Alma Mater Studiorum – Università di Bologna

DOTTORATO DI RICERCA IN

SCIENZE E TECNOLOGIE AGRARIE, AMBIENTALI E ALIMENTARI

Ciclo 35

**Settore Concorsuale:** 07/E1 – CHIMICA AGRARIA, GENETICA AGRARIA E PEDOLOGIA

**Settore Scientifico Disciplinare:** AGR/07 - GENETICA AGRARIA

# CHARACTERIZATION OF AN INTERNATIONAL TETRAPLOID WHEAT GERMPLASM INCLUDING LANDRACES AND PRIMITIVE WHEAT TOWARDS IMPROVED RESILIENCE TO ABIOTIC AND BIOTIC STRESSES AND QUALITY

**Presentata da:** Matteo Campana

**Coordinatore Dottorato** — Massimiliano Petracci

**Supervisore** — Roberto Tuberosa

**Esame finale anno 2023**

# Sommario

# ABSTRACT

This thesis aimed to characterise two large tetraploid germplasm collections. The Global Durum Panel, involving modern cultivars and landrances and the Tetraploid Global Collection which comprises all the tetraploid wheat subgroups. Two distinct parallel studies were carried out.

The first is focused on the characterisation of both collection for yield and quality related traits. The panel were phenotyped for two consecutive years each. In this phase the following traits were collected: the number of fertile spikelets per spike, the number of fertile florets of central spikelet for the spike-related traits. The following grain related traits were also phenotyped: the thousand kernel weight, the average grain area, average grain length, average grain width, grain brightness, grain redness, grain yellowness. GWAS analysis were performed for each collected trait and major QTLs were subjected to candidate gene analysis. Major QTLs emerging from GWA study were located on chromosome 2A with a strong bibliographic evidence for grain number-related traits such as the fertile spikelet number, the number of fertile florets per central spikelet. On the other hand two evident peaks were detected on chromosomes 6A and 7B for grain size and weight related traits.

The second work was focused on the characterisation of the Global Durum Panel for root system architecture components, namely the root growth angle. GWAS analysis was perfomed and three major QTLs were detected on chromosome 2A, 6A and 7A. These three QTLs all have a bibliographic evidence.

# 1 INTRODUCTION

## 1.1 Taxonomy of durum wheat

Durum wheat (*Triticum turgidum* ssp. *durum*) is a tetraploid wheat species belonging to the Poaceae family. More specifically it is part the Triticeae, a tribe of the Pooideae subfamily, which comprises more than 300 species, including grass crops of remarkable relevance such as rye (*Secale cereale*) and barley (*Hordeum vulgare*). Six are the species included in the *Triticum* genus: *T. monococcum* L. (AA genome), *T. urartu* Tumanian ex Gandilyan (AA genome), *T. turgidum* L. (AABB genome), *T. timopheevi* (AAGG genome), *T. aestivum* L. (AABBDD genome). and *T. zhukovskyi* (AAAAGG genome). (Matsuoka., 2011).

## 1.2 Durum wheat as a domesticated species

A prime and fundamental distinction between wheat species is between wild and domesticated species. The durum wheat belongs to the domesticated species. On the phenotypical level, the wild and the domesticated species differ in three main aspects: firstly, the wild wheats present smaller seeds in comparison with the domesticated forms, which have wider seeds. The second main difference regards the rachis toughness. On the one hand, wild forms have a brittle rachis which leads to the fragmentation of the spikelets during the crops ripening phase. On the other, domesticated forms present tougher rachis which prevents the ear from being shattered during the crops' ripening, as a result these second forms turn out to be more practical and harvestable. Thirdly, wild and domesticated forms are also distinguishable from one another because of the position of the seeds and the glumes: a tight bond between the seeds and the glumes usually characterizes wild forms and the domesticated forms tend to be free-threshing, as they release the seeds from the glumes (Salamini et al., 2002). Therefore, two larger groups are identified as the hulled wheat to which the wild forms usually belong and the free-threshing wheats to which domesticated wheat forms generally pertain.

## 1.3 The history domesticated tetraploid species: Durum wheat

Wheats is also classified in three different sections according to ploidy level: Sect. Monococcon (mainly diploid species), Sect. Dicoccoidea (chiefly tetraploid species), Sect. Triticum (principally hexaploidy species) (Matsuoka., 2011). The Durum wheat belongs to the Sect. Dicoccoidea. If we were to trace a brief history of history of the domesticated wheat species, *T. monococcum* (AA diploid genome) and *T. Urartu* (AA diploid genome) occupy a unique position. *T. monococcum* was one of the first wheat species to be domesticated in the Karacadag mountain range, in South-eastern Turkey and in the Northern part of the Levantine region. *T. monococcum* was obtained directly from its wild form *T. boeticum* (Feldman, 2001). According to Matsuoka, approximately one million years ago a wild wheat species, *T. Urartu* (AA diploid genome), come to a genetic divergence, contemporarily in the North- western Iraq and in Eastern Turkey. Polyploidy is widely known as a noteworthy tool for evolution: the occurrence of genetic diploidization and dosage compensation in polyploid wheats mainly entails genes coding for structural or storage protein, while enzyme-coding loci remain active (Feldman., 2001). Hence, polyploids can tolerate an increased amount of genetic variation caused by mutations, in fact polyploids usually show a better adaptation to a larger number of environments and a wider range of morphological traits. The appearance of the first tetraploid wheat occurred about 300.000-500.000 years ago in the Fertile Crescent due to a hybridization leading to an allopolyploidization between *T. urartu* (2n = 2x = 14, genome AA) and *Aegilops speltoides* (2n = 2x = 14, genome SS), thus originating wild emmer wheat (*T. turgidum* ssp. *dicoccoides*). It is generally accepted that two separated hybridization events had occurred between *T. urartu* and *A. speltoides,* generating two different tetraploid wheat as a result. The first hybridization determined the genesis of *T. turgidum* (2n = 4x = 28, AABB genome), and the second that of *T. timopheevi* (2n = 4x = 28, AAGG genome) (Feldman., 2001). However, because hybrids between these species have a high sterility rate, it has been highlighted the possibility that the B and G genomes could have diverged at the tetraploid stage if not before the hybridization, thus coming from two different crosses between *T. urartu* and diploid species (Feldman., 2001) as a result. According to molecular and cytological studies, the B genome, which donor has been identified in *A. speltoides* (genome SS), may have been evolved from a different species strictly

related to *A. speltoides* (Feldman., 2001).

Durum wheat wild progenitor is wild emmer wheat (*Triticum turgidum* ssp *dicoccoides*). Phytogeographical studies highlighted the Fertile Crescent as the wild wheat point of origin. The Fertile Crescent is a wide area placed in the Near East Asia, extending from the East Mediterranean basin to western Iran, including South-eastern Turkey, Lebanon, Syria, Israel, Jordan and the Tigris-Euphrates basin (Fig. 1). Wild emmer wheat's cultivation had started as early as the Pre-Pottery Neolithic, nearly 10.000 years ago and wild emmer's domestication is referred back to the Pre-Pottery Neolithic B (9000 years ago) in the Levantine corridor (Feldman., 2001; Matsuoka., 2011). This region is the western section of the Fertile Crescent, extending from South-eastern Turkey to Israel along the Mediterranean Sea and includes the Nile banks. Wild emmer domestication led to *T. dicoccum* (domesticated emmer), characterized by a non- brittle rachis and hulled seeds (Sahri et al., 2014). Human selection, then, brought to the appearance of several physiological traits of the utmost agronomic relevance like larger seeds, an increased apical dominance and a decreased seeds dormancy leading to a distinguished tetraploid wheat *T. turgidum* ssp *durum*, the durum wheat. A pivotal trait emerging besides the previous was the free-threshing trait. Wild and domesticated emmer both possess hulled seed and hard glumes which required a harsh threshing process during the harvest, Durum wheat on the other hand has softer glumes allowing an easier threshing. This characteristic is affected by mutations at many loci: *Tg* (tenacious glume) and *Q*, which interaction has an epistatic nature. *Tg* affects mainly the glume toughness while *Q* is involved in determining spike shape, glume toughness, plant height and the spike emergence time in a pleiotropic behaviour. (Simons et al., 2006).


Figure 1 Map of the Near East: the red dashed line delimitates the region of the Fertile Crescent, which is considere to be the site of origin of the wheat species. Source: Salamini et al., 2002)

# 2. CHAPTER I

## 2.1 Introduction

### 2.1.1 Yield, a complex quantitative trait

The main characteristic of quantitative traits is that they can be measured and oftentimes they have a remarkable importance from an economic point of view. These traits are controlled by a complex genetic network and are known as metric traits or polygenic traits. Their expression occurs through multiple loci called QTL (quantitative trait loci). Each QTL contributes to the final phenotype with a plus or minus effect respect to phenotypic mean, while on the other hand in qualitative traits loci effects are either absence or presence. Moreover, polygenic traits are characterized by a continue variability, following a normal phenotypic distribution that does not allow to divide them in distinct categories. In addition to that, quantitative traits are affected by the environment and this may hide the genetic effects. So, it is of the utmost importance to evaluate environmental components with the aim to reduce its effects by performing experiments in multiple replicates and multiple environments.

Grain yield improving has always been characterized by many constraints as it is a typical quantitative trait controlled by a plethora of genes, which are largely affected by environmental factors and human management. During the last fifty years, genetic improvements both in bread and durum wheat have been mostly accomplished by enhancing harvest index as well as decreasing plant height (Mangini et al., 2018). Grain yield in wheat is usually reported to be associated with grain number and the achievement of significant improvements in yield without increasing grain number seems unreachable. Even though this method resulted to be very beneficial, it only took in consideration the final number of grains set, thus putting the focus on the final grain amount produced which determines spike fertility (Guo et al., 2016).

### 2.1.2 Major yield components traits

Grain yield is conveyed as the sum of a many traits known as “yield components”. The usually refer to the number of spikes per surface unit and grain yield per spike. In grain yield per spike (GYS) are included the number of kernels per spike (KNS) and the kernel weight, normally expressed as thousand kernel weight (TKW). Grain weight itself is characterized by further sub-components connected to seed morphology such as grain area, grain length and grain width (Mangini et al., 2021). Moreover TKW is a trait of the utmost relevance because of its direct connection to industrial quality. Traits like the number of kernels per spike and grain weight are inherited quantitively, while the number of spikes within the surface unit, especially under conventional cropping systems, depends mostly on planting density (Mangini et al., 2018).

More in detail, the correlation between KNS and TKW phenotypes has been usually found negative but not always consistent, while the correlations between GYS and KNS and TKW have been always resulted to be positives (Mangini et al., 2018).

The aforementioned correlations among phenotypes might be attributed to many elements: genetic linkage, pleiotropy, environmental factors, and competition between yield components for a scarce nutrient element in common. A likely hypothesis for the negative correlation between kernel weight and number could be that the improvement of grain number production results in a lower disposal of nutrients during for each grain, which in turn creates a reduction in single grain weight due to competition effects (Mangini et al., 2018).

These traits usually show normal distribution trends and the wide variation underlines the polygenic control of yield components (Fig. 5).

Grain number is a determinant component for grain yield, and grain number determinantion relies heavily on floret fertility. Floret fertility is restrained by the allocation of assimilates to spikes and their distribution (Guo et al., 2017). It is of utmost importance to better understand assimilate supply to the grains; furthermore, it is also a prerequisite to learn about the critical traits and genes controlling assimilate distribution. In order to comprehend assimilate partitioning, Guo et al. (2017) studied five patterns of dry weight distribution.

The first step is the tiller-to-main shoot: one of the grain yield components in several crops is tillering because tiller number is a relevant factor in determining the competition for assimilate supply between tillers and main shoot.

The second step is spike-to-stem: the introduction of *Rht* genes has greatly mitigated assimilate competition occurring between spikes and stems.

The third refers to the spikelet-to-spikelet within a spike competition: spikelet fertility can be exploited to investigate the competition for assimilates between spikelets in a spike.

Floret-to-floret within individual spikelets is the fourth step: competition for assimilates between florets is possibly determined by a large loss in grain number.

The fifth stage is the grain-to-spike chaff: the spike fertility index, the ratio between grain number per spike and the weight of spike chaff (called spike fertility index), is an important indicator of dry weight distribution between grains and spike chaff.

### 2.1.3 Putative genes with direct effects on yield

Usually yield-related traits are determined by major genes which can be divided in many groups. Transcription factors, that could affect grain number due to spike development regulation; genes involved in growth regulators signalling thus determining plant architecture; genes affecting cell division, involved in grain size changes; genes which regulate inflorescence architecture and seed number; and genes involved in carbohydrate metabolism, affecting plant architecture and grain yield (Mangini et al., 2021).

### 2.1.4 Grain weight and grain size genes

*TaGS5* is the wheat ortholog of rice *OsGS5* (Brinton & Uauy, 2019) and it is located on chromosome 3 group in wheat, usually expressed in developing grains and young spikes. *OsGS5* codes for a putative serine carboxypeptidase and its overexpression in rice is associated with a positive regulation of mitotic cell division which leads to a pericarp cell expansion (Li et al., 2011). *TaGS5-3A„* has been extensively studied and a single nucleotide polymorphism transition T/G which leads to a missense mutation hence to a amino acid change of alanine to serine, was detected. The *TaGS5-3A-T* allele bearing RILs showed a significant 2% increase in grain weight when compared to the *TaGS5-3A-G* allele with higher gene expression and enzyme activity (Me et al., 2015).

The centromeric region of chromosome 6A in wheat bears a locus affecting grain size and yield and it is shown to be the ortholog of rice OsGW2. *TaGW2* codes for a Ring-type E3 ligase which is involved in ubiquination activitiy and it is a negative regulator of grain weight. Wheat homeologues usually show a similar expression pattern, but it tends to change in relation to the development phase: *TaGW2-A* and D are most expressed up to the anthesis, while *TaGW2-B* shows an enhanced expression level during late grain filling (Tillet et al., 2022).

RNAi silencing of TaGW2 led to the detection of an enhanced transcript levels of cytokinin synthesis genes like *TaIPT2* and lower expression patterns of cytokinin degradation related genes such as *TaCKX1* (Geng et al., 2017). A TILLING study performed across multiple environments and years on *TaGW2-6A* mutants, showed a 6.6% increase in GW (Simmonds et al., 2016).

### 2.1.5 Grain number and floral architecture genes

The number of grains per spikelet is determined by each single floret fertility. A wheat spikelet generally produces up to 12 floret primordia but after anthesis, several florets undergo abortion thus leading to a reduced final number of grains in the spikelet (Guo et al., 2016).

In a study conducted by Sakuma et al., (2019), a single major QTL has been mapped on chromosome 2A, accounting for 61% of the phenotypic variance. This *GNI* gene codes for an HD-zip transcription factor and its expression has been associated negatively with floret fertility (Sakuma et al., 2019). *GNI* is more active during the development of the apical florets and the rachilla. Three diverse allele were studied for *GNI-A* through a haplotype analysis of 111 accessions comprehensive of wild emmer and modern durum cultivars. *GNI-A105N,* the wild type allele, codes for an asparagine at the 105th amino acid position within the protein, *GNI-A105K* codes for a lysine and *GNI-A105Y* for a tyrosine. RILs has been tested in field experiments and plants carrying the *105Y* allele showed a yield advantage of 10 to 30% more when compared to the wild type allele. In this test Grain number per spike was increased but with no negative consequences for Grain weight (Sakuma et al., 2019).

*WAPO1* is gene that codes for an F-box protein which, in wheat, is part of the Skp1-Cullin1-F-box complex involved in ubiquination of target substrates and their subsequent degradation by the proteosome system (Tillet et al., 2022). In rice, the ortholog *APO1* has been observed to be involved with C-class MADS box genes, which codes for transcription factors that regulate floral tissue determination and when overexpressed, *WAPO1* has been associated with an enhanced spikelet number per spike (Wittern et al., 2022). It is likely that WAPO1 acts as a delayer of the termination of the inflorescence growth, thus leading to more branching and finally more spikelets in the spike.

## 2.2 MATERIALS & METHODS

### 2.2.1 Global Durum Panel

Genetic variability is not deemed as a relevant element on its own, but recently breeders and researchers are exploiting useful genetic variability aiming for certain genomic regions which are well known to be relevant (Tuberosa and Pozniak., 2014). Thus, in order to identify useful alleles and subsequently make them available for pre-breeding and breeding efforts, the Global Durum Panel (GDP) was created. The GDP was designed starting from the Durum Wheat Reference Collection (DWRC), which is composed of 2.503 tetraploid wheat accessions, provided by 25 worldwide institutions and partners (Mazzucotelli et al., 2020). The DWRC comprised *T. durum* modern cultivars, an Evolutionary Pre-breeding pOpulation from INRA, France (EPO, David et al., 2014), *T. durum* landraces and wild tetraploid wheat subspecies.
From the starting set, 762 accessions were selected basing on molecular data to constitute the GDP, thus capturing 94-97% of the starting variability. All the accessions were then genotyped with the Illumina iSelect 90K SNP array technology (Wang et al., 2014). This generated a total of 42.520 polymorphic SNPs, which were then filtered for missing data (Mazzucotelli et al., 2020).

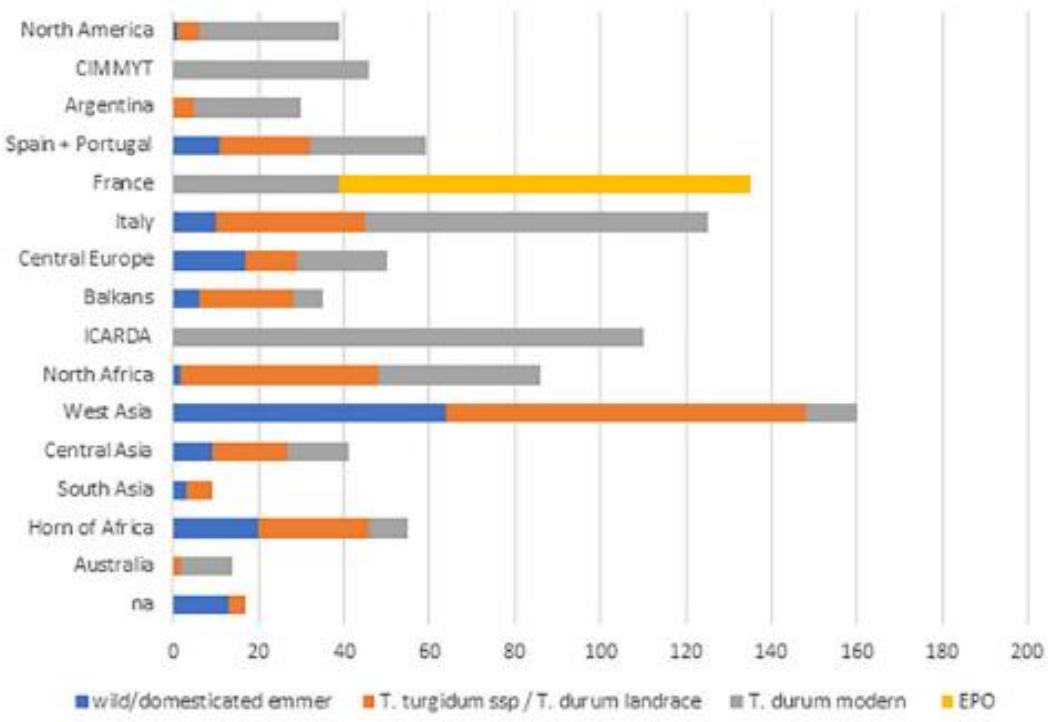


Figure 2 Geographic origin of the accessions belonging to the GDP (Mazzucotelli,2020)

GDP was grown in Pian del Volpi (Grosseto, Italy, 42° 57' 54.226" N, 11° 5' 37.152" E) for two consecutive seasons (2020 and 2021) at the APSOV experimental station.

### 2.2.2 Tetraploid Global Collection

Up to now, many factors such as the migration of man, modern agriculture and trade were involved in the spread of the main taxa of *T. turgidum* ssp. *dicoccum* (domesticated emmer) and *T. turgidum* ssp. *durum* from Fertile Crescent to Africa, Europe and India. This diffusion, along with the selection carried out by man for many domestication, adaptation and quality-related traits, led to a germplasm marked out by a very large biodiversity, which is considered to be the foundation of the pursuit for future wheat improvements.

Hence, in order to investigate the unravelled genetic variability in tetraploid wheat, a comprehensive panel of wheat genotypes, including all the major tetraploid germplasm pools (modern durum elite, durum landraces, wild and domesticated emmer and, had been assembled and measured for genetic diversity through the Illumina iSelect 90K SNP genotyping platform.

This work was performed by gathering already genotyped collections and new sets of genotypes to enhance the representativeness of the panel (Maccaferri et al., 2019). All the wheat accessions were refined through single seed descent (SSD) generations in greenhouse and then genotyped. Overall, 90K SNP genotypic data were produced for a total of 2.558 accessions (Maccaferri et al., 2015, Wang et al., 2014). Raw genotyping information related to modern durum cultivars, durum landraces and emmer were provided by AgriBio, CREA, University of Bologna, University of Saskatchewan and USDA-ARS. In addition to that, 490 tetraploid wheat accessions from the areas of domestication (Mediterranean Basin, Fertile Crescent, East and West Asia) were added to improve the representativeness of the collection.

TGC was grown over two seasons in two different environments: APSOV experimental station in Pian del Volpi (Grosseto, Italy) and Cadriano (Bologna, Italy, 44° 33' 8.933" N, 11° 24' 51.458" E).

### 2.2.3 Phenotyping

#### *2.2.3.1 Spike phenotyping*

The phenotyping protocol was carried out at DISTAL, Bologna on both collections. Approximately twelve spikes were harvested at physiological maturity, then stored as a bundle. Six spikes were subsequently selected based on phenotypic homogeneity and then scored for spike morphology and spike fertility traits. Data were collected in two different excel sheets.

In the first sheet were listed all the qualitative non-numeric traits, mainly used for accession identification, which are based on a direct assessment of the spike morphology-related characteristics:

1. **Spike shape** assessed according to six different categories: Square shape (SQR), Spear shape (SPR), Pyramid shape (PYR), Clavate shape (CV), Ethiopian shape with long spikes (ETH1), Ethiopian shape with weak awns (ETH2)

2. **Spike compactness** visually ranked as follow: Very Low (LL), Low (L), Medium (M) High (H), Very High (HH)
3. **Glume pubescence/hairiness** (YES or NO)
4. **Glume colour** has been assessed based on the different colours detected: White (W), Bronze/Red (BRZ), Brownish (BRN), Black Veined (BV), Black/Blue (B)
5. **Awn colour** was classified according to the different colours detected: White (W), Bronze/Red (BRZ), Brown (BRN), Black (B)

The second sheet included the quantitative data related to fertility traits and spike length. Every phenotypic trait was collected for six selected spikes for each accession:

1. **Spike length** measured as the average of the six spikes collected
2. **Sterile spikelet** per spike (number)
3. **Fertile spikelet** per spike (number)
4. **Fertile florets per central spikelet** (number), recorded as the ratio between the number of fertile florets and the total number of florets.

The phenotyping protocol was standardised for each accession. Here below are listed the steps followed in the process:

1. Six spikes showing a homogeneous and consistent phenotype were selected
2. Qualitative data collection
3. Image acquisition of the spikes, including a ruler and a label showing the genotype code for identification


Figure 3. Spike photograph

4. Quantitative spike fertility-related data collection in the following order:

4.1. Average spike length

4.2. Sterile spikelet number

4.3. Fertile spikelet number

Fertile florets out of the total number of florets in the central spikelet

5. Image acquisition of approximately twenty seeds of the phenotyped spikes following the same procedure employed in the step number 3

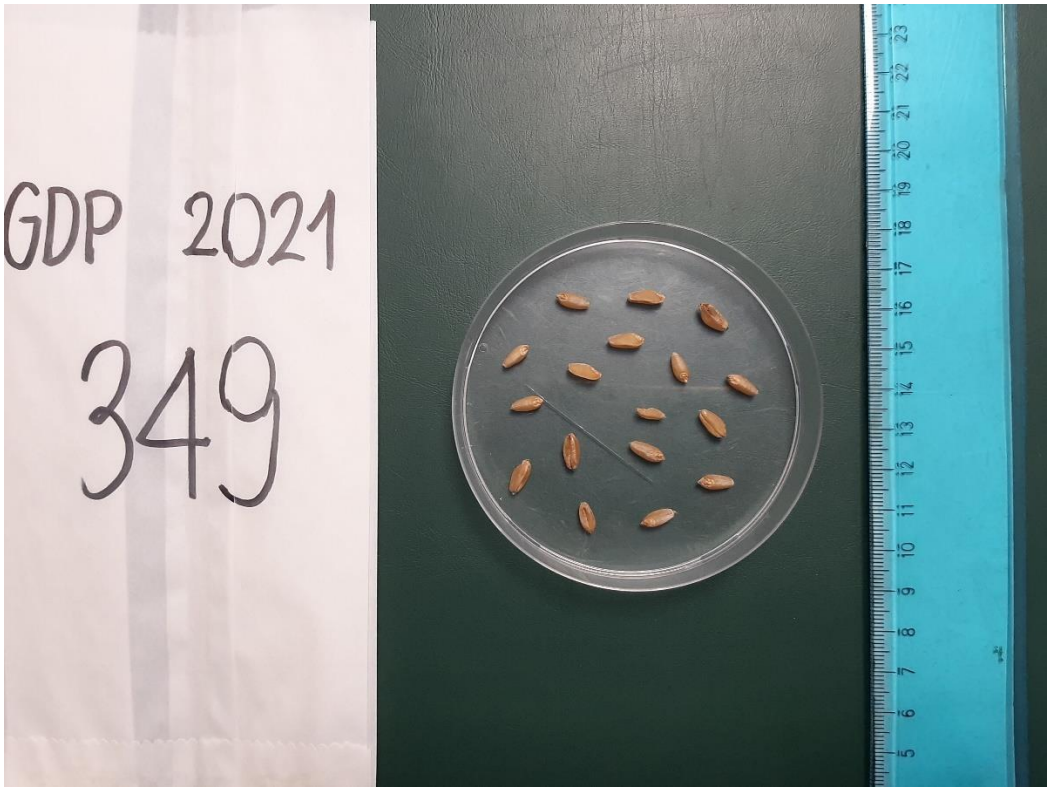


Figure 4. Seeds photograph

6. Seeds were then stored in a small paper bag for consecutive phenotype analysis regarding grain traits.

#### *2.2.3.2 Grain phenotyping*

Seeds obtained previously from the spike destructive data collection, were measured for several grain yield and quality related traits such as Thousand Kernel Weight (TKW), Grain Surface (Area), Grain Perimeter, Grain Length, Grain Width and Grain Colour.

Seeds were weighted beforehand to record the TKW, then fifty seeds were selected based on size and homogeneity: broken seeds, off-types, white chalky and shrivelled seeds were discarded. These seeds were subjected to digital image analysis. Image were obtained through a flatbed colour image scanner CanoScan LiDE 400 (Canon Inc., Tokyo, Japan) with an optical resolution of 4800 dpi. Accessions seeds were scattered on the flatbed scanner keeping them separated for precise measurements.

A black cardboard was placed on the scanner in order to enhance contrast. Images were taken at 300 dpi and saved in JPEG format. Digital images were then analysed through GrainScan (Whan et al., 2014), a software specifically developed for grain size and colour measurements. The average value of fifty seeds for each accession was measured.

Grain length and grain width were obtained using the default threshold provided by the software. Colours were recorded by the scanner in raw RGB values, which were subsequently converted in CIEL*a*b* values, a space colour characterised by three dimensions: L* indicates brightness, a* positive values represent redness and negative values indicate greenness while b* positive and negative values indicate yellowness and blueness respectively.

### 2.3 Field experimental design

Both panels were field tested following a modified unreplicated augmented design with eight checks replicated in each block. Accessions were grown in $2m^2$ plots with 6 rows and 0.5 m spacing between plots. Checks employed are listed as follows: Karim, Iride, Trouvè=Nachit, Saragolla, Monastir, Faraj, Cham-1, Altar84.

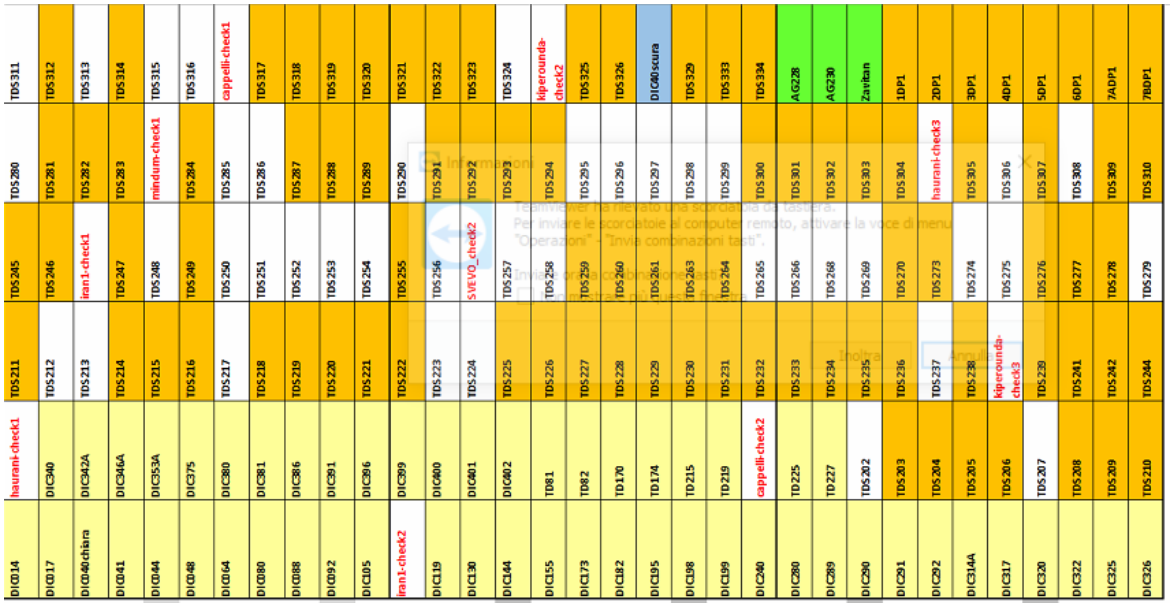


Figure 5 Field map of TGC in Cadriano during the 2019 season. Checks are hghlighted in red and reapeated within each block

## 2.4 Statistical Analysis

Statistical analysis was carried out with RStudio software (RStudio Team, 2020). Heritability was also computed for every investigated trait with R package *repeatability*.

R package *lme4* was used to produce best linear unbiased estimators (BLUEs) for each phenotypic data in every environment. Different parameters were considered in each environment and cluster of environments.

Clusters of environments were analysed with the following variables:

A. **~ Genotype + Block + Heading date + Environment + Genotype:Environment**

In this model genotype and heading date were treated as fixed variable, while block, environment and interaction between genotype and environment (GxE) were considered as random variables. ANOVA was performed to detect significant environment and GxE interactions using a 0.05 p-value threshold.

BLUEs were obtained for single environments including the following variables in the model:

A. **~ Genotype + Block + Heading date**

## 2.5 Imputation and LD decay

Polymorphic information (PIC) content was determined for the dataset using the following formula (Serrote et al., 2020):

- **1 – (MAF)2 – ((1-MAF)2**

PIC measures the ability of a marker to find polymorphisms; for this reason it has huge relevance in selecting markers suitable for genetic studies (Serrote et al., 2020).

An R script developed at UNIBO was employed to filter the HapMap file based on the following parameters: Minor Allele Frequency (MAF) higher than 0.01, SNPs missing call higher than 0.3, samples with a missing rate above 50%. Following the filtering process, the ultimate HapMap file included 23423 SNPs. After that the genotyping dataset was imputed with Beagle v5.4 (Browning et al., 2021) to assign A/B variants to missing SNPs based on their position and nearest SNPs (Beagle 5.4 uses a linkage disequilibrium-based algorithm).

The imputed vcf file was employed to compute the Linkage Disequilibrium decay in the durum germplasm with the software Tassel 5 (Bradbury et al., 2007). The LD decay was then plotted through three linkage thresholds (r 2 equals 0.3, 0.5, and 0.8).

## 2.6 Pruning and Population structure analysis

PLINK software (Chang et al., 2015) was employed for the pruning phase in order to remove redundant SNPs in the HapMap file and create three files for the different $r^2$ thresholds (0.3, 0.5, and 0.8).

Output files were then subjected to population structure analysis with ADMIXTURE, a model-based likelihood method. ADMIXTURE was ameliorated with the block relaxation algorithm, the quasi-Newton convergence acceleration method, and q = 3 secants (Alexander et al., 2009), defining the sub-population memberships from k=2 to k=20. In order to detect the best number of subpopulations to be subjected to the analysis, the cross-validated error rate, delta cv error, minimum group size, maximum admixed lines in a group, and admixed lines percentage were taken into account. The minimum k for the best parameters was chosen, and, as for the reported dataset, k=10 with a $r^2$ = 0.5 was used.

TASSEL 5 was employed in order to convert the imputed HapMap file into a distance matrix and thus create the kinship data frame through the convertion of the values in genetic relatedness. Heatmap and ward clustering (Ward.D2 algorithm) were computed on the kinship matrix with the R (R Core team, 2020) packages pheatmap v1.0.12 and dendextend v1.15.2.

Neighbour Joining Tree was calculated with the R package adegenet v2.1.5

## 2.7 GWAS Analysis

Genome Wide Association Study (GWAS) was carried out with the R package GAPIT3 (Wang and Zhang., 2021).

Threshold of permutations was compared to the Bonferroni adjusted threshold, which has been calculated by dividing the significant p-value of 0.05 with the number of markers at a $r^2$ threshold of 0.8 and calculating the negative logarithm in base 10. On average, the Bonferroni threshold calculated via the permutation steps varied between 4 and 5. Thus examined peaks above the threshold had an enhanced probability of $10^4$-$10^5$ in resulting associated with phenotypic variance. GAPIT3 R package was used to carry out GWAS analysis including the following model: GLM (naive, MLM + K, MLMM + K, FarmCPU and Blink.

Within every model, the PCA number was set to 0 and model selection to false.

Final GWAS output were portrayed as Manhattan plot graphs while data for every trait were merged in a single file including each model considered.

## 2.3 Results

Descriptive statistics for each environment were obtained. Histograms for each phenotypic data distribution are showed here as well as descriptive statistics data, heritability and ANOVA results.

### 2.3.1 GDP 2020

Table 1 Descriptive statistics for GDP 2020 for spike related traits

| | SS | FS | FF | UF |
|---|---|---|---|---|
| **min** | -0.2 | 13.47 | 1.9 | 0.71 |
| **max** | 1.92 | 24.96 | 6.03 | 2.52 |
| **range** | 2.12 | 11.49 | 4.13 | 1.81 |
| **median** | 0.1 | 18.48 | 3.94 | 1.49 |
| **mean** | 0.29 | 18.53 | 4.06 | 1.55 |
| **SE.mean** | 0.02 | 0.06 | 0.03 | 0.02 |
| **var** | 0.17 | 2.84 | 0.62 | 0.17 |
| **std.dev** | 0.42 | 1.68 | 0.79 | 0.41 |
| **coef.var** | 1.45 | 0.09 | 0.19 | 0.26 |
| $h^2$ | 0.78434 | 0.696403 | 0.831716 | 0.653571 |

Table 2 Descriptive statistics for GDP 2020 for grain related traits

| | TKW | Area | Perimeter | Length | Width | L* | a* | b* |
|---|---|---|---|---|---|---|---|---|
| **min** | 27.18 | 15.64 | 20.66 | 6.63 | 2.78 | 48.27 | 5.96 | 14.52 |
| **max** | 75.04 | 25.5 | 26.88 | 9.31 | 3.98 | 62.84 | 12.04 | 26.6 |
| **range** | 47.87 | 9.86 | 6.22 | 2.67 | 1.2 | 14.57 | 6.07 | 12.08 |
| **median** | 49.99 | 20.52 | 23.46 | 7.83 | 3.37 | 55.59 | 7.94 | 20.64 |
| **mean** | 50.31 | 20.51 | 23.44 | 7.82 | 3.37 | 55.63 | 8.27 | 20.53 |
| **SE.mean** | 0.39 | 0.09 | 0.06 | 0.02 | 0.01 | 0.12 | 0.06 | 0.1 |
| **var** | 61.37 | 3.34 | 1.31 | 0.21 | 0.04 | 5.55 | 1.57 | 4.37 |
| **std.dev** | 7.83 | 1.83 | 1.15 | 0.45 | 0.2 | 2.36 | 1.25 | 2.09 |
| **coef.var** | 0.16 | 0.09 | 0.05 | 0.06 | 0.06 | 0.04 | 0.15 | 0.1 |
| $h^2$ | 0.631367 | 0.480435 | 0.687766 | 0.821772 | 0.460608 | 0.681824 | 0.195367 | 0.818884 |

Table 3 ANOVA results for GDP 2020

| Trait | Variables | Sum Sq | Df | F value | Pr(>F) | |
|---|---|---|---|---|---|---|
| SS | **Genotype** | 1851.43 | 686 | 2.9852 | 0.0000000000000002 | *** |
| SS | **Block** | 15.97 | 10 | 1.7668 | 0.06825 | |
| SS | **Heading date** | 5.46 | 10 | 0.6034 | 0.81012 | |
| SS | **Residuals** | 191.67 | 212 | | | |
| FS | **Genotype** | 482.72 | 684 | 5.5603 | 2.00E-16 | *** |
| FS | **Block** | 2.08 | 10 | 1.6374 | 0.09773 | |
| FS | **Heading date** | 0.85 | 10 | 0.6661 | 0.75511 | |
| FS | **Residuals** | 26.91 | 212 | | | |
| FF | **Genotype** | 125.286 | 688 | 2.8515 | <2e-16 | *** |
| FF | **Block** | 0.852 | 10 | 1.3339 | 0.2139 | |
| FF | **Heading date** | 0.881 | 10 | 1.3793 | 0.1914 | |
| FF | **Residuals** | 13.539 | 212 | | | |
| TKW | **Genotype** | 27197.5 | 402 | 2.6403 | 9.20E-12 | *** |
| TKW | **Block** | 371.1 | 10 | 1.4482 | 0.164 | |
| TKW | **Heading date** | 207 | 10 | 0.8078 | 0.6215 | |
| TKW | **Residuals** | 4048.6 | 158 | | | |
| Area | **Genotype** | 1422.5 | 399 | 2.0115 | 5.62E-07 | *** |
| Area | **Block** | 18.3 | 10 | 1.0325 | 0.41887 | |
| Area | **Heading date** | 34.4 | 10 | 1.9408 | 0.04372 | * |
| Area | **Residuals** | 269.4 | 152 | | | |
| Perimeter | **Genotype** | 544.23 | 399 | 3.2321 | 2.30E-15 | *** |
| Perimeter | **Block** | 1.69 | 10 | 0.3995 | 0.945233 | |
| Perimeter | **Heading date** | 11.08 | 10 | 2.6261 | 0.005639 | ** |
| Perimeter | **Residuals** | 64.15 | 152 | | | |
| Length | **Genotype** | 86.597 | 399 | 5.5204 | 2.20E-16 | *** |
| Length | **Block** | 0.193 | 10 | 0.4915 | 0.893597 | |
| Length | **Heading date** | 1.18 | 10 | 3.001 | 0.001744 | ** |
| Length | **Residuals** | 5.976 | 152 | | | |
| Width | **Genotype** | 17.2997 | 401 | 2.0027 | 6.86E-07 | *** |
| Width | **Block** | 0.4127 | 10 | 1.9156 | 0.04702 | * |
| Width | **Heading date** | 0.46 | 10 | 2.1354 | 0.02492 | * |
| Width | **Residuals** | 3.2528 | 151 | | | |
| L* | **Genotype** | 2254.2 | 397 | 2.8131 | 1.33E-12 | *** |
| L* | **Block** | 21.3 | 10 | 1.0553 | 0.4004 | |
| L* | **Heading date** | 11.7 | 10 | 0.5797 | 0.8288 | |
| L* | **Residuals** | 306.8 | 152 | | | |
| A* | **Genotype** | 525.17 | 392 | 1.1318 | 0.1876 | |
| A* | **Block** | 11.57 | 10 | 0.9777 | 0.4652 | |
| A* | **Heading date** | 3.28 | 10 | 0.2769 | 0.9854 | |
| A* | **Residuals** | 179.93 | 152 | | | |
| B* | **Genotype** | 2056.81 | 399 | 5.3193 | 2.00E-16 | *** |
| B* | **Block** | 17.25 | 10 | 1.7797 | 0.06868 | . |

| | | | | | |
|---|---|---|---|---|---|
| B* | **Heading date** | 5.37 | 10 | 0.5542 | 0.84879 |
| B* | **Residuals** | 147.3 | 152 | | |

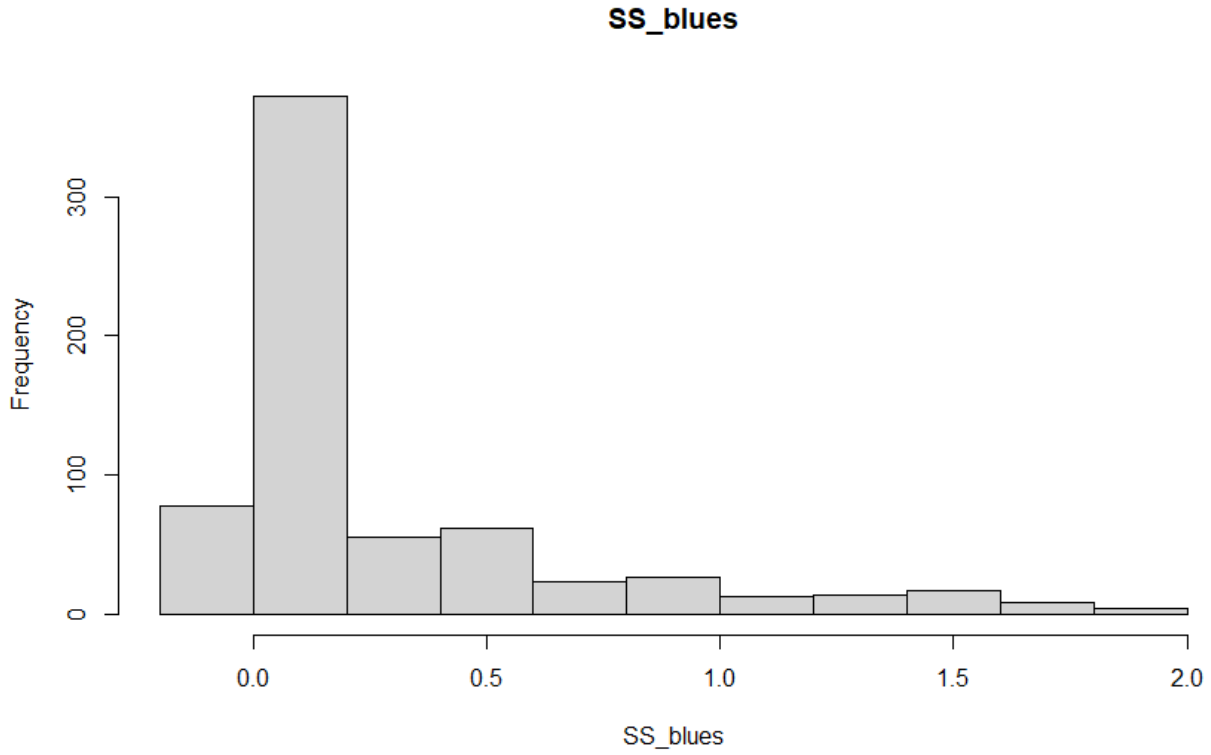


Figure 6 Sterile spikelets distribution frequencies in the GDP 2020

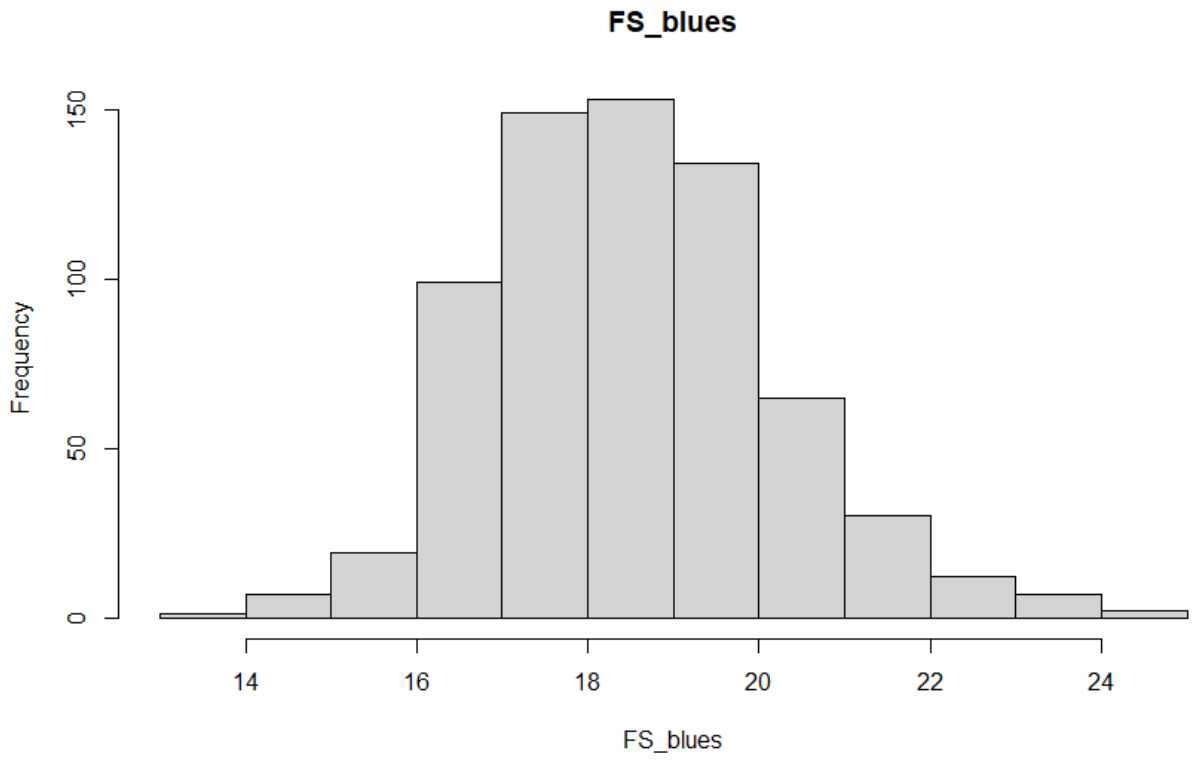


Figure 7 Fertile spikelet distribution frequency in the GDP 2020

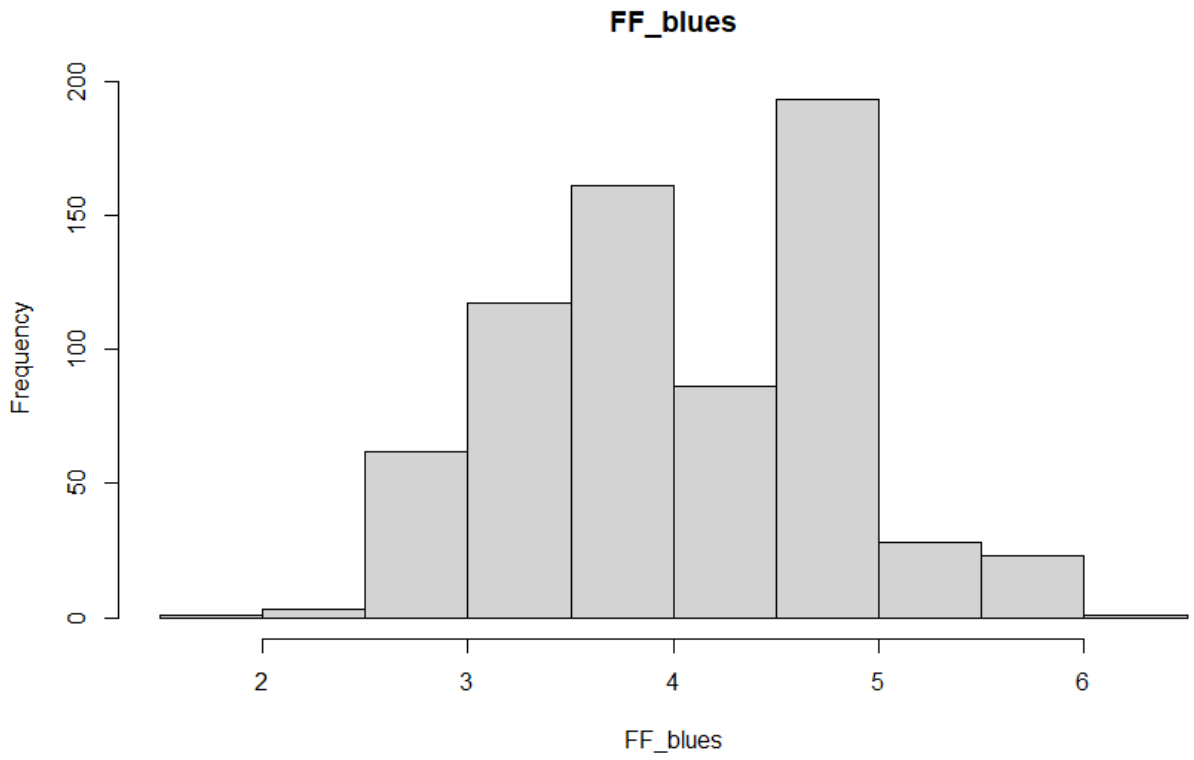


Figure 8 Fertile florets per central spikelet distribution frequency in the GDP 2020

