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

IDENTIFICATION OF RESISTANCE FACTORS TO THE MAIN WHEAT DISEASES IN THE MEDITERRANEAN AREA

Presentata da: Jad B Novi

Coordinatore Dottorato

Massimiliano Petracci

Supervisore

Marco Maccaferri

Esame finale anno 2023

Summary

Abstract

Several diseases challenge bread and durum wheat productions worldwide. The importance of these cereals requires adequate protection to pathogens that can cause strong yield and grain quality losses. The main chapter is about *Puccinia striiformis* f.sp. *tritici,* the agent of yellow rust on wheat, is one of the major causes of production losses worldwide that imposes the breeders to research new sources of durable resistance in the germplasm globally. The collaborative Global Durum Resource collection (https://wheat.pw.usda.gov/GG3/global_durum_genomic_resources) composed by a wide range of durum cultivars, landraces, and pre-breeding lines, was assessed in multi-environment conducive nurseries in the Mediterranean region (Italy, Egypt, Lebanon, Morocco and Turkey) and Argentina across three years (2019-2021) for yellow rust resistance (infection type and severity). The Global Durum Resource is provided with the Illumina wheat 90K SNP array and population structure has already been described (Maccaferri et al 2019; Mazzucotelli et al 2020). GWAS using different models (GLM, MLM, MLMM, FarmCPU and BLINK) and including the kinship matrix as covariate identified at least 15 genomic regions harboring loci involved in YR response in the tetraploid wheat germplasm. In particular, loci in chromosome 1B, 2B, 4B, 5A, 6A, 7B showed high significance across nurseries/years, with various patterns of GxE. The wide range of observed infection types and the population numerosity allowed also to perform a differential GWAS analysis for complete (IT = 0-3) or partial resistance (IT = 4-6), revealing loci with IT-specificity, allowing further analyses for the identification of QTLs useful in breeding process. The second chapter is about *Zymoseptoria tritici*, agent of STB (Septoria Tritici Blotch), a foliar pathogen that yearly causes high damages if not controlled. Recurrent epidemics are observed in northern Europe for bread wheat and in the Mediterranean countries for durum wheat, particularly in Tunisia, Italy and Southern France/Spain. Strong and environmentally unsustainable fungicide treatments are required to control the pathogens. Moreover, this approach is further complicated by the resistance that this pathogen easily develop to the most diffused fungicide classes. On the other side, modern cultivar stocks are depleted in effective resistances. For all these reasons, in recent years research in durum wheat breeding is focused on the identification of novel, underexploited resistance genes to be subsequently and conveniently moved into the pre-breeding and breeding stream. The experiment carried out in this thesis concerned the phenotyping of a landrace subset, which were inoculated in the UNIBO experimental field, in Cadriano (BO). The plants were phenotyped for disease height characters, infection type at the flag leaf and infection type at the level of the canopy below the flag leaf. This experiment opens up a rich scenario of analysis and opportunities to investigate and discover new loci of resistance to STB. Third chapter is about Fusarium head blight (FHB) is a fungal disease caused by pathogens belonging to the genus *Fusarium*. In particular, *Fusarium culmorum* and *Fusarium graminearum* species cause severe grain yield losses and accumulation of mycotoxins (e.g. deoxynivalenol or DON) in wheat that compromise food safety and animal health. Three main types of resistance to FHB have been reported in wheat: Type I acts against initial infection, while type II against spread within the head, Type III (DON content) is directly related to the final kernel contamination levels and it is therefore worth to assess Durum wheat *Triticum turgidum* ssp durum is known to be highly susceptible to FHB compared to bread wheat *Triticum aestivum*. None of the selected durum wheat varieties or lines show a level of resistance to FHB comparable to Sumai 3. Over 250 QTL/genes for FHB resistance have been identified in bread wheat, such as Fhb 1 and Fhb 5 but only a small number of FHB resistance loci have been mapped in durum wheat. The aim of this work is to find loci of partial resistance to FHB already present in durum and bread wheat germplasm and therefore easily cumulative. This experiment involved the inoculation and phenotyping of a hard and a bread panel. The durum wheat

panel was phenotyped in one location in 2021 and was mainly composed of modern and landrace varieties. The bread wheat panel, on the other hand, was sown in three different locations and is made up of modern varieties supplied by different breeding companies. The experiment included the detection of incidence and severity, which will then be used to carry out a GWAS analysis for the identification of new resistance loci.

Figures and Tables summary

1 Introduction

1.1 Taxonomy of durum wheat

Durum wheat (*Triticum turgidum* ssp*. durum*) is a cereal belonging to the Poacea family, more specifically to the tribe of Triticeae and the genus *Triticum*. The genus *Triticum* includes six species of wheat, all of which have multiples of a basic set of seven different chromosomes (n = x = 7) on gametic cells and two set in somatic cells (2n =2x = 14). The two species with the smaller number of chromosomes are *Triticum monococcum* (AA genome) and *triticum urartu* (AA genome). *Triticum turgidum* (AABB) and *Triticum timopheevii* (AAGG) are tetraploids, while *Triticum aestivum* (AABBDD genome) and *Triticum zhukovskyi* (AABBGG genome) are hexaploids. Modern cultivars of durum wheat belong to the ssp. *durum* (Blanco, 2007; Matsuoka, 2011).

1.2 Origin and evolution of durum wheat

Man lived a nomadic lifestyle as a hunter-gatherer, relying on hunting wild animals and collecting wild plants for his food up to the Neolithic revolution which happened ca. 10.000 years ago when human evolved to an agrarian lifestyle. The Neolithic revolution took root in the Levantine Corridor and spread through the Fertile Crescent, which is located in the Middle East and includes a region stretching from Jordan, Isreal, Lebanon and Syria through southeastern Turkey and along the Tigris and Euphrates rivers through Iraq and western Iran (Tuberosa, Graner, & Frison, 2014). This area was the center of domestication of wheat. Domestication is the process of genetic selection that, transforms wild forms into domesticated varieties of crops altering key traits. The transition from wild to domesticated forms of crops has mainly interested changes to the three principal morphological features that make the crop easier to harvest: seed size, ear rachis stiffness and the easeness with which the seeds are released from its enclosing glumes, i.e., threshability (Salamini et al., 2002). Durum wheat has a tetraploid genome (AABB) due to the cross between *Triticum urartu (AA)* and *Aegilops speltoides (BB genome)* (Akhunov et al., 2005) which originated the ancestral wild wheat (*Triticum turgidum* ssp. *dicoccoides*) from which the emmer wheat (*Triticum turgidum* ssp. *dicoccum*) and then the durum wheat were derived, again about 10,000 years ago and later (Özkan et al., 2002).

1.3 Wheat importance in world and Italy

According to FAO, the world wheat production is near 771 million of tons and Asia is the first producer with 287 million (43,6%). After we found Europe with 217 million of tons (33%) and Americas (16%), Africa (3,4%) and Oceania (3,4%). Main producer countries are China, India, USA and Russia. Wheat is a major source of energy, protein, and dietary fiber in human nutrition and animal feeding. Currently, about 95% of the wheat grown worldwide is hexaploid bread wheat, with most of the remaining 5% being tetraploid durum wheat. (Žilić et al., 2011). Wheat is the source of 20% of the calories consumed by humans (Brenchley et al., 2012) and is the most cultivated cereal after maize. Wheat optimum is between the latitudes of 30° and 60° N and 27° and 40° S. In altitude, the crop is grown from sea level to more than 3 000 m.a.s.l., and it has been reported at 4 570 m.a.s.l. in Tibet (FAO). The EU produced an annual average of 147.1 million tons of soft wheat and 8.0 million tons of durum wheat during the 2012-2016 time period. Most of the durum is produced in Italy, Greece, and France. Italy accounted for nearly 58% of EU durum production in 2016, followed by Greece (21%) and France (12%) but also Germany and United Kingdom are great producers (Taylor, 2017). Durum wheat is mostly cultivated in semi-arid areas of Mediterranean and it is used for pasta production, one of the most important products in Italy. Usually, gluten and protein are considered the most important factors that influence pasta quality, but the global quality of durum wheat is a complex of environmental factors, yield characteristics, genetics factors and technological requirements (Fagnano et al., 2012). Wheat production in Italy was almost 7 million tons in 2017 (FAO), of which 4,2 is durum wheat. In Italy, about 67% of durum wheat production comes from the Southern regions and it is mainly used for pasta production (Fagnano et al., 2012) and it is mostly cultivated in Puglia (943.000 tons), Sicily (807.000 tons), Emilia-Romagna (461.000) and Marche (455.000) (ISTAT, 2017).

1.4 Evolution of wheat breeding in Italy and new perspective

During XIX and the beginning of XX century, the cultivation of durum wheat in Italy was based on the utilization of a huge of landraces, being the only source of germplasm available (Motzo et al., 2001). A landrace is a cultivated, genetically heterogeneous variety that has evolved in a certain ecogeographical area and is therefore adapted to the edaphic and climatic conditions and to its traditional management and uses (Casañas et al., 2017). Between years ’20 and ’50, durum wheat

was subjected to breeding program, with the selection of pure inbred lines and with the first adoption of the artificial crossing. Some breeders were doing interspecific cross followed by selection (Motzo et al., 2001). The most important Italian breeder of this period was Nazareno Strampelli, he was mainly working to make up a variety with simultaneous resistance to allurement, rust and early ripening. Nazareno Strampelli was very smart because he got 250 varieties of wheat from all over the world, to be crossed with the variety "Rieti", the best one variety resistant to rust. In 1913 Nazareno Strampelli did his masterpiece: the cross between two lines, “aristate 21” and “mutica 67” (obtained crossing the varieties Rieti and Wilhelmina) with the Japanese early wheat Akakomughi. The varieties obtained with this cross were resistant to rust (inherited by Rieti, gene Lr34), with early maturation (inherited by Akakomughi, gene *Ppd-D1*) and resistant to allurement (obtained thanks to a lucky case of linkage between *Ppd-D1* and *Rht8*, associated to semi-dwarf phenotype) (Lorenzetti et al., 2018). Nazareno Strampelli was also the constitutor of Senatore Cappelli: this cultivar was released in 1915 and it is still cultivated in Italy, after almost 90 years, it was without any doubt the most important constitution of this period, covering up to 60% of the Italian durum wheat surface. The origin of Cappelli is traced back to the North African population Jean Retifah. Even though Senatore Cappelli is a cultivar susceptible to rust, lodging and with late maturation, it has important characteristics linked to high yield and high semolina quality (Motzo et al., 2001).

Nowadays we have cultivated modern wheat varieties from modern breeding activities that must respond to the requirements of distinctiveness, uniformity, stability and superior agronomic value. However, excessive uniformity towards the top ideotypes in the elite germplasm is also associated with negative aspects, especially regarding to the impact of the loss of resilience to abiotic and biotic stress, and genetic erosion due to inbreeding and limited genetic diversity due to intense selection in front of non-adequate pre-breeding activities (Motzo et al., 2001). According to FAO, genetic erosion may thus be defined as a permanent reduction in richness or evenness of common or localized alleles or the loss of combination of alleles over time in a defined area. This happens mostly for the resistance to pathogens because hosts and pathogens survive together in a complex balance that sees no one dominate, thanks basically to the genetic heterogeneity of natural populations and the discontinuity in spatial distribution. With genetic erosion, there is the necessity to find and store new genetic variability from wild variety and landraces (Motzo et al., 2001). At the same time, there is also the need to find and preserve new sources of genetic variability, and to achieve this goal, there is germplasm conservation. Germplasm conservation is the most successful method to

conserve the genetic traits of endangered and commercially valuable species. Germplasm is a live information source for all the genes present in the respective plant, which can be conserved for long periods and regenerated whenever it is required in the future. Germplasm conservation can be done *in-situ*, which means conservation of genetic resources in the form of natural populations by establishing biosphere reserves such as national parks and sanctuaries, or *ex-situ*, which deals with conservation of an endangered species outside its natural habitat. In this method genetic information of cultivated and wild plant species is preserved in the form of in vitro cultures and seeds, which are stored as gene banks for long-term use. This type of conservation creates a bank of genes/DNA, seeds, and germplasms and forms a genetic information library (Bhatia, 2015). The material stored can be used for breeding and the introduction of new genetic variability into modern varieties for fight pathogen evolution.

1.5 Overview of crop protection

After the industrial revolution, there has been a development in agriculture linked to the emergence of new disciplines and technologies that have led to a significant increase in production, mainly thanks to progress in agronomy, breeding, advances in the industry of fertilizers and protection products, greatly reducing the risks to which crops are subject. However, these advances are associated with costs and side effects such as greenhouse gas emissions, air and water pollution, reduction of genetic variability and "loss" of useful genes under stress conditions (Brodt et al., 2011). The concept of sustainable agriculture has thus been developed, with environmental health, social equity and economic viability among its fundamental objectives (Brodt et al., 2011). Breeding improvement, linked to good management practices, plays an essential role (Richards et al., 2007) in exercising sustainable farming practices. The search for resistance genes against pathogens plays a key role in reducing the use of pesticides and, consequently, limiting costs and pollution.

In view of the climate changes that are affecting our planet, it is necessary to identify genes useful for increase the resilience of today's varieties. Climate change is the main cause of: alteration and formation of new microclimates, changes in the plant-pathogenic relationships, causing the emergence and spread of new diseases that the previously selected varieties may not be able to deal with. Higher temperatures and changes in rainfall patterns can lead to genetic variation and changes in pathogen populations, e.g. changes in geographical distribution, new population dynamics (e.g. number of generations) and changes in crop phenology, with greater synchronization

of host-pathogen cycles (Juroszek & von Tiedemann, 2013). From this derives the need for genetic improvement to identify and use genes of resistance useful to the changed needs, which can be found in the non-elite germplasm, especially among the populations of landraces: traditional varieties grown and adapted to a particular environment that, thanks to ecological adaptation, can be considered to recover genetic variability useful against the pathogens (Pinheiro de Carvalho et al., 2013).

Pathogen resistance can be divided into vertical resistance (given by a single gene) and horizontal resistance (Kou & Wang, 2010). Vertical resistance is given by single genes, called R-genes (or resistance genes) (Poland et al., 2009), which on the basis of a gene-per-gene interaction with pathogen proteins, trigger hypersensitive responses of high resistance. It is characterized by high effectiveness but is easily overcome because it is based on individual genes. Horizontal resistance, on the other hand, made up of the additive effect of several genes, presents greater flexibility as regards the mutability of the pathogen, being effective also against several races of the pathogen and it is more durable (Parlevliet & Zadoks, 1977; Poland et al., 2009). It is important for resistance that this is "durable", in other words that it is maintained in the cultivar even after it has been widely cultivated in an environment (Johnson, 1983). To obtain it, it may be necessary to pyramid several resistance factors, even from different sources, in order to obtain a broad-spectrum resistance (Fuchs, 2017).

1.6 Most important foliar and spike diseases in wheat

The main fungal disease in wheat are represented by rusts, Fusarium and Septoria. Most important rusts are: yellow rust (*Puccinia striiformis* f. sp. *Tritici*), brown rust (*Puccinia triticina* Erikson = *P. recondita* Roberge ex Desmaz f. sp. *tritici*) and black rust (*Puccinia graminis* f.sp. *tritici*). Yellow rust is the most aggressive and can cause a drop-in production of up to 100% in the event of major epidemics, although losses usually range from 10 to 70% (Chen, 2005). Resistance to this pathogen is usually race-specific, inducing to wheat a strong resistance (Xu et al., 2013), based on HR reactions (hypersensitive response), that is not durable being easily overcome by pathogens. According to recent observations by CIMMYT, in recent years yellow rust has been spreading throughout Italy and the Mediterranean, both on durum wheat and on common wheat. The CIMMYT also states that on bread wheat, new Warrior races have been developed starting from the regions of Northern Europe, adapted and partly helped by higher temperatures, as the fungus can adapt and benefit

from a temperature increase (Gautam et al., 2013). Brown rust is a pathogen that in the Mediterranean area becomes highly dangerous for durum wheat, as it is endemic and every year capable of inducing particularly widespread epidemic episodes. It occurs annually in a wide range of environments causing significant yield losses (more than 50%) under favourable field conditions on the susceptible wheat genotypes (El-Orabey, 2018). The third species of rust on wheat is black rust, that has sparked worldwide attention for cases of virulence towards important resistance genes because of the Ug99 race (Pretorius et al., 2010; Singh et al., 2011). In 2016 was reported an outbreak of black rust in Sicily, caused by the new strain TTTTF-like, which destroyed hundreds of hectares (Bhattacharya, 2017). According to FAO, without adequate control, the pathogen could spread very quickly along the Mediterranean area and the Adriatic coast, endangering the cultivation of wheat. Another important fungal disease on wheat is Fusarium Head Blight, caused by *Fusarium* sp. (the most important is *Fusarium graminearum*). Epidemics of Fusarium head blight occur frequently in the USA, the UK, and China, causing billions of dollars losses not only from yield and quality reduction but also from the contamination of trichothecene mycotoxins, which threaten the health of humans and animals. Sources for Fusarium head blight resistance are limited, and thus far only one resistance gene, PFT, has been identified in wheat, and no fully resistant cultivars are yet available (Jia et al., 2019). The most important foliar fungal disease in the context of this thesis is Septoria Tritici Blotch, caused by *Zymoseptoria tritici*, which will be discussed later in this paper.

1.7 Association mapping and genome-wide association studies (GWAS)

The study of association mapping consists in identifying a correlation between genotype and phenotype in each sample of individuals, based on their level of linkage disequilibrium (LD) (Breseghello & Sorrells, 2006). Goode, 2017, define linkage disequilibrium as the non-random association of alleles at two or more loci in a general population. When alleles are in linkage disequilibrium, frequencies of haplotypes do not occur like expected. Linkage disequilibrium between two alleles is related to the time of the mutation events, genetic distance, and population history (Goode, 2017). The density of the molecular markers to be used is determined by the level of LD (Flint-Garcia et al., 2003). The two most widely used data sets for studying genetic variability are those derived from biparental crosses (like F2 populations or recombinant inbred lines (RILs)) and those that consist of individuals assembled with complex relatedness or geographical origin (diversity populations, also called diversity panels). These two sets of data differ in one very

important characteristic: the number of recombination events they manage to capture. Biparental crossings only exploit recombination events that occurred during the constitution of the population, so only the most recent ones are studied. Instead, the populations of diversity can provide more information, having captured all the historical events of recombination that occurred during the evolution of the individuals used. The populations of diversity are typically obtained by collecting as many information samples as possible to make them highly representative of the population under study. The most used association mapping method is GWAS, a technique that involves testing for association most of the segments of the genome, using molecular markers (i.e. SNPs) densely distributed in the genome (Rafalski, 2010). Despite its high effectiveness, low cost and speed, genome-wide association analysis is not free of weaknesses, related do false positive and negative. The main problem related to false positive is population structure correlation. The population structure creates genome-wide linkage disequilibrium between unrelated loci. When, due to factors such as domestication, genetic drift or selection, the allele frequencies between subpopulations of a species are significantly different, loci that have no effect on the trait can co-segregate with the locus related of the trait of interest, giving rise to false positives for their high statistical power. False negatives, on the other hand, occur when there are low-frequency causal alleles or with QTLs with low effect (Korte & Farlow, 2013). While the basic GWAS analysis involved the analysis of SNPs by studying two allelic variants, now GWAS is moving towards the analysis of haplotypes. Maldonado et al. (2019) defines the haplotype as a set of SNPs that have a strong linkage disequilibrium between them. This type of GWAS can be more efficient than the use of individual markers, also avoid the limit related to biallelic markers, being able to increase the efficiency of detection of QTLs and the knowledge about genetic traits that cannot be studied with a single marker approach (Maldonado et al., 2019).

1.8 GWAS technologies in wheat

To perform GWAS in wheat, one of the most used methods is to use the 90K SNPs array with 90 000 gene-associated SNPs that provides dense coverage of the wheat genome (Wang et al., 2014). You et al. (2018) defin SNP array like a type of DNA microarray containing designed probes harboring the SNP positions, which is hybridized with fragmented DNA to determine the specific alleles of all SNPs on the array for the hybridized DNA sample. This method is widely used also to detect marker–trait associations in quantitative trait locus mapping experiments (Wang et al., 2014). Recently, a

system that can perform sequencing and genotyping has been developed and named genotyping-by-sequencing (GBS) to be able to genotype and simultaneously find new SNPs. GBS is a novel one-step approach and constitutes a rapid, efficient, high-throughput and cost-effective tool for GWAS (Baloch et al., 2017). Other tools of interest are represented by DArT, exome capture and KASP® assays. DArT simultaneously types several thousand loci in a single test and generates fingerprints of the entire genome by assessing the presence or absence of DNA fragments in genomic representations generated by genomic DNA samples (Akbari et al., 2006). Exome capture is a method used to extract and sequence the exome. This technique restricts the attention only on the portions of DNA that encode mRNA and, eventually, lead to a certain phenotype, going to explain the genetic variation and is a good strategy for finding rare alleles and sometimes complex traits (Bamshad et al., 2011; Kaur & Gaikwad, 2017). On the other hand, KASP® is a homogeneous genotyping technique that is founded on fluorescence (Semagn et al., 2014) and it is based on allele-specific oligo extension and fluorescence resonance energy transfer (FRET) for signal generation (Kumpatla et al., 2012; Semagn et al., 2013). The fluorescent reporting system comprises of four single labelled oligonucleotides that hybridize to one another in free solution to form a fluorescent quenched pair which upon introduction of complementary sequences generates a measurable signal (Kumpatla et al., 2012). In 2019, the annotated sequence (where genes and their transcripts were identified) of Svevo (durum wheat) was discovered (Maccaferri et al., 2019), providing the basis for all gene expression and genetic diversity analyses that can also be used for pathogen resistance research.

1.9 Genotyping technologies

As a result of genome wide association, it is essential to develop a laboratory assay to predict the allelic status of the identified accessions, which can then also be used later in applied breeding. A genotyping that today is performed in wheat predominantly through single nucleotide polymorphisms (SNPs). A single nucleotide is a variation in the gene material at a single nucleotide such that which the two polymorphic alleles are found to be present in the reference population in well- characterized (= mutation by nucleotide substitution). SNP polymorphism is the type of

most frequent mutation by far and therefore characteristically allows the whole genome to be probed, including near or within the causative genes of interest.

There are two large families of technologies for characterizing SNPs that differ in the amount of data they can provide: high-throughput and low-throughput technologies. High-throughput includes high-processivity techniques that result in fine characterization over many individuals: a lot of information is evaluated in a single reaction, such as the iSelect SNP array technology, which allows two million random SNPs to be analyzed and plumb the entire genotype of an individual. Low-throughput techniques allow little information to be obtained on many individuals, when less detail is needed, as in the case of selection for the resistance character to a particular disease. In this context, screening is performed on thousands of individuals for a few genomic regions. Among these assays, the most widely used are HRM assays (high-resolution melting) and KASP® with allele-specific primers.

1.10 KASP® markers

KASP® (kompetitive allele specific pcr) technology enables genotyping of single nucleotide polymorphisms and is based on the principle of competitive PCR.

Competitive allele-specific PCR is an end-point genotyping technology that allows detection of which base is present at the base of the SNP, in a simple and cost-effective manner. Unlike the normal PCR protocol, three primers are designed here: two allele-specific primers, one for each allele of the SNP of interest (these are biallelic varietal SNPs), and one genome target-specific (common) primer. Each allele-specific forward has a 5' free tail to which fluorophores (Fret cassette) will be paired. It should be pointed out that in a diploid species it is necessary to make sure that there are no paralogs: otherwise

otherwise the common primer should be specific for the copy of interest, to identify the right genome. There are two allele-specific varietal primers, with two different colored fluorescents.

During the first round of PCR, the allele-specific primer appends with the target SNP, along with the common primer, amplifying the target region. Only the allele-specific allele terminating with the right base will pair up and, in subsequent pcr cycles, increase exponentially. The tails labeled by the FRET cassettes will pair with the new complementary sequences and dissociate from the quencer, emitting fluorescence. Eventually you will have fluorescence from the right primer in the PCR products. This is called competitive PCR precisely because there are two variety-specific primers,

with two different fluorescence colors in competition: the primer that ends with the correct base (that of the SNP present) wins.

2. Chapter I: Yellow rust (*Pst*)

2.1 Introduction

2.1.1 Yellow rust: the disease

Stripe or yellow rust, caused by *Puccinia striiformis* Westend. (Ps), it is one of the most destructive and widespread diseases in cereal production. The special form that infects wheat is called *P. striiformis* f.sp. *tritici* (Pst) (Wellings, 2011). The pathogen is widespread in almost all areas of the world, where wheat is widely cultivated, and losses reach up to 5.47 million tons annually, both for durum and soft wheat (Lin et al., 2018). The pathogen that causes yellow rust initially infects the green tissues of the wheat and infection can occur at any time, from the juvenile stage to the maturity of the plants, as long as they are still green. In optimal conditions for the pathogen, symptoms appear around a week after the infection has occurred, while sporulation begins after about two weeks. The fungus forms small yellow-orange pustules, called uredia, which contain thousands of uredinospores.

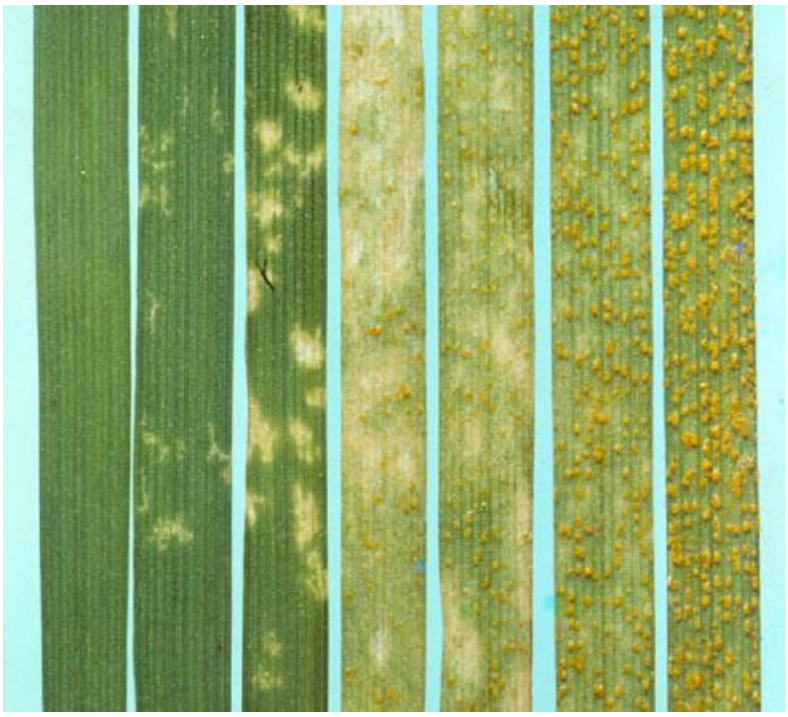


FIGURE 1: SYMPTOMS OF YELLOW RUST ON WHEAT LEAF

The continuous streaks of uredia and necrosis do not form at the juvenile stage, but after the plant begins to elongate the stem. According to the level of resistance of the plant to the pathogen and to the temperature, different amounts of chlorosis or necrosis (hypersensitive response or hypersensitive response, HR) may appear, with or without sporulation. Plants attacked by this pathogen are weakened as Pst uses water and nutrients (Chen, 2005).

Wheat yield losses due to Pst infections are generally the result of a limited number of kernels per ear of lower weight and quality (Chen et al., 2014).

Yellow rust has always been considered a recurring disease of temperate regions with low average temperatures and high humidity conditions, but even in the hottest and driest regions, where Pst had never occurred, there have been serious epidemics due to this pathogen (Hovmøllerer et al., 2010). The urediniospores are capable of germinating between 0 ° C and 20 ° C, but the optimal growth range for germination is between 7 ° C and 12 ° C. Sporulation can take place between 0 ° C and 28 °C but occurs more rapidly between 12 ° C and 20 ° C where the pathogen develops even more rapidly (Chen et al., 2013). The urediniospores are dicaryotic (n + n) and remain on the primary host (wheat) producing more than one secondary infection. This phase is responsible for Pst outbreaks on wheat. As temperatures begin to rise, Pst begins to produce teliospores; these are bicellular spores with very thick walls, which possess a diploid nucleus (2n) following karyogamy. Meiosis produces basiodospores (n) which germinate and infect Barberis spp. thus completing the cycle.

Pst has as an alternative host Berberis spp. (B. chinensis, B. holstii, B. koreana, B. vulgaris). On Berberis spp. Pst produces first pycniospores and then eciospores; the latter are those that will produce primary infection on durum wheat (Jin et al., 2010; Wang et al., 2015).

2.1.2 Life cycle and biology

To protect themselves from pathogens, plants have developed both passive and active defense systems. Pathogens that can overcome passive defenses (pre-existing physical and biochemical barriers) must deal with an active protection system composed of two levels (Dodds and Rathjen, 2010).

The first level of active defense is set up to detect the presence of PAMPs (pathogen-associeted molecular patterns), molecules that result from the breakdown of both pathogen and plant structures during infection (Sørensen et al., 2017). PAMPs are recognized by specific plant receptors

called PRRs (pattern recognition receptors). The stimulation of the PRRs leads to the activation of a cellular defense system called PTI (PRRs triggered immunity). It is important to remember that PTI has the ability to reject multiple types of microorganisms, pathogenic or not, according to the degree of conservation of the PAMP structure (eg: chitins, liposaccharides, etc.) in the different species, genera, families or classes (Zipfel, 2014). When the system is active, the plant is perceived as non-host by the intruding pathogen. Pathogens try to bypass the pathogen's active defense system to feed, killing or parasitizing the plant. Biotrophic parasitic pathogens (organisms, such as rusts belonging to the genus Puccinia) have evolved to be able to create compatibility with their host thanks to the release of effectors (small proteins) (Sørensen et al., 2017). In the second level of active defense, the plant's intercellular receptors can recognize these effectors and produce another type of response called ETI or effector triggered immunity. PTI and ETI are similar reactions, but ETI is stronger and faster, often also involving another type of defense induced called hypersensitive response (HR or hypersensitive response), which leads to the death of the attacked cells, thus preventing the pathogen from feeding and spreading. in host tissues (Dodds and Rathjen, 2010). When even this type of defense is bypassed, the plant is defenseless, and the pathogen can feed and complete its life cycle on the host.

2.1.3 Main Pst resistance genes

Long-term disease control depends on the characterization and use of different genetic resistance resources (Wu et al., 2016). As discussed extensively in the previous chapter, the less specific nature of APR and HTAP genes makes them preferable to all-stage resistance due to rapid virulence changes in pathogenic populations (mutations and migrations of new races). With the sole contribution of HATP and other types of quantitative resistance, many cultivars and accessions may not have the same effectiveness in fighting yellow rust if introduced in an environment other than that in which they were selected or evolved, due to sensitivity. HTAP resistance to environmental variations between different wheat growth areas (Chen et al. 2013). Pyramiding (i.e., obtaining a favorable combination of genes in the final ideotype) genes or QTLs from different resources can lead to high levels of lasting resistance. This goal can be achieved with the sole use of phenotypic selection, but the use of MAS (Marked Assisted Selection) is a valid tool that facilitates and speeds up the work of the breeder (Zheng et al. 2017). It is important to remember that combining genes with each other HTAP and all-stage resistance is a method of achieving more complete resistance. When the all-

stage resistance is active, the cultivars that possess both the types described above, will be completely protected by a specific breed of Pst; following the overcoming of race-specific resistance by new pathotypes, the plant will be able to remain resistant thanks to HTAP, albeit to a lesser extent (Chen et al 2013).

The rust resistance genes of wheat, both class R and APR, are represented by abbreviations: Lr for leaf or brown rust, Sr for black rust (Steam or black rust) and Yr for yellow rust (Yellow or stripe rust), with no class distinctions between R and APR and with an increasing number based on new findings (Ellis et al. 2014).

The 2013-2014 Catalog of Gene Symbols For Wheat (McIntosh et al. 2013) includes 67 officially recognized Yr genes (Yr1-Yr67) and 42 temporarily designated Yr genes (Maccaferri et al. 2015).

2.2 Materials and methods

2.2.1 Plant materials

As plant materials, we used two representative panels for the genetic heritage of tetraploid wheat:

- Global Durum Panel (GDP): The GDP is a breeding dedicated tool developed with a *forward* approach. This panel contains approximately 1000 genotypes, representative of the genetic diversity in durum wheat. This panel contains mostly modern varieties and landraces, with a small selection of emmer and primitive tetraploids. In particular, we used a set composed of about 500 durum cultivars and 200 durum landraces. (Mazzucotelli et al., 2020)

Landrace group	N° accessions	Breeding group	N° lines
Turkey-Transcaucasian	29	Italy	80
Central Asia	18	Central Asia	14
Arabian Peninsula	9	France	39
Iberian Peninsula	21	South America	25
Central Europe	18	Spain	27
South Asia	6	Central Europe	25
Greece	16	South Mediterranean	53
Italy	34	Ethiopia	8
Ethiopia	26	North America	33
North Africa	47	ICARDA	110
Argentina	5	CIMMYT	46
Levant	46	Australia	12
North America	5		
Australia	2		

FIGURE 2: ORIGIN OF THE GENOTYPES IN THE GDP PANEL (MAZZUCOTELLI ET AL., 2020)

- Tetraploid Global Collection (TGC): Unlike the GDP, the TGC was developed using a *reverse* approach. This panel is composed of approximately 1800 genotypes, representative of all the genetic variability of tetraploid wheat, including the four main germplasm groups used in breeding and wheat domestication: wild emmer wheat, domesticated emmer wheat, durum wheat landraces and durum wheat cultivars (Maccaferri et al., 2019). Analysis on TGC are still ongoing and will be further released in the publication.

By representing the global germplasm of durum wheat, also including genotypes used in the domestication process, these two panels were especially assembled and studied in search of sources of disease resistance. Modern germplasm is often limited and requires new sources of genetic variability, which can be identified in these panels, as proposed by this experiment. To do this, both panels were genotyped with the wheat high-density Illumina *iSelect* 90K SNP assay to provide a common genotype framework.

2.2.2 The environments

This PhD project is part of a close collaboration with the CEREALMED project and with the seed company "Apsov sementi S.P.A", located in Northern Italy. Actually the focus will be on GDP panel, since analysis on TGC are still ongoing. Thanks to this close cooperation we were able to obtain data for GDP panel from 8 different environments and different years, mainly in the Mediterranean area:

- Grosseto, Italy (2019, 2020 and 2021)
- Baucina, Italy (2019)
- Foggia, Italy (2020)
- Marchouch, Morocco (2019)
- Izmir, Turkey (2019)
- Lari, Lebanon (2020)
- Beni-Suef, Egypt (2021)
- La Dulce, Argentina (2021)

The Mediterranean area is the main production area of durum wheat, hence the need to screen the aforementioned panels to evaluate the effects caused by diseases in different areas of the Mediterranean area. The multi-environment trial can highlight differences in responses to different pathogenic races, in such a way as to be able to identify more reliable / stronger loci of resistance to the main races involved in infections affecting the Mediterranean area. Argentina, on the other hand, is the only environment outside the Mediterranean area and the Country is an emerging producer that is increasing its production of durum wheat. La Dulce is a locality in the province of Buenos Aires near the sea, identifying it with characteristics like that of the area we are considering, with the possibility of highlighting resistance to different strains, potentially useful for other production areas. The GDP has been phenotyped in all these environments and in all these years, different speech for the TGC, which was instead phenotyped only in Grosseto, in all three years. This

is because the GDP has been widely used within the CEREALMED project and shared with all partners.

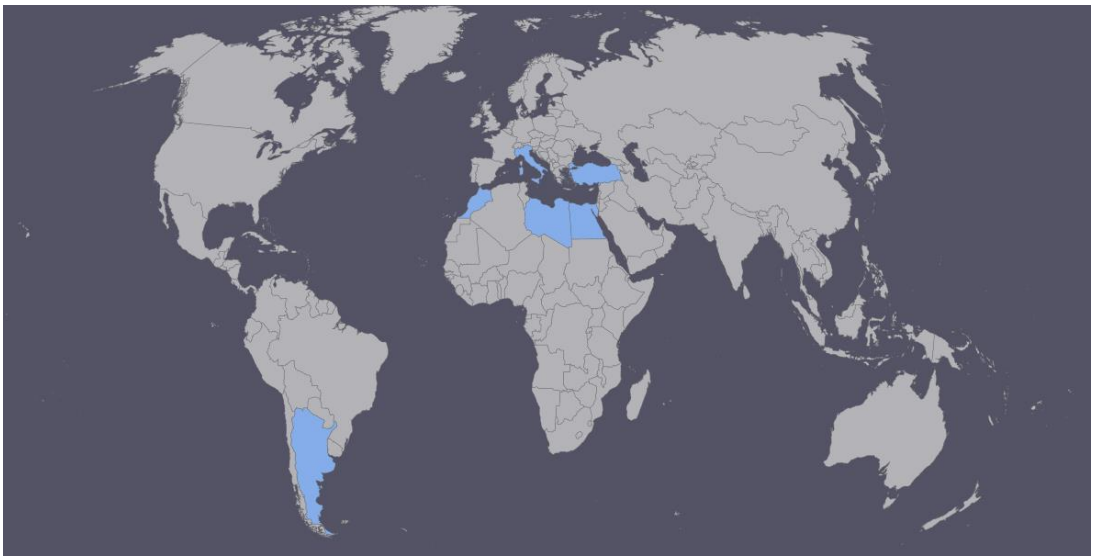


FIGURE 3: WORLD MAP HIGHLIGHTING THE STATES FROM WHICH THE DATA WAS TAKEN.

