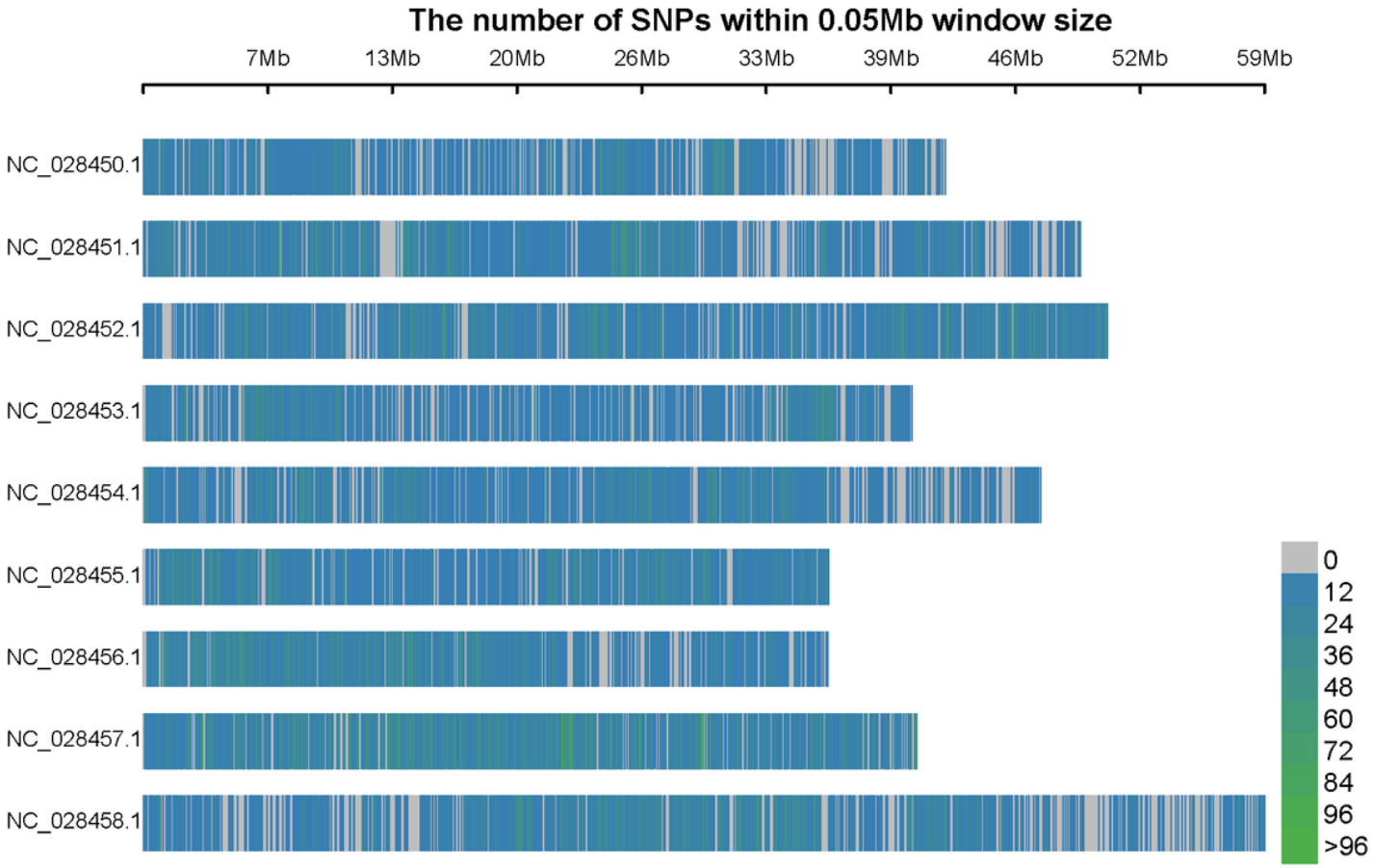

Figure 1

Distribution of the SNPs across the nine chromosomes

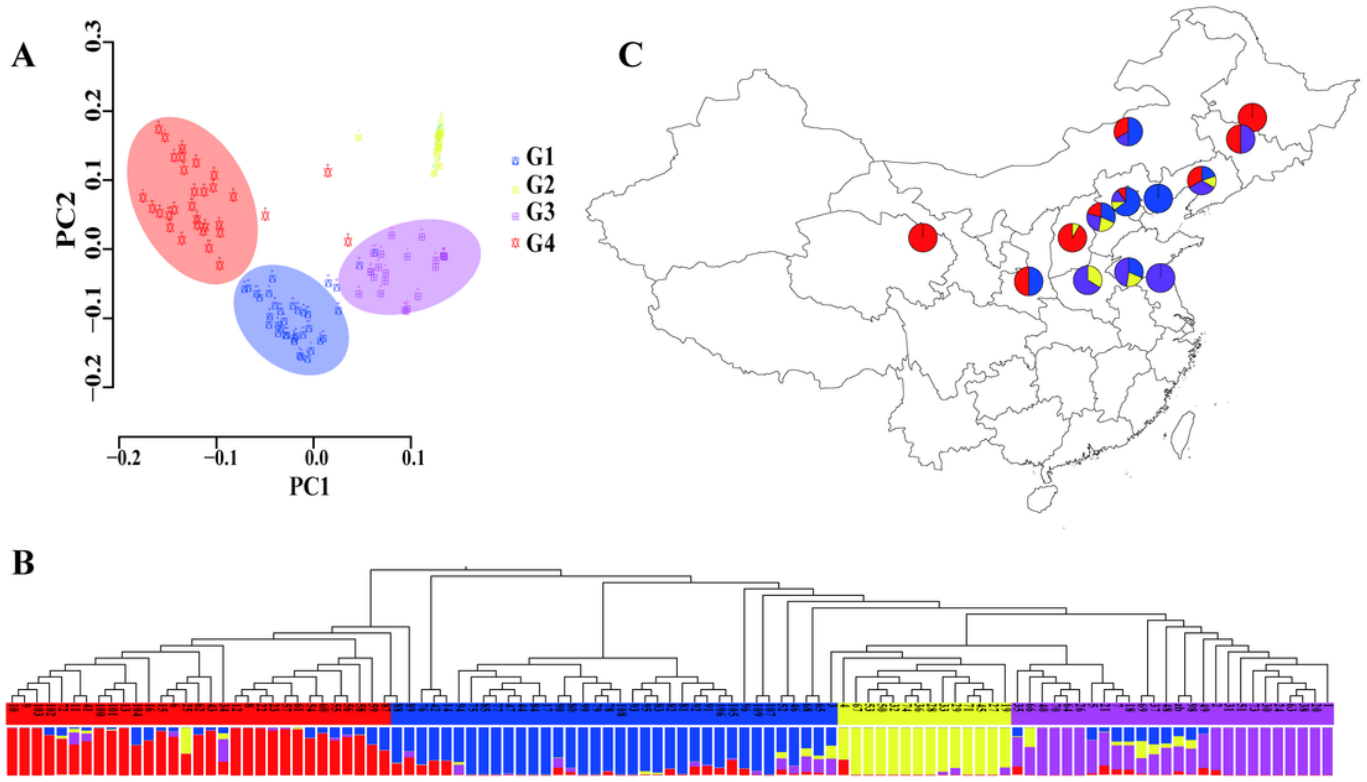

Figure 2

Geographic distribution, principal components analysis (PCA), neighbour-joining (NJ) phylogenetic tree, and ancestry structure of the 107 foxtail millet accessions examined here.