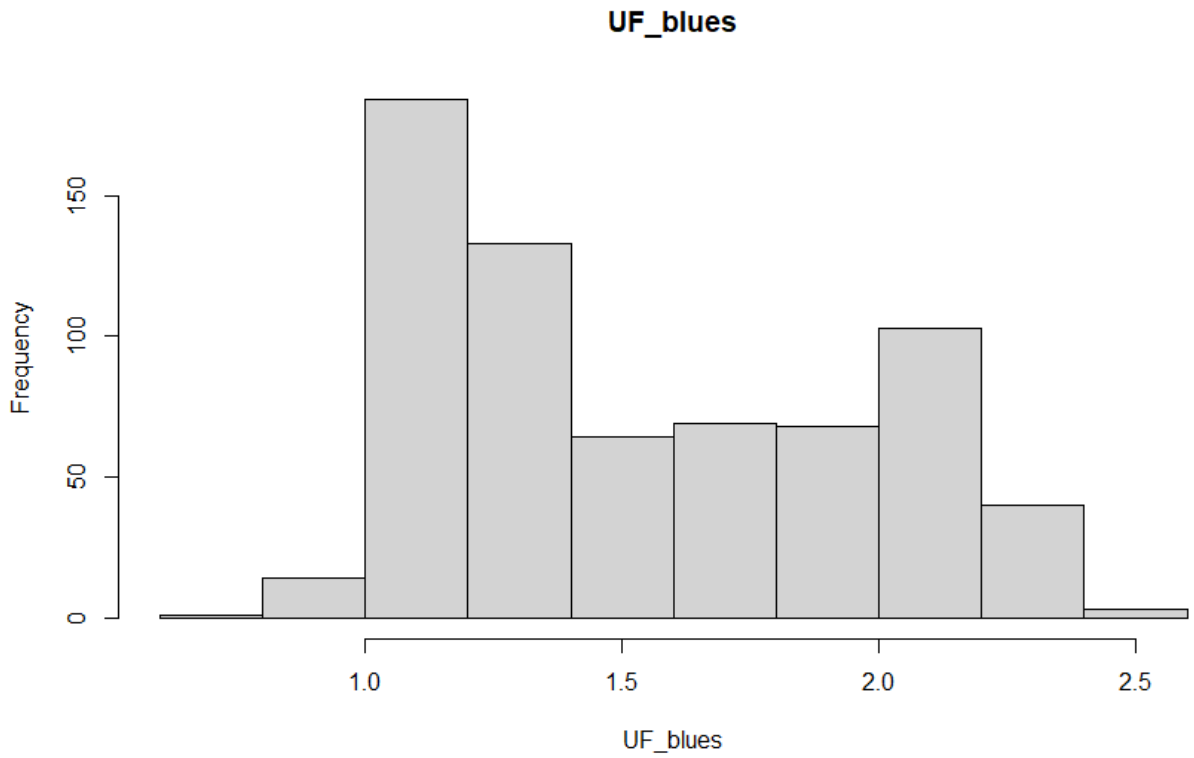


Figure 9 Unfertile florets distribution frequency in the GDP 2020

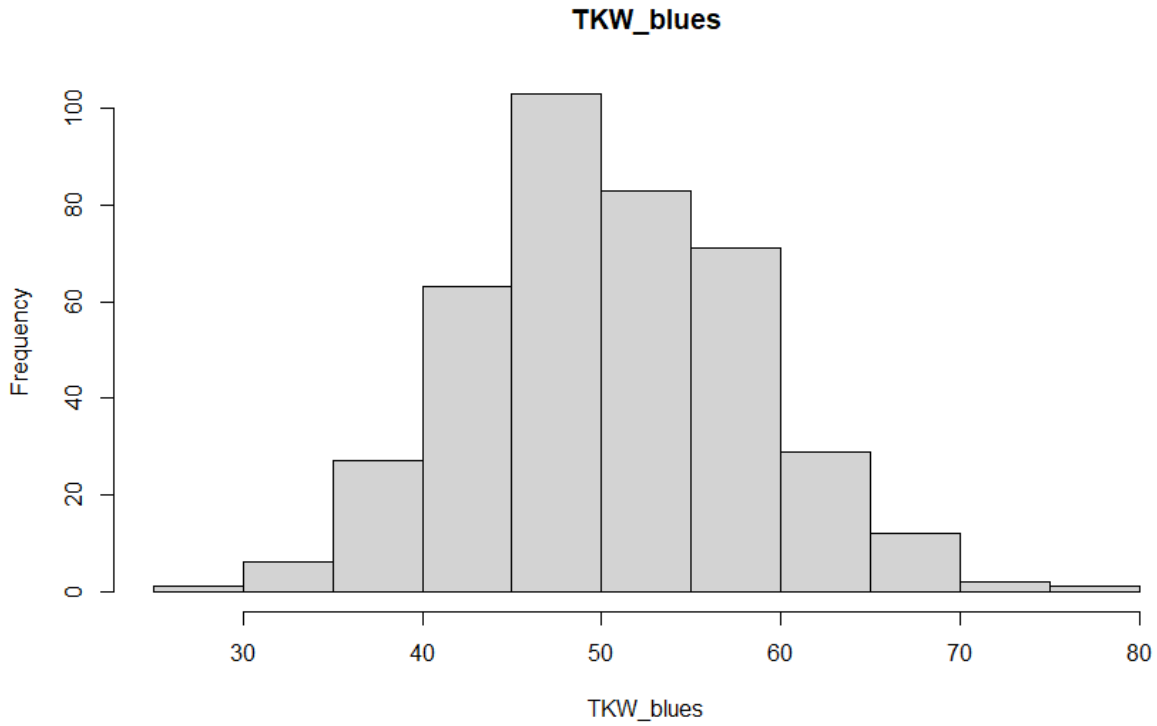


Figure 10 Thounsand kernel weight distribution frequency in the GDP 2020

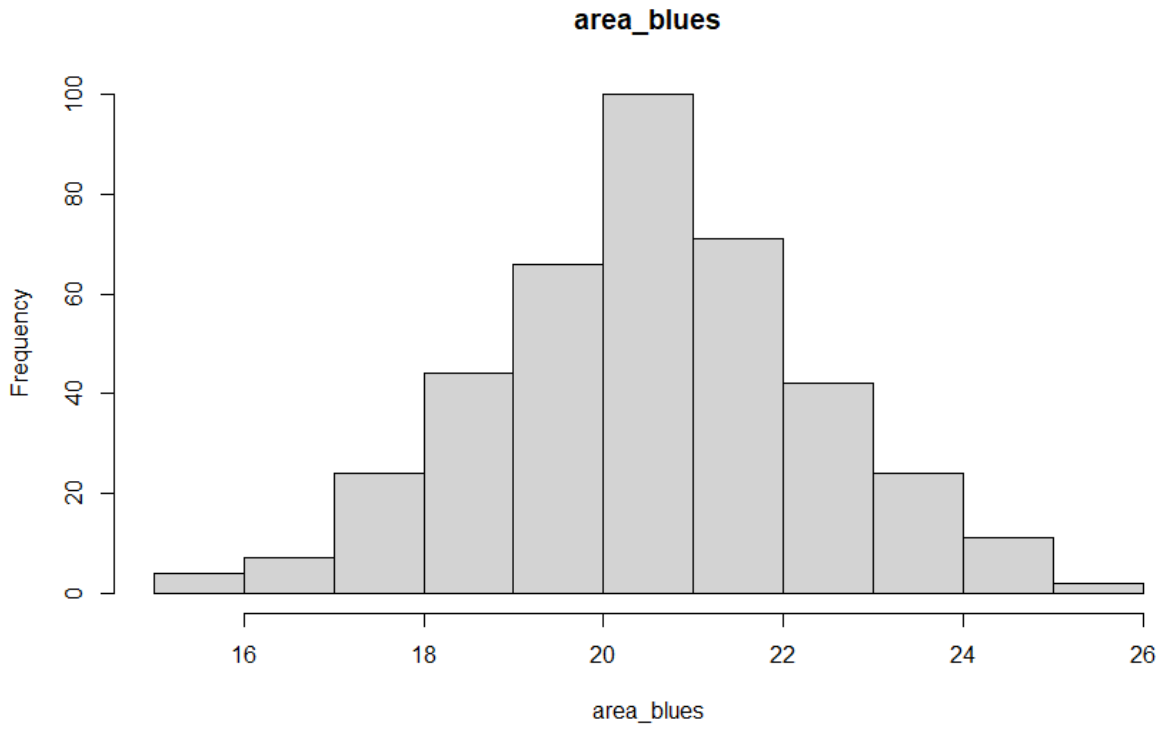


Figure 11 Grain area ditribution frequency in the GDP 2020

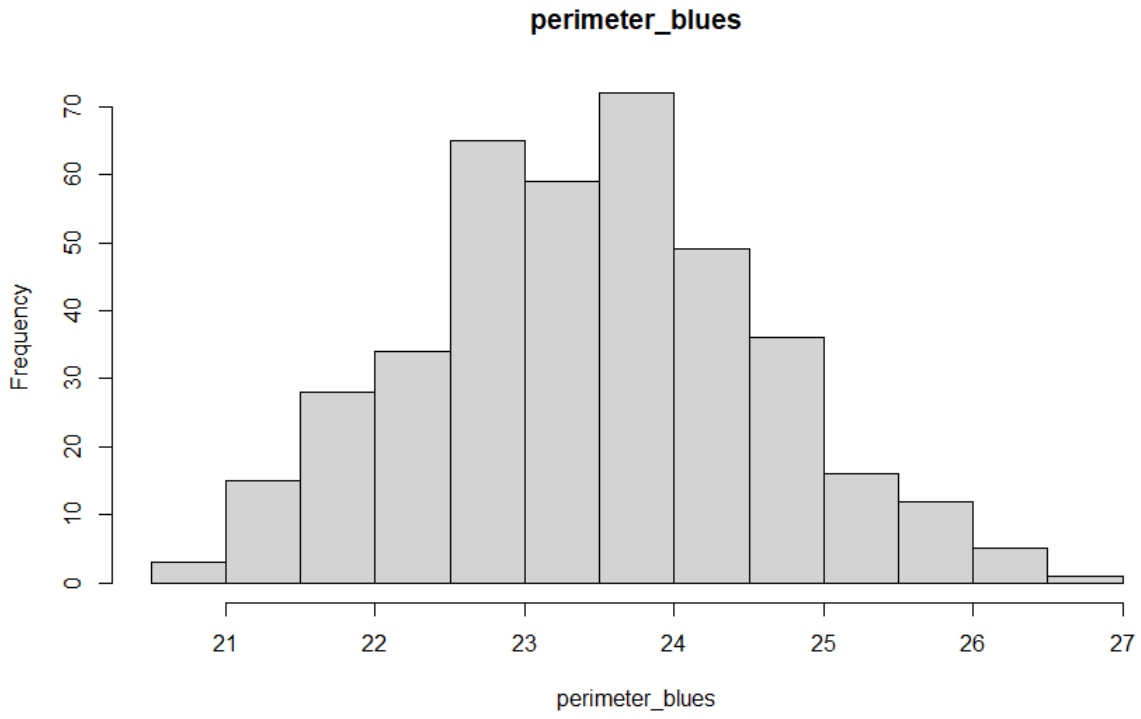


Figure 12 Grain perimeter ditribution frequency in the GDP 2020

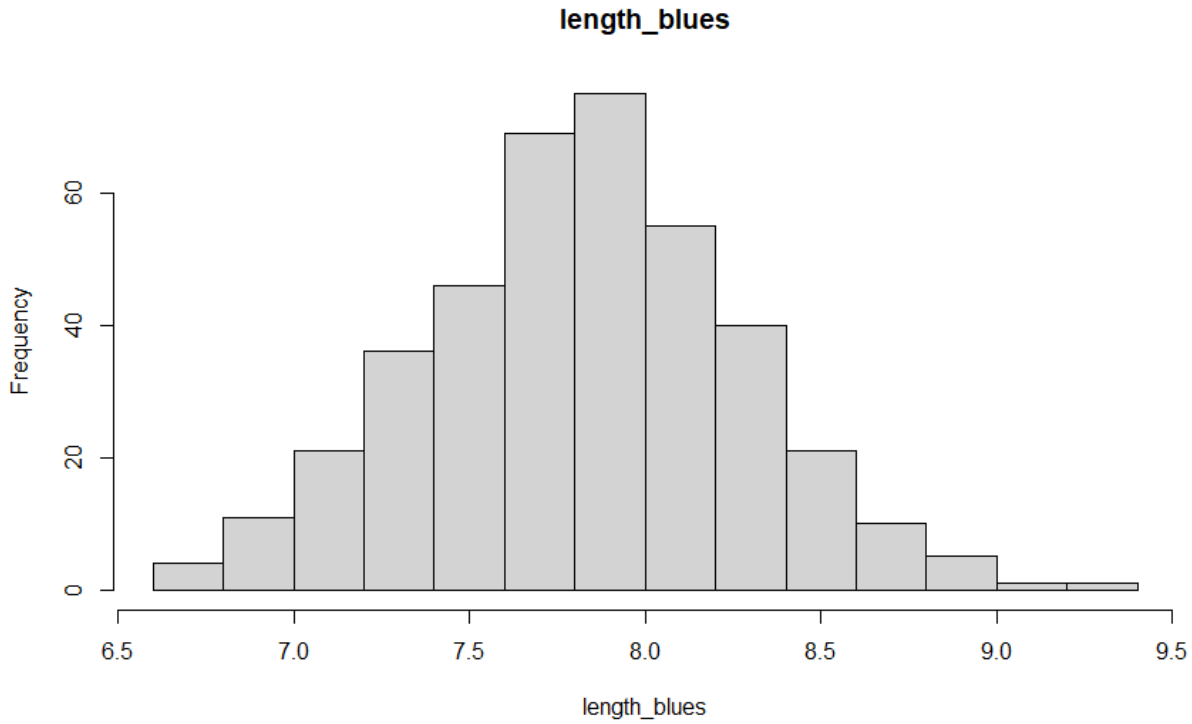


Figure 13 Grain length distribution frequency in the GDP 2020

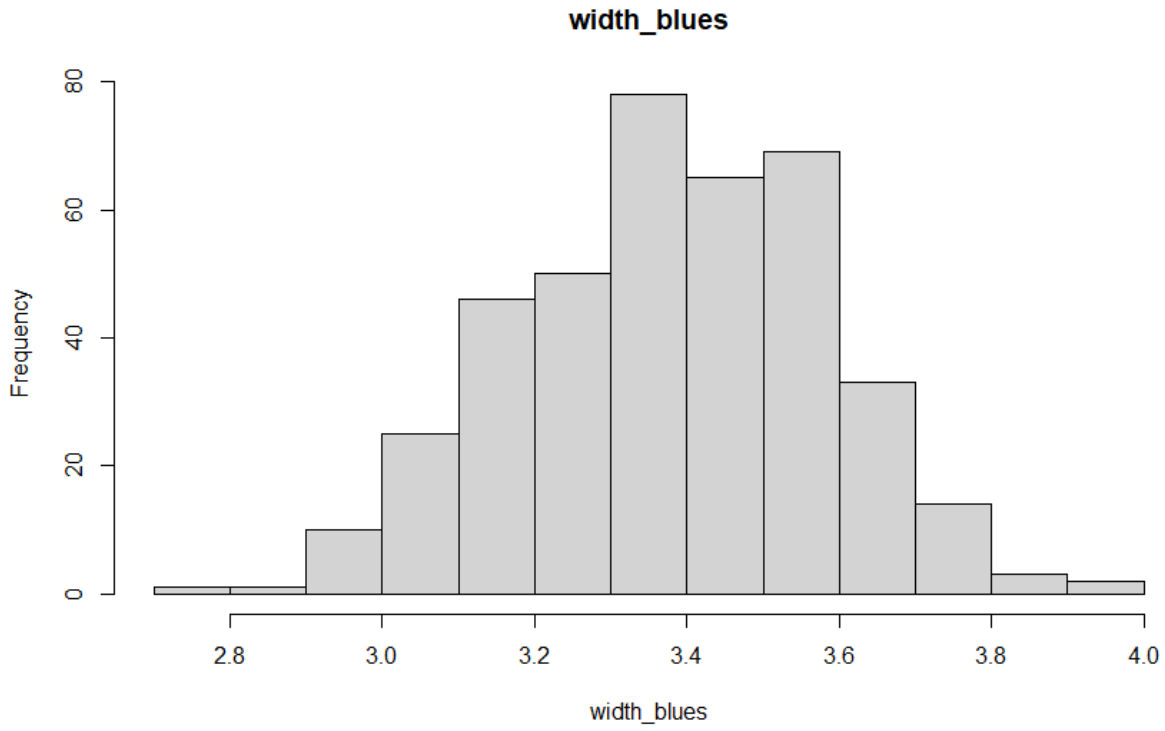


Figure 14 Grain width distribution frequency in the GDP 2020

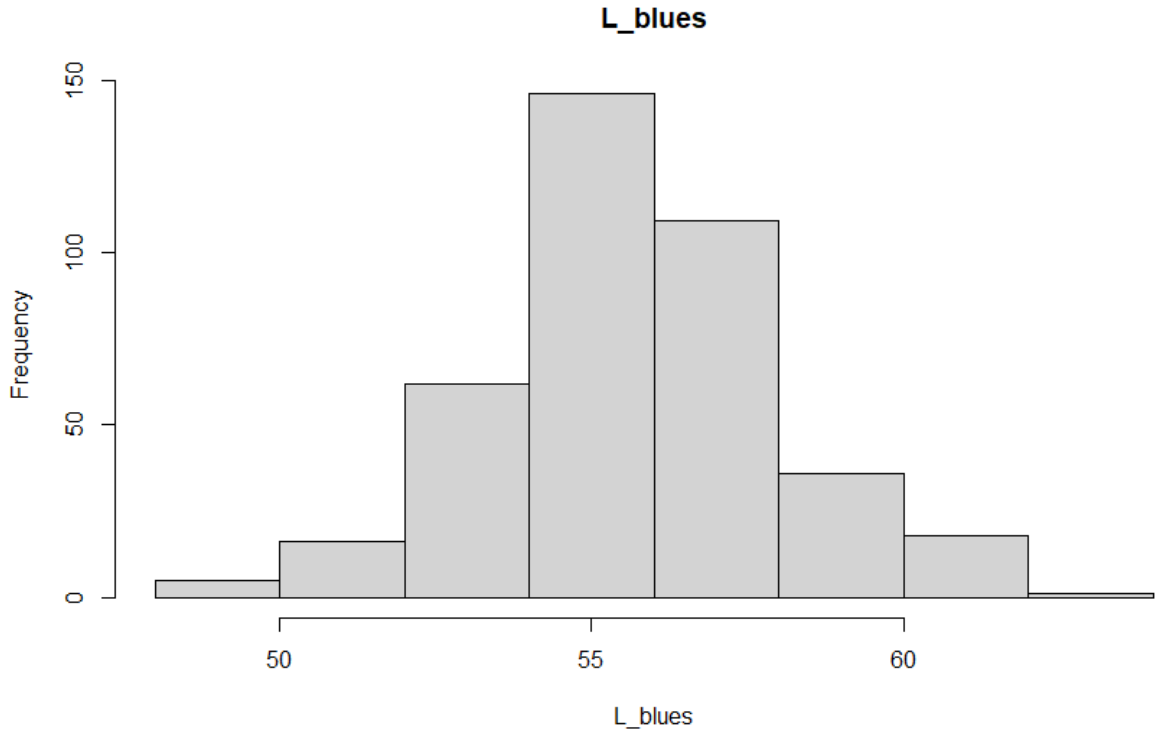


Figure 15 Grain brightness distribution frequency in the GDP 202

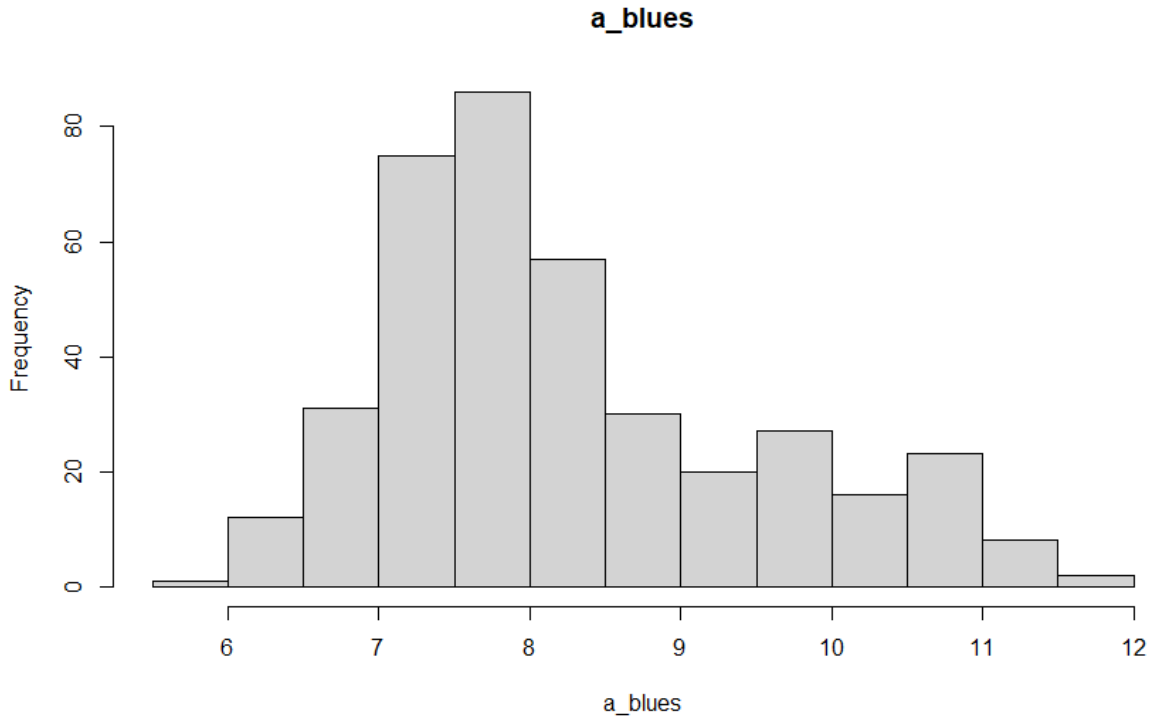


Figure 16 Grain redness distribution frequency in the GDP 2020

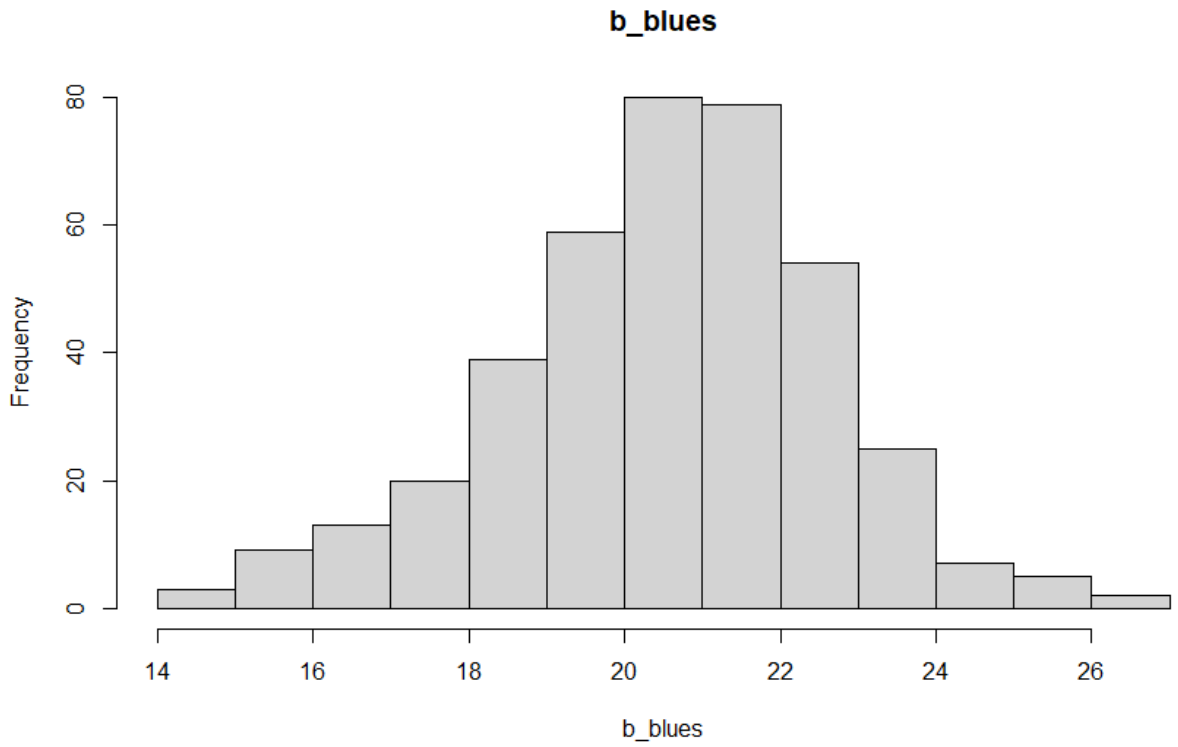


Figure 17 Grain yellowness ditribution frequency in the GDP 202

### 2.3.2 GDP 2021

Table 4 Descriptive statistics for GDP 2021 for spike related traits

| | SS | FS | FF | UF |
|---|---|---|---|---|
| **min** | -0.27 | 17.73 | 1.05 | -0.38 |
| **max** | 2.93 | 28.62 | 5.71 | 3.48 |
| **range** | 3.19 | 10.89 | 4.66 | 3.86 |
| **median** | 0.52 | 23.09 | 3.36 | 1.67 |
| **mean** | 0.68 | 23.19 | 3.33 | 1.68 |
| **SE.mean** | 0.02 | 0.07 | 0.03 | 0.02 |

| | | | | |
|---|---|---|---|---|
| **var** | 0.3 | 3.93 | 0.56 | 0.26 |
| **std.dev** | 0.55 | 1.98 | 0.75 | 0.51 |
| **coef.var** | 0.81 | 0.09 | 0.23 | 0.31 |
| $h^2$ | 0.865927 | 0.826054 | 0.789289 | 0.53102 |

Table 5 Descriptive statistics for GDP 2021 for grain related traits

| | TKW | Area | Perimeter | Length | Width | L* | a* | b* |
|---|---|---|---|---|---|---|---|---|
| **min** | 31.22 | 14.28 | 20.29 | 6.32 | 2.66 | 45.34 | 6.67 | 14.81 |
| **max** | 87.06 | 27.76 | 28.99 | 9.77 | 4.19 | 61.47 | 13.45 | 24.77 |
| **range** | 55.84 | 13.48 | 8.7 | 3.46 | 1.53 | 16.13 | 6.78 | 9.97 |
| **median** | 59.87 | 21.4 | 24.36 | 7.79 | 3.49 | 52.97 | 10.35 | 20.61 |
| **mean** | 59.47 | 21.34 | 24.3 | 7.78 | 3.49 | 53.03 | 10.12 | 20.46 |
| **SE.mean** | 0.37 | 0.09 | 0.05 | 0.02 | 0.01 | 0.07 | 0.05 | 0.06 |
| **var** | 97.54 | 5.5 | 1.61 | 0.25 | 0.07 | 3.9 | 1.7 | 2.79 |
| **std.dev** | 9.88 | 2.35 | 1.27 | 0.5 | 0.26 | 1.97 | 1.31 | 1.67 |
| **coef.var** | 0.17 | 0.11 | 0.05 | 0.06 | 0.07 | 0.04 | 0.13 | 0.08 |
| $h^2$ | 0.746399 | 0.691539 | 0.766108 | 0.755733 | 0.776966 | 0.737745 | 0.367573 | 0.86956 |

Table 6 ANOVA results for GDP 2021

| Trait | Variables | Sum Sq | Df | F value | Pr(>F) | |
|---|---|---|---|---|---|---|
| SS | **Genotype** | 170.113 | 709 | 5.3228 | <2e-16 | *** |
| SS | **Block** | 0.561 | 9 | 1.384 | 0.2014 | |
| SS | **Heading date** | 0.389 | 12 | 0.72 | 0.7299 | |
| SS | **Residuals** | 5.995 | 133 | | | |
| FS | **Genotype** | 2521.24 | 712 | 4.5173 | 2.00E-16 | *** |
| FS | **Block** | 13.87 | 9 | 1.9663 | 0.04803 | * |
| FS | **Heading date** | 8.94 | 12 | 0.9504 | 0.49945 | |
| FS | **Residuals** | 104.26 | 133 | | | |
| FF | **Genotype** | 332.56 | 713 | 3.9018 | 2.00E-16 | *** |
| FF | **Block** | 2.29 | 9 | 2.133 | 0.03085 | * |
| FF | **Heading date** | 2.15 | 12 | 1.4987 | 0.13207 | |
| FF | **Residuals** | 15.9 | 133 | | | |
| TKW | **Genotype** | 50301 | 695 | 2.9566 | 6.74E-13 | *** |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| TKW | **Block** | 271 | 9 | 1.2314 | | 0.2812 | |
| TKW | **Heading date** | 285 | 12 | 0.9713 | | 0.4794 | |
| TKW | **Residuals** | 3231 | 132 | | | | |
| Area | **Genotype** | 3077.03 | 693 | 2.415 | | 3.94E-09 | *** |
| Area | **Block** | 15.62 | 9 | 0.9439 | | 0.4898 | |
| Area | **Heading date** | 11.84 | 12 | 0.5368 | | 0.8871 | |
| Area | **Residuals** | 231.66 | 126 | | | | |
| Perimeter | **Genotype** | 1013.31 | 693 | 3.9167 | | <2e-16 | *** |
| Perimeter | **Block** | 2.42 | 9 | 0.7215 | | 0.6884 | |
| Perimeter | **Heading date** | 4.81 | 12 | 1.0742 | | 0.3873 | |
| Perimeter | **Residuals** | 47.04 | 126 | | | | |
| Length | **Genotype** | 174.458 | 693 | 3.8048 | | <2e-16 | *** |
| Length | **Block** | 0.691 | 9 | 1.1609 | | 0.3258 | |
| Length | **Heading date** | 0.595 | 12 | 0.7498 | | 0.7003 | |
| Length | **Residuals** | 8.337 | 126 | | | | |
| Width | **Genotype** | 31.718 | 693 | 3.1185 | | 2.31E-13 | *** |
| Width | **Block** | 0.146 | 9 | 1.1054 | | 0.3638 | |
| Width | **Heading date** | 0.071 | 12 | 0.4023 | | 0.9605 | |
| Width | **Residuals** | 1.849 | 126 | | | | |
| L* | **Genotype** | 2360.03 | 691 | 3.7088 | 2.00E-16 | | *** |
| L* | **Block** | 20.74 | 9 | 2.5027 | 0.01144 | | * |
| L* | **Heading date** | 18.05 | 12 | 1.6338 | 0.09022 | | . |
| L* | **Residuals** | 116.03 | 126 | | | | |
| A* | **Genotype** | 980.98 | 693 | 1.2835 | 0.04116 | | * |
| A* | **Block** | 5.48 | 9 | 0.5521 | 0.8337 | | |
| A* | **Heading date** | 15.01 | 12 | 1.134 | 0.33888 | | |
| A* | **Residuals** | 138.96 | 126 | | | | |
| B* | **Genotype** | 1822.69 | 691 | 6.8342 | 2.00E-16 | | *** |
| B* | **Block** | 5.97 | 9 | 1.7198 | 0.09119 | | . |
| B* | **Heading date** | 3.08 | 12 | 0.6642 | 0.78268 | | |
| B* | **Residuals** | 47.86 | 124 | | | | |

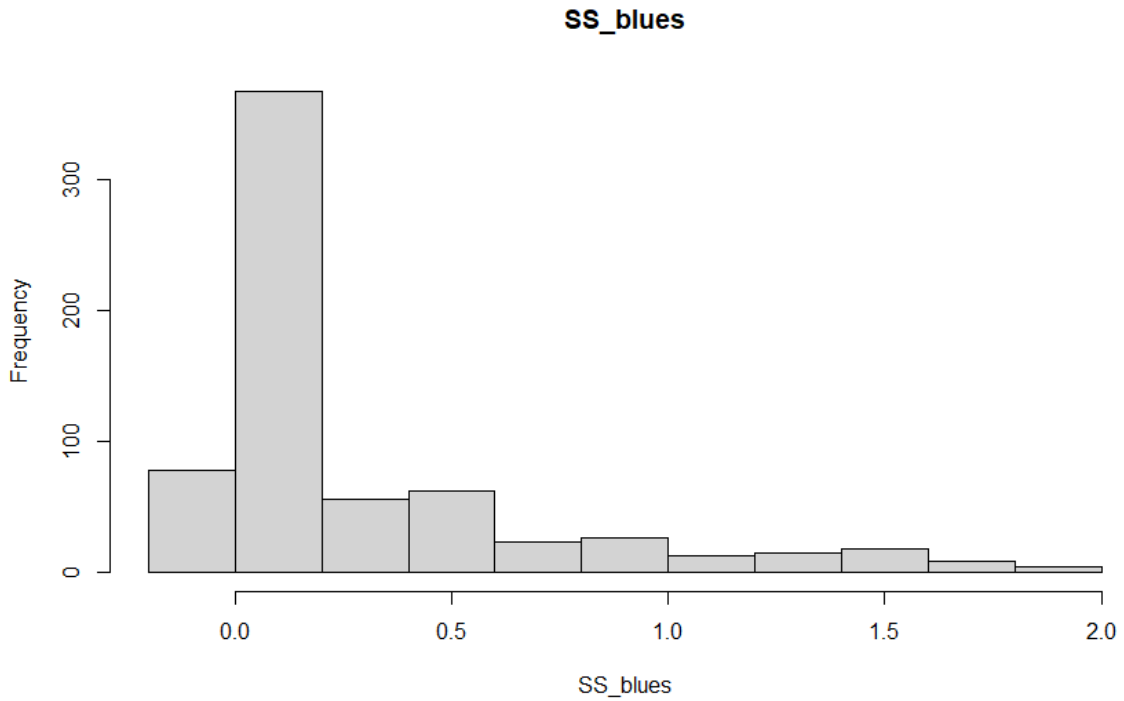


Figure 18 Sterile spikelets distribution frequency for GDP 2021

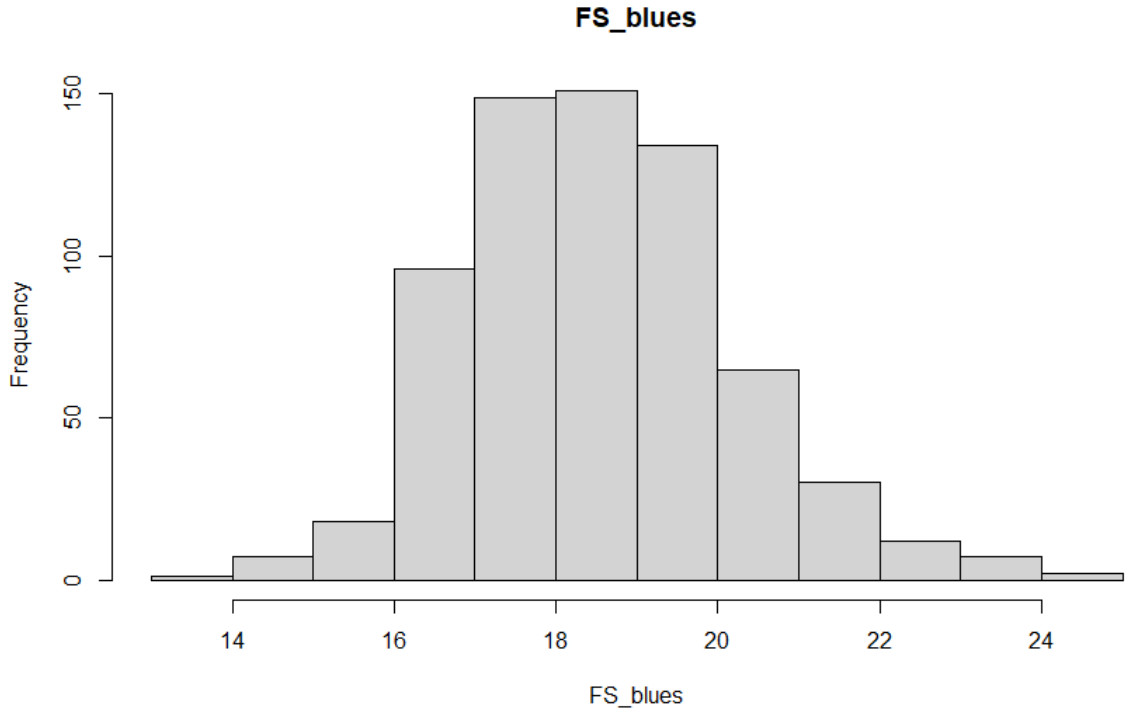


Figure 19 fertile spikelets distribution frequency for GDP 2021

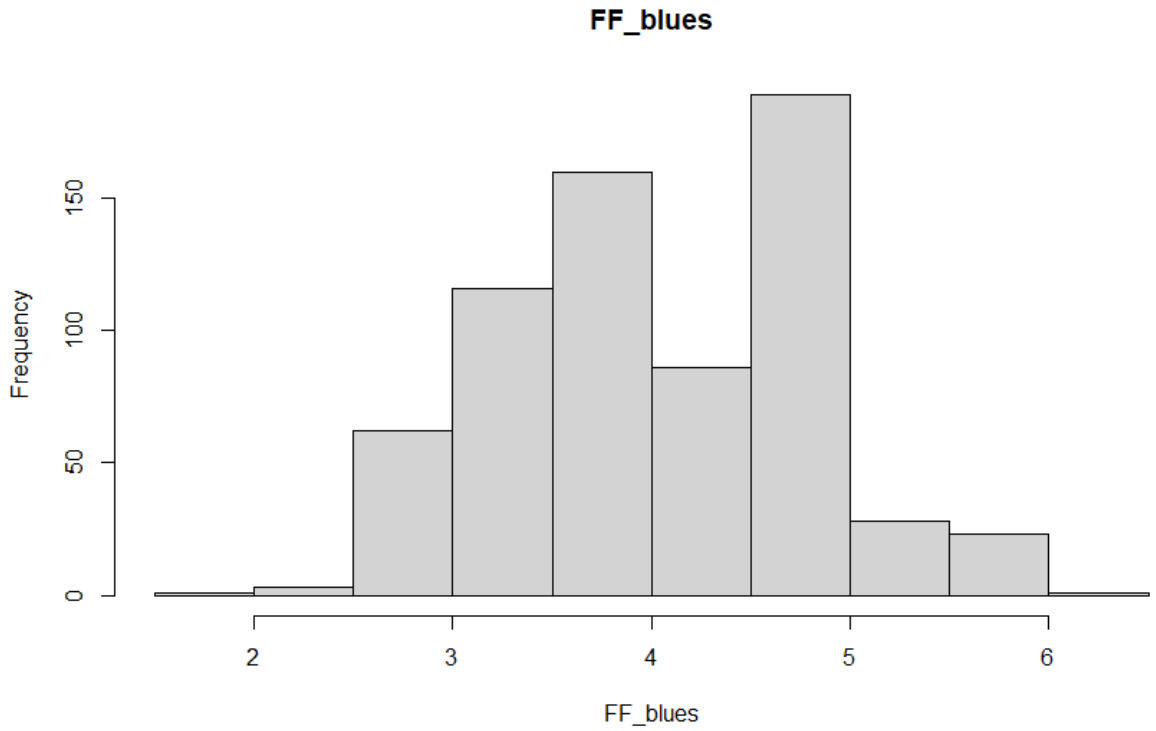


Figure 20 Fertile florets per central spikelet ditribution frequency in the GDP 2021

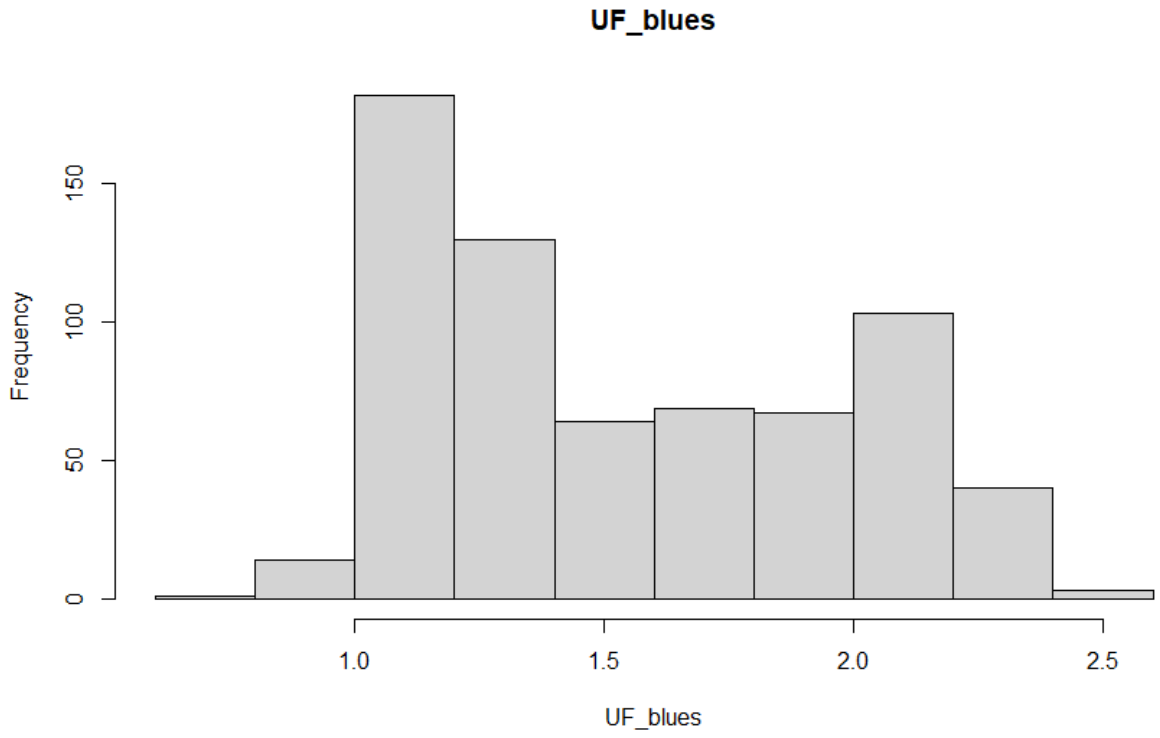


Figure 21 unfertile florets distribution frequency in the GDP 2021

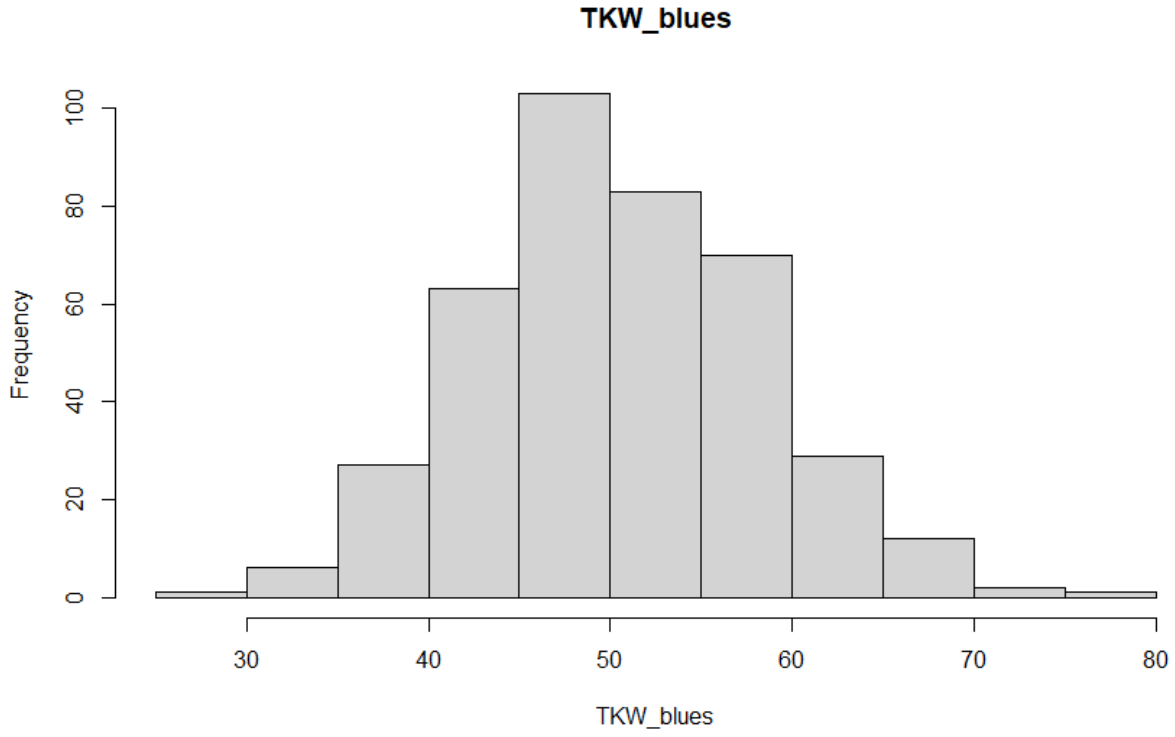


Figure 22 Thousand kernel weight distribution frequency in the GDP 2021

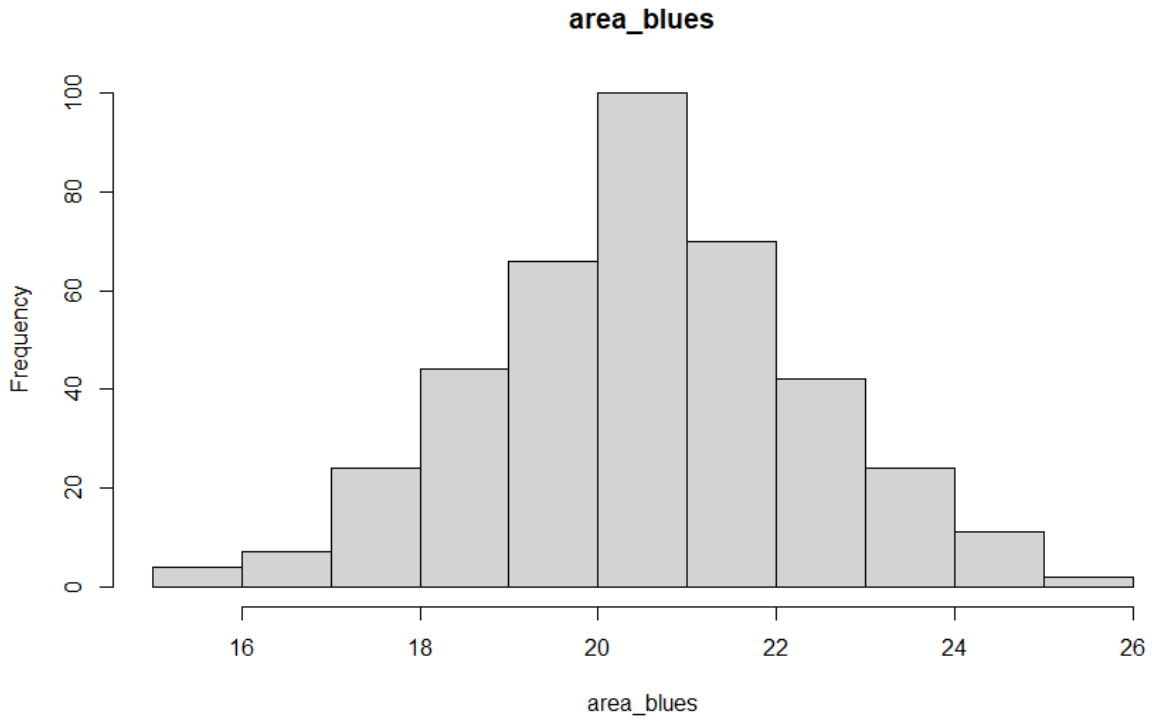


Figure 23 Grain area distribution frequency in the GDP 2021

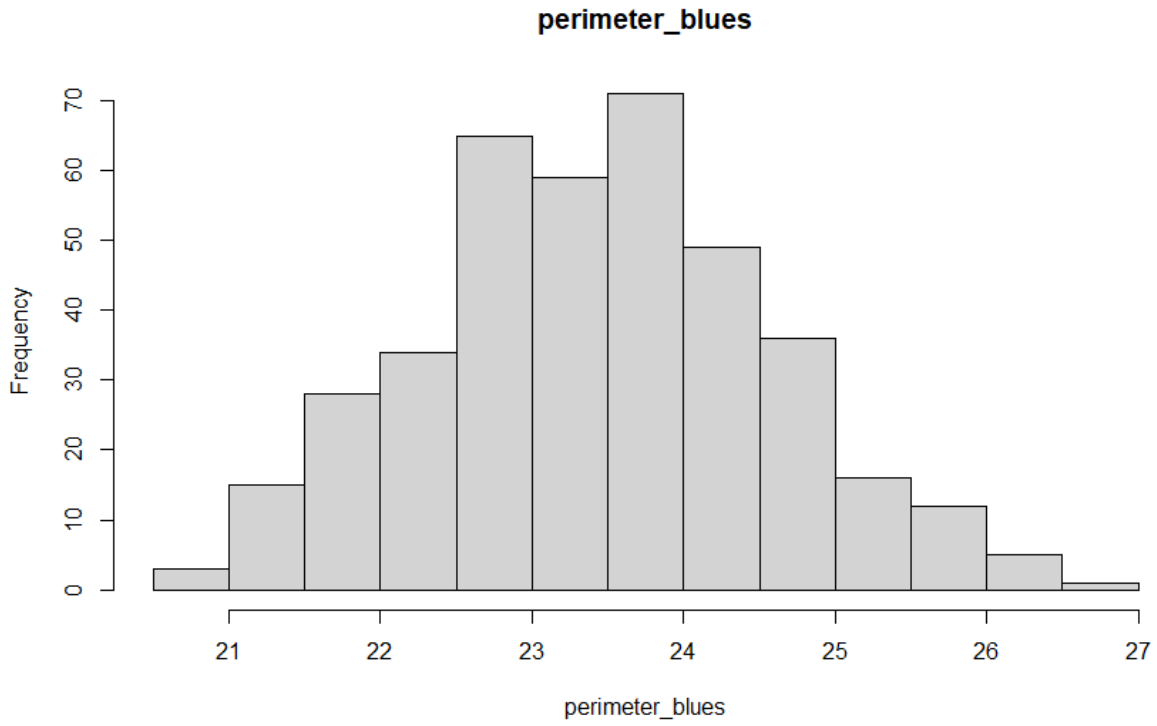


Figure 24 Grain perimeter distribution frequency in the GDP 2021

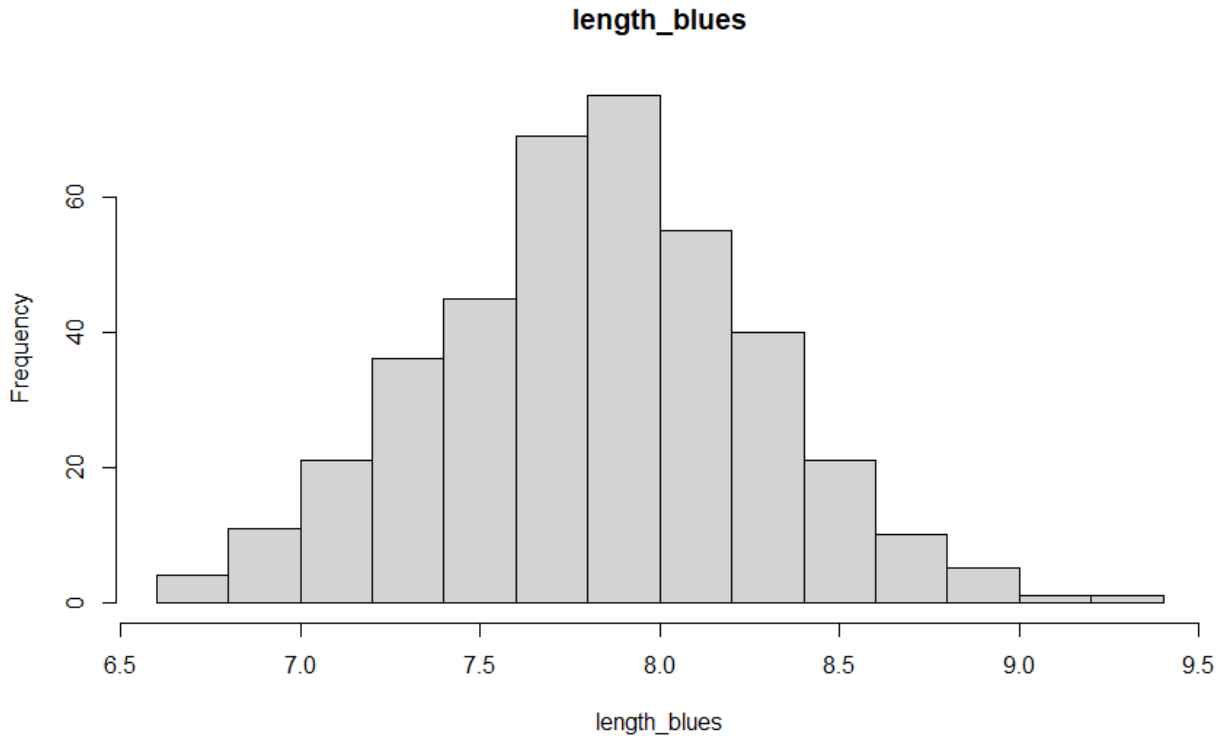


Figure 25 Grain length distribution frequency in the GDP 2021

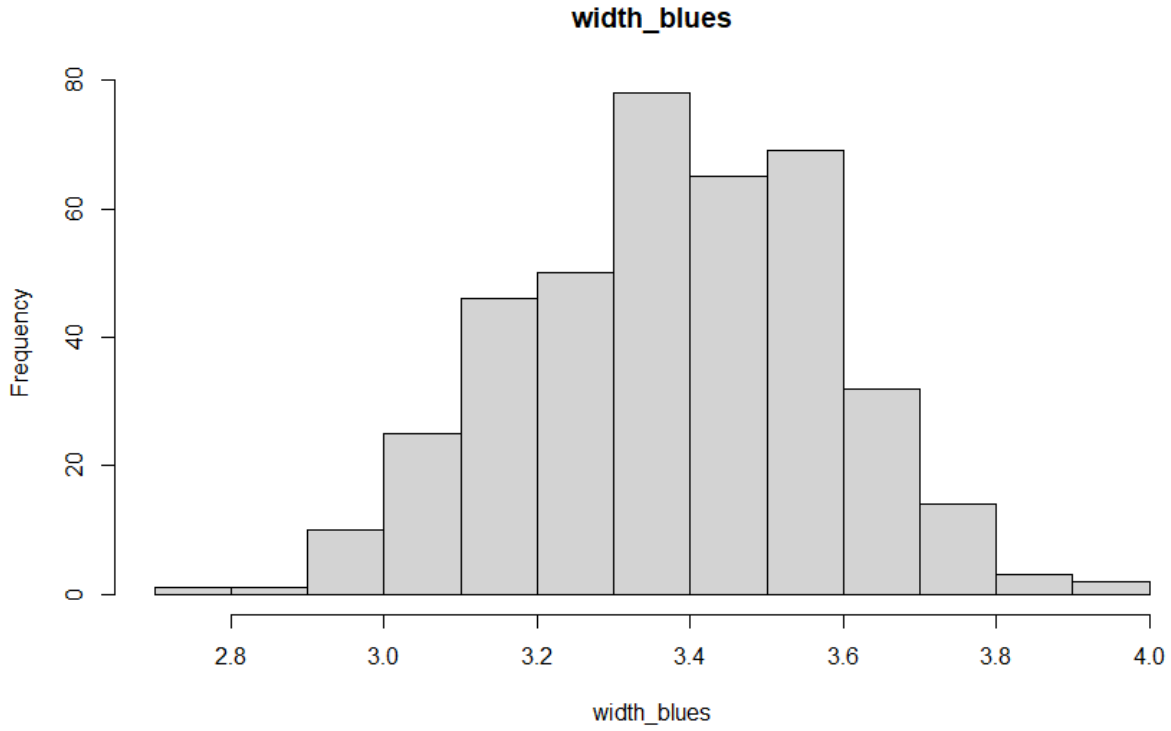


Figure 26 Grain width distribution frequency in the GDP 2021

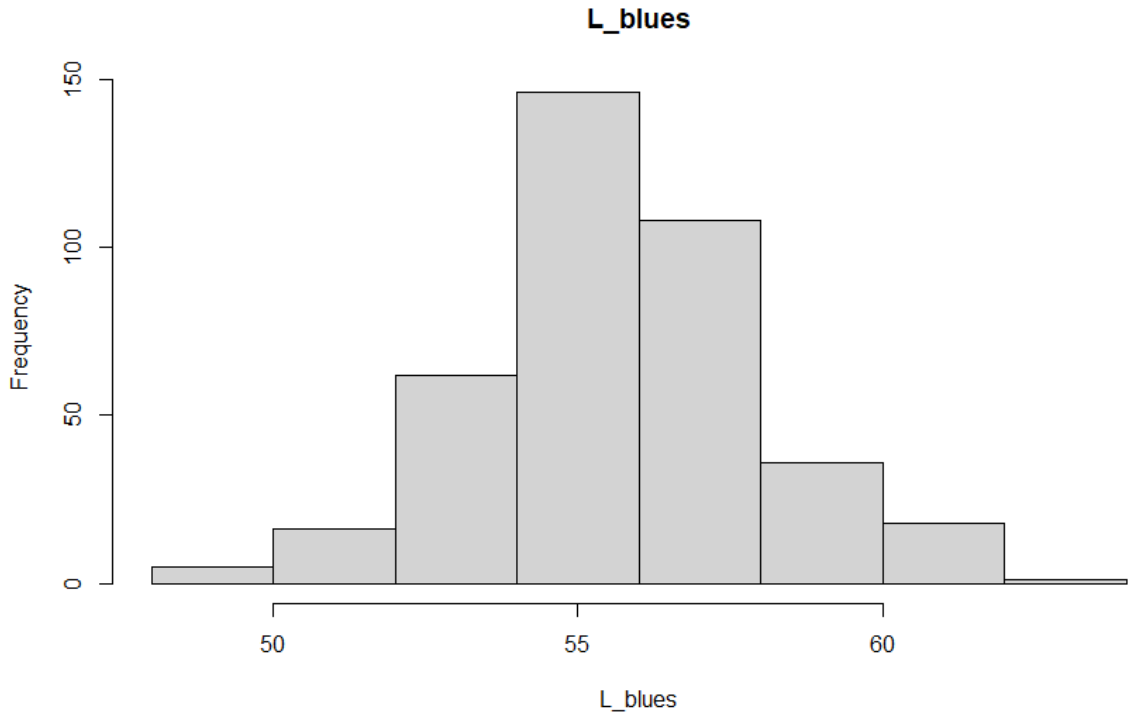


Figure 27 Grain brightness distribution frequency in the GDP 2021

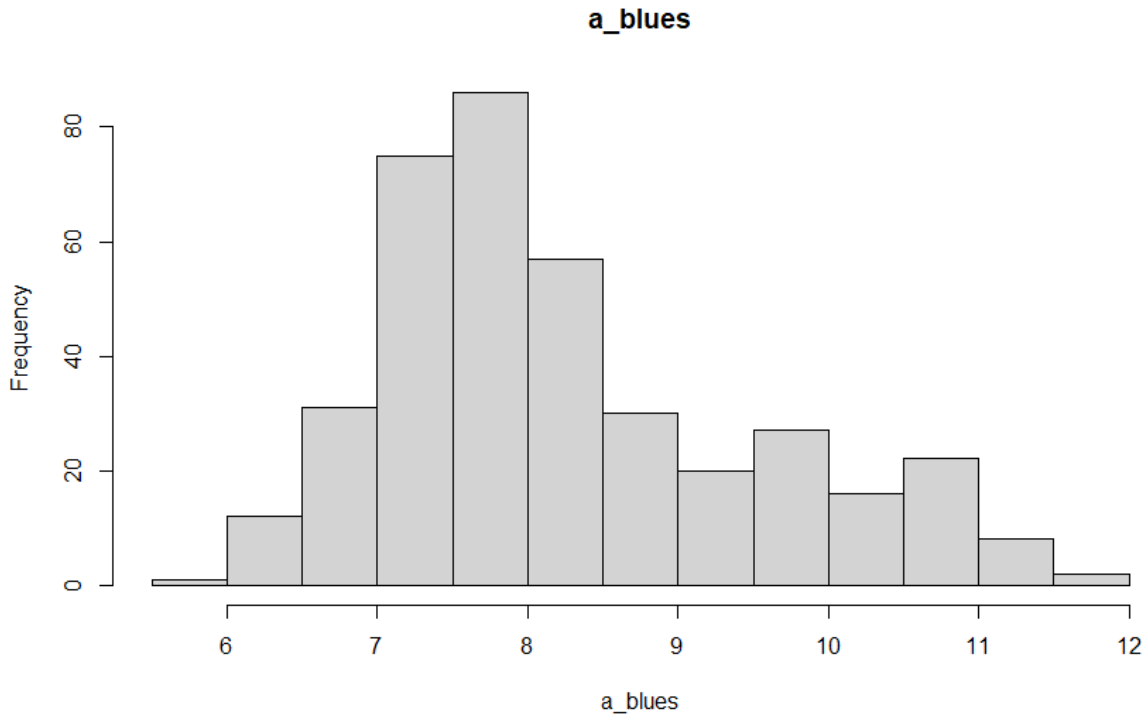


Figure 28 Grain redness distribution frequency in the GDP 2021

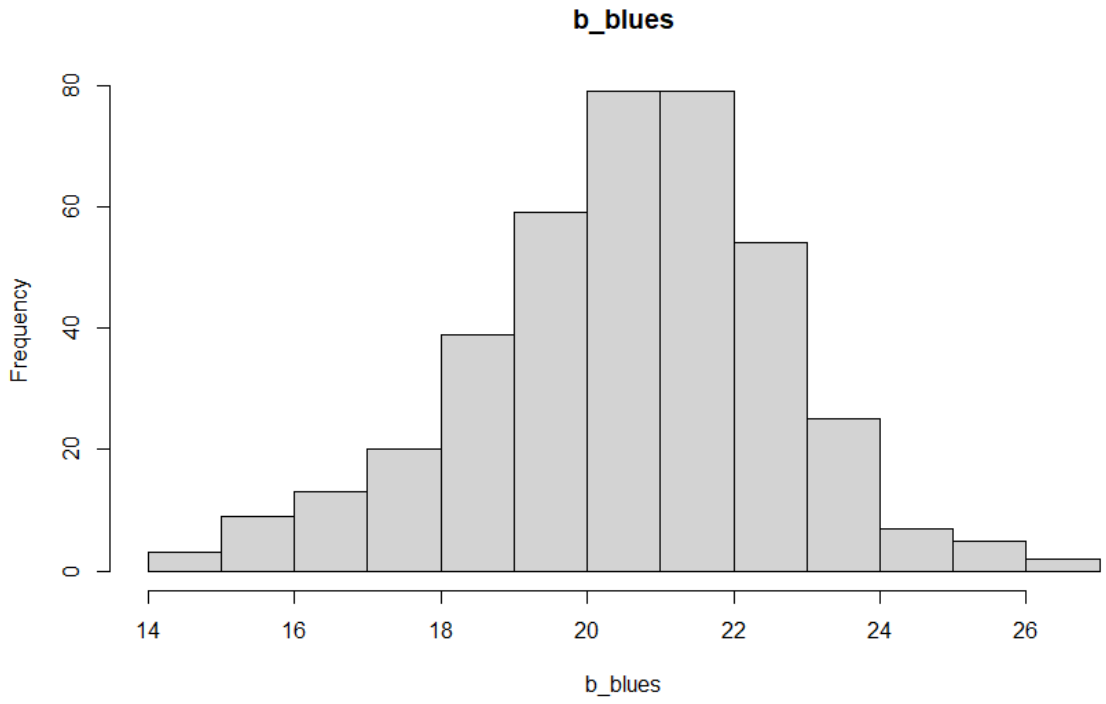


Figure 29 Grain yellowness distribution frequency in the GDP 2021

### 2.3.3 TGC 2019

Table 7 Descriptive statstics for TGC 2019

| | SS | FS | FF | UF |
|---|---|---|---|---|
| **min** | -0.2 | 10.33 | 1.06 | -0.03 |
| **max** | 3.28 | 29.83 | 5.4 | 3.35 |
| **range** | 3.48 | 19.5 | 4.34 | 3.38 |
| **median** | 0.53 | 20 | 2.85 | 1.33 |
| **mean** | 0.69 | 20.21 | 2.82 | 1.37 |
| **SE.mean** | 0.02 | 0.07 | 0.02 | 0.02 |
| **var** | 0.42 | 7.28 | 0.58 | 0.45 |
| **std.dev** | 0.65 | 2.7 | 0.76 | 0.67 |
| **coef.var** | 0.95 | 0.13 | 0.27 | 0.49 |
| $h^2$ | 0.469002 | 0.629284 | 0.681618 | 0.343806 |

Table 8 ANOVA results for TGC 2019

| Trait | Variables | Sum Sq | Df | F value | Pr(>F) | |
|---|---|---|---|---|---|---|
| SS | **Genotype** | 533.21 | 1362 | 1.9314 | 0.0001 | *** |
| SS | **Block** | 5.65 | 17 | 1.6399 | 0.07204 | . |
| SS | **Residuals** | 17.03 | 84 | | | |
| FS | **Genotype** | 9536.4 | 1362 | 2.4122 | 6.45E-07 | *** |
| FS | **Block** | 33.9 | 17 | 0.688 | 0.8061 | |
| FS | **Residuals** | 243.8 | 84 | | | |
| FF | **Genotype** | 534.87 | 1367 | 2.3385 | 1.37E-06 | *** |
| FF | **Block** | 4.71 | 17 | 1.6561 | 0.06817 | . |
| FF | **Residuals** | 14.05 | 84 | | | |
| **UF** | **Genotype** | 465.48 | 1357 | 1.1687 | 0.1813 | |

| | | | | | |
|---|---|---|---|---|---|
| **UF** | **Block** | 5.52 | 17 | 1.107 | 0.3613 |
| **UF** | **Residuals** | 24.65 | 84 | | |