2.2.3 Field trials

The field design, in most of the experiments, consists of an augmented block design. In this experimental design the field is divided into blocks, and within each block common checks are included. This allows to assess many entries/accessions as unreplicated and still be able to carry out experimental tests to be carried out even when the number of entries is huge and the amount of available seed is limited.

FIGURE 4: EXAMPLE OF FIELD DESIGN WITH AUGMENTED BLOCK DESIGN. IN THIS CASE WE HAVE THREE CHECKS (CIRCLED) WHICH ARE REPEATED IN ALL BLOCKS.

Two environments (Foggia and Egypt) have a different experimental design, having the entire panel replicated completely twice.

2.2.4 Phenotyping

Both panels were phenotyped for yellow rust, taking infection type (IT) and severity (SEV) as data. The infection type represents the type of infection caused on the plant and was recorded on a scale from 0 to 9. The severity instead represents the percentage of leaf area affected within the parcel.

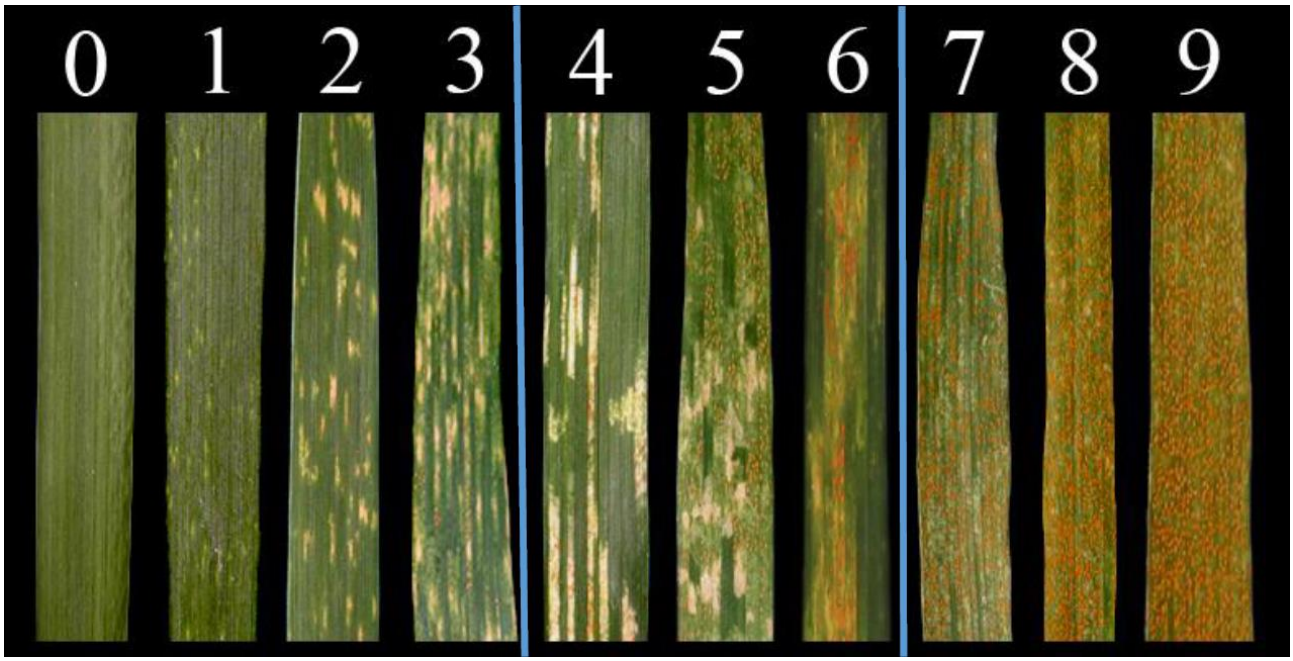


FIGURE 5: INFECTION TYPE SCALE RELATED TO YELLOW RUST INFECTION

The infection type scale is here described:

- **0 No visible sign or symptoms.**
- **1 Hypersensitive necrotic and/or chlorotic reactions, with typical flecks.**
- **2 Hypersensitive necrotic and/or chlorotic reactions with blotches or stripes; No sporulation.**
- **3 Hypersensitive necrotic and/or chlorotic reactions with blotches or stripes; Trace sporulation.**
- **4 Hypersensitive necrotic and/or chlorotic reactions with blotches or stripes; Light sporulation.**
- **5 Hypersensitive necrotic and/or chlorotic reactions with blotches or stripes; Intermediate sporulation.**
- **6 Hypersensitive necrotic and/or chlorotic reactions with blotches or stripes; Moderate sporulation.**
- **7 Hypersensitive necrotic or chlorotic reactions with blotches or stripes; Abundant sporulation.**
- **8 Chlorosis behind sporulation areas, abundant sporulation.**
- **9 No necrosis, neither chlorosis; intense sporulation (full compatible reaction).**

If the plant shows a reaction between 0 to 3 (highlighted in green), it is considered resistant. 4 to 6 (highlighted in orange) partially resistant and 7 to 9 susceptible (highlighted in red).

In most environments, data was taken for IT and SEV on multiple dates, tending to be 7 days apart to create solid data and see disease progression over time. To measure the course over time, the Under Disease Progress Curve Area (AUDPC) was measured. AUDPC is a useful quantitative

summary of disease intensity over time, for comparison across years, locations, or management tactics (American Phytopathological Society). AUDPC is calculated as follows:

$$\boldsymbol{AUDPC} = \sum_{\boldsymbol{i=1}}^{\boldsymbol{n-1}} \frac{\boldsymbol{y_i + y_{i+1}}}{\boldsymbol{2}} * (\boldsymbol{t_{i+1} - t_1})$$

Where *y* are the values of the phenotypic data, and *t* is the time between one score and another. The AUDPC has also been relativized, using the min-max rescaling as follows, to calculate the Relative AUDPC (RAUDPC).

$$x' = \frac{x - \min(x)}{\max(x) - \min(x)}$$

x represents the value to be relativized, taking into consideration the maximum and minimum value of the scoring.

FIGURE 6: YELLOW RUST SYMPTOMS ON WHEAT LEAVES

In addition to the data of the disease, data concerning the heading date (HD) and the plant height (Ph) were recorded in the various environments. HD was noted in days, starting with the sowing date. The Ph is instead expressed in cm.

2.2.5 Statistical analysis

The statistical phenotypic analysis was performed using the software R and RStudio (RStudio Team, 2020). The heritability for each trait was calculated using the package *repeatability* in Rstudio.
Best linear unbiased estimations (BLUEs) were obtained using the R package *lme4* by mixed model analysis (lmer). The parameters introduced in the model were different based single environments and cluster of environments
The variables included in the model for BLUEs extraction across the cluster of environments trials were the following:

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

Based on the following model, interaction with environment (E), across genotype and environment (GxE) were evaluated and considered during standardization of phenotypic values obtaining the BLUEs. The significant interactions between E and GxE were obtained via ANOVA analysis using the significant p-value threshold of 0.05.
For the single environments, the same analysis was performed with the following model to obtain the BLUEs:

B. **~ Genotype + Block (or Rep when we have the replicas)**

Imputation and LD decay:

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

- **1 – (MAF)2 – ((1-MAF)2**

PIC measures the ability of a marker to detect polymorphisms and therefore has enormous importance in selecting markers for genetic studies (Serrote et al., 2020).

An UNIBO internally developed R script was used to filter the HapMap file based on the following parameters: Minor Allele Frequency (MAF) greater than 0.01, SNPs missing call greater than 0.3, samples with a missing rate above 50%. After filtering, the final HapMap file included 23423 SNPs. The genotyping dataset was then imputed using Beagle v5.4 (Browning et al., 2021) to assign A/B variants to miss SNPs based on their position and closer SNPs (Beagle 5.4 uses a linkage disequilibrium-based algorithm).

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

2.2.6 Pruning and Population structure analysis

PLINK software (Chang et al., 2015) was used for the pruning step, removing redundant SNPs in the HapMap file and generating three output files for the different r^2 thresholds (0.3, 0.5, and 0.8).

The pruned output files were used for population structure analysis using the model-based likelihood method ADMIXTURE. The method was optimized using 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. To detect the best number of subpopulations to be analyzed, the cross-validated error rate, delta cv error, minimum group size, maximum admixed lines in a group, and admixed lines percentage were considered. 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 used to convert the imputed HapMap file into a distance matrix and thus obtain the kinship data frame by converting the values in genetic similarities. Heatmap and ward clustering (Ward.D2 algorithm) were performed on the kinship matrix using the R (R Core team, 2020) packages pheatmap v1.0.12 and dendextend v1.15.2.

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

2.2.7 GWAS analysis

GWAS analysis was performed using the R package GAPIT3 (Wang and Zhang, 2021) with few edits to the pipeline.

Basically, the final imputed hapmap was filtered in separate files for GDP, and separate kinship ad population structure analysis were performed for the dataset.

The permutation threshold was compared to the Bonferroni adjusted threshold, obtained by dividing the significant p value of 0.05 with the number of pruned markers at a r^2 threshold of 0.8 and calculating the negative logarithm in base 10. In average, for the different phenotypes, the Bonferroni threshold calculated via the permutation steps ranged between 4 and 5. This means that, the considered peaks above the threshold had an increased probability of 10^4-10^5 to associated with phenotypic variance.

The GWAS analysis was performed with GAPIT3 using the following methods: GLM, MLM, MLMM, FarmCPU and Blink. The edited GAPIT3 pipeline was divided in two step analysis: the first step included the kinship matrix derived from TASSEL, the second included both the kinship matrix and the Q file derived from the population structure analysis with ADMIXTURE, selecting the correct number of Ks based on the cross-validated error rate, Delta cv error, minimum group size, maximum admixed lines in a group, and admixed lines percentage were considered.

The first analysis included the following edited models: GLM (naïve), MLM + K, MLMM + K, FarmCPU and Blink.

The second analysis included the following models, where Q file from ADMIXTURE analysis were used as covariates: GLM + Q, MLM+ Q+K, MLMM Q+K, FarmCPU, Blink.

As for both pipelines, the number of dimensions for PCA was set to 0 and model selection to false.

The final Manhattan plots and raw data were merged in unique file including all the models, for each trait.

2.2.8 Drawing of KASP® primers

This work is in the context of a development of KASP® that have been validated through data collected in Grosseto, so I will report some details and results as they are closely related to this thesis work.

For this work, genotypes were selected from the GDP collection aswell, identified based on genotyping results using the 90K SNP array at loci on chromosome 1B. Once the lines to be assayed

were identified, KASP® primers were drawn for the most discriminating. In chromosome 1B, 5 adjacent SNP markers were identified in the genetic and physical map of durum wheat, which in the GDP panel are highly associated with resistance to rust yellow rust.

2.2.9 Procedure for transformation of SNP array polymorphisms into single assays of the KASP®-type assays for fluorescent PCR diagnostics

1) Sequence identification (*query*) in the reference genome. Once the following were identified the discriminatory blocks and their location, the marker sequence was identified in the reference genome, in this case the genome of the Svevo variety, in the *BLAST* of *Graingenes* (https://wheat.pw.usda.gov/cgi-bin/seqserve/blast_wheat.cgi). This site allows, by entering the sequence of the SNP, to go and locate it in the genome: it locates the largest portion because of that small sequence in FASTA format. For example, the marker IWB858 is located at chromosome 7B with the SNP [A/G]: you took the sequence and pasted it on *BLAST*: *QUERY* is the copied sequence. This is aligned with the *SUBJECT* which is the reference genome and sequences are returned that are present in the *SUBJECT* that are very similar to the QUERY (the site gives various options of the E value, expectation value, which the lower the better, because it restitutes an estimate of the sequences identical to the query that might be found by chance in the genome queried. Another parameter is identity, which must be very high, possibly close to 100%).
2) Once the sequence was identified, it was entered into the *Graingenes* Browser, typing the number of the chromosome and the exact sequences, saving the found sequence, through save track date. (100 bases to the right of the SNP and 100 to the left).
3) We proceeded with multiple alignment: a new FASTA was created and inserted the sequence of the SNP through a multiple MAFFT (online version) algorithm. https://mafft.cbrc.jp/alignment/server/
4) We visualized the alignment through the MEGA (molecular evolutionary genetics analisys https://www.megasoftware.net/mega4/) software: the software aligns the bases and displays them.

To draw the primers, it was decided to use Primer3plus software. Primer3plus (http://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi), "primer check" task, by entering the chosen sequence, based on precise parameters-which will be discussed in the next

chapter-assesses the suitability of primers for the Kasp reaction. In the sequence aligned in MEGA, the SNP (the exact position and number of bases) was displayed. First the allele-specific primers were drawn: two forwards (to the left of the mutation) and the two reverse (right), one terminating with one version of the SNP and the other with the other known allelic version. After that, at least three common primers were drawn from the opposite side: for the specific-forward the common is reverse while for the specific-reverse the common will be forward.

2.3 Results

The first part of the results will focus on the descriptive statistics for each environment, also showing the various distribution histograms for the most important characters of this analysis, IT and SEV of the last and most significant score. For each environment we find descriptive statistics data, including heritability (h^2), ANOVA results and distribution graphs for IT and SEV of the blues of the respective phenotypes.

2.3.1 La Dulce, Argentina

The analysis of the descriptive statistics (Tab. 1) shows very high heritability values with values of 0.96 for HD and Ph, while the infection type and severity values stand at 0.81.

	HD	Ph	YR_IT	YR_SEV
min	89.00	64.78	-0.01	-0.61
max	121.00	168.78	9.00	89.39
range	32.00	104.00	9.02	90.00
median	104.00	92.78	3.99	10.55
mean	103.67	100.43	4.22	16.48
SE.mean	0.22	0.68	0.07	0.47
var	38.35	362.55	3.40	174.52
std.dev	6.19	19.04	1.84	13.21
coef.var	0.06	0.19	0.44	0.80
h^2	0.96	0.96	0.81	0.81

TABLE 1: DESCRIPTIVE STATISTIC TABLE OF ARGENTINA EXPERIMENT, FOR HEADING DATE (HD), PLANT HEIGHT (PH), INFECTION TYPE AND SEVERITY

The ANOVA (Tab. 2) shows significant differences between genotypes for all characters, as far as the plant height is concerned, here we also have an interaction with the blocks.

Trait	Variables	Sum Sq	Df	F value	Pr(>F)	
HD	**Genotype**	24899.5	774	21.3797	<2e-16	***
HD	**Block**	2.7	3	0.5884	0.625	
Plant height	**Genotype**	228118	773	24.5536	<0.00000000000000022	***
Plant height	**Block**	155	3	4.2917	0.008385	**
YR_IT	**Genotype**	2756.81	774	5.0808	7.72E-12	***
YR_IT	**Block**	2.73	3	1.2999	0.2832	
YR_SEV	**Genotype**	135530	774	5.2621	3.41E-12	***
YR_SEV	**Block**	182	3	1.8278	0.1521	

TABLE 2: ANOVA RESULTS FOR ARGENTINA EXPERIMENT FOR ALL TRAITS CONSIDERED, *USING GENOTYPE AND BLOCK AS VARIABLES*

The distribution histograms show a non-normal trend but flattened towards resistance values (Fig. 7) for the SEV, while for the infection type the distribution is almost normal (Fig. 8).

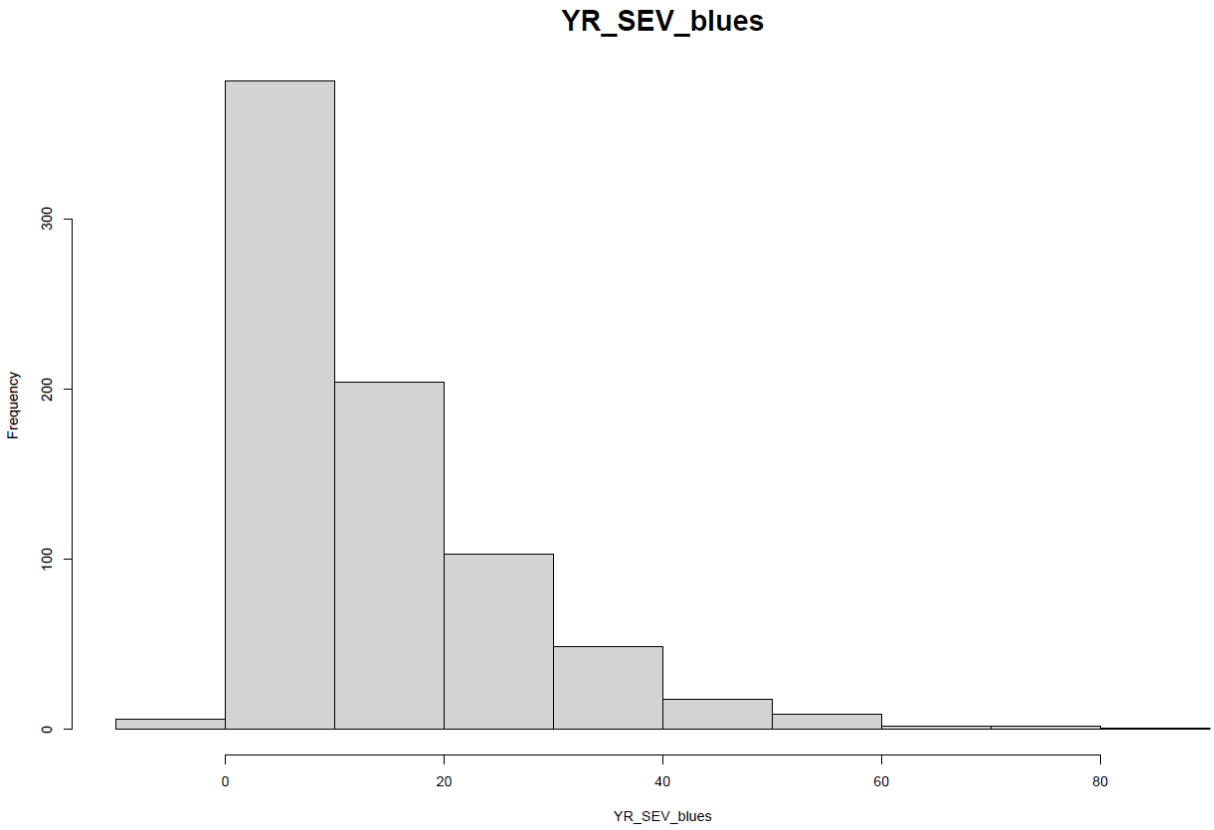


FIGURE 7: DISTRIBUTION HISTOGRAM FOR BLUES OBTAINED FOR SEVERITY IN ARGENTINA

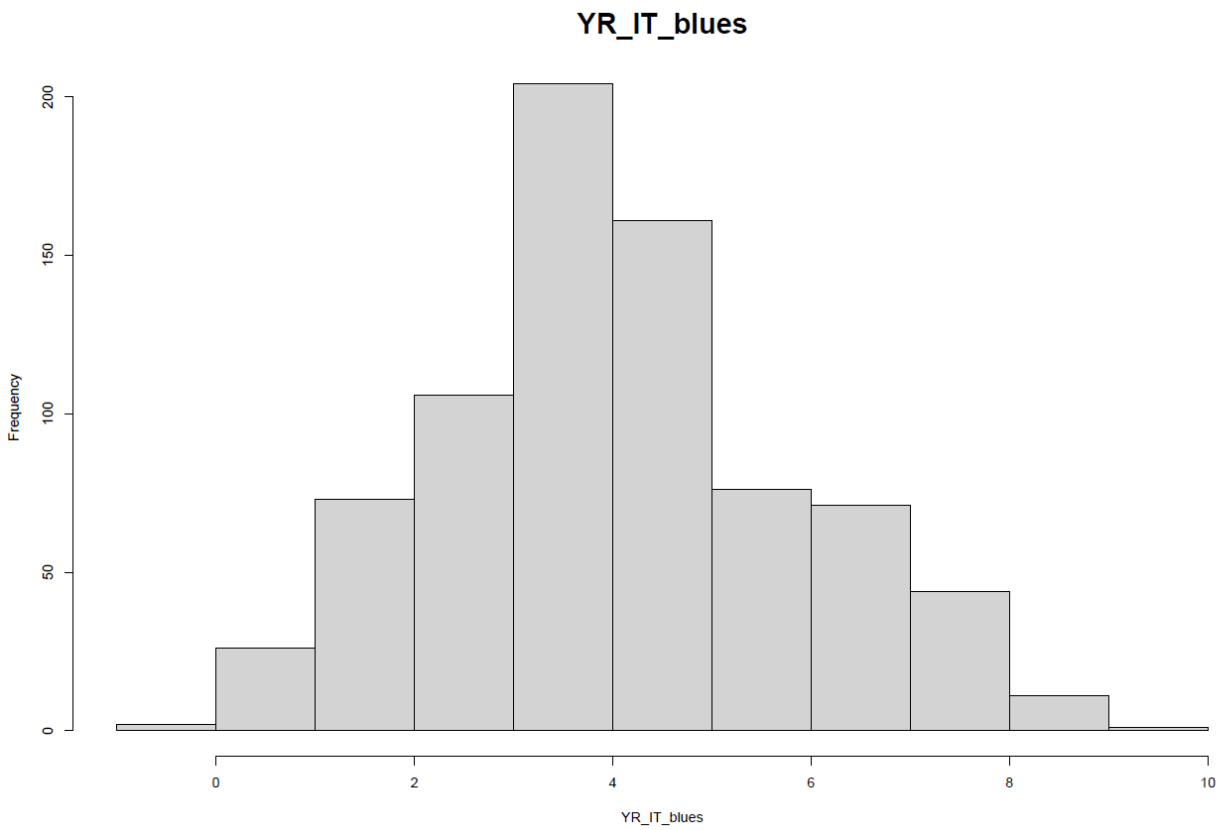


FIGURE 8: DISTRIBUTION HISTOGRAM FOR BLUES OBTAINED FOR INFECTION TYPE IN ARGENTINA

2.3.2 Baucina, Italy

In Tab. 3 we can see the descriptive statistic and the related heritability, that is pretty high for the traits of interest. h^2 is 0.89 for HD, and 0.78 for the infection type. The higher value (0.90) is for the SEV.

	HD	YR_IT	YR_SEV
min	129.90	0.99	5.00
max	141.14	9.02	100.00
range	11.24	8.03	95.00
median	133.98	5.99	60.00
mean	134.03	5.57	55.86
SE.mean	0.09	0.07	0.89
var	5.91	3.83	602.40
std.dev	2.43	1.96	24.54
coef.var	0.02	0.35	0.44
h^2	0.89	0.78	0.90

TABLE 3: DESCRIPTIVE STATISTIC AND HERITABILITY FOR THE MAIN TRAITS SCORED IN BAUCINA

Regarding the ANOVA (Tab. 4) we only have significance between genotypes.

Trait	Variables	Sum Sq	Df	F value	Pr(>F)	
HD	**Genotype**	4367.8	738	9.0391	1.17E-06	***
HD	**Block**	2.7	3	1.3818	0.2804	
YR_IT	**Genotype**	2719.19	766	4.1851	0.0004216	***
YR_IT	**Block**	2.67	3	1.0491	0.3949885	
YR_SEV	**Genotype**	413215	766	8.483	1.95E-06	***
YR_SEV	**Block**	174	3	0.9126	0.4545	

TABLE 4: ANOVA RESULTS FOR BAUCINA EXPERIMENT FOR ALL TRAITS CONSIDERED, USING GENOTYPE AND BLOCK AS VARIABLES

The distribution histograms show a normal trend for both the traits, infection type (Fig. 9) and SEV (Fig. 10)

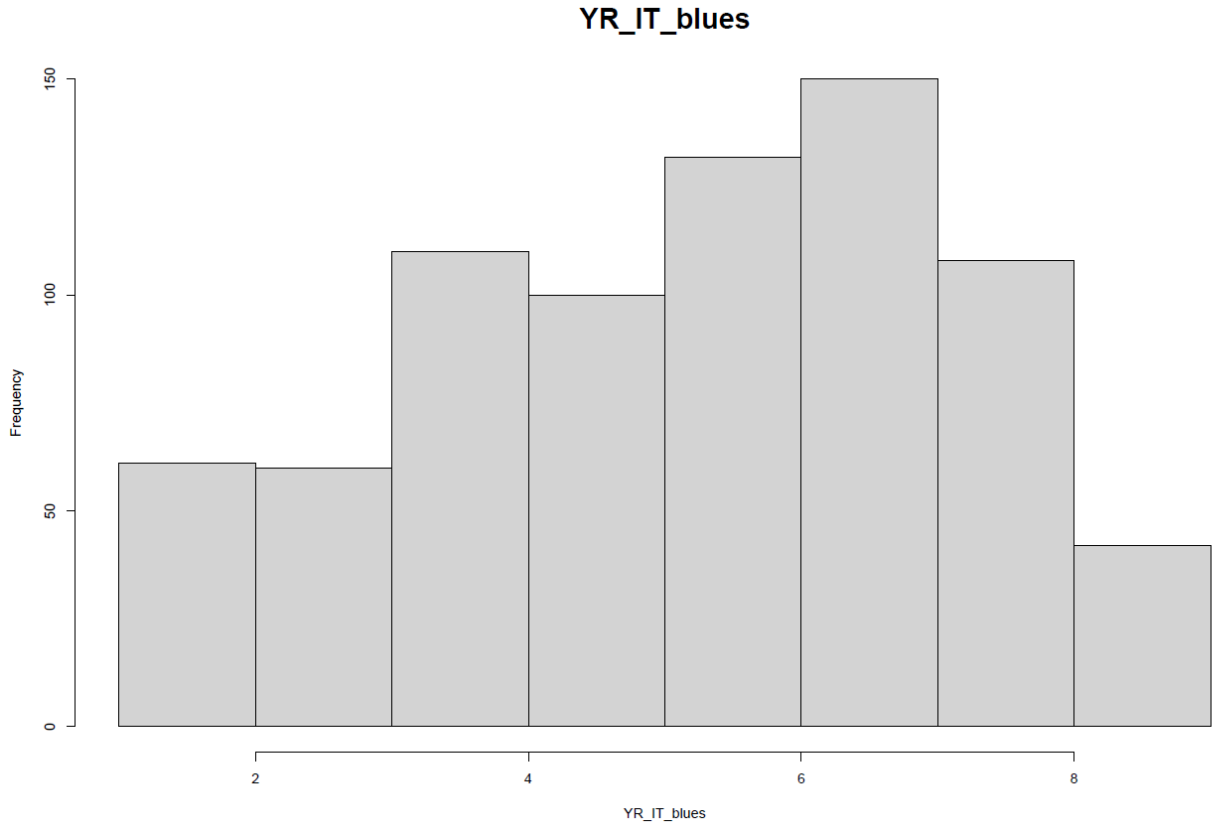


FIGURE 9: DISTRIBUTION HISTOGRAM FOR BLUES OBTAINED FOR INFECTION TYPE IN BAUCINA

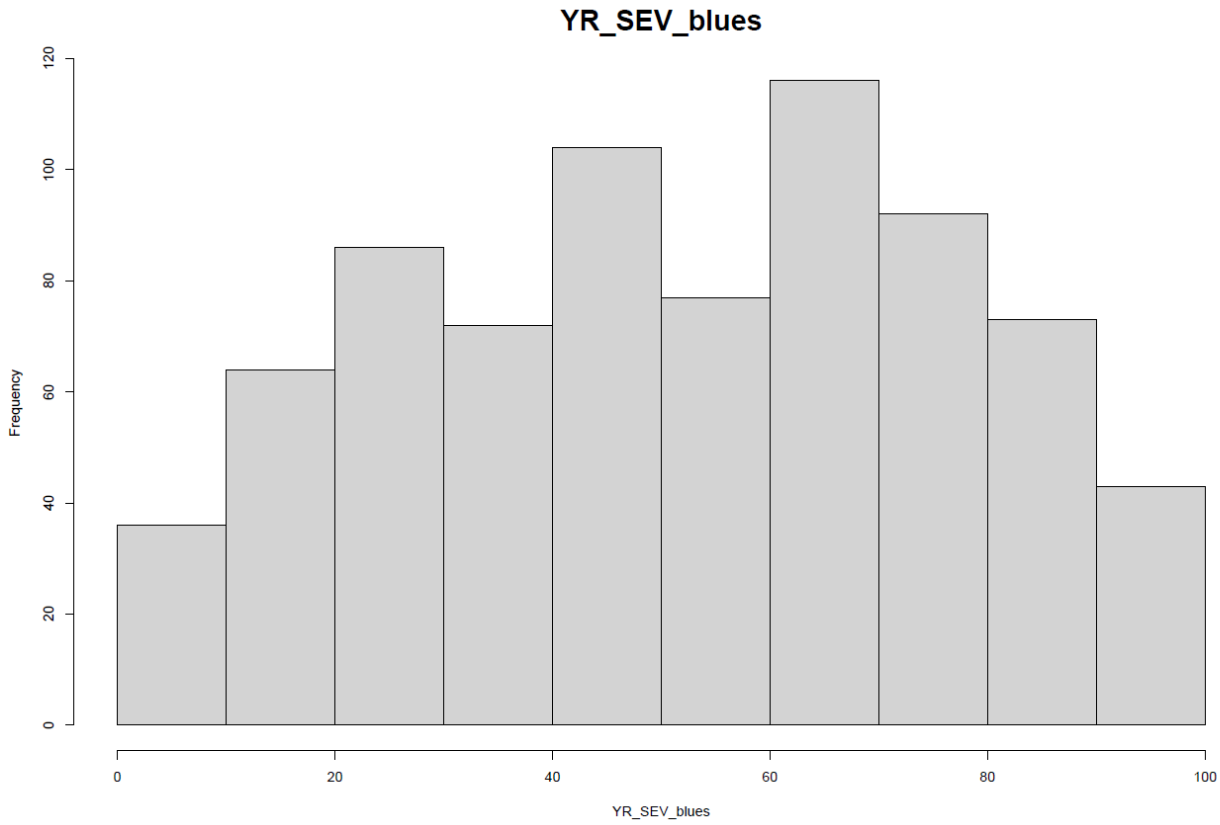


FIGURE 10: DISTRIBUTION HISTOGRAM FOR BLUES OBTAINED FOR SEVERITY IN BAUCINA

2.3.3 Beni-Suef, Egypt

Egypt traits analyzed show a high heritability with HD at 0.99, infection type at 0.93 and SEV at 0.92, showing a solid factor (Tab. 5)

	HD	YR_IT	YR_SEV
min	70	0	0
max	110.5	9	95
range	40.5	9	95
median	89.5	0	0
mean	91.2	2.39	15.99
SE.mean	0.41	0.19	1.47
var	47.78	10.81	611.5
std.dev	6.91	3.29	24.73
coef.var	0.08	1.38	1.55
h^2	0.99	0.93	0.92

TABLE 5: DESCRIPTIVE STATISTIC AND HERITABILITY FOR THE MAIN TRAITS SCORED IN EGYPT

The ANOVA shows high differences for all traits on Genotype variable, while Rep is significative only for HD trait.

Trait	Variables	Sum Sq	Df	F value	Pr(>F)	
HD	**Genotype**	26754.7	280	450.45	< 0.00000000000000022	***
HD	**Rep**	72.6	1	342.27	< 0.00000000000000022	***
YR_IT	**Genotype**	6111.4	284	15.3299	<2e-16	***
YR_IT	**Rep**	0	1	0.0154	0.9013	
YR_SEV	**Genotype**	344929	283	12.7427	<2e-16	***
YR_SEV	**Rep**	164	1	1.7183	0.191	

TABLE 6: ANOVA RESULTS FOR EGYPT EXPERIMENT FOR ALL TRAITS CONSIDERED, USING GENOTYPE AND REP AS VARIABLES

Distribution histograms in this case have for both traits the same trend, showing low level of disease and a non-normal distribution (Fig. 11 and 12).

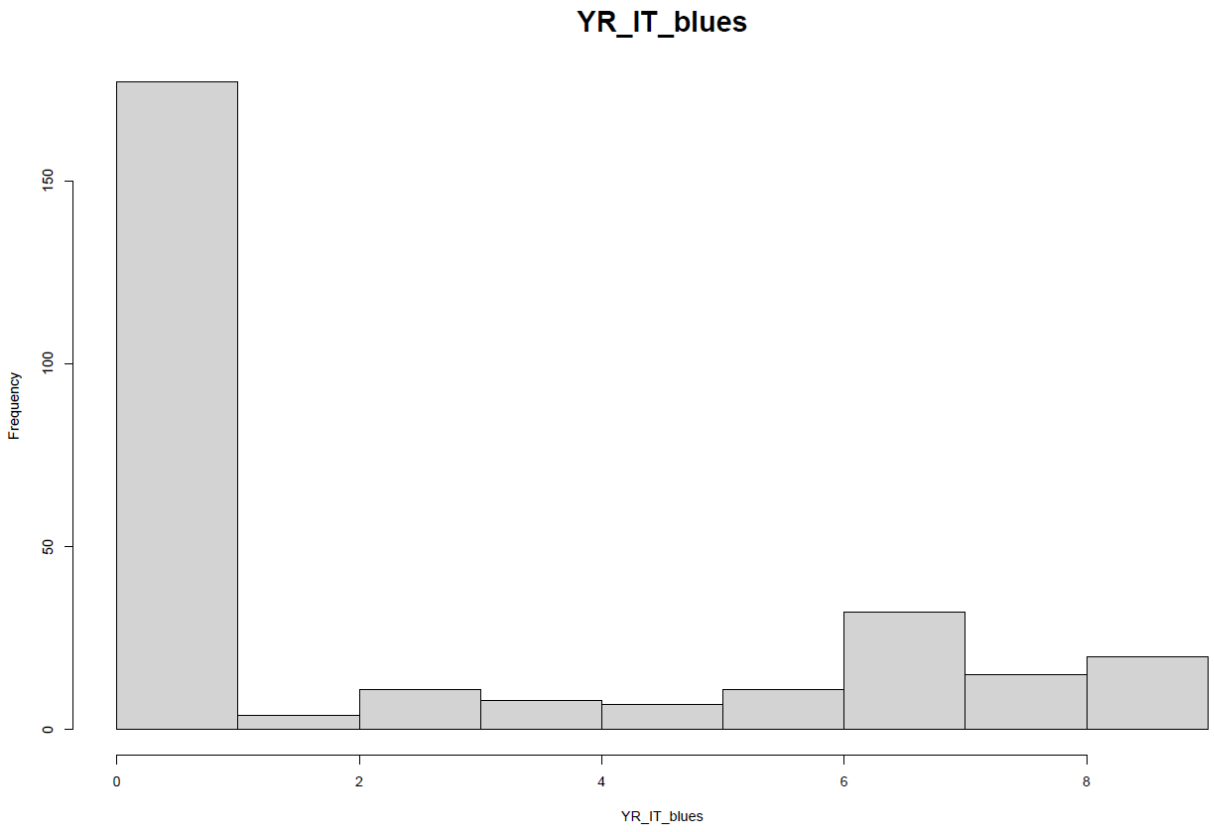


FIGURE 11: DISTRIBUTION HISTOGRAM FOR BLUES OBTAINED FOR INFECTION TYPE IN EGYPT

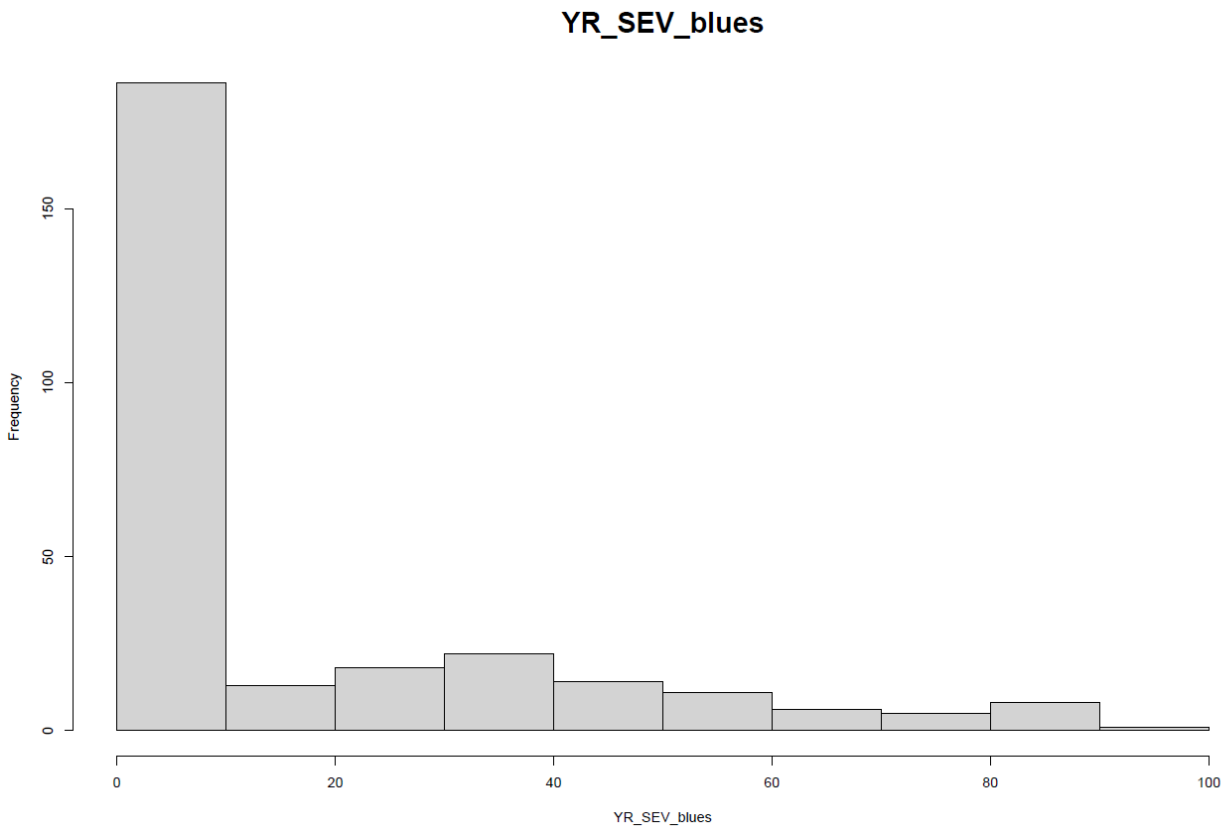


FIGURE 12: DISTRIBUTION HISTOGRAM FOR BLUES OBTAINED FOR SEVERITY IN EGYPT

2.3.4 Foggia, Italy

Foggia traits analyzed show a high heritability, even if a bit lower compared to the others, with HD as higher value, with 0.82. Infection type at 0.0.72 and SEV at 0.75. For this environment we also have data related to AUDPC and RAUDPC. AUDPC shows 0.72 for IT and 0.77 for SEV. RAUDPC h^2 values are the same, compared to AUDPC (Tab. 7).