(a) PCA plot of the first two components. (b) Upper panel, NJ-phylogenetic tree of the 107 accessions used in this study; lower panels, results of ancestry structure analysis with K = 4. (c) Geographic distribution of the 105 accessions from China. The pie charts on the map show the proportions of the four groups (G1, G2, G3, and G4) at the geographical locations. Blue, yellow, purple, and red represent G1, G2, G3, and G4, respectively

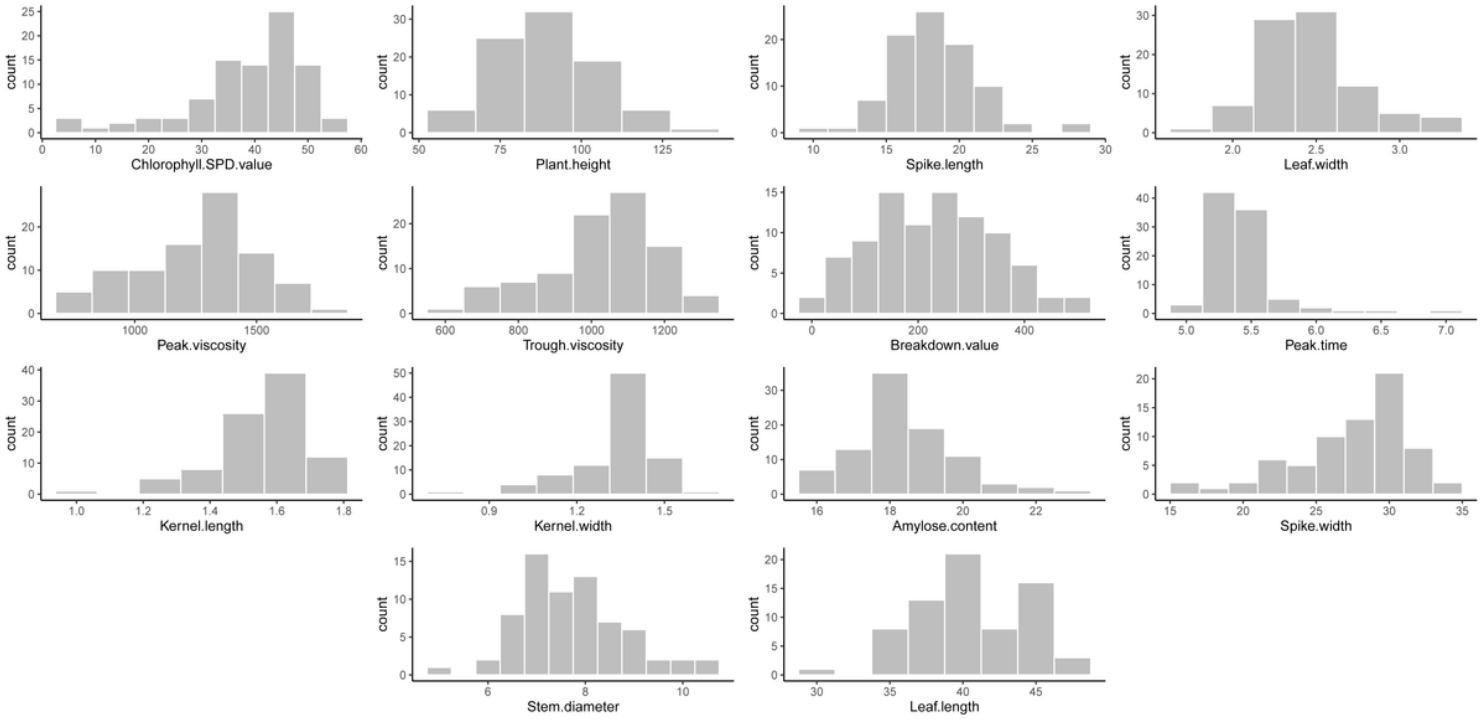

Figure 3

Frequency distributions of the phenotypic values of the 16 traits of the 107 foxtail millet accessions