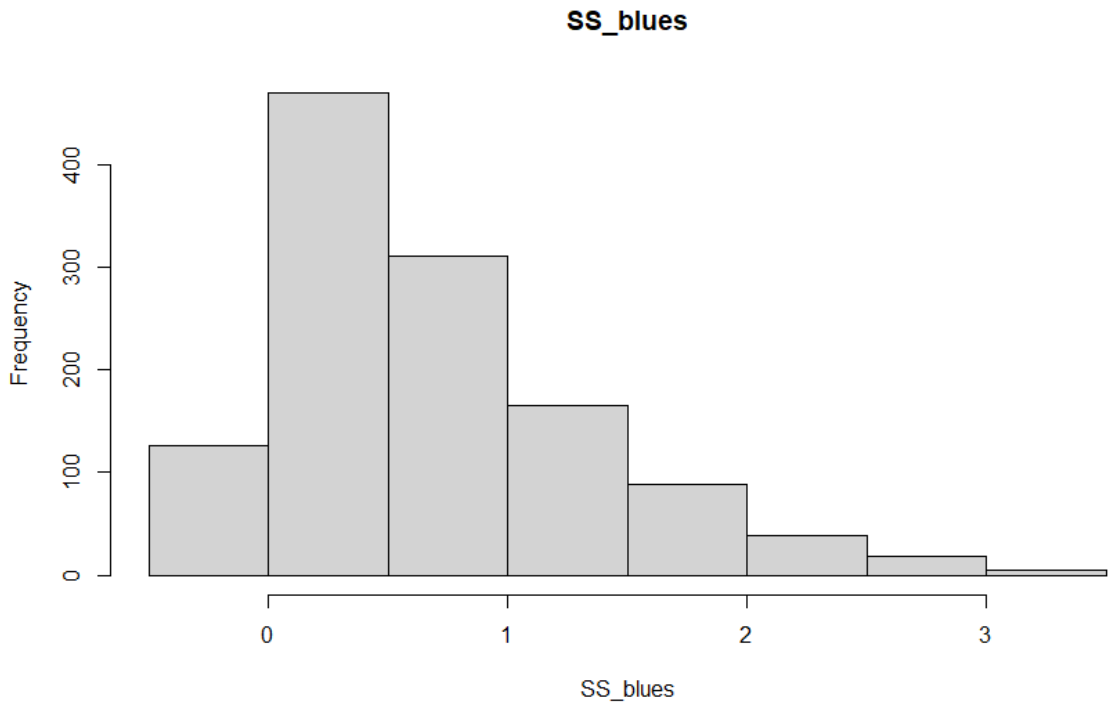


Figure 30 Sterile spikelets distribution frequency in TGC 2019

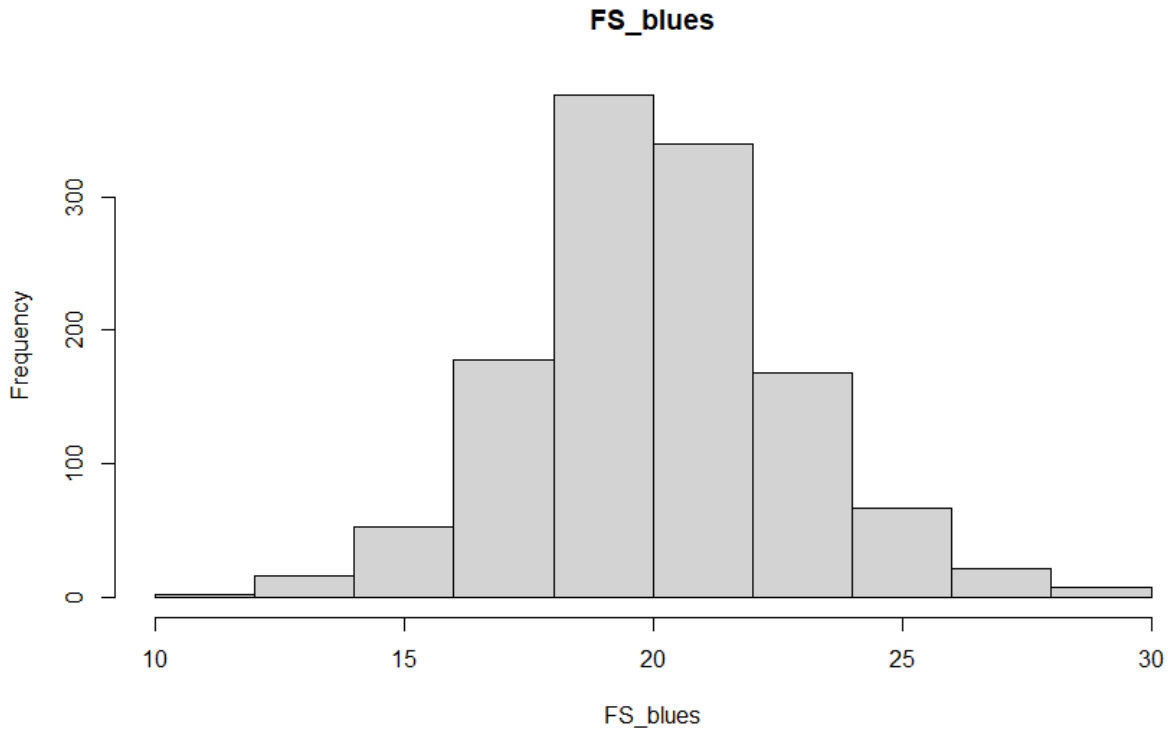


Figure 31 Fertile spikelets distribution frequency in TGC 2019

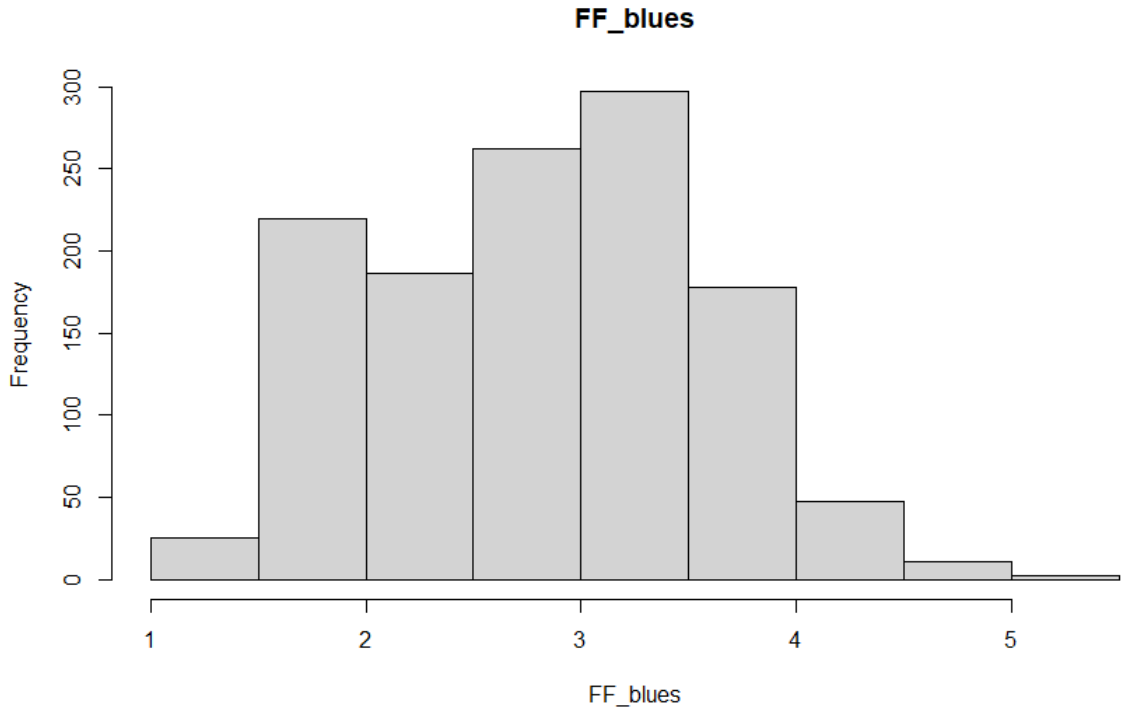


Figure 32 Fertile florets per central spikelet distribution frequency in TGC 2019

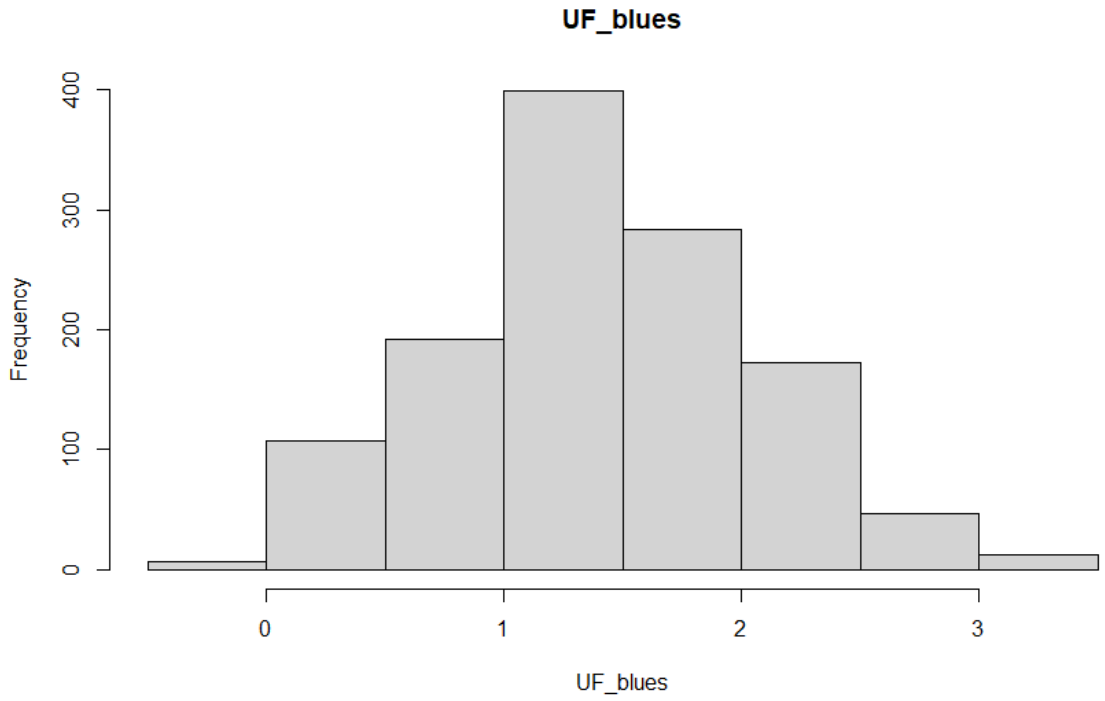


Figure 33 Unfertile florets distribution frequency for TGC 2019

### 2.3.4 TGC 2020

Table 9 Descriptive statistics for TGC 2020 for spike related traits

| | SS | FS | FF | UF |
|---|---|---|---|---|
| **min** | 0 | 11.86 | 1.17 | -0.39 |
| **max** | 3.68 | 25.94 | 4.62 | 2.51 |
| **range** | 3.68 | 14.08 | 3.45 | 2.9 |
| **median** | 1.32 | 18.97 | 2.84 | 0.93 |
| **mean** | 1.32 | 19.03 | 2.83 | 0.94 |
| **SE.mean** | 0.02 | 0.08 | 0.02 | 0.02 |
| **var** | 0.34 | 5.8 | 0.4 | 0.23 |

| | | | | |
|---|---|---|---|---|
| **std.dev** | 0.59 | 2.41 | 0.63 | 0.48 |
| **coef.var** | 0.45 | 0.13 | 0.22 | 0.51 |
| $h^2$ | 0.626325 | 0.504504 | 0.23892 | 0.228585 |

Table 10 Dscriptive statistics for grain related traits in TGC 2020

| | TKW | Area | Perimeter | Length | Width | L* | a* | b* |
|---|---|---|---|---|---|---|---|---|
| **min** | 20.85 | 10.7 | 18.66 | 4.94 | 2.41 | 48.7 | 4.1 | 11.1 |
| **max** | 93.24 | 28.67 | 29.77 | 10.64 | 4.05 | 71.41 | 13.84 | 27.38 |
| **range** | 72.4 | 17.97 | 11.11 | 5.7 | 1.64 | 22.7 | 9.74 | 16.28 |
| **median** | 55.17 | 21.01 | 24.04 | 7.91 | 3.41 | 59.22 | 7.65 | 19.83 |
| **mean** | 55.18 | 20.77 | 24.06 | 7.92 | 3.37 | 59.14 | 8.02 | 19.34 |
| **SE.mean** | 0.35 | 0.1 | 0.06 | 0.03 | 0.01 | 0.12 | 0.06 | 0.1 |
| **var** | 100.11 | 8.21 | 3.08 | 0.53 | 0.08 | 12.46 | 2.86 | 8.75 |
| **std.dev** | 10.01 | 2.87 | 1.76 | 0.73 | 0.29 | 3.53 | 1.69 | 2.96 |
| **coef.var** | 0.18 | 0.14 | 0.07 | 0.09 | 0.08 | 0.06 | 0.21 | 0.15 |
| $h^2$ | 0.865192 | 0.924448 | 0.92307 | 0.942496 | 0.923175 | 0.838604 | 0.84798 | 0.903673 |

Table 11 ANOVA results in TGC 2020

| Trait | Variables | Sum Sq | Df | F value | Pr(>F) | |
|---|---|---|---|---|---|---|
| SS | **Genotype** | 303.172 | 943 | 2.0645 | 0.00054 | *** |
| SS | **Block** | 0.111 | 4 | 0.1774 | 0.949123 | |
| SS | **Heading date** | 0.733 | 11 | 0.4281 | 0.936879 | |
| SS | **Residuals** | 8.565 | 55 | | | |
| FS | **Genotype** | 3484.3 | 943 | 1.4871 | 0.03226 | * |
| FS | **Block** | 3.6 | 4 | 0.3589 | 0.83677 | |
| FS | **Heading date** | 10.6 | 11 | 0.3871 | 0.95567 | |
| FS | **Residuals** | 136.7 | 55 | | | |
| FF | **Genotype** | 321.04 | 942 | 1.1273 | 0.2943 | |
| FF | **Block** | 0.8 | 4 | 0.6589 | 0.6232 | |
| FF | **Heading date** | 2.25 | 11 | 0.6777 | 0.7533 | |
| FF | **Residuals** | 16.63 | 55 | | | |
| UF | **Genotype** | 125.92 | 938 | 1.1339 | 0.2866 | |
| UF | **Block** | 0.288 | 4 | 0.6083 | 0.6584 | |
| UF | **Heading date** | 1.151 | 11 | 0.8836 | 0.5614 | |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| UF | **Residuals** | 6.393 | 54 | | | | |
| TKW | **Genotype** | 57866 | 785 | 5.8664 | | 1.58E-10 | *** |
| TKW | **Block** | 123 | 4 | 2.4497 | | 0.06006 | . |
| TKW | **Heading date** | 62 | 11 | 0.4486 | | 0.92418 | |
| TKW | **Residuals** | 553 | 44 | | | | |
| Area | **Genotype** | 5011 | 809 | 12.4885 | | <2e-16 | *** |
| Area | **Block** | 1.1 | 4 | 0.5445 | | 0.7039 | |
| Area | **Heading date** | 9.4 | 11 | 1.724 | | 0.0971 | . |
| Area | **Residuals** | 23.3 | 47 | | | | |
| Perimeter | **Genotype** | 2133.71 | 811 | 13.6522 | | 2.00E-16 | *** |
| Perimeter | **Block** | 0.24 | 4 | 0.3147 | | 0.86671 | |
| Perimeter | **Heading date** | 4.25 | 11 | 2.0062 | | 0.04914 | * |
| Perimeter | **Residuals** | 9.06 | 47 | | | | |
| Length | **Genotype** | 372.91 | 812 | 18.4639 | | 2.00E-16 | *** |
| Length | **Block** | 0.04 | 4 | 0.445 | | 0.77545 | |
| Length | **Heading date** | 0.61 | 11 | 2.2305 | | 0.02831 | * |
| Length | **Residuals** | 1.17 | 47 | | | | |
| Width | **Genotype** | 42.969 | 810 | 10.3004 | | 3.94E-16 | *** |
| Width | **Block** | 0.044 | 4 | 2.1455 | | 0.08991 | . |
| Width | **Heading date** | 0.06 | 11 | 1.0673 | | 0.40691 | |
| Width | **Residuals** | 0.242 | 47 | | | | |
| L* | **Genotype** | 8177.6 | 790 | 6.9006 | 1.58E-12 | | *** |
| L* | **Block** | 16.2 | 4 | 2.6957 | 0.042 | | * |
| L* | **Heading date** | 31.6 | 11 | 1.9165 | 0.06115 | | . |
| L* | **Residuals** | 70.5 | 47 | | | | |
| A* | **Genotype** | 1794.49 | 811 | 7.4128 | 3.67E-13 | | *** |
| A* | **Block** | 2 | 4 | 1.671 | 0.1725 | | |
| A* | **Heading date** | 5.32 | 11 | 1.6212 | 0.1237 | | |
| A* | **Residuals** | 14.03 | 47 | | | | |
| B* | **Genotype** | 5630.9 | 812 | 7.7513 | 1.48E-13 | | *** |
| B* | **Block** | 2.5 | 4 | 0.7035 | 0.5935 | | |
| B* | **Heading date** | 8.3 | 11 | 0.8451 | 0.5974 | | |
| B* | **Residuals** | 42 | 47 | | | | |

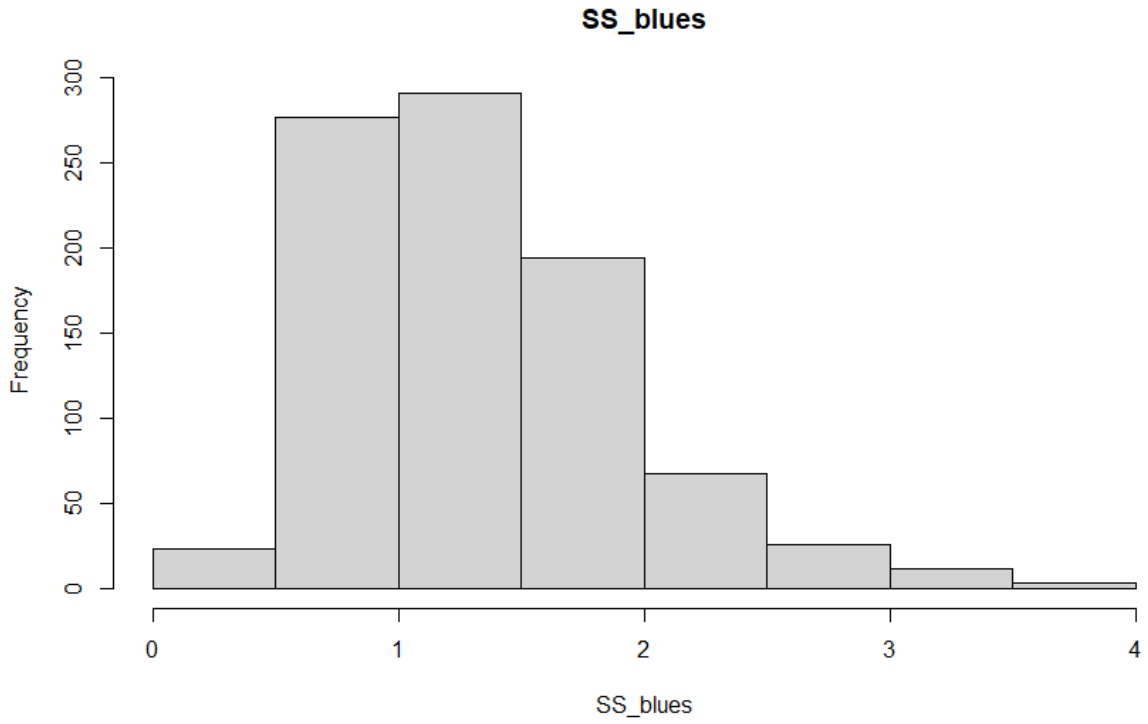


Figure 34Sterile spikelelts distribution frequency in TGC 2020

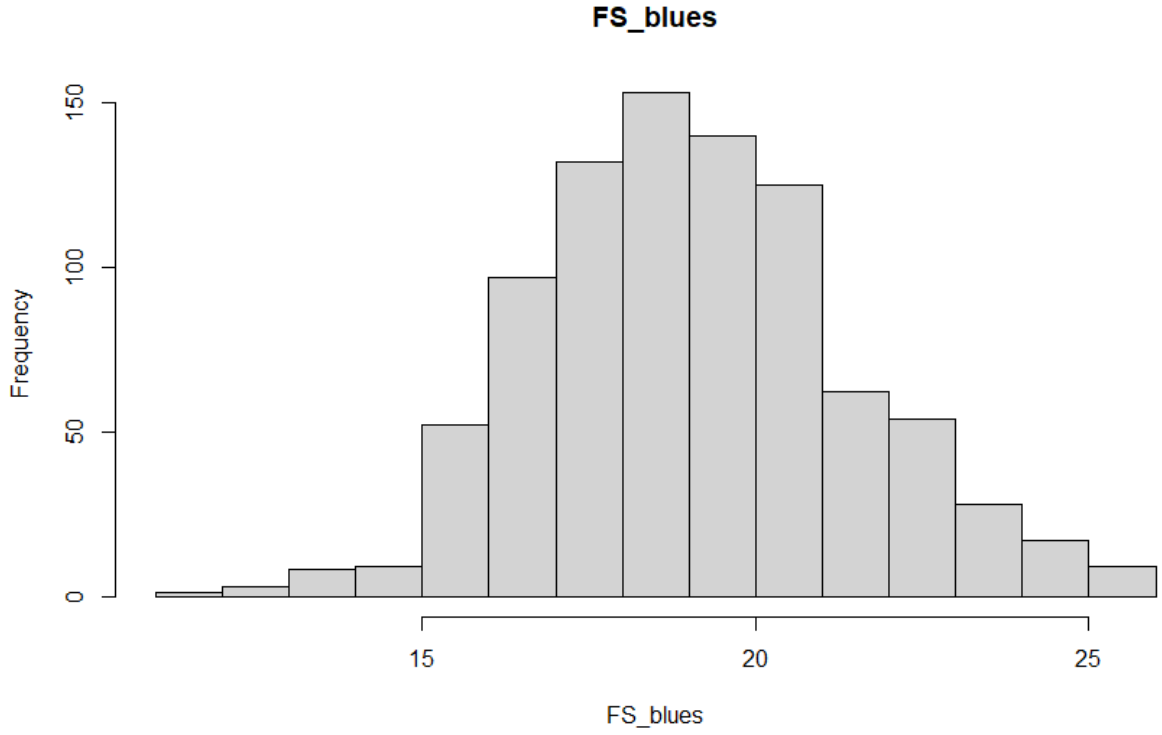


Figure 35 Fertile spikelets distribution frequency in TGC 2020

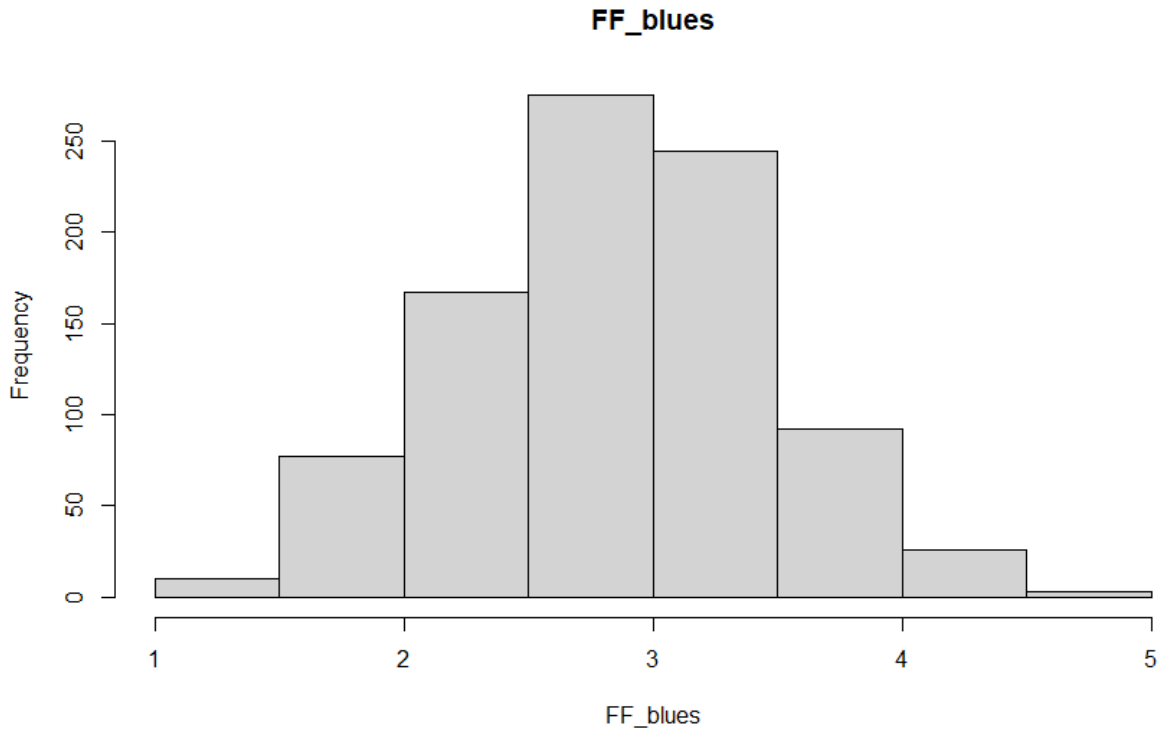


Figure 36 Fertile florets per central spikelet distribution frequency in TGC 2020

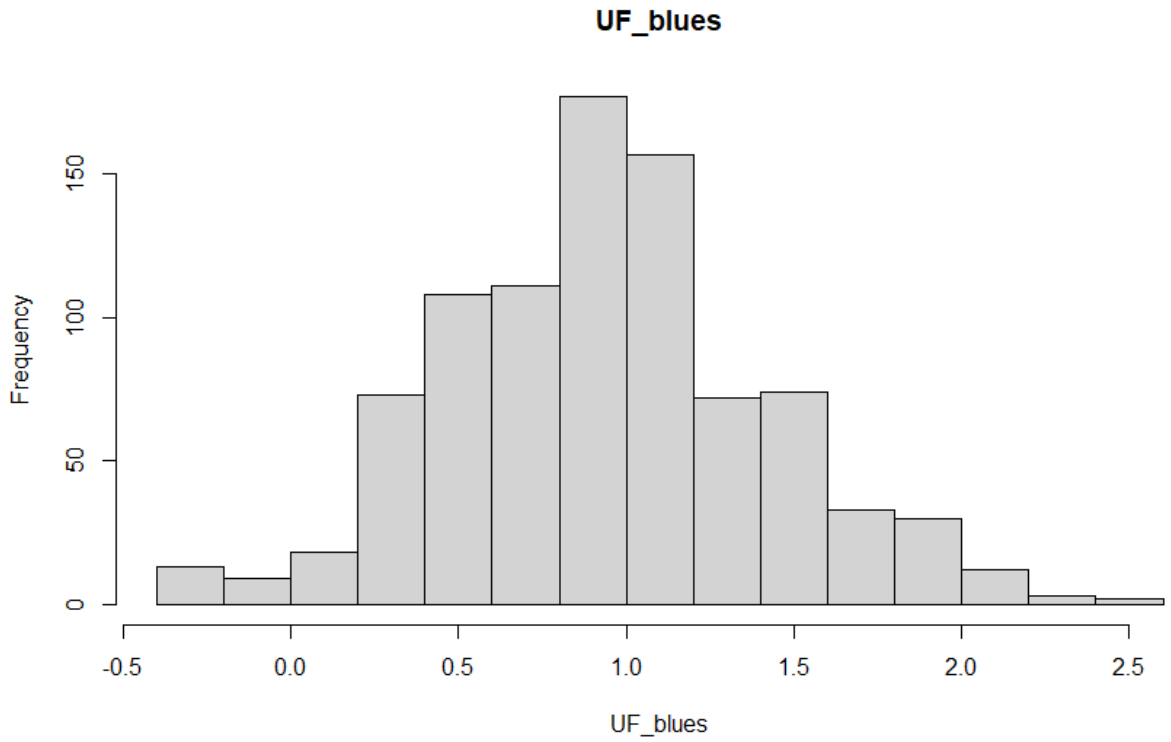


Figure 37 Unfertile florets ditribution frequency in TGC 2020

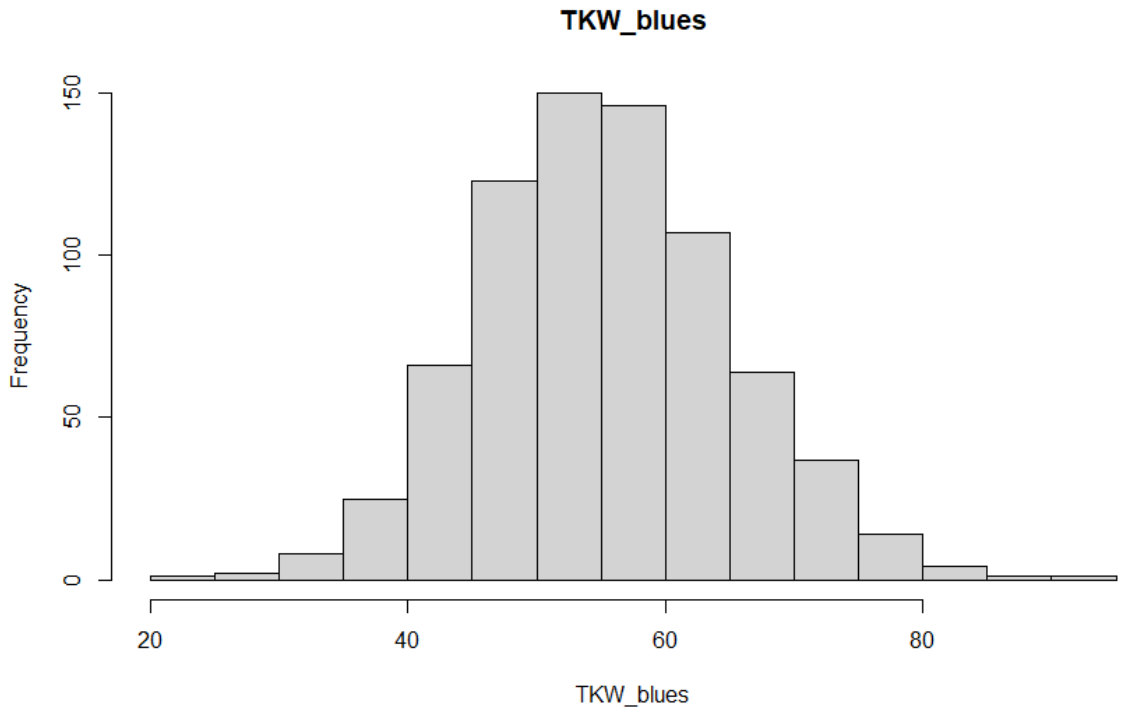


Figure 38 Thousand kernel weight distribution frequency in TGC 2020

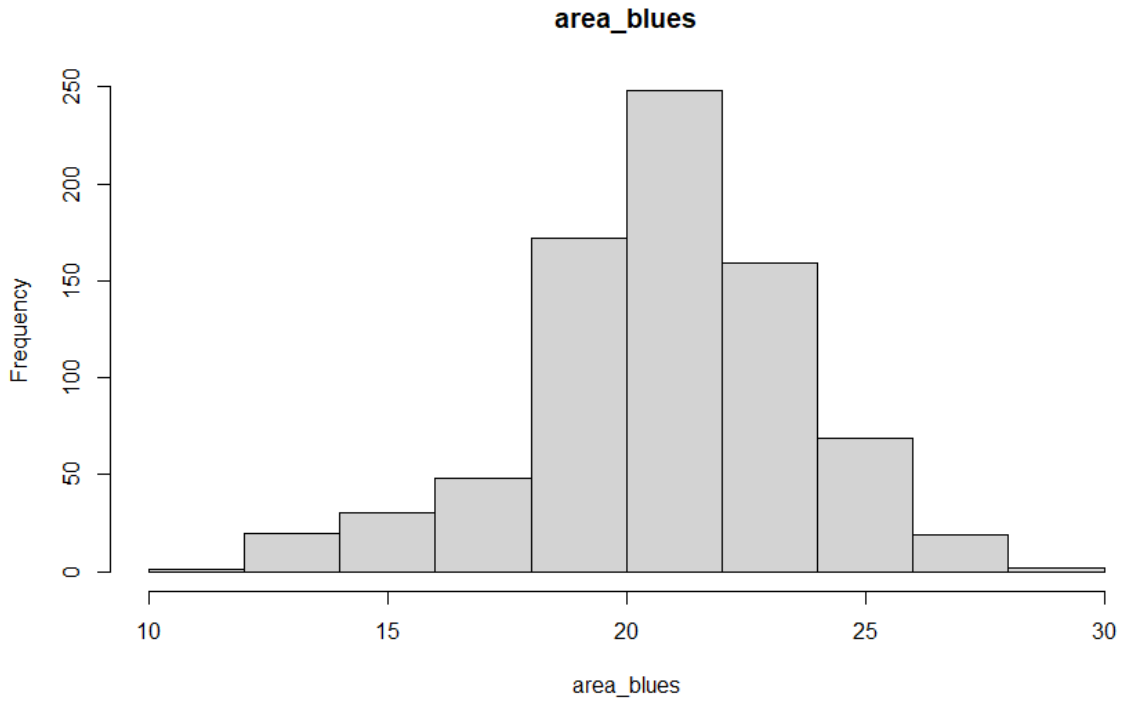


Figure 39 Grain area distribution frequency in TGC 2020

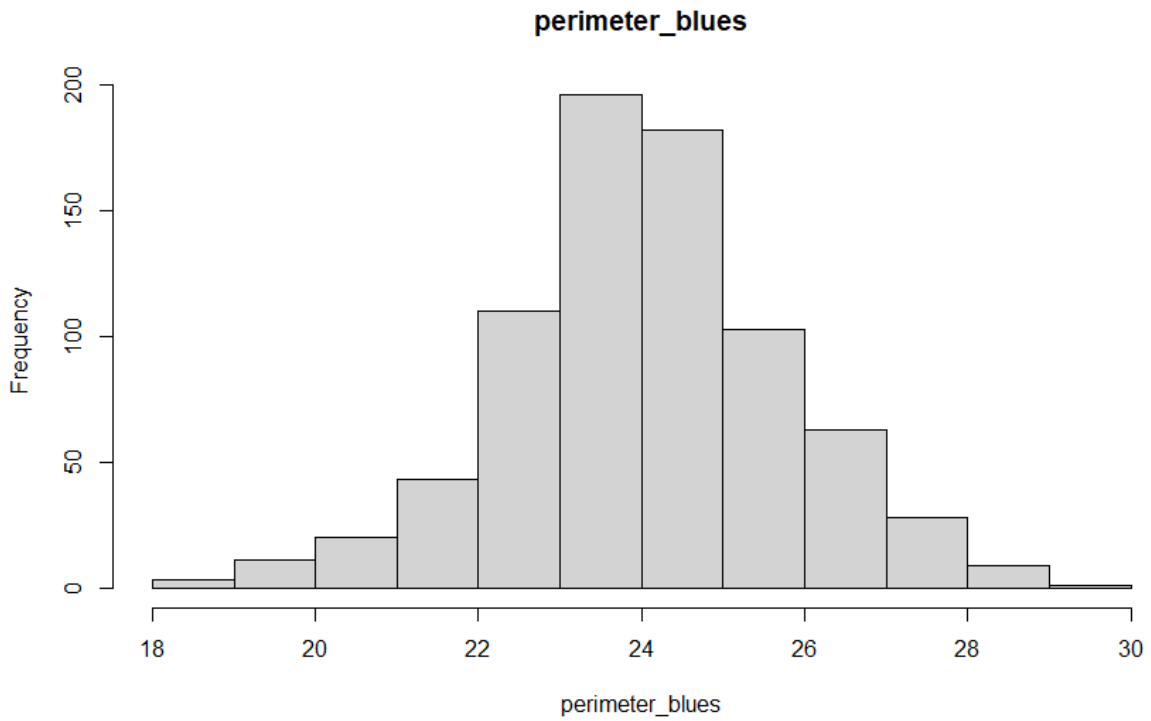


Figure 40 Grain perimeter distribution frequency in TGC 2020

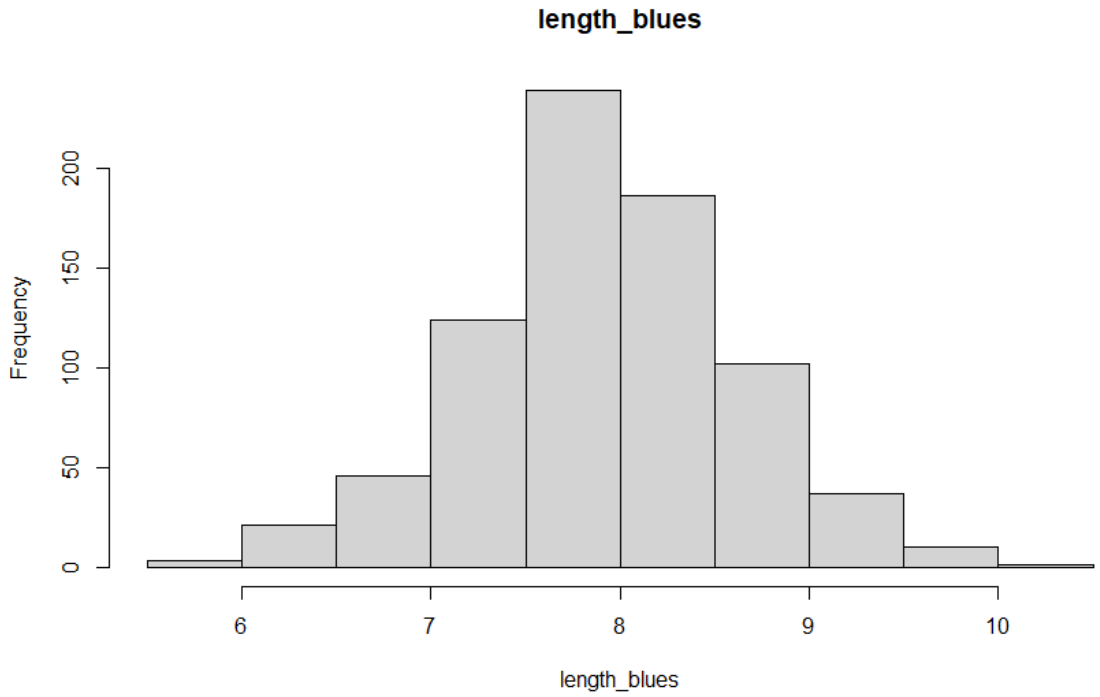


Figure 41 Grain length distribution frequency in TGC 2020

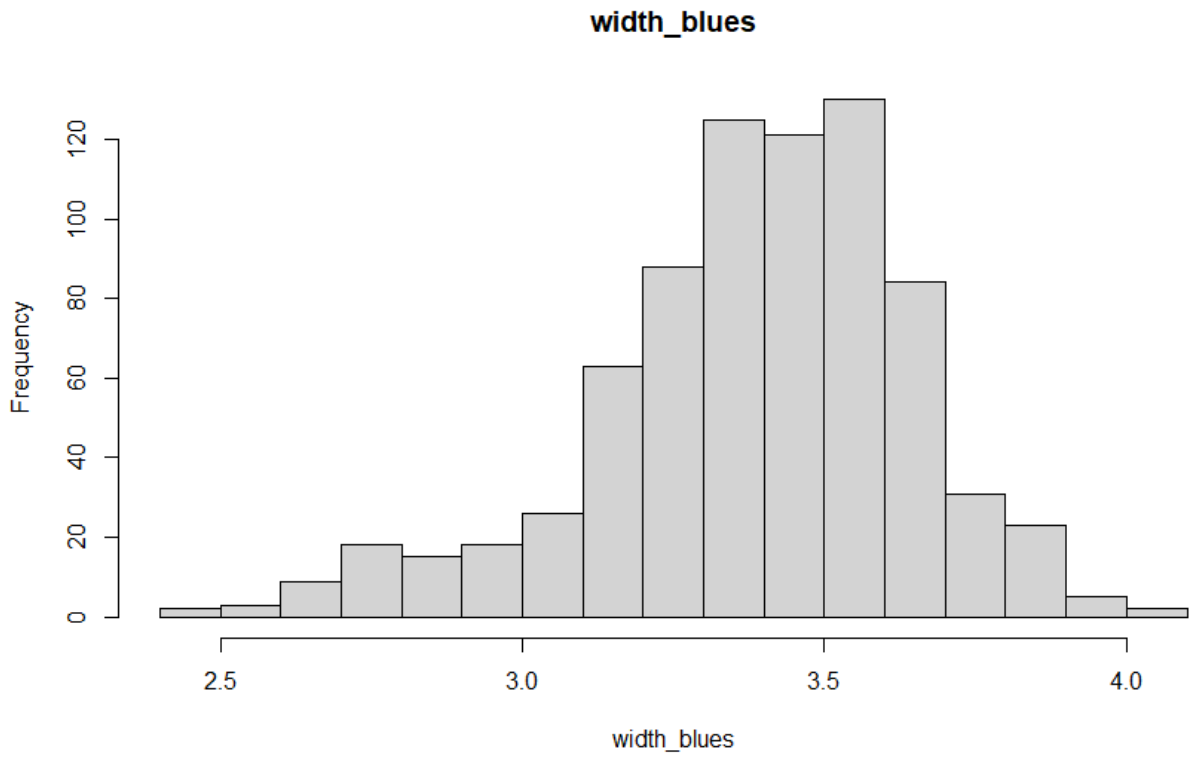


Figure 42 Grain width distribution frequency in TGC 2020

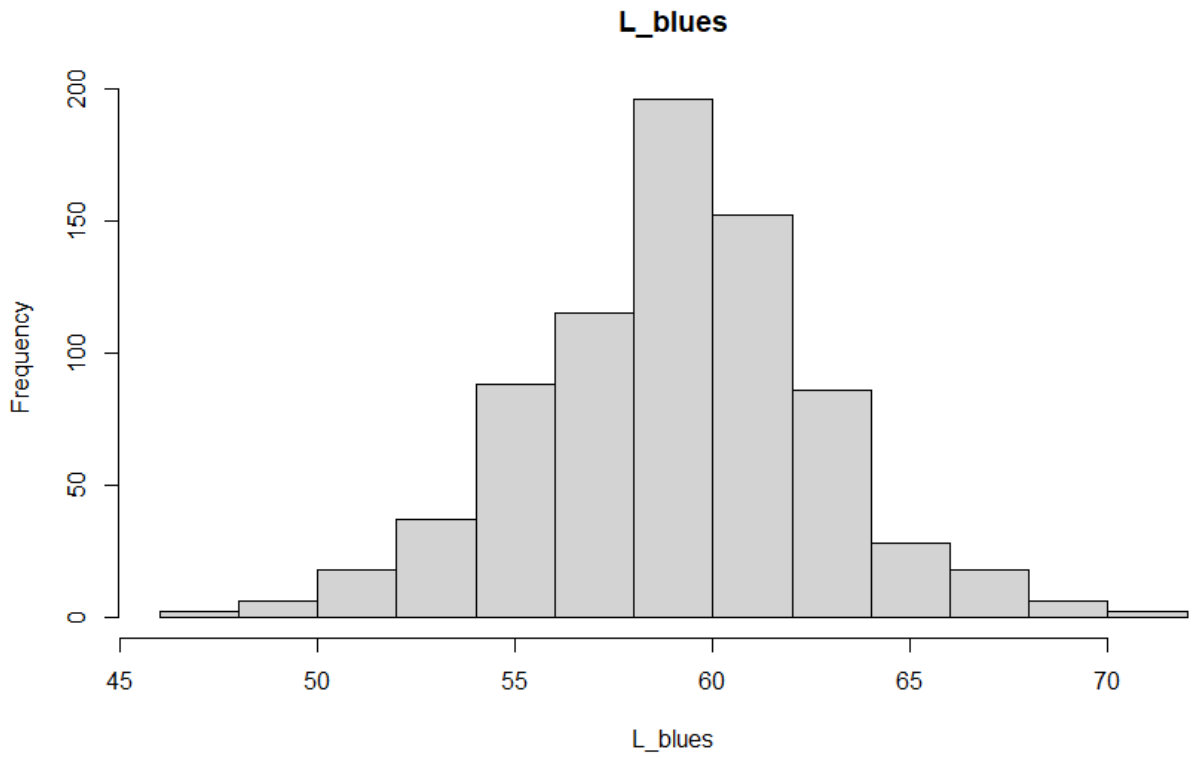


Figure 43 Grain brightness distribution frequency in TGC 2020

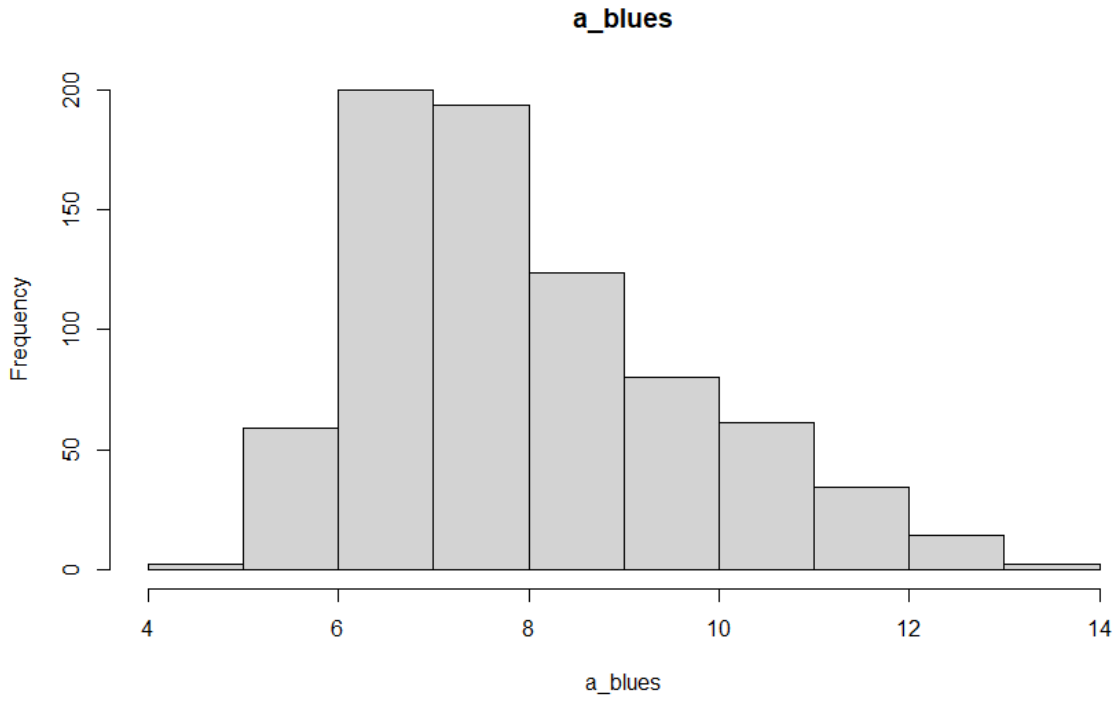


Figure 44 Grain redness distribution frequency in TGC 2020

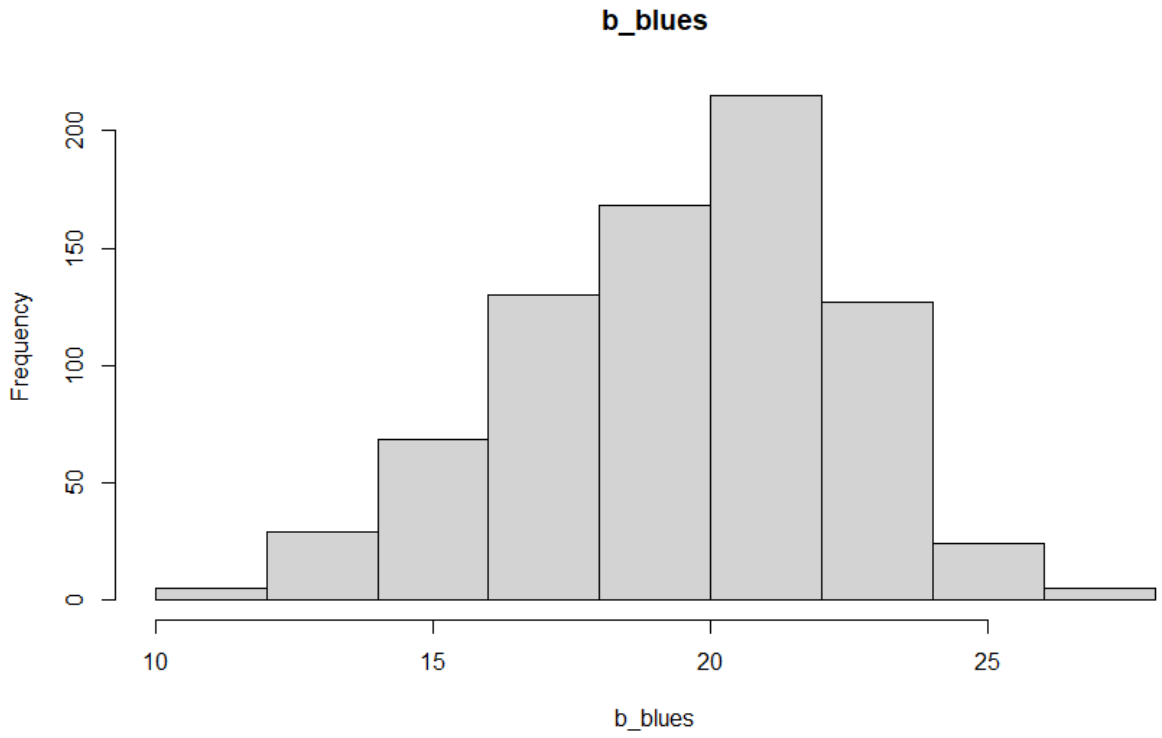


Figure 45 Grain yellowness distribution frequency in TGC 2020

### 2.3.5 Single Environments summary

Differences among each single environment were investigated herein. In order to accomplish this task, heritability scores, ANOVA results and distribution of frequencies were considered. As regards heritability, GDP 2021 showed higher values for every trait when compared to GDP 2020, while TGC 2019 showed the same trend compared to TGC 2020.

It can be noticed that TGC 2020 and GDP 2020, both grown in Grosseto during the same season, show a lower heritability probably due to the harsh drought conditions that occurred on that specific season.

ANOVA showed significance in:

- Grain width in GDP Grosseto 2020, Fertile spikelet number, fertile floret number per central spikelet and grain brightness in GDP Grosseto 2021 and grain brightness in TGC 2020 for blocks
- Grain area, perimeter, length and width in GDP 2020 and grain perimeter and width in TGC 2020 for heading date

Histogram distribution showed strong asymmetrical trend especially for the sterile spikelet number trait across all environments. Unfertile florets number per central spikelet also showed a similar trend, while the a* trait which is related to the redness of the grain, showed a bimodal trend in both season of GDP

### 2.3.6 GDP 2020 and GDP 2021

Table 12 Descriptive statistics for spike related traits in GDP cluster

| | SS | FS | FF | UF |
|---|---|---|---|---|
| **min** | 0.95 | 18.07 | 1.41 | 0.14 |
| **max** | 4.67 | 26.9 | 5.18 | 2.28 |
| **range** | 3.71 | 8.83 | 3.76 | 2.13 |
| **median** | 1.89 | 22.17 | 3.27 | 1.15 |
| **mean** | 2.07 | 22.3 | 3.23 | 1.18 |
| **SE.mean** | 0.02 | 0.06 | 0.03 | 0.01 |
| **var** | 0.21 | 2.36 | 0.47 | 0.12 |
| **std.dev** | 0.46 | 1.54 | 0.69 | 0.35 |
| **coef.var** | 0.22 | 0.07 | 0.21 | 0.29 |
| $h^2$ | 0.708575 | 0.299664 | 0.61306 | 0 |

Table 13 Descriptive statistics for grain related traits in GDP cluster

| | TKW | Area | Perimeter | Length | Width | L* | a* | b* |
|---|---|---|---|---|---|---|---|---|
| **min** | 19.09 | 11.65 | 18.21 | 5.82 | 2.63 | 42.61 | 6.66 | 17.78 |
| **max** | 71.21 | 24.51 | 25.04 | 9.11 | 3.99 | 59.31 | 13.12 | 27.21 |
| **range** | 52.12 | 12.87 | 6.83 | 3.29 | 1.37 | 16.7 | 6.46 | 9.43 |
| **median** | 44.56 | 18.61 | 21.92 | 7.13 | 3.31 | 51.14 | 9.6 | 23.69 |
| **mean** | 44.72 | 18.57 | 21.88 | 7.12 | 3.31 | 51.21 | 9.54 | 23.48 |
| **SE.mean** | 0.3 | 0.07 | 0.04 | 0.02 | 0.01 | 0.07 | 0.04 | 0.06 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **var** | 64.93 | 3.75 | 1.13 | 0.2 | 0.05 | 3.88 | 1.26 | 2.91 |
| **std.dev** | 8.06 | 1.94 | 1.06 | 0.45 | 0.21 | 1.97 | 1.12 | 1.71 |
| **coef.var** | 0.18 | 0.1 | 0.05 | 0.06 | 0.06 | 0.04 | 0.12 | 0.07 |
| $h^2$ | 0.566096 | 0.592654 | 0.663375 | 0.780284 | 0.568512 | 0.530054 | 0.044489 | 0.794221 |

Phenotypic traits in this environmental cluster show lower values compared to each single environment, this is especially true for the fertile spikelet number (FS)

Table 14 ANOVA results for GDP cluster

| Trait | Variables | Sum Sq | Df | F value | Pr(>F) | |
|---|---|---|---|---|---|---|
| SS | **Genotype** | 271.011 | 744 | 8.8938 | <2e-16 | *** |
| SS | **Environment** | 10.135 | 1 | 247.4647 | <2e-16 | *** |
| SS | **Block** | 0.312 | 10 | 0.7629 | 0.6647 | |
| SS | **Heading date** | 0.636 | 14 | 1.11 | 0.3471 | |
| SS | **Genotype x Environment** | 83.004 | 658 | 3.08 | <2e-16 | *** |
| SS | **Residuals** | 14.826 | 362 | | | |
| FS | **Genotype** | 3185.8 | 741 | 4.984 | 2.00E-16 | *** |
| FS | **Environment** | 1176.9 | 1 | 1364.343 | 2.00E-16 | *** |
| FS | **Block** | 17.4 | 10 | 2.0157 | 0.03089 | * |
| FS | **Heading date** | 11.5 | 14 | 0.9537 | 0.50094 | |
| FS | **Genotype x Environment** | 1258.3 | 657 | 2.2203 | 2.00E-16 | *** |
| FS | **Residuals** | 312.3 | 362 | | | |
| FF | **Genotype** | 697.4 | 743 | 7.5943 | 2.00E-16 | *** |
| FF | **Environment** | 83.06 | 1 | 672.002 | 2.00E-16 | *** |
| FF | **Block** | 3.17 | 10 | 2.5664 | 0.00522 | ** |
| FF | **Heading date** | 2.21 | 14 | 1.2775 | 0.21865 | |
| FF | **Genotype x Environment** | 196.18 | 660 | 2.405 | 2.00E-16 | *** |
| FF | **Residuals** | 44.74 | 362 | | | |
| UF | **Genotype** | 175.091 | 744 | 2.5136 | <2e-16 | *** |
| UF | **Environment** | 57.089 | 1 | 609.7572 | <2e-16 | *** |
| UF | **Block** | 1.029 | 10 | 1.0992 | 0.3616 | |
| UF | **Heading date** | 1.178 | 14 | 0.8987 | 0.5605 | |
| UF | **Genotype x Environment** | 138.172 | 655 | 2.2531 | <2e-16 | *** |
| UF | **Residuals** | 33.893 | 362 | | | |
| TKW | **Genotype** | 65604 | 725 | 3.4018 | 2.20E-16 | *** |
| TKW | **Environment** | 263 | 1 | 9.894 | 0.00182 | ** |
| TKW | **Block** | 359 | 10 | 1.3506 | 0.20272 | |
| TKW | **Heading date** | 506 | 14 | 1.36 | 0.1715 | |
| TKW | **Genotype x Environment** | 20757 | 377 | 2.0699 | 3.07E-11 | *** |
| TKW | **Residuals** | 8193 | 308 | | | |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Area | **Genotype** | 3943.6 | 723 | 2.9367 | | 2.20E-16 | *** | |
| Area | **Environment** | 73.1 | 1 | 39.3667 | | 1.25E-09 | *** | |
| Area | **Block** | 10.2 | 10 | 0.5467 | | 0.8561 | | |
| Area | **Heading date** | 32.2 | 14 | 1.2367 | | 0.24752 | | |
| Area | **Genotype x Environment** | 891.6 | 374 | 1.2835 | | 0.01233 | * | |
| Area | **Residuals** | 547.9 | 295 | | | | | |
| Perimeter | **Genotype** | 1296.36 | 721 | 4.374 | | 2.20E-16 | *** | |
| Perimeter | **Environment** | 9.14 | 1 | 22.2431 | | 3.71E-06 | *** | |
| Perimeter | **Block** | 0.94 | 10 | 0.2288 | | 0.99337 | | |
| Perimeter | **Heading date** | 10.85 | 14 | 1.8847 | | 0.02761 | | * |
| Perimeter | **Genotype x Environment** | 262.45 | 373 | 1.7117 | 8.26E-07 | | *** | |
| Perimeter | **Residuals** | 121.26 | 295 | | | | | |
| Length | **Genotype** | 242.028 | 722 | 6.3379 | | 2.20E-16 | *** | |
| Length | **Environment** | 0.754 | 1 | 14.2647 | | 0.000192 | *** | |
| Length | **Block** | 0.298 | 10 | 0.5636 | | 0.843104 | | |
| Length | **Heading date** | 1.219 | 14 | 1.6461 | | 0.066459 | . | |
| Length | **Genotype x Environment** | 33.095 | 377 | 1.6597 | | 2.94E-06 | *** | |
| Length | **Residuals** | 15.603 | 295 | | | | | |
| Width | **Genotype** | 44.899 | 723 | 2.9467 | | 2.20E-16 | *** | |
| Width | **Environment** | 1.286 | 1 | 61.0345 | | 9.92E-14 | *** | |
| Width | **Block** | 0.175 | 10 | 0.831 | | 0.599056 | | |
| Width | **Heading date** | 0.254 | 14 | 0.8606 | | 0.60251 | | |
| Width | **Genotype x Environment** | 10.287 | 377 | 1.2947 | 0.009958 | | ** | |
| Width | **Residuals** | 6.217 | 295 | | | | | |
| L* | **Genotype** | 4081.7 | 722 | 3.7025 | 2.20E-16 | | *** | |
| L* | **Environment** | 396.7 | 1 | 259.8398 | 2.20E-16 | | *** | |
| L* | **Block** | 20.6 | 10 | 1.3508 | 0.202896 | | | |
| L* | **Heading date** | 19.6 | 14 | 0.9182 | 0.539425 | | | |
| L* | **Genotype x Environment** | 835.8 | 373 | 1.4675 | 0.000293 | | *** | |
| L* | **Residuals** | 450.4 | 295 | | | | | |
| A* | **Genotype** | 1301.12 | 723 | 1.6238 | 9.44E-07 | | *** | |
| A* | **Environment** | 295.36 | 1 | 266.5067 | 2.20E-16 | | *** | |
| A* | **Block** | 12.7 | 10 | 1.1463 | 0.3274 | | | |
| A* | **Heading date** | 16.19 | 14 | 1.0431 | 0.4102 | | | |
| A* | **Genotype x Environment** | 438.45 | 377 | 1.0494 | 0.3325 | | | |
| A* | **Residuals** | 326.94 | 295 | | | | | |
| B* | **Genotype** | 3588.7 | 719 | 6.8936 | <2e-16 | | *** | |
| B* | **Environment** | 81.7 | 1 | 112.8529 | <2e-16 | | *** | |
| B* | **Block** | 8.9 | 10 | 1.229 | 0.2718 | | | |
| B* | **Heading date** | 4.8 | 14 | 0.4772 | 0.9443 | | | |
| B* | **Genotype x Environment** | 302.4 | 372 | 1.1226 | 0.1494 | | | |
| B* | **Residuals** | 212.1 | 293 | | | | | |

It can be noticed that spike related traits show a more significant GxE interaction rather than grain-related traits, which are however significantly affected by environment conditions.