	HD	YR_IT	YR_DS	AUDPC_IT	RAUDPC_IT	AUDPC_SEV	RAUDPC_SEV
min	141.76	1.50	15.00	10.50	0.07	105.00	0.07
max	159.24	7.14	75.00	52.50	0.93	525.00	0.93
range	17.47	5.64	60.00	42.00	0.86	420.00	0.86
median	150.00	3.00	45.00	17.50	0.21	315.00	0.50
mean	150.09	3.16	42.57	21.79	0.30	290.48	0.45
SE.mean	0.19	0.07	1.16	0.49	0.01	7.94	0.02
var	11.54	1.57	413.29	72.37	0.03	19244.41	0.08
std.dev	3.40	1.25	20.33	8.51	0.17	138.72	0.28
coef.var	0.02	0.40	0.48	0.39	0.58	0.48	0.63
h^2	0.82	0.72	0.75	0.72	0.72	0.77	0.77

TABLE 7: DESCRIPTIVE STATISTIC AND HERITABILITY FOR THE MAIN TRAITS SCORED IN FOGGIA

Regarding ANOVA we used Genotype and Rep as variables, showing a good significance into the analysis for all the traits except for the SEV related traits, where Rep was not significative (Tab. 8).

Trait	Variables	Sum Sq	Df	F value	Pr(>F)	
HD	**Genotype**	6678.1	303	5.8152	<0.00000000000000022	***
HD	**Rep**	67.6	1	17.8356	2.96E-05	***
YR_IT	**Genotype**	1128.6	304	3.7701	<0.00000000000000022	***
YR_IT	**Rep**	24.43	1	24.8115	9.24E-07	***
YR_SEV	**Genotype**	320418	304	4.075	<2e-16	***
YR_SEV	**Rep**	242	1	0.9342	0.3343	
AUDPC_IT	**Genotype**	54252	304	3.8324	<0.00000000000000022	***
AUDPC_IT	**Rep**	1551	1	33.3039	1.52E-08	***
RAUDPC_IT	**Genotype**	22.5955	304	3.8324	<0.00000000000000022	***
RAUDPC_IT	**Rep**	0.6459	1	33.3039	1.52E-08	***
AUDPC_SEV	**Genotype**	15378627	304	4.2841	<2e-16	***
AUDPC_SEV	**Rep**	21427	1	1.8146	0.1787	
RAUDPC_SEV	**Genotype**	64.051	304	4.2841	<2e-16	***
RAUDPC_SEV	**Rep**	0.089	1	1.8146	0.1787	

TABLE 8: ANOVA RESULTS FOR FOGGIA EXPERIMENT FOR ALL TRAITS CONSIDERED, USING GENOTYPE AND REP AS VARIABLES

Distribution histograms, for IT show low levels of disease, for SEV we have a bimodal distribution with high level of genotypes with low and high severity scores (Fig. 13 and 14).

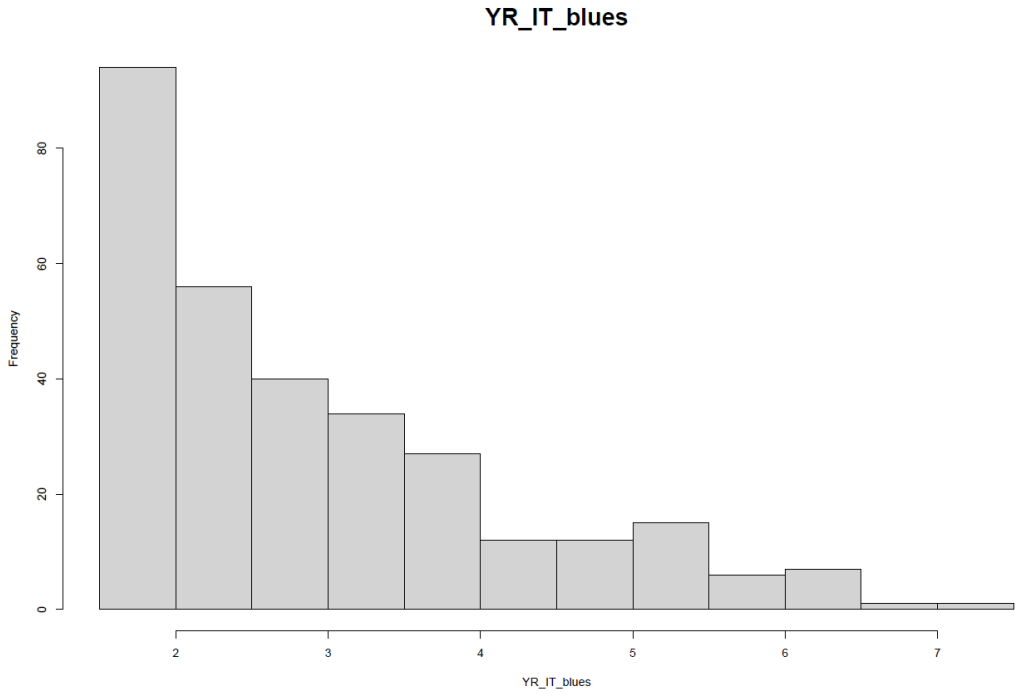


FIGURE 13: DISTRIBUTION HISTOGRAM FOR BLUES OBTAINED FOR INFECTION TYPE IN FOGGIA

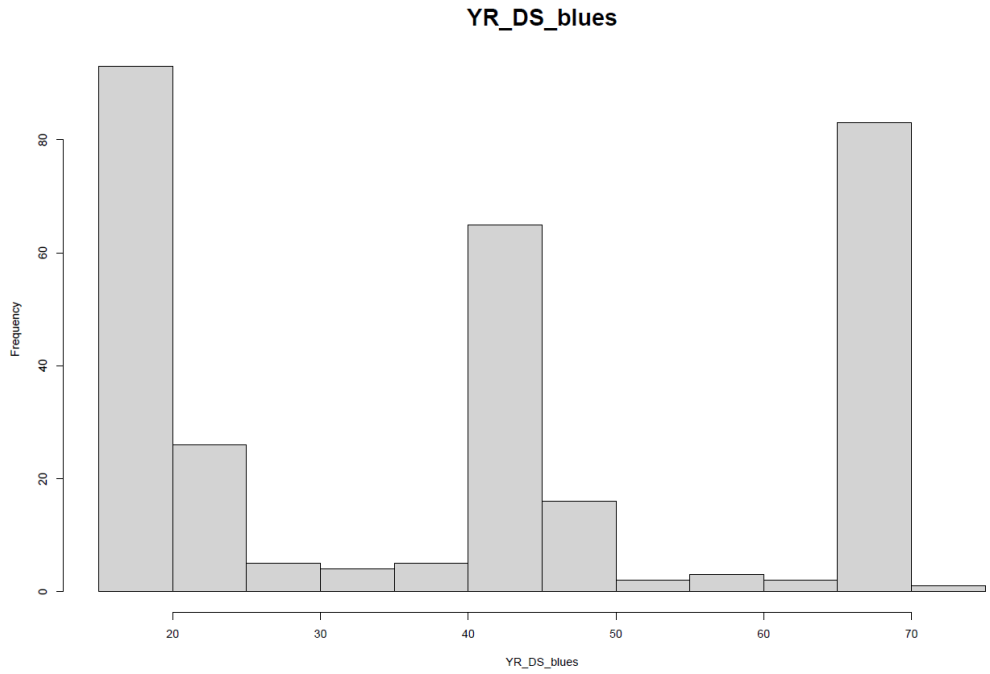


FIGURE 14: DISTRIBUTION HISTOGRAM FOR BLUES OBTAINED FOR SEVERITY IN FOGGIA

2.3.5 Grosseto 2019, Italy

For Grosseto 2019 we have data from AUDPC and RAUDPC of the traits taken into consideration. For this environment we have a lower heritability work referred to HD, 0.77. The highest heritability is the SEV (0.85) and the related AUDPC and RAUDPC (0.88 both). IT shows slightly lower heritability (0.80), with AUDPC and RAUDPC having both value of 0.87 (Tab. 9).

	HD	YR_IT	YR_SEV	AUDPC_IT	RAUDPC_IT	AUDPC_SEV	RAUDPC_SEV
min	148.00	0.85	7.11	4.00	0.02	86.62	-0.01
max	172.00	9.09	103.30	36.00	0.40	1124.76	1.02
range	24.00	8.24	96.20	32.00	0.37	1038.14	1.03
median	158.00	5.09	65.35	20.00	0.21	631.62	0.54
mean	158.96	5.37	62.88	20.37	0.21	625.62	0.53
SE.mean	0.34	0.09	1.56	0.33	0.00	14.39	0.01
var	36.91	2.58	815.29	36.13	0.00	69761.52	0.07
std.dev	6.08	1.61	28.55	6.01	0.07	264.12	0.27
coef.var	0.04	0.30	0.45	0.30	0.33	0.42	0.52
h^2	0.77	0.80	0.85	0.87	0.87	0.88	0.88

TABLE 9: DESCRIPTIVE STATISTIC AND HERITABILITY FOR THE MAIN TRAITS SCORED IN GROSSETO 2019

ANOVA showed significance by genotype for all traits, while the interaction between blocks was not significant, except slightly for SEV and AUDPC_SEV (Tab. 10).

Trait	Variables	Sum Sq	Df	F value	Pr(>F)	
HD	**Genotype**	13535.7	325	4.3109	7.74E-07	***
HD	**Block**	17.9	2	0.9289	0.4043	
YR_IT	**Genotype**	975.68	336	5.2335	6.12E-10	***
YR_IT	**Block**	2.42	2	2.1825	0.1241	
YR_SEV	**Genotype**	298015	336	7.1834	1.31E-12	***
YR_SEV	**Block**	801	2	3.2436	0.04791	*
AUDPC_IT	**Genotype**	13997.3	336	7.613	4.06E-13	***
AUDPC_IT	**Block**	8.1	2	0.7443	0.4806	
RAUDPC_IT	**Genotype**	1.89255	336	7.613	4.06E-13	***
RAUDPC_IT	**Block**	0.0011	2	0.7443	0.4806	
AUDPC_SEV	**Genotype**	25737407	336	8.9535	1.47E-14	***
AUDPC_SEV	**Block**	50272	2	2.9381	0.06276	.
RAUDPC_SEV	**Genotype**	27.648	336	8.7477	2.38E-14	***
RAUDPC_SEV	**Block**	0.038	2	2.022	0.1437	

TABLE 10: ANOVA RESULTS FOR GROSSETO 2019 EXPERIMENT FOR ALL TRAITS CONSIDERED, USING GENOTYPE AND BLOCK AS VARIABLES

The distribution graph for the IT shows a normal distribution with a very flattened curve in the center of the Gaussian (Fig. 15), as for the SEV we have a non-normal distribution, with the blues values very similar for most of the bands considered (Fig. 16).

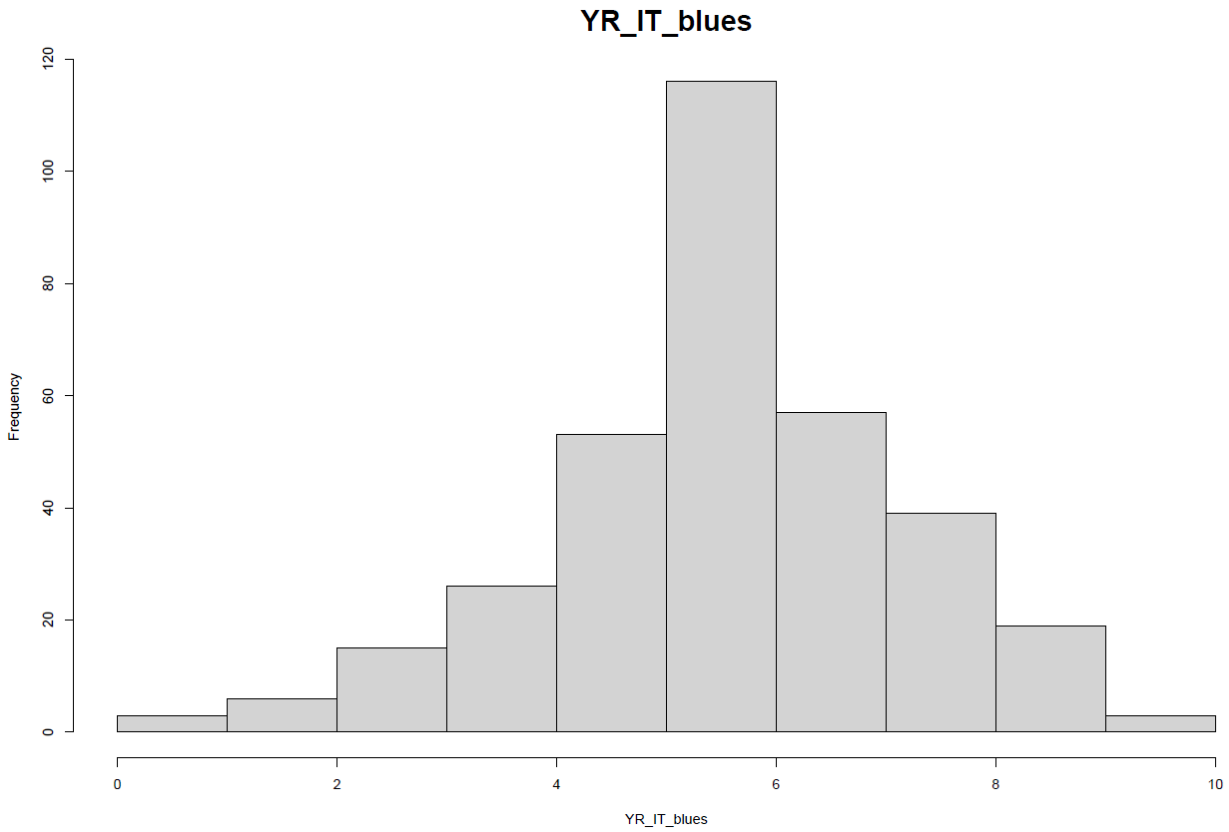


FIGURE 15: DISTRIBUTION HISTOGRAM FOR BLUES OBTAINED FOR INFECTION TYPE IN GROSSETO 2019

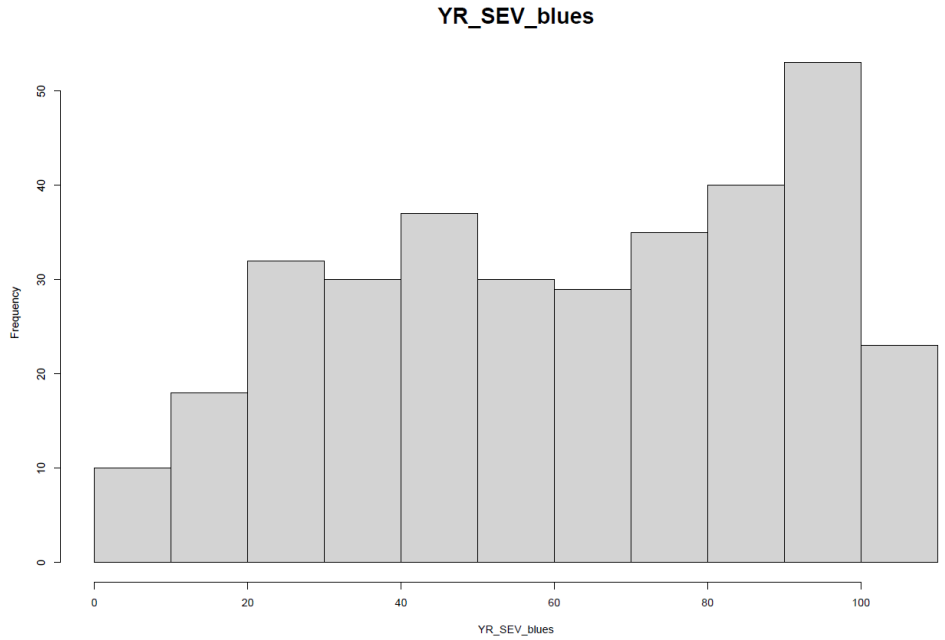


FIGURE 16: DISTRIBUTION HISTOGRAM FOR BLUES OBTAINED FOR SEVERITY IN GROSSETO 2019

2.3.6 Grosseto 2020, Italy

For Grosseto 2020 we have data from AUDPC and RAUDPC of the traits taken into consideration like for Grosseto 2019. HD heritability is higher compared to the previous year (0.88). SEV in this case has the lowest value of h^2 (0.59), with the related traits with a 0.60 value both. IT shows higher heritability (0.79), with AUDPC and RAUDPC having both value of 0.78 (Tab. 11).

	HD	YR_IT	YR_SEV	AUDPC_IT	RAUDPC_IT	AUDPC_SEV	RADUPC_SEV
min	118.91	-0.15	-2.72	-0.67	-0.01	-15.84	-0.03
max	130.72	8.62	100.21	42.91	1.00	499.27	1.00
range	11.81	8.77	102.93	43.58	1.01	515.11	1.03
median	122.87	2.71	10.21	12.77	0.25	51.61	0.10
mean	123.16	2.66	14.25	12.54	0.26	68.63	0.14
SE.mean	0.10	0.07	0.58	0.35	0.01	2.89	0.01
var	6.77	3.14	216.93	79.84	0.04	5395.22	0.02
std.dev	2.60	1.77	14.73	8.94	0.19	73.45	0.15
coef.var	0.02	0.67	1.03	0.71	0.74	1.07	1.07
h^2	0.88	0.79	0.59	0.78	0.78	0.60	0.60

TABLE 11: DESCRIPTIVE STATISTIC AND HERITABILITY FOR THE MAIN TRAITS SCORED IN GROSSETO 2020

Regarding the interactions of ANOVA, there are evident differences between genotypes for all the traits, while for the blocks, significance is highlighted for the traits of HD, SEV and related values.

Trait	Variables	Sum Sq	Df	F value	Pr(>F)	
HD	**Genotype**	6709.5	710	8.1646	<0.00000000000000022	***
HD	**Block**	30.3	10	2.6163	0.005211	**
YR_IT	**Genotype**	2678.11	675	4.6429	<2e-16	***
YR_IT	**Block**	11.66	10	1.3642	0.1997	
YR_SEV	**Genotype**	154672	662	2.4726	8.70E-13	***
YR_SEV	**Block**	1977	10	2.0918	0.02697	*
AUDPC_IT	**Genotype**	69350	675	4.6091	<2e-16	***
AUDPC_IT	**Block**	288	10	1.2918	0.2376	
RAUDPC_IT	**Genotype**	30.8376	675	4.5957	<2e-16	***
RAUDPC_IT	**Block**	0.1141	10	1.1481	0.3287	
AUDPC_SEV	**Genotype**	3833709	657	2.6429	2.77E-14	***
AUDPC_SEV	**Block**	56619	10	2.5644	0.006283	**
RAUDPC_SEV	**Genotype**	15.3348	657	2.6429	2.77E-14	***
RAUDPC_SEV	**Block**	0.2265	10	2.5644	0.006283	**

TABLE 12: ANOVA RESULTS FOR GROSSETO 2020 EXPERIMENT FOR ALL TRAITS CONSIDERED, USING GENOTYPE AND BLOCK AS VARIABLES

Also, in this case the IT distribution graph shows a normal trend (Fig. 17), while the severity one shows a trend of the curve squeezed towards low severity values, this can be attributed to the fact that the season was very dry which did not allow adequate development of the disease (Fig. 18).

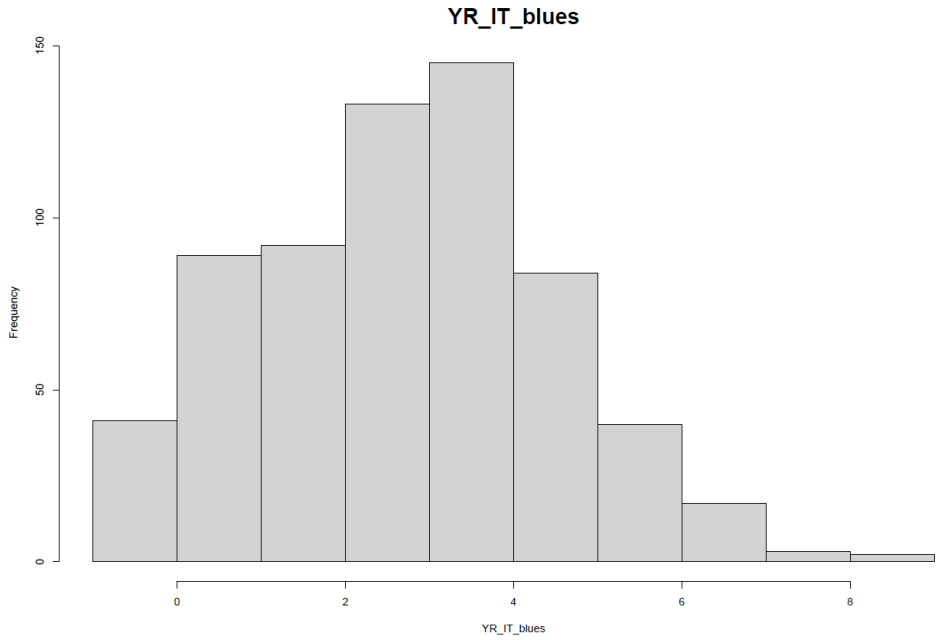


FIGURE 17: DISTRIBUTION HISTOGRAM FOR BLUES OBTAINED FOR INFECTION TYPE IN GROSSETO 2020

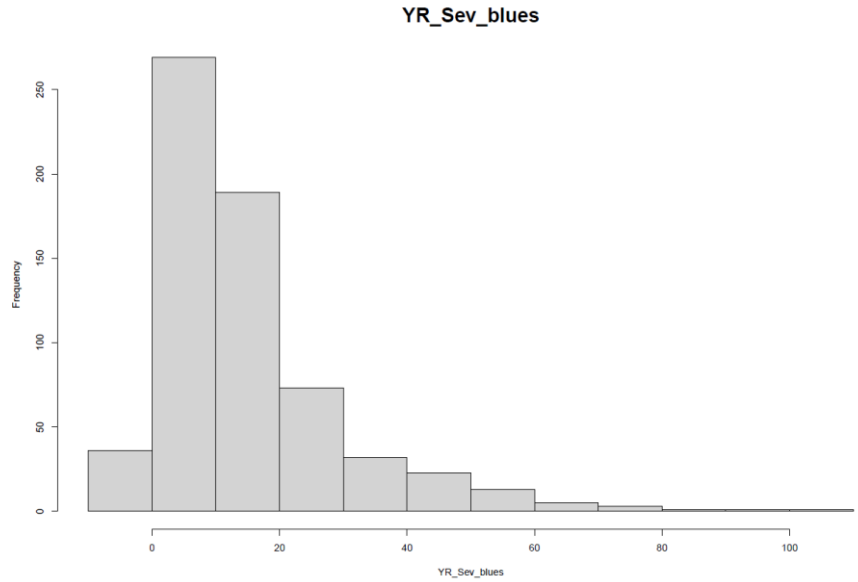


FIGURE 18: DISTRIBUTION HISTOGRAM FOR BLUES OBTAINED FOR SEVERITY IN GROSSETO 2020

2.3.7 Grosseto 2021, Italy

For the last year of Grosseto, the heritability values tend to be a bit more uniform with HD at 0.76, IT at 0.87 (AUDPC and RAUDPC at 0.84), while SEV has a value of 0.78, with AUDPC and RAUDPC at 0.76 (Tab. 13).

	HD	YR_IT	YR_SEV	AUDPC_IT	RAUDPC_IT	AUDPC_SEV	RAUDPC_SEV
min	130.73	0.00	-0.15	0.00	0.00	-3.35	0.00
max	147.03	9.00	100.97	117.00	1.00	1320.73	1.02
range	16.31	9.00	101.12	117.00	1.00	1324.08	1.02
median	141.67	6.00	57.31	78.00	0.67	670.73	0.52
mean	141.63	5.70	52.81	72.91	0.62	672.33	0.52
SE.mean	0.14	0.06	0.88	0.84	0.01	11.55	0.01
var	7.06	3.27	608.88	554.11	0.04	103995.48	0.06
std.dev	2.66	1.81	24.68	23.54	0.20	322.48	0.25
coef.var	0.02	0.32	0.47	0.32	0.32	0.48	0.48
h^2	0.76	0.87	0.78	0.84	0.84	0.76	0.76

TABLE 13: DESCRIPTIVE STATISTIC AND HERITABILITY FOR THE MAIN TRAITS SCORED IN GROSSETO 2021

ANOVA does not show particular significance beyond the differences between genotypes, there is significance between blocks only for the HD trait (Tab. 14).

Trait	Variables	Sum Sq	Df	F value	Pr(>F)	
HD	**Genotype**	3041.33	381	4.3597	<0.00000000000000022	***
HD	**Block**	61.19	9	3.7133	0.0003807	***
YR_IT	**Genotype**	2727	778	6.4593	<2e-16	***
YR_IT	**Block**	1.07	9	0.2184	0.9915	
YR_SEV	**Genotype**	539211	778	4.2615	<2e-16	***
YR_SEV	**Block**	1702	9	1.1629	0.3225	
AUDPC_IT	**Genotype**	456799	778	5.3339	<2e-16	***
AUDPC_IT	**Block**	474	9	0.4782	0.8877	
RAUDPC_IT	**Genotype**	33.37	778	5.3339	<2e-16	***
RAUDPC_IT	**Block**	0.035	9	0.4782	0.8877	
AUDPC_SEV	**Genotype**	92806465	778	4.0005	<2e-16	***
AUDPC_SEV	**Block**	370066	9	1.379	0.2018	
RAUDPC_SEV	**Genotype**	54.915	778	4.0005	<2e-16	***
RAUDPC_SEV	**Block**	0.219	9	1.379	0.2018	

TABLE 14: ANOVA RESULTS FOR GROSSETO 2021 EXPERIMENT FOR ALL TRAITS CONSIDERED, USING GENOTYPE AND BLOCK AS VARIABLES

The distribution of IT tends to normal even if it is more shifted towards high infection type values (Fig. 19), while the values of the SEV are almost normally distributed (Fig. 20).

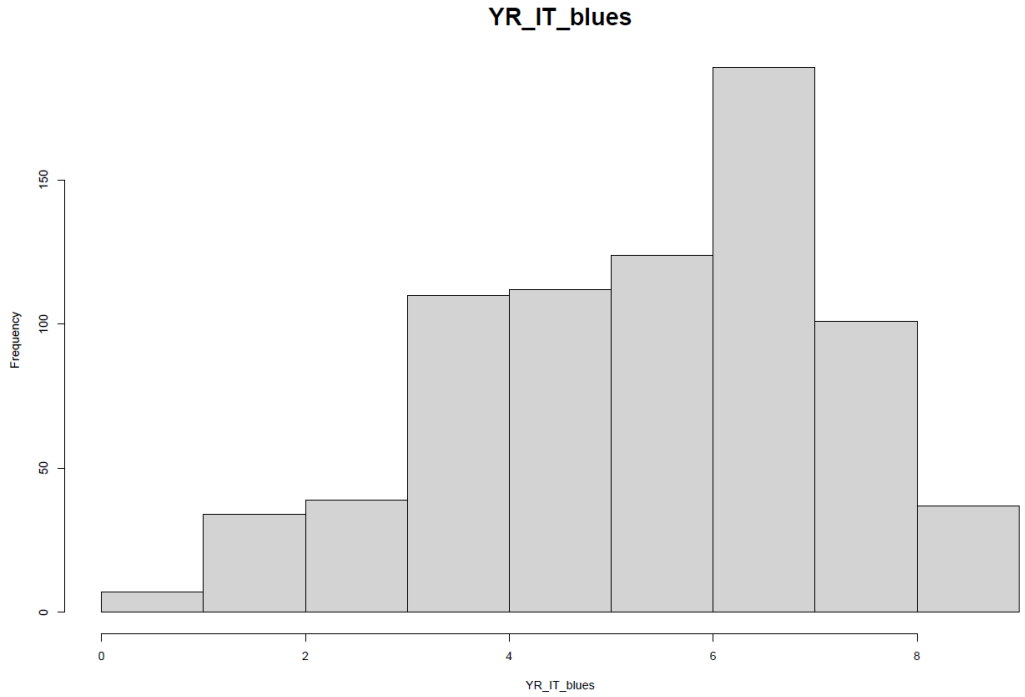


FIGURE 19: DISTRIBUTION HISTOGRAM FOR BLUES OBTAINED FOR INFECTION TYPE IN GROSSETO 2021

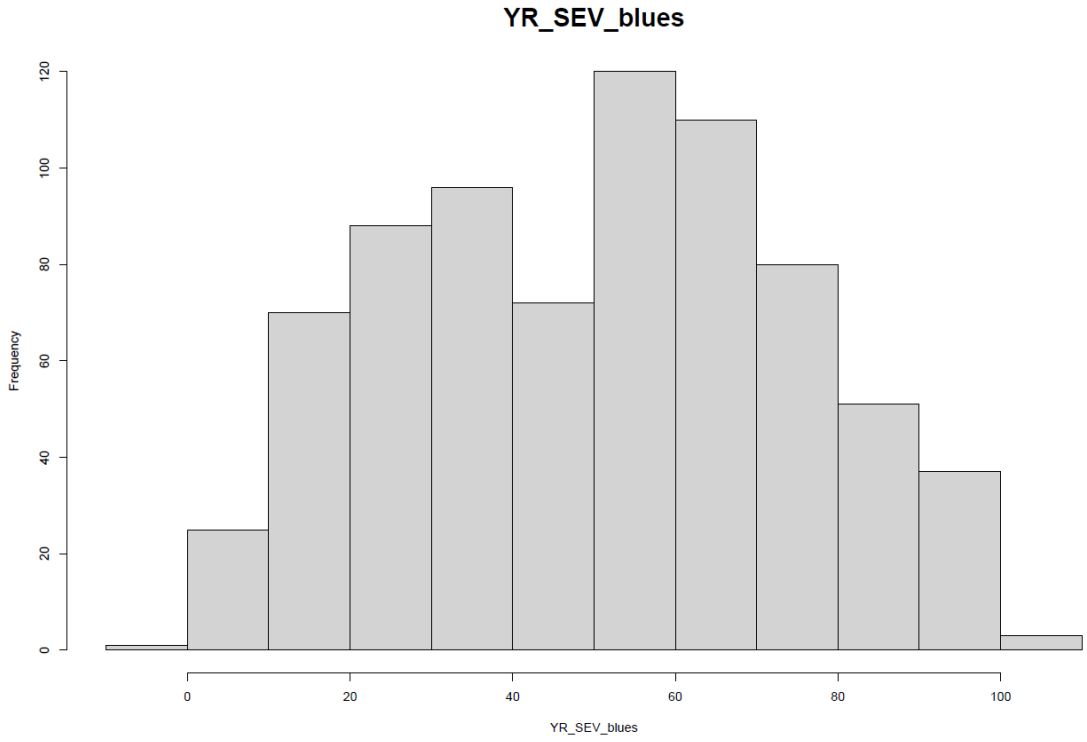


FIGURE 20: DISTRIBUTION HISTOGRAM FOR BLUES OBTAINED FOR SEVERITY IN GROSSETO 2021

2.3.8 LARI, Lebanon

For Lebanon there are only three traits, HD, IT and SEV. For these the heritability values are particularly high: HD 0.99, IT 0.91 and SEV, lower, 0.78 (Tab.15).

	HD	YR_IT	YR_SEV
min	136.00	0.00	0.00
max	165.00	9.00	100.00
range	29.00	9.00	100.00
median	150.00	5.00	30.00
mean	149.63	4.26	34.43
SE.mean	0.25	0.09	0.76
var	60.38	7.91	563.82
std.dev	7.77	2.81	23.74
coef.var	0.05	0.66	0.69
h^2	0.99	0.91	0.78

TABLE 15: DESCRIPTIVE STATISTIC AND HERITABILITY FOR THE MAIN TRAITS SCORED IN LEBANON

The ANOVA in this case shows significance only between genotypes (Tab. 16).

Trait	Variables	Sum Sq	Df	F value	Pr(>F)	
HD	**Genotype**	42737	982	81.7419	<2e-16	***
HD	**Block**	1	5	0.4383	0.8188	
YR_IT	**Genotype**	7484	981	9.2837	4.79E-12	***
YR_IT	**Block**	1.4	5	0.3448	0.8822	
YR_SEV	**Genotype**	541569	981	4.2379	6.94E-07	***
YR_SEV	**Block**	560	5	0.8604	0.5169	

TABLE 16: ANOVA RESULTS FOR LEBANON EXPERIMENT FOR ALL TRAITS CONSIDERED, USING GENOTYPE AND BLOCK AS VARIABLES

The distribution graphs show, for IT, an almost normal trend, with the lack of intermediate values (Fig. 21). As for the SEV, the trend is definitely not normal, with a flattening towards low values of SEV, showing a poor diffusion of the pathogen, probably due to the climatic conditions (Fig. 22)

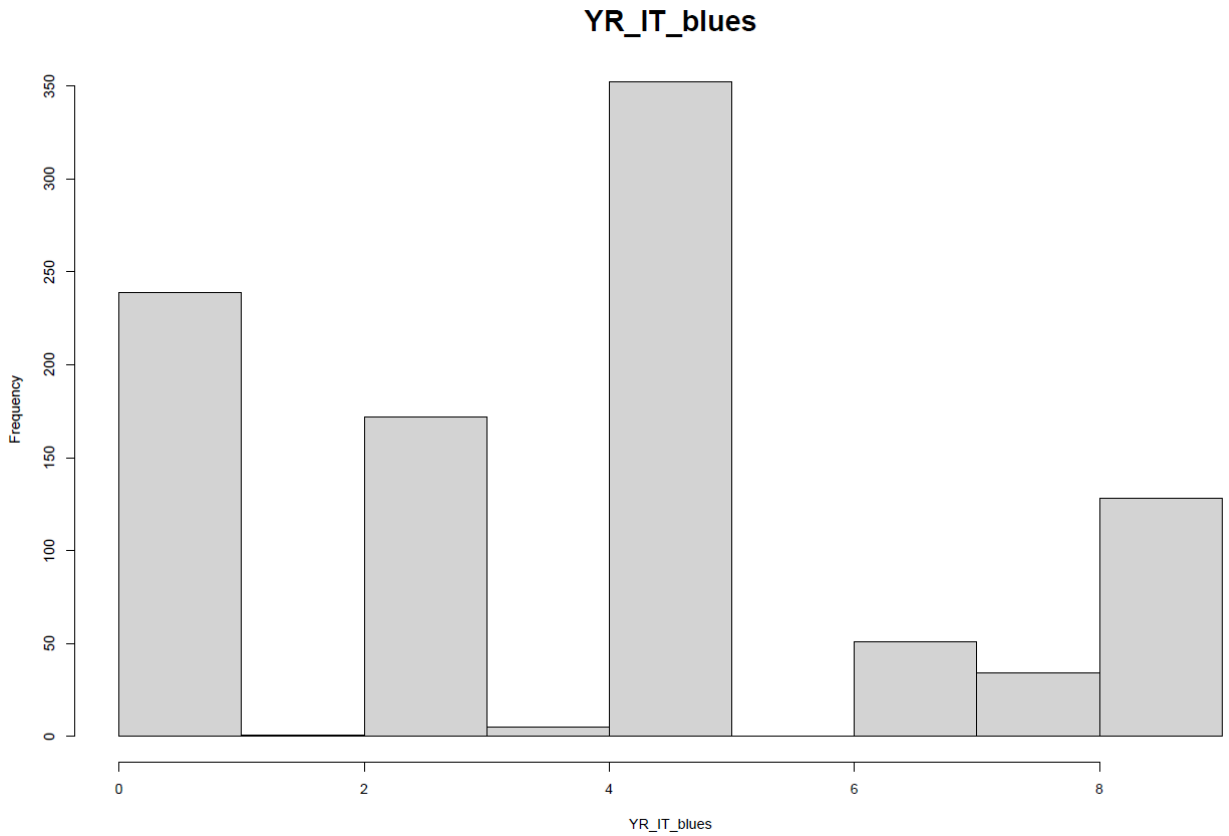


FIGURE 21: DISTRIBUTION HISTOGRAM FOR BLUES OBTAINED FOR INFECTION TYPE IN LEBANON

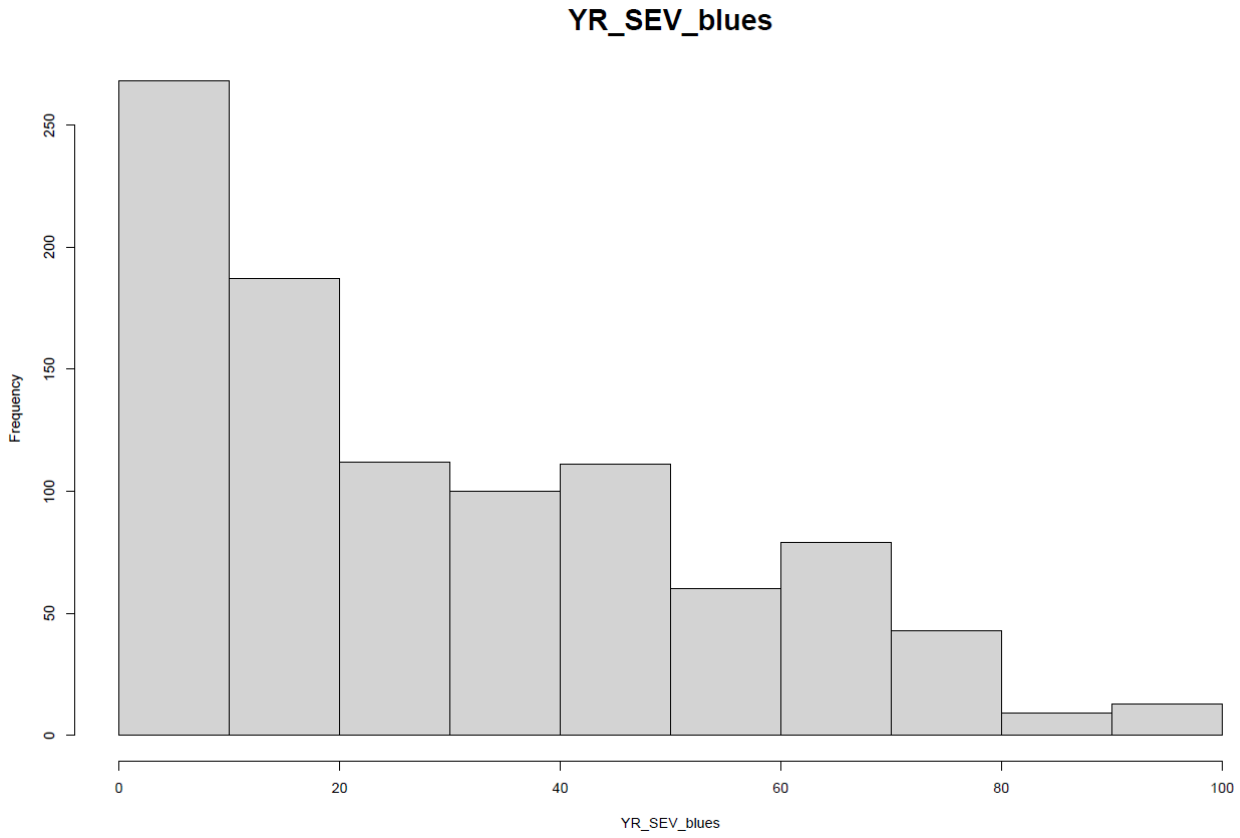


FIGURE 22: DISTRIBUTION HISTOGRAM FOR BLUES OBTAINED FOR SEVERITY IN LEBANON

2.3.9 Marchouch, Morocco

Morocco, in addition to the data from HD, IT and SEV, also has data relating to the height of the plant. The heritability for this value is 0.82, HD has the highest value of 0.93, followed by IT with 0.87 and the lowest value is attributed to the SEV with 0.75 (Tab. 17).