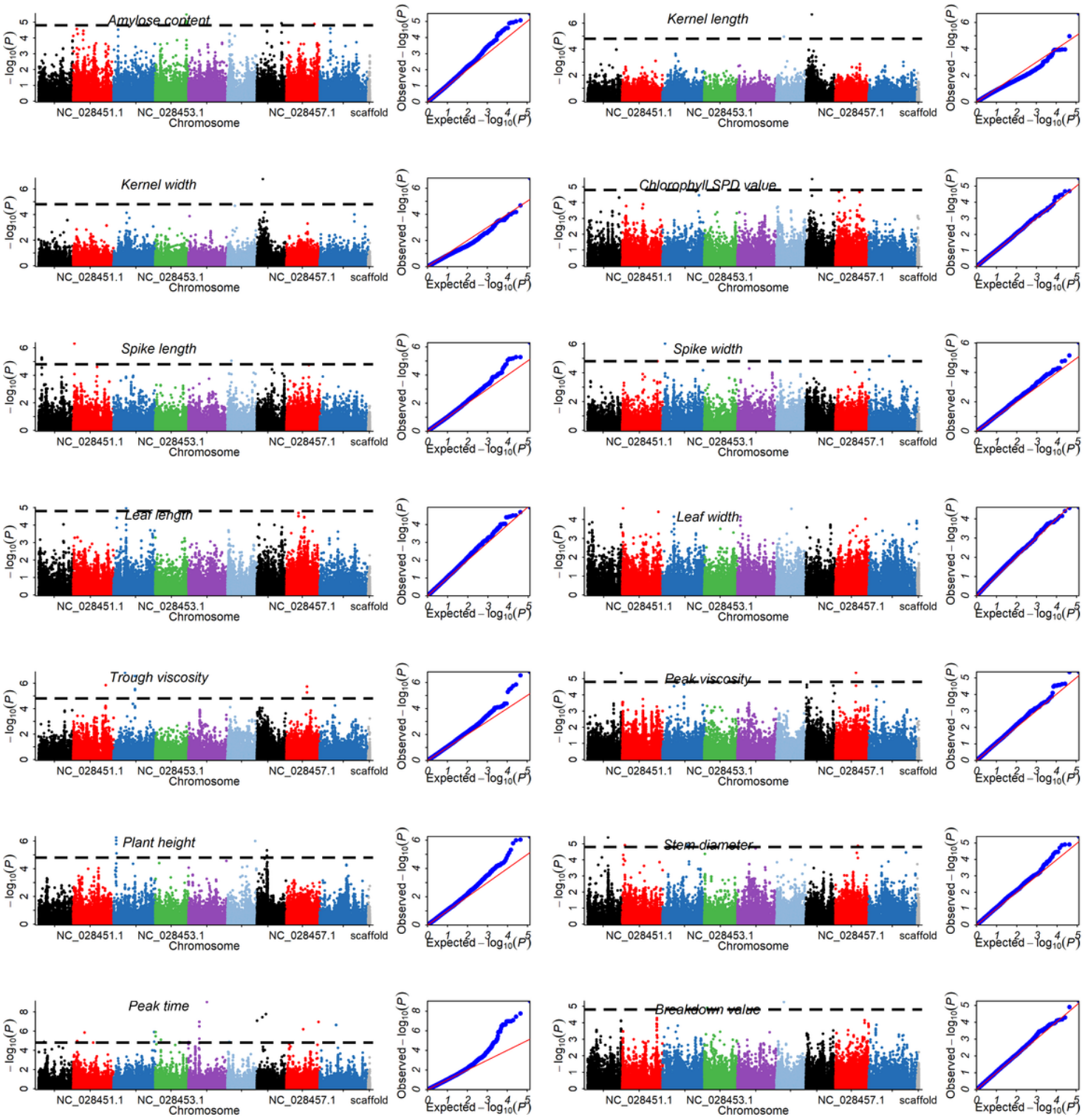

Figure 4

Genome-wide association study of the 16 traits of the 107 foxtail millet accessions