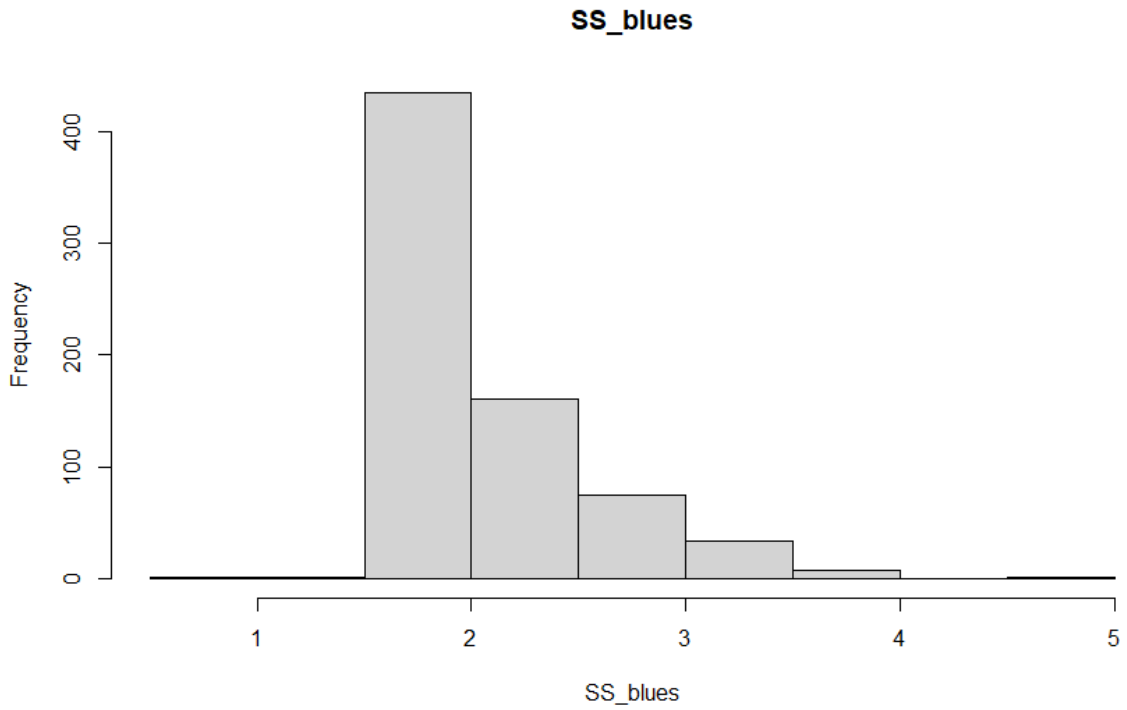


Figure 46 Sterile spikelets distribution frequency for GDP cluster

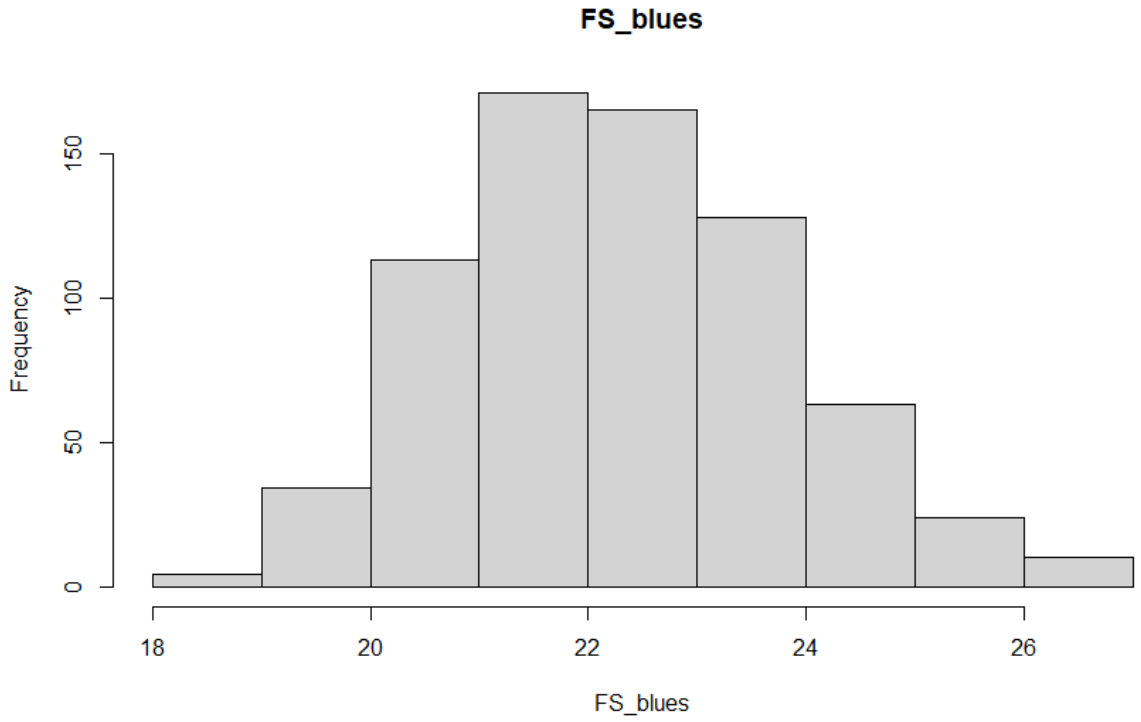


Figure 47 Fertile spikelets distribution frequency for GDP cluster

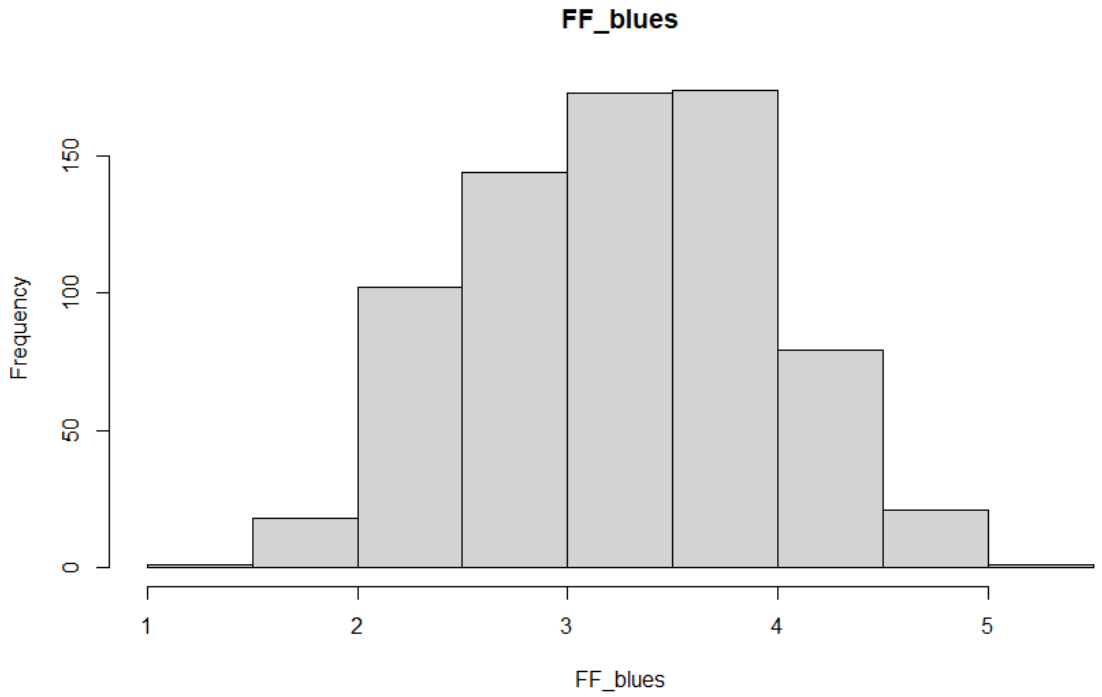


Figure 48 Fertile florets per central spikelet distribution frequency in GDP custer

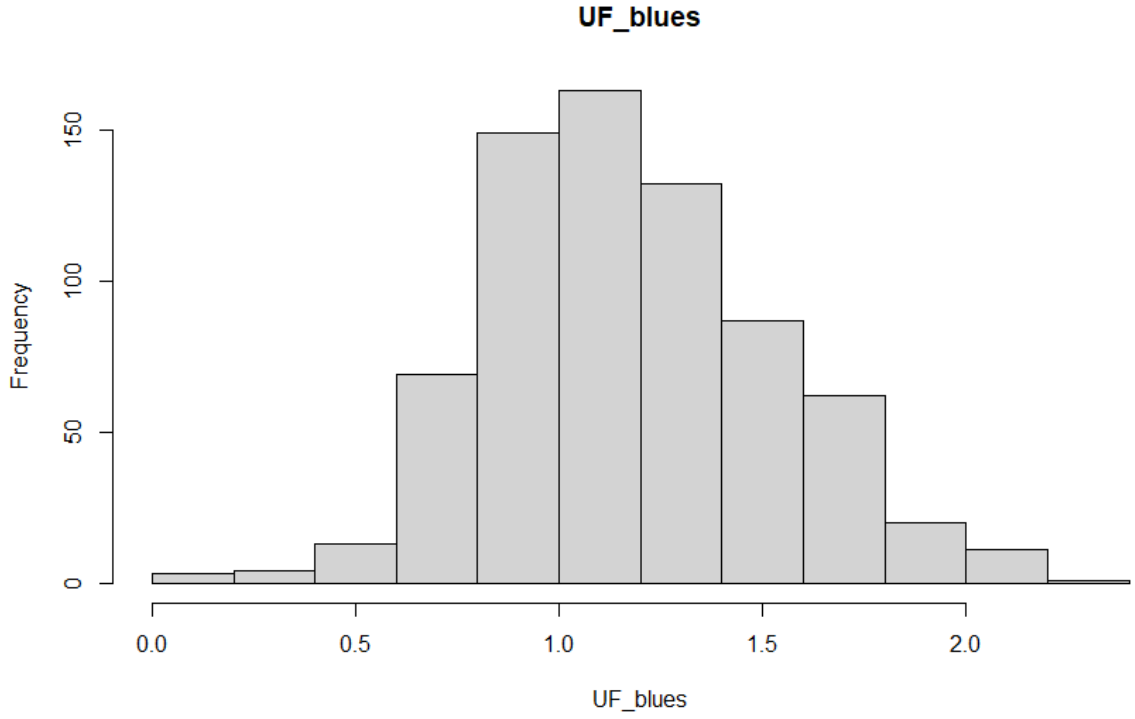


Figure 49 Unfertile florets distribution frequency in GDP cluster

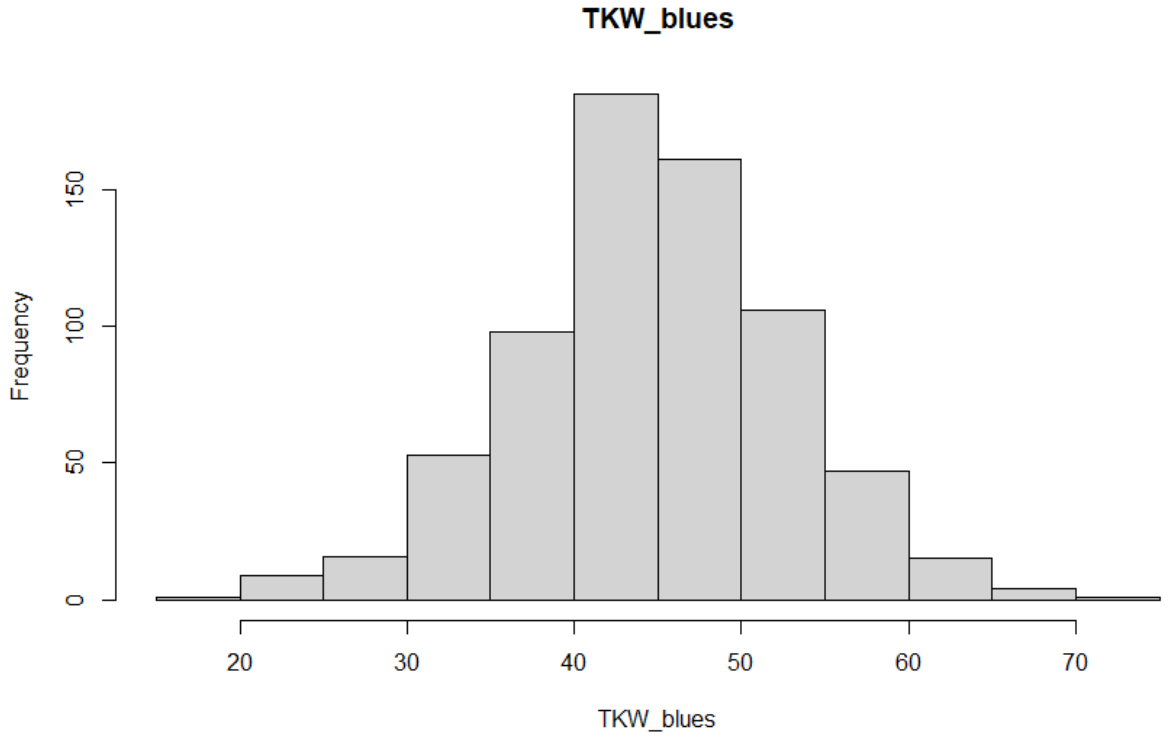


Figure 50 Thousand kernel weight distribution frequency in GDP cluster

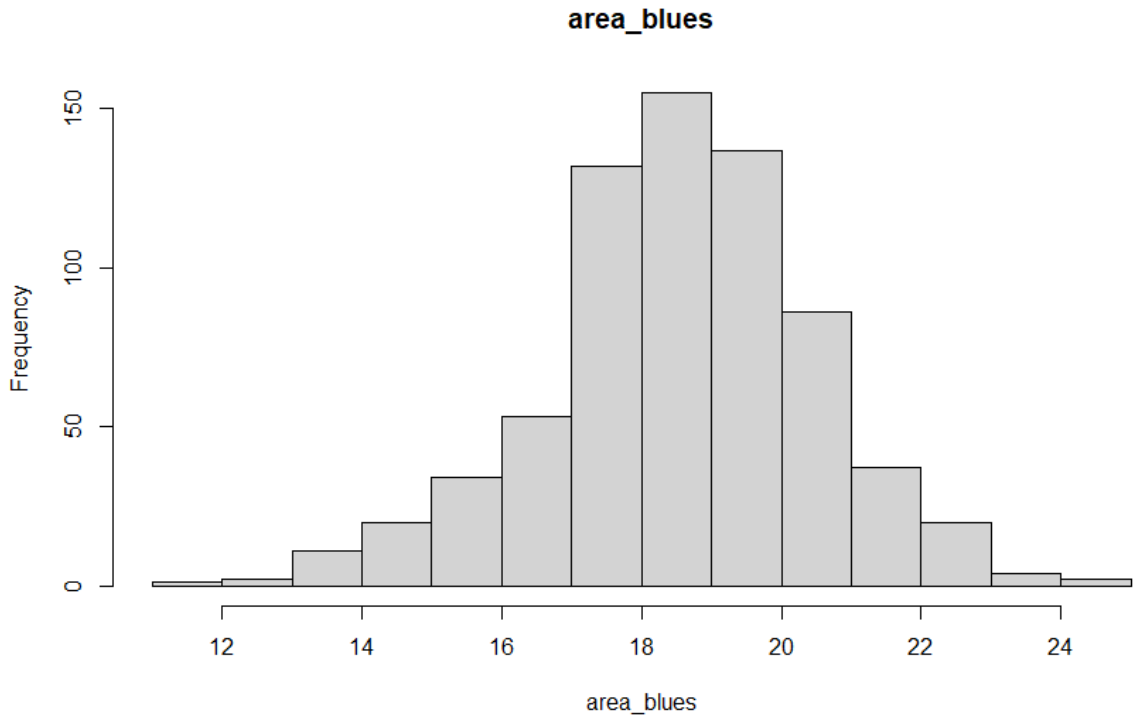


Figure 51 Grain area distribution frequency for GDP cluster

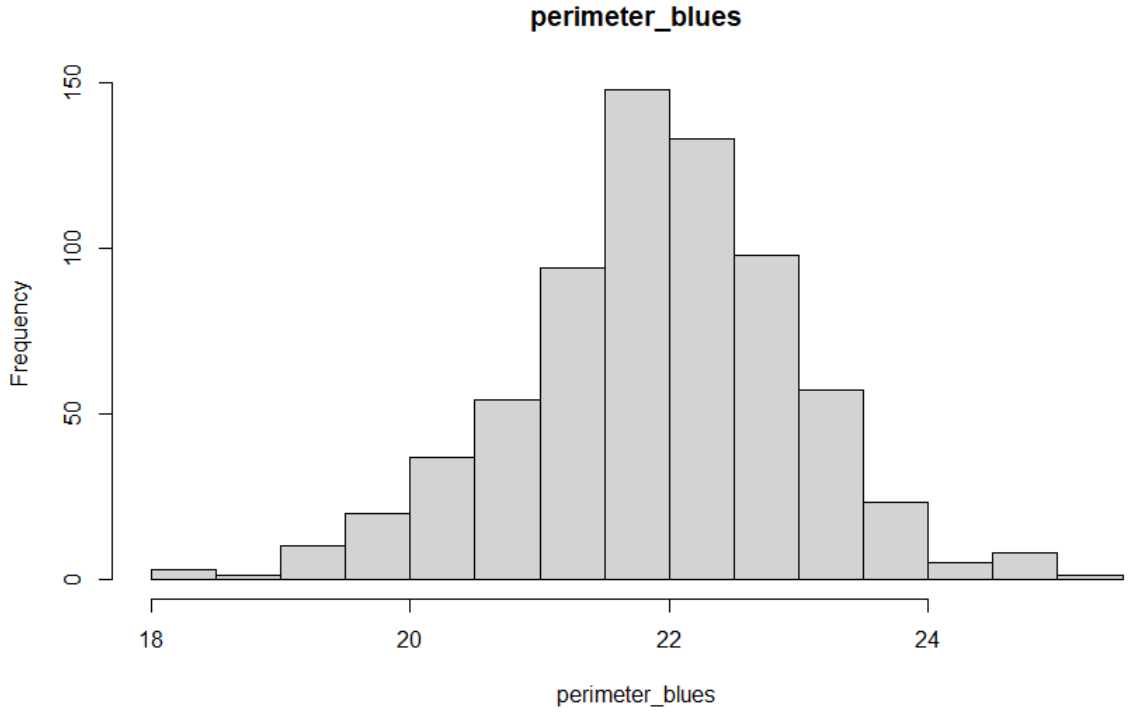


Figure 52 Grain perimeter distribution frequencies for GDP cluster

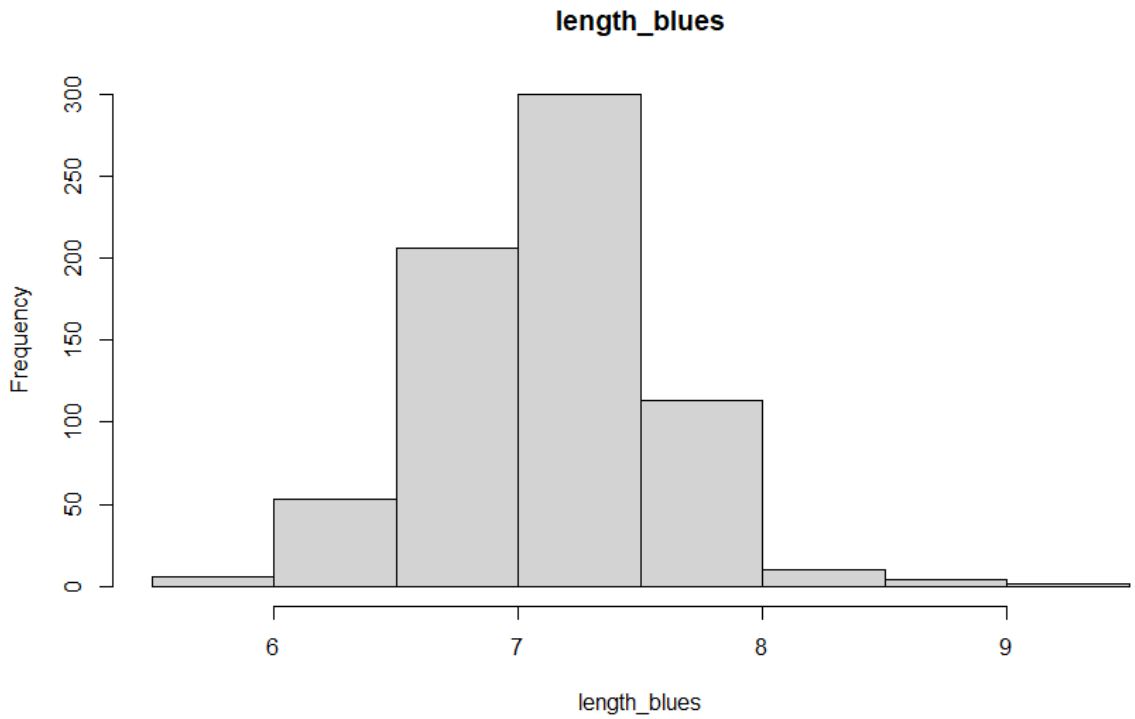


Figure 53 Grain length distribution frequency for GDP cluster

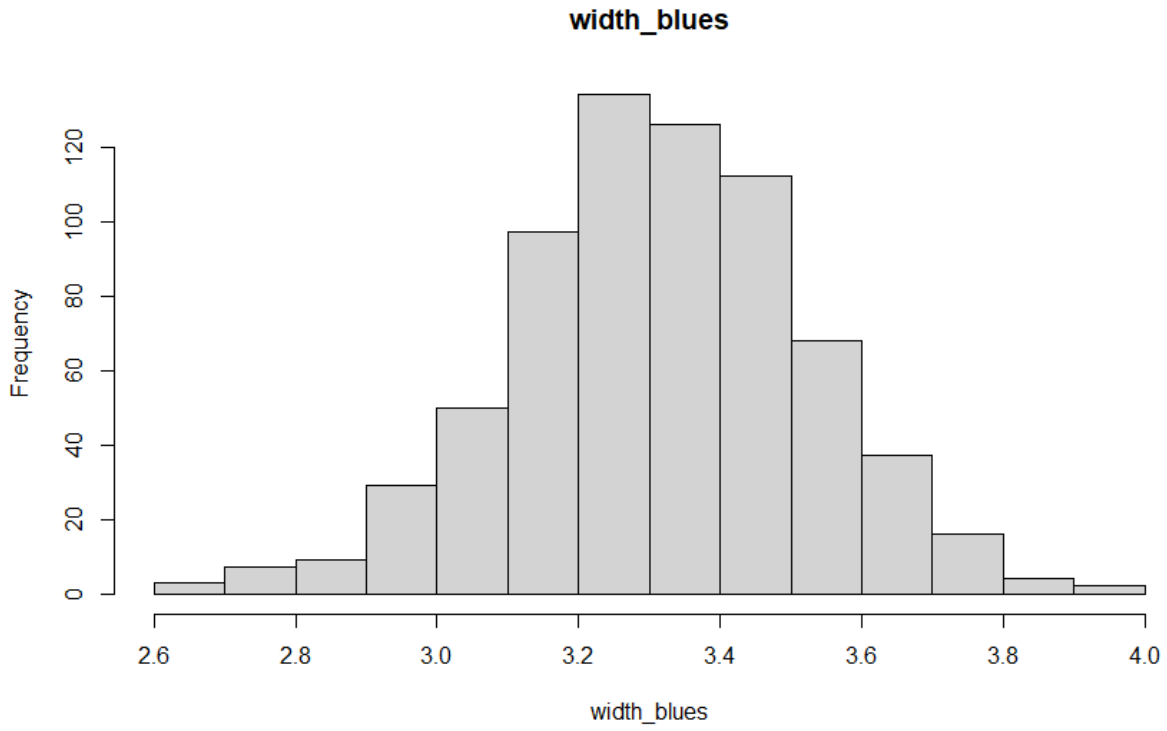


Figure 54 Grin width distribution frequency for GDP cluster

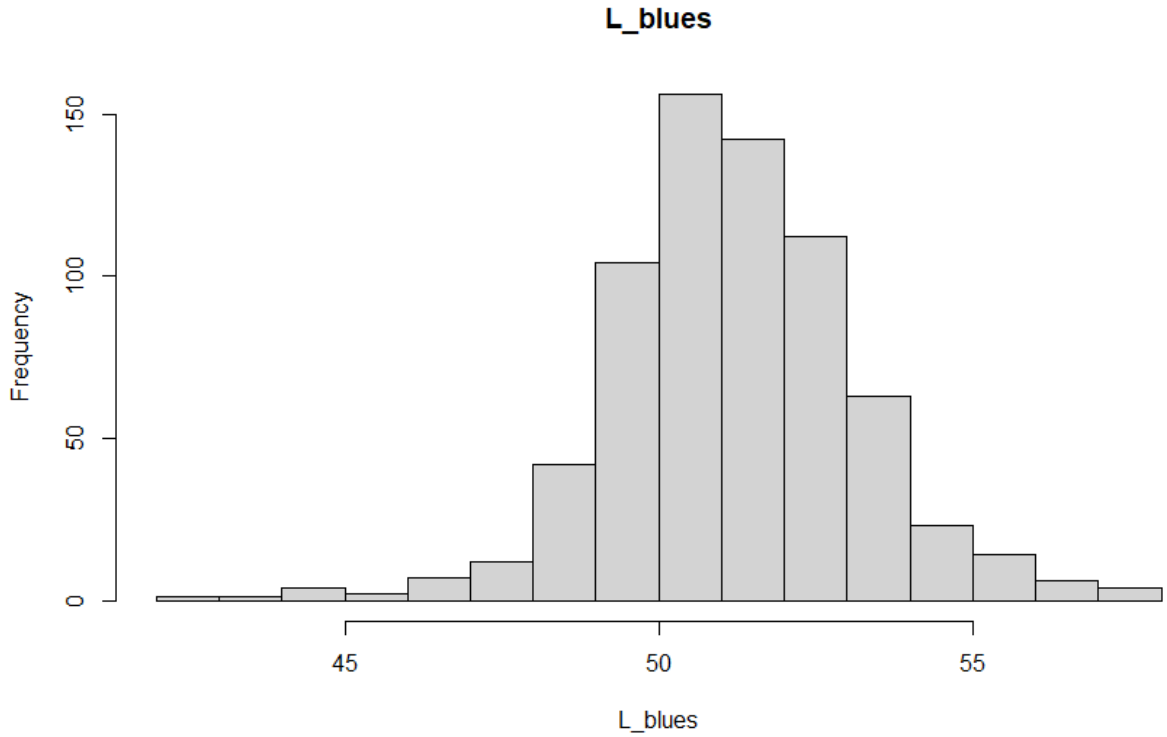


Figure 55 Grain brightness distribution frequency for GDP cluster

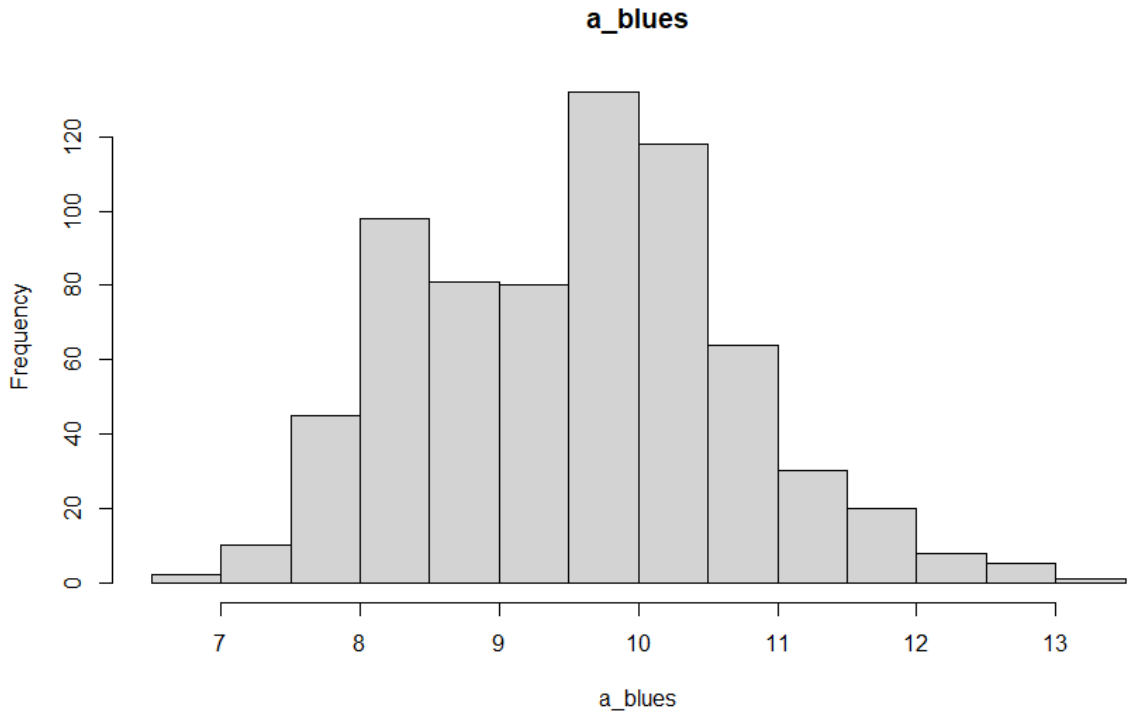


Figure 56 Grain redness distribution frequency for GDP cluster

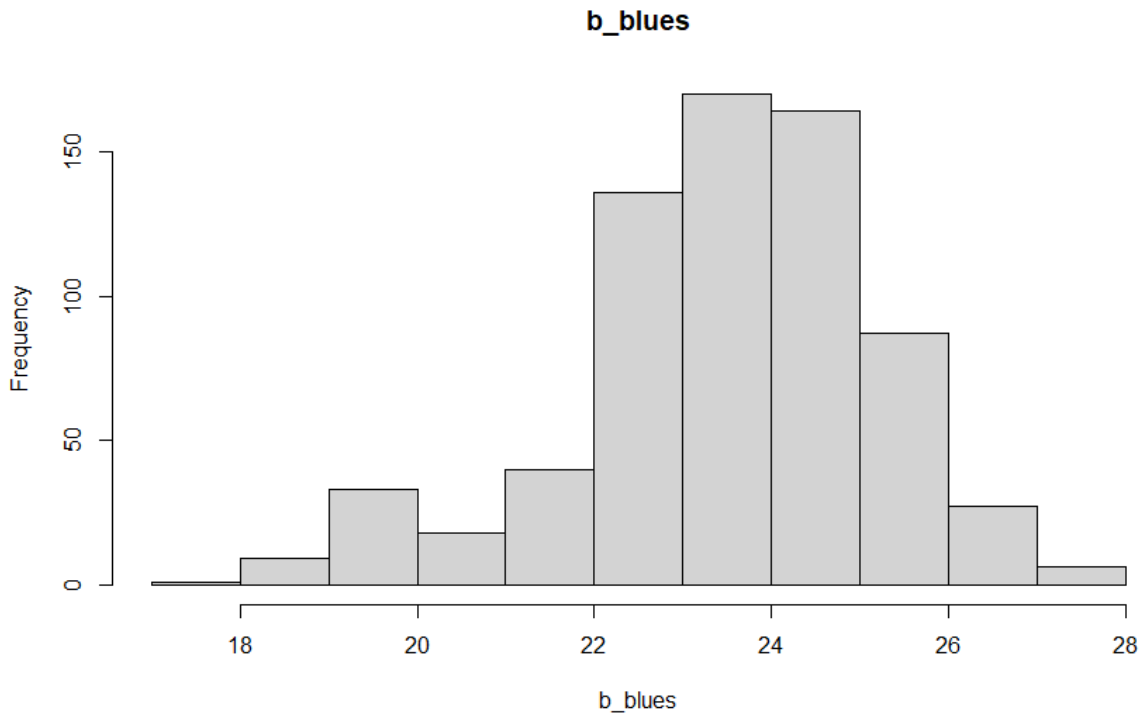


Figure 57 Grain yellowness distribution frequency for GDP cluster

### 2.3.7 TGC 2019 and TGC 2020, combined analysis

Table 15 Descriptive statistics for spike related traits in TGC cluster

| | SS | FS | FF | UF |
|---|---|---|---|---|
| **min** | -0.18 | 13.1 | -0.27 | 0.01 |
| **max** | 4.2 | 39.41 | 4.75 | 5.67 |
| **range** | 4.38 | 26.3 | 5.02 | 5.66 |
| **median** | 0.88 | 26.85 | 1.91 | 3.95 |
| **mean** | 1 | 26.92 | 1.82 | 3.9 |
| **SE.mean** | 0.02 | 0.07 | 0.02 | 0.02 |
| **var** | 0.41 | 7.21 | 0.6 | 0.41 |
| **std.dev** | 0.64 | 2.69 | 0.77 | 0.64 |

| | | | | |
|---|---|---|---|---|
| **coef.var** | 0.65 | 0.1 | 0.42 | 0.16 |
| $h^2$ | 0.629677 | 0.720234 | 0.696186 | 0.372234 |

TGC multienvironmental analysis highlighted an enhanced heritability score for all traits compared to each single environment.

Table 16 ANOVA results for TGC cluster

| Trait | Variables | Sum Sq | Df | F value | Pr(>F) | |
|---|---|---|---|---|---|---|
| SS | **Genotype** | 694.43 | 1458 | 5.3398 | 2.20E-16 | *** |
| SS | **Environment** | 1.16 | 1 | 13.0404 | 0.000358 | *** |
| SS | **Block** | 2.42 | 17 | 1.5961 | 0.064063 | . |
| SS | **Heading date** | 3.78 | 22 | 1.9274 | 0.008384 | ** |
| SS | **Genotype x Environment** | 166.86 | 848 | 2.206 | 6.93E-15 | *** |
| SS | **Residuals** | 26.31 | 295 | | | |
| FS | **Genotype** | 10722.7 | 1459 | 5.2251 | 2.20E-16 | *** |
| FS | **Environment** | 1.4 | 1 | 0.9659 | 0.326506 | |
| FS | **Block** | 21.6 | 17 | 0.9054 | 0.56806 | |
| FS | **Heading date** | 67.1 | 22 | 2.1676 | 0.002177 | ** |
| FS | **Genotype x Environment** | 2156.1 | 852 | 1.7992 | 2.90E-09 | *** |
| FS | **Residuals** | 414.9 | 295 | | | |
| FF | **Genotype** | 925.7 | 1460 | 5.1617 | 2.20E-16 | *** |
| FF | **Environment** | 5.9 | 1 | 48.0077 | 2.67E-11 | *** |
| FF | **Block** | 2.42 | 17 | 1.1588 | 0.29796 | |
| FF | **Heading date** | 4.75 | 22 | 1.7571 | 0.02065 | * |
| FF | **Genotype x Environment** | 235.85 | 856 | 2.2431 | 2.02E-15 | *** |

| | | | | | | |
|---|---|---|---|---|---|---|
| FF | **Residuals** | 36.24 | 295 | | | |
| UF | **Genotype** | 413.14 | 1456 | 2.2574 | 2.20E-16 | *** |
| UF | **Environment** | 0.68 | 1 | 5.3868 | 0.020972 | * |
| UF | **Block** | 3.37 | 17 | 1.5749 | 0.069673 | . |
| UF | **Heading date** | 7.28 | 22 | 2.6311 | 0.000135 | *** |
| UF | **Genotype x Environment** | 221.37 | 850 | 2.0719 | 5.03E-13 | *** |
| UF | **Residuals** | 37.08 | 295 | | | |

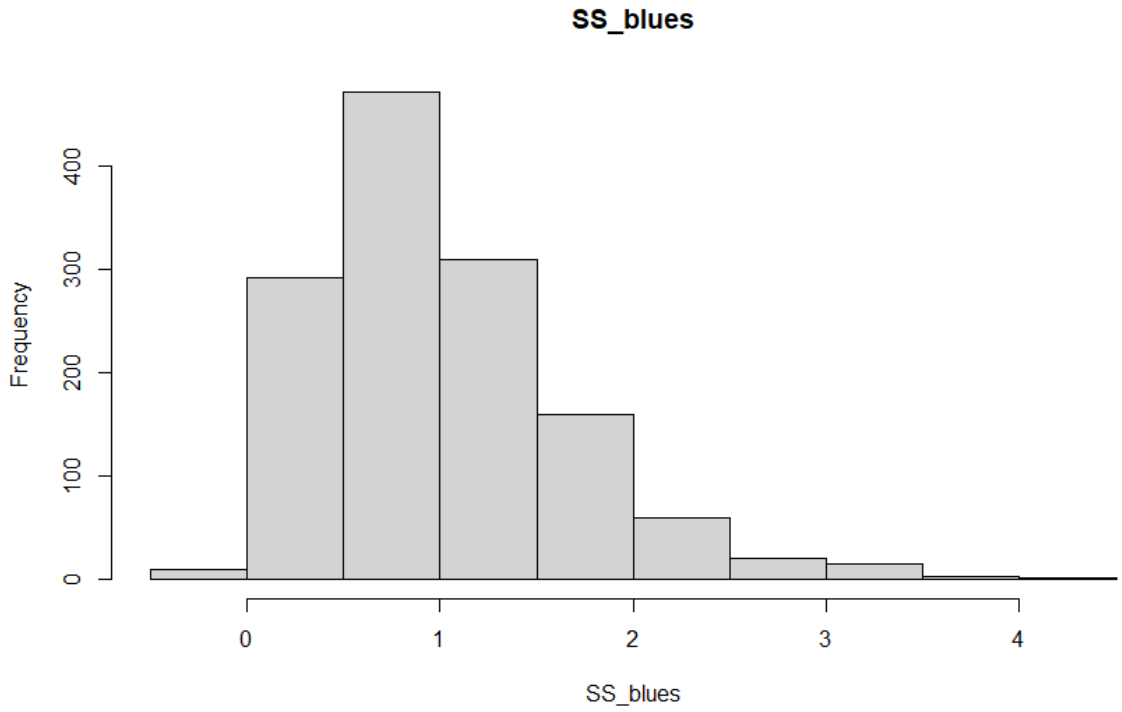


Figure 58 Sterle spikelets distribution frequency for TGC lcuster

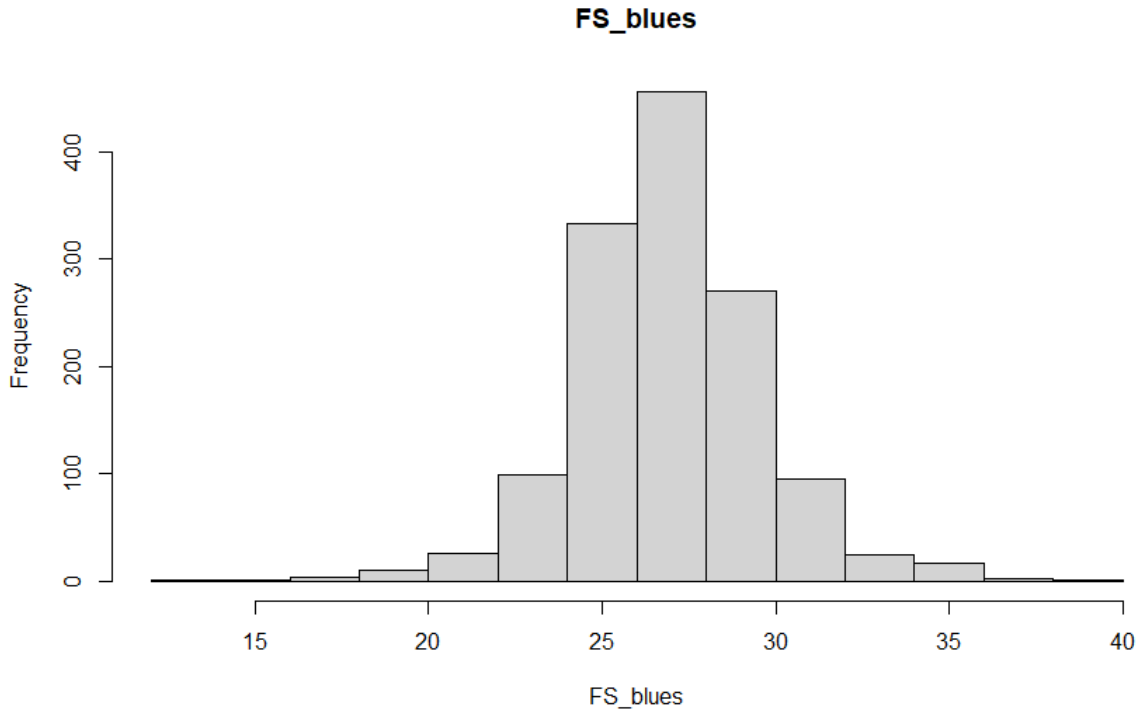


Figure 59 Fertile spikelet distribution frequency for TGC cluster

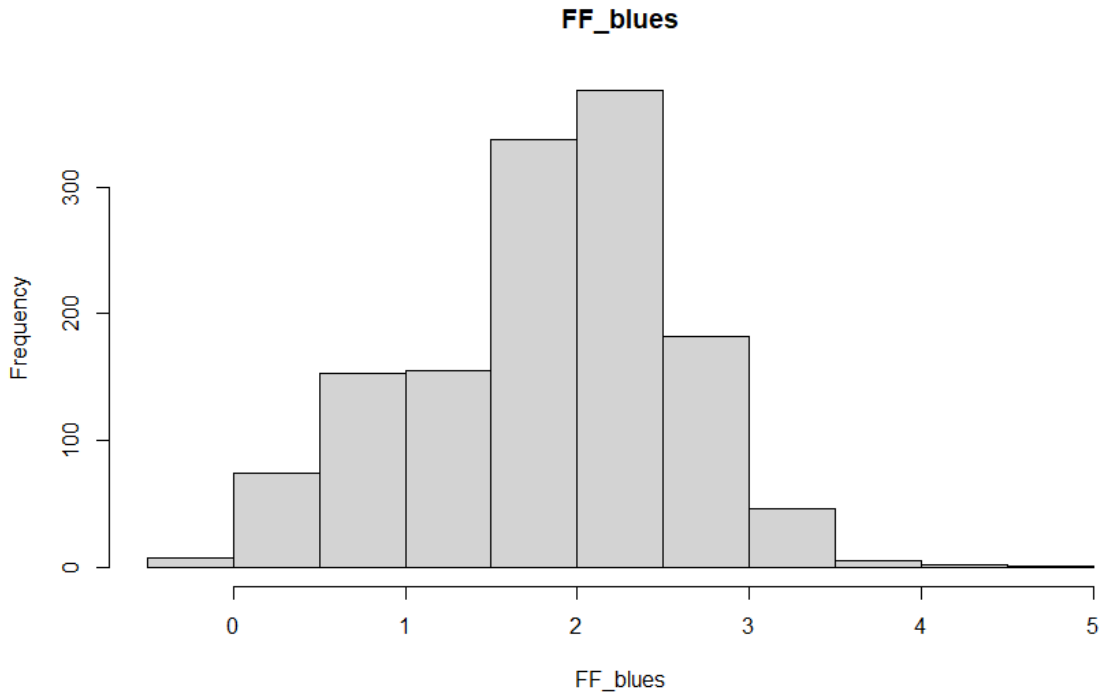


Figure 60 Fertile florets per central spikelet distribution frequency for TGC cluster

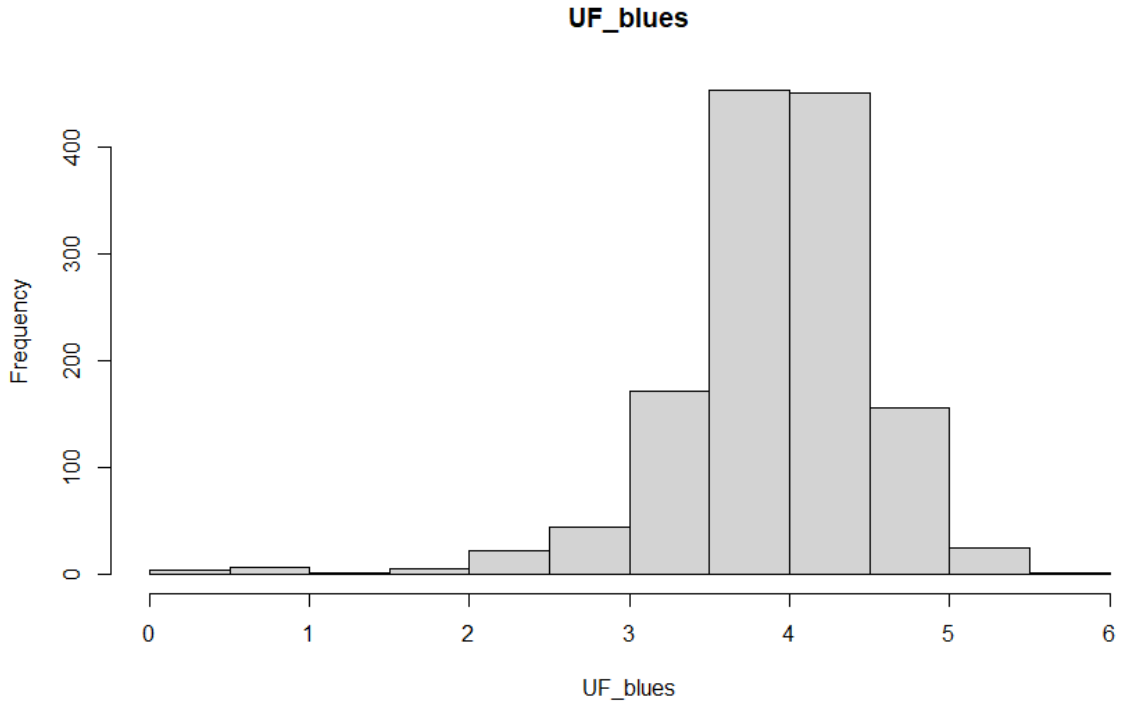


Figure 61 Unfertile florets distribution frequency for TGC cluster

### 2.3.8 Fertility GWAS results

GWAS analysis was performed using BLUEs data from the GDP and TGC with GAPIT3 software. The filtered hapmap was used for LD decay and kinship matrix computing with TASSEL 5.
At 0.3 $r^2$ threshold the LD decay was nearly 1 Mbp.

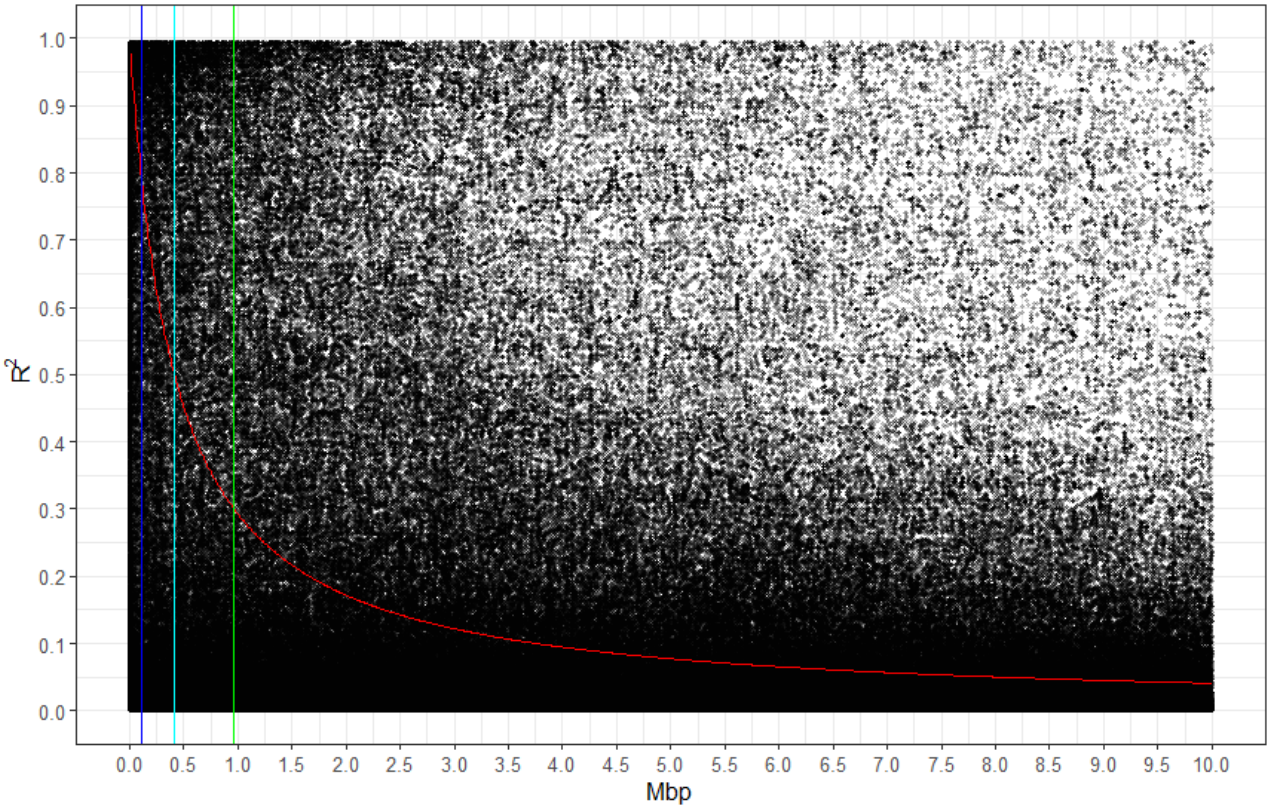


Figure 62 LD decay plot for GDP and TGC

For the different GWAS models used in this work, Bonferroni threshold was computed by performing 1000 permutations with FarmCPU model. The significant p-value threshold (0.05) was divided for the number of SNP markers after the pruning process. The Bonferroni threshold ranged on average from 4 to 5 and in order to detect significant trait-marker associations, the cutoff value of 4 was adopted.

Manhattan plots herein presented are referred to BLINK model only even though the GWAS analysis was carried out for all the aforementioned models.

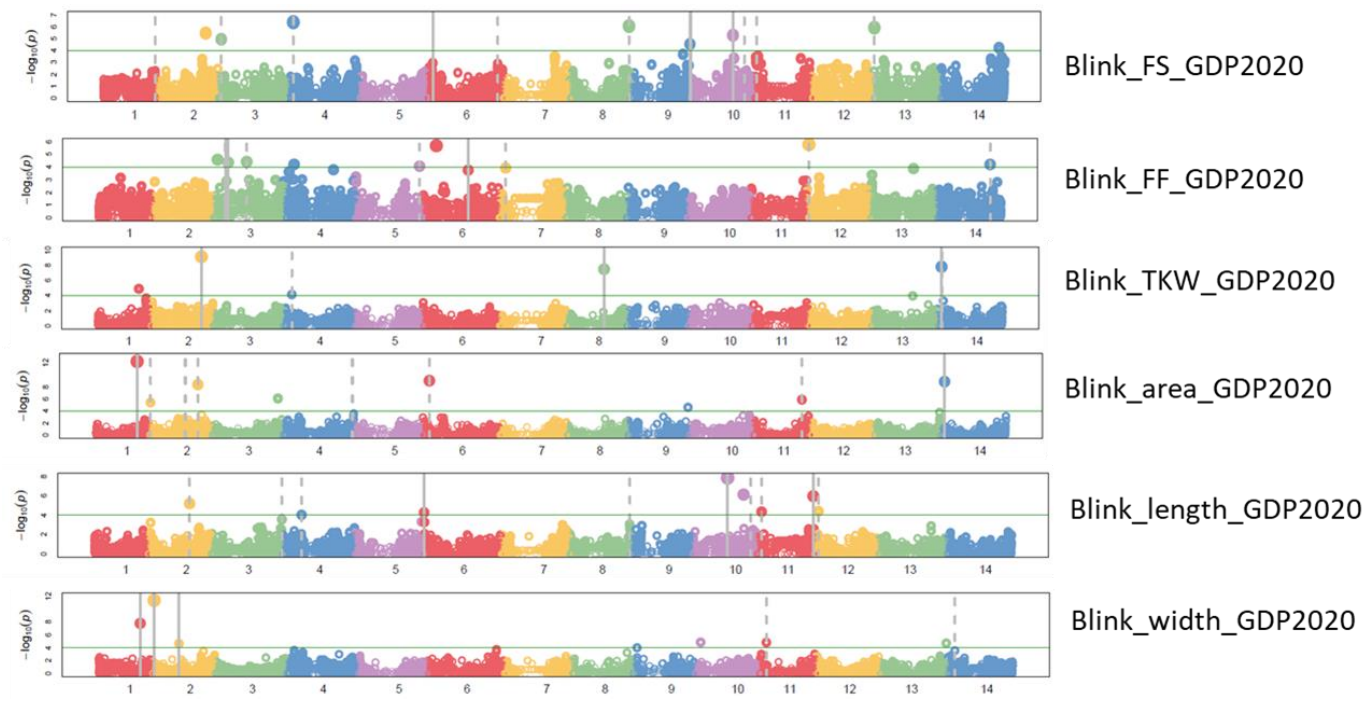


Figure 63 Manhattan plot for the investigated traits in GDP 2020.

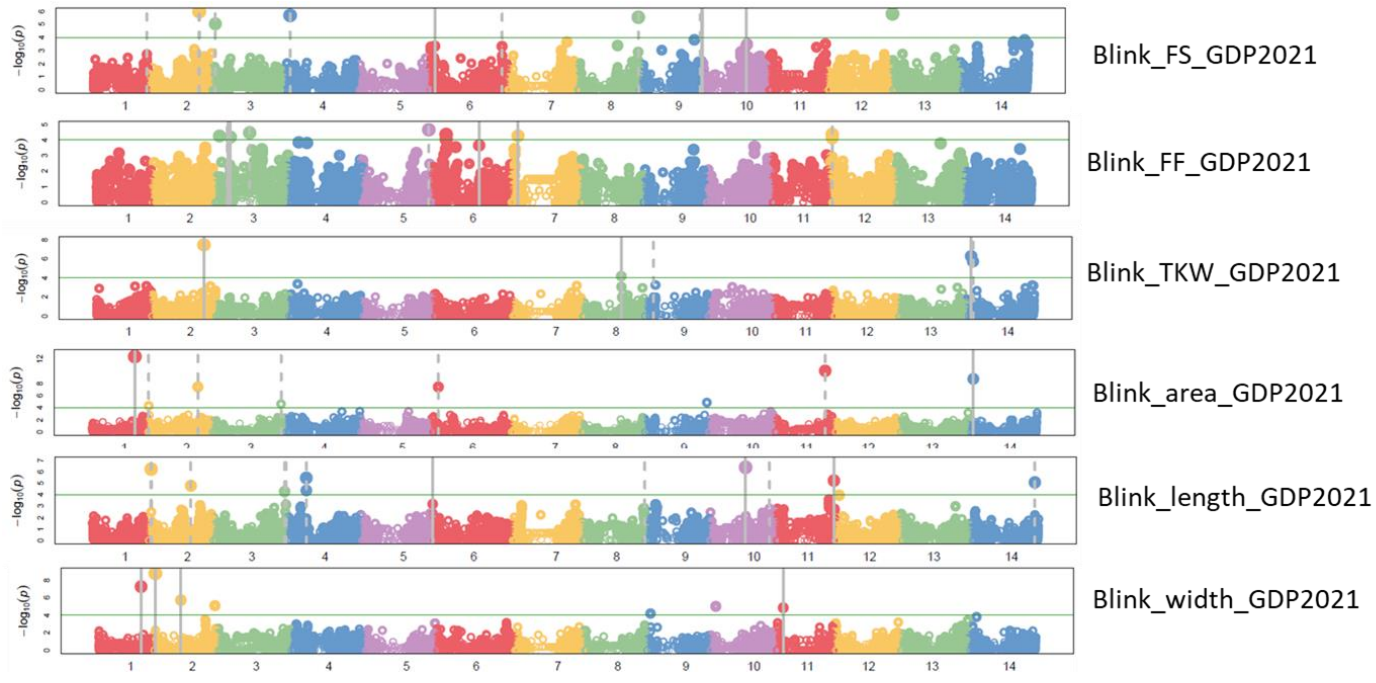


Figure 64 Manhattan plot GWAS output for GDP 2021

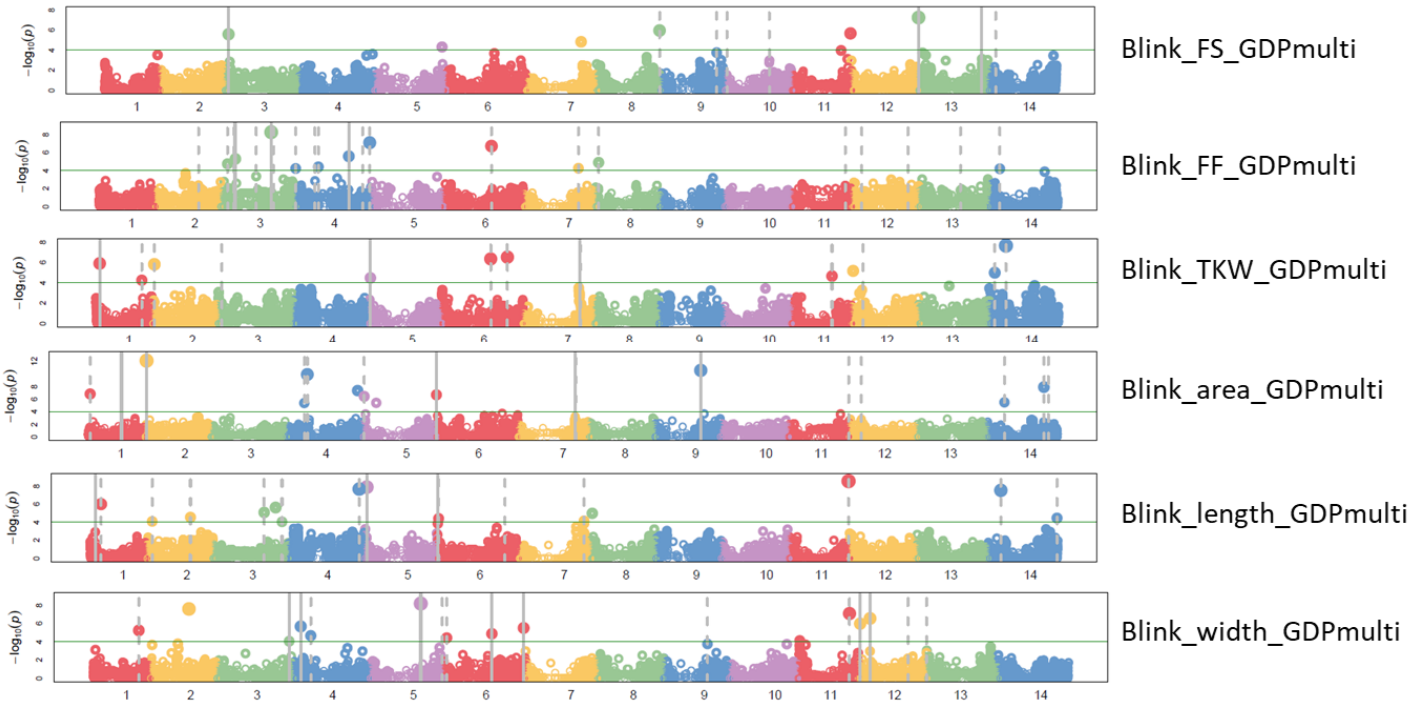


Figure 65 Manhattan plot GWAS output for GDP cluster

Manhattan plots, here depicted for the two environments of the GDP and the combined analysis, show the peaks of significance on the whole tetraploid genome. Every coloured section represents a different chromosome. Bonferroni threshold was used to set the threshold of significance to 4.