	HD	Ph	IT	SEV
min	76.00	60.09	-0.23	0.00
max	110.00	130.09	9.21	100.00
range	34.00	70.00	9.44	100.00
median	92.50	89.58	4.77	5.00
mean	92.18	91.35	3.52	17.48
SE.mean	0.23	0.56	0.12	0.90
var	40.68	240.63	11.55	612.94
std.dev	6.38	15.51	3.40	24.76
coef.var	0.07	0.17	0.97	1.42
h^2	0.93	0.82	0.87	0.75

TABLE 17: DESCRIPTIVE STATISTIC AND HERITABILITY FOR THE MAIN TRAITS SCORED IN MOROCCO

ANOVA shows high significance between genotypes, with a slightly lower significance on the plant height.

Trait	Variables	Sum Sq	Df	F value	Pr(>F)	
HD	**Genotype**	24265.9	759	10.4159	1.96E-08	***
HD	**Block**	5.5	3	0.5943	0.6254	
YR_IT	**Genotype**	432384	759	5.2271	1.45E-05	***
YR_IT	**Block**	296	3	0.9056	0.4543	
YR_SEV	**Genotype**	8712	758	8.2481	1.97E-07	***
YR_SEV	**Block**	6.6	3	1.5773	0.2233	
Ph	**Genotype**	125244	757	2.6938	0.003504	**
Ph	**Block**	175	3	0.9486	0.434265	

TABLE 18: ANOVA RESULTS FOR MOROCCO EXPERIMENT FOR ALL TRAITS CONSIDERED, USING GENOTYPE AND BLOCK AS VARIABLES

The distribution graphs show a non-normal trend for both IT and SEV, with a distribution more shifted towards low values (Fig. 23 and 24).

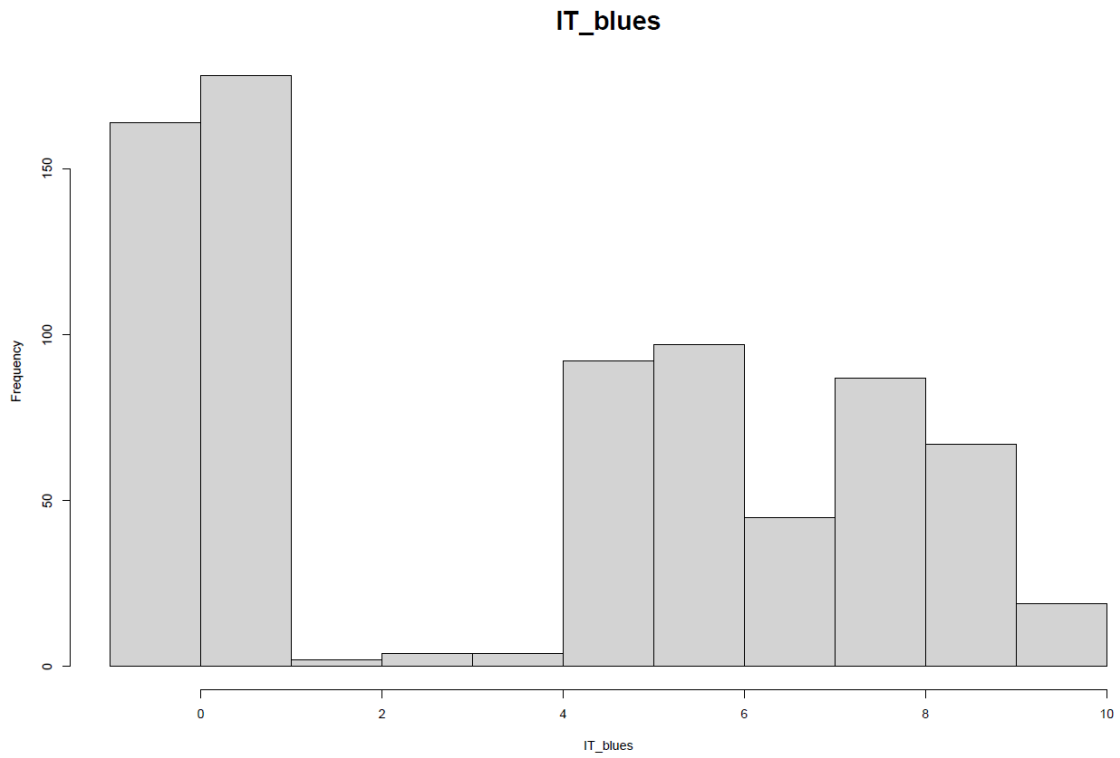


FIGURE 23: DISTRIBUTION HISTOGRAM FOR BLUES OBTAINED FOR INFECTION TYPE IN MOROCCO

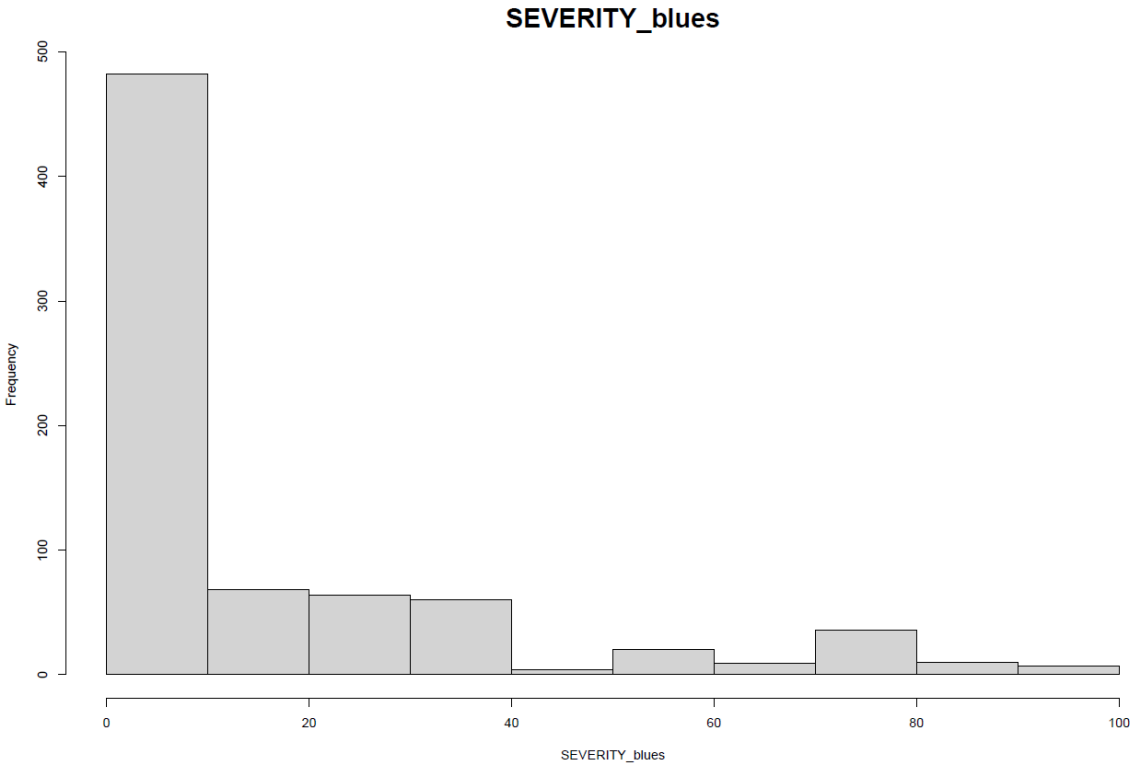


FIGURE 24: DISTRIBUTION HISTOGRAM FOR BLUES OBTAINED FOR SEVERITY IN MOROCCO

2.3.10 Izmir, Turkey

For Turkey, no HD data is available. For the data of IT and SEV we have very high heritability values with 0.90 for the IT and 0.87 for the SEV (Tab. 19), which show significant differences for the ANOVA between the genotypes for both traits and between the blocks for the trait IT (Tab. 20).

	YR_IT	YR_SEV
min	-0.61	0.09
max	9.54	99.09
range	10.15	99.00
median	3.13	19.58
mean	3.24	29.02
SE.mean	0.09	0.81
var	8.17	639.42
std.dev	2.86	25.29
coef.var	0.88	0.87
h^2	0.90	0.87

TABLE 19: DESCRIPTIVE STATISTIC AND HERITABILITY FOR THE MAIN TRAITS SCORED IN TURKEY

Trait	Variables	Sum Sq	Df	F value	Pr(>F)	
YR_IT	**Genotype**	7185.9	968	14.092	4.87E-12	***
YR_IT	**Block**	12.8	4	6.0983	0.001165	**
YR_SEV	**Genotype**	590120	968	7.458	1.55E-08	***
YR_SEV	**Block**	441	4	1.3495	0.2765	

TABLE 20: ANOVA RESULTS FOR TURKEY EXPERIMENT FOR ALL TRAITS CONSIDERED, USING GENOTYPE AND BLOCK AS VARIABLES

The distribution of the characters of IT and SEV is not normal with values of IT more oriented towards low values and values of SEV decidedly flattened to zero (Fig. 25 and 26).

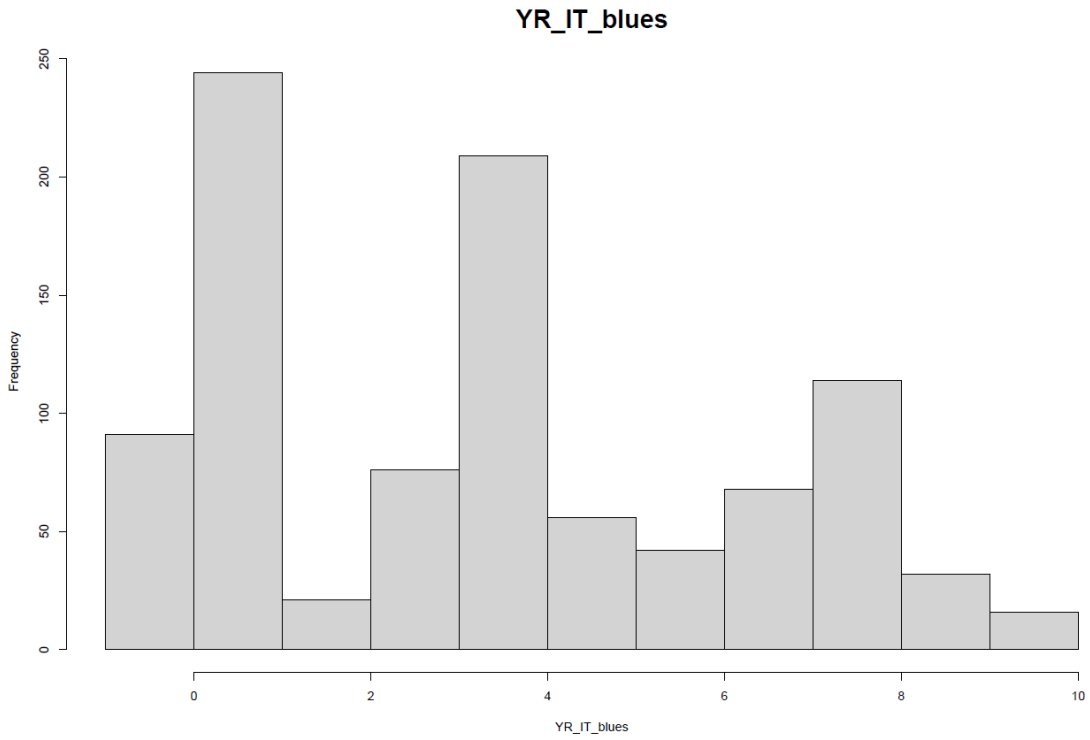


FIGURE 25: DISTRIBUTION HISTOGRAM FOR BLUES OBTAINED FOR INFECTION TYPE IN TURKEY

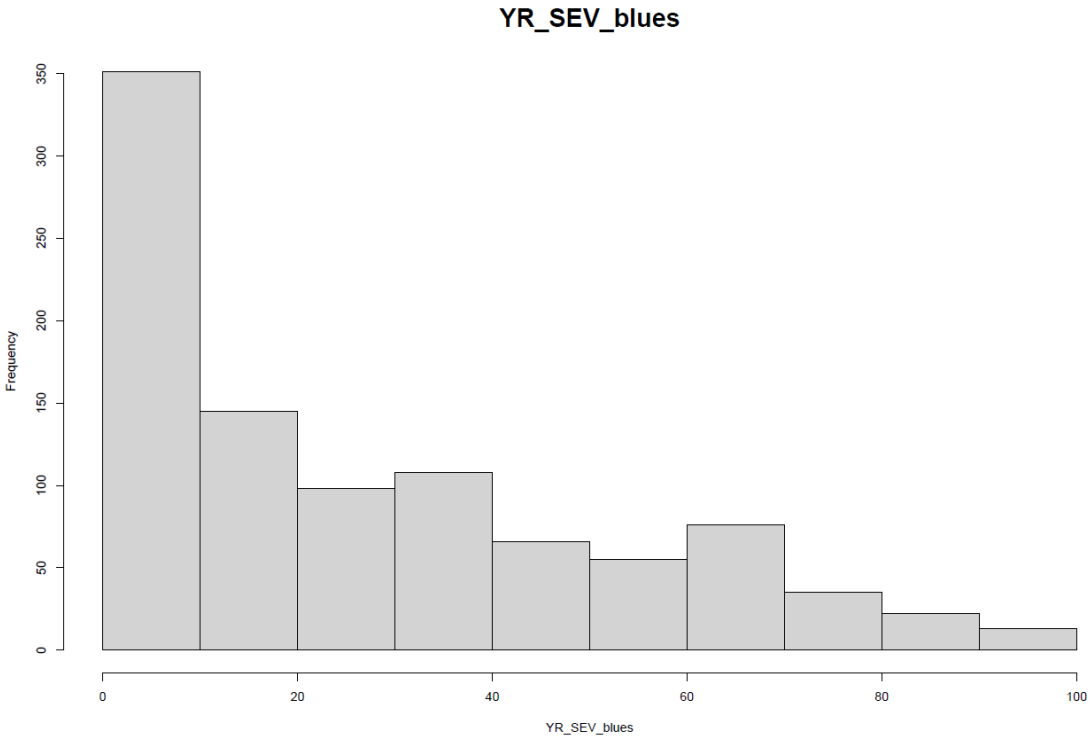


FIGURE 26: DISTRIBUTION HISTOGRAM FOR BLUES OBTAINED FOR SEVERITY IN TURKEY

2.3.11 Single environments summary

Here a short summary will be presented to highlight the differences among the different environments by taking into consideration the most important data: heritability, ANOVA and distribution histograms:

- Heritability: Heritability values show solid for all phenotypes in all environments. The environment with the lowest heritability (0,59-0,60 as the lower values) for different phenotypes is Grosseto 2020, showing a different trend compared to Grosseto 2019 and 2021. This may be caused by the fact that 2020, in Grosseto, was a dry year which prevented the proliferation of the disease.
- ANOVA: Has highlighted the following significance for blocks in these environments;
    1. Argentina for plant height
    2. Egitto, Grosseto 2021 and Turkey for heading date.
    3. Grosseto 2019 for Severity and AUDPC SEV
    4. Foggia for HD, infection type, AUDPC IT and RAUDPC IT
    5. Grosseto 2020 for HD, severity, AUDPC SEV and RAUDPC SEV

The significance found for these data indicates that there were differences between the genotypes within the different blocks, leading to a spatial correction based on them.

- Distribution histograms: Some environments showed non-normal disease distributions, particularly severity (Argentina, Grosseto 2020 and Lebanon). Since the disease requires certain conditions to develop, it can be assumed that non-normal distributions are the result of less favorable or better less inductive environmental conditions for infection development. In cases where both Infection type and Severity show a non-normal trend (Egypt, Morocco and Turkey), it can be assumed that the conditions were such that the pathogen had difficulty growing and developing.

2.3.12 PCA

With all the available environments we carried out a PCA to see if some environments clustered together according to the phenotype. To do this we used the ggplot2 package.

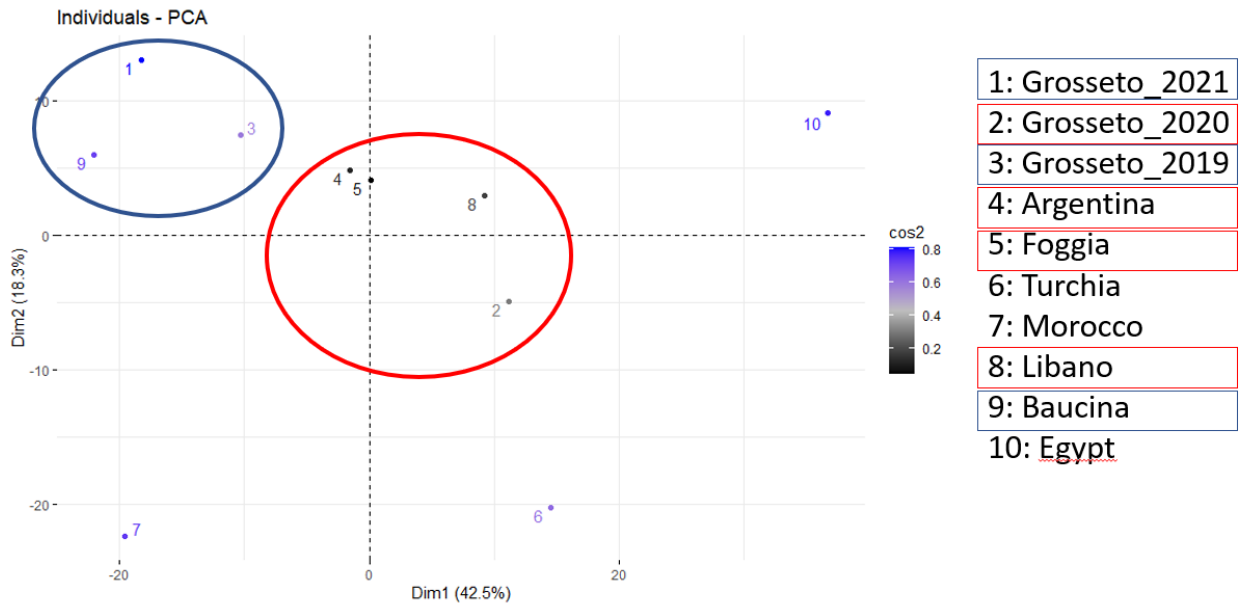


FIGURE 27: PCA REPRESENTING THE DIFFERENT ENVIRONMENTAL CLUSTER FOUND. IN BLUE THREE ENVIRONMENTS THAT CLUSTERIZE TOGETHER (GROSSETO 2019, GROSSETO 2021 AND BAUCINA), AND IN RED 4 DIFFERENT ENVIRONMENTS (GROSSETO 2020, ARGENTINA, FOGGIA AND LEBANON)

The PCA has identified two clusters of environments (Fig. 27), the first consisting of Grosseto 2019, 2021 and Baucina. The second includes Grosseto 2020, Lebanon, Argentina and Foggia. Environments that do not group with others were individually analsed only (Egypt, Morocco and Turkey).

2.3.13 Grosseto 2019, Grosseto 2021 and Baucina

	YR_IT	YR_SEV
min	0.93	2.59
max	9.11	107.14
range	8.17	104.55
median	5.60	57.05
mean	5.36	54.16
SE.mean	0.06	0.78
var	2.60	481.58
std.dev	1.61	21.94
coef.var	0.30	0.41
h^2	0.67	0.65

TABLE 21: DESCRIPTIVE STATISTIC AND HERITABILITY FOR THE MAIN TRAITS FOR THREE ENVIRONMENTS CLUSTER

The first cluster of environments shows an average high heritability (Tab. 21), with h2 for IT equal to 0.67 and 0.65 for SEV. Although they are similar environments, considering the fact that these are different years, the data is solid.

Trait	Variables	Sum Sq	Df	F value	Pr(>F)	
HD	**Genotype**	13762	788	4.21E+00	1.95E-09	***
HD	**Environment**	1289435	2	1.56E+05	<0.00000000000000022	***
HD	**Block**	65	9	1.75E+00	0.1002	
HD	**Genotype:Block**	372	138	6.51E-01	0.9759	
HD	**Genotype:Environment**	16765	106	3.82E+01	<0.00000000000000022	***
YR_IT	**Genotype**	4443.5	798	7.234	<0.00000000000000022	***
YR_IT	**Environment**	43.2	2	28.0802	9.02E-10	***
YR_IT	**Block**	1.8	9	0.2593	0.9833121	
YR_IT	**Genotype:Block**	89.8	171	0.6821	0.977366	
YR_IT	**Genotype:Environment**	234.8	147	2.0753	0.0003328	***
YR_SEV	**Genotype**	770036	798	12.1893	<0.00000000000000022	***
YR_SEV	**Environment**	1397	2	8.8214	0.0003704	***
YR_SEV	**Block**	1409	9	1.9776	0.0541741	.
YR_SEV	**Genotype:Block**	26211	171	1.9362	0.0008284	***
YR_SEV	**Genotype:Environment**	50249	147	4.318	5.75E-11	***

TABLE 22: ANOVA RESULTS FOR THREE ENVIRONMENTS CLUSTER FOR ALL TRAITS CONSIDERED, USING GENOTYPE, ENVIRONMENTS AND BLOCK AS VARIABLES, USING ALSO INTERACTION BETWEEN GENOTYPE X ENVIRONMENT AND GENOTUPE X BLOCK

Looking at the combined ANOVA (Tab. 22) the environment has significance for all characters, as well as the GxE (genotype X environment) interaction. In the case of SEV, we have interaction for every factor, including blocks and the interaction between genotype and blocks.

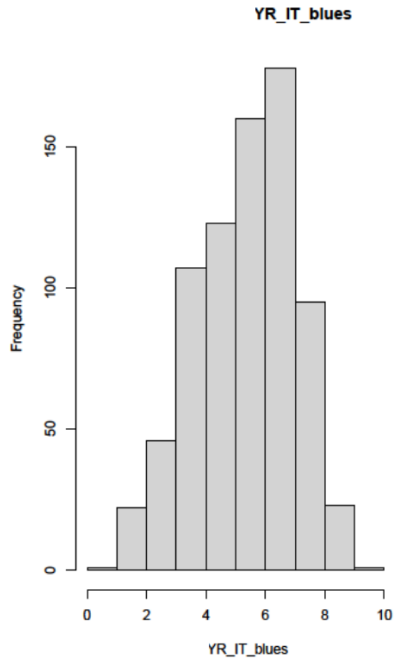


FIGURE 28: DISTRIBUTION HISTOGRAM FOR BLUES OBTAINED FOR INFECTION TYPE IN THE FIRST CLUSTER (THREE ENVIRONMENTS)

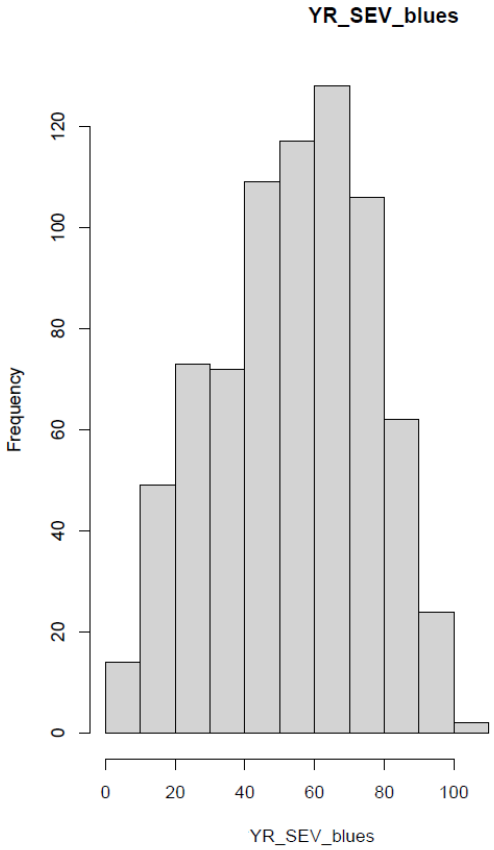


FIGURE 29: DISTRIBUTION HISTOGRAM FOR BLUES OBTAINED FOR SEVERITY IN THE FIRST CLUSTER (THREE ENVIRONMENTS)

The histograms of this environment have a normal distribution for the blues of IT and SEV (Fig. 28 and 29).

2.3.14 Grosseto 2020, Argentina, Foggia and Lebanon

	YR_IT	YR_SEV
min	-0.78	1.63
max	8.74	98.60
range	9.52	96.97
median	3.67	24.83
mean	3.72	28.56
SE.mean	0.07	0.55
var	4.45	306.44
std.dev	2.11	17.51
coef.var	0.57	0.61
h^2	0.75	0.64

TABLE 23: DESCRIPTIVE STATISTIC AND HERITABILITY FOR THE MAIN TRAITS FOR FOUR ENVIRONMENTS CLUSTER

In this cluster of environments, we have higher inheritability values for IT, with a value of 0.75 (compared to 0.67 in the other cluster), while the value of SEV is practically identical (0.64 vs 0.65) (Tab. 23).

Trait	Variables	Sum Sq	Df	F value	Pr(>F)	
HD	**Genotype**	69663	998	31.8481	<0.00000000000000022	***
HD	**Environment**	551918	3	83939.69	<0.00000000000000022	***
HD	**Block**	768	10	35.0573	<0.00000000000000022	***
HD	**Genotype:Block**	14887	452	15.0276	<0.00000000000000022	***
HD	**Genotype:Environment**	5959	702	3.8731	<0.00000000000000022	***
YR_IT	**Genotype**	9955	997	10.9412	<0.00000000000000022	***
YR_IT	**Environment**	146.1	3	53.3793	<0.00000000000000022	***
YR_IT	**Block**	35.2	10	3.8561	6.77E-05	***
YR_IT	**Genotype:Block**	474.5	449	1.1579	0.09299	.
YR_IT	**Genotype:Environment**	1605.8	700	2.5138	<0.00000000000000022	***
YR_SEV	**Genotype**	790162	997	9.6304	<0.00000000000000022	***
YR_SEV	**Environment**	134389	3	544.3335	<0.00000000000000022	***
YR_SEV	**Block**	2025	10	2.4606	0.007916	**
YR_SEV	**Genotype:Block**	112978	449	3.0575	<0.00000000000000022	***
YR_SEV	**Genotype:Environment**	158521	694	2.7756	<0.00000000000000022	***

TABLE 24: ANOVA RESULTS FOR FOUR ENVIRONMENTS CLUSTER FOR ALL TRAITS CONSIDERED, USING GENOTYPE, ENVIRONMENTS AND BLOCK AS VARIABLES, USING ALSO INTERACTION BETWEEN GENOTYPE X ENVIRONMENT AND GENOTUPE X BLOCK

In this case we have significance for each phenotype with each factor of the ANOVA, indicating interaction between all the components analyzed and a strong GxE interaction (Tab. 24).

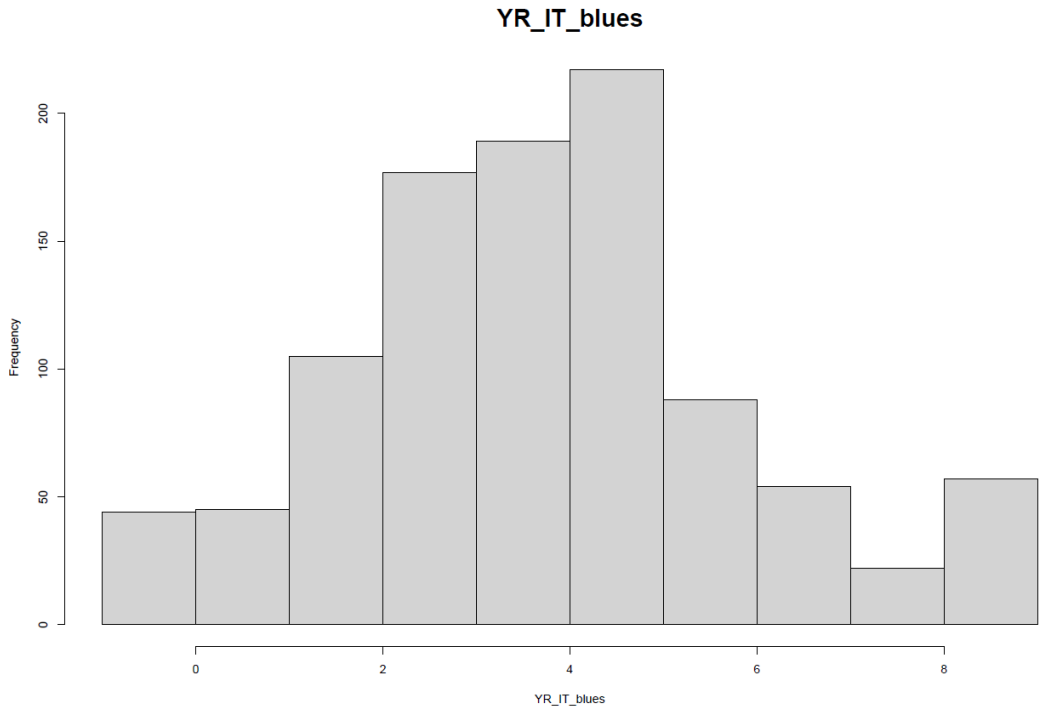


FIGURE 30: DISTRIBUTION HISTOGRAM FOR BLUES OBTAINED FOR INFECTION TYPE IN THE SECOND CLUSTER (FOUR ENVIRONMENTS)

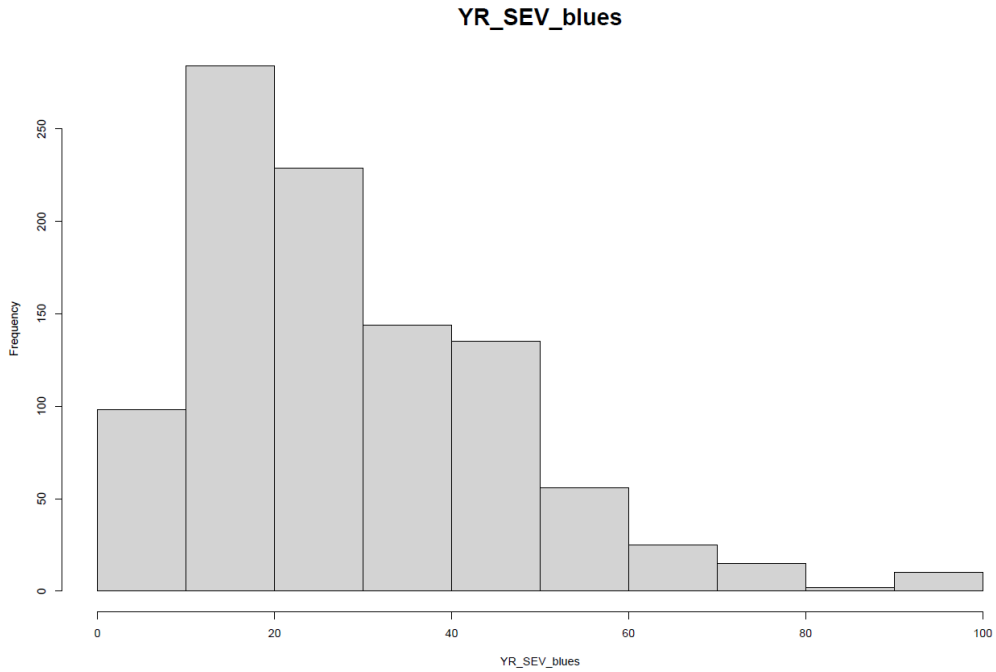


FIGURE 31: DISTRIBUTION HISTOGRAM FOR BLUES OBTAINED FOR SEVERITY IN THE SECOND CLUSTER (FOUR ENVIRONMENTS)

The histograms show that for IT the distribution is almost normal, while for SEV we have a narrower distribution towards a low spread of the disease (Fig. 30 and 31).

2.3.15 GWAS results

The blues were used to carry out the GWAS using 5 models: BLink, FARMCPU, MLMM, MLM and GLM. To do this, a hapmap (already imputed) filtered for the correctly mapped SNPs and the kinship matrix (produced using TASSEL5) were used. The hapmap was initially used to calculate the LD decay.

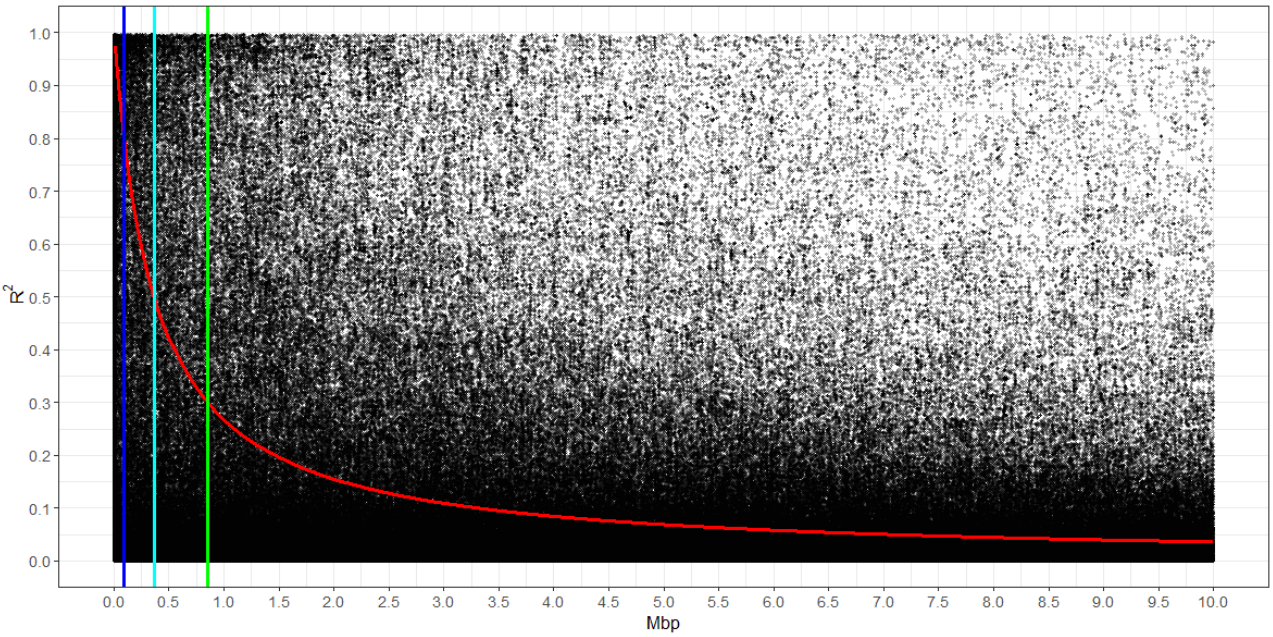


FIGURE 32: LD DECAY WITH THREE DIFFERENT THRESHOLDS: 0,3 (GREEN), 0,5 (LIGHT BLUE) AND 0,8 (BLUE)

The computed LD decay at r2 threshold of 0.3 was approximately 1Mbp (Fig. 32).