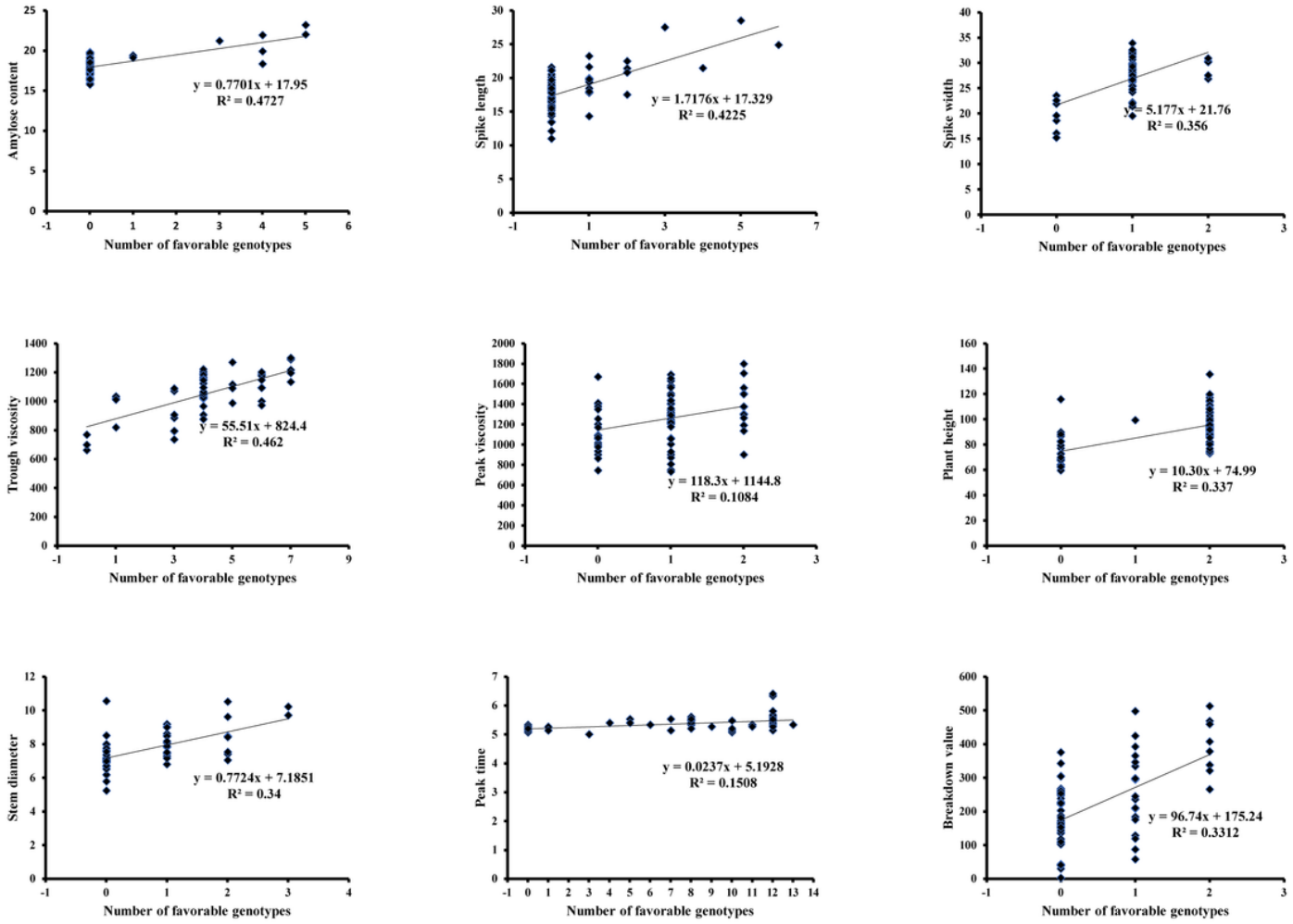

Figure 5

Linear regression analyses of the numbers of favourable genotypes on the different phenotypic values