Single environment analysis show very similar peaks for nearly every trait, while the combined GWAS highlights different peaks. A strong signal on chromosome 2A is very clear across the environments and also in the combined analysis. This could be due to the presence of the *GNI-2* locus, already studied. There is a stable signal on chromosome 4B for the trait FS which also correspond to a similar peak in the combied analysis. Chromosome 7B has a coincidence peak among all three analysis for the TKW trait.

There are also peak coincidence on chromosomes 1A, 1B and 6A for the grain width trait, which is clearly detectable through all three analysis. Grain length is characterised by a peak on chromosome 6A.

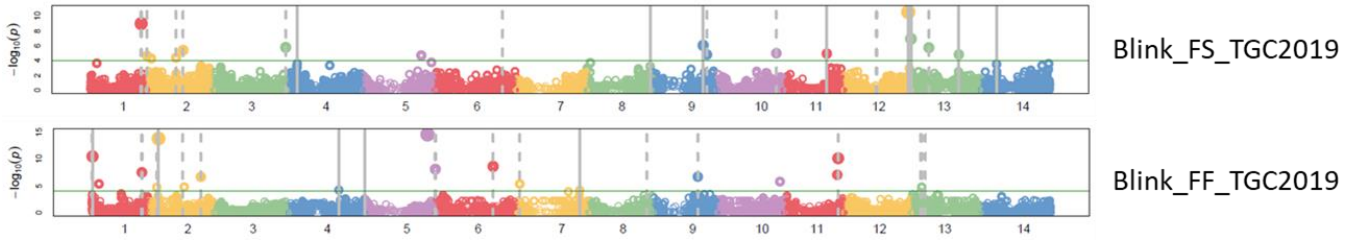


Figure 66 Manhattan plot for the spike fertility traits in TGC 2019

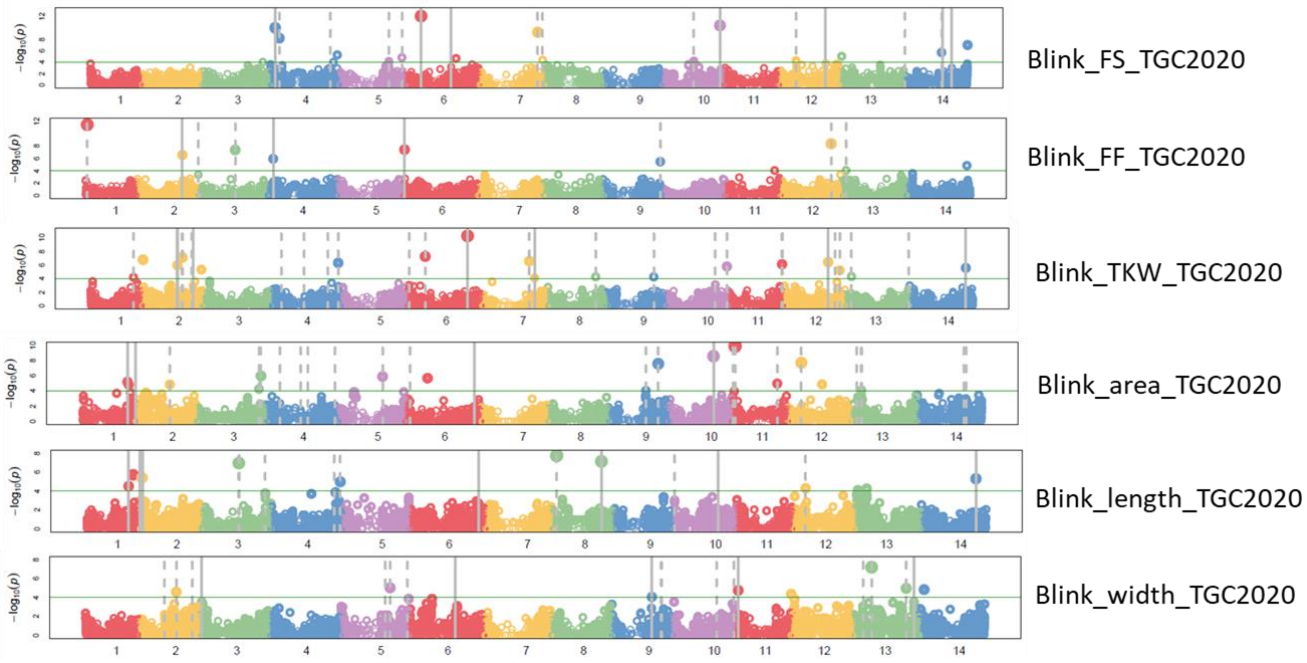


Figure 67 Manhattan plot output for the TGC 2020

Since TGC 2019 has been phenotyped only for spike-fertility related traits, here are reported only two manhattan plots.

Following the GWAS analysis, peaks resulting from the Manhattan plots were summarized in the Table 17 for the GDP and Table 18 for TGC. Data are referred to cluster environment analysis. Peaks are indicated by the single most signicatevely associated SNP alongside with the confidence interval which was determined using the LD decay (1.0 Mb) at both sides of the tag SNP. Significant markers here reported , are mapped on the Svevo RefSeq v1.0 reference genome. The -log(p) value is referred to the BLINK model.

Table 17 Most associated SNP from GWAS results in GDP

| SNP | Chromosome | Position | -log(P) | C.I (+/-LD decay) | Trait |
|---|---|---|---|---|---|
| IWB67308 | 2A | 35700735 | 5.58 | IWB67308 - IWB51686 | FS |
| IWA6465 | 4B | 656489913 | 5.95 | IWB72184 - IWB74054 | FS |
| IWB21158 | 7A | 5590055 | 7.24 | IWB71146 - IWB34436 | FS |
| IWB67301 | 2B | 2559456 | 4.22 | IWB66351 - IWB7677 | FF |
| IWA6850 | 4B | 36326487 | 4.88 | IWB70449 - IWB61488 | FF |
| IWB25495 | 3B | 675968918 | 6.51 | IWA3046 - IWB65507 | TKW |
| IWB73924 | 7B | 171541547 | 7.65 | IWB73924 - IWB71851 | TKW |
| IWB26242 | 1B | 19537743 | 12.07 | IWB8104 - IWB44700 | Grain area |
| IWB73924 | 7B | 171541547 | 5.47 | IWB73924 - IWB71851 | Grain area |
| IWB37079 | 3A | 17214656 | 7.86 | IWB35874 - IWB72257 | Grain length |
| IWB67460 | 6A | 606825167 | 8.52 | IWB65928 - IWB16508 | Grain length |
| IWB73249 | 2B | 79053945 | 5.66 | IWB45339 - IWB67029 | Grain width |
| IWB5996 | 6A | 524333330 | 7.07 | IWB9600 - IWB33872 | Grain width |
| IWB69456 | 7A | 61263506 | 8.43 | IWB40391 - IWB22591 | Grain brightness |
| IWB1030 | 7B | 171576452 | 15.30 | IWB73924 - IWB71851 | Grain brightness |
| IWA7148 | 2A | 704236759 | 7.40 | IWA5216 - IWB7166 | Grain redness |
| IWA2644 | 5A | 667286036 | 6.17 | IWB71094 - IWA2646 | Grain redness |
| IWB23681 | 3B | 768866167 | 7.42 | IWB60646 - IWB23680 | Grain yellowness |
| IWB72251 | 7A | 3034881 | 9.04 | IWB66267 - IWB21994 | Grain yellowness |

Table 18 Most associated SNP from GWAS results in TGC

| SNP | Chromosome | Position | -log(P) | C.I (+/-LD decay) | Trait |
|---|---|---|---|---|---|
| IWB11011 | 1B | 431037305 | 3.50 | IWB71872 - IWB7028 | FS |
| IWB60297 | 2A | 562062720 | 7.69 | IWB44801 -IWB13477 | FS |
| IWB8932 | 6A | 444333914 | 9.16 | IWA2416 - IWA428 | FS |
| IWB24065 | 1B | 14429522 | 3.94 | IWB2188 - IWB73279 | FF |
| IWB32738 | 3A | 735562087 | 4.81 | IWB65706 - IWB50704 | FF |
| IWB7107 | 5B | 661651875 | 5.47 | IWB10034 - IWB50537 | FF |
| IWB13742 | 1B | 592592522 | 7.07 | IWB8867 - IWB7410 | TKW |
| IWB21895 | 2B | 555610066 | 5.11 | IWB46098 - IWA5141 | TKW |
| IWA429 | 2B | 145635634 | 4.77 | IWB32296 - IWB27957 | Grain area |
| IWB23124 | 6B | 146354976 | 6.27 | IWA3424 - IWB10696 | Grain area |
| IWB10610 | 2B | 765476835 | 4.97 | IWB70506 - IWA3474 | Grain length |
| IWB49696 | 4B | 546670085 | 7.13 | IWA5955 - IWB6922 | Grain length |
| IWB14408 | 7B | 606338863 | 5.30 | IWB61109 - IWB14408 | Grain length |
| IWB53342 | 5B | 668490739 | 6.64 | IWB29437 - IWB25892 | Grain width |
| IWB11722 | 6A | 30332275 | 5.91 | IWB69175 - IWB26178 | Grain width |
| IWA3037 | 2B | 596004366 | 4.59 | IWB50067 - IWA2189 | Grain brightness |
| IWB73963 | 5A | 451516943 | 9.08 | IWB14493 - IWB71451 | Grain brightness |
| IWB23681 | 3B | 768866167 | 8.20 | IWB60646 - IWB23680 | Grain redness |
| IWB42829 | 6A | 26766423 | 7.83 | IWB22480 - IWB70424 | Grain redness |
| IWA628 | 3B | 260580846 | 4.91 | IWB1111 - IWB42046 | Grain yellowness |
| IWB23612 | 6A | 29437066 | 8.92 | IWB43285 - IWB72838 | Grain yellowness |

For each considered trait, the main QTLs were studied for candidate genes analysis within the confidence interval, which was determined starting from the LD decay at both sides (1.0 Mb). The confidence interval was investigated based on the Triticum turgidum cv Svevo *RefSeq* v1.0.

For the candidate genes, both position and gene description were obtained with Ensembl Plants Biomart

Table 19 Fertile spikelet per spike candidate genes. In orage the peaks identified in the GDP, in blue in the TGC

| Gene stable ID | Gene start (bp) | Gene end (bp) | Gene description |
|---|---|---|---|
| **TRITD2Av1G019050** | 36289642 | 36293816 | Kinase interacting (KIP1-like) family protein |
| **TRITD4Bv1G199180** | 655246374 | 655253180 | F-box and Leucine Rich Repeat domains containing protein, putative isoform 1 TE? |
| **TRITD2Av1G202930** | 562242917 | 562246563 | WRKY transcription factor |
| **TRITD6Av1G153220** | 445179612 | 445185200 | Tetratricopeptide repeat protein 1-like |

For the fertile spikelet (FS) trait, the candidate genes were comprised in the kinase family (GDP-2A), F-box and leucine rich (GDP- 4B), WRKY transcription factor (TGC- 2A) and tetratricopeptide repeat protein1-like (TGC-6A).

From the Svevo genes, the Chinese Spring orthologues were obtained in order to elucidate the protein function and their respective phenotype using Knetminer

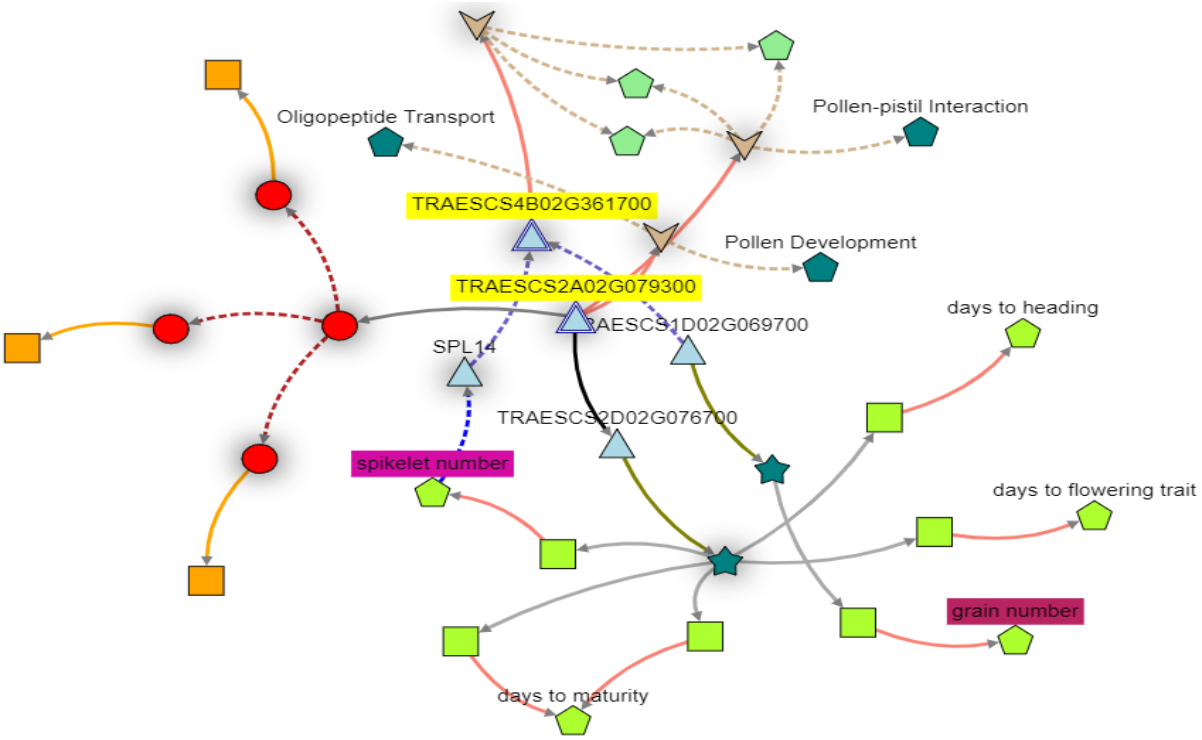


Figure 68 Knetminer network of the genes function for the FS trait in the GDP

The protein function found here are mostly related to the spikelet number itself, but also to the grain number and heading and flowering traits.

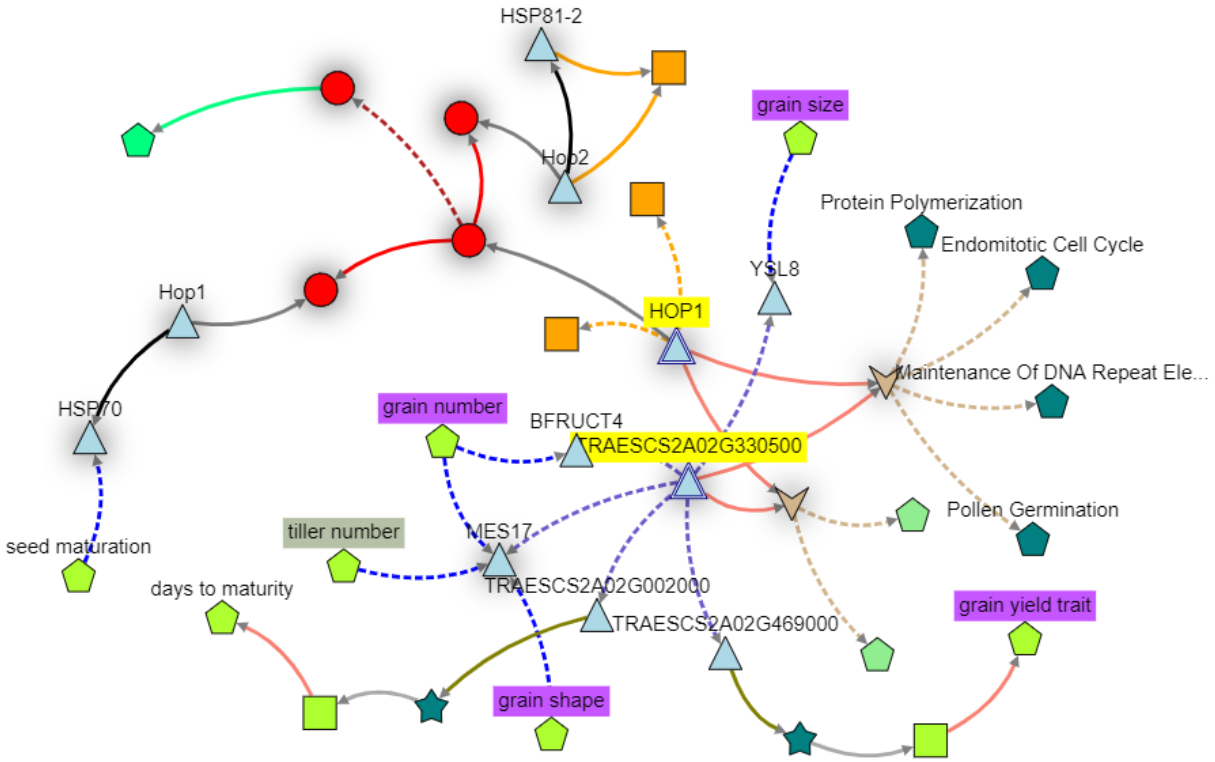


Figure 69Knet miner network of the genes function for the FS trait in the TGC

Table 20 Fertile florets per central spikelet candidate genes

| Gene stable ID | Gene start (bp) | Gene end (bp) | Gene description |
|---|---|---|---|
| **TRITD2Av1G134810** | 370285942 | 370287380 | Receptor-like kinase |
| **TRITD2Bv1G001000** | 1341694 | 1344094 | Ethylene receptor |
| **TRITD4Bv1G014770** | 36326599 | 36329970 | 26S proteasome non-atpase regulatory subunit, putative |

Table 21 TKW candidate genes. In orage the peaks identified in the GDP, in blue in the TGC

| Gene stable ID | Gene start (bp) | Gene end (bp) | Gene description |
|---|---|---|---|
| **TRITD3Bv1G222100** | 676283650 | 676284462 | Leucine-rich repeat receptor-like protein kinase |
| **TRITD7Bv1G060720** | 171577502 | 171588038 | Alpha-glucan water dikinase |
| **TRITD1Bv1G193520** | 592543714 | 592560282 | NADH-quinone oxidoreductase subunit C/D |

For the TKW trait the discovered genes functions belong to leucine rich repeat receptor-like protein kinase (GDP-3B) and a NADH-quinone oxidoreductase (TGC-1B). The alpha-glucan water dikinase is an enzyme with an important role in stanch degradation in source tissues as pointed out in previous study (Ral et al., 2021).

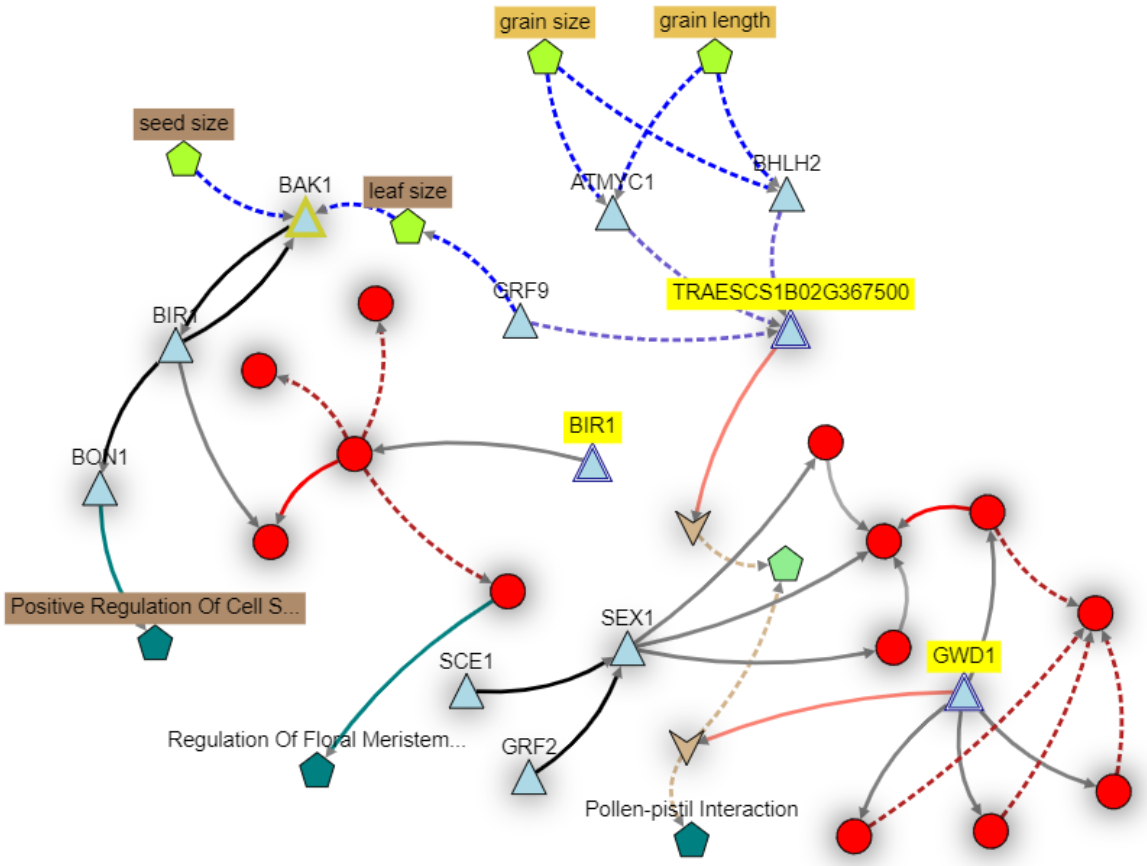


Figure 70 Knetminer network of gene functions for TKW trai

Table 22 Grain area candidate genes

| Gene stable ID | Gene start (bp) | Gene end (bp) | Gene description |
|---|---|---|---|
| **TRITD7Bv1G060720** | 171577502 | 171588038 | Alpha-glucan water dikinase |
| **TRITD2Bv1G056620** | 145746061 | 145749219 | Small nuclear RNA activating complex (SNAPc), subunit SNAP43 protein |
| **TRITD6Bv1G052260** | 146510219 | 146516463 | Protein kinase, putative |

The grain area candidate gene detected in the GDP overlapped with the same candidate gene for TKW

Table 23 Grain length candidate genes

| Gene stable ID | Gene start (bp) | Gene end (bp) | Gene description |
|---|---|---|---|
| **TRITD6Av1G223280** | 606954781 | 606960627 | NAC domain protein, G |

Table 24 Grain width candidate genes

| Gene stable ID | Gene start (bp) | Gene end (bp) | Gene description |
|---|---|---|---|
| **TRITD2Bv1G033260** | 79527960 | 79532729 | Alpha-L-arabinofuranosidase 1 |
| **TRITD6Av1G185050** | 525603565 | 525607101 | Glutamine synthetase |
| **TRITD5Bv1G238370** | 668212339 | 668218151 | Squamosa promoter-binding-like protein |
| **TRITD6Av1G012970** | 30130259 | 30132189 | F-box family protein |

Gran width candidate genes belong to arabinofuranosidase (2B-GDP), Glutamine synthetase (6A-GDP), squamosa promoter biding-like protein (TGC-5B) and F-box protein (TGC-6A). Glutamine synthetase is involved in nitrogen assimilation especially during grain filling stage, thus becoming important for grain grain size (Wei et al., 2021)

*Table 25 Grain redness candidate genes*

| Gene stable ID | Gene start (bp) | Gene end (bp) | Gene description |
|---|---|---|---|
| **TRITD6Av1G010380** | 25918178 | 25921574 | GDSL esterase/lipase |

An esterase LIP4 involved in the flavonoid biosynthesis was detected on chromosome 6A in the TGC collection

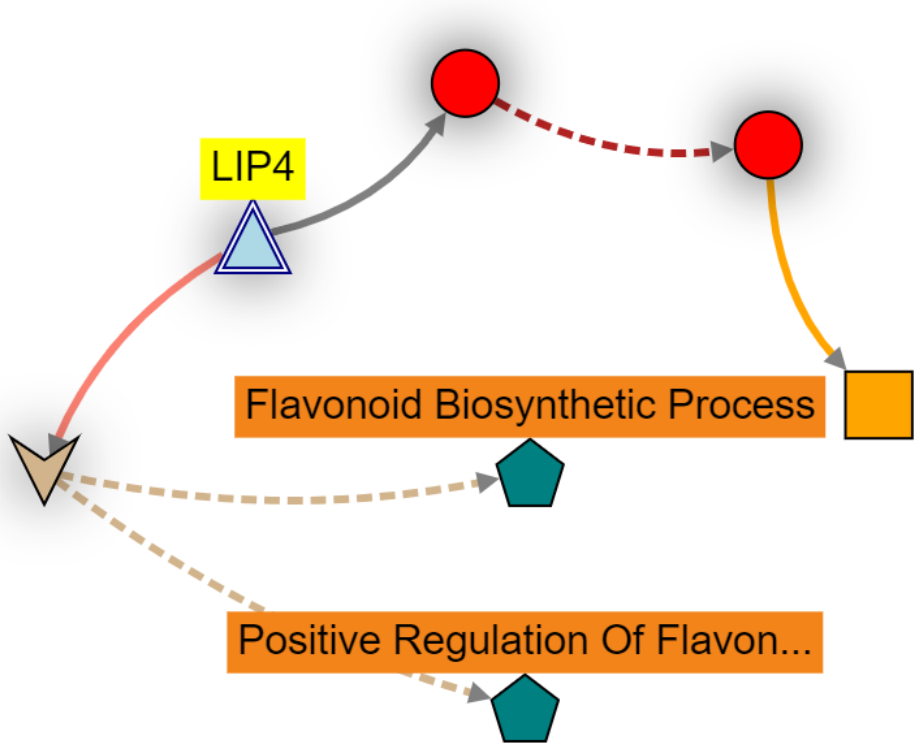


Figure 71 Knetminer network of genes functions for Grain redness trait in TGC

Table 26 Candidate gene for grain yellowness in TGC

| Gene stable ID | Gene start (bp) | Gene end (bp) | Gene description |
|---|---|---|---|
| **TRITD7Bv1G026240** | 72110348 | 72116598 | Pentatricopeptide repeat-containing protein |

A pentatricopeptide repeat-containing protein was identified and it is involved in the carotenoid biosynthetic process.

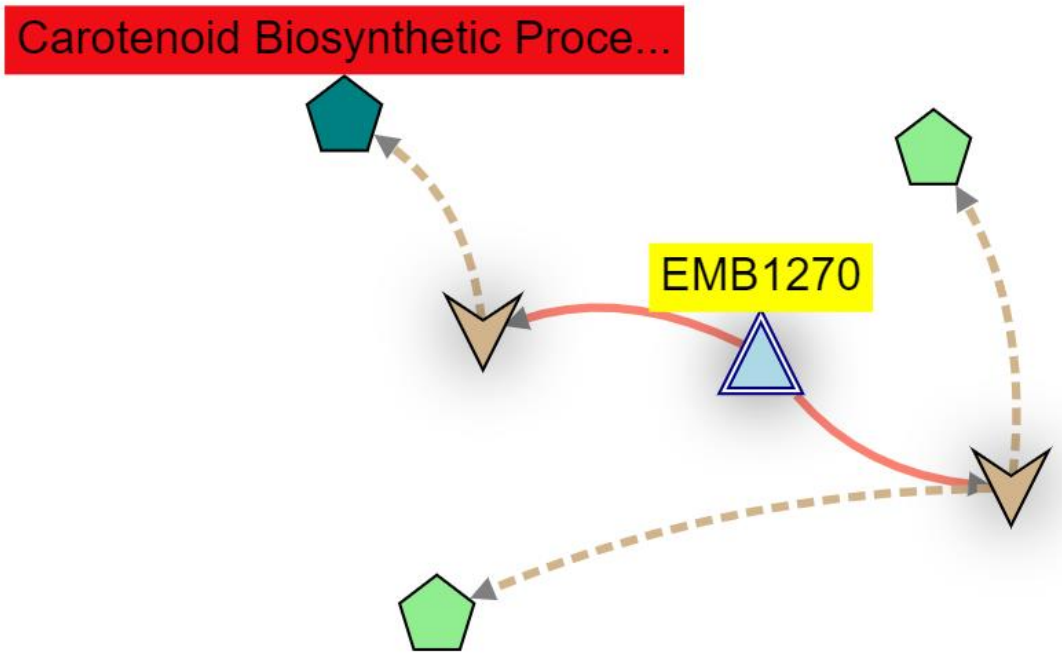


Figure 72Knetminer network of gene function in TGC for grain yellowness

## 2.4 Discussion

Both panel, GDP and TGC, were evaluated in two years each in order to dissect the genetic control for loci of interest regarding yield components and grain quality. As expected variability within each collection was very high, especially in the TGC since its larger dimensions as a panel and the inclusion of a comprehensive group of durum landraces and both domesticated and wild emmer.

From the GWAS analysis, peaks were considered addressing a major QTL when the phenotypic variance explained reached 10%.

An extensive meta-analysis was performed for the QTLs regarding yield traits and its cvomponents. On chromosome 2A a coincidence peak for FS trait among GDP 2020, 2021 and cluster analysis was detected. Its function was linked to a protein kinase, which is usually addressed to grain size and number regulation (Khan et al., 2022). There is bibliographic evidence with QTL1622_2A (Milner et al., 2016), QTL1387_2A (Graziani et al., 2014), QTL0964_2A (Blanco et al., 2012) which all related to kernel number per spike, spikelet number per spike and grain yield. An additional peak on the same chromosome in TGC cluster GWAS resulted coding for a WRKY transcription factor, which has been demonstrated to be linked to an enhanced spikelet number per spike (Khan et al., 2022). The same chromosome region coincides with an already studied QTL from Mangini et al., 2018. For the same trait there was evidence of a third coincidence peak among environments for GDP on the chromosome 4B, which found a validation on the bibliography matching with QTL1076_4B (Distelfeld, unpublished) and QTL0696_KNS (Mangini et al., 2018) for the kernel number per spike trait. For the fertile florets number per central spikelet (FF) trait, a coincidence peak was noticed on chromosome 2A across all GWAS results in the GDP. This found evidence in bibliography as it is close to QTL1842_2A (Roncallo et al., 2017). Additionally, a second QTL regarding this traits was found to be coincidente with a QTL studied by Roncallo et al (2017) on chromosome 7A. Two QTLs regarding TKW were detected on chromosome 1A: one finds coincidence peaks in two environments (Grosseto2021 and cluster analysis) while the second only in Grosseto2020. Both are mentioned by Faris et al (2014). On chromosome 1B a peak regarding TKW was noticed on all GDP environment and TGC 2020. This coincides with QTL1735_1B (Peng et al., 2003). Overlaps on chromosome 2A were found for a QTL on Grosseto2021 and cluster analysis for a QTL studied by Fatiukha et al (2020) and another QTL found on chromosome 2A was reported by Avni et al (2018). Russo et al (2014)

previously detected a QTL on chromosome 3B which overlaps with a QTL detected in the cluster analysis. An additional QTL regarding TKW detected on chromosome 4B in Grosseto2020 was already reported by Patil et al (2013) and Peleg et al (2011). On chromosome 7B a peak was identified in the GDP cluster GWAS and it overlaps with QTL0734_TKW (Mangini et al., 2018), while another peak observed in both Grosseto2021 and cluster analysis was reported by Roncallo et al (2017) and Fatiukha et al (2020). For the Grain Area (KA) trait an overlap between a QTL on chromosome 2A and a QTL reported by Mangini et al (2021) was observed. The same study also reported additional QTLs that matched with QTL here found on chromosomes 3A and 7B. GA QTLs studied by Haugrud et al (2022) overlapped in this work with QTLs on chromosome 4A and 5A, while a QTL found coincidence with Desiderio et al (2019) on chromosome 2B. Three peaks for Grain perimeter on chromosome 2B were already described by Desiderio et al (2019) and a fourth QTL was reported on the same chromosome by Russo et al (2014). As regard Grain length A QTL detected on chromosome 2A in both Grosseto2021 and the cluster analysis was reported previously by Haugrud et al (2022) and Mangini et al (2021). In Grosseto2021, two QTLs found on chromosomes 2B and 4A matched with QTLs reported by Desiderio et al (2019). A peak was found consistently on chromosome 6A in the GDP panel and there is bibliographic evidence close to QTL1361_6A (Golabadi et al., 2011). An additional peak found in both GDP 2020 and GDP 2021 but also in TGC 2020 is located on chromosome 6A, which is a close position to QTL1680_6A (Patil et al, 2013) related to grain yield, QTL1414_6A (Graziani et al., 2014) involved in test weight and QTL0719_TKW, QTL1729_6A (Mangini et al., 2018; Peleg et al., 2011) both associated with kernel weight. Regarding Grain width, two peaks on chromosome 1B detected in Grosseto2020 and the cluster analysis were already reported by Mangini et al (2021). The same work also reported a QTL found in the cluster analysis on chromosome 6B. Haugrud et al (2022) reported two QTLs found on 7B and one on 3A, while Russo et al (2014) reported a QTL detected on 3B. In this study peaks that did not match with QTLs previously studied within the literature were also detected: QFF.ubo_2B.5_multiENV, QFF.ubo_4B_Gro2021_&_multiENV for the fertile florets trait. QTKW.ubo.2A.2_multiENV, QTKW.ubo_3B.3_multiENV for the TKW. QKA.ubo_1A_Gro2021_&_multiENV, QKA.ubo_7B.2_Gro2021_&_multiENV for Grain area. QKP.ubo_2A.2_Gro2021_&_multiENV, QKP.ubo_6A_Gro2021_&_multiENV for Kernel perimeter. QKL.ubo_3A_Gro2021_&_multiENV , QKL.ubo_6A_Gro2021_&_multiENV for kernel length. QKW.ubo_2B.2_Gro2021_&_multiENV,

QKW.ubo_3A.2_Gro2021_&_multiENV, QKW.ubo_3B.3_Gro2021_&_multiENV, QKW.ubo_6°.2_Gro2021_&_multiENV for kernel width.

## 2.5 Bibliography

Alexander, David H., John Novembre, and Kenneth Lange. "Fast model-based estimation of ancestry in unrelated individuals." Genome research 19.9 (2009): 1655-1664.

Avni, Raz, et al. "Genome based meta-QTL analysis of grain weight in tetraploid wheat identifies rare alleles of GRF4 associated with larger grains." Genes 9.12 (2018): 636.

Blanco, A., et al. "Relationships between grain protein content and grain yield components through quantitative trait locus analyses in a recombinant inbred line population derived from two elite durum wheat cultivars." Molecular Breeding 30.1 (2012): 79-92.

Bradbury, Peter J., et al. "TASSEL: software for association mapping of complex traits in diverse samples." Bioinformatics 23.19 (2007): 2633-2635.

Brinton, J., & Uauy, C. (2019). A reductionist approach to dissecting grain weight and yield in wheat. In Journal of Integrative Plant Biology (Vol. 61, Issue 3, pp. 337–358). Blackwell Publishing Ltd. https://doi.org/10.1111/jipb.12741

Browning, Brian L., et al. "Fast two-stage phasing of large-scale sequence data." The American Journal of Human Genetics 108.10 (2021): 1880-1890.

Chang, Christopher C., et al. "Second-generation PLINK: rising to the challenge of larger and richer datasets." Gigascience 4.1 (2015): s13742-015.

David, Jacques, et al. "Genotyping by sequencing transcriptomes in an evolutionary pre-breeding durum wheat population." Molecular breeding 34.4 (2014): 1531-1548.

Desiderio, Francesca, et al. "Genomic regions from an Iranian landrace increase kernel size in durum wheat." Frontiers in plant science 10 (2019): 448.

Faris, Justin D., et al. "Analysis of agronomic and domestication traits in a durum× cultivated emmer wheat population using a high-density single nucleotide polymorphism-based linkage map." Theoretical and applied genetics 127 (2014): 2333-2348.

Fatiukha, Andrii, et al. "Grain protein content and thousand kernel weight QTLs identified in a durum× wild emmer wheat mapping population tested in five environments." Theoretical and Applied Genetics 133 (2020): 119-131.

Geng, J., Li, L., Lv, Q. et al. TaGW2-6A allelic variation contributes to grain size possibly by regulating the expression of cytokinins and starch-related genes in wheat. Planta 246, 1153–1163 (2017). https://doi.org/10.1007/s00425-017-2759-8

Golabadi, M., et al. "Identification of microsatellite markers linked with yield components under drought stress at terminal growth stages in durum wheat." Euphytica 177.2 (2011): 207-221.

Guo, Zifeng, Gustavo A. Slafer, and Thorsten Schnurbusch. "Genotypic variation in spike fertility traits and ovary size as determinants of floret and grain survival rate in wheat." Journal of experimental botany 67.14 (2016): 4221-4230.

Graziani, M., et al. "QTL dissection of yield components and morpho-physiological traits in a durum wheat elite population tested in contrasting thermo-pluviometric conditions." Crop and Pasture Science 65.1 (2014): 80-95

Khan, Hanif, et al. "Genome-wide association study for grain yield and component traits in bread wheat (Triticum aestivum L.)." Frontiers in Genetics 13 (2022): 982589

Khan, Nadia, et al. "TaGSNE, a WRKY transcription factor overcomes the tradeoff between grain size and grain number and associates with root development in common wheat." Journal of experimental botany (2022): erac327.

Li, Y., Fan, C., Xing, Y. et al. Natural variation in GS5 plays an important role in regulating grain size and yield in rice. Nat Genet 43, 1266–1269 (2011). https://doi.org/10.1038/ng.977

Ma, L.; Li, T.; Hao, C.; Wang, Y.; Chen, X.; Zhang, X. TaGS5-3A, a grain size gene selected during wheat improvement for larger kernel and yield. Plant Biotechnol. J. 2015, 14, 1269–1280.

M. Feldman, “Origin of Cultivated Wheat,” In: A. P. Bonjean and W. J. Angus, Ed., The World Wheat Book: A History of Wheat Breeding, Intercept Ltd., London, 2001, pp. 3-56.

Maccaferri, Marco, et al. "A genome-wide association study of resistance to stripe rust (Puccinia striiformis f. sp. tritici) in a worldwide collection of hexaploid spring wheat (Triticum aestivum L.)." G3: Genes, Genomes, Genetics 5.3 (2015): 449-465.

Maccaferri, Marco, et al. "Durum wheat genome highlights past domestication signatures and future improvement targets." Nature genetics 51.5 (2019): 885-895.

Mangini, Giacomo, et al. "Genetic dissection of the relationships between grain yield components by genome-wide association mapping in a collection of tetraploid wheats." PloS one 13.1 (2018): e0190162.

Mangini, Giacomo, et al. "Candidate genes and quantitative trait loci for grain yield and seed size in durum wheat." Plants 10.2 (2021): 312

Matsuoka, Yoshihiro. "Evolution of polyploid Triticum wheats under cultivation: the role of domestication, natural hybridization and allopolyploid speciation in their diversification." Plant and cell physiology 52.5 (2011): 750-764.

Mazzucotelli, Elisabetta, et al. "The Global Durum Wheat Panel (GDP): An international platform to identify and exchange beneficial alleles." Frontiers in plant science 11 (2020): 569905.

Milner, Sara Giulia, et al. "A multiparental cross population for mapping QTL for agronomic traits in durum wheat (T riticum turgidum ssp. durum)." Plant Biotechnology Journal 14.2 (2016): 735-748.

Patil, R. M., et al. "Mapping of QTL for agronomic traits and kernel characters in durum wheat (Triticum durum Desf.)." Euphytica 190.1 (2013): 117-129.

Peleg, Zvi, et al. "Genetic analysis of wheat domestication and evolution under domestication." Journal of experimental botany 62.14 (2011): 5051-5061.

Peng, Junhua, et al. "Domestication quantitative trait loci in Triticum dicoccoides, the progenitor of wheat." Proceedings of the National Academy of Sciences 100.5 (2003): 2489-2494.

Peters Haugrud, Amanda R., et al. "Identification of stable QTL controlling multiple yield components in a durum× cultivated emmer wheat population under field and greenhouse conditions." G3 13.2 (2023): jkac281.

Ral, Jean-Philippe, et al. "Down-regulation of Glucan, Water-Dikinase activity in wheat endosperm increases vegetative biomass and yield." Plant biotechnology journal 10.7 (2012): 871-882.

Roncallo, Pablo F., et al. "QTL mapping and analysis of epistatic interactions for grain yield and yield-related traits in Triticum turgidum L. var. durum." Euphytica 213.12 (2017): 1-20.

Russo, Maria Anna, et al. "A dense durum wheat× T. dicoccum linkage map based on SNP markers for the study of seed morphology." Molecular breeding 34 (2014): 1579-1597.

Sahri, Ali, et al. "Towards a comprehensive characterization of durum wheat landraces in Moroccan traditional agrosystems: analysing genetic diversity in the light of geography, farmers' taxonomy and tetraploid wheat domestication history." BMC evolutionary biology 14.1 (2014): 1-18.

Sakuma, S.; Golan, G.; Guo, Z.; Ogawa, T.; Tagiri, A.; Sugimoto, K.; Bernhardt, N.; Brassac, J.; Mascher, M.; Hensel, G.; et al. Unleashing floret fertility in wheat through the mutation of a homeobox gene. Proc. Natl. Acad. Sci. USA 2019, 116, 5182–5187

Salamini, Francesco, et al. "Genetics and geography of wild cereal domestication in the near east." Nature Reviews Genetics 3.6 (2002): 429-441.

Serrote, Caetano Miguel Lemos, et al. "Determining the Polymorphism Information Content of a molecular marker." Gene 726 (2020): 144175.

Simmonds, J., Scott, P., Brinton, J. et al. A splice acceptor site mutation in TaGW2-A1 increases thousand grain weight in tetraploid and hexaploid wheat through wider and longer grains. Theor Appl Genet 129, 1099–1112 (2016). https://doi.org/10.1007/s00122-016-2686-2

Simons, Kristin J., et al. "Molecular characterization of the major wheat domestication gene Q." Genetics 172.1 (2006): 547-555.

Tillett, Brandon J et al. "Genes Impacting Grain Weight and Number in Wheat (Triticum aestivum L. ssp. aestivum)." Plants (Basel, Switzerland) vol. 11,13 1772. 4 Jul. 2022, doi:10.3390/plants11131772

Wang, Jiabo, and Zhiwu Zhang. "GAPIT Version 3: boosting power and accuracy for genomic association and prediction." Genomics, proteomics & bioinformatics 19.4 (2021): 629-640.

Wang, Shichen, et al. "Characterization of polyploid wheat genomic diversity using a high-density 90 000 single nucleotide polymorphism array." Plant biotechnology journal 12.6 (2014): 787-796.

Wei, Yihao, et al. "Localization, gene expression, and functions of glutamine synthetase isozymes in wheat grain (Triticum aestivum L.)." Frontiers in plant science 12 (2021): 580405.

Whan, Alex P., et al. "GrainScan: a low cost, fast method for grain size and colour measurements." Plant methods 10.1 (2014): 1-10.

Wittern, L.M.; Barrero, J.M.; Bovill, W.D.; Verbyla, K.L.; Hughes, T.; Swain, S.M.; Steed, G.; Webb, A.A.; Gardner, K.; Jacobs, J.; et al. Overexpression of the WAPO-A1 gene increases the number of spikelets per spike in bread wheat. bioRxiv 2022

# 3 CHAPTER II

## 3.1 Introduction

### 3.1.1 Adaptive traits for sustainable agriculture

In recent years, the world agricultural production has substantially increased because of the rise in food demand determined by the significant growth in world population. This process has fostered an intensification of agricultural practices in high-input environments granted with larger usage of irrigation and fertilizers and paramount cultivation of plant breeding with higher yields (Koevoets et al., 2016). However, this tendency would lead to irreversible damages to the environment. Several are the factors at play. Firstly, the current, agricultural, irrigation and fertilization management is not sustainable, as it is responsible for almost 70% of freshwater withdrawals in the world (Rosengrant et al., 2009). The water consumption caused by this unsustainable irrigation system combined with the growing non-agricultural request of fresh-water, will lead to a dangerous scarcity in agro-systems. It also might be mention in addition to this, that excessive irrigation creates fertilizers leaching. Another unsustainable agricultural practice concerns deep tilling cultivation. Deep tilling practices cause massive greenhouse gas emission (Snyder et al., 2009). A boost towards a more sustainable agriculture has become indispensable in order to avoid further harm to the environment. A feasible more sustainable alternative to current agricultural systems is Conservation agriculture or CA. Conservation agriculture is defined as an approach to agriculture aimed at minimalizing soil disturbance through permanent soil cover and crop rotations (Hobbs et al., 2008). CA practices have proven to lead to an improvement in soil health and relative biotic factors, and a decrease in fertilizers' employment. This means that agriculture will have to face crop production under suboptimal conditions, forming a gap between the yield obtained under high-input traditional agriculture (called potential yield) and the current yield (Koevoets et al., 2016). In this context, landraces could be exploited as a genetic source of favourable genes, in order to incorporate them into elite cultivars.

### 3.1.2 Landraces: general features

For what the domesticated species are concerned, the Landraces, or Traditional Varieties, are widely recognised as dynamic entities characterized by genetic diversity. According to Zeven (1998), the complex nature of these entities constitutes an insurmountable obstacle to the formulation of a coherent and conclusive definition of Landraces. However, recently Villa et al. (2005) introduced the following definition: "a landrace is a dynamic population(s) of a cultivated plant that has historical origin, distinct identity and lacks formal crop improvement, as well as often being genetically diverse, locally adapted and associated with traditional farming systems" If we follow this definition, Landraces are individuated by the characteristic long time of development and their relation to specific geographical areas. It is in these locations that the Landraces adapt to the local specific agroecosystems, which in most cases are typified by restrictive environmental conditions. These processes of adaptation cause changes in genotypes frequencies and, as a result, modifications in phenotypes. This adaptability made landraces more suitable for cultivation in suboptimal conditions- e.g. under abiotic, biotic and, human factors- than any other modern cultivars. Landraces are distinguished from modern (or elite) cultivars. Modern (or elite) cultivars, in fact, result from a breeding programme, involving controlled artificial crosses and subsequent progeny selection up to development of superior, pure and homogeneous varieties. More specifically, in the case of self-pollinating species - e.g. durum wheat- cultivars are bred to be genetically homogeneous, pure lines. They are developed to present an increased yield and thus they are employed in a traditional, high-input based agriculture. On the opposite, landraces are genetically diverse, comprising many, different, homozygous lines. Their genetic diversity is twofold: diversity between site is caused by reproductive isolation and diversity within sites which is associated with climate changes and biological factors (Villa et al., 2005). Also, Landraces differ from modern cultivars in terms of origin, the selection that originates Landraces characteristically lacks a formal genetic improvement and Landraces undergo a natural selection subjected to unintentional human contribution, e.g. for seed traits (Villa et al., 2005). Finally, Landraces' high genetic diversity- higher compared to the elite cultivars'- is effective against abiotic as well as biotic stresses (Sahri et al. 2014).

In these terms, a thorough examination of the landraces' diversity becomes relevant in order to identify the genes which are responsible for yield stability and consequently for resistance to diseases and resilience to drought and low nutrient environments

### 3.1.3 Root system anatomy

A fibrous root system is a distinctive characteristic of the Poaceae family and wheat, as representative of all the small grain cereals in general. More in detail, wheat has an embryonic and a post-embryonic root system (Fig.4). The embryonic portion generates from the embryo and emerges at germination, showing a changeable number of seminal roots and the primary root. Embryonic and post-embryonic roots develop lateral roots. The post-embryonic part comprises the crown roots, also known as nodal roots, which originate from the lower part of the stem.

From a transversal prospective, a single root's primary structure can be observed. It comprises three sections: the epidermis, the cortex and the vascular cylinder (stele). The epidermis consists of a single cells layer, whereas the cortex presents multiple layers of parenchymatic cells (Rossini et al., 2018). The innermost section of the cortex is the endodermis layer. What characterises the Endodermis layer of absorbing roots is the presence of the Casparian strips, a region characterised by hydrophobic properties. The suberin and the lignin contained in the Casparian strips hinder with the passage of water and solutes through the endodermis (Esau et al., 2006). The outermost layer of the vascular cylinder is called pericycle and it is surrounded by endodermis. Finally, xylematic vessels can be observed in the inner section of the stele.

A lengthwise dissection from the root tip of each root, instead, shows firstly, a meristematic section in the root cap zone, consisting of meristematic cells in active division; secondly, an elongation section; and, thirdly, a differentiation section, which presents lateral roots formation

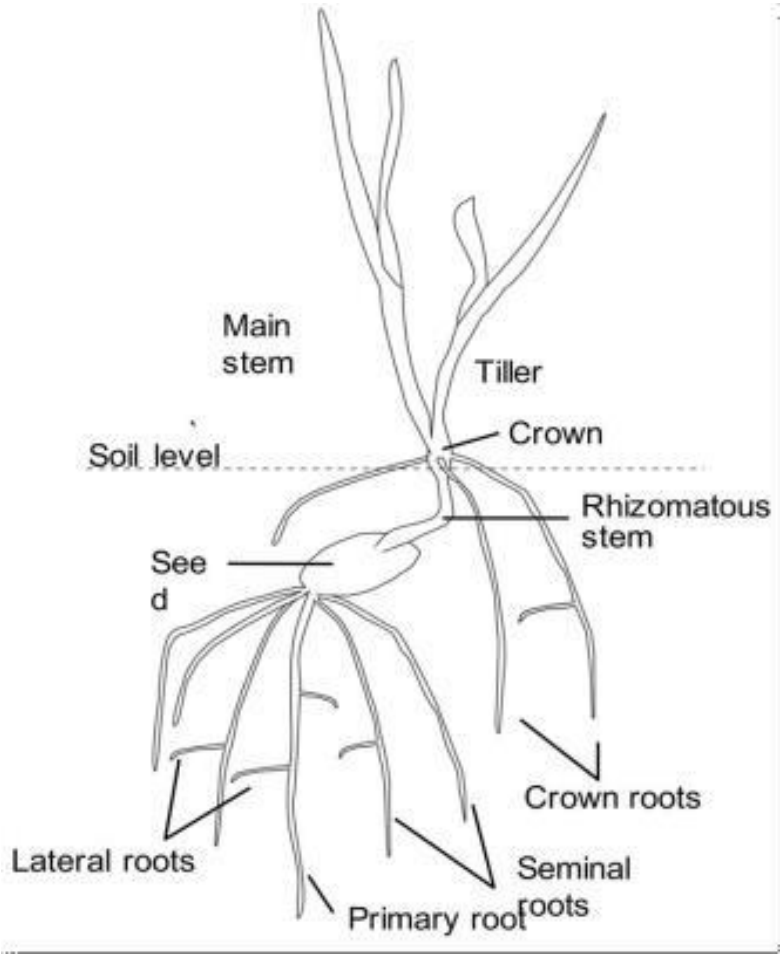


Figure 73 Plant root system morphology. Source: Rossini et al., 2018

### 3.1.4 Root system architecture

The main breeding programmes have mostly focused on shoots' selection, neglecting the portions of the plants growing below the ground's surface. However, research into breeding has shown that elite cultivars tend to present small root systems, smaller than landraces, and that they tend to produce higher yield if cultivated under optimal nutritive conditions. Significantly, Siddique et al., (1990), observed that the root-shoot ratio is significantly lower in modern wheat cultivar when compared to the landraces. Because a lower ratio implies a smaller root system size, modern breeding practices switched their focus on increasing the Harvest Index, selecting new varieties with a faster and earlier growth, to favour the development of the shoots rather than the roots themselves. With the aim of increasing yield's stability under suboptimal conditions, more attention has been dedicated to the study of the optimization of root system architecture (Koevoets et al., 2016).

The root system architecture (RSA) is the three-dimensional configuration of the roots of a plant (Lynch., 1995). The root architecture is strictly dependant on "roots' distribution" and "roots' topology". By "roots' distribution" we regard the root manifestation along a positional gradient, and by "roots' topology" we mean the degree of articulation of each singular root axe. The spatial disposition of the roots is determined by several factors: root length, root growth angle. However, the factors, which affect the root system architecture dramatically, are the location in the vertical gradient and the number of the roots (Fig.5) (Koevoets et al., 2016).

The root system architecture has multiple features affecting the root functions. First of all, as the soil's resources- e.g., water and nutrients- are usually unevenly distributed, the root system architecture, determing the roots disposition in a certain volume of soil, directly affects the plant's capacity to adjust to and to exploit the soil resources (i.e., root plasticity Lynch., 1995). Secondly, RSA is directly entailed in the mechanical support of the above- ground part of the plant (Ennos & Fitter,, 1992), determining lodging resistance/susceptibility. Thirdly, it is strictly associated with water and solutes transport capacity. Finally, it might affect the extent and amount of root interactions with soil micro-organisms, thus creating C fluxes (Wullschleger et al., 1994).

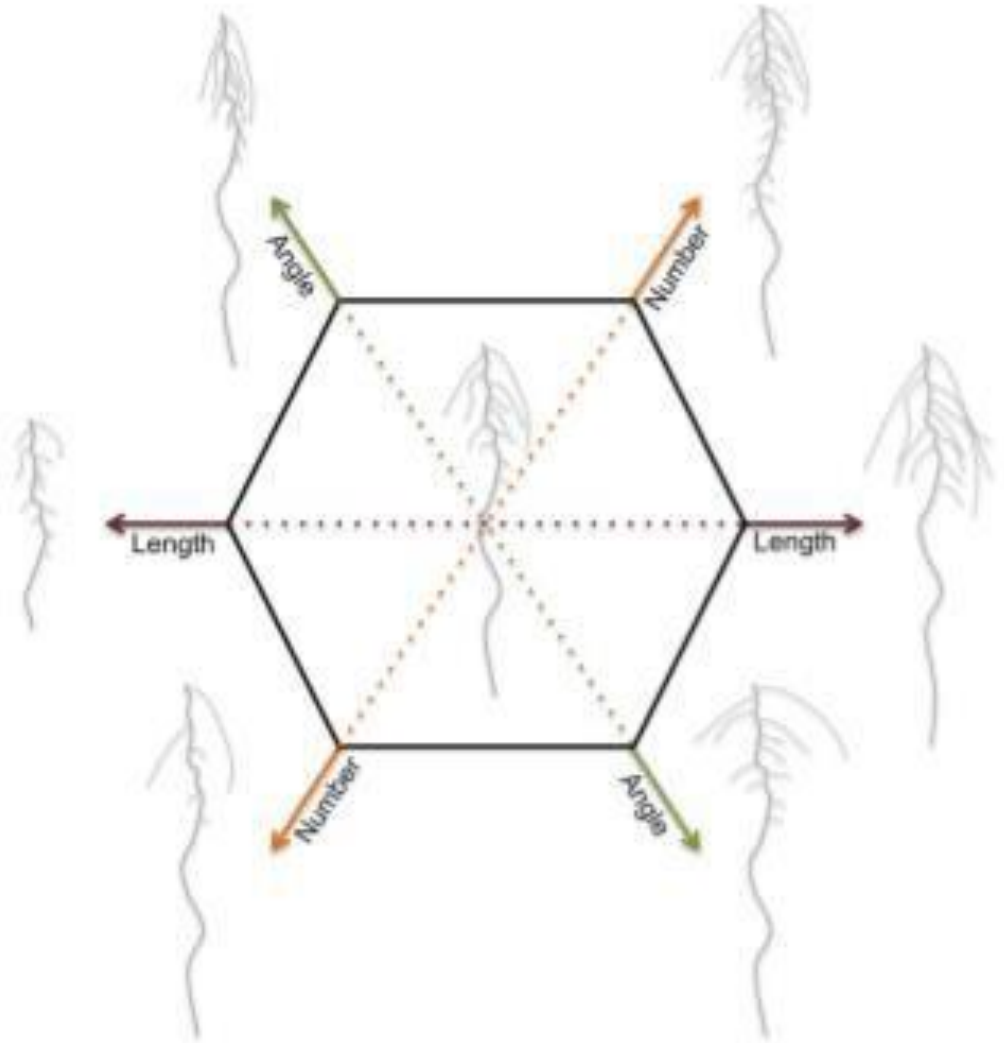


Figure 74 Main traits affecting root spatial configuration or RSA. Each trait is represented in both extreme phenotype forms. Source: Koevoets et al., 2016

### 3.1.5 Root growth angle and relative ideotypes

Very recently, the growing interest in the RSA fostered the development of different ideotypes. Central concern regarding these ideotypes is the root growth angle (RGA). By RGA we understand the distance calculated between the two outermost roots of the whole root system of a single plant (Maccaferri et al., 2016) The Root growth angle is a trait relevant for cereal crops water uptake and nitrogen foraging, since it allows a double exploration of the soil: vertically and horizontally. As previously mentioned, soil resources, which comprises water and nutrients, are present variable patterns of distribution in the soil profile. Soil resources follow in a heterogeneous scheme a vertical and horizontal distribution (Koevoets et al., 2016). The vertical distribution usually involves nutrients' accumulation under the aboveground part of the plants, while the horizontal refers to the distribution of nutritive elements caused by leaching and plant cycling. Nutrients characterised by a low mobility, like phosphate (PO43-) tends to concentrate in the topsoil layer, whereas water and mobile nutrients like nitrate (NO3-) are prone to leaching, thus hoarding in the deeper soil layers (Jobbàgy & Jackson., 2001). A first ideotype suitable for cereal crops is the "Topsoil foraging". The "Topsoil foraging" (Fig.6) is a successful plant adaptation to low-phosphorous environments, as detected in Phaseolus vulgaris (Lynch-Brown., 2001). In common bean, several tools for phosphate mobilization and uptake enhancing have been observed as symbiotic formation with soil microbiota (mycorrhizas) or as the exudation of organic acid and phosphatases (Lynch., 1995). A different root system architecture was detected and improved here, and P's uptake efficiency was demonstrated. In this case, the root system undergoes a shallow distribution showing a wide root growth angle, thus exploring the uppermost soil zone where P is pre-eminently accumulated, as well as presenting a strong lateral root growth together with root hairs proliferation (Williamson., 2001). A second ideotype suitable for cereal crops is the "Steep, cheap and deep" (Fig 7) (Lynch., 2013). This ideotype is characterised by a long, thick primary root with few, long, lateral roots and a seminal root system with a narrow growth angle. The primary root's thickness inables a proper soil penetration, particularly effective trait in case of hard soils. The long, but few lateral roots, instead, are useful for a co-optimization strategy to acquire both nitrates and phosphates (Lynch., 2013). The seminal roots in this study case might present two options: on the one hand, seminal roots show a broad growth

angle and a proliferation of root hairs which enables to exploit top-soil resources, on the other, roots are thicker and grow at a narrower

angle with a lesser lateral branching, hence exploring the deeper soil layers (earlier than crown roots) contributing to the water uptake (Manschadi et al., 2013). In this second eventuality, a proper topsoil development of crown root is required in order to provide an appropriate phosphate acquirement. The deep rooting ideotype aforementioned is the resulting adaptation to environments characterised by the scarcity of mobile nutrients such as nitrates and water. Hence the SCD ideotype is perfectly adequate in rainfed agricultural conditions or in nitrate limited environments (Koevoets et al., 2016).