The population structure was also identified for the panel, 10 subgroups were identified: 5 groups formed by cultivars (EPO, Cimmyt '80, Icarda temperate, North_American/French, Italian CV), and 5 landraces groups (Ethiopian landraces, Western Asian founders, Western Asian landraces, Mediterranean landraces and Central Mediterranean landraces). To obtain this, distance and kinship matrix was computed with Tassel 5, and dendrogram tree was performed with Ward.D2 algorithm (Fig. 33).

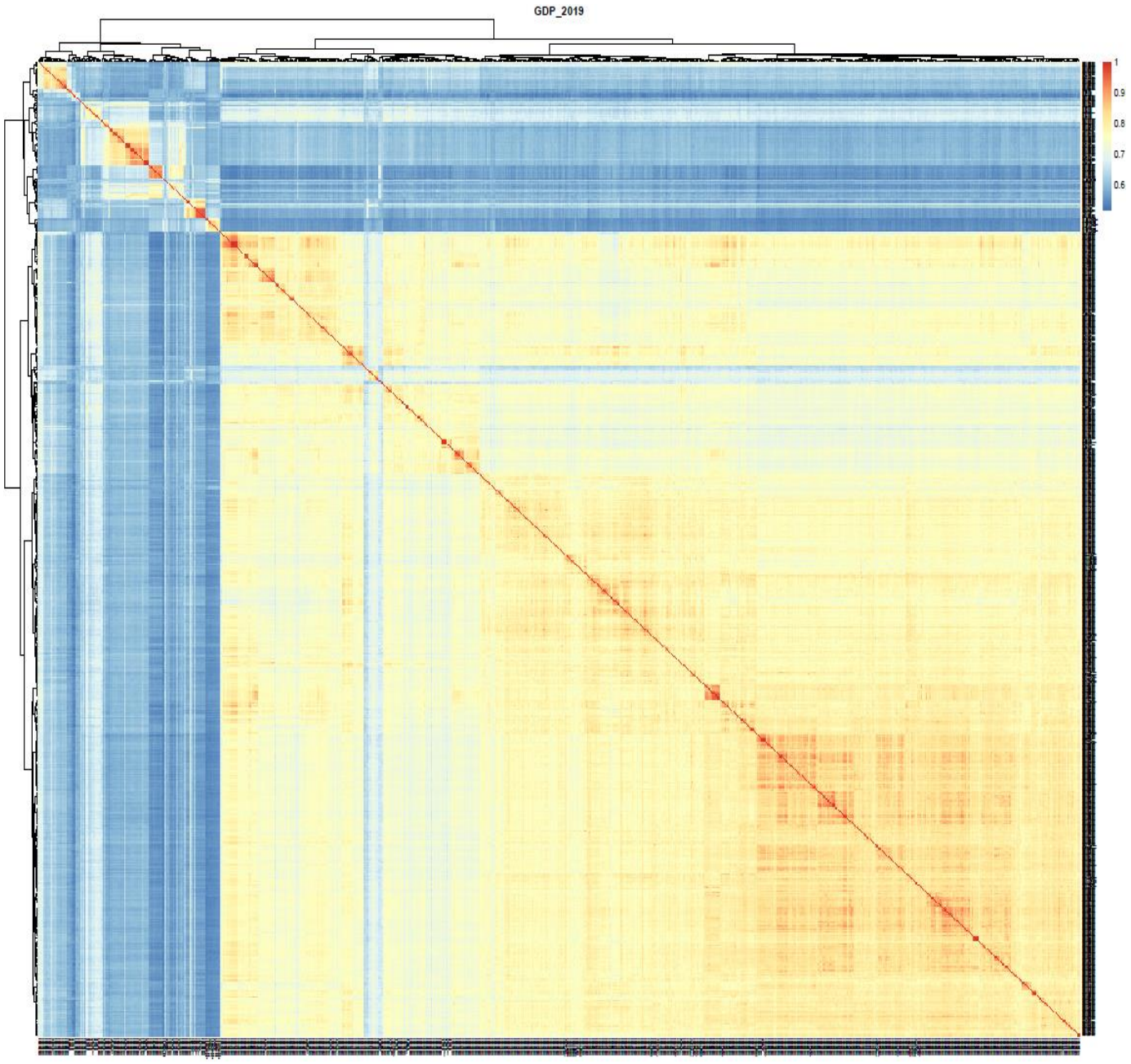


FIGURE 33: POPULATION STRUCTURE FOR ALL THE GENOTYPES USED IN THIS MULTIENVIRONMENTAL EXPERIMENT

For the 10 subgroups the distribution of the infection type within the cluster of three environments was checked to note the distributions of the phenotypes (Fig. 34).

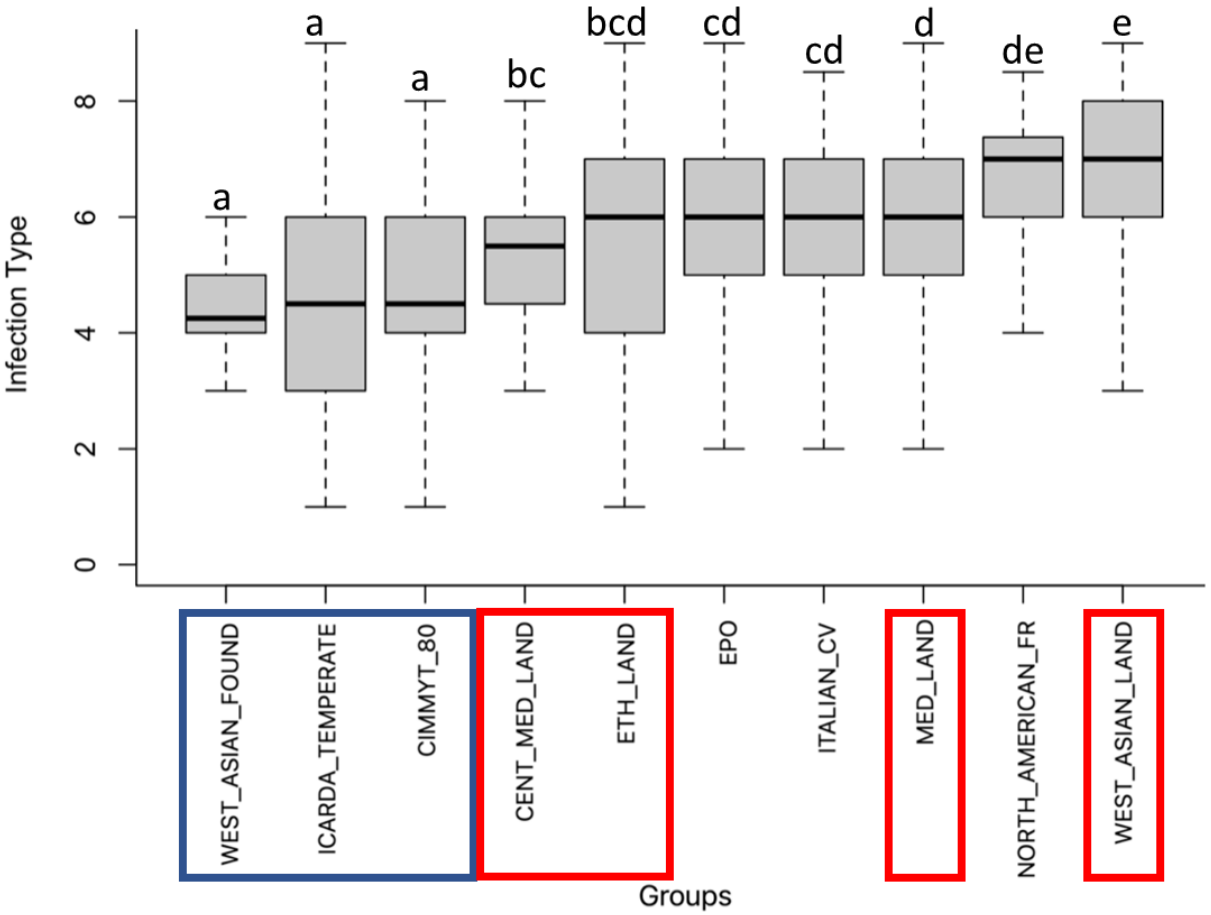


FIGURE 34: DISTRIBUTION BOXPLOT FOR ALL THE SUBGROUPS FOUND USING ADMIXTURE. IN BLUE THE MOST RESISTANT GROUPS. IN RED THE LANDRACE GROUPS SHOWING MORE SUSCEPTIBILITY.

From the boxplot it can be seen how the three most resistant groups (highlighted in blue) are made up of cultivars from the cimmyt and the icarda breeding programs. The other resistant group, as well as the most resistant, is constituted by the landraces which gave rise to the modern varieties of West Asia. In red we can see the other groups of landraces which shows more susceptibility than the cultivars. However the landraces are the groups carrying partial resistance IT that are different from the resistances present in modern germplasm and that are actually focus of active research today.

For each environment a GWAS has been carried out which will be compared below for the Infection type and severity characters (Blink model, Fig35).

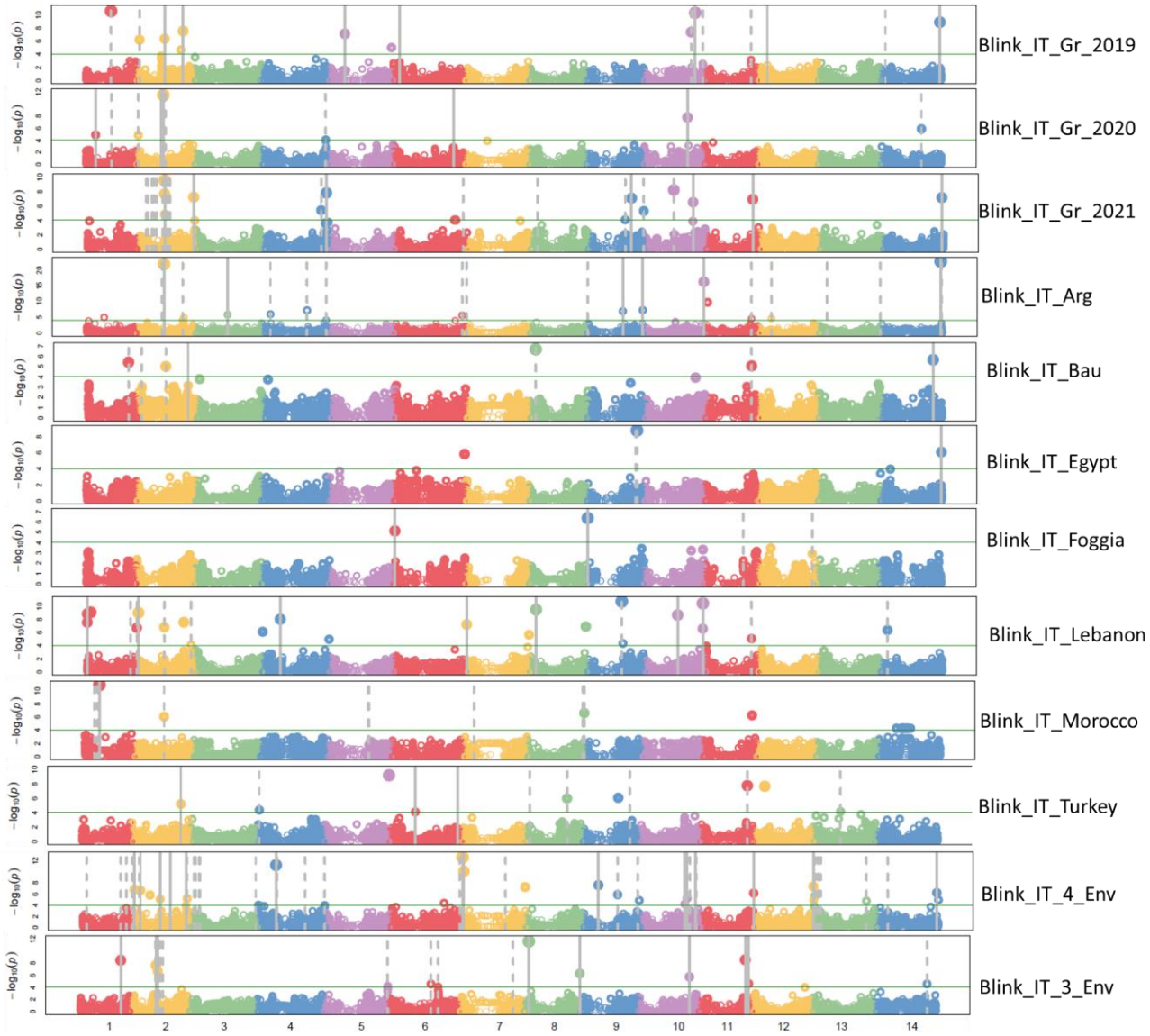


FIGURE 35: MANHATTAN PLOT OF THE BLINK MODEL FOR ALL THE ENVIRONMENTS AND THE CLUSTERS, FOR IT TRAIT

The manhattan plots show the peaks on each chromosome and their significance, the threshold of which is set to 4 using the Bonferroni threshold. Comparing all the environments we can see how on chromosome 1A we have peaks in the majority of the environments, mostly in the proximal and distal region. On chromosome 1B we can see several peaks representing *Yr24/26*, a QTL already studied by UNIBO, and probably other loci in almost all the environments. Chromosome 2A has a different shape, with few peaks in Argentina only for Blink model. Chromosome 2B is almost empty as well with some effects on the 4 environments combined, Lebanon and Argentina plot. Chromosome 3A has a peak only for one environment, Grosseto 2019. Chromosome 3B has few peaks, one in Foggia and one in Turkey. Chromosome 4A has two peaks, one in the 4 environments cluster and one in Lebanon. Chromosome 4B has few coincident peaks, the proximal is in common

with the second cluster and Lebanon, while the distal peak is in common again between the second cluster, Morocco and Lebanon. On 5A we have some coincidence peaks with the 4 environments cluster and Foggia (proximal peak), Turkey, Lebanon and Grosseto 2021 (middle peak), and Egypt, Argentina and Grosseto 2021 (distal peak). Chromosome 5B has a peak coincidence between Lebanon and Grosseto 2021, and another coincidence peak for five environments. Chromosome 6A has a distal peak in common between several environments. Chromosome 6B has just few peaks, one in Turkey and one in the first cluster of environments. On 7A we have just two peaks, one in the 4 environments cluster and one in Turkey. On the last chromosome, 7B, we have a high distal peak in common with several environments.

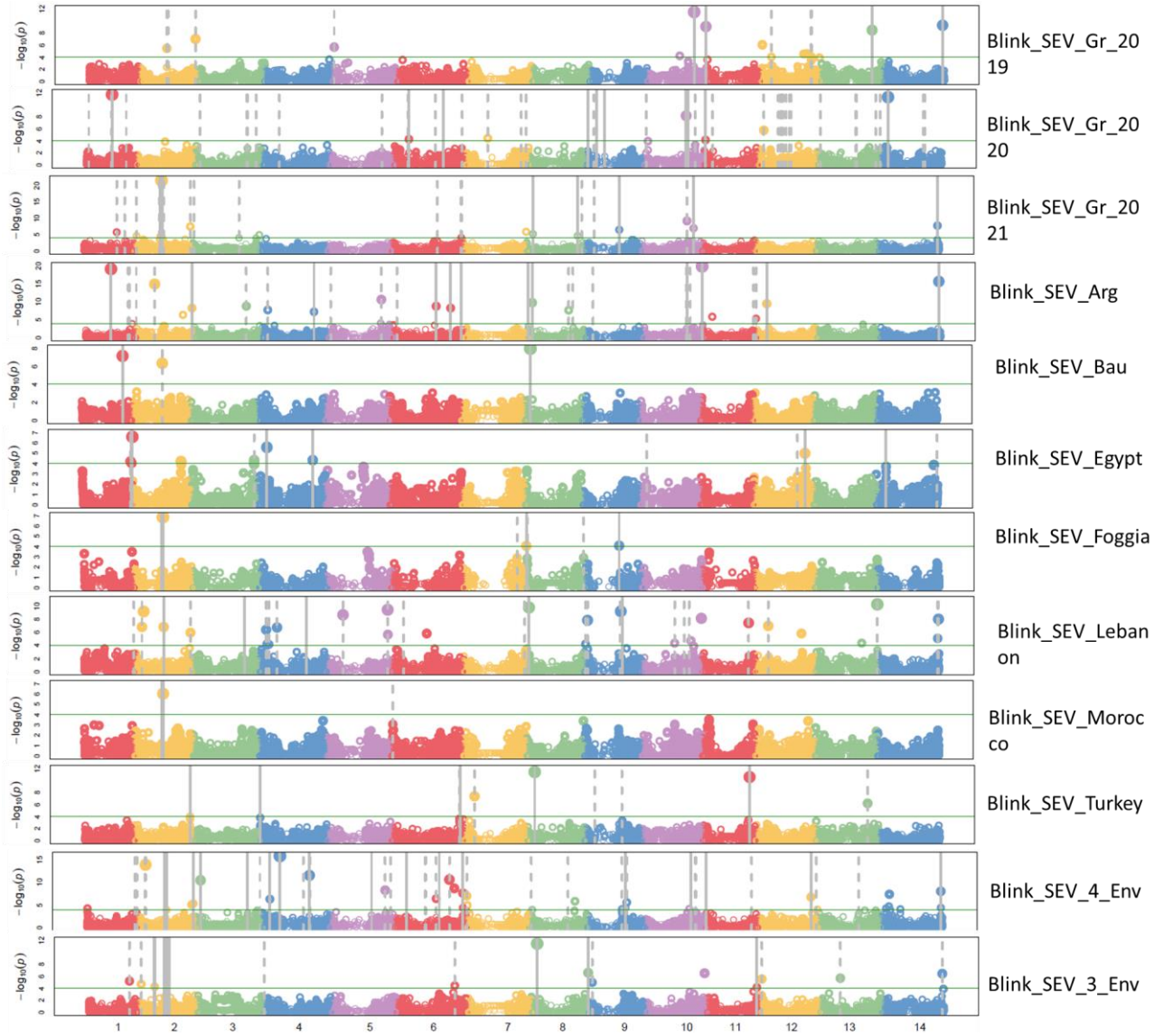


FIGURE 36: MANHATTAN PLOT OF THE BLINK MODEL FOR ALL THE ENVIRONMENTS AND THE CLUSTERS, FOR SEV TRAIT

For SEV (Fig. 36), we have a distal peak on chromosome 1A for three environments, and a middle peak for two environments. On chromosome 1B we have the same peak for IT. On chromosome 2A we have one proximal peak on the first cluster and a distal peak in Egypt and Argentina. On 2B we have a proximal peak on 4 environments, Egypt and Argentina. We also have a middle peak on the first cluster and a distal peak in Egypt and Argentina. On 3A we can see a proximal peak for Lebanon and Grosseto 2019, and a distal peak on 4 environments, Lebanon and Argentina. Peaks on 3B we have a peak in common for the first cluster and Argentina, On 4A we just have three different peaks, one in the first cluster, one in Turkey and one in Grosseto 2020. On chromosome 4B we have a proximal peak on the second cluster, Turkey, Lebanon, Baucina, Argentina and Grosseto 2021. On 5A we have a proximal peak on the second cluster and in Lebanon, and another one, in the same cluster, in the distal region. On chromosome 5B we have some distal peak in the second environment cluster, Lebanon, Argentina, Grosseto 2021, Grosseto 2020 and Grosseto 2019. On chromosome 6A we have a proximal peak in Grosseto 2020 and Argentina and a distal peak in the second cluster, Turkey, Lebanon and Argentina. Chromosome 6B has a proximal peak for the 3 environments, Lebanon, Argentina, Grosseto 2020 and Grosseto 2019. There is also a distal peak for the 4 environments, Lebanon, Egypt and Grosseto 2019. Chromosome 7A has a single peak in the 3 environments cluster and a distal peak for Turkey, Lebanon and Grosseto 2019. On chromosome 7B we have a peak in common for the first cluster of environments and Grosseto 2020, plus a distal peak peak in common with 6 environments. We can see several peaks in common between IT and SEV, needing further analysis to understand how they can be explained by common or different loci on the response pathways. The analysis reported a strong overlap in several environments with *Yr65* (Cheng et al., 2014), which is why we went to precisely investigate the candidate genes within its range. His range showed, using knetminers (Fig. 37), several genes associated with resistance against yellow rust, with several proteins acting on the resistance metabolic pathway that require further study. This, in any case, can be considered a milestone to further investigate candidate genes for this strong gene and all others.

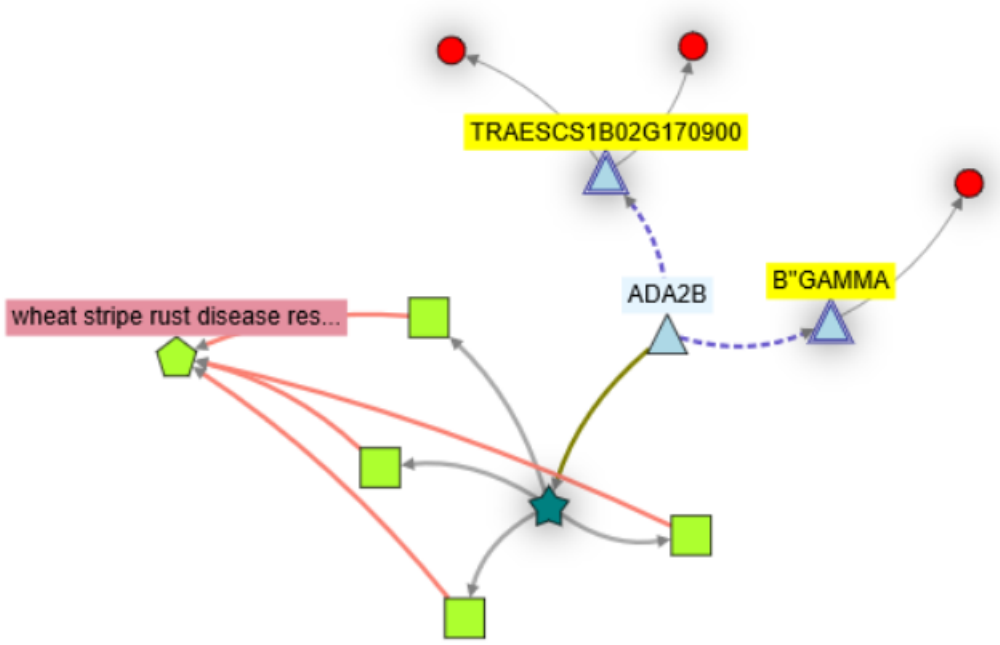


FIGURE 37: *YR65* CANDIDATE GENES FOR STRIPE RUST RESISTANCE

2.3.16 KASP® essays

Following the phenotypic and haplotypic results of GWAS, 31 assays were designed on a total of 16 SNP, named KUBO (Kasp-UniBO). In chromosome 1Bs is IWB31756 (GENE-0227_217), at position 36.9 cM on the durum consensus map, on which primer K108 was drawn. On IWB74352 (tplb0029b01_1022 at 37.2 cM) the K109, K110 on IWB51681 (Ra_c27867_685 at 37.4 cM), IWB31942 with K111. IWB54040 with K126, IWB62955 with K127, IWB38604 (Ku_c17486_152 at 39.1 cM) with K128, IWA8337 (wsnp_RFL_Contig2446_2008649 at 39.1 cM) with K129, IWB52670 (Ra_c7745_1437 at 37.1 cM) with K130 and K131. IWB74352 (tplb0029b01_1022 at 37.2 cM) with K132, IWB38604 (Ku_c17486_152 at 39.1 cM) for K133, K134, K135, K136, and finally IWB12179 (BS00101500_51 at 41.2 cM) with K137 and K138. The KASP® that will be addressed in this work (K109 and K129) are those only the best ones that came out of the study, as well as found for a gene of great interest for yellow rust resistance: Yr24/26.

Kubo109, drawn on SNP IWB74352 in chromosome 1B, is perfectly consistent with the expected results and divides the genotypes into two clusters. We find the lines Gdp-024, Gdp-214, Gdp-365, Gdp-236, Gdp-367, Gdp-104, Gdp-327, Gdp-373, Gdp-748, Gdp-262, Gdp-378, Gdp-118, Gdp-269, Gdp-119, Gdp-280, Gdp-051, Gdp-342, Gdp-237 with susceptible scores grouped at allele 2 while

the remaining resistant Gdp-296, Gdp-031, Gdp-097, Gdp-408, Gdp-419, Gdp-042, Gdp-180, Gdp-045, Gdp-252, Gdp-272, Gdp-027, Gdp-153, Gdp-062, Gdp-295, Gdp-191 grouped at allele 1. Gdp-064, Gdp-48 and Gdp-278 are heterozygous while Gdp-367 and Gdp-035 indeterminate.

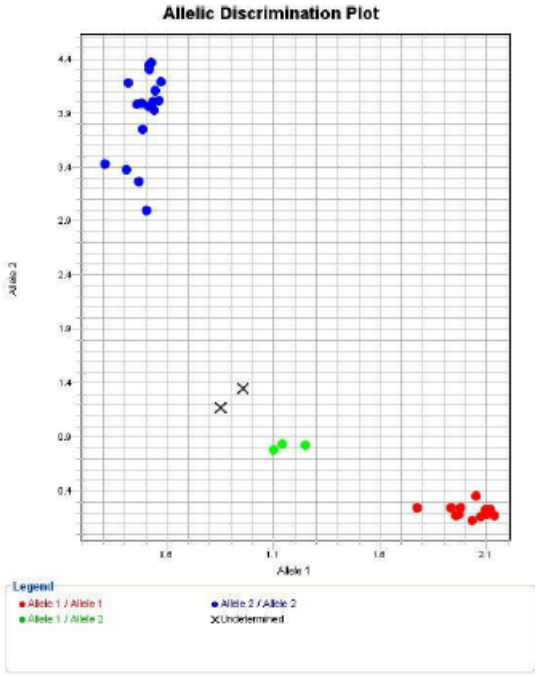


FIGURE 38: K109 RESULTS

K129 drawn in SNP IWA8337 on chromosome 1B discriminates into two clusters the genotypes. Gdp-024, Gdp-051, Gdp-104, Gdp-118, Gdp-119, Gdp-214, Gdp-236, Gdp-237, Gdp-252, Gdp-262, Gdp-269, Gdp-280, Gdp-327, Gdp-342, Gdp-365, Gdp-373 at allele 2, Gdp-031, Gdp-035, Gdp-042, Gdp-045, Gdp-048, Gdp-062, Gdp-064, Gdp-097, Gdp-153, Gdp-180, Gdp-191, Gdp-278, Gdp-295, Gdp-296, Gdp-408, Gdp-419 at allele 1.

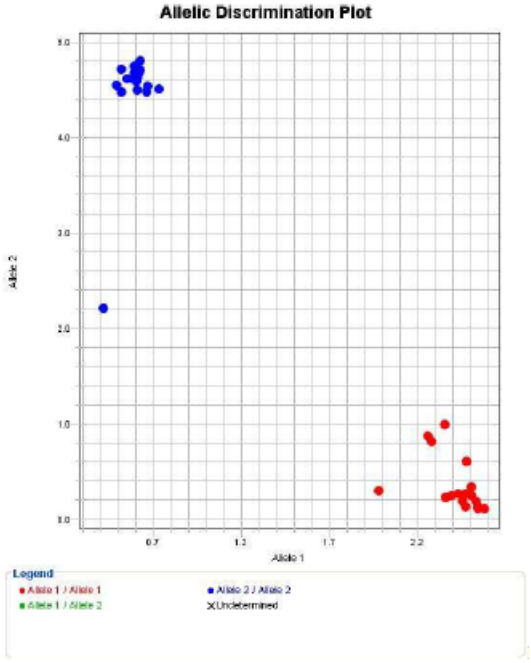


FIGURE 39: K129 RESULTS

2.4 Discussion

This experiment was conducted by taking data from ten different environments of the Mediterranean area. The use of a representative panel such as that of the GDP aimed to identify new genetic resources that can help genetic improvement for an important area for durum wheat such as the Mediterranean area. In this area, novel varieties better adapted to climate change and the new disease outbreaks that will arise are needed. The GWAS analysis conducted revealed a considerable number of peaks that explained the phenotypic variance for at least 10%. Some of these peaks were common to multiple traits or multiple environments. An intensive meta-QTL-analysis was carried out for this work, which showed the following results. On chromosome 1A, overlaps were found for QTL *QYr.tam-1A_Avocet-YrA* (Basnet et al., 2014a) *QYr.tam-1AS_TAM111* (Basnet et al., 2014a; Maccaferri et al., 2015), and QTL reported by Bariana et al. 2010 and Liu et al., 2017c. two peaks coinciding with the *IWB35957* marker were found, in common for the cluster of three environments, also in common with Baucina 2019. Several QTL found on 1B are coincident with *Yr9* (McIntosh et al., 1998), *Yr65* (Cheng et al., 2014), *QYr.cym-1BS_Pastor* (Rosewarne et al., 2012), a QTL reported by A. A. Naz et al., 2012, *YrExp1* (Lin & Chen, 2008; Maccaferri et al., 2015), *QYrex.wgp-1BL_Express* (Lin & Chen, 2009), *QYr.sun-1B_Kukri* (Bariana et al., 2010), *Yr24* (Huerta-Espino & Singh, 2017), *Yr26* (Cheng et al., 2014), *Yr29/Lr46* (Cobo et al., 2019). Peaks found on chromosome 2A find coincidence with *QYr.inra-2BS_Renan* (Dedryver et al., 2009) and *QYrst.orr-2B.1_Stephens* (Vazquez et al., 2012). On chromosome 2B, several overlaps were found for Yr31 (Singh et al., 2003), Yr41 (Maccaferri et al., 2015), QYr.inra-2B.1_CampRemy (Mallard et al., 2005), Yr43 (Maccaferri et al., 2015), Yr3 (Maccaferri et al., 2015). Only two coincidence peaks were found on chromosome 3B, which in the bibliography were matching with *QYr.sun-3B_Kukri* (Bariana et al., 2010) and *QYrid.ui-3B_RioBlanco* (Chen et al., 2012). On chromosome 4A several peaks overlap with QTL found by Liu et al., 2017c. On 4B the peaks are almost coincident, once again for the cluster of three environments and Baucina, with one peak coincidence with *QYr-4B_Sachem* (Singh et al., 2013). On chromosome 5A we have a peak in the proximal part, linked to the infection type, found in the Foggia experiment, and two peaks on the distal part, rather close, again for the infection type of Egypt and Argentina. Again, bibliographic evidence showed coincidence with QTLs from Liu et al., 2017c and with *QYr.cim-5AL_Francolin* (Lan et al., 2014). On chromosome 5B there is a coincident peak for severity and infection type found in the Grosseto 2019 experiment, with bibliographic evidence not too far from the QTL *YrEXP-2* (*Graingenes Yr* database). The other peak found is always

for the severity of the same experimental test in a slightly more distal position. Coincidence peaks were found for *QYr.ufs-5B_Cappelle-Desprez* (Agenbag et al., 2012), *QYr.sun-5B_Wollaroi* (Bansal et al., 2014) and one is very close with *Yr48* (Maccaferri et al. 2015; Lowe et al. 2011). On 6A we have two almost coincident peaks for two traits in the two clusters of environments, which have no bibliographic confirmation. On 6B clusters overlap with *QYr.inra-6B_Renan* (Dedryver et al., 2009) and *QYr.ucw-6B* (Dong et al., 2017), while 7A has a peak coincidence with *Yrxy1* (Maccaferri et al. 2015). Also, on the 7B chromosome we found two almost coincident peaks, one for Argentina which explains both infection type and severity and one for the infection type of the experiment held in Egypt. For peaks on 7B, we've found an overlap for *Yr67* (Bariana et al., 2022), *QYr-7BL_Strongfield* (Singh et al., 2013) and QTL *Yrw.wgb.7B-3* (Mu et al., 2020). As mentioned earlier, Yr65 was one of the genes with the most hits for different environments. This led us to focus on this gene for the characterization of candidate genes that, however, require further study. *Knetminer* and *Graingenes* confirm an association, obviously, between the gene and plant response. In all of this, however, further investigation of the role of interval-related proteins that has already been previously identified is needed to carry out appropriate studies for the characterization of this important and widespread gene within durum wheat. In addition to the bibliographic hits made, this study also found some intervals that were not matched in the bibliography, indicating the possibility that these could be considered putative novel genes that could be of particular interest within modern breeding and that could open up a range of challenges to characterize them so that they could bring new genetic variability, in terms of yellow rust resistance, to the global durum wheat community. The putative novel genes that have been identified are respectively *QYr.uboIT_1A_Lebanon_CV_&_L* on chromosome 1A, *QYr.ubo_IT_3A_3MULTIENVIRONMENT_CV_&_L* on 3A *QYr.ubo_IT_3B_4MULTIENVIRONMENT_CV_&_L* on 3B, *QYr.ubo_IT_4A.2_FOGGIA_CV_&_L*, *QYr.ubo_IT_4A_GRO_2020_CV_&_L*, *QYr.ubo_IT_4A.3_TURKEY_CV_&_L* on 4A, *QYr.ubo_IT_4B.2_LIBAN_CV_&_L* on 4B, *QYr.ubo_IT_5A_ARG_CV_&_L* on 5A, *QYr.ubo_IT_5B.3_4MULTIENVIRONMENT_CV_&_L*, *QYr.ubo_IT_5B.4_3MULTIENVIRONMENT_CV_&_L* on the 5B, *QYr.ubo_IT_6A_LIBANO_CV_&_L*, *QYr.ubo_IT_6A.2_TURKEY_CV_&_L* on the 6A, *QYr.ubo_IT_7A_4MULTIENVIRONMENT_CV_&_L* on the 7A and finally *QYr.ubo_IT_7B_LIBANO_CV_&_L* on the 7B.

3 Chapter II: Septoria Tritici Blotch (STB)

3.1 Introduction

3.1.1 *Zymoseptoria tritici*

Zymoseptoria tritici (syns. *Mycosphaerella graminicola*, *Septoria tritici*) is an Ascomycete fungus (Quaedvlieg et al., 2011) and it is the causal agent of septoria tritici blotch (STB), the main leaf disease of wheat in temperate regions (Fones & Gurr, 2015) and a major threat for wheat production globally. The initial symptoms of STB are small chlorotic spots on the leaves that appear soon after seedlings emerge in the fall or spring. As they enlarge, the lesions become light tan and develop darker colored fruiting bodies. Lesions on mature leaves most often are long, narrow and delimited by leaf veins but also can be shaped irregularly or can be elliptical, particularly on seedlings or leaves that were young when infected. Mature lesions contain black or brown fruiting structures, the asexual pycnidia or sexual pseudothecia. The pycnidia or pseudothecia develop in the substomatal cavities of the host so are spaced regularly within the lesions (Ponomarenko et al., 2011). The presence of pycnidia on lesions of the foliar tissue is also an important diagnostic element of the disease in the field because the symptoms related to the septoria complex can be easily confused with those related to other diseases of different fungal origin (CAEMILIA, 2008). STB is especially damaging in humid and temperate areas where yield losses may reach up to 50% (Eyal et al., 1987). In Europe, STB is the most economically damaging disease of wheat, with an estimated 1 billion euros per year in fungicide expenditure directed toward its control (Kettles & Kanyuka, 2016). *Zymoseptoria tritici* causes considerable damage in many countries of the Mediterranean area: in France, losses due to Z. tritici are estimated each year to 1.7 tons/hectare on average (Cheval et al., 2017). In UK annual average losses are around 20% in untreated fields, while losses of 5 to 10% are more typical in fields planted to resistant cultivars or treated with fungicides (McDonald & Mundt, 2016; Fones & Gurr, 2015). *Z. tritici* has become more important in Tunisia since the introduction of early maturing, semi dwarf, high yielding varieties. In some Tunisian fields located at Jendouba, Bizerte, Beja, and Kef were recorded STB incidence at 60% and severity 40% (Chedli et al., 2018).