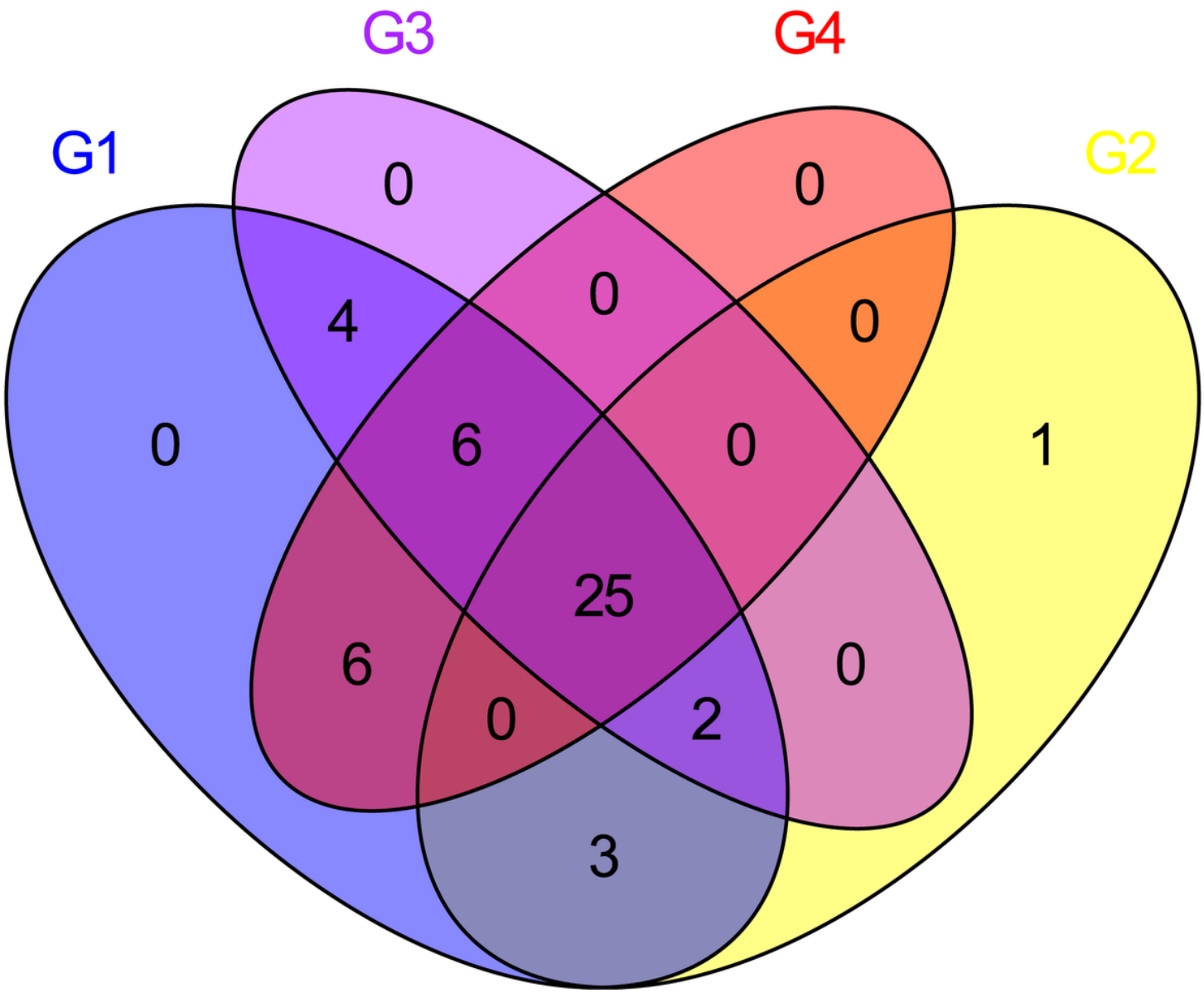

Figure 6

Distribution of favourable genotypes in G1, G2, G3, and G4 groups

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryfileS1.xlsx
- SupplementaryfileS2.xlsx
- SupplementaryfileS3.xlsx
- SupplementaryfileS4.xls
- SupplementaryfileS5.xlsx

- SupplementaryfileS6.xlsx