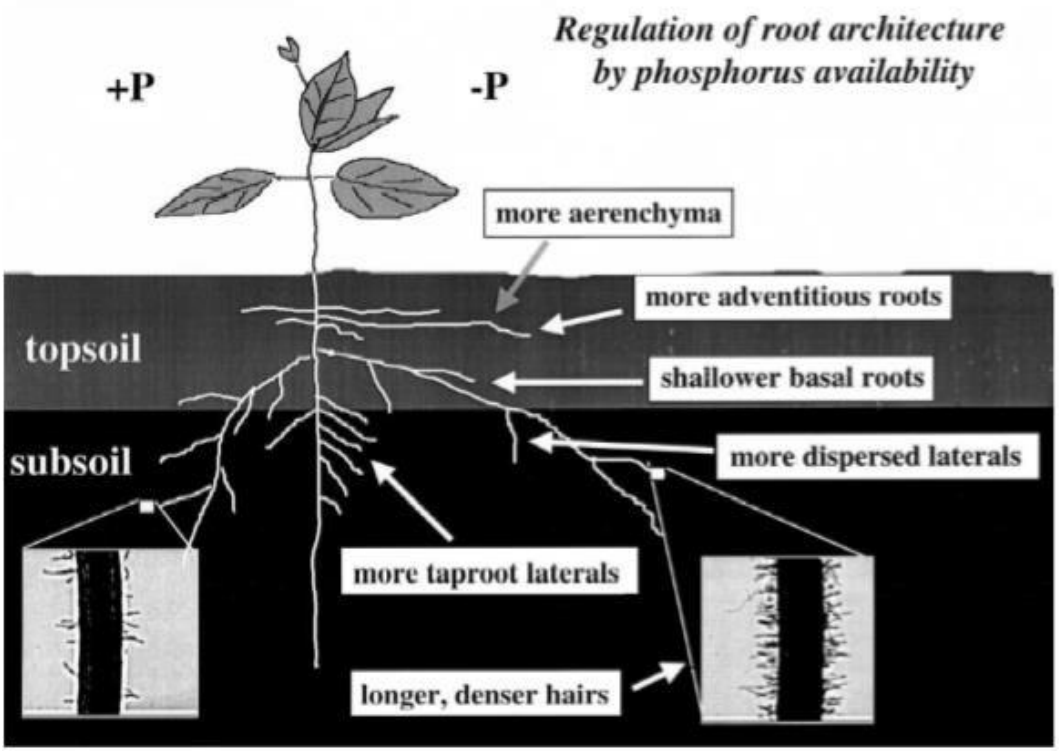


Figure 75 Topsoil foraging ideotype for P acquirement. Source: Lynch., 2001

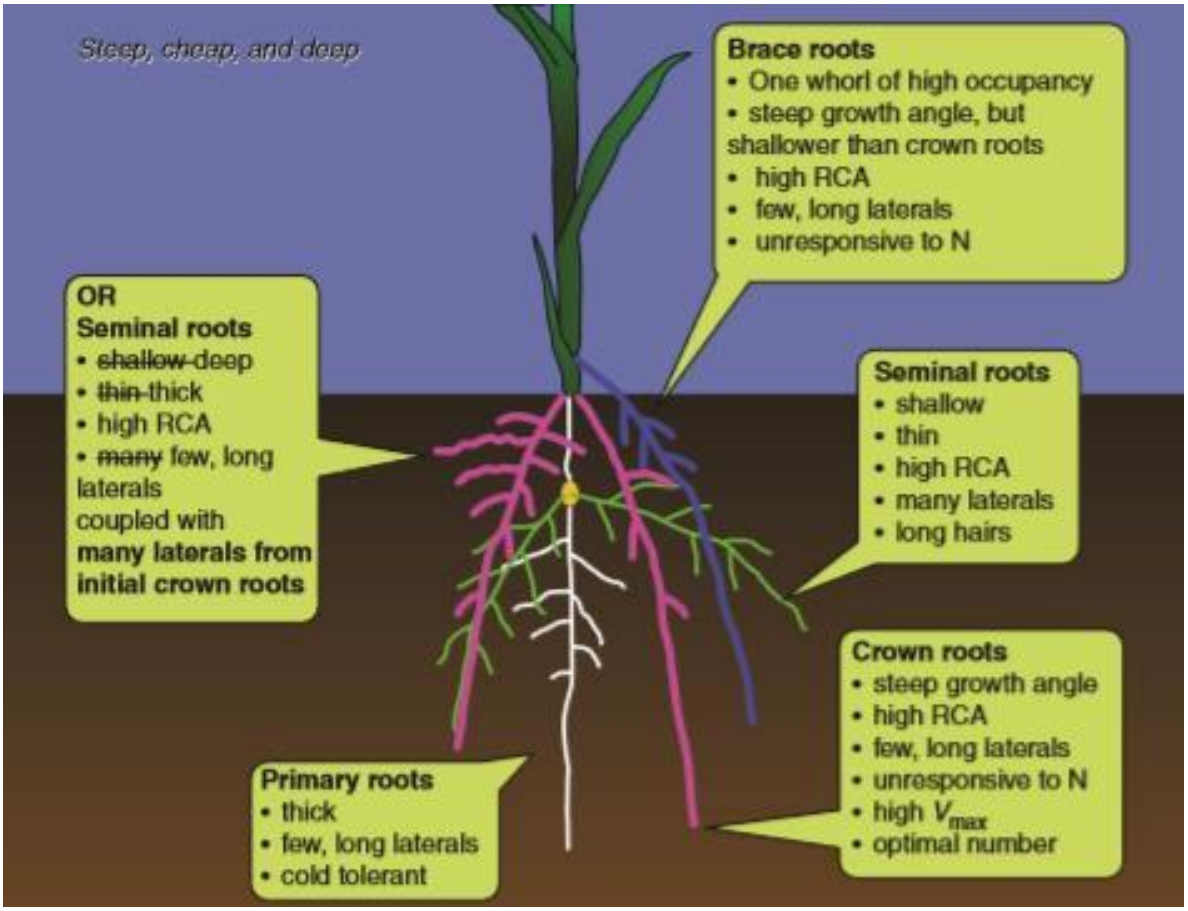


Figure 76 Steep, cheap and deep ideotype for N and water uptake. Source: Lynch, J.P., 2013

### 3.1.6 Root system architecture phenotyping methods

Phenotypic root assessments are needed to learn more about the root system architecture. In facts roots develop in a belowground solid substrate which is the soil, thus, hindering a proper phenotypic evaluation and often leading to damaging the original root system structure. As a consequence, it becomes impossible to carry out further phenotypic assessments on that same individual. New phenotyping techniques are developed with the aim of allowing proper observations. Phenotypic root assessments' methods are usually classified in two categories: ex situ or in situ. However, two additional distinctions could be added to categorise such methodologies: static (single individual screening) or dynamic (several evaluations on the same individual in different times) (Meister et al., 2014). Ex situ methodologies allow fast evaluations of root measures, by extracting the roots from the substrate of cultivation. Ex situ methodologies are thus considered in a static. On the other hand, in situ methodologies provide root images directly from the growth medium, therefore enabling dynamic assessments. The latter comprise novel platforms based on transparent growth media which facilitate a harmless roots removal. Alternatively, there are the hydroponics and aeroponics methods (Zobel et al., 1976). Both of them suit roots with high-throughput phenotyping, but that lack substrate resistance to roots development. In other word, those roots that do not reflect the actual behaviour of the plant in field conditions (Koevoets et al., 2016).

Finally, soil-filled rhizotrons are employed for RSA characterization, because they provide both more realistic roots development data and more accurate measurements during further phenologic phases (Meister et al., 2014).

To conclude, new methodologies -e.g. X-ray tomography and nuclear magnetic resonance (NMR)- were introduced (Hillnhutter et al., 2012), but their diffusion is still limited due to the costs involved in the purchase of the equipment.

## 3.2 MATERIALS AND METHODS

### 3.2.1 Phenotypic analysis

The Global Durum Panel (GDP), already genotyped extensively with the Infinium iSelect Illumina 90K SNP array thus allowing QTL analysis as a result. Accessions were characterized for RGA during the 2019 and 2020 years at seedling stage. Eleven seeds were selected based on kernel uniformity, then sterilised in a 5% sodium hypochlorite solution for 5 minutes and rinsed in distilled water. The seeds were then placed in Petri dishes imbued with distilled water and submitted to pre-germination in incubator for 24 h at 28°C.

Once the pre-germination time was over, the sprouting seeds were removed from the incubator to be subsequently grown on blue cardboard in a semi-hydroponic fashion. One line was drawn on each paper sheet 2 cm from the top border of the sheet. For each genotype 6 seeds were selected from the starting 11 submitted to the pre-germination protocol. Bigger seeds were preferred because they present a more abundant nutrient storage. However, the selection also took in consideration the length of the seminal root. In fact, for a seminal root not to receive damages when placed on the plate or during the growth phase, it has to be short and slightly emerged. For each genotype selected seeds were arranged on the drawn line of one cardboard spaced 8 cm from each other and 5 cm from the lateral border. They were arranged with the ventral furrow towards the surface of the paper sheet and the seminal sprouting root pointed downward. Then a second thinner filter paper sheet soaked in distilled water was laid on the cardboard, therefore covering the germinated seeds. Thereafter the two layer of filter paper were fastened to each other with clip supports.

Wet cardboard were then placed vertically in plastic boxes and attached to each other in order to avoid the passing of light. Accessions were then grown in growth chamber for 7 days at 22°C under 16 h light photoperiod. After the growth period, photos were taken of the plants root system, which were subsequently analysed through ImageJ software to acquire data regarding root angle. RGA was then acquired as a mean among the 6 plants for each genotype in each replicate.

The experiment was carried out following a randomized complete block design for two replicates. Accessions were divided into 34 blocks and 3 checks were included in each block. Colosseo, Lloyd and Svevo were chosen as checks in this experiment. Colosseo is an Italian cultivar showing a good yield potential but with a low level of adaptation to arid environmental conditions typical of the Southern Mediterranean basin. Lloyd is a Northern American cultivar adapted to low input agricultural conditions, while Svevo is an Italian early-flowering cultivar and well adapted to the Mediterranean environment.

ANOVA was performed considering blocks, replicates and technical replicates for each genotype and data were linear-adjusted for block effect with R, thus obtaining BLUEs for a subsequent GWAS analysis.

### 3.2.2 GWAS

R package GAPIT3 was used to perform GWAS on RGA trait and the following models were included: GLM (naive, MLM + K, MLMM + K, FarmCPU and Blink.

Model selection was set to false and PCA number was set on zero. Results of the GWAS were portrayed in Manhattan plot graphs.

## 3.3 RESULTS

### 3.3.1 Phenotypic analysis

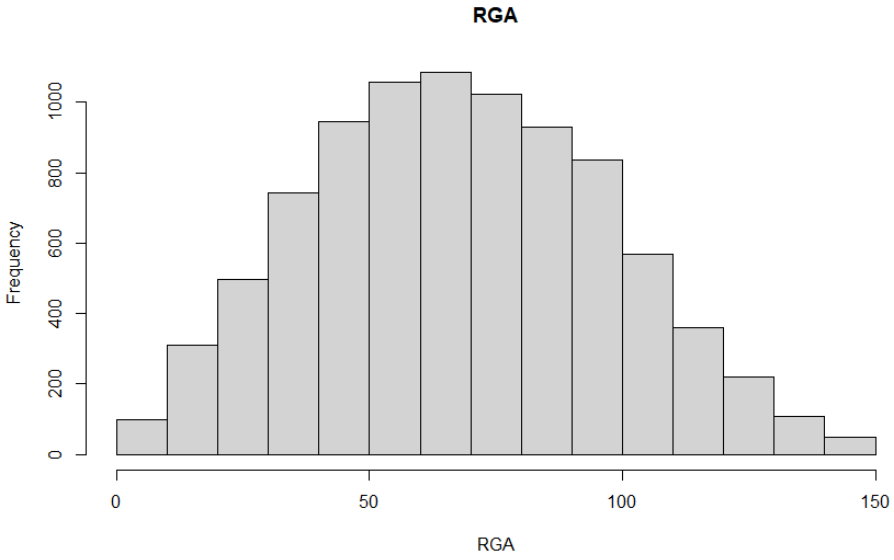


Figure 77 Distribution frequency of Global Durum Panel RGA raw data

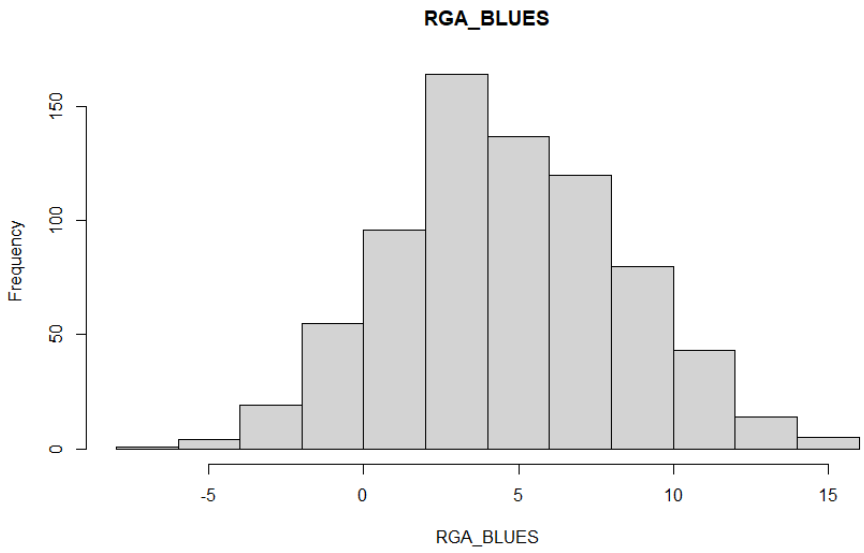


Figure 78 Distribution frequency of Global Durum Panel RGA BLUEs data

Table 27 Descriptive statistics for RGA linear adjusted data

| | RGA |
|---|---|
| min | 20.15 |
| max | 116.39 |
| range | 96.24 |
| median | 67.19 |
| mean | 67.77 |
| SE.mean | 0.59 |
| var | 265.85 |
| std.dev | 16.3 |
| coef.var | 0.24 |
| $h^2$ | 0.806976294 |

Table 28 ANOVA results for RGA data

| Trait | Variables | Sum Sq | Df | F value | Pr(>F) | |
|---|---|---|---|---|---|---|
| **RGA** | **Genotype** | 2397465 | 764 | 4.8969 | 2.20E-16 | *** |
| RGA | **Plant** | 72243 | 5 | 22.5472 | 2.20E-16 | *** |
| RGA | **Block** | 256079 | 32 | 12.4879 | 2.20E-16 | *** |
| RGA | **Replicate** | 33709 | 1 | 52.6035 | 4.46E-13 | *** |
| RGA | **Residual** | 5191264 | 8101 | | | |

In table 24 a summary of the descriptive statistic is shown as well as ANOVA result in table 25. RGA in the GDP has been demonstrated to follow a normal trend as depicted in histogram distribution of the trait. Both raw data and BLUEs distribution frequencies are represented.

Statistics indices are referred to the linear adjusted values. The highest RGA value is 116.39 while the minimum recorded RGA is 20.15. The panel showed a high heritability value (0.80).

ANOVA results showed that technical replicates within genotype (plant), blocking and replicates had a very significant interaction with the genotype.

### 3.3.2 GWAS

For the RGA trait, a GWAS was performed using the GAPIT3 pipeline, which takes into account the kinship among accessions. Graphical results of the analysis are portrayed here in Manhattan plots in Figures 81-85.

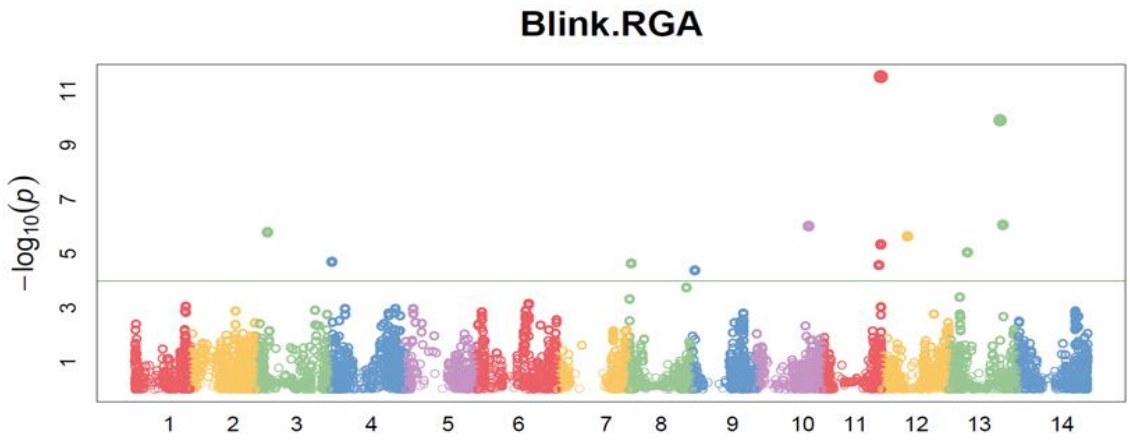


Figure 79 Blink Manhattan plot of GWAS for RGA.

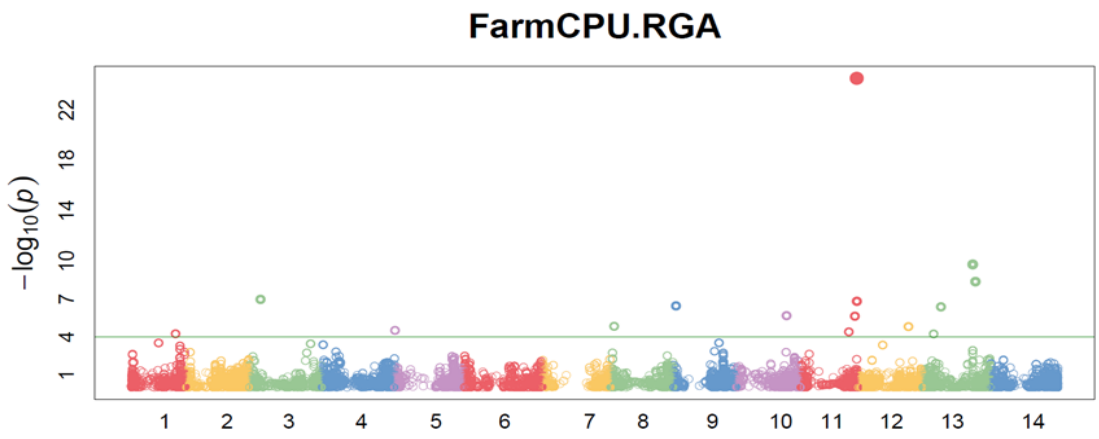


Figure 80 FarmCPU Manhattan plot of GWAS for RGA

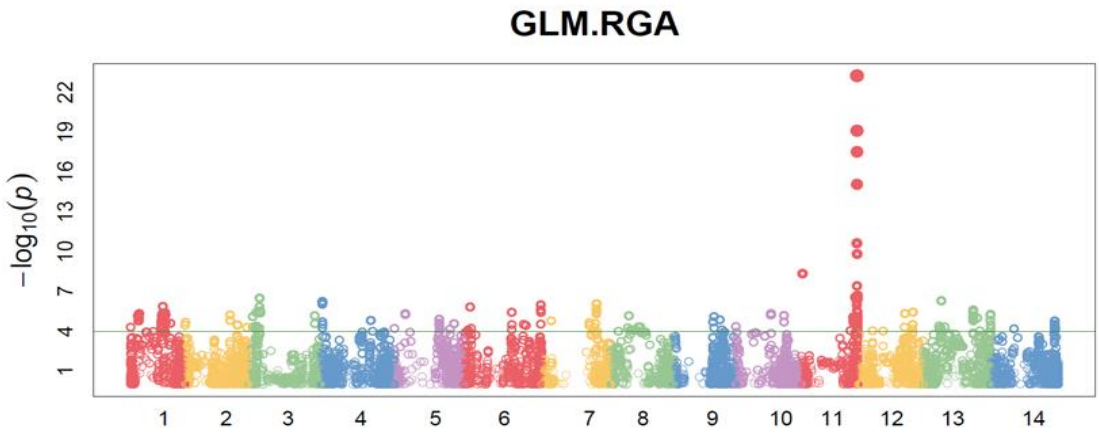


Figure 81 GLM Manhattan plot of GWAS for RGA

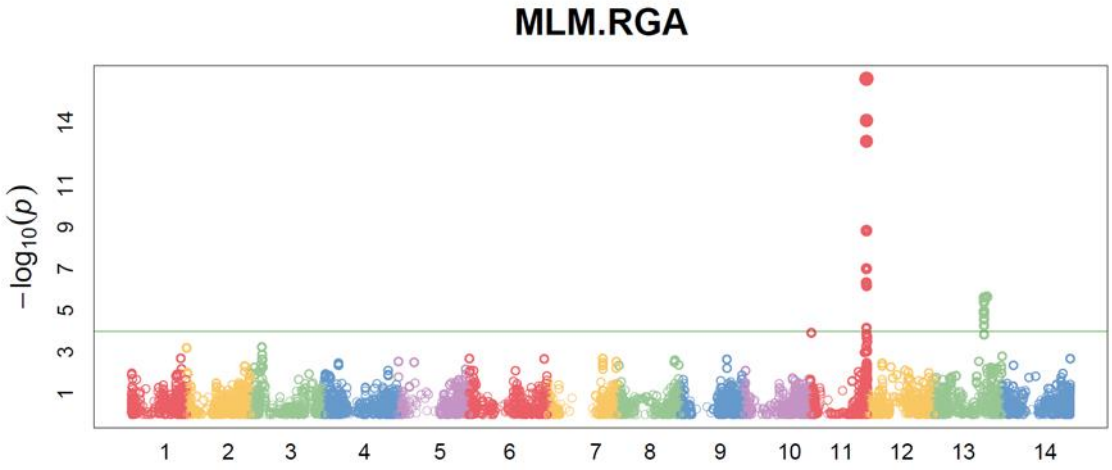


Figure 82 MLM Manhattan plot of GWAS for RGA

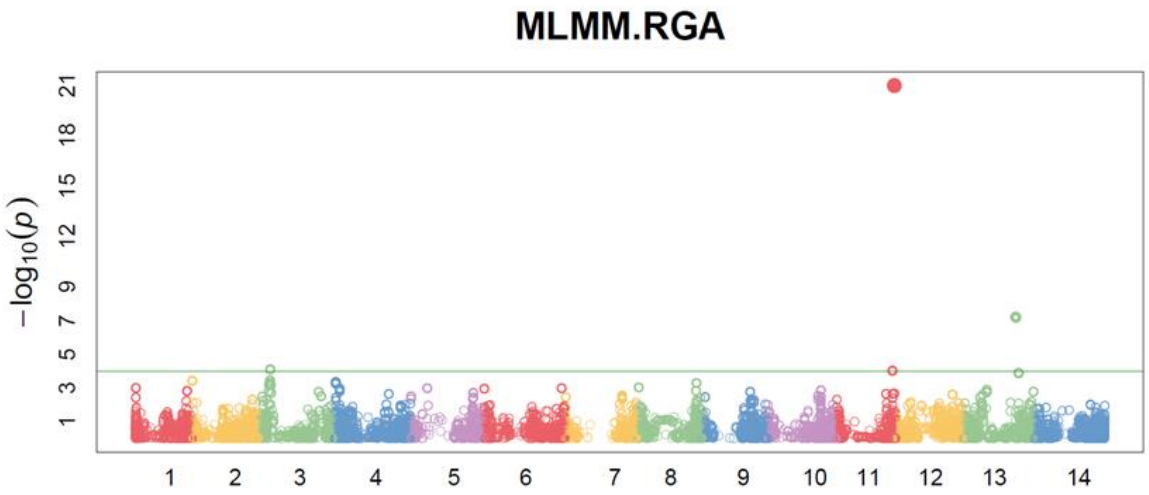


Figure 83 MLMM Manhattan plot of GWAS for RG

Following the GWAS analysis, a Manhattan plot as a output was obtained. The most associated markers were summarized in Table 26 using the most associated marker for every peak with the confidence interval, computed based on the LD decay (1.0 Mb) on each side of the tag SNP.

Table 29 Most associated SNP in GDP regarding RGA trait

| **SNP** | **Chromosome** | **Position** | **-log(P)** | **C.I (+/-LD decay)** | **Trait** |
|---|---|---|---|---|---|
| wsnp_CAP12_c948_496702 | 2A | 120,418,376. | 5.79 | 120418376 - 122727254 | RGA |
| Tdurum_contig42418_2618 | 6A | 600,464,983. | 11.52 | 599799143 - 601490086 | RGA |
| Kukri_c17556_411 | 7A | 193,844,745. | 5.05 | 193844745 - 194295012 | RGA |

| | | | | | |
|---|---|---|---|---|---|
| BS00015354_51 | 7A | 535,191,977. | 9.92 | 535078794 - 535969373 | RGA |

Four peaks were detected by the BLINK model, herein reported: on the chromosome 2A, on the 6A and two significant peaks on the 7A

Table 30 Candidate genes for RGA in GDP

| Gene stable ID | Gene start (bp) | Gene end (bp) | Gene description |
|---|---|---|---|
| **TRITD2Av1G053980** | 120416514 | 120418491 | 40S ribosomal protein SA |
| **TRITD6Av1G219530** | 600457694 | 600470784 | Myosin |
| **TRITD7Av1G197420** | 535187757 | 535193509 | Serine/threonine-protein phosphatase |

For main peaks detected in the GWAS analysis, candidate genes were studied based on the Triticum turgidum cv Svevo *RefSeq* v1.0.

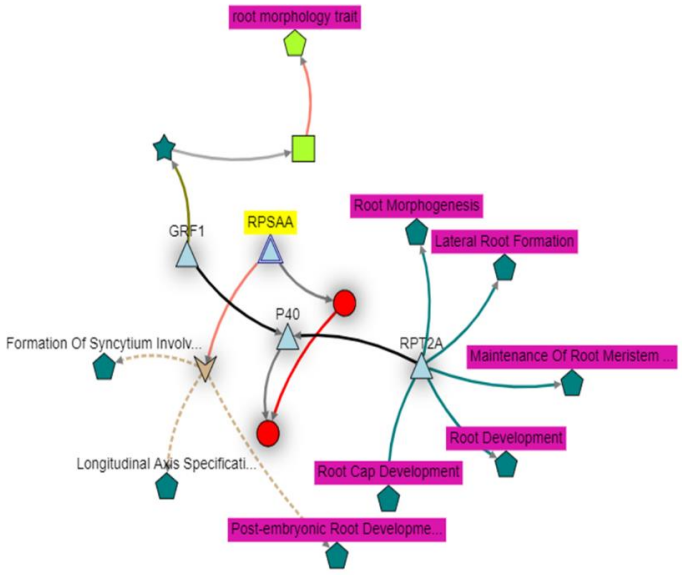


Figure 84 Knetminer network of RGA candidate gene function on chromosome 2A

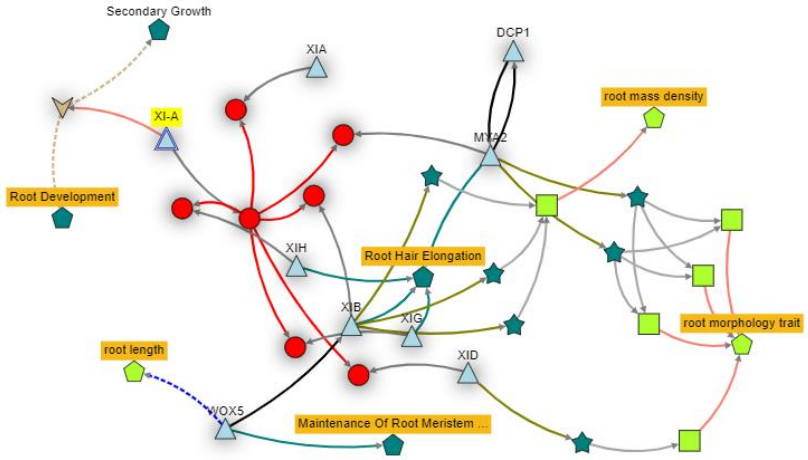


Figure 85 Knetminer network of RGA candidate gene function on chromosome 6A

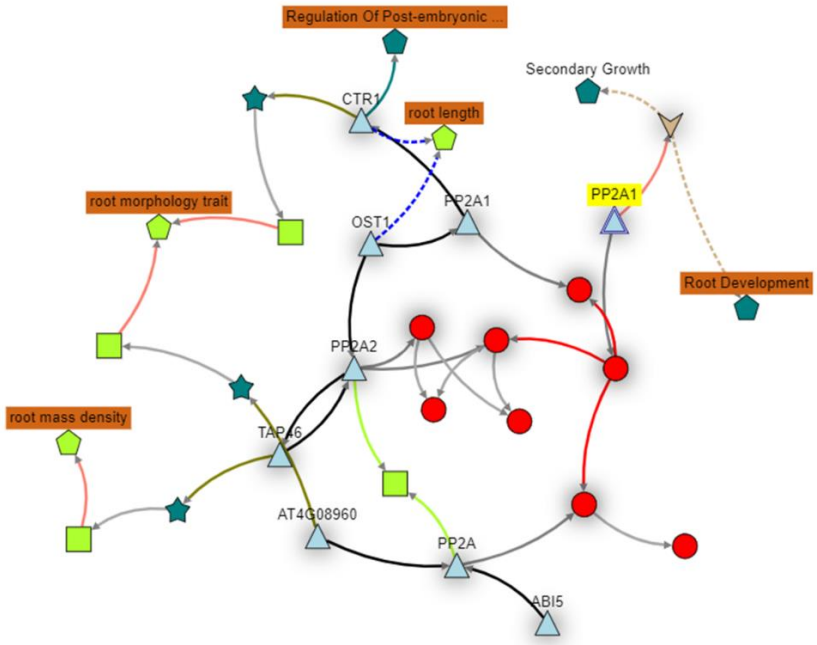


Figure 86 Knetminer network of RGA candidate gene function on chromosome 7A

## 3.4 DISCUSSION

From the peaks found in the GWAS, candidate genes analysis was carried out, then they were searched within the bibliography. Candidate gene found on the chromosome 7A based on Knetminer results, has been linked to root development, since its function is related to a PP2A which is involved in abiotic stress response (Pais et al 2009). This QTL has already been studied by Maccaferri et al 2016. In addition the function of the candidate gene on the chromosome 6A is linked to a myosin which has a role in root organogenesis (Abu-abied 2018). This peak is very close to the chromosome region of QRga.UniboDP-6A.2 (Maccaferri et al 2016) related to the root growth angle and QRl.SMxMC-6A which on the other hand is linked to root length (Iannucci et al 2017).

## 3.5 BIBLIOGRAPHY

Abu-Abied, Mohamad, et al. "Myosin XI-K is involved in root organogenesis, polar auxin transport, and cell division." Journal of experimental botany 69.12 (2018): 2869-2881.

Ennos, A. R., and A. H. Fitter. "Comparative functional morphology of the anchorage systems of annual dicots." Functional ecology (1992): 71-78.

Evert, Ray F. Esau's plant anatomy: meristems, cells, and tissues of the plant body: their structure, function, and development. John Wiley & Sons, 2006.

Hillnhütter, C., et al. "Nuclear magnetic resonance: a tool for imaging belowground damage caused by Heterodera schachtii and Rhizoctonia solani on sugar beet." Journal of Experimental Botany 63.1 (2012): 319-327.

Hobbs, Peter R., Ken Sayre, and Raj Gupta. "The role of conservation agriculture in sustainable agriculture." Philosophical Transactions of the Royal Society B: Biological Sciences 363.1491 (2008): 543-555.

Iannucci, Anna, et al. "Mapping QTL for root and shoot morphological traits in a durum wheat× T. dicoccum segregating population at seedling stage." International Journal of Genomics 2017 (2017).

Jobbagy, Esteban G., and Robert B. Jackson. "The distribution of soil nutrients with depth: global patterns and the imprint of plants." Biogeochemistry 53.1 (2001): 51-77.

Koevoets, Iko T., et al. "Roots withstanding their environment: exploiting root system architecture responses to abiotic stress to improve crop tolerance." Frontiers in plant science 7 (2016): 1335.

Lynch, Jonathan. "Root architecture and plant productivity." Plant physiology 109.1 (1995): 7.

Lynch, Jonathan P., and Kathleen M. Brown. "Topsoil foraging–an architectural adaptation of plants to low phosphorus availability." Plant and Soil 237.2 (2001): 225-237.

Lynch, Jonathan P., and Kathleen M. Brown. "Topsoil foraging–an architectural adaptation of plants to low phosphorus availability." Plant and Soil 237.2 (2001): 225-237.

Maccaferri, Marco, et al. "Prioritizing quantitative trait loci for root system architecture in tetraploid wheat." Journal of experimental botany 67.4 (2016): 1161-1178..

Manschadi, Ahmad M., et al. "The role of root architectural traits in adaptation of wheat to water-limited environments." Functional plant biology 33.9 (2006): 823-837.

Meister, Robert, et al. "Challenges of modifying root traits in crops for agriculture." Trends in plant science 19.12 (2014): 779-788.

País, Silvia Marina, María Teresa Téllez-Iñón, and Daniela Andrea Capiati. "Serine/threonine protein phosphatases type 2A and their roles in stress signaling." Plant signaling & behavior 4.11 (2009): 1013-1015.

Rosegrant, Mark W., Claudia Ringler, and Tingju Zhu. "Water for agriculture: maintaining food security under growing scarcity." Annual review of Environment and resources 34.1 (2009): 205-222.

Rossini, Laura, et al. "Genetics of whole plant morphology and architecture." The Barley Genome. Springer, Cham, 2018. 209-231.

Sahri, Ali, et al. "Towards a comprehensive characterization of durum wheat landraces in Moroccan traditional agrosystems: analysing genetic diversity in the light of geography, farmers' taxonomy and tetraploid wheat domestication history." BMC evolutionary biology 14.1 (2014): 1-18.

Siddique, K. H. M., R. K. Belford, and D1 Tennant. "Root: shoot ratios of old and modern, tall and semi-dwarf wheats in a Mediterranean environment." Plant and soil 121.1 (1990): 89-98.

Snyder, Clifford S., et al. "Review of greenhouse gas emissions from crop production systems and fertilizer management effects." Agriculture, Ecosystems & Environment 133.3-4 (2009): 247-266.

Villa, Tania Carolina Camacho, et al. "Defining and identifying crop landraces." Plant genetic resources 3.3 (2005): 373-384.

Williamson, Lisa C., et al. "Phosphate availability regulates root system architecture in Arabidopsis." Plant physiology 126.2 (2001): 875-882.

Wullschleger, Stan D., Lewis H. Ziska, and James A. Bunce. "Respiratory responses of higher plants to atmospheric CO2 enrichment." Physiologia Plantarum 90.1 (1994): 221-229.

Zeven, A.C. (1998). Landraces: A review of definitions and classifications. Euphytica, 104, 127-139.

Zobel, Richard W., Peter Del Tredici, and John G. Torrey. "Method for growing plants aeroponically." Plant Physiology 57.3 (1976): 344-346.

# 4 CONCLUSIONS AND PERSPECTIVES

In this work two large collection were characterised for different traits related to yield and its components. These two collections showed a remarkable diversity for the yield related traits, meaning that they could be a valid source of novel allelic variants. The Global Durum Panel has been also characterised for the root growth angle, a root system architecture component, which is directly linked to abiotic stress resilience such as water or nutritive elements limited conditions. Nowadays yield but also its stability has become a major point to focus on in a global context of climate change. Therefore it becomes very important not only to enhance yield itself, but also to improve the ability of the crop to withstand the biotic and abiotic stresses caused by the new restricting environmental conditions mentioned before. Hence the exploration of novel allelic diversity is the main goal to achieve new genes discovery that can allow new breeding strategies and programmes. To accomplish this task, the characterisation of comprehensive germplasm sources involving all the tetraploid wheat subgroups is probably one of the main tools for new useful genes discovery. In this work marker-trait associations were detected in two comprehensive tetraploid collections and the results found stability in putative QTLs already reported in literature, However part of these results are also highlighted as novel loci of interest. Stable MTAs identified in two or more analysis could be considered for further validation and characterisation. To conclude, the Global Durum resources represented by the Global Durum Panel and the Tetraploid Global Collection need to be investigated and characterised more deeply for gene of interest in order to be included in marker-assisted selection programs for the development of yield-enhanced wheat varieties.