3.1.2 Biology and cycle

Z. tritici is heterothallic with a bipolar mating system with two mating type alleles, *mat 1-1* and *mat 1-2* (Neddaf et al., 2017) and the cycle generally starts with ascospores (sexual reproduction), which are transported over long distances in the wind and when the spores are carried on the leaf surface the first phase of infection begins (Siah et al., 2014). *Zymoseptoria tritici* is an hemibiotrophic fungus, with two distinct phases of infection. The first is the biotrophic one, which begins with the spores transported on the leaf surface by the rain splashes or wind. The spores then germinate and start the penetration in the cellular tissues through the stomata, giving life to an asymptomatic phase lasting 7-10 days. This phase is asymptomatic because *Z. tritici* can suppress or avoid the plant's defence mechanisms. During the biotrophic stage the pathogen use lipids and fatty acids from the host cell as suggested by the increase of expression of genes encoding cutinase and lipase proteins during this phase (Duba et al., 2018) the necrotic phase is characterized by a strong defence response and reprogramming of transcriptomes that leads to apoptosis and the release of nutrients in the apoplast, thus going to promote the growth of the fungus. The dead areas of the leaf expand to form lesions in which pycnids, the asexual structures of the fungus, form. These pycnids will release pycnidiospores from the stomatal openings to allow the fungus to carry out multiple cycles of infection via rain splash (Kettles & Kanyuka, 2016). Ascospores, however, are also the cause of the progress of the infection during the vegetative season as there are more cycles of sexual reproduction during the year, which leads to *Z. tritici* to have a wide genetic variability (Siah et al., 2014). The pathogen requires temperatures around 20° and more than 85% RH for his infection progress; in suboptimal condition the latency is quite longer but it doesn't reduce the damage. During the growing season it is important to take rainfall in consideration as a high number of rainy days may be responsible for a high rate of infection since rainfalls are crucial for temperature-dependent polycyclic disease progression (Chaloner et al., 2019).

3.1.3 Host-pathogen interaction in *Z. tritici*

For the plant, to be able to trigger its own defence system, is necessary to recognizes non-self-molecules that are called PAMPs (Pathogen-associated molecular patterns), which range from nucleic acids to lipids, peptides, and glycans (Fermaintt et al., 2019; Marshall et al., 2011; Kettles & Kanyuka, 2016). These PAMPs, if recognized, lead to PAMP-triggered immunity (PTI) (Marshall et

al., 2011; Kettles & Kanyuka, 2016). One of the most important PAMPs in fungal pathogens is chitin, as it is one of the main components of their cell wall and is not present in plants. To overcome the problem of chitin recognition, some fungal pathogens (such as *Z. tritici*) secrete proteins containing LysM (carbohydrate-binding protein domains) that have ultrahigh chitin-binding affinity and can interfere with chitin detection by the host (Kombrink & Tomma, 2013; Kettles & Kanyuka, 2016). *Z. tritici*, during its biotrophic phase, encodes three LysM-containing proteins (Mg1LysM, Mg3LysM and MgxLysM). Both Mg1LysM and Mg3LysM are transcriptionally highly up regulated during symptomless colonization of wheat (peaking at day 4 for Mg3LysM and at day 9 for Mg1LysM) and both bind chitin. However, only Mg3LysM can avoid plant's defenses by blocking chitin-induced elicitation. Both Mg3LysM and Mg1LysM also have a unique protective effect for *Z. tritici*, shielding the fungal cell wall from digestion by host chitinases (Marshall et al., 2011; Kettles & Kanyuka, 2016). In other words, these proteins with LysM domains prevent the pathogen to be recognised by the host and the activation of plant immune system, avoiding HR and allowing the biotrophic phase of *Z. tritici*. During the switch from asymptomatic phase to symptomatic infection phase, or during the symptomatic phase, the pathogen upregulates the production of some effectors (like NEP1, a member of the Necrosis and Ethylene-inducing Peptide 1) that can induce necrosis in the host (Duba et al., 2018, Kettles & Kanyuka, 2016). Recently, a study from M'Barek et al. (2015) identified two putative 'Necrosis-Inducing Proteins' designated as ZtNIP1 and ZtNIP2, but the mechanism behind the symptomatic phase is still not well known. Susceptibility to necrotrophs was found to depend on the presence of dominant sensitivity genes in the host (M'Barek et al., 2015).

3.1.4 Resistance genes against *Z. tritici*

As a consequence of EU restrictions on the use of some of the most potent fungicides and the emergence of resistance phenomena, it is vital for wheat breeding to detect resistance genes against major wheat diseases. Against *Z. tritici*, at least 20 distinct genetic loci have been identified (*Stb* loci) that confer qualitative, often isolate-specific resistance to *Z. tritici* (Kettles & Kanyuka, 2016). Most of them were identified in bread wheat. Qualitative resistance is often preferred in breeding programs because of the large phenotypic effect, which aid accurate and fast selection of resistant progeny (Yang et al., 2018). In addition, a large number of QTLs have been mapped, which make a small but more durable contribution to resistance to this pathogen (Kettles & Kanyuka, 2016). The first genes named for Z. tritici resistance were *Stb*1, *Stb*2 and *Stb*3 and they have been

mapped to chromosome arms 5BL, 1BS and 7AS respectively. *Stb*4 was found in cv. Tadinia and was the first gene to be identified by controlled inoculation with a single isolate of Z. tritici and it was subsequently mapped to chromosome arm 7DS. The first gene to be mapped for Z. tritici resistance was *Stb*5, originated from a highly resistant hexaploid line, Synthetic 6x, derived from Triticum dicoccoides (AABB genomes) and Triticum tauschii (DD genome). *Stb*5 was mapped to the pericentromeric region of chromosome arm 7DS. After *Stb*5, *Stb*8 was identified in another synthetic line, W7984, bred by CIMMYT (the International Maize and Wheat Improvement Centre). It was mapped to the long arm of chromosome 7B and gave resistance to a USA isolate. Two more genes, *Stb*16q and *Stb*17, were discovered in the synthetic hexaploid line M3. *Stb*16q is one of the few major genes that confers resistance in seedling stages and has a significant effect on the variation in the necrotic area of the leaf, the area covered with pycnidae and the latent phase of *Z. tritici*. *Stb*16q was mapped on chromosome 3DL and because it was not possible to determine if there was a single gene at the locus it was designed as a quantitative locus. *Stb*17 is situated on chromosome 5AL and was detected only at the adult plant stage and was less potent than *Stb*16q. Two other important genes are *Stb*7 and *Stb*12. *Stb*7, on chromosome 4AL, was first identified in cultivar ST6 as conferring resistance to the Canadian isolate MG2. *Stb*12 is closely linked to *Stb*7 on chromosome 4AL and gave resistance to isolate *Isr*398 from Israel. Recently, Yang et al. (2018) mapped a new resistance gene, *Stb*19, on the distal region of chromosome 1DS. *Stb*19 confers resistance to three isolates of *Z. tritici* at seedling stages. Furthermore, two KASP® markers were developed and validated and are ready for immediate use (Yang et al., 2018). In addition, other genes have also been identified in common wheat: *Stb*9, *Stb*10, *Stb*11, *Stb*13, *Stb*15 and *Stb*18 (Brown et al., 2015). Regarding quantitative resistance, 89 QTLs of interest have been identified (Karlstedt et al., 2019) that can be combined to give high and durable resistance of strong interest to breeders (Karisto et al., 2018). Although STB is a major problem in durum wheat, *Triticum durum* is more susceptible to STB than *Triticum aestivum*, suggesting on the one hand the adaptation of the strains of *Z. tritici* to durum wheat or a lower genetic resistance of modern varieties in durum wheat compared to soft for narrowing the genetic base of the varieties (Chedli et al., 2018; Brown et al, 2015).

3.2 Materials and methods

3.2.1 The panel

The panel used for this experiment is a subpanel of the CEREALMED landrace panel. The subpanel used in question consists only of durum wheat landraces (around 350) sampled from the Tetraploid Global Collection and from the Global Durum Panel.

3.2.2 Experimental layout and location

This experiment was held in Cadriano (BO) in the experimental fields of the University of Bologna. The experimental design involved the sowing of two randomized replicas inside which there are nine checks repeated eight times in each replica. The checks have been chosen to provide a wide range of susceptibility and resistance and are: Cappelli, Iran1, Karim, Kirmizi, Kunduru, Meridiano, Svevo, Trinakria and Valnova. Each parcel was 1.50m long and 30cm wide.

3.2.3 Inoculation protocol and scoring

The field was prepared by placing sprinklers covering a circumference of 11 meters to provide adequate sprinkler irrigation for the spread of the pathogen. Prior to inoculation, the field was wet for two days to create adequate humidity levels. The inoculum, supplied by UNIBO, was then distributed using a spray bar. Two inoculation passages were carried out four days apart, in the flag leaf phase. Once inoculated, irrigation was suspended, to allow adherence of the inoculum, for one day and then restarted with two irrigation rounds, the first at 17:00 lasting 30 minutes, while the second was carried out at 07:00, always 30 minutes. After 20 and 27 days the surveys of the traits of interest were carried out: Height, Symptoms on the flag leaf and symptoms on the canopy under the flag leaf, following the protocol followed by INRAE BIOGER, a French research institute. Each trait was recorded on a scale of 0 to 9. As regards height, 0 to 5 indicates the presence of the disease at a lower level of the flag leaf, while 6 to 9 indicate various levels on the leaf. flag. For the score of the other two traits, the photos of the various parts of the plant were observed to establish the

average score within the parcel. refer to the picture? to see the difference in symptoms based on the different scores.

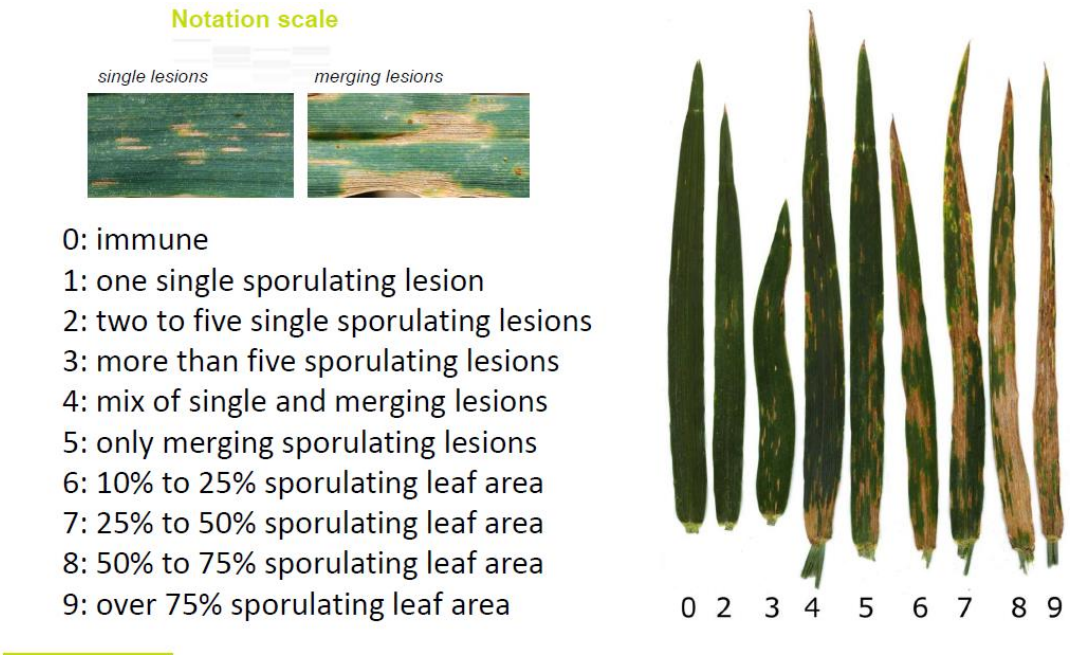


FIGURE 40: EVALUATION SCALE FOR STB INFECTION TYPE

3.2.4 Statistical analysis

For this paragraph, refer to the paragraph already presented for the rust experiment, number 2.2.5

3.3 Results

The first part of the results will focus on the descriptive statistics of the experiment, also showing the various distribution histograms for the most important traits of this analysis, height, infection type at the flag leaf and infection type in the canopy below the flag leaf (infection type flag -1). For the experiment we find descriptive statistics data, including heritability (h2), ANOVA results and distribution graphs for the traits of the blues of the respective phenotypes.

3.3.1 Descriptive statistics, ANOVA and histograms

Table 28 shows the descriptive statistic of this experiment with the relative heritability, remarkably high for HD (0.89) and in any case very high for height (0.71). Heritability is medium but similar as regards the two infection types (0.54 and 0.57 respectively).

	HD	Height	Infection.type.flag.leaf	Infection.type.flag-1
min	30	-0.73	-0.66	-0.76
max	46	9.73	9.66	9.76
range	16	10.47	10.33	10.53
median	40	8.27	4.66	5.76
mean	39.9	6.84	4.16	5.43
SE.mean	0.12	0.1	0.1	0.09
var	9.97	7.42	6.96	5.55
std.dev	3.16	2.72	2.64	2.35
coef.var	0.08	0.4	0.63	0.43
h^2	0.89	0.71	0.54	0.57

TABLE 25: DESCRIPTIVE STATISTIC AND H2 VALUES FOR THE MAIN TRAITS ANALYZED FOR SEPTORIA EXPERIMENT

the ANOVA shows strong between genotypes and between the two replicates for all the traits examined, except HD, which is evidently much more stable within the replicates (Tab. 29).

Trait	Variables	Sum Sq	Df	F value	Pr(>F)	
HD	**Genotype**	7891.5	696	9.663	<2e-16	***
HD	**Rep**	0.5	1	0.4261	0.5163	
Height	**Genotype**	5087.6	677	4.7449	1.35E-11	***
Height	**Rep**	41.4	1	26.1218	3.41E-06	***
Infection Type Flag Leaf	**Genotype**	4821.8	677	2.6156	5.14E-06	***
Infection Type Flag Leaf	**Rep**	36.5	1	13.405	0.000526	***
Infection Type Flag -1	**Genotype**	3824.5	677	3.0795	2.08E-07	***
Infection Type Flag -1	**Rep**	44.9	1	24.4968	6.18E-06	***

TABLE 26: ANOVA RESULTS FOR THE MAIN TRAITS RECORDED IN THE SEPTORIA EXPERIMENT USING AS VARIABLES GENOTYPE AND REP

Let us now examine the distribution histograms. The first is related to height. The blues trend of the histogram is evidently not normal, with the curve flattened towards high values of height, highlighting how on average the disease has reached the flag leaf in most varieties. This is certainly also due to the height of the plants, which caused a strong lodging inside the field (Fig. 40).

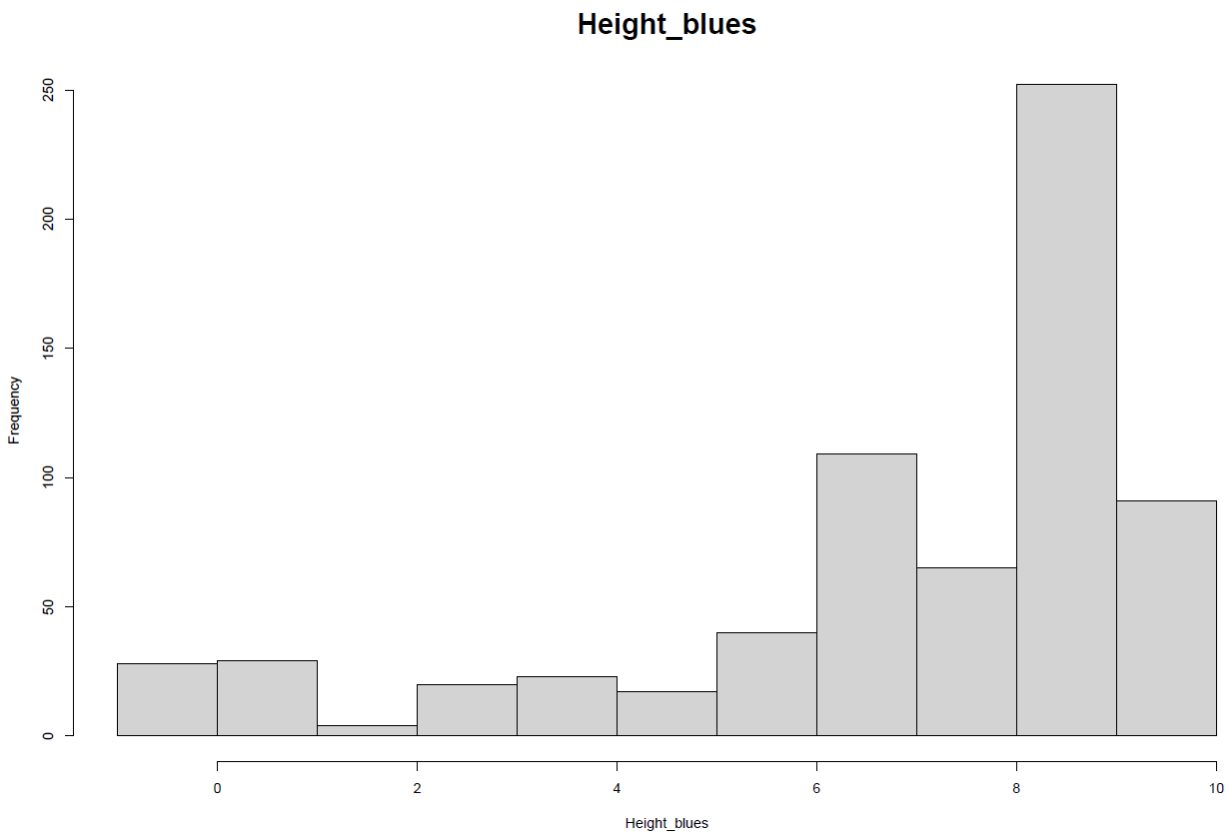


FIGURE 41: DISTRIBUTION HISTOGRAMS FOR BLUES VALUES RELATED TO INFECTION HEIGHT

The histogram referring to the infection on the flag leaf, on the other hand, has a much more normal trend, indicating a wide distribution of symptoms (Fig. 41).

.

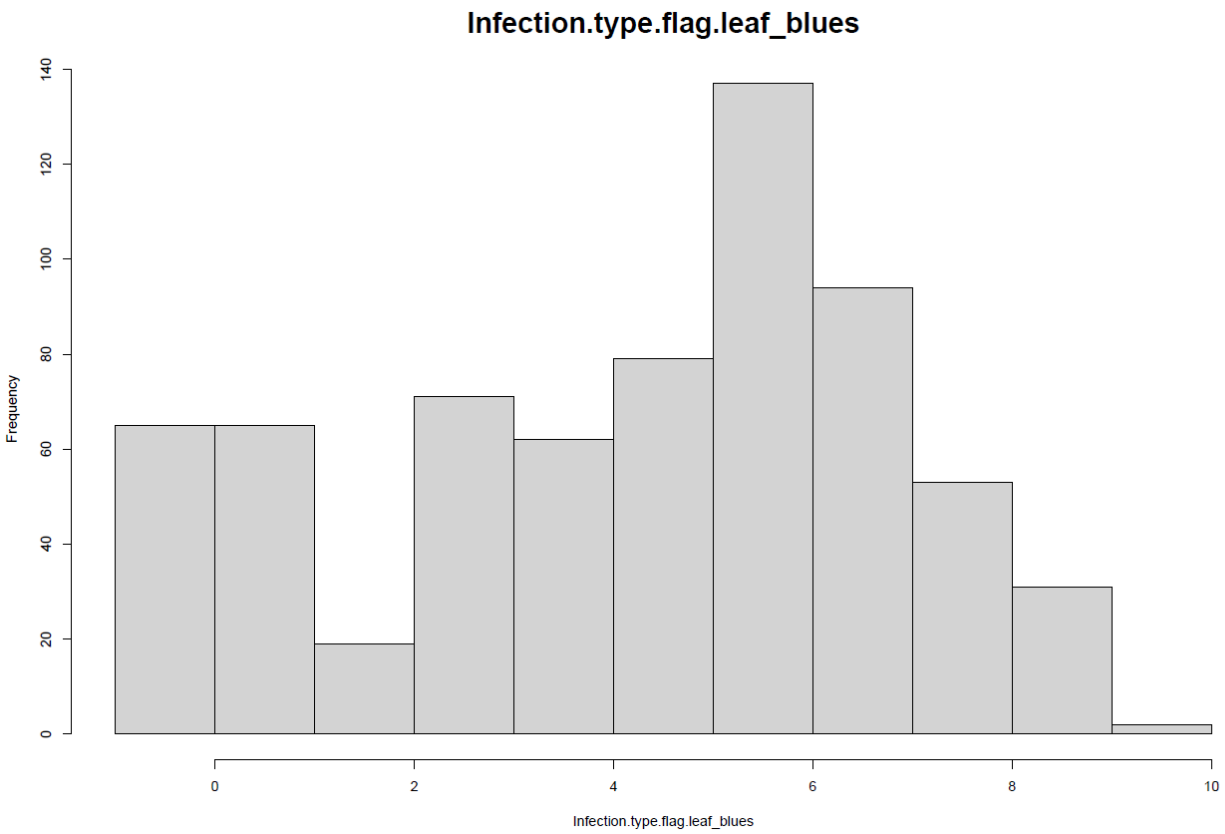


FIGURE 42: DISTRIBUTION HISTOGRAMS FOR BLUES VALUES RELATED TO FLAG LEAF INFECTION

The last histogram, referring to the infection of the canopy under the flag leaf, has a similar trend with the previous one, even if this is slightly more shifted towards high susceptibility values. These two graphs indicate a good distribution of the levels of infection that may be useful for the search for useful genes and to be able to discriminate intermediate levels of resistance (Fig. 42).

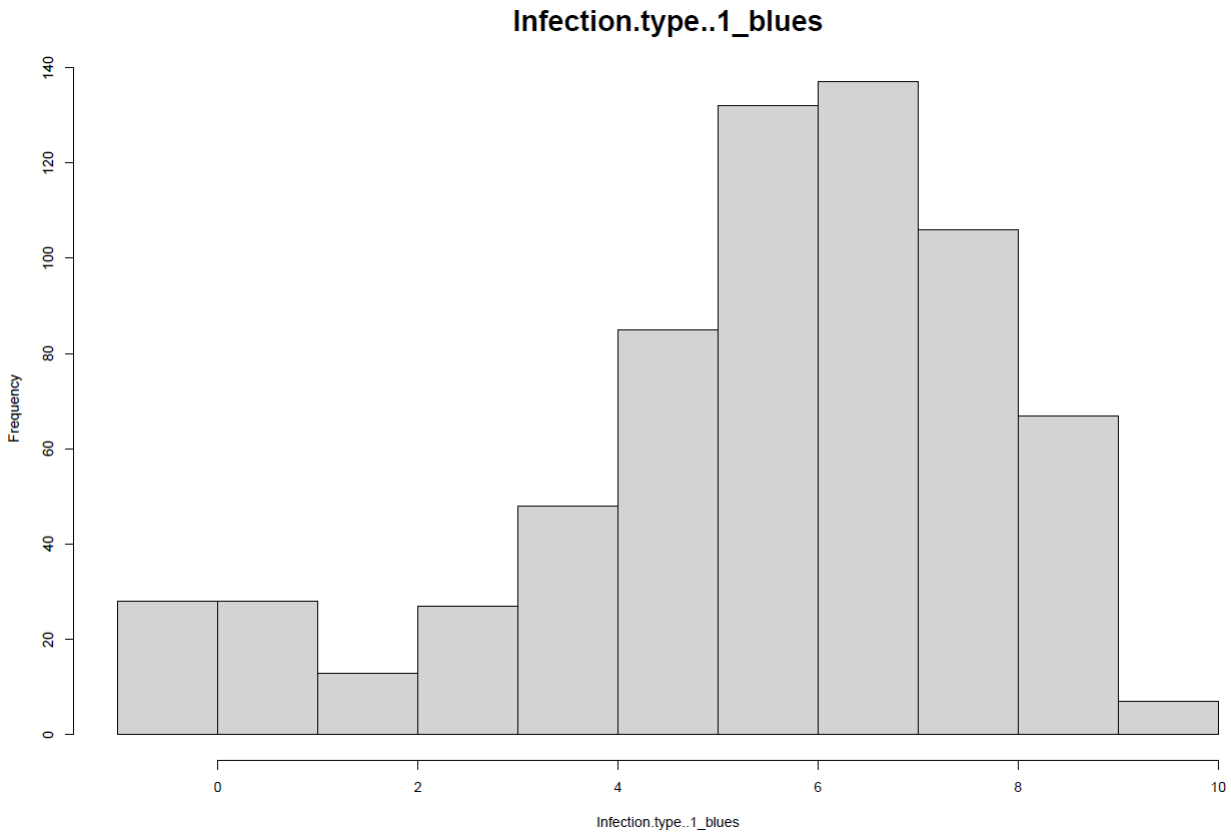


FIGURE 43: DISTRIBUTION HISTOGRAMS FOR BLUES VALUES RELATED TO CANOPY INFECTION

3.4 Discussion

The experiment held for STB has the same objectives, in terms of analysis, as the one done for yellow rust. At present the experiment needs to be repeated in order to have a solid dataset on which to carry out a GWAS and possible individual QTLs. The analyzes carried out so far have shown how the experiment, however, gave the desired results to be able to be continued. The inoculum has shown efficacy, in fact looking at the distribution graphs we can see a wide spectrum of results that manages to highlight resistant genotypes, partially resistant genotypes and susceptible genotypes. Obtaining this data bodes well to be able to identify both resistance and partial resistance genes, which can be pyramidal to counter this dangerous wheat disease. Ultimately, this experiment needs to be traced back so that we can continue the analysis to perform solid GWAS that can allow the identification of resistance loci that can be investigated to open the doors to new sources of resistance.

4. Chapter III: Fusarium Head Blight (FHB)

4.1 Introduction

4.1.1 *Fusarium graminearum*

The fungal pathogen *Fusarium graminearum* Schwabe [teleomorph *Gibberella zeae* (Schweinitz) Petch], is the most common causal agent of Fusarium head blight (FHB) in many parts of the world and it affects several cereals as wheat, barley and other small grains both in temperate and in semitropical areas. Host resistance to FHB is composed of multiple mechanisms:

- Resistance to initial infection (type I)
- Disease spread (type II)
- Toxin accumulation (type III)
- Kernel infection (type IV)
- Yield loss (type V)

Of which the last three have been less studied (Gong et al., 2020). FHB causes several types of damage at market level. The most evident damage is caused by a diffuse sterility in the ear, with the formation of discoloured, withered and light test-weight kernels causing problems related to the marketability of the seed (Goswami and Kistler, 2004). Additionally, infected grains may contain significant levels of trichothecenes toxins such as deoxynivalenol (DON), commonly known as vomitoxin, that cause several problems for human and animal health because they persist during storage and are heat resistant (Leplat et al., 2012). Acute adverse effects of the toxin in animals include food refusal, emesis, diarrhoea, alimentary hemorrhaging and contact dermatitis (Bennett and Klich, 2003). In humans, *F. graminearum* has been linked to alimentary toxic aleukia and Akakabi toxicosis, illnesses characterized by nausea, vomiting, anorexia and convulsions (Goswami and Kistler, 2004). There have been several outbreaks of this disease in recent years, mainly due to warmer temperatures due to climate change. These temperatures, connected to good levels of humidity and rainfall, allow a greater proliferation of the pathogen. A single control strategy against this disease has not proved effective, showing the need to combine more than one, especially genetic resistance. Genetic resistance represents the most sustainable form of resistance, for a more sustainable form of agriculture. At present, about 52 QTLs have been identified for resistance

to FHB, present in more or less the entire genome with the exception of chromosome 7D (Dweba et al., 2016). To date, Fhb1 is the most stable and effective resistant gene against FHB. This gene was initially found on landrace Sumai 3, showing a significantly lower percentage of infected spikelets and FDK than other genotypes without this resistance gene (Gong et al., 2020).

4.2 Materials and methods

4.2.1 The panels

Field data from two different panels were collected for this experiment. The first is a panel of durum wheat made up of varieties tested at the University of BOKU, whose experiment was replicated in Cadriano (BO) in 2021, in the experimental field of the University of Bologna. The panel is composed by durum modern cultivar, durum landraces, emmer accessions and dicoccoides selected from GDP and TGC for a total of 154 accessions.

ID	Trait	ID	Trait	ID	Trait	ID	Trait
100DP2	tall	GDPv1-562	tall	GDPv1-035	modern semidwarf	DGE-1	modern semidwarf
136DP2	tall	GDPv1-567	tall	GDPv1-223	modern semidwarf	DUPRI	modern semidwarf
148DP2	tall	GDPv1-583	tall	GDPv1-229	modern semidwarf	DYLAN	modern semidwarf
154DP2	tall	GDPv1-606	tall	GDPv1-243	modern semidwarf	EDMORE	modern semidwarf
165DP2	tall	GDPv1-613	tall	GDPv1-316	modern semidwarf	FURIO_CAMILLO	modern semidwarf
49DP1	tall	GDPv1-623	tall	GDPv1-352	modern semidwarf	GIBRALTAR	modern semidwarf
54DP1	tall	GDPv1-625	tall	GDPv1-362	modern semidwarf	GRAZIA	modern semidwarf
Aus26347	tall	GDPv1-658	tall	GDPv1-363	modern semidwarf	HERAKLION	modern semidwarf
Aus26350	tall	GDPv1-666	tall	GDPv1-364	modern semidwarf	IDEFIX	modern semidwarf
Aus26352	tall	GDPv1-669	tall	GDPv1-385	modern semidwarf	IRIDE	modern semidwarf
Aus26430	tall	GDPv1-692	tall	GDPv1-386	modern semidwarf	ISILDUR	modern semidwarf
Aus26450	tall	GDPv1-721	tall	GDPv1-388	modern semidwarf	JOYAU	modern semidwarf
Aus26457	tall	GDPv1-725	tall	GDPv1-391	modern semidwarf	KARIM	modern semidwarf
Aus26519	tall	GDPv1-728	tall	GDPv1-394	modern semidwarf	KARUR	modern semidwarf
Aus26533	tall	GDPv1-734	tall	GDPv1-399	modern semidwarf	KOFA	modern semidwarf
Aus26543	tall	GDPv1-735	tall	GDPv1-403	modern semidwarf	KRONOS	modern semidwarf
Aus26544	tall	GDPv1-755	tall	GDPv1-460	modern semidwarf	KUBANKA	modern semidwarf
Aus26551	tall	L37	tall	GDPv1-516	modern semidwarf	KUNDURU	modern semidwarf
Aus26572	tall	LD108	tall	GDPv1-531	modern semidwarf	LEVANTE	modern semidwarf
Aus26602	tall	LD178	tall	GDPv1-543	modern semidwarf	LLOYD	modern semidwarf
Aus26609	tall	PI94749	tall	GDPv1-544	modern semidwarf	MERIDIANO	modern semidwarf
Aus26708	tall	SSD_118	tall	AC_MORSE	modern semidwarf	MINDUM	modern semidwarf
Aus27250	tall	SSD_122	tall	ACHILLE	modern semidwarf	MONASTIR	modern semidwarf
Aus28778	tall	SSD_303	tall	ANCO_MARZIO	modern semidwarf	NEODUR	modern semidwarf
Aus28784	tall	SSD_335	tall	ANTALIS	modern semidwarf	OBELIX	modern semidwarf
Aus28785	tall	SSD_346	tall	ASTERIX	modern semidwarf	OVIDIO	modern semidwarf
Aus28791	tall	TC4	tall	ATIL	modern semidwarf	PESCADOU	modern semidwarf
DIC-unibo-136	tall	TC9	tall	AWALI-1	modern semidwarf	PR22D89	modern semidwarf
GDPv1-558	tall	TDS217	tall	BABYLONE	modern semidwarf	PROVENZAL	modern semidwarf
		TDS219	tall	BELZER	modern semidwarf	RASCON/TARRO	modern semidwarf
		TDS228	tall	BIENSUR	modern semidwarf	RGT_VOILUR	modern semidwarf
		TDS229	tall	BRAVADUR	modern semidwarf	RUSSELLO_SG7	modern semidwarf
		TDS333	tall	CAPPELLI	modern semidwarf	SERAFO_NICK	modern semidwarf
		tetra-IPK322	tall	CEEDUR	modern semidwarf	SHABHA	modern semidwarf
		Xu037	tall	CESARE	modern semidwarf	SY_CISCO	modern semidwarf
		TC3	tall	CLAUDIO	modern semidwarf	SVEVO	modern semidwarf
		TC5	tall	COLOMBO	modern semidwarf	SUMAI-3	modern semidwarf
		DIC126B	tall	COLOSSEO	modern semidwarf	TEODORICO	modern semidwarf
		TDS211	tall	DAURUR	modern semidwarf	TITO_FLAVIO	modern semidwarf
		TDS231	tall			TIZIANA	modern semidwarf
		TDS235	tall			WB_TURBO	modern semidwarf
		TDS239	tall			ZARDAK	modern semidwarf
		TDS310	tall				
		TDS314	tall				

FIGURE 44: ACCESSIONS LIST OF DURUM WHEAT PANEL USED FUSARIUM EXPERIMENT

The other panel is made up of 243 modern varieties of bread wheat belonging to different breeders who have joined the INNOVAR project for the constitution of this panel.

4.2.2 Experimental layout and locations

The durum wheat panel was sown in 2020 to be phenotyped in 2021 in Cadriano (BO) in the experimental field of the University of Bologna. This panel was seeded in two randomized replicates with replicated checks to have more solid results. Each of the two replicas had one side of the field made up of tall plants (mainly landraces and dicoccoides) while the other was made up of low, mostly modern cultivars. Each replicas had three checks (Normanno, Simeto and Saragolla) repeated eight times. For this experiment, bines with a length of 1.50 m and a width of about 20 cm were sown. The Innovar panel was evaluated over 2 different years in different locations. In 2021 it was sown in Malalbergo (BO) in collaboration with SIS - Società Italiana Sementi, with bines of 1 m length and 20 cm width. In 2022 this panel was sown in three different locations: Cadriano (BO), in the experimental fields of the University of Bologna (length of 1.50 m and a width of about 20 cm), in Malalbergo (BO), again in the fields in collaboration with SIS (bines of 1 m length and 20 cm width), and in Ravenna, in the experimental farm Cà Bosco (plots of 2x1 m), in collaboration with Horta srl. In 2021, the INNOVAR panel was seeded with an augmented block design, with repeated checks within each genotype block created at the time of field design, using six different checks: PR22R58, Altamira, Blasco, Aubusson, Bologna and Ilaria. Checks were selected based on recorded phenotype reaction to FHB, to have resistant, partial resistant and susceptible checks. In 2022, the same experimental design was used for Malalbergo and Ravenna, while two randomized replicas of the entire panel were sown in Cadriano, using the same checks.

FIGURE 45: FUSARIUM EXPERIMENTAL FIELD IN CADRIANO (BO)

4.2.3 Inoculation protocol and scoring

Each field was managed following the same inoculation protocol:

- The genotypes were divided into three (INNOVAR panel) or four (durum wheat panel) different groups, each characterized by a different flowering time to be able to take the moment of maximum susceptibility of the spike.

- Each group was inoculated twice with *Fusarium graminearum* (concentration of the spore 10^6), based on its flowering time with 4 days between the first and the second inoculation. The inoculum was provided by UNIBO and distributed with a spray bar to allow a uniform distribution.

- Irrigation was performed every day from booting stage (Zadoks 41) until late milk stage (Zadoks 77) during nighttime every three hours, starting from 06:00 pm until 06:00 am.

Irrigation was carried out using sprinklers that could recreate natural rain conditions as much as possible, to favor the conditions for the spread and growth of the pathogen.

- The scoring was performed observing every accession plot and it was recorded as a double evaluation with two indexes: fusarium incidence and fusarium severity.

- Fusarium incidence is expressed as the average percentage of infected spikes in the plot, while fusarium severity as the average percentage of infected spikelet in each spike. Severity was scored using this scale: 0, 2, 5, 10, 25, 50, 75 or 90% of infected spikelet.

- The incidence and severity data were used to calculate the index ((incidence * severity)/ 100) and the AUDPC, already calculated also for the yellow rust experiment.

- Two scores has been performed, one after 20 days from the first inoculation (Zadoks 77 late milk) and the second after 27 days from the first inoculation (Zadoks 83-85 soft dough stage).

- The central part of the plots was collected to allow analysis to be carried out to evaluate the quantity of DON.

FIGURE 46: FHB SYMPTOMPS INTO THE FIELD

For each experiment, HD was also taken, to be used both for the division into groups and as a potential covariate / data within the analysis. HD was recorded as days starting from April 1.

4.3 Results

The first part of the results will focus on the descriptive statistics for each experiment, also showing the various distribution histograms for the most important characters of this analysis, incidence and SEV of the last and most significant score. For each experiment we find descriptive statistics data, including heritability (h^2), ANOVA results and distribution graphs for IT and SEV of the blues of the respective phenotypes.

4.3.1 Fusarium durum panel

	Incidence	Severity	Index	Inc_AUDPC	Sev_AUDPC	Ind_AUDPC	Inc_RAUDPC	Sev_RAUDPC	Ind_RAUDPC
min	0	0	0	0	0	0	0	0	0
max	100.29	80.8	80.8	705.02	521.85	521.85	1.01	0.9	0.9
range	100.29	80.8	80.8	705.02	521.85	521.85	1.01	0.9	0.9
median	100	47.82	47.82	686	279.8	270.25	1	0.5	0.5
mean	97.23	46.55	46.24	663.18	276.2	270.55	0.95	0.49	0.48
SE.mean	1.1	1.43	1.45	7.97	9.27	9.45	0.01	0.02	0.02
var	162.9	275.15	284.65	8567.59	11606.54	12055.56	0.02	0.04	0.04
std.dev	12.76	16.59	16.87	92.56	107.73	109.8	0.13	0.19	0.19
coef.var	0.13	0.36	0.36	0.14	0.39	0.41	0.14	0.39	0.41
h^2	0.98	0.87	0.87	0.97	0.87	0.87	0.95	0.87	0.87

TABLE 27: DESCRIPTIVE STATISTIC OF THE FUSARIUM DURUM PANEL FOR ALL THE TRAITS RECORDED WITH THE RELATED H^2

Regarding this experiment, we can note a very high heritability for all characters, especially regarding incidence. This means that in the two replicates the same genotypes showed the same predisposition to infection. the higher heritability is the one we can see on the incidence data (0.98), while for the other data, the lower heritability is equal to 0.87 for the severity and index data (Tab. 30).

Trait	Variables	Sum Sq	Df	F value	Pr(>F)	
Incidence	**Genotype**	43440	134	61.6735	<0.0000000000000002	***
Incidence	**Rep**	34	1	6.3877	0.01241	*
Severity	**Genotype**	80260	134	7.8032	<0.0000000000000002	***
Severity	**Rep**	236	1	3.0811	0.08102	.
Index	**Genotype**	82636	134	7.8973	<2e-16	***
Index	**Rep**	183	1	2.3469	0.1274	
Inc_AUDPC	**Genotype**	2269309	134	36.902	<0.0000000000000002	***
Inc_AUDPC	**Rep**	8255	1	17.989	3.65E-05	***
Sev_AUDPC	**Genotype**	3183863	134	7.794	<2e-16	***
Sev_AUDPC	**Rep**	4935	1	1.6189	0.205	
Ind_AUDPC	**Genotype**	3281440	134	7.874	<2e-16	***
Ind_AUDPC	**Rep**	2217	1	0.713	0.3996	
Inc_RAUDPC	**Genotype**	4.7763	134	22.567	<0.00000000000000022	***
Inc_RAUDPC	**Rep**	0.0247	1	15.658	0.0001116	***
Sev_RAUDPC	**Genotype**	10.1169	134	7.8178	<2e-16	***
Sev_RAUDPC	**Rep**	0.0123	1	1.2718	0.261	
Ind_RAUDPC	**Genotype**	10.2363	134	7.8221	<2e-16	***
Ind_RAUDPC	**Rep**	0.0046	1	0.4667	0.4955	

TABLE 28: ANOVA RESULTS FOR FUSARIUM DURUM PANEL USING GENOTYPE AND REP AS VARIABLES

ANOVA (Tab.31) showed high significant relationships among accessions for all traits, and also a high significant correlation between reps for Inc_AUDPC and Inc_RAUDPC, while the correlation was lower, but still significant for incidence and severity traits.

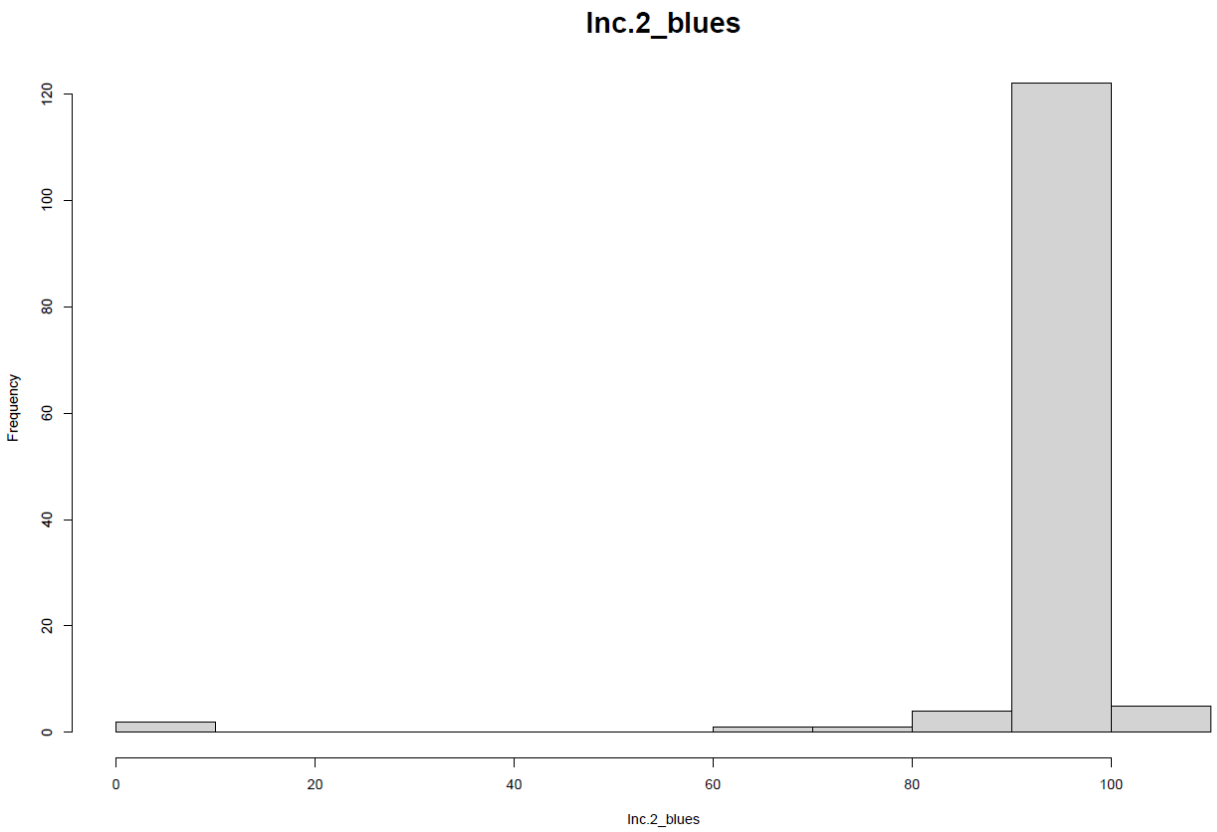


FIGURE 47: DISTRIBUTION HISTOGRAMS RELATED TO FUSARIUM INCIDENCE IN CADRIANO (BO)

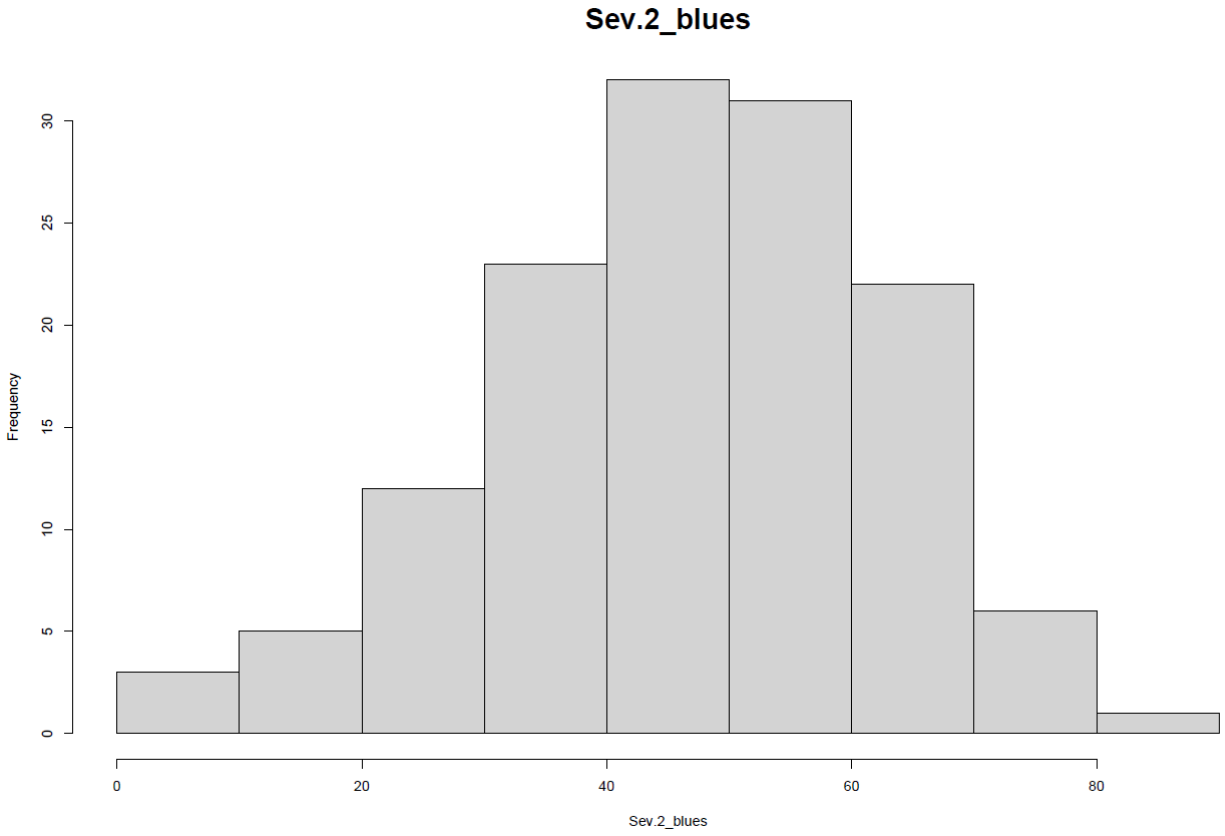


FIGURE 48: DISTRIBUTION HISTOGRAMS RELATED TO FUSARIUM SEVERITY IN CADRIANO (BO)

Looking at the distribution histograms we can see how the incidence is very high within the panel, with a non-normal distribution. This is because practically all plots have a very high incidence. We can see a single genotype that has zero incidence, and it is Sumai 3, the bread wheat most resistant check. Within this panel of hard wheat, this single genotype of common wheat has also been placed in order to compare its reaction. As far as severity is concerned, the results are different. The histogram shows an almost normal trend, indicating that although the incidence is very high within our plots, the severity in the different genotypes is very wide.

4.3.2 INNOVAR Panel Cadriano 2022 (bread wheat)

	Inc	Sev	Index	Inc_AUDPC	Sev_AUDPC	Index_AUDPC	Inc_RAUDPC	Sev_RAUDPC	Index_RAUDPC
min	-18.24	-27.71	-22.38	-111.58	-20.85	-112.77	-0.17	-0.02	-0.26
max	72.08	88.45	71.17	582.98	873.55	480.65	0.67	0.76	0.81
range	90.31	116.17	93.54	694.57	894.39	593.42	0.84	0.78	1.07
median	29.06	45.06	11.56	174.97	302.19	78	0.19	0.24	0.13
mean	27.98	43.39	14.81	182.63	297.23	96.34	0.2	0.23	0.16
SE.mean	1.59	1.63	1.63	7.58	9.4	7.06	0.01	0.01	0.01
var	633.59	672.58	670.33	14467.23	22288.74	12563.47	0.02	0.02	0.04
std.dev	25.17	25.93	25.89	120.28	149.29	112.09	0.15	0.14	0.2
coef.var	0.9	0.6	1.75	0.66	0.5	1.16	0.73	0.59	1.25
h^2	0.86	0.84	0.86	0.80	0.87	0.84	0.79	0.84	0.84

TABLE 29: DESCRIPTIVE STATISTIC OF THE INNOVAR PANEL IN CADRIANO (BO) FOR ALL THE TRAITS RECORDED WITH THE RELATED H^2

In this case the heritability is slightly lower than the previous experiment, albeit still having very high values. In this case the highest heritability value is that of the AUDPC severity (0.87), linked to other heritability values of 0.84 (RAUDPC severity and severity). the incidence has a value of 0.86 which goes down slightly for the AUDPC (0.80) and RAUDPC (0.79), which represent the lowest values of heritability for this experiment. The index has heritability values ranging from 0.86 (index) to 0.84 (AUDPC and RAUDPC) (Tab. 32). Heading data was also carefully taken for the INNOVAR experiments due to its strong correlation with FHB. For this we went to see the correlation with the phenotypic data taken regarding FHB, to see if we can use HD as a covariate within our analyzes. The correlation graph shows an inversely proportional correlation between HD and several of the data taken into

consideration (Fig. 47), due to this HD was used as a covariate within our ANOVA and pipeline to obtain blues and blups.

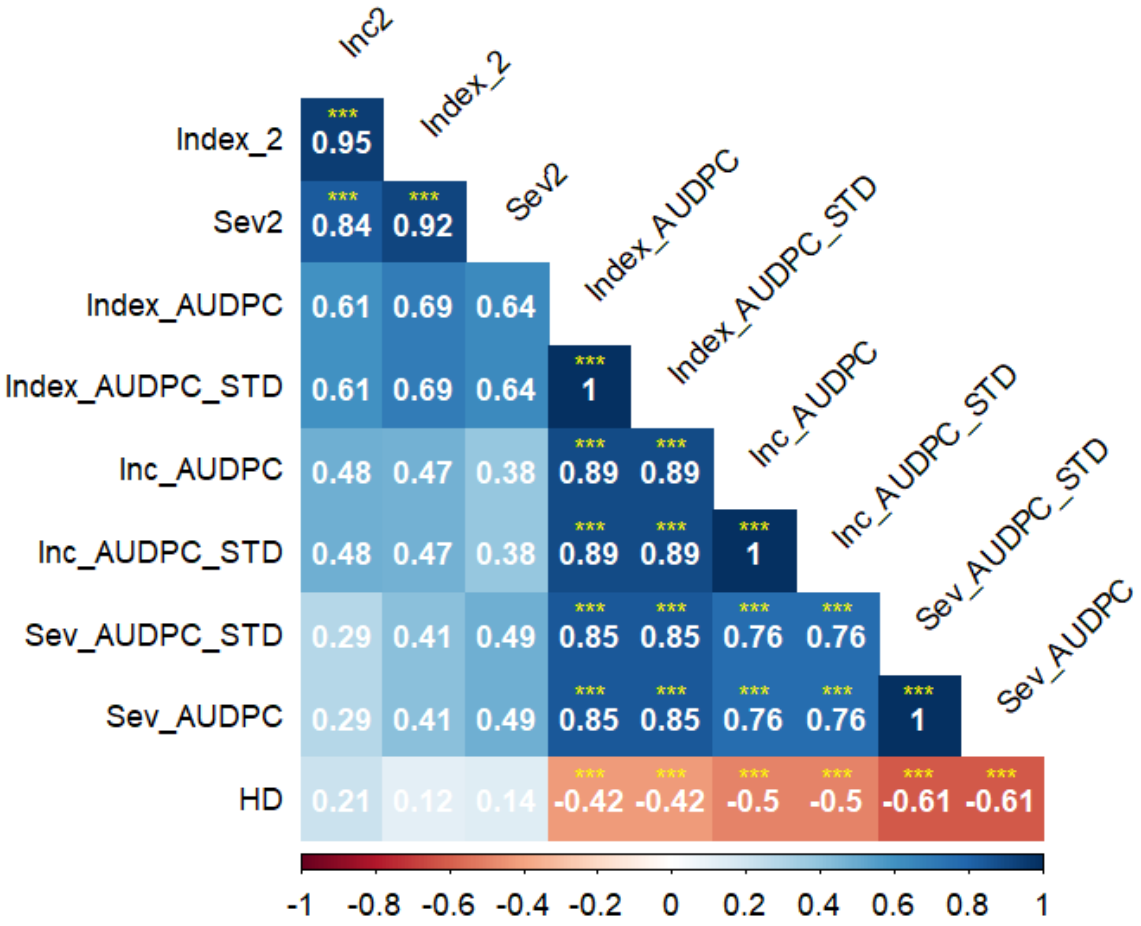


FIGURE 49: CORRELOGRAM FOR THE INNOVAR PANEL TESTED IN CADRIANO (BO) FOR ALL THE TRAITS OF INTEREST

The correlogram shows also high levels of correlation between index, severity and incidence. These three traits have a less significant correlation with the AUDPC and RAUDPC (AUDPC_STD in the correlogram), on the other hand all the AUDPC traits are highly correlated among them.

Trait	Variables	Sum Sq	Df	F value	Pr(>F)	
Incidence	**Genotype**	216213	250	4.2693	<2e-16	***
Incidence	**Rep**	22	1	0.1068	0.7441	
Severity	**Genotype**	204811	250	4.668	<0.0000000000000002	***
Severity	**Rep**	974	1	5.5516	0.01945	*
Index	**Genotype**	232016	250	4.7746	<2e-16	***
Index	**Rep**	158	1	0.8133	0.3683	
Inc_AUDPC	**Genotype**	4831304	250	3.0564	1.23E-15	***
Inc_AUDPC	**Rep**	649	1	0.1027	0.749	
Sev_AUDPC	**Genotype**	6208697	250	3.9247	<0.00000000000000022	***
Sev_AUDPC	**Rep**	184262	1	29.119	1.94E-07	***
Ind_AUDPC	**Genotype**	4806157	250	4.4707	<0.0000000000000002	***
Ind_AUDPC	**Rep**	30608	1	7.1179	0.008267	**
Inc_RAUDPC	**Genotype**	7.693	250	3.0384	1.69E-15	***
Inc_RAUDPC	**Rep**	0.0004	1	0.0388	0.8441	
Sev_RAUDPC	**Genotype**	5.4118	250	3.2239	<0.00000000000000022	***
Sev_RAUDPC	**Rep**	0.1893	1	28.1983	2.94E-07	***
Ind_RAUDPC	**Genotype**	14.0986	250	4.3433	<0.00000000000000022	***
Ind_RAUDPC	**Rep**	0.1121	1	8.6371	0.003687	**

TABLE 30: ANOVA RESULTS FOR THE MAIN TRAITS RECORDED IN CADRIANO (BO) USING AS VARIABLES GENOTYPE AND REP

ANOVA (Tab. 33) showed high significant relationships among accessions for all traits, and also a high significant correlation between reps for Sev_RAUDPC trait, while the correlation was lower, but still significant for Ind_RAUDPC, Ind_AUDPC and severity.

As for the distribution graphs, in this case for the incidence (Fig. 48) we have a much more normal distribution than the previous experiment and a worse distributed severity (Fig. 49), with severity levels more shifted towards higher infection values. This panel, being made up of modern varieties only, could demonstrate how modern varieties tend to be more susceptible overall than a mixed panel of cultivars and landraces.

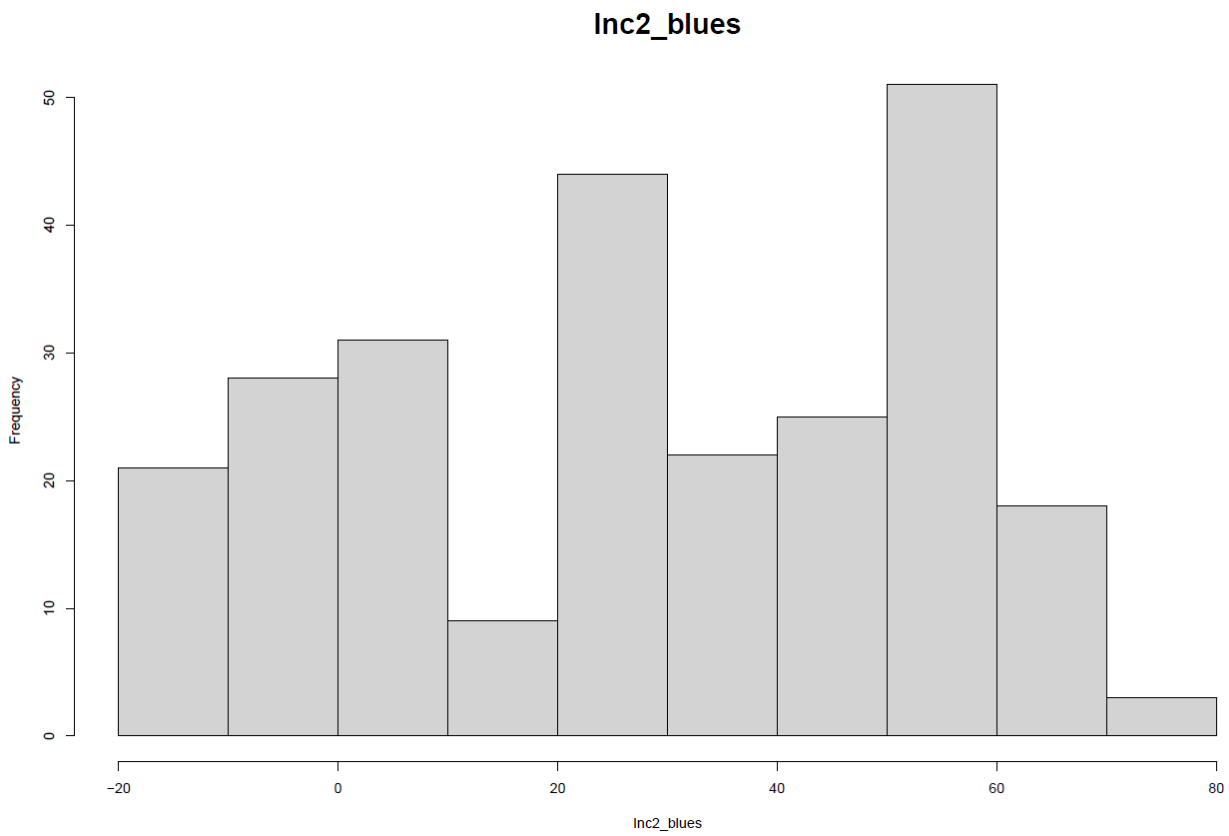


FIGURE 50: DISTRIBUTION HISTOGRAMS RELATED TO FUSARIUM INCIDENCE IN CADRIANO (BO)

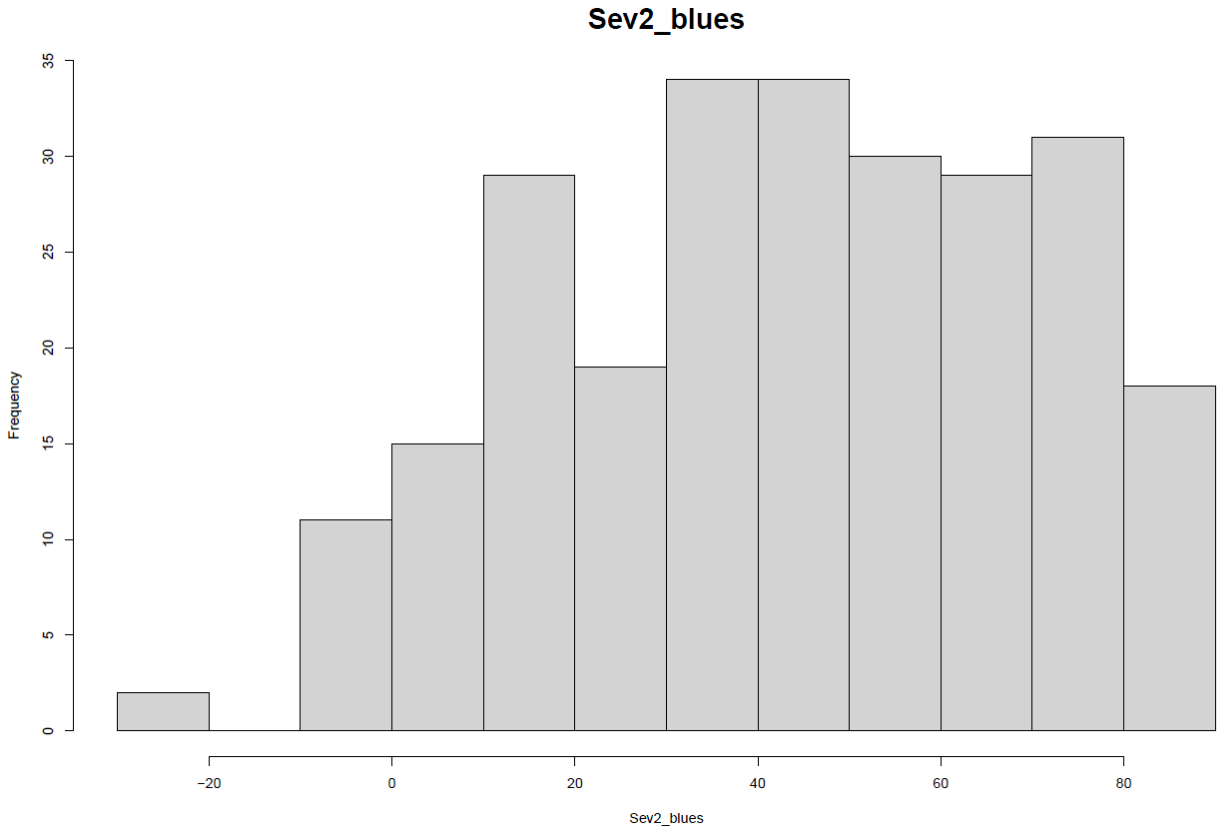


FIGURE 51: DISTRIBUTION HISTOGRAMS RELATED TO FUSARIUM SEVERITY IN CADRIANO (BO)

4.3.3 Innovar Panel Malalbergo (SIS)

	Inc	Sev	Index	Inc_AUDPC	Sev_AUDPC	Ind_AUDPC	Inc_RAUDPC	Sev_RAUDPC	Ind_RAUDPC
min	8.1	-1.08	-4.93	59.4	-61.75	-68.69	-0.09	-0.12	-0.16
max	167.31	95.97	83.75	1258.63	835.03	793.81	1.2	1.1	1.1
range	159.21	97.05	88.68	1199.23	896.78	862.49	1.29	1.21	1.26
median	71.26	29.92	20.58	610	271.46	198.4	0.5	0.32	0.25
mean	71.06	36.19	27.2	613.38	295.41	221.33	0.54	0.36	0.3
SE.mean	1.96	1.55	1.47	16.71	10.95	10.87	0.02	0.01	0.02
var	854.57	532.71	481.84	61953.19	26641.7	26222.04	0.07	0.05	0.05
std.dev	29.23	23.08	21.95	248.9	163.22	161.93	0.26	0.22	0.23
coef.var	0.41	0.64	0.81	0.41	0.55	0.73	0.48	0.62	0.77
h^2	0.53	0.55	0.51	0.33	0.38	0.35	0.25	0.36	0.31

TABLE 31: DESCRIPTIVE STATISTIC OF THE INNOVAR PANEL IN MALALBERGO (BO) FOR ALL THE TRAITS RECORDED WITH THE RELATED H^2

From the test held in Malalbergo it is evident that the heritability values are considerably lower (Tab. 34), showing differences within the checks blocks. Heritability has slightly higher values for the data relating to incidence, severity and index, and then significantly lower for the relative AUPDC and RAUDPC. Incidence has a heritability value of 0.53, which for AUDPC and RAUDPC drops to 0.33 and 0.25 respectively. For the severity the heritability is 0.55 which is lowered to 0.38 and to 0.36 for AUDPC and RAUDPC. Finally, for the index the value of h2 is 0.51, to be lowered to 0.35 for the AUDPC and to 0.31 for the RAUDPC. As for the correlations, in this case we have a significant correlation for all characters except for the last severity measurement which has no significance with the incidence measurements, with the first severity measurement, with the index of the first. relief and the related AUDPC. The HD shows a negative correlation with all characters except for the last index and the last severity, which is why it was used as a covariate of the analysis (Fig. 50).

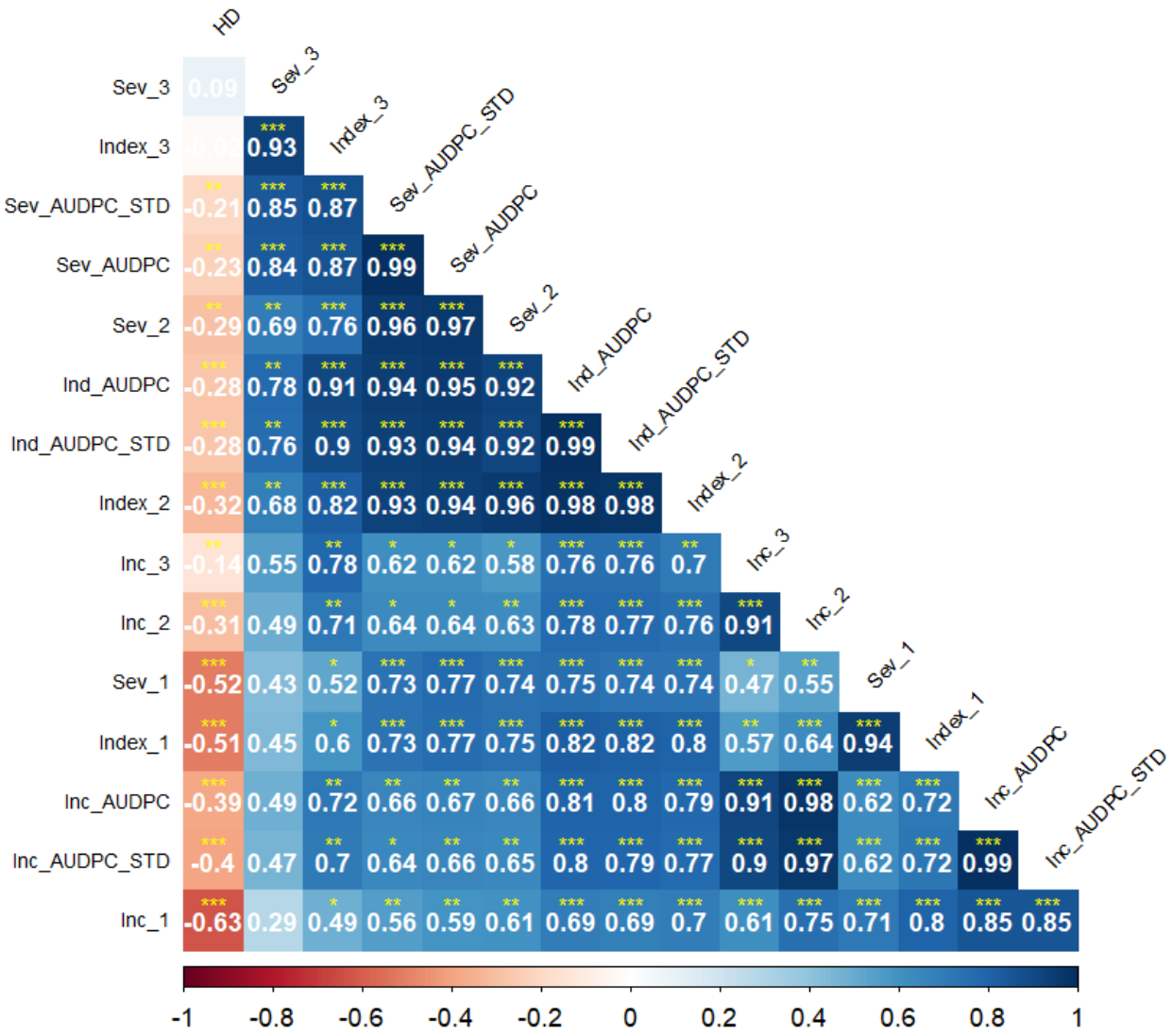


FIGURE 52: CORRELOGRAM FOR THE INNOVAR PANEL TESTED IN MALALBERGO (BO) FOR ALL THE TRAITS OF INTEREST

The ANOVA (Tab. 35) in this case shows no significance for the variables used to correct the raw data, if not a slight significance between genotypes as regards the incidence.

Trait	Variables	Sum Sq	Df	F value	Pr(>F)
Incidence	**Genotype**	80375	187	2.4491	0.05847 .
Incidence	**Block**	546	2	1.5565	0.25793
Severity	**Genotype**	70100	187	1.6707	0.1861
Severity	**Block**	1033	2	2.3012	0.1506
Index	**Genotype**	62279	187	1.534	0.2323
Index	**Block**	960	2	2.2104	0.1603
Inc_AUDPC	**Genotype**	8179691	187	0.9995	0.5554
Inc_AUDPC	**Block**	57735	2	0.6596	0.5382
Sev_AUDPC	**Genotype**	5272791	187	0.9537	0.5954
Sev_AUDPC	**Block**	69574	2	1.1766	0.3476
Ind_AUDPC	**Genotype**	4559554	187	1.0439	0.5183
Ind_AUDPC	**Block**	65227	2	1.3962	0.2919
Inc_RAUDPC	**Genotype**	8.2759	187	0.978	0.5739
Inc_RAUDPC	**Block**	0.0694	2	0.7674	0.4897
Sev_RAUDPC	**Genotype**	9.7598	187	0.8662	0.6756
Sev_RAUDPC	**Block**	0.1259	2	1.0447	0.3872
Ind_RAUDPC	**Genotype**	9.6899	187	1.0131	0.5439
Ind_RAUDPC	**Block**	0.179	2	1.7497	0.2231

TABLE 32: ANOVA RESULTS FOR THE MAIN TRAITS RECORDED IN MALALBERGO (BO) USING AS VARIABLES GENOTYPE AND BLOCK

The distribution of histograms shows a normal trend as regards incidence (Fig. 51), while for severity (Fig. 52) the trend is more squeezed towards low values, indicating a lower distribution of the disease within the field.