# 5 SUPPLEMENTARY MATERIALS

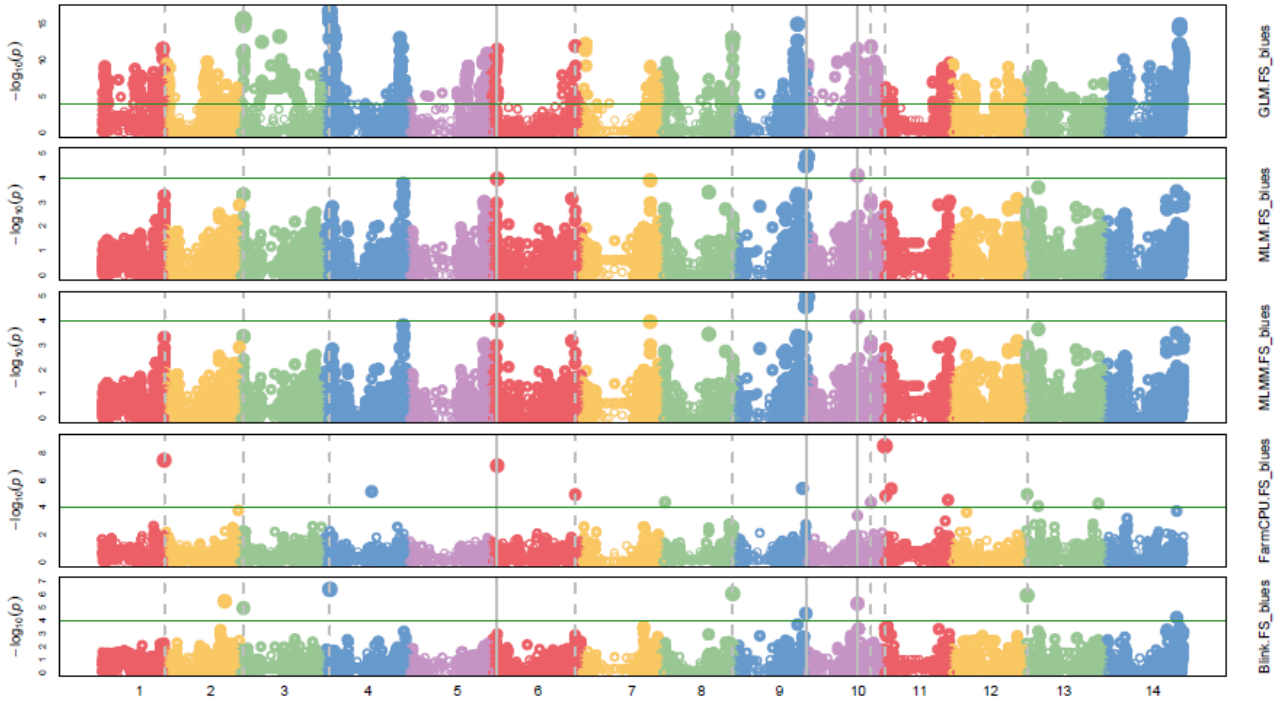


Figure 87 Manhattan plots for FS trait in GDP 2020

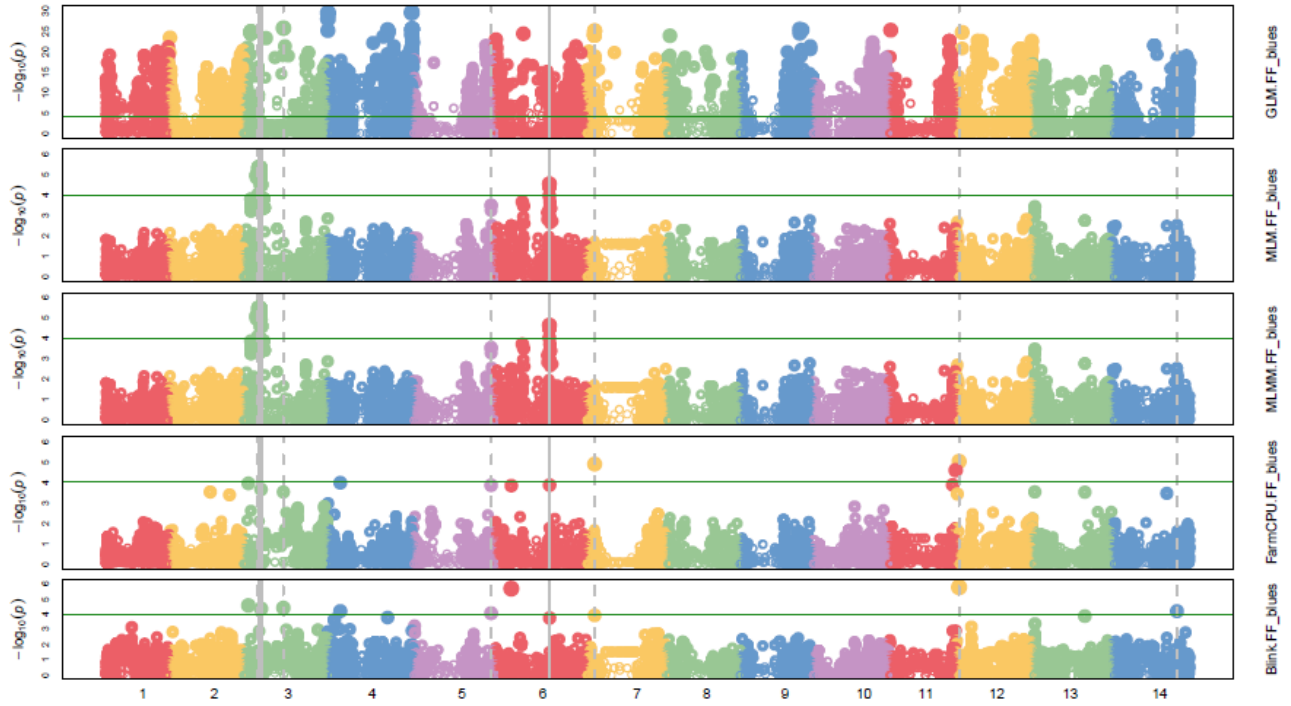


Figure 88 Manhattan plots for FF trait in GDP 2020

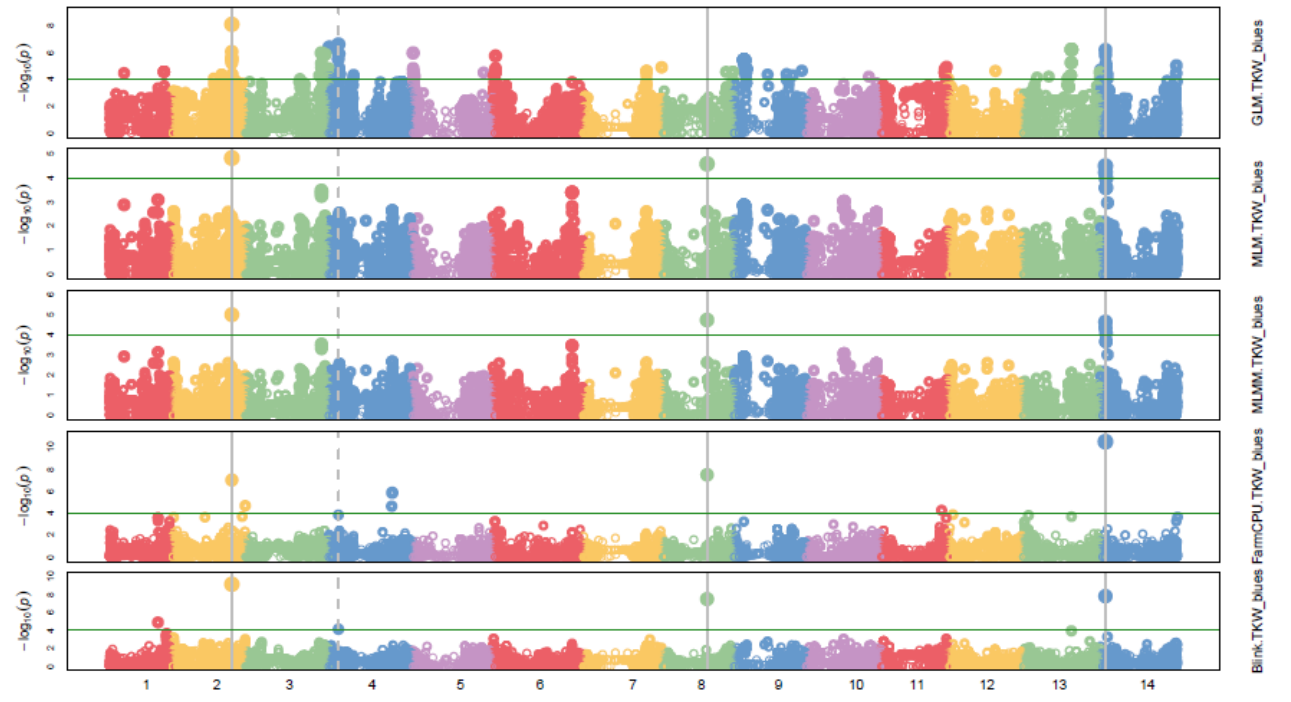


Figure 89 Manhattan plots for TKW in GDP 2020

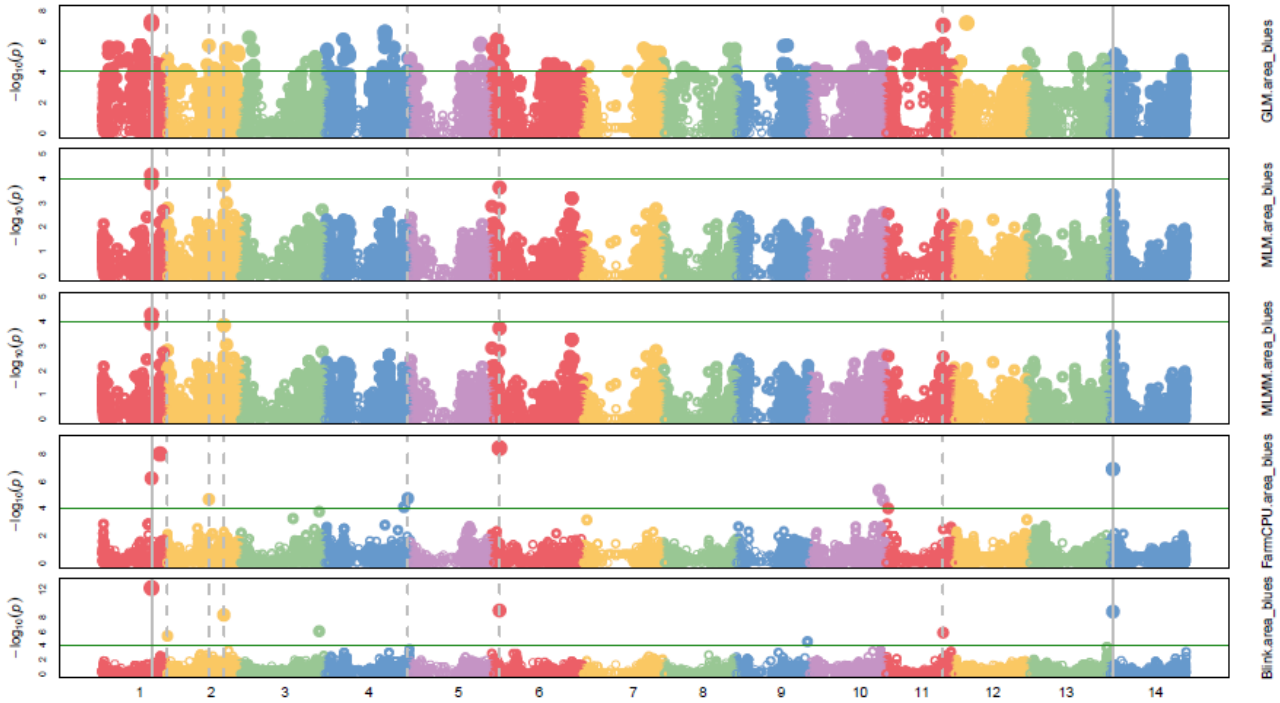


Figure 90 Manhattan plots for Grain area in GDP 2020

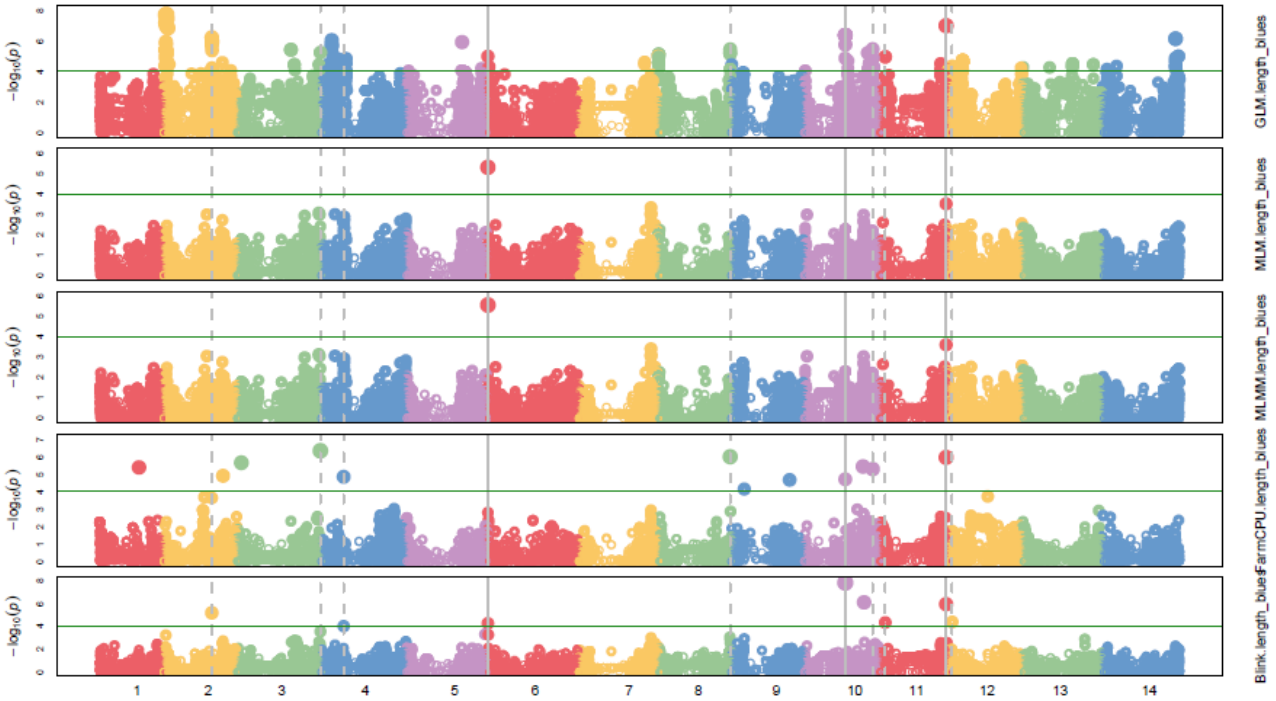


Figure 91 Manhattan plots for Grain length in GDP 2020

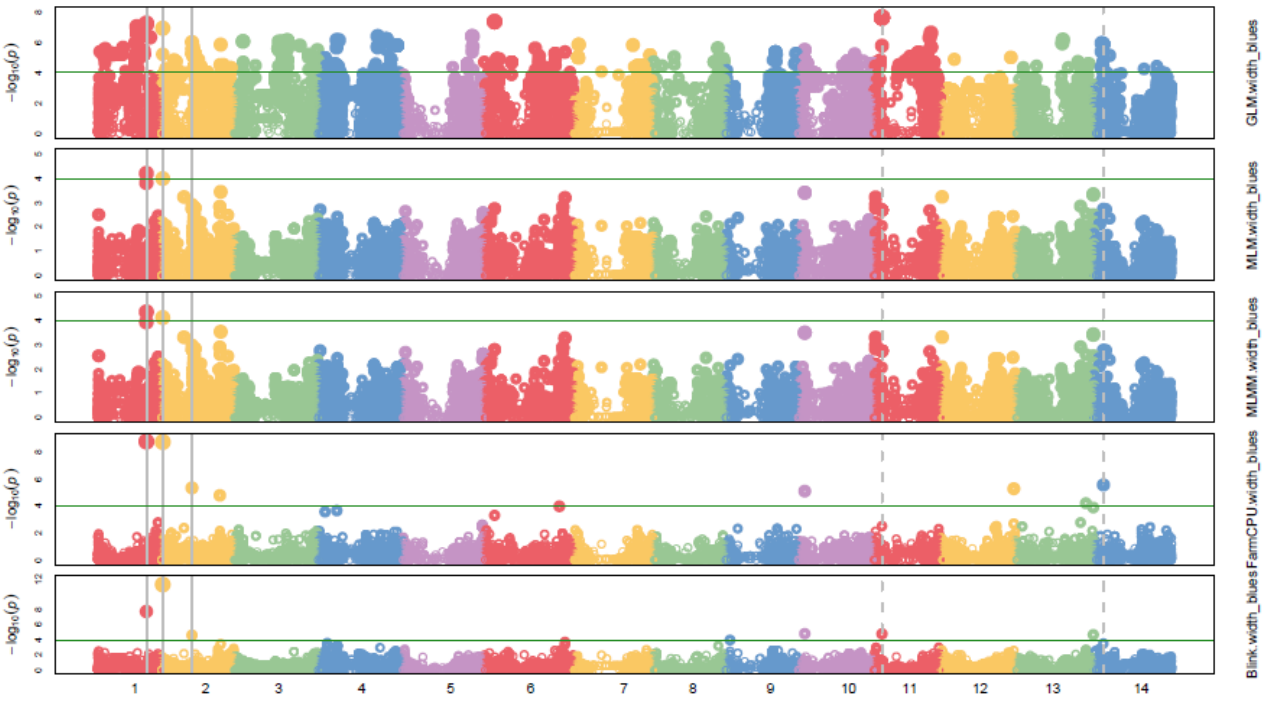


Figure 92 Manhattan plots for Grain width in GDP 2020

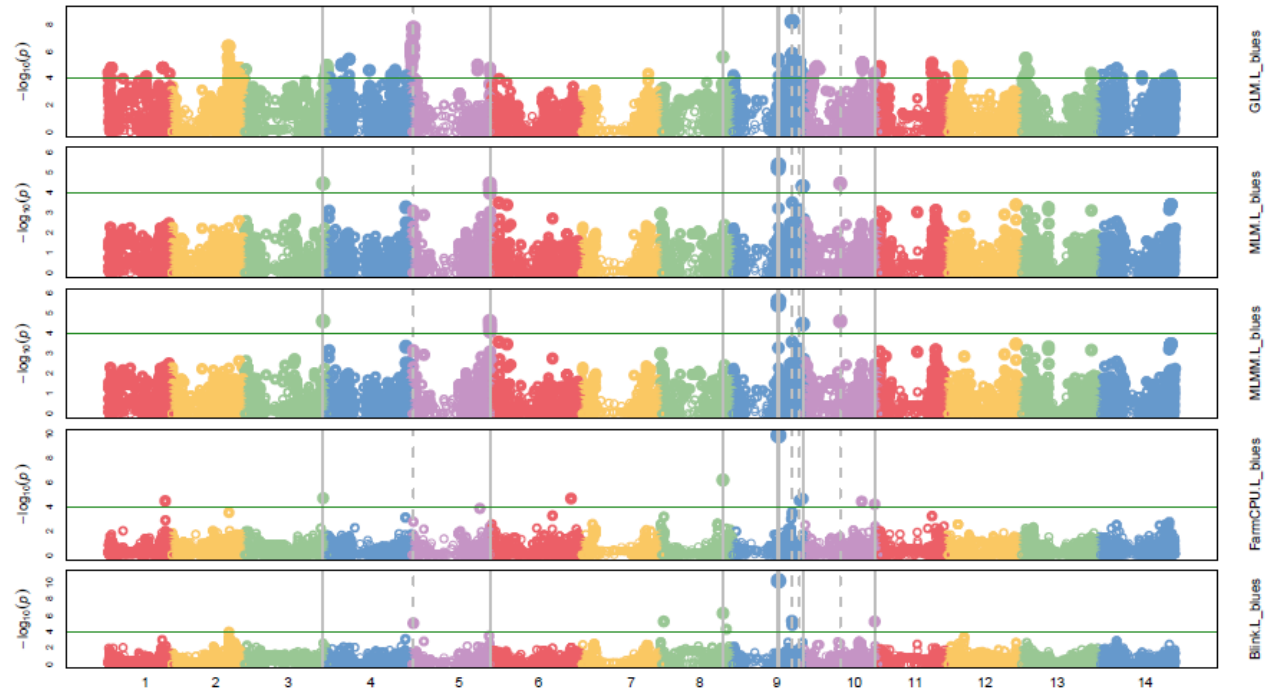


Figure 93 Manhattan plots for Grain brightness in GDP 2020

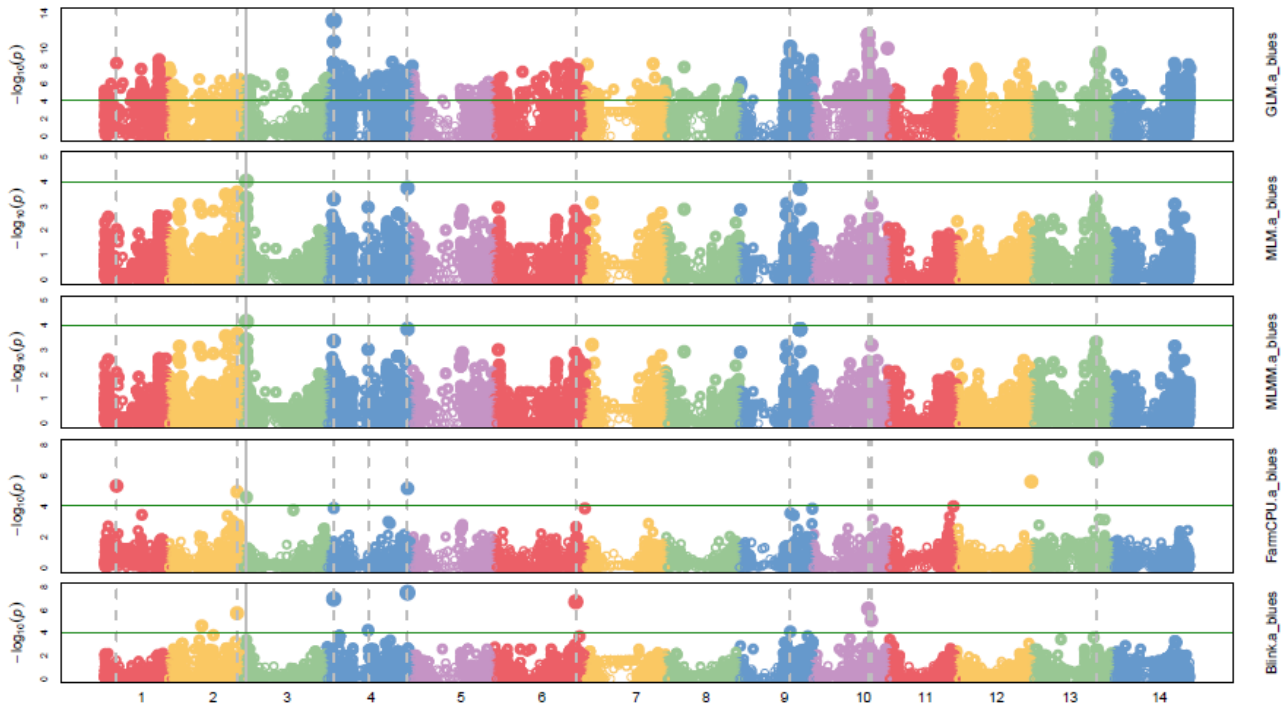


Figure 94 Manhattan plots for Grain redness in GDP 2020

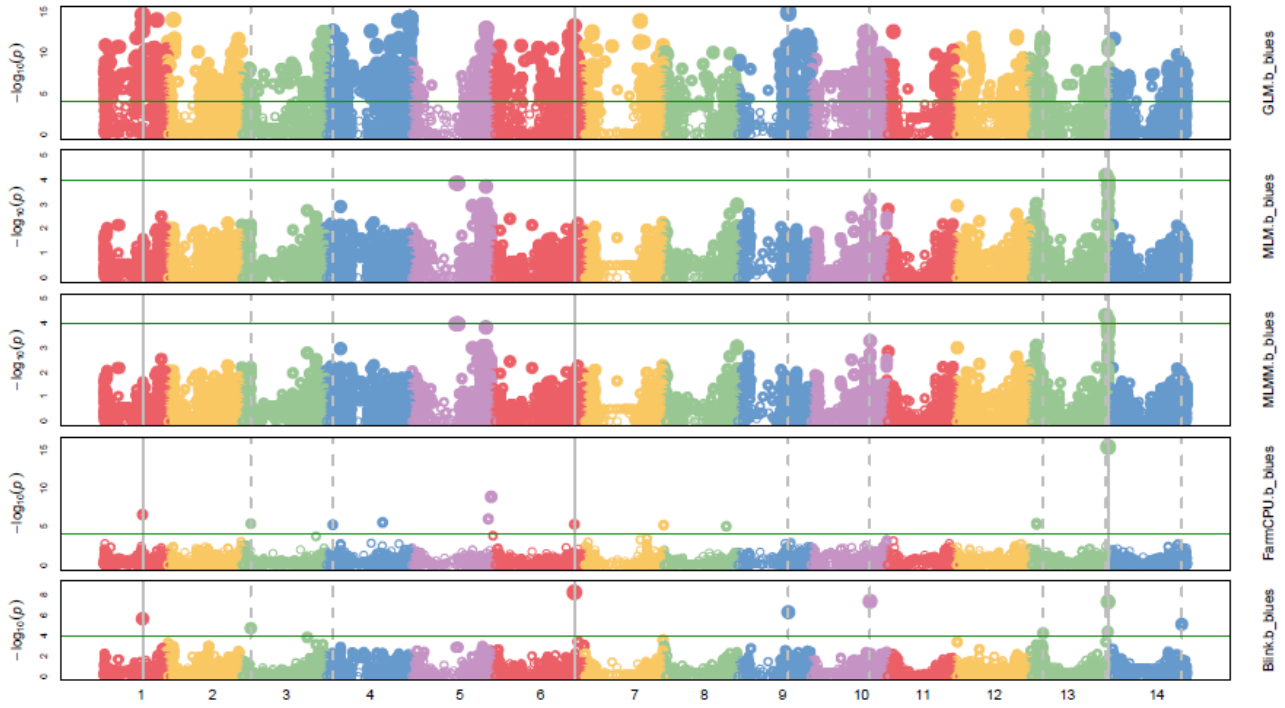


Figure 95 Manhattan plots for Grain yellowness in GDP 2020

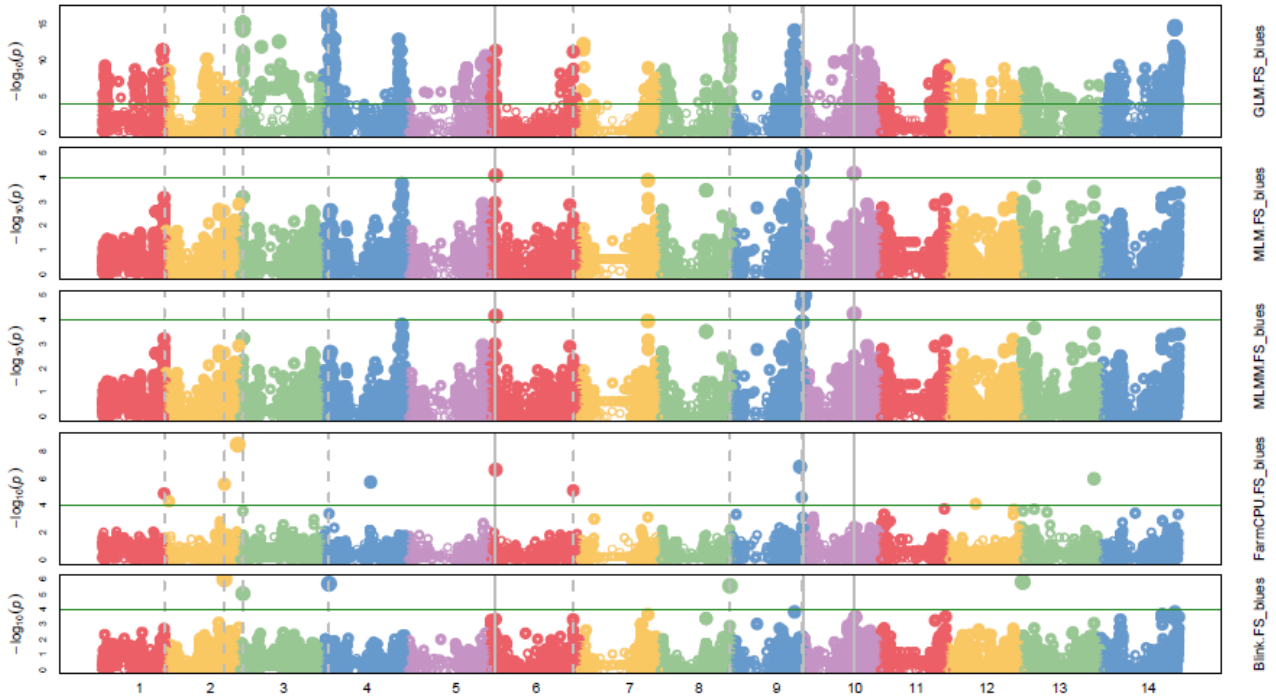


Figure 96 Manhattan plots for FS in GDP 2021

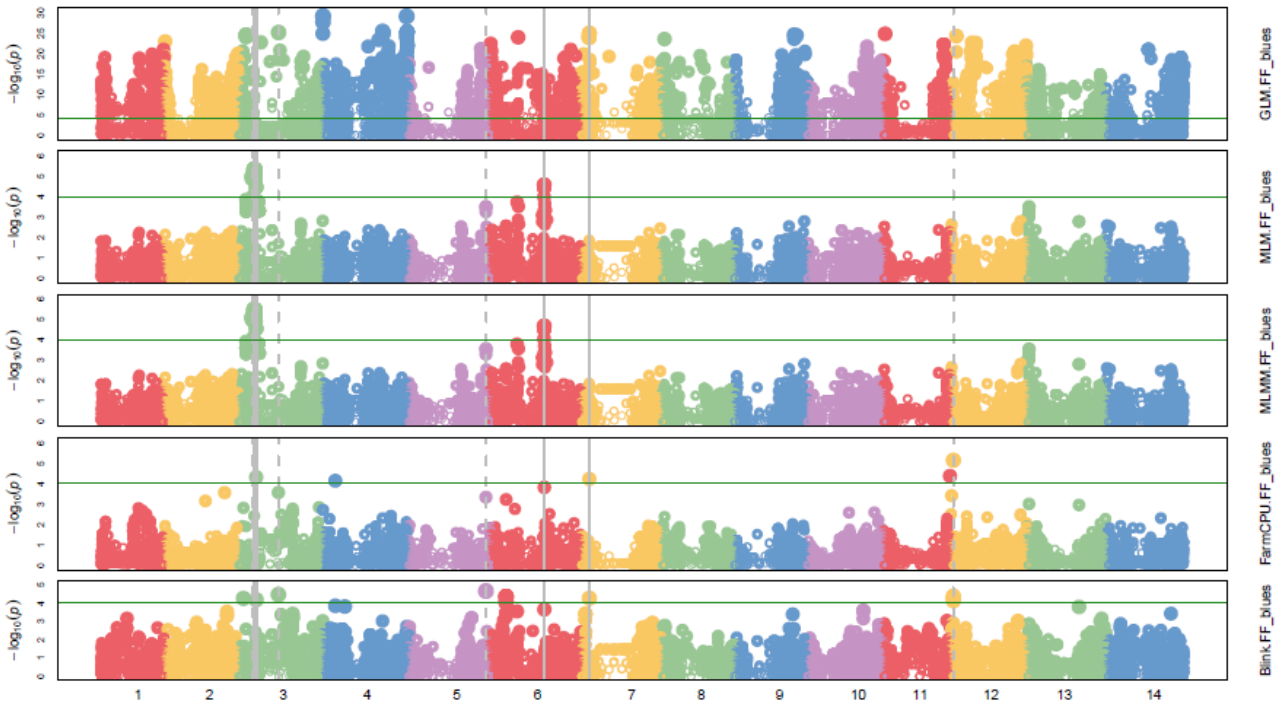


Figure 97 Manhattan plots for FF in GDP 2021

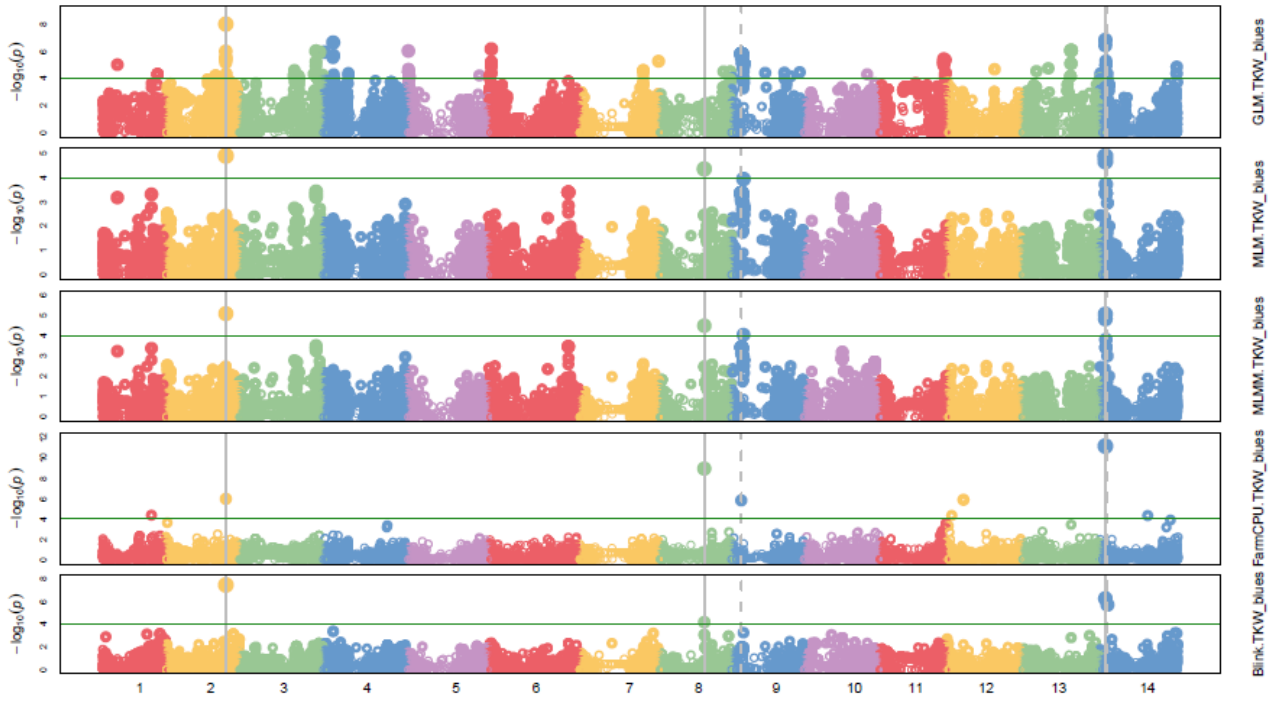


Figure 98 Manhattan plots for TKW in GDP 2021

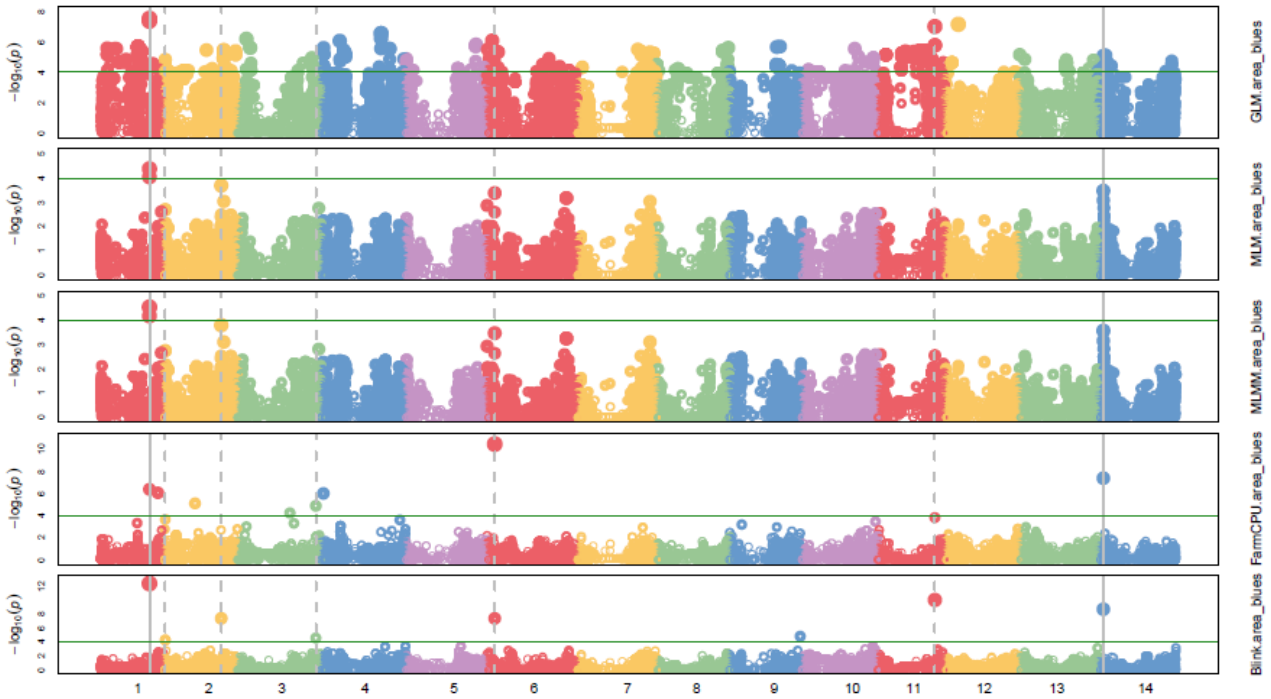


Figure 99 Manhattan plots for Grain area in GDP 2021

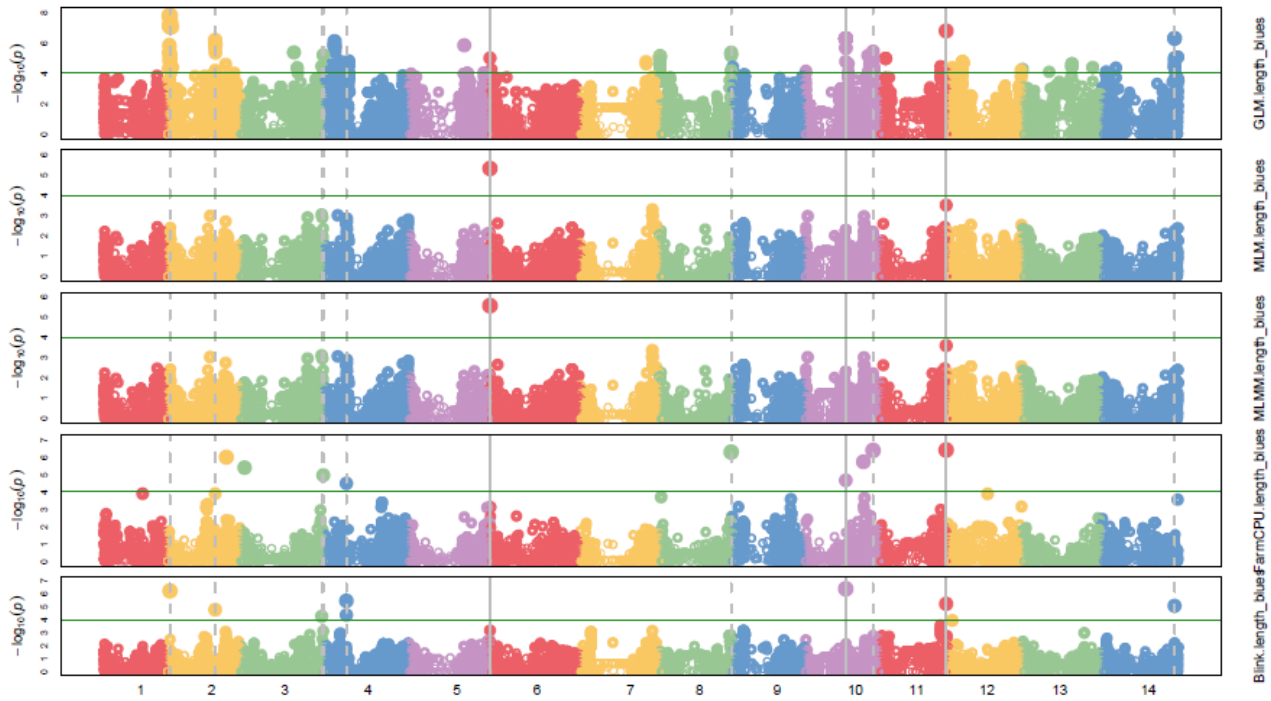


Figure 100 Manhattan plots for Grain length in GDP 2021

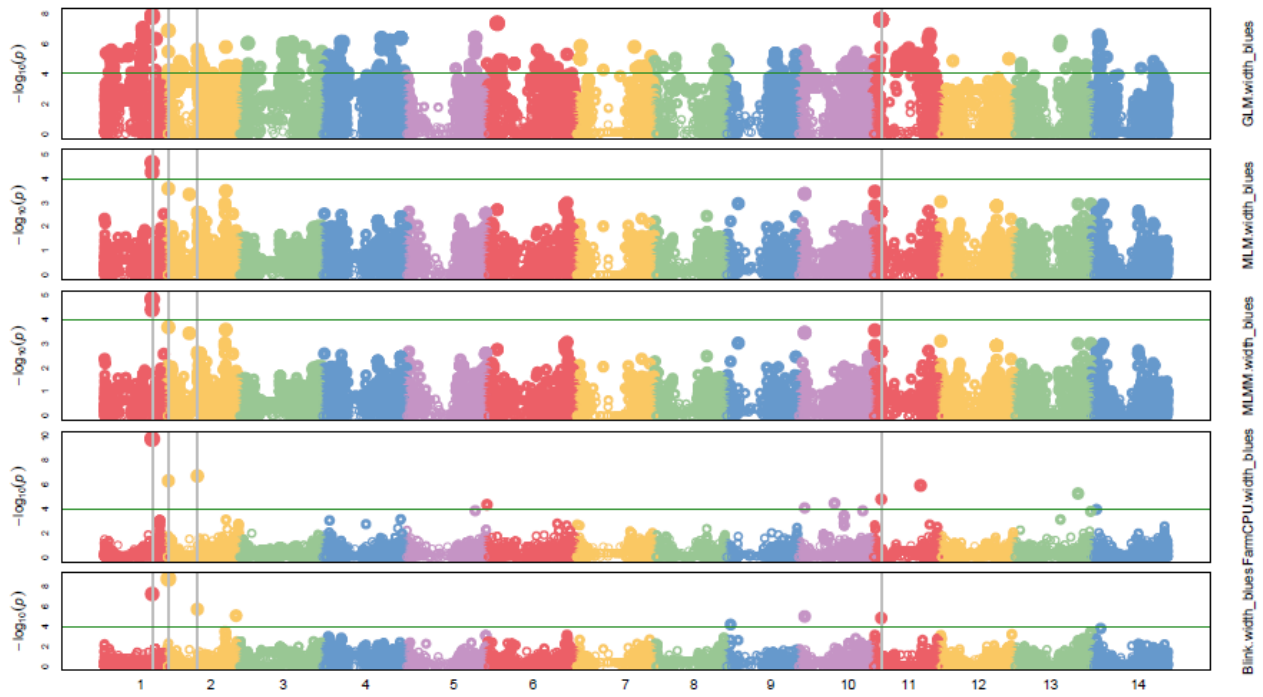


Figure 101 Manhattan plots for Grain width in GDP 2021

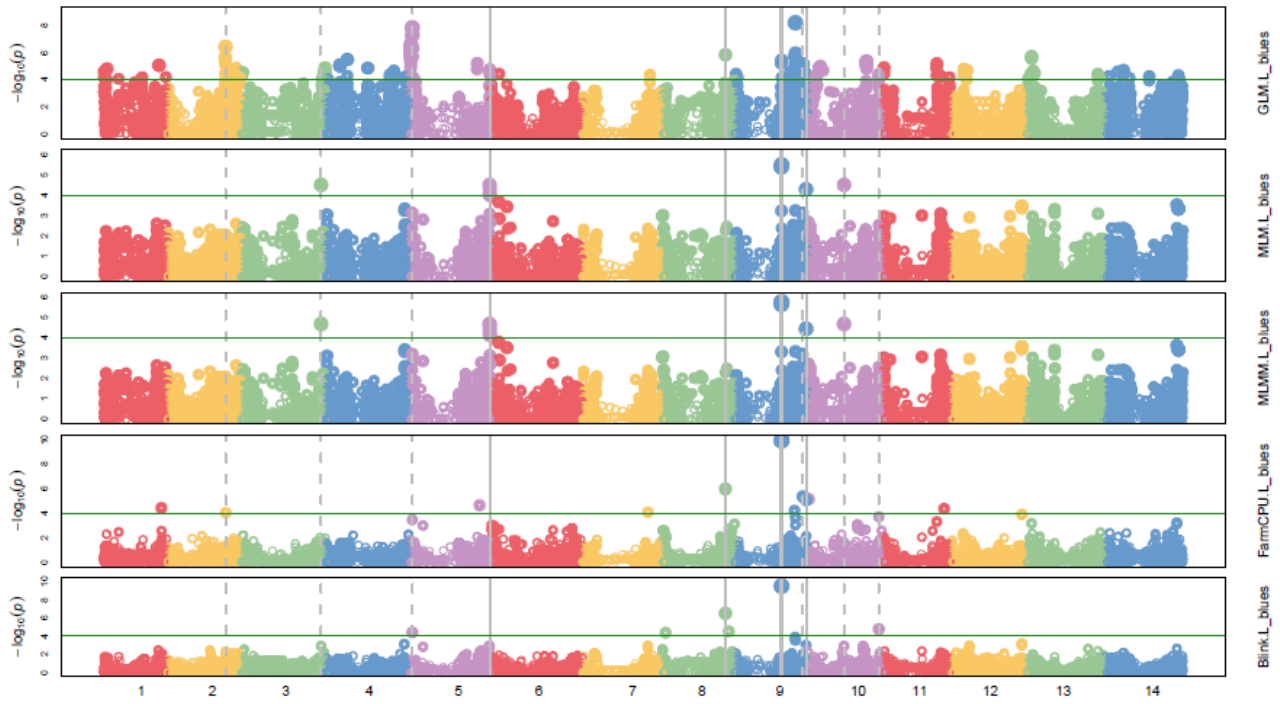


Figure 102 Manhattan plots for Grain brightness in GDP 2021

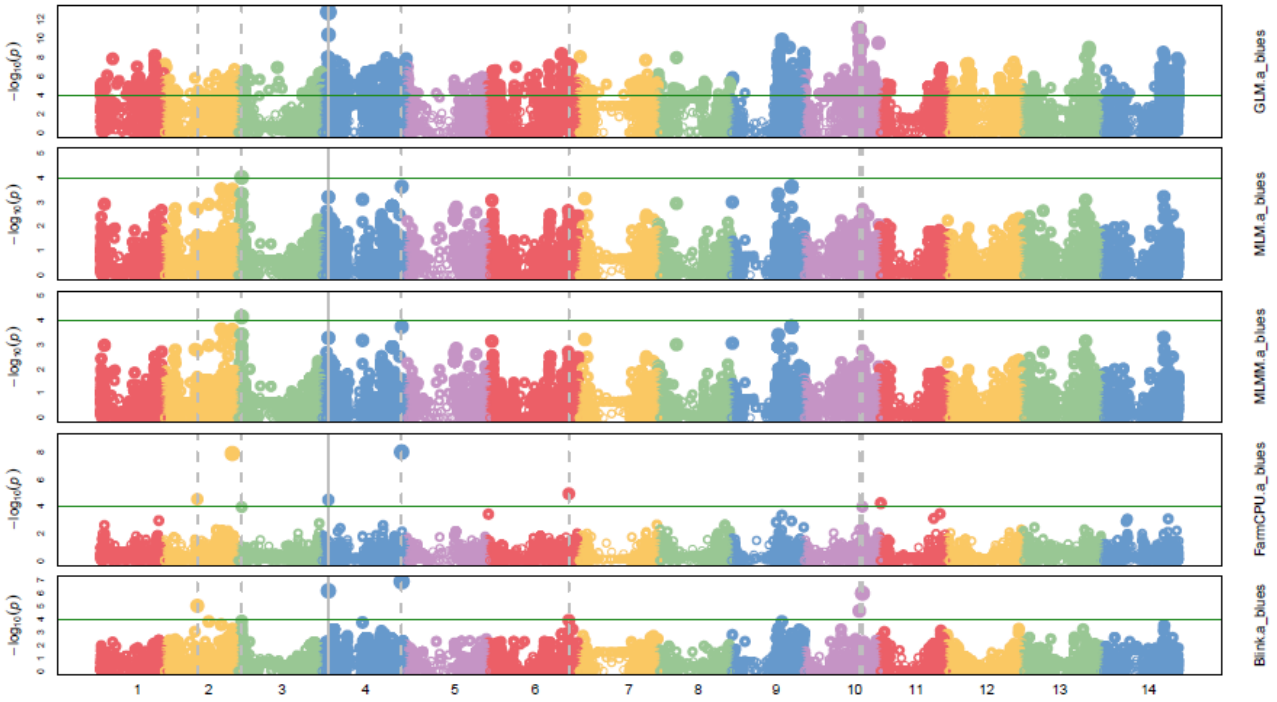


Figure 103 Manhattan plots for Grain redness in GDP 2021

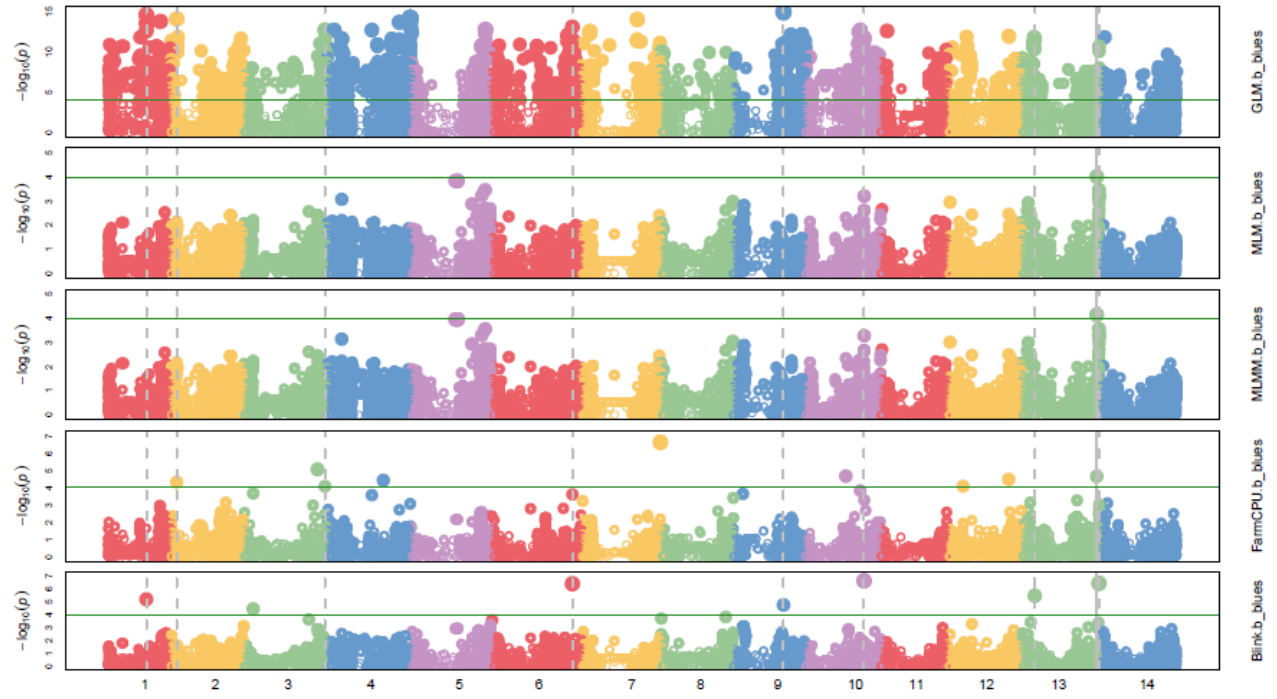


Figure 104 Manhattan plot for Grain yellowness in GDP 2021

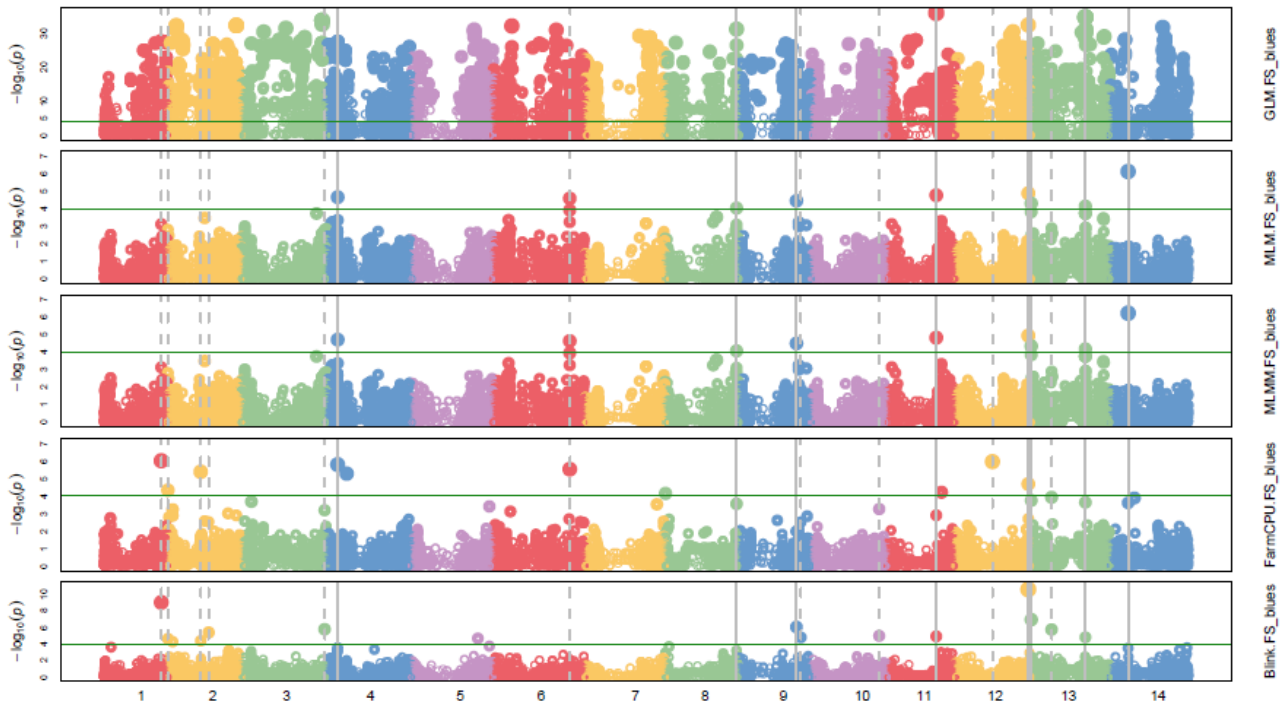


Figure 105 Manhattan plots for FS in TGC 2019

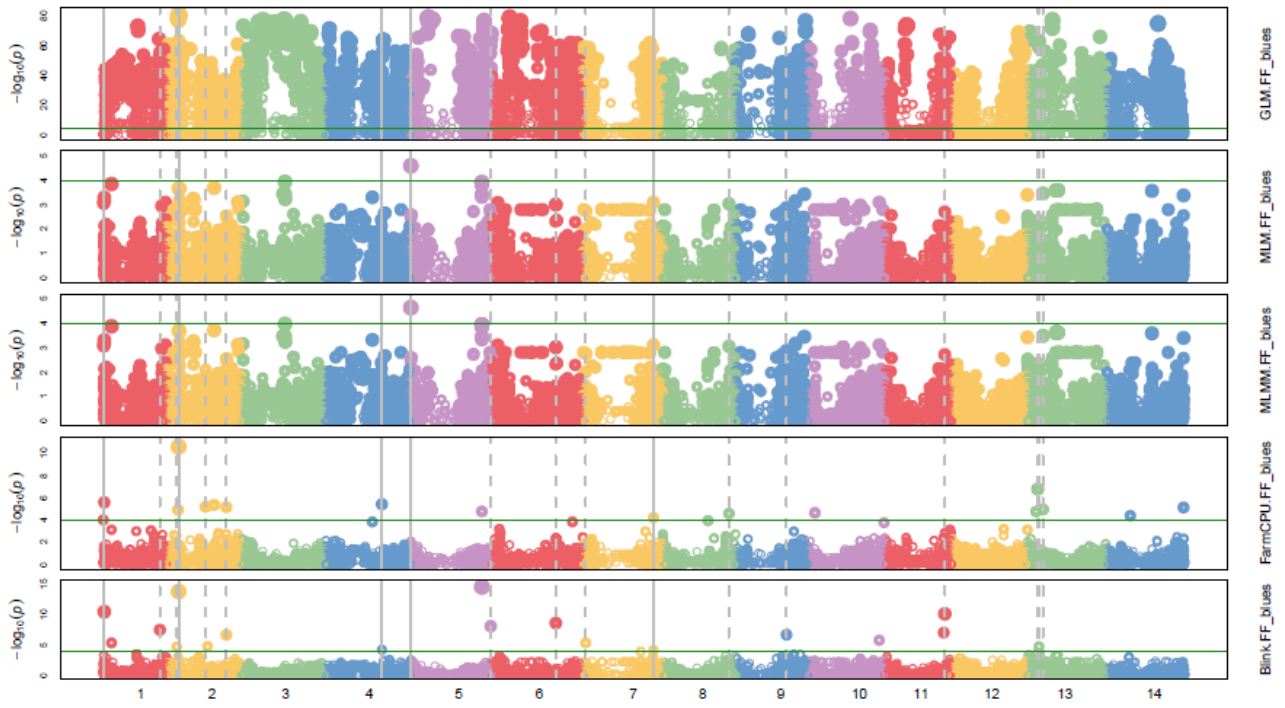


Figure 106 Manhattan plots for FF in TGC 2019

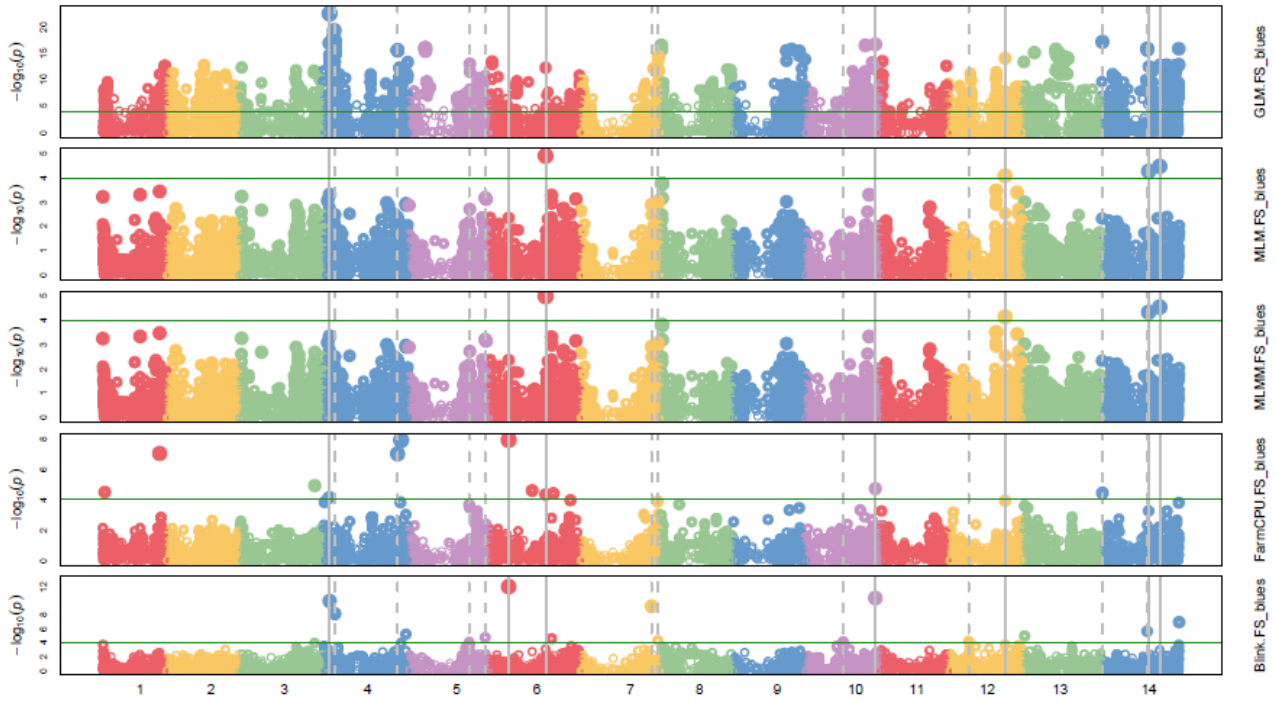


Figure 107 Manhattan plots for FS in TGC 2020

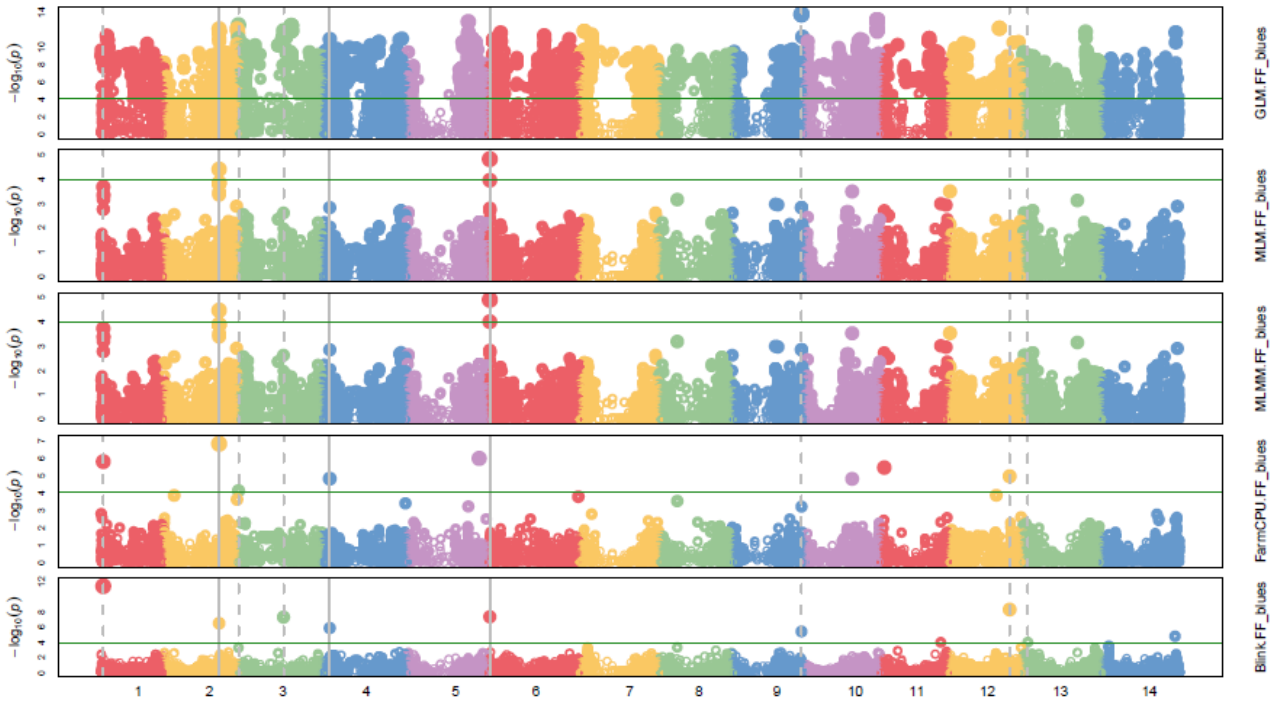


Figure 108 Manhattan plots for FF in TGC 2020

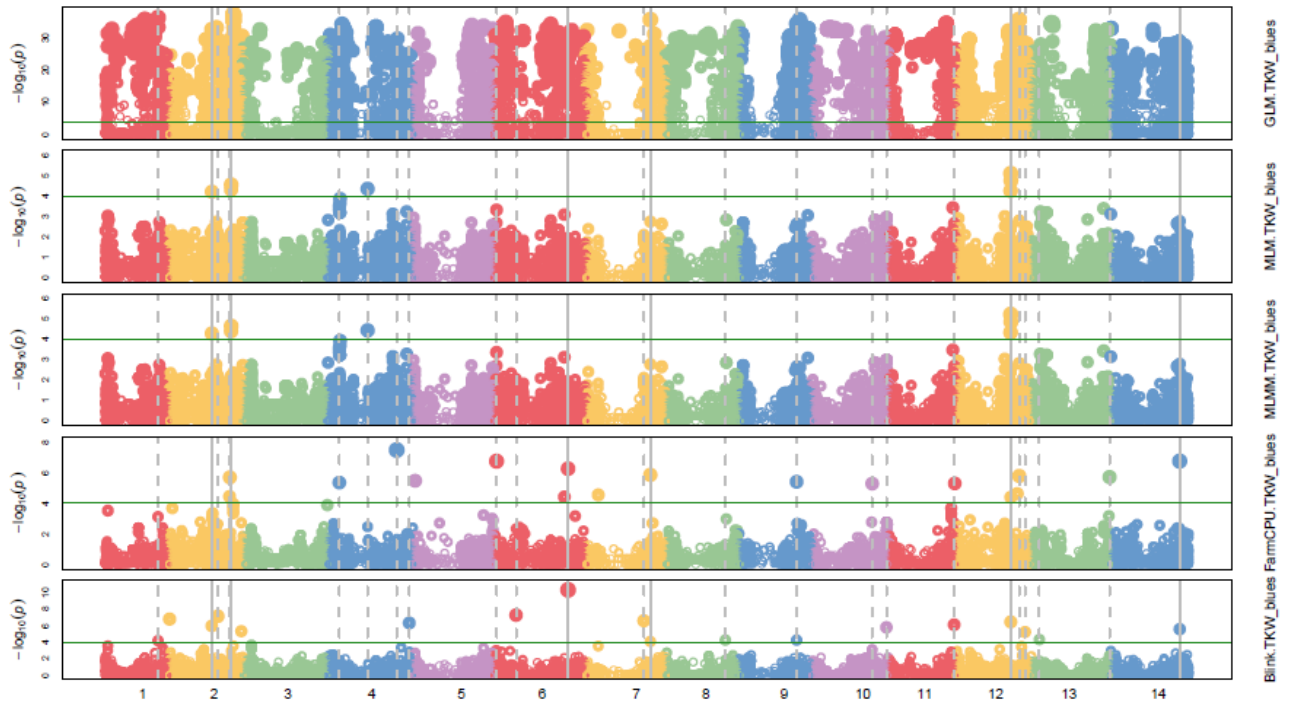


Figure 109 Manhattan plost for TKW in TGC 2020

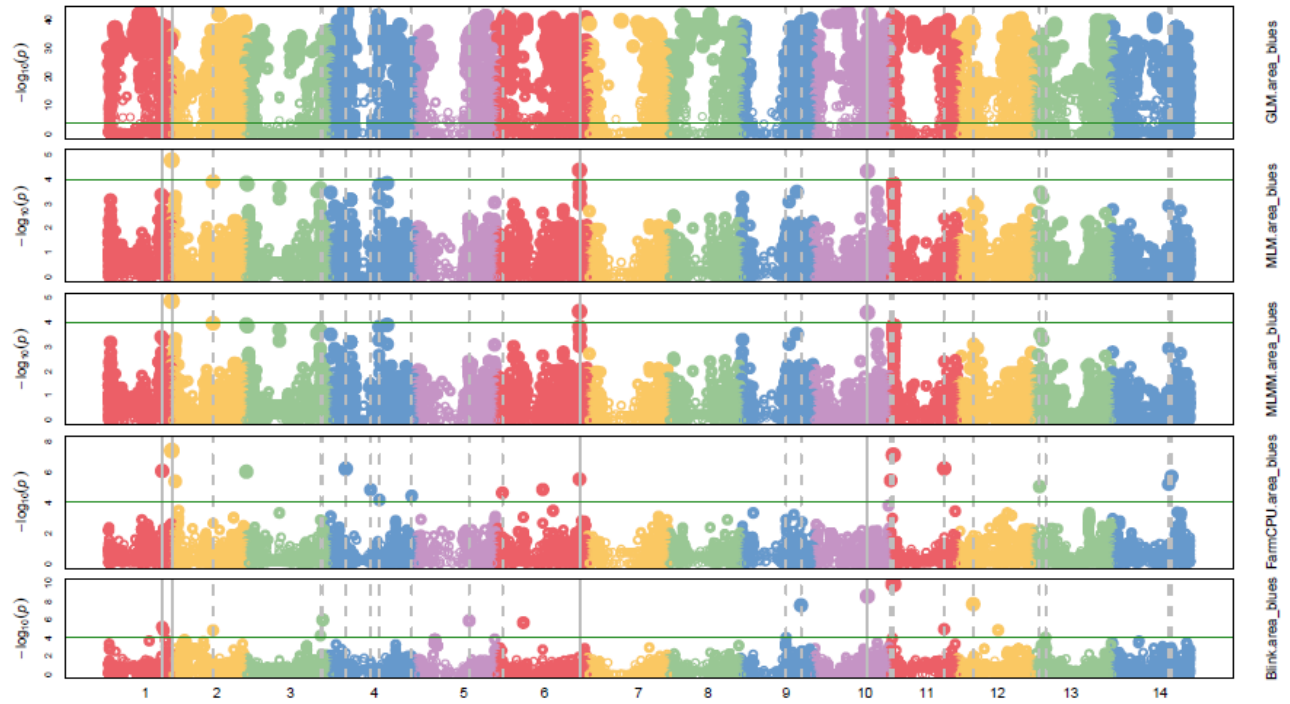


Figure 110 Manhattan plots for Grain area in TGC 2020

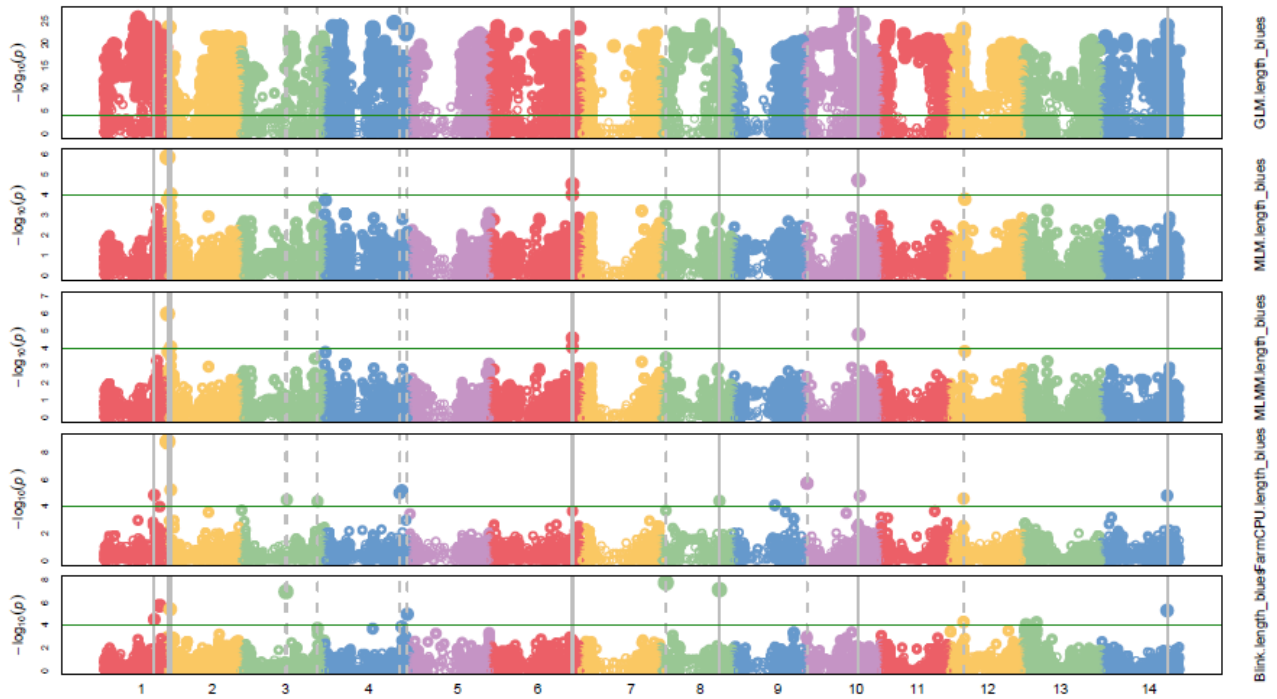


Figure 111 Manhattan plots for Grain length in TGC 2020

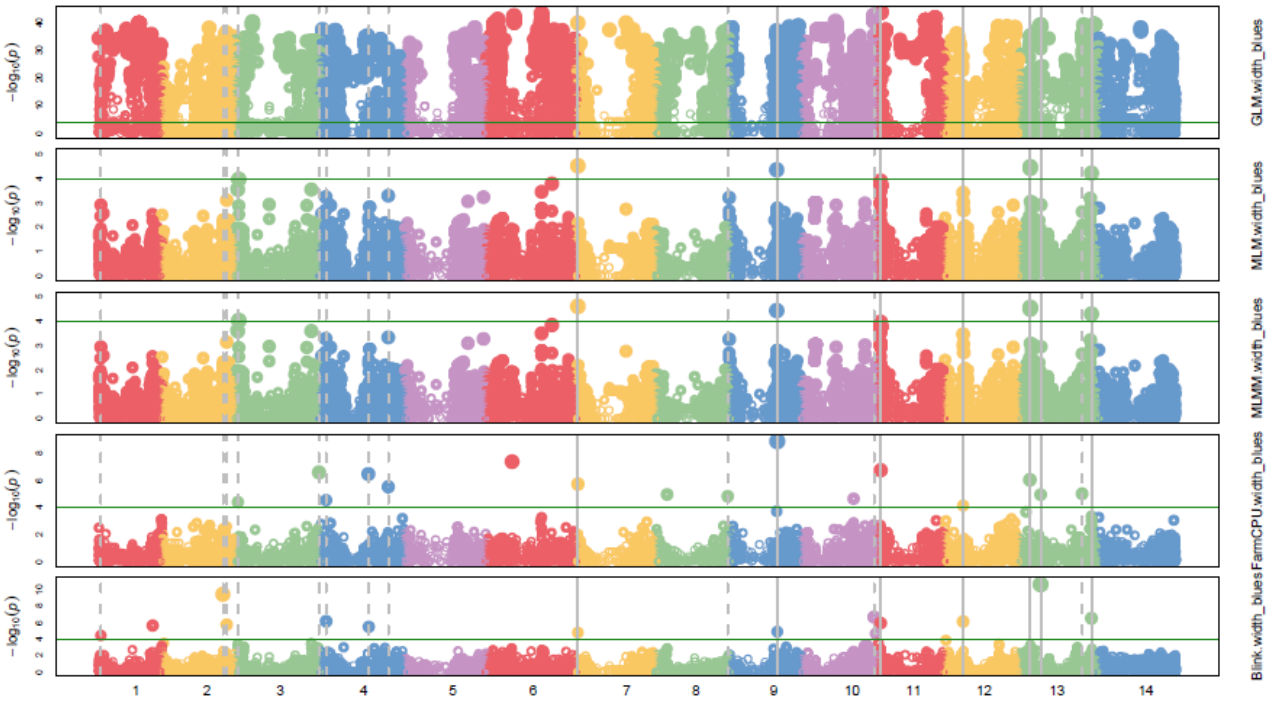


Figure 112 Manhattan plots for Grain width in TGC 2020

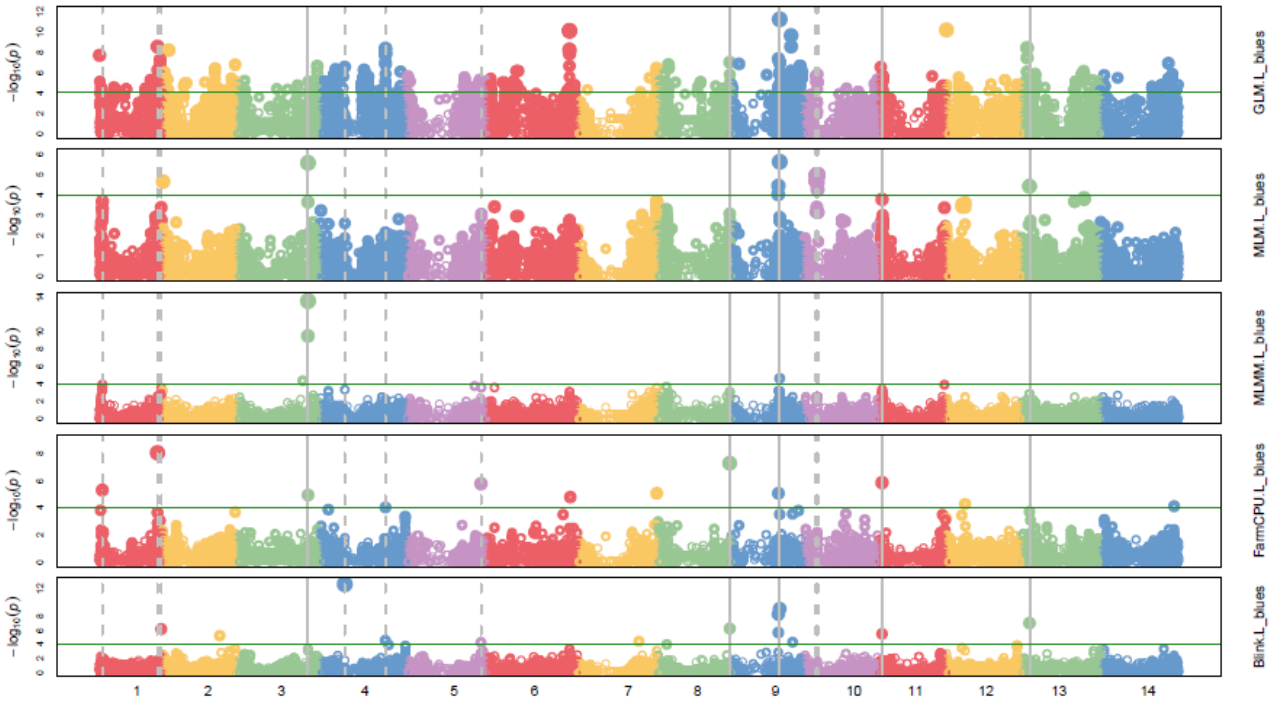


Figure 113 Manhattan plots for Grain brightness in TGC 2020

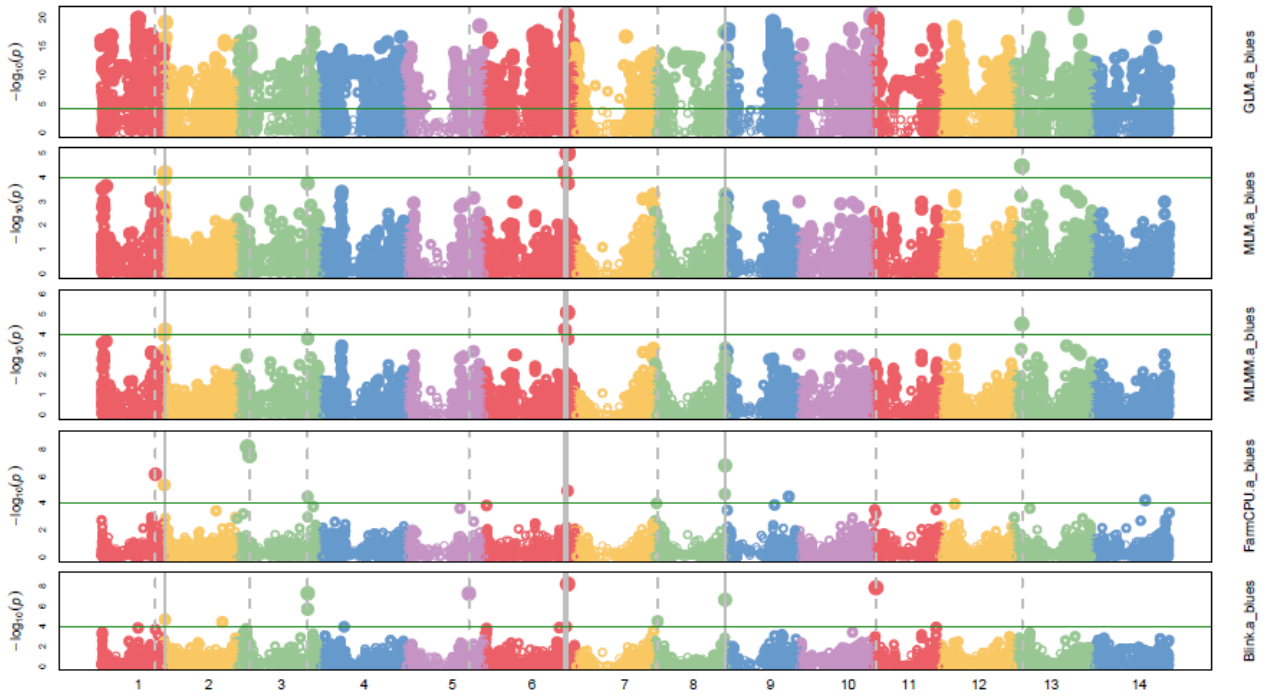


Figure 114 Manhattan plots for Grain redness in TGC 2020

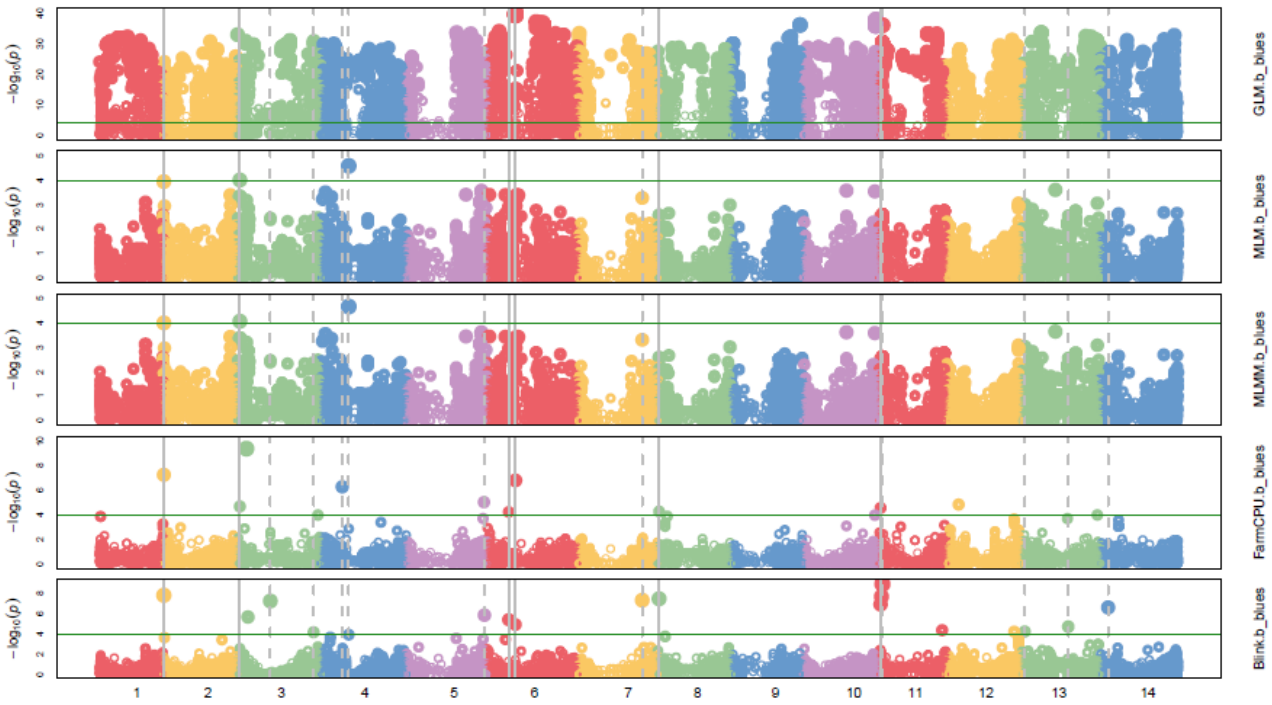


Figure 115 Manhattan plots for Grain yellowness in TGC 2020

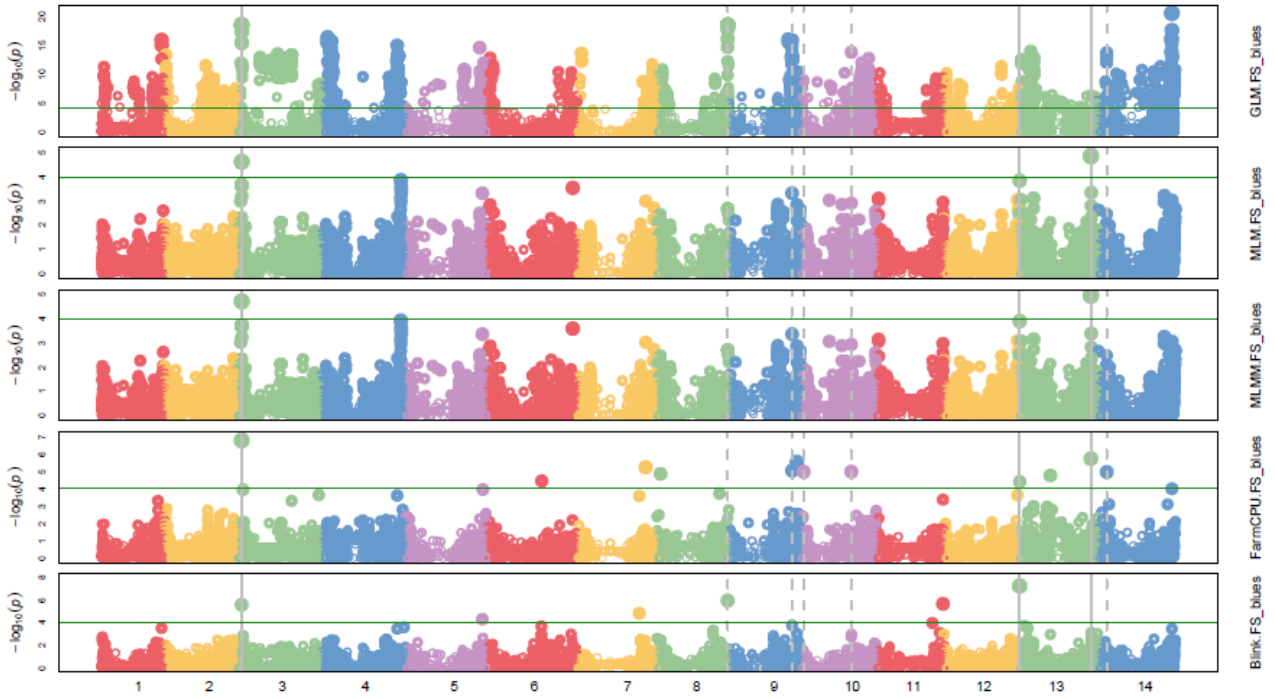


Figure 116 Manhattan plots for FS in GDP cluster

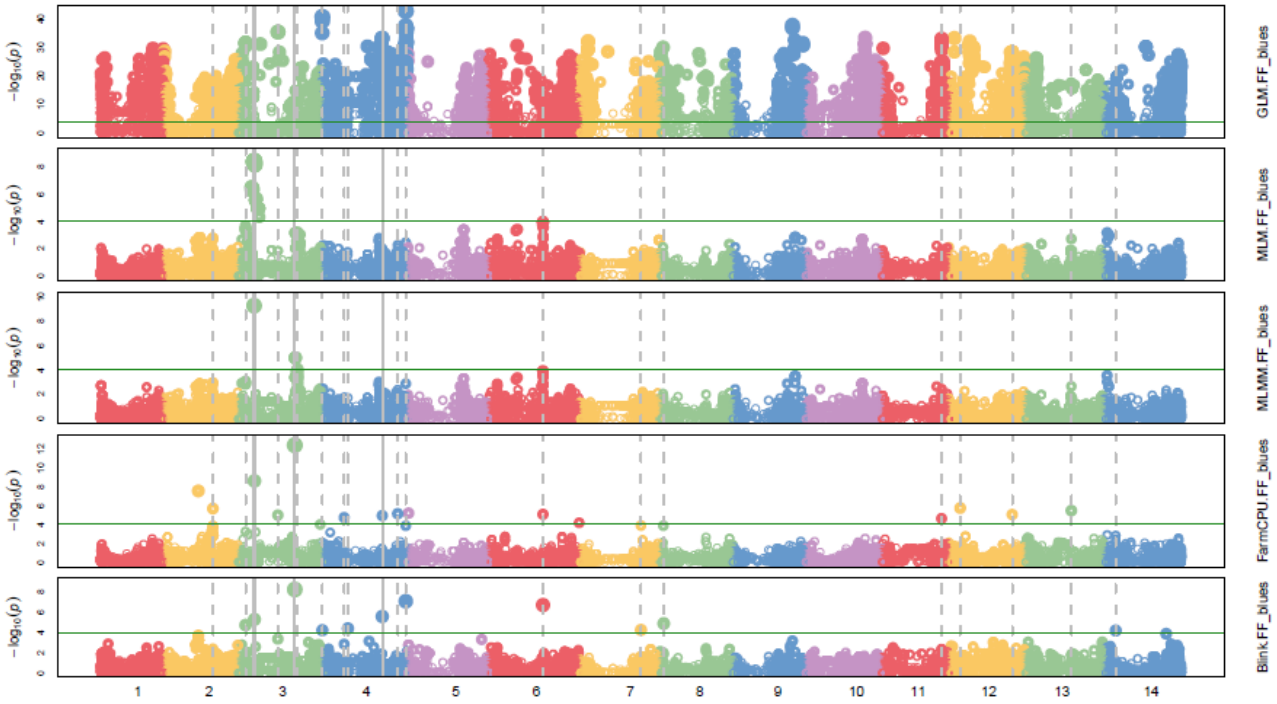


Figure 117 Manhattan plots for FF in GDP cluster

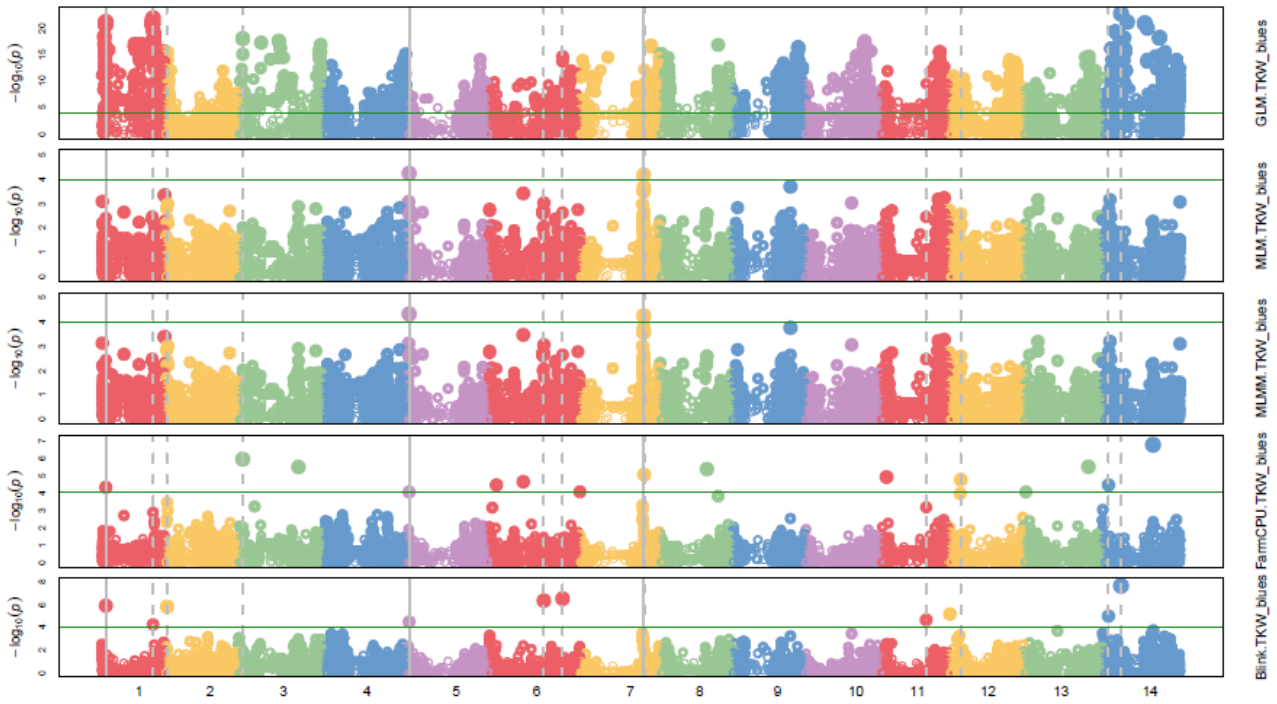


Figure 118 Manhattan plot for TKW in GDP cluster

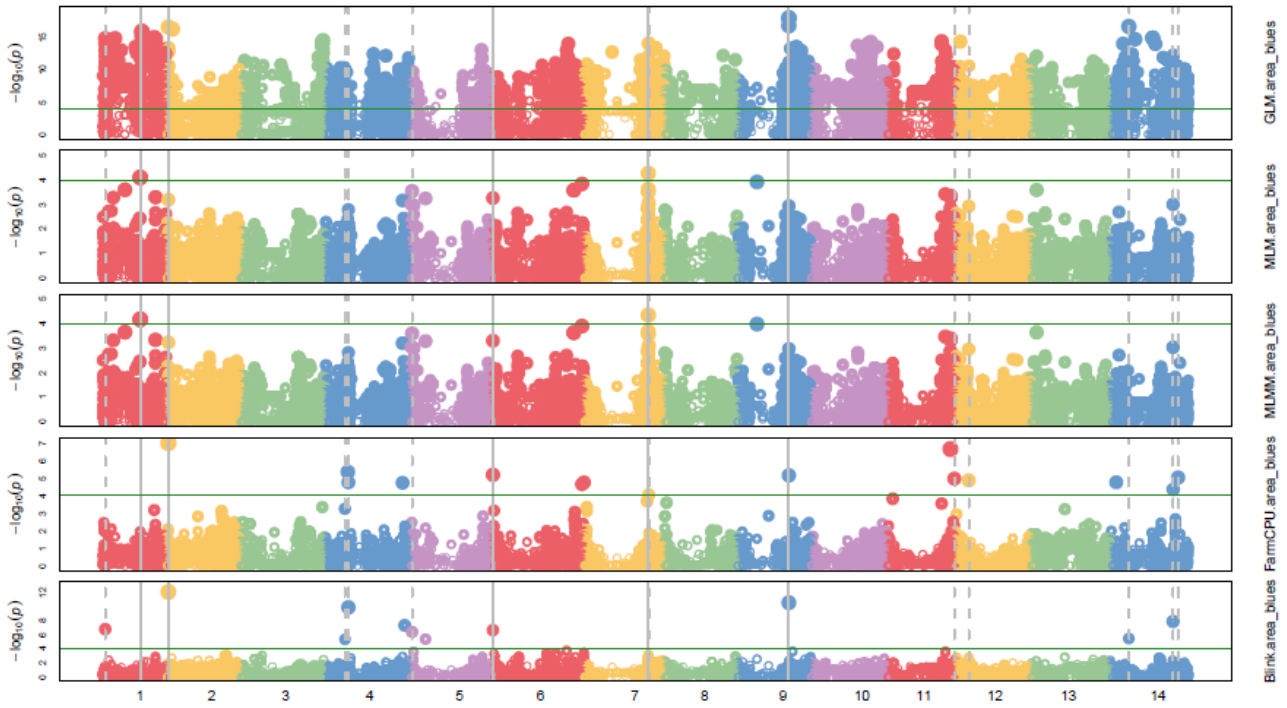


Figure 119 Manhattan plots for Grain area in GDP cluster

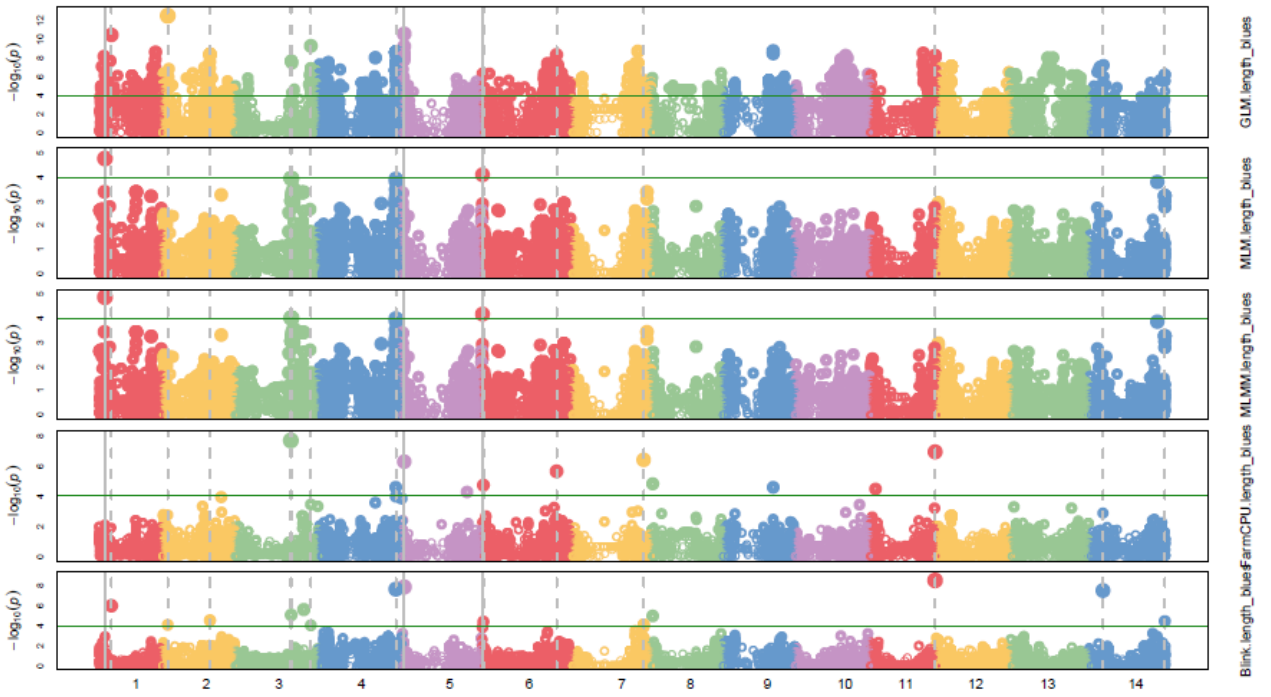


Figure 120 Manhattan plots for Grain length in GDP cluster

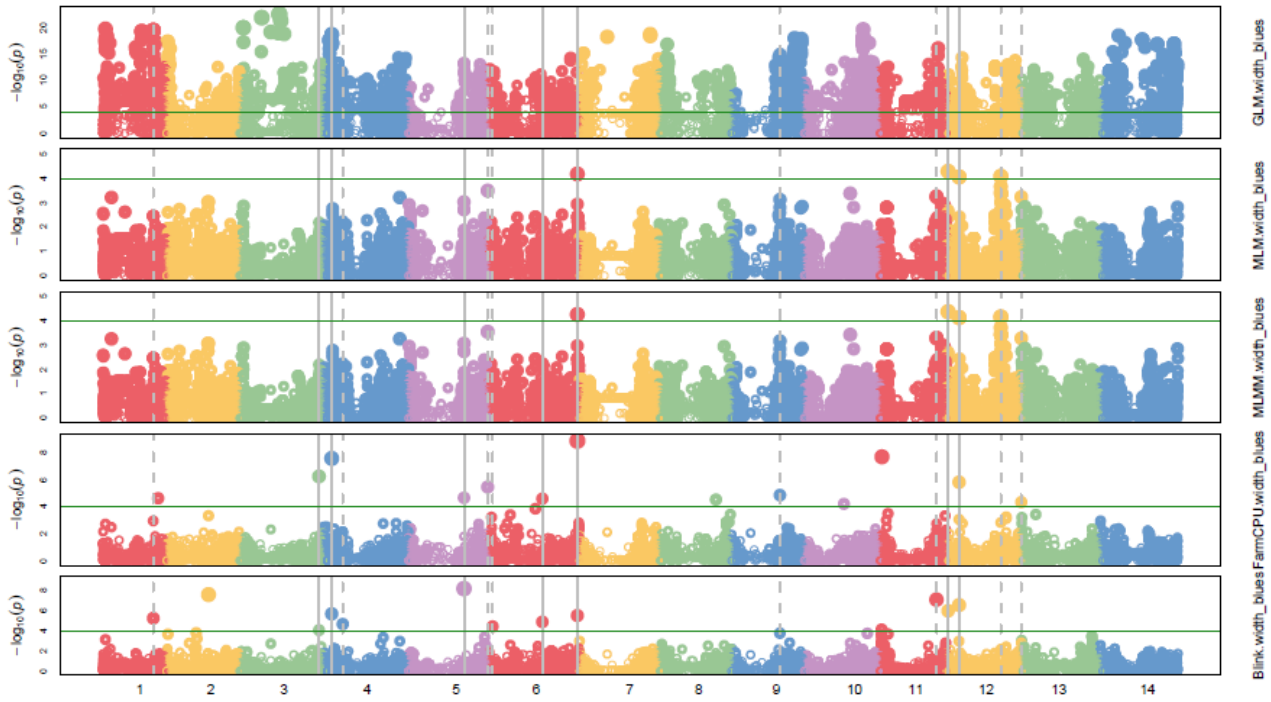


Figure 121 Manhattan plots for Grain width in GDP cluster

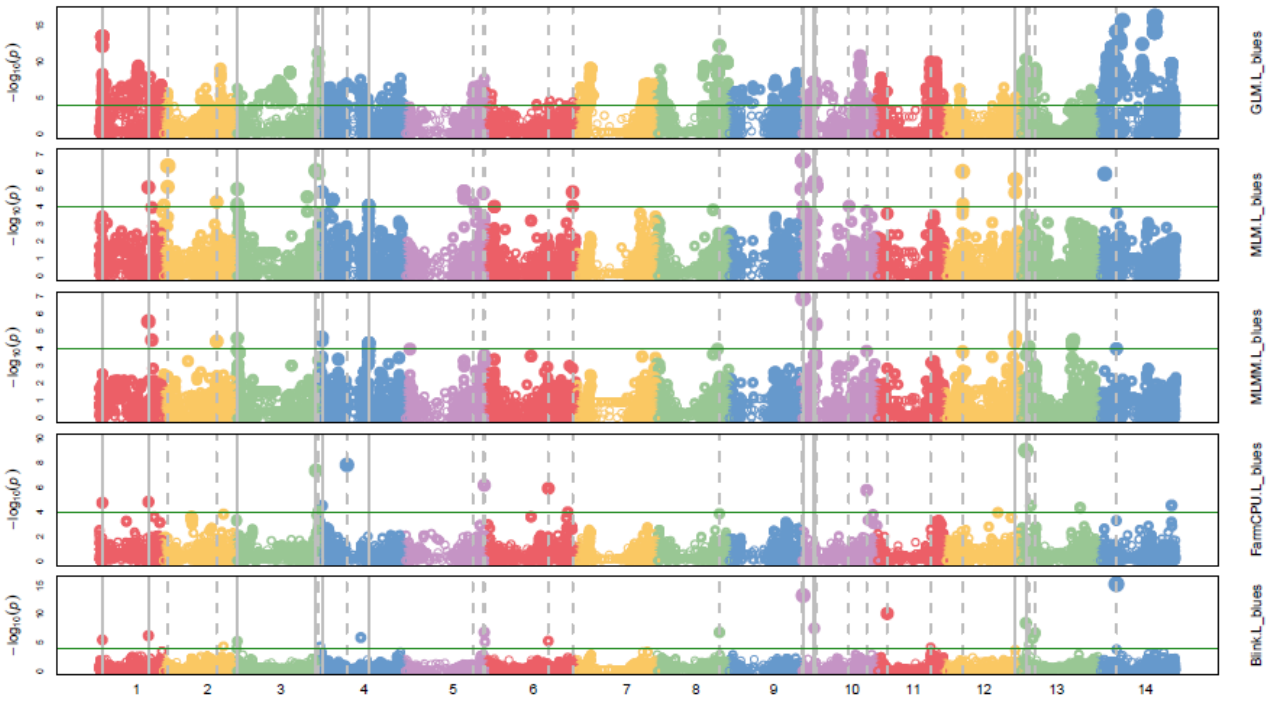


Figure 122 Manhattan plots for Grain brightness in GDP cluster

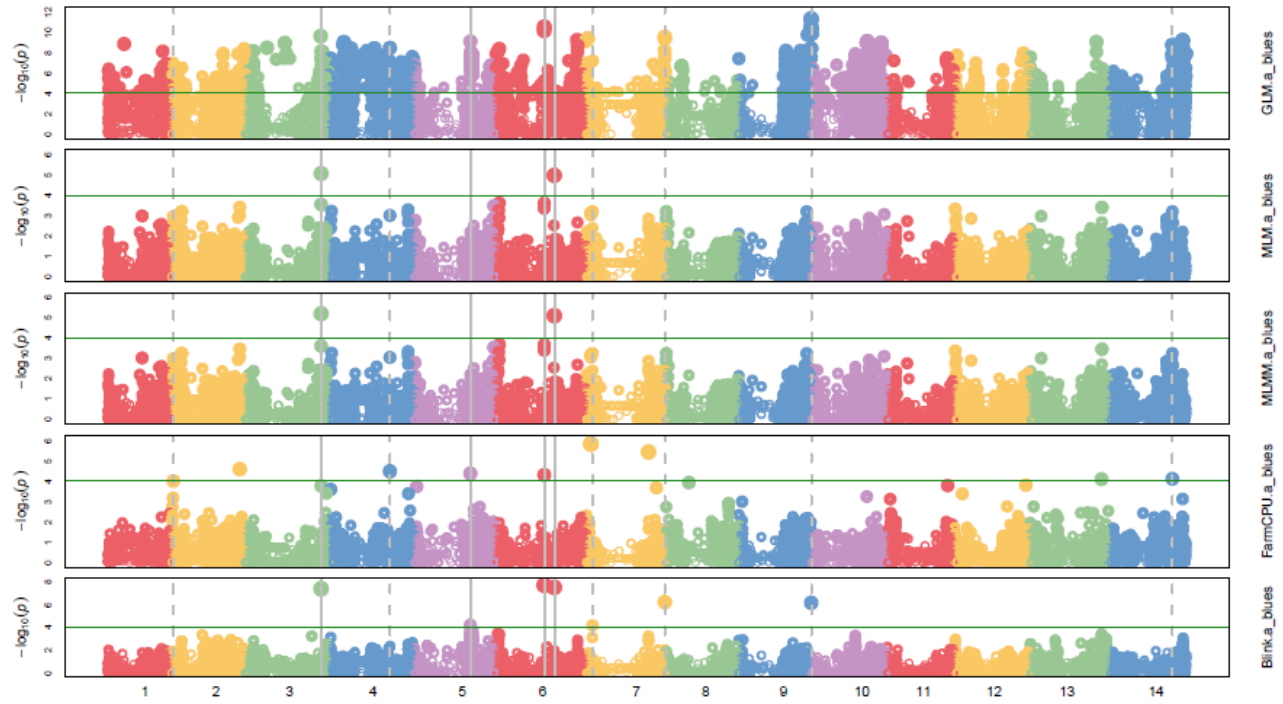


Figure 123 Manhattan plots for Grain redness in GDP cluster

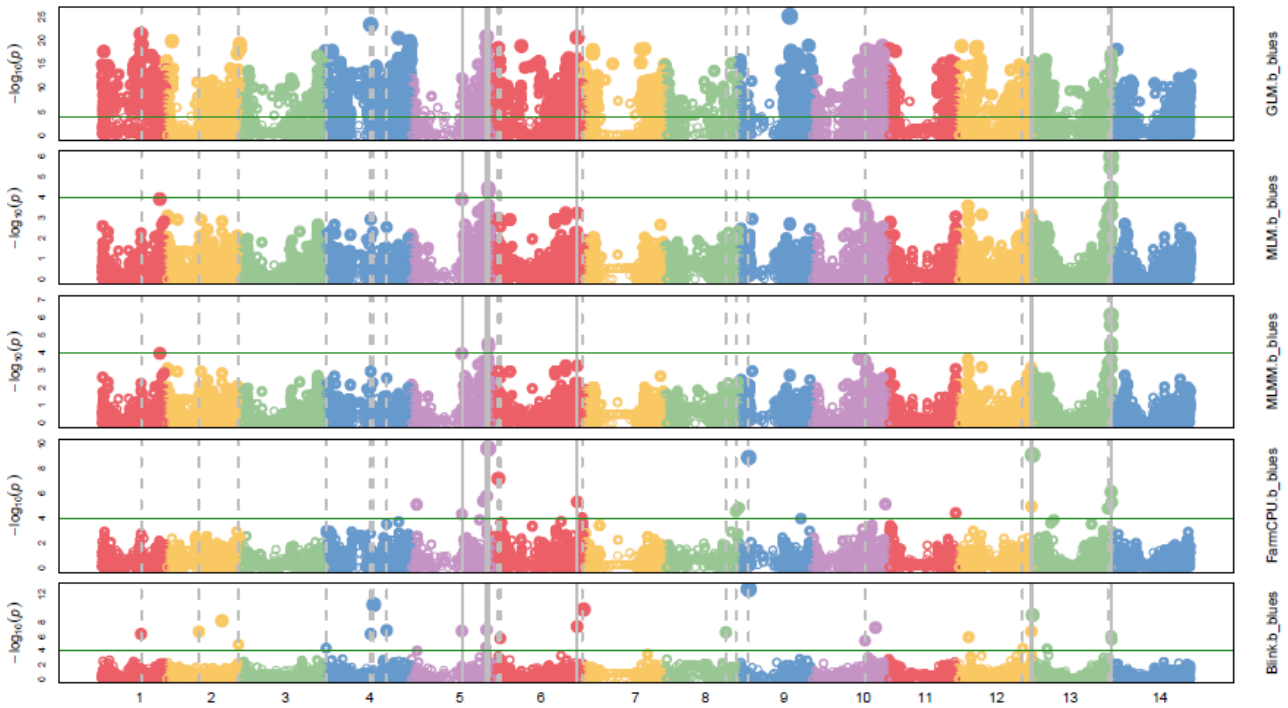


Figure 124 Manhattan plot for Grain yellowness in GDP cluster

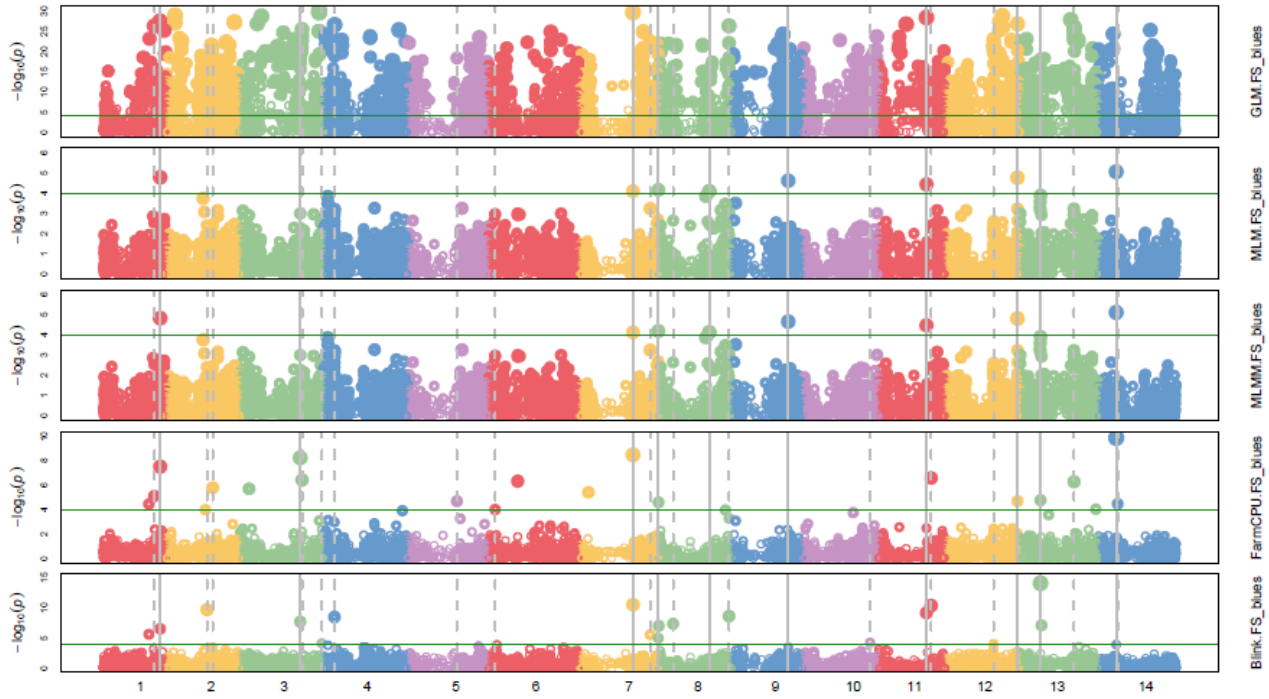


Figure 125 Manhattan plots for FS in TGC cluster

| TAG SNP | Chromosome | Position (bp) | -logP BLINK | -logP Farm CPU | Confidence Interval | Trait-Environment |
|---|---|---|---|---|---|---|
| IWB67307 | 2A | 35700820 | 4.94 | | IWB67308 - IWB51686 | FSGDP2020 |
| IWB70422 | 2B | 56659272 | 6.37 | | IWB70422 - IWA1093 | FSGDP2020 |
| IWA2595 | 4B | 656495107 | 6.04 | | IWB60914 - IWB74054 | FSGDP2020 |
| IWB65869 | 5B | 461264815 | 5.28 | | IWB2149 - IWA6291 | FSGDP2020 |
| IWB72132 | 2A | 368007148 | 4.41 | 3.52 | IWB52947 - IWB71845 | FFGDP2020 |
| IWB24626 | 3B | 499053436 | 3.75 | 3.86 | IWB65116 - IWB39029 | FFGDP2020 |
| IWB21625 | 4A | 75901226 | 3.92 | 4.88 | IWA7124 - IWA1320 | FFGDP2020 |
| IWA7725 | 6B | 23559843 | 5.77 | 5.02 | IWB47927 - IWB7667 | FFGDP2020 |
| IWB22186 | 1B | 556587625 | 9.12 | 7.01 | IWB66244 - IWB36872 | TKWGDP2020 |
| IWB21302 | 4B | 420592493 | 7.48 | 7.48 | IWB21302 - IWB15003 | TKWGDP2020 |
| IWB67149 | 7B | 49057316 | 7.80 | 10.50 | IWA3539 - IWB67149 | TKWGDP2020 |
| IWB37094 | 1A | 453222978 | 12.20 | 6.21 | IWB36443 - IWB51724 | areaGDP2020 |
| IWB9457 | 1B | 532189394 | 8.37 | | IWB35930 - IWA515 | areaGDP2020 |
| IWA3726 | 3B | 71641611 | 9.00 | 8.43 | IWB62905 - IWB24234 | areaGDP2020 |
| IWB7281 | 6A | 529249256 | 5.85 | | IWB7281 - IWB65994 | areaGDP2020 |
| IWB67149 | 7B | 49057316 | 8.84 | 6.88 | IWA3539 - IWB67149 | areaGDP2020 |
| IWB10413 | 1B | 449298414 | 5.20 | | IWB10413 - IWB43497 | lengthGDP2020 |
| IWB35796 | 5B | 378823034 | 7.84 | 4.70 | IWB27185 - IWB33255 | lengthGDP2020 |
| IWB8323 | 6A | 606862202 | 5.96 | 5.96 | IWA3909 - IWB72460 | lengthGDP2020 |
| IWB37094 | 1A | 453222978 | 7.76 | 8.77 | IWB36443 - IWB51724 | widthGDP2020 |
| IWB72208 | 6A | 75863240 | 4.80 | | IWB72207 - IWB51699 | widthGDP2020 |
| IWB67307 | 2A | 35700820 | 5.03 | | IWB67308 - IWB51686 | FSGDP2021 |
| IWB70422 | 2B | 56659272 | 5.68 | | IWB43273 - IWA1093 | FSGDP2021 |

| | | | | | | |
|---|---|---|---|---|---|---|
| IWA2595 | 4B | 656495107 | 5.53 | | IWB60914 - IWB74054 | FSGDP2021 |
| IWB72132 | 2A | 368007148 | 4.46 | | IWB52947 - IWB71845 | FFGDP2021 |
| IWB21625 | 4A | 75901226 | 4.27 | 4.22 | IWA7124 - IWA1320 | FFGDP2021 |
| IWA7725 | 6B | 23559843 | 4.37 | 5.14 | IWB26058 - IWB7667 | FFGDP2021 |
| IWB22186 | 1B | 556587625 | 7.46 | 5.97 | IWB66244 - IWB36872 | TKWGDP2021 |
| IWB21302 | 4B | 420592493 | 4.18 | 8.96 | IWB21302 - IWB15003 | TKWGDP2021 |
| IWB67149 | 7B | 49057316 | 6.27 | 11.18 | IWA3539 - IWB67149 | TKWGDP2021 |
| IWB37094 | 1A | 453222978 | 12.39 | 6.37 | IWB36443 - IWB51724 | areaGDP2021 |
| IWB9457 | 1B | 532189394 | 7.40 | | IWB35930 - IWA515 | areaGDP2021 |
| IWA3726 | 3B | 71641611 | 7.39 | 10.47 | IWB62905 - IWB24234 | areaGDP2021 |
| IWB7281 | 6A | 529249256 | 10.04 | 3.81 | IWB7281 - IWB65994 | areaGDP2021 |
| IWB67149 | 7B | 49057316 | 8.71 | 7.38 | IWA3539 - IWB67149 | areaGDP2021 |
| IWB67670 | 1B | 31235425 | 6.22 | | IWB71165 - IWB73610 | lengthGDP2021 |
| IWB21498 | 2A | 754200846 | 4.26 | | IWB21498 - IWB61340 | lengthGDP2021 |
| IWB27735 | 4B | 661727032 | | 6.28 | IWB8229 - IWB9880 | lengthGDP2021 |
| IWB41039 | 5B | 378824189 | 6.38 | 4.66 | IWB35796 - IWB33255 | lengthGDP2021 |
| IWB49960 | 7B | 681552072 | 5.07 | | IWA2191 - IWB73409 | lengthGDP2021 |
| IWB37094 | 1A | 453222978 | 7.24 | 9.80 | IWB36443 - IWB51724 | widthGDP2021 |
| IWB26242 | 1B | 19537743 | 8.76 | 6.33 | IWB72106 - IWB44700 | widthGDP2021 |
| IWB72207 | 6A | 75863121 | 4.84 | 4.80 | IWB72207 - IWB51699 | widthGDP2021 |
| IWB12994 | 3A | 15396933 | 4.41 | | IWB12994 - IWB52332 | LGDP2021 |
| IWB51790 | 5A | 428392124 | 9.50 | 9.96 | IWA4477 - IWB73898 | LGDP2021 |
| IWB51370 | 2B | 64884424 | 6.17 | 4.48 | IWB51370 - IWB44381 | aGDP2021 |
| IWB35598 | 5B | 511520004 | 4.63 | | IWB4569 - IWB7598 | aGDP2021 |
| IWB12984 | 1A | 358935623 | 5.19 | | IWA8026 - IWB15964 | bGDP2021 |

| | | | | | | |
|---|---|---|---|---|---|---|
| IWB8645 | 3B | 747334437 | 6.39 | | IWB73646 - IWB35001 | bGDP2021 |
| IWA6574 | 5A | 465821610 | 4.75 | | IWB22035 - IWB70649 | bGDP2021 |
| IWB11840 | 7A | 122485891 | 5.46 | | IWB65337 - IWB49474 | bGDP2021 |
| IWB67308 | 2A | 35700735 | 5.58 | 6.78 | IWB67308 - IWB51686 | FSGDPmulti |
| IWA6465 | 4B | 656489913 | 5.95 | | IWB72184 - IWB74054 | FSGDPmulti |
| IWA582 | 5A | 578511465 | | 5.04 | IWA3623 - IWB10414 | FSGDPmulti |
| IWB44988 | 5B | 459072112 | | 4.98 | IWB48406 - IWB7880 | FSGDPmulti |
| IWB21158 | 7A | 5590055 | 7.24 | 4.40 | IWB71146 - IWB34436 | FSGDPmulti |
| IWA3193 | 2A | 70572573 | 4.73 | | IWA5893 - IWB72480 | FFGDPmulti |
| IWB67301 | 2B | 2559456 | 4.22 | | IWB66351 - IWB7677 | FFGDPmulti |
| IWB67292 | 2B | 553730755 | 5.57 | 4.94 | IWB874 - IWA244 | FFGDPmulti |
| IWB2227 | 2B | 770410989 | 7.07 | | IWB58206 - IWB62759 | FFGDPmulti |
| IWA6850 | 4B | 36326487 | 4.88 | | IWB70449 - IWB61488 | FFGDPmulti |
| IWB35377 | 6B | 584331734 | | 5.06 | IWB44084 - IWA3636 | FFGDPmulti |
| IWB35066 | 1A | 39715027 | 5.90 | 4.30 | IWB35066 - IWB11970 | TKWGDPmulti |
| IWB5406 | 1A | 475115250 | 4.26 | 2.86 | IWB5807 - IWB31604 | TKWGDPmulti |
| IWB26242 | 1B | 19537743 | 5.82 | 3.43 | IWB72106 - IWB44700 | TKWGDPmulti |
| IWB54293 | 2A | 36290593 | | 5.92 | IWB67308 - IWB51686 | TKWGDPmulti |
| IWB25495 | 3B | 675968918 | 6.52 | | IWA3046 - IWB65507 | TKWGDPmulti |
| IWB73924 | 7B | 171541547 | 7.65 | | IWB73924 - IWB71851 | TKWGDPmulti |
| IWB70650 | 1A | 26062632 | 6.81 | | IWB3682 - IWB33537 | areaGDPmulti |
| IWB26242 | 1B | 19537743 | 12.08 | 7.00 | IWB8104 - IWB44700 | areaGDPmulti |
| IWB8334 | 2B | 181230947 | 5.36 | | IWB40225 - IWB8099 | areaGDPmulti |
| IWA6573 | 5A | 465822267 | 10.55 | 5.15 | IWB22035 - IWB70649 | areaGDPmulti |
| IWB52504 | 6A | 495888740 | | 3.54 | IWA8592 - IWB33680 | areaGDPmulti |

| | | | | | | |
|---|---|---|---|---|---|---|
| IWB73924 | 7B | 171541547 | 5.48 | | IWB73924 - IWB71851 | areaGDPmulti |
| IWB46974 | 1B | 450244694 | 4.54 | | IWB10413 - IWB2709 | lengthGDPmulti |
| IWB66417 | 2B | 728573174 | 7.65 | | IWB74 - IWB7129 | lengthGDPmulti |
| IWB37079 | 3A | 17214656 | 7.86 | 6.27 | IWB35874 - IWB72257 | lengthGDPmulti |
| IWB67460 | 6A | 606825167 | 8.53 | | IWB65928 - IWB16508 | lengthGDPmulti |
| IWB61977 | 7B | 710031002 | 4.45 | | IWB5972 - IWB62681 | lengthGDPmulti |
| IWA6835 | 1A | 468957024 | 5.23 | | IWB55805 - IWA6378 | widthGDPmulti |
| IWB73249 | 2B | 79053945 | 5.67 | 7.56 | IWB45339 - IWB67029 | widthGDPmulti |
| IWB50348 | 5A | 445388953 | | 4.84 | IWB40506 - IWB43738 | widthGDPmulti |
| IWB5996 | 6A | 524333330 | 7.07 | | IWB9600 - IWB33872 | widthGDPmulti |
| IWA8380 | 6B | 118588722 | 6.54 | | IWA6978 - IWB59110 | widthGDPmulti |
| IWB8941 | 2A | 759800800 | | 4.06 | IWB9423 - IWB66205 | LGDPmulti |
| IWB69456 | 7A | 61263506 | 8.44 | 9.05 | IWB40391 - IWB22591 | LGDPmulti |
| IWB1030 | 7B | 171576452 | 15.30 | | IWB73924 - IWB71851 | LGDPmulti |
| IWA7148 | 2A | 704236759 | 7.40 | | IWA5216 - IWB7166 | aGDPmulti |
| IWB59368 | 4A | 727533909 | 6.23 | | IWB2634 - IWB29720 | aGDPmulti |
| IWA2644 | 5A | 667286036 | 6.18 | | IWB71094 - IWA2646 | aGDPmulti |
| IWB12984 | 1A | 358935623 | 6.38 | | IWA8026 - IWB15964 | bGDPmulti |
| IWB23450 | 3A | 687444737 | 6.97 | | IWB23450 - IWB7306 | bGDPmulti |
| IWB23681 | 3B | 768866167 | 7.43 | | IWB60646 - IWB23680 | bGDPmulti |
| IWB22561 | 6B | 689678858 | 6.73 | | IWB2097 - IWB66694 | bGDPmulti |
| IWB72251 | 7A | 3034881 | 9.04 | 9.14 | IWB66267 - IWB21994 | bGDPmulti |
| IWB11725 | 2B | 238229658 | 16.33 | 10.06 | IWB23529 - IWB7772 | SSTGC2019 |
| IWA6680 | 7A | 668674104 | 14.20 | 13.07 | IWB5961 - IWA1032 | SSTGC2019 |

| | | | | | | |
|---|---|---|---|---|---|---|
| IWB50438 | 2B | 105765256 | | 5.79 | IWB68761 - IWA8381 | FSTGC2019 |
| IWB39508 | 3B | 698961638 | | 5.52 | IWB39915 - IWB44729 | FSTGC2019 |
| IWB34911 | 5A | 531602689 | 6.04 | | IWB33312 - IWB49700 | FSTGC2019 |
| IWB8581 | 5B | 622084007 | 4.98 | | IWB22266 - IWB56071 | FSTGC2019 |
| IWB60160 | 6B | 673212505 | 10.55 | 4.68 | IWB10268 - IWB55191 | FSTGC2019 |
| IWB5842 | 1A | 9931400 | 10.45 | 5.58 | IWB46412 - IWB1201 | FFTGC2019 |
| IWA6610 | 1B | 108722229 | 13.76 | 10.59 | IWB10085 - IWA1567 | FFTGC2019 |
| IWA3237 | 2B | 523912377 | 4.21 | 5.42 | IWB71212 - IWB32838 | FFTGC2019 |
| IWA6216 | 2B | 436048922 | | 3.82 | IWA5256 - IWB43933 | FFTGC2019 |
| IWA8290 | 3B | 588723900 | 8.59 | | IWB24473 - IWB27739 | FFTGC2019 |
| IWB70443 | 4A | 28430854 | 5.33 | | IWB26155 - IWB67723 | FFTGC2019 |
| IWB14901 | 7A | 106204557 | 4.74 | | IWB31199 - IWA7205 | FFTGC2019 |
| IWB70422 | 2B | 56659272 | 6.67 | 6.57 | IWB43273 - IWA546 | FSTGC2020 |
| IWB3569 | 4A | 718036011 | 5.31 | 5.10 | IWB62395 - IWA4651 | FSTGC2020 |
| IWB63365 | 1B | 505494440 | 5.12 | | IWB13329 - IWB31661 | FFTGC2020 |
| IWB10033 | 2A | 3321605 | 4.22 | | IWB41956 - IWA6745 | FFTGC2020 |
| IWB54322 | 3B | 10876002 | | 6.06 | IWB985 - IWB64002 | FFTGC2020 |
| IWB54164 | 4A | 113864372 | | 4.26 | IWA7271 - IWA3361 | FFTGC2020 |
| IWB5882 | 5B | 497440892 | 5.09 | | IWB10247 - IWB5882 | FFTGC2020 |
| IWB10413 | 1B | 449298414 | 6.96 | | IWB10413 - IWB46974 | TKWTGC2020 |
| IWB13742 | 1B | 592592522 | 7.07 | 5.87 | IWB8867 - IWB7410 | TKWTGC2020 |
| IWB21895 | 2B | 555610066 | 5.11 | | IWB46098 - IWA5141 | TKWTGC2020 |
| IWB29377 | 3B | 17838629 | 8.76 | 4.46 | IWA289 - IWB34925 | TKWTGC2020 |
| IWB45865 | 4A | 534094345 | 8.73 | 6.50 | IWB12211 - IWB55257 | TKWTGC2020 |

| | | | | | | |
|---|---|---|---|---|---|---|
| IWA3353 | 6B | 508079151 | 6.82 | 5.94 | IWA7084 - IWA5722 | TKWTGC2020 |
| IWB34065 | 7B | 610970231 | 6.79 | | IWB56081 - IWA5706 | TKWTGC2020 |
| IWA605 | 1A | 499298643 | 5.28 | | IWB43647 - IWB58517 | areaTGC2020 |
| IWB70456 | 1B | 381740389 | 5.87 | | IWB69041 - IWB66462 | areaTGC2020 |
| IWB11614 | 2A | 688361442 | 4.21 | 4.23 | IWB61299 - IWB7479 | areaTGC2020 |
| IWA429 | 2B | 145635634 | 4.77 | 3.93 | IWB32296 - IWB27957 | areaTGC2020 |
| IWB8291 | 3B | 758987014 | | 6.93 | IWB38921 - IWB12260 | areaTGC2020 |
| IWB35961 | 5A | 550109914 | 8.42 | | IWA46 - IWB44169 | areaTGC2020 |
| IWB69199 | 6A | 28196978 | 8.72 | | IWB70424 - IWB72985 | areaTGC2020 |
| IWB23124 | 6B | 146354976 | 6.28 | | IWA3424 - IWB10696 | areaTGC2020 |
| IWB55460 | 1A | 464975664 | | 4.83 | IWA5740 - IWB8994 | lengthTGC2020 |
| IWB1557 | 2A | 417851835 | 6.96 | | IWB57229 - IWA5293 | lengthTGC2020 |
| IWB10610 | 2B | 765476835 | 4.98 | | IWB70506 - IWA3474 | lengthTGC2020 |
| IWB49696 | 4B | 546670085 | 7.13 | | IWA5955 - IWB6922 | lengthTGC2020 |
| IWB11140 | 5B | 14079779 | | 5.71 | IWB73824 - IWB9179 | lengthTGC2020 |
| IWB14408 | 7B | 606338863 | 5.30 | 4.79 | IWB61109 - IWB14408 | lengthTGC2020 |
| IWA4008 | 1A | 21971024 | 4.43 | | IWB22004 - IWA7050 | widthTGC2020 |
| IWA6479 | 1B | 563620824 | 9.40 | | IWB7846 - IWB65886 | widthTGC2020 |
| IWB22672 | 2A | 770496662 | | 6.57 | IWB7101 - IWB7326 | widthTGC2020 |
| IWA4541 | 2B | 454923383 | 5.48 | | IWA4517 - IWB41706 | widthTGC2020 |
| IWB12116 | 5A | 444851220 | 4.88 | 8.87 | IWB40506 - IWB43738 | widthTGC2020 |
| IWB53342 | 5B | 668490739 | 6.64 | | IWB29437 - IWB25892 | widthTGC2020 |
| IWB11722 | 6A | 30332275 | 5.92 | 6.73 | IWB69175 - IWB26178 | widthTGC2020 |
| IWB35738 | 7A | 194295012 | 10.60 | 4.93 | IWB41777 - IWB35738 | widthTGC2020 |

| | | | | | | |
|---|---|---|---|---|---|---|
| IWA3037 | 2B | 596004366 | 4.59 | | IWB50067 - IWA2189 | LTGC2020 |
| IWB32997 | 4B | 664950191 | 6.28 | 7.28 | IWB8859 - IWB32997 | LTGC2020 |
| IWB73963 | 5A | 451516943 | 9.09 | | IWB14493 - IWB71451 | LTGC2020 |
| IWB46448 | 1A | 505711246 | | 6.13 | IWA3406 - IWA6145 | LTGC2020 |
| IWB21864 | 2A | 117552097 | | 7.51 | IWB45503 - IWB66712 | aTGC2020 |
| IWA2649 | 3A | 595302964 | 7.25 | | IWA1462 - IWA2649 | aTGC2020 |
| IWB23681 | 3B | 768866167 | 8.20 | | IWB60646 - IWB23680 | aTGC2020 |
| IWB42829 | 6A | 26766423 | 7.83 | | IWB22480 - IWB70424 | aTGC2020 |
| IWA6489 | 1A | 9061864 | 7.80 | 7.27 | IWB33789 - IWB10312 | bTGC2020 |
| IWB30429 | 3A | 720847800 | 5.84 | | IWB14695 - IWB60694 | bTGC2020 |
| IWA628 | 3B | 260580846 | 4.92 | 6.81 | IWB1111 - IWB42046 | bTGC2020 |
| IWB34259 | 4B | 13403079 | 7.46 | 4.27 | IWB63893 - IWB72103 | bTGC2020 |
| IWB23612 | 6A | 29437066 | 8.92 | | IWB43285 - IWB72838 | bTGC2020 |
| IWB34211 | 7B | 72062296 | 6.60 | | IWB35358 - IWB67435 | bTGC2020 |
| IWB11011 | 1B | 431037305 | 3.51 | 5.80 | IWB71872 - IWB7028 | FSTGC_multi |
| IWB60297 | 2A | 562062720 | 7.69 | 8.27 | IWB44801 -IWB13477 | FSTGC_multi |
| IWB11072 | 2B | 104336022 | 8.48 | | IWB72913 - IWB68761 | FSTGC_multi |
| IWA1100 | 4B | 656790359 | 8.61 | | IWB74189 - IWB48353 | FSTGC_multi |
| IWB34911 | 5A | 531602689 | 3.48 | | IWB33312 - IWB49700 | FSTGC_multi |
| IWB8932 | 6A | 444333914 | 9.17 | | IWA2416 - IWA428 | FSTGC_multi |
| IWB36146 | 6B | 451473732 | 4.01 | | IWA1251 - IWB38147 | FSTGC_multi |
| IWB56168 | 7A | 500533708 | | 6.29 | IWB7506 - IWB21762 | FSTGC_multi |
| IWB5842 | 1A | 9931400 | 4.65 | 4.29 | IWB3088 - IWB1201 | FFTGC_multi |
| IWB24065 | 1B | 14429522 | 3.95 | 5.63 | IWB2188 - IWB73279 | FFTGC_multi |
| IWB38008 | 2A | 64253822 | 3.59 | 7.02 | IWB38008 - IWB3684 | FFTGC_multi |

| | | | | | | |
|---|---|---|---|---|---|---|
| IWB5529 | 2B | 43053687 | 5.68 | 4.55 | IWB60077 - IWB23606 | FFTGC_multi |
| IWB32738 | 3A | 735562087 | 4.82 | 7.16 | IWB65706 - IWB50704 | FFTGC_multi |
| IWB29632 | 5A | 10559440 | 3.67 | | IWB43705 - IWB50392 | FFTGC_multi |
| IWB7107 | 5B | 661651875 | 5.48 | 11.45 | IWB10034 - IWB50537 | FFTGC_multi |

*Figure 126 Main peaks detected through BLINK and Farm CPU model in GWAS*