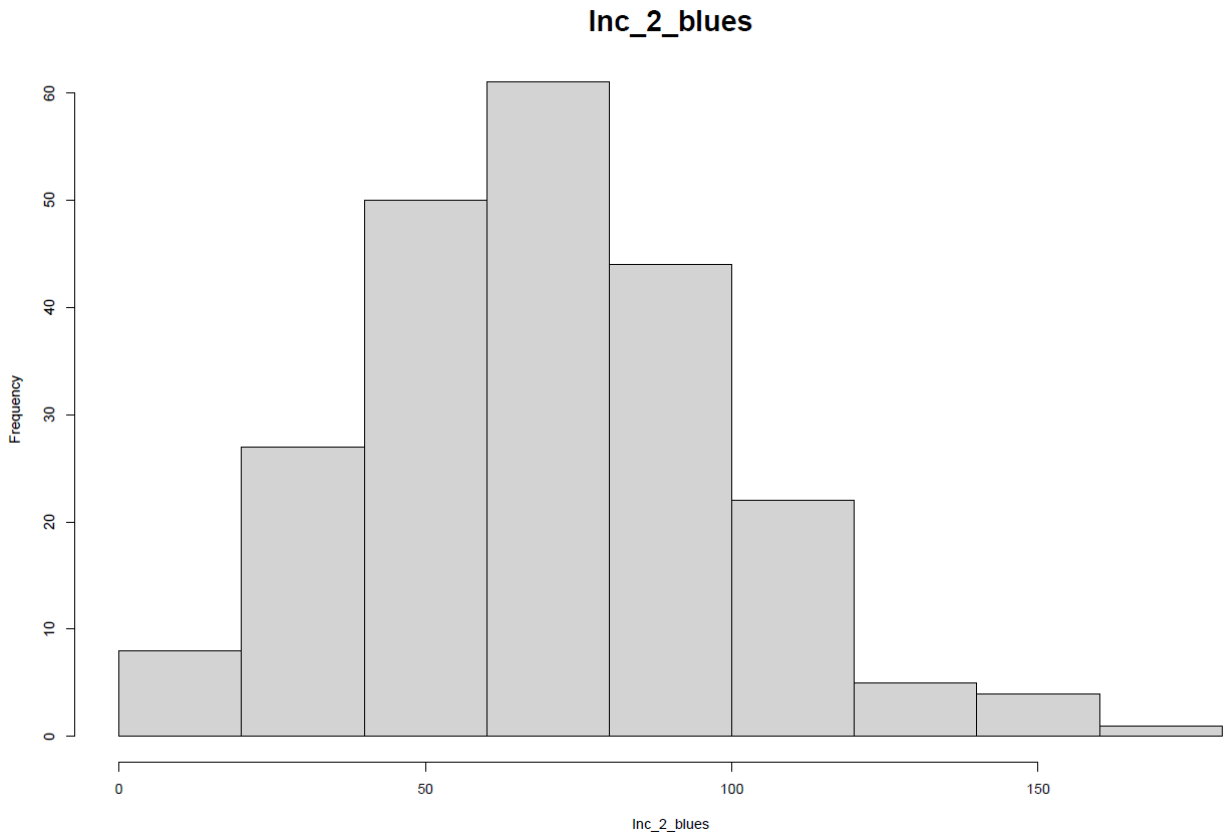


FIGURE 53: DISTRIBUTION HISTOGRAMS RELATED TO FUSARIUM INCIDENCE IN MALALBERGO (BO)

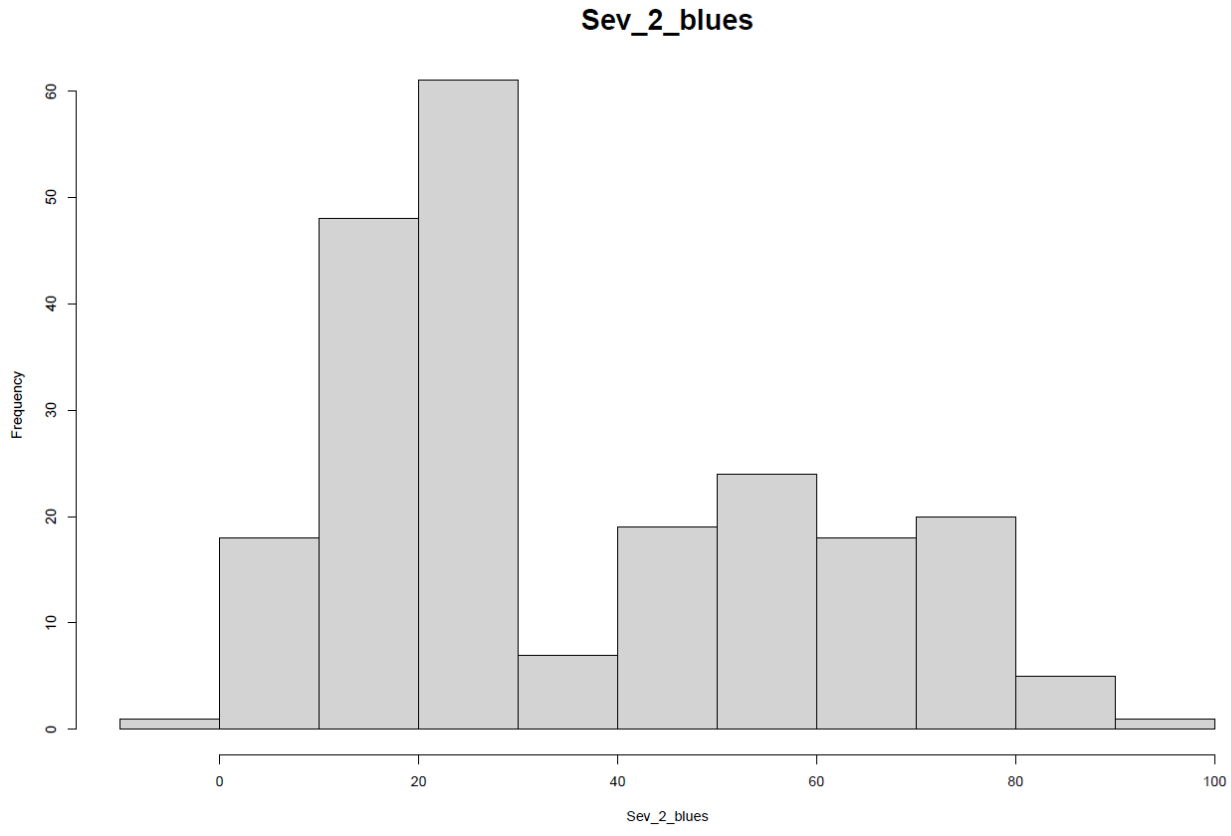


FIGURE 54: DISTRIBUTION HISTOGRAMS RELATED TO FUSARIUM INCIDENCE IN MALALBERGO (BO)

4.3.4 INNOVAR Panel Ravenna (HORTA)

	Inc	Sev	Index	Inc_AUDPC	Sev_AUDPC	Index_AUDPC	Inc_RAUDPC	Sev_RAUDPC	Index_RAUDPC
min	1.52	0.42	-4.19	34.37	0	-12.63	0	0	-0.01
max	91.52	101.07	92.51	927.95	1085	826.41	1	0.63	0.38
range	90	100.65	96.7	893.57	1085	839.03	1	0.63	0.39
median	50.79	70.42	35.31	296.87	353.5	198.6	0.29	0.09	0.07
mean	51.06	66.18	38.51	328.48	412.82	226.8	0.33	0.11	0.08
SE.mean	1.52	1.64	1.7	11.48	12.64	9.76	0.01	0	0
var	533.22	615.94	662.67	30299.03	36733.12	21891.75	0.04	0	0
std.dev	23.09	24.82	25.74	174.07	191.66	147.96	0.2	0.06	0.06
coef.var	0.45	0.38	0.67	0.53	0.46	0.65	0.59	0.56	0.73
h^2	0.75	0.33	0.56	0.52	0.36	0.48	0.52	0.56	0.54

TABLE 33: DESCRIPTIVE STATISTIC OF THE INNOVAR PANEL IN RAVENNA FOR ALL THE TRAITS RECORDED WITH THE RELATED H^2

In the experiment held in Ravenna, the heritability values are variable (Tab. 36), but in any case, not very high. The highest value belongs to the incidence (0.75), which drops to 0.52 for the AUDPC and RAUDPC. The index is 0.56, to drop to 0.48 for the AUDPC and rise to 0.54 for the RAUDPC. The lowest value is related to severity (0.33) which rises considerably for AUDPC (0.36) and RAUDPC (0.56). The correlations in this case tend to be different from the previous experiment, with very high correlations between the incidence, severity and index measurements, but with many non-significant correlations between the AUDPC / RAUDPC values and the phenotypic values taken in the field. In all this, however, HD is negatively correlated (Fig. 53) with most of the traits detected, so also in this experiment it was used as a covariate to correct our data.

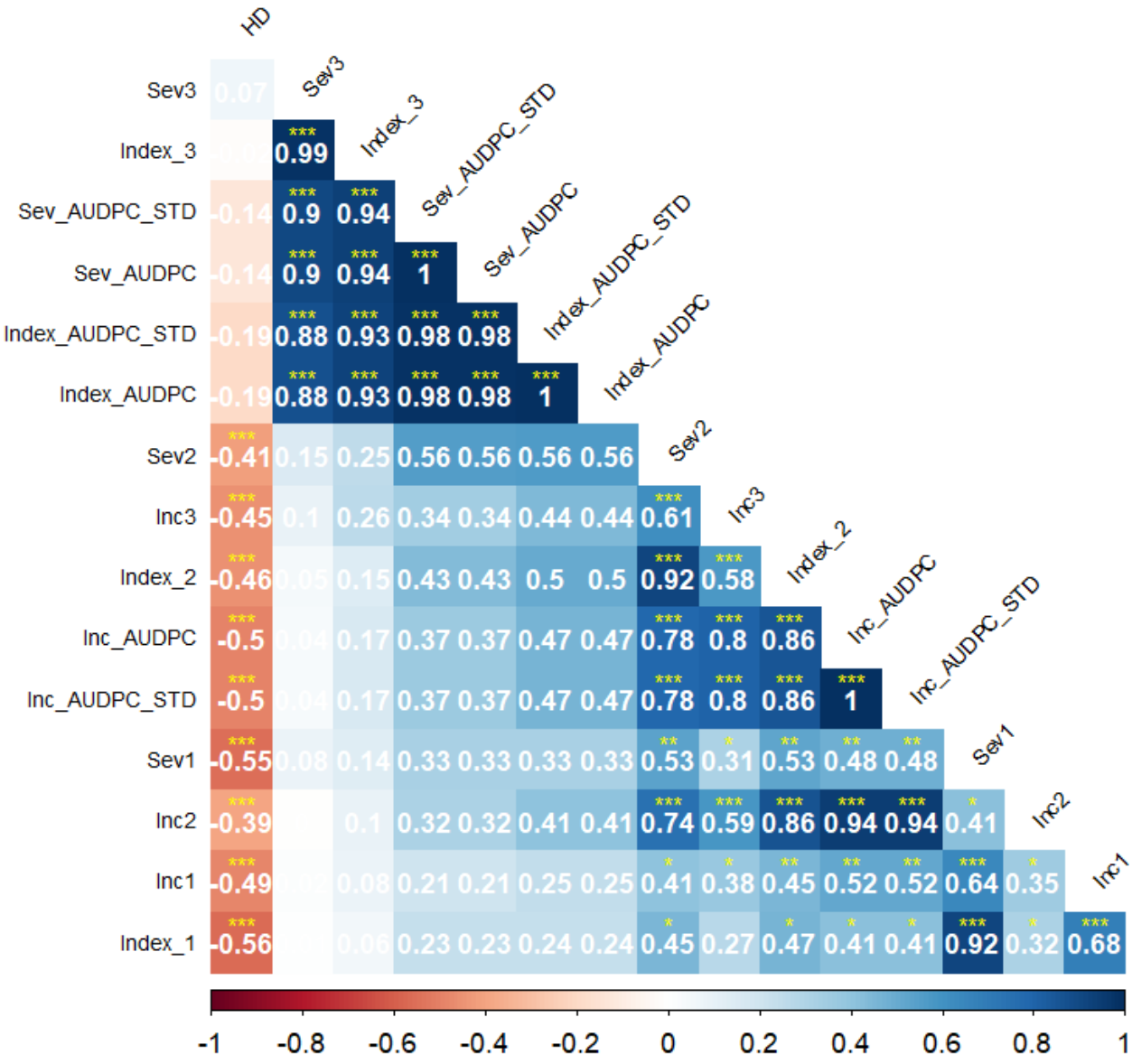


FIGURE 55: CORRELOGRAM FOR THE INNOVAR PANEL TESTED IN RAVENNA FOR ALL THE TRAITS OF INTEREST

ANOVA shows similar results to the previous experiment, with significance only within the genotypes and only for the character of severity and RAUDPC index (Tab. 37), showing a different trend compared to the experiments held in Cadriano.

Trait	Variables	Sum Sq	Df	F value	Pr(>F)	
Incidence	**Genotype**	75638	188	1.2789	0.2802	
Incidence	**Block**	1241	2	1.973	0.168	
Severity	**Genotype**	61956	188	2.5242	0.01259	*
Severity	**Block**	548	2	2.097	0.15182	
Index	**Genotype**	81850	188	1.6228	0.1156	
Index	**Block**	1323	2	2.466	0.1131	
Inc_AUDPC	**Genotype**	4477551	188	1.5793	0.1295	
Inc_AUDPC	**Block**	32695	2	1.084	0.3593	
Sev_AUDPC	**Genotype**	5032946	188	0.9519	0.594	
Sev_AUDPC	**Block**	44867	2	0.7977	0.4657	
Ind_AUDPC	**Genotype**	3673879	188	1.6217	0.116	
Ind_AUDPC	**Block**	49672	2	2.061	0.1563	
Inc_RAUDPC	**Genotype**	5.6211	188	1.5793	0.1295	
Inc_RAUDPC	**Block**	0.041	2	1.084	0.3593	
Sev_RAUDPC	**Genotype**	0.51407	188	1.4808	0.1672	
Sev_RAUDPC	**Block**	0.00295	2	0.7977	0.4657	
Ind_RAUDPC	**Genotype**	0.57523	188	1.8374	0.0665	.
Ind_RAUDPC	**Block**	0.00686	2	2.061	0.1563	

TABLE 34: ANOVA RESULTS FOR THE MAIN TRAITS RECORDED IN RAVENNA USING AS VARIABLES GENOTYPE AND BLOCK

In line with what we have seen previously, the distribution of incidence is normal (Fig. 54) as in the previous experiment. However, severity shows a completely opposite trend (Fig. 55), being squeezed towards susceptibility rather than towards resistance.

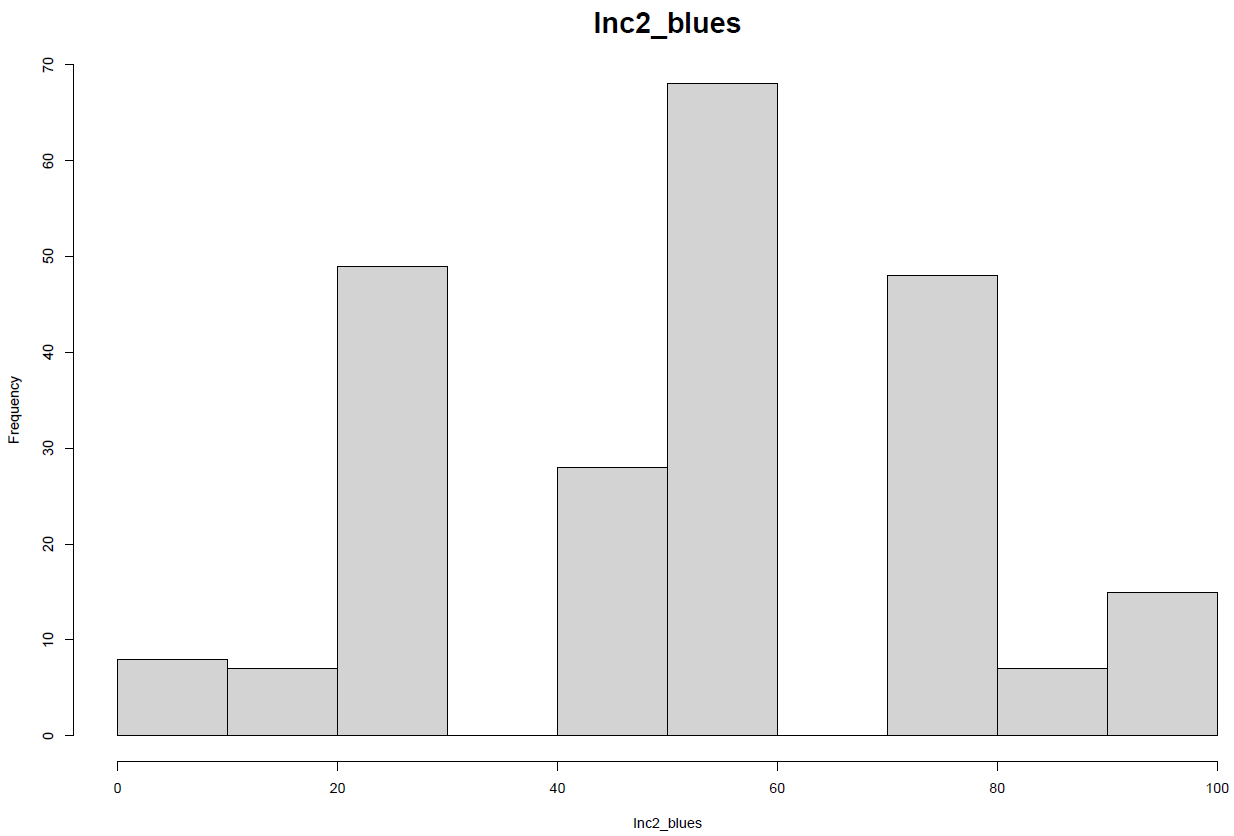


FIGURE 56: DISTRIBUTION HISTOGRAMS RELATED TO FUSARIUM INCIDENCE IN RAVENNA

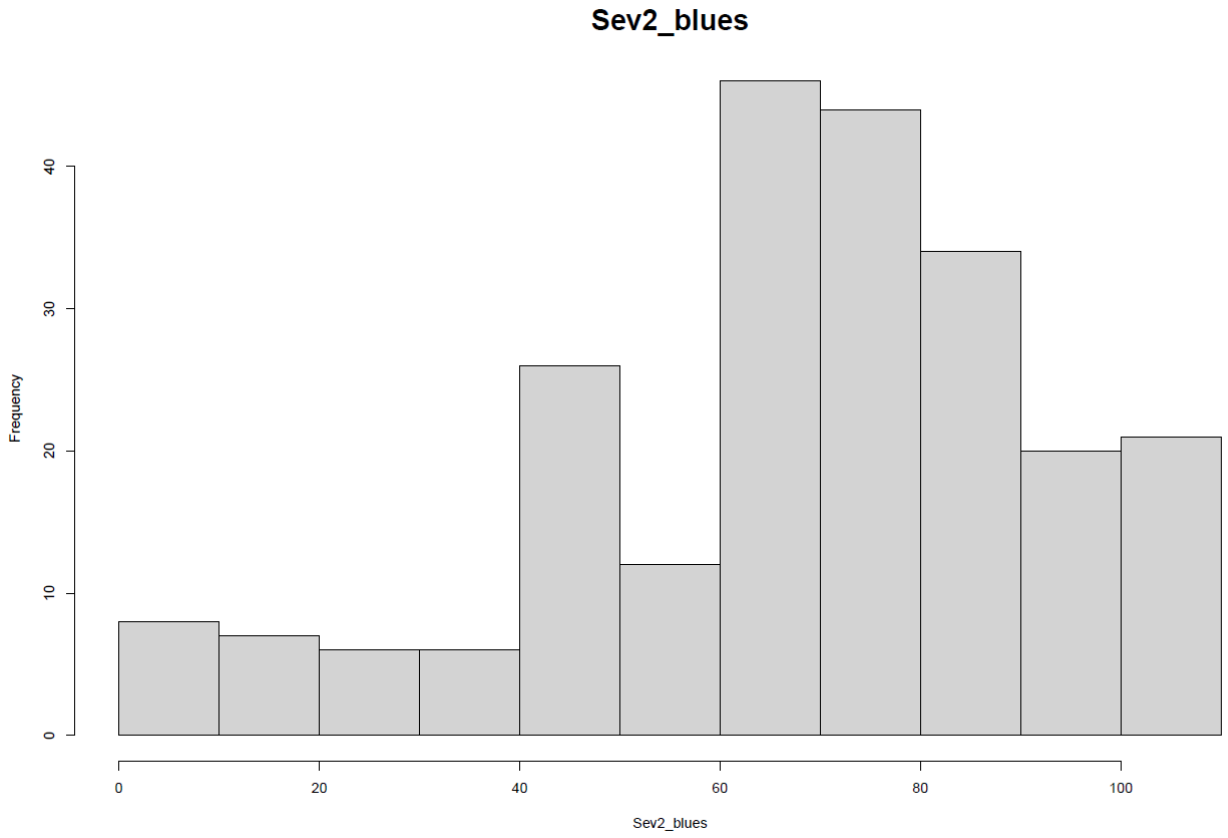


FIGURE 57: DISTRIBUTION HISTOGRAMS RELATED TO FUSARIUM SEVERITY IN RAVENNA

4.4 Discussion

As with the STB experiment, this experiment, focused on FHB, requires data taken over several years in order to have greater data robustness. Also in this case the inoculation mechanism was effective in ensuring a good distribution of the disease, providing a wide range of usable data. For the fusarium experiment carried out on hard, despite a large spread of the disease, it is important to emphasize again the resistance shown by Sumai 3, the primary source of resistance against FHB. The high levels of spread of the disease, albeit with very different levels of severity, can be attributed to a genetic background that is not necessarily performing to combat this disease. On the other hand, the tests carried out on soft wheat cultivars also show this a wide range of susceptibility / resistance, as working on varieties registered by different companies, for different locations, it is easy to assume that some work has been carried out of specific breeding to combat FHB, while for others it completely lacks a background of resistance. With the basis of this experiment, the need is to proceed as soon as possible in repeating the experiment in order to then be able to continue the analyzes, carry out a GWAS and go to identify and investigate loci of resistance for this disease which, with the help of the change climatic, risks bringing some cultivation areas to their knees.

Conclusions and perspectives

These experiments are part of the same research context: finding new loci of resistance to the main diseases of wheat. The bases of these experiments are common, starting from phenotyping the panels in various environments and years to be able to discriminate loci of resistance as accurately as possible, even to different races of emerging pathogens. The experiments on Septoria and Fusarium are still young but have already shown considerable relevance for the phenotypic data taken. These experiments showed a wide phenotypic diversity following the inoculation of the pathogen, highlighting the possibility of successfully discriminating for resistant, susceptible, and partially resistant genotypes, so as to be able to identify new resistance loci. Subsequently, on these two experiments what has been done on yellow rust will be done, going to perform a GWAS analysis that can identify potential loci of interest. An analysis of the candidate genes will then be carried out in order to understand the functions of probable genes responsible for the resistance response. Regarding the yellow rust experiment, many spikes have been identified on durum wheat that may represent potential loci of resistance to this dangerous wheat disease. The peaks studied showed an association to putative candidate genes that have a resistance function in the literature, making the results concrete. With these results, the future prospects are to investigate in detail everything that has been found so as to have a picture of all the potential loci found in these experiments. Following the identification of the candidate genes, these must be validated, making overexpression or knockout tests in order to be able to see the effects of these mutations. Once these genes have been validated, the goal is to build molecular markers that can be used for MAS, to be able to quickly introduce resistance genes into breeding programs and speed up the development of new resistant varieties.

Contributions to the work

This work involved the collaboration of several internal UNIBO collaborators and several external collaborators, both Italian and foreign. The UNIBO team collaborated in all phases of the project. Thanks to the collection of data and the support in some parts of the analysis of external partners, the work was able to conclude as best as possible.

J. Novi, G. Sciara, E. Mazzucotelli, F. Desiderio, F. De Sario, S. Stefanelli, C. Invernizzi, P. Viola, F. Oliveri, B. Randazzo, A.M. Mastrangelo, I. Omar, A. Baidani, H. Özkan, M. El-Areed, I. Bashour, R. Tuberosa, M. Maccaferri.

Conception of work: R. Tuberosa, M. Maccaferri, J. Novi, E. Mazzucotelli.

Field management and data collection UNIBO: J. Novi, M. Campana, M. Bozzoli, P. Viola, C. Invernizzi, F. Oliveri, S. Stefanelli, E. Mazzucotelli, F. Desiderio.

Field management and data collection other locations: B. Randazzo, A.M. Mastrangelo, I. Omar, A. Baidani, H. Özkan, M. El-Areed, I. Bashour.

Genetic and statistical analyses and the genomic associations: J. Novi, G. Sciara, M. Bozzoli, M. Campana, F. De Sario, F. Desiderio.

KASP development: G. Sciara and J. Novi

Bibliography

Agenbag, G. M., et al. "Identification of adult plant resistance to stripe rust in the wheat cultivar Cappelle-Desprez." *Theoretical and Applied Genetics* 125 (2012): 109-120.

Akbari, Mona, et al. "Diversity arrays technology (DArT) for high-throughput profiling of the hexaploid wheat genome." Theoretical and applied genetics 113.8 (2006): 1409-1420.

Akhunov, E.D., Akhunova, A.R. & Dvořák, J. Theor Appl Genet (2005) 111: 1617. https://doi-org.ezproxy.unibo.it/10.1007/s00122-005-0093-1

Alexander, D. H., Novembre, J., & Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. Genome Research, 19(9), 1655–1664. https://doi.org/10.1101/gr.094052.109

Baloch, Faheem Shehzad, et al. "A whole genome DArTseq and SNP analysis for genetic diversity assessment in durum wheat from central fertile crescent." PloS one 12.1 (2017): e0167821.

Bamshad, Michael J., et al. "Exome sequencing as a tool for Mendelian disease gene discovery." Nature Reviews Genetics 12.11 (2011): 745.

Bansal, Urmil, et al. "Molecular mapping of an adult plant stem rust resistance gene Sr56 in winter wheat cultivar Arina." Theoretical and applied genetics 127 (2014): 1441-1448.

Bariana, H. S., et al. "Molecular mapping of adult plant stripe rust resistance in wheat and identification of pyramided QTL genotypes." Euphytica 176 (2010): 251-260.

Bariana, Harbans S., et al. "Discovery of the new leaf rust resistance gene Lr82 in wheat: Molecular mapping and marker development." Genes 13.6 (2022): 964.

Basnet, B. R., et al. "Characterization of Yr54 and other genes associated with adult plant resistance to yellow rust and leaf rust in common wheat Quaiu 3." *Molecular Breeding* 33 (2014): 385-399.

Bennett, J W, and M Klich. “Mycotoxins.” *Clinical microbiology reviews* vol. 16,3 (2003): 497-516. doi:10.1128/CMR.16.3.497-516.2003

Bhatia S. 2015. Application of plant biotechnology. In: Modern applications of plant biotechnology in pharmaceutical sciences. Amsterdam: Elsevier Inc. 157-207

Bhattacharya, S. (2017) Deadly new wheat disease threatens Europe's crops. Nature, 542, 145– 146.

Blanco, A., Il grano, Script editore, 2007. Collana: Coltura & Cultura, capitolo: Origine ed evoluzione, pp 170-195

Bradbury, P. J., Zhang, Z., Kroon, D. E., Casstevens, T. M., Ramdoss, Y., & Buckler, E. S. (2007). TASSEL: software for association mapping of complex traits in diverse samples. Bioinformatics, 23(19), 2633 2635.https://doi.org/10.1093/bioinformatics/btm308

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

Brenchley R. et al., Analysis of the bread wheat genome using whole-genome shotgun sequencing. Nature 491, 705–710 (2012). doi:10.1038/nature11650; pmid:23192148

Breseghello, Flavio, and Mark E. Sorrells. "Association mapping of kernel size and milling quality in wheat (Triticum aestivum L.) cultivars." Genetics 172.2 (2006): 1165-1177.

Brodt, S., Six, J., Feenstra, G., Ingels, C. & Campbell, D. (2011) Sustainable Agriculture. Nature Education Knowledge 3(10):1

Brown, James KM, et al. "Genetics of resistance to Zymoseptoria tritici and applications to wheat breeding." Fungal Genetics and Biology 79 (2015): 33-41.

Browning, B. L., Zhou, Y., & Browning, S. R. (2018). A One-Penny Imputed Genome from Next-Generation Reference Panels. The American Journal of Human Genetics, 103(3), 338–348. https://doi.org/10.1016/J.AJHG.2018.07.015

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.

Casañas F, Simó J, Casals J, Prohens J. Toward an Evolved Concept of Landrace. Front Plant Sci. 2017;8:145. Published 2017 Feb 8. doi:10.3389/fpls.2017.00145

Chaloner Thomas M., Fones Helen N., Varma Varun, Bebber Daniel P. and Gurr Sarah J. A new mechanistic model of weather-dependent Septoria tritici blotch disease risk374Phil. Trans. R. Soc. B. http://doi.org/10.1098/rstb.2018.0266

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

Chedli, Rim Bel Hadj, et al. "Occurrence of Septoria tritici blotch (Zymoseptoria tritici) disease on durum wheat, triticale, and bread wheat in Northern Tunisia." Chilean journal of agricultural research 78.4 (2018): 559-568.

Chen, Jianli, et al. "Genome-wide identification of QTL conferring high-temperature adult-plant (HTAP) resistance to stripe rust (Puccinia striiformis f. sp. tritici) in wheat." *Molecular Breeding* 29 (2012): 791-800.

Chen, W., Wellings, C., Chen, X., Kang, Z., & Liu, T. (2014). Wheat stripe (yellow) rust caused by Puccinia striiformis f. Sp. Tritici. MOLECULAR PLANT PATHOLOGY, 14.

Chen, X. M. "Epidemiology and control of stripe rust [Puccinia striiformis f. sp. tritici] on wheat." Canadian Journal of Plant Pathology 27.3 (2005): 314-337.

Cheng, P., et al. "Molecular mapping of genes Yr64 and Yr65 for stripe rust resistance in hexaploid derivatives of durum wheat accessions PI 331260 and PI 480016." *Theoretical and Applied Genetics* 127 (2014): 2267-2277.

Cheval, Pénélope, et al. "Evolution of QoI resistance of the wheat pathogen Zymoseptoria tritici in Northern France." Crop protection 92 (2017): 131-133.

Cobo, Nicolas, et al. "A high-resolution map of wheat QYr. ucw-1BL, an adult plant stripe rust resistance locus in the same chromosomal region as Yr29." *The plant genome* 12.1 (2019): 180055.

de Carvalho, Miguel AA Pinheiro, et al. "Cereal landraces genetic resources in worldwide GeneBanks. A review." Agronomy for Sustainable Development 33.1 (2013): 177-203.

Dedryver, F., et al. "Characterization of genetic components involved in durable resistance to stripe rust in the bread wheat 'Renan'." *Phytopathology* 99.8 (2009): 968-973.

Dodds, P. N., & Rathjen, J. P. (2010). Plant immunity: Towards an integrated view of plant–pathogen interactions. Nature Reviews Genetics, 11(8), 539–548. https://doi.org/10.1038/nrg2812.

Dolores Vazquez, M., et al. "Genetic analysis of adult plant, quantitative resistance to stripe rust in wheat cultivar 'Stephens' in multi-environment trials." *Theoretical and Applied Genetics* 124 (2012): 1-11.

Dong, Zhenzhen, et al. "Validation and characterization of a QTL for adult plant resistance to stripe rust on wheat chromosome arm 6BS (Yr78)." Theoretical and Applied Genetics 130 (2017): 2127-2137.

Duba, A., Goriewa-Duba, K., & Wachowska, U. (2018). A review of the interactions between wheat and wheat pathogens: Zymoseptoria tritici, fusarium spp. and parastagonospora nodorum. International Journal of Molecular Sciences, 19(4), 1–21. https://doi.org/10.3390/ijms19041138

Dweba, C. C., et al. "Fusarium head blight of wheat: Pathogenesis and control strategies." *Crop protection* 91 (2017): 114-122.

Ellis, J. G., Lagudah, E. S., Spielmeyer, W., & Dodds, P. N. (2014). The past, present and future of breeding rust resistant wheat. Frontiers in Plant Science, 5. https://doi.org/10.3389/fpls.2014.00641.

El-Orabey, W. M. (2018). Virulence of some Puccinia triticina races to the effective wheat leaf rust resistant genes Lr 9 and Lr 19 under Egyptian field conditions. Physiological and Molecular Plant Pathology, 102, 163–172. https://doi.org/10.1016/j.pmpp.2017.12.006

Eyal, Z. The Septoria diseases of wheat: concepts and methods of disease management. Cimmyt, 1987.

Fagnano, M., Fiorentino, N., D'Egidio, M. G., Quaranta, F., Ritieni, A., Ferracane, R., & Raimondi, G. (2012). Durum wheat in conventional and organic farming: Yield amount and pasta quality in Southern Italy. The Scientific World Journal, 2012. https://doi.org/10.1100/2012/973058

Fermaintt, C. S., Sano, K., Liu, Z., Ishii, N., Seino, J., Dobbs, N., … Yan, N. (2019). A bioactive mammalian disaccharide associated with autoimmunity activates STING-TBK1-dependent immune response. Nature Communications, 10(1). https://doi.org/10.1038/s41467-019-10319-5

Flint-Garcia, Sherry A., Jeffry M. Thornsberry, and Edward S. Buckler IV. "Structure of linkage disequilibrium in plants." Annual review of plant biology 54.1 (2003): 357-374.

Fones, Helen, and Sarah Gurr. "The impact of Septoria tritici Blotch disease on wheat: An EU perspective." Fungal Genetics and Biology 79 (2015): 3-7.

Fuchs M. Pyramiding resistance-conferring gene sequences in crops. Curr Opin Virol. 2017 Oct;26:36-42. doi: 10.1016/j.coviro.2017.07.004. Epub 2017 Jul 26. Review. PubMed PMID: 28755651.

Gautam, H. R., M. L. Bhardwaj, and Rohitashw Kumar. "Climate change and its impact on plant diseases." Current Science (2013): 1685-1691.

Gong, Xuan, et al. "Development of an evaluation system for Fusarium resistance in wheat grains and its application in assessment of the corresponding effects of Fhb1." *Plant Disease* 104.8 (2020): 2210-2216.

Goode E.L. (2011) Linkage Disequilibrium. In: Schwab M. (eds) Encyclopedia of Cancer. Springer, Berlin, Heidelberg

Goswami, Rubella S., and H. Corby Kistler. "Heading for disaster: Fusarium graminearum on cereal crops." *Molecular plant pathology* 5.6 (2004): 515-525.

H. Özkan, A. Brandolini, R. Schäfer-Pregl, F. Salamini, AFLP Analysis of a Collection of Tetraploid Wheats Indicates the Origin of Emmer and Hard Wheat Domestication in Southeast Turkey, Molecular Biology and Evolution, Volume 19, Issue 10, October 2002, Pages 1797–1801, https://doi.org/10.1093/oxfordjournals.molbev.a004002

Hovmoller, M. S., Walter, S., & Justesen, A. F. (2010). Escalating Threat of Wheat Rusts. Science, 329(5990), 369–369. https://doi.org/10.1126/science.1194925.

Huerta-Espino, J., and R. P. Singh. "First detection of virulence in Puccinia striiformis f. sp. tritici to wheat resistance genes Yr10 and Yr24 (= Yr26) in Mexico." *Plant Disease* 101.9 (2017): 1676-1676.

Jia, L. J., Tang, H. Y., Wang, W. Q., Yuan, T. L., Wei, W. Q., Pang, B., … Tang, W. H. (2019). A linear nonribosomal octapeptide from Fusarium graminearum facilitates cell-to-cell invasion of wheat. Nature Communications, 10(1). https://doi.org/10.1038/s41467-019-08726-9

Jin, Y., Szabo, L. J., & Carson, M. (2010). Century-Old Mystery of Puccinia striiformis Life History Solved with the Identification of Berberis as an Alternate Host. Phytopathology, 100(5), 432–435. https://doi.org/10.1094/PHYTO-100-5-0432.

Johnson R. (1983) Genetic Background of Durable Resistance. In: Lamberti F., Waller J.M., Van der Graaff N.A. (eds) Durable Resistance in Crops. NATO Advanced Science Institutes Series (Series A: Life Sciences), vol 55. Springer, Boston, MA

Juliana, Philomin, et al. "Genome-wide association mapping for resistance to leaf rust, stripe rust and tan spot in wheat reveals potential candidate genes." *Theoretical and applied genetics* 131.7 (2018): 1405-1422.

Juroszek, Peter, and Andreas von Tiedemann. "Climate change and potential future risks through wheat diseases: a review." European Journal of Plant Pathology 136.1 (2013): 21-33.

Karisto, Petteri, et al. "Ranking quantitative resistance to Septoria tritici blotch in elite wheat cultivars using automated image analysis." Phytopathology 108.5 (2018): 568-581.

Karlstedt, Frances, et al. "Mapping of quantitative trait loci (QTL) for resistance against Zymoseptoria tritici in the winter spelt wheat accession HTRI1410 (Triticum aestivum subsp. spelta)." Euphytica 215.6 (2019): 108.

Kaur, Parampreet, and Kishor Gaikwad. "From genomes to GENE-omes: exome sequencing concept and applications in crop improvement." Frontiers in Plant Science 8 (2017): 2164.

Kettles, G. J., & Kanyuka, K. (2016). Dissecting the molecular interactions between wheat and the fungal pathogen Zymoseptoria tritici. Frontiers in Plant Science, 7(APR2016), 1–7. https://doi.org/10.3389/fpls.2016.00508

Kombrink, Anja, and Bart PHJ Thomma. "LysM effectors: secreted proteins supporting fungal life." PLoS pathogens 9.12 (2013): e1003769.

Korte, Arthur, and Ashley Farlow. "The advantages and limitations of trait analysis with GWAS: a review." Plant methods 9.1 (2013): 29.

Kumpatla SP, Buyyarapu R, Abdurakhmonov IY, Mammadov JA (2012) Genomics-assisted plant breeding in the 21st century: technological advances and progress. In: Abdurakhmonov I (ed) Plant breeding.

Lan, Caixia, et al. "QTL characterization of resistance to leaf rust and stripe rust in the spring wheat line Francolin# 1." *Molecular breeding* 34 (2014): 789-803.

Leplat, Johann, et al. "Survival of Fusarium graminearum, the causal agent of Fusarium head blight. A review." *Agronomy for sustainable development* 33.1 (2013): 97-111.

Li, Jing, et al. "Genome-wide association study reveals the genetic architecture of stripe rust resistance at the adult plant stage in Chinese endemic wheat." *Frontiers in Plant Science* 11 (2020): 625.

Lin, F., and X. M. Chen. "Molecular mapping of genes for race-specific overall resistance to stripe rust in wheat cultivar Express." *Theoretical and Applied Genetics* 116 (2008): 797-806.

Lin, X., N'Diaye, A., Walkowiak, S., Nilsen, K. T., Cory, A. T., Haile, J., … Pozniak, C. J. (s.d.). Genetic analysis of resistance to stripe rust in durum wheat (Triticum turgidum L. var. Durum). 2018.

Liu W. Z., Maccaferri M., Bulli P., Rynearson S., Tuberosa R., Chen X. M., et al. (2017). Genome-wide association mapping for seedling and field resistance to *Puccinia striiformis f. sp. tritici* in elite durum wheat. *Theor. Appl. Genet.* 130 649–667. 10.1007/s00122-016-2841-9

Lowe, Iago, Dario Cantu, and Jorge Dubcovsky. "Durable resistance to the wheat rusts: integrating systems biology and traditional phenotype-based research methods to guide the deployment of resistance genes." Euphytica 179 (2011): 69-79.

M'Barek, S. B., J. H. Cordewener, S. M. T. Ghaffary, T. A. van der Lee, Z. Liu, A. M. Gohari, R. Mehrabi, A. H. America, O. Robert and T. L. Friesen (2015). "FPLC and liquid-chromatography mass spectrometry identify candidate necrosis-inducing proteins from culture filtrates of the fungal wheat pathogen Zymoseptoria tritici." Fungal Genetics and Biology 79: 54-62.

Maccaferri, M., Zhang, J., Bulli, P., Abate, Z., Chao, S., Cantu, D., … Dubcovsky, J. (2015). 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), 449–465. https://doi.org/10.1534/g3.114.014563.

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

Maldonado, Carlos, et al. "Genome-wide haplotype-based association analysis of key traits of plant lodging and architecture of maize identifies major determinants for leaf angle: hapLA4." PloS one 14.3 (2019): e0212925.

Mallard, Stephanie, et al. "Genetic analysis of durable resistance to yellow rust in bread wheat." Theoretical and Applied Genetics 110 (2005): 1401-1409.

Marshall, R., Kombrink, A., Motteram, J., Loza-Reyes, E., Lucas, J., Hammond-Kosack, K. E., … Rudd, J. J. (2011). Analysis of two in planta expressed LysM effector homologs from the fungus mycosphaerella graminicola reveals novel functional properties and varying contributions to virulence on wheat. Plant Physiology, 156(2), 756–769. https://doi.org/10.1104/pp.111.176347

Matsuoka, Y. (2011) 'Evolution of polyploid Triticum wheats under cultivation: the role of domestication, natural hybridization and allopolyploid speciation in their diversification. Plant Cell Physiol', Vol.52: 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.

McDonald, B. A., & Mundt, C. C. (2016). How knowledge of pathogen population biology informs management of Septoria tritici blotch. Phytopathology, 106(9), 948–955. https://doi.org/10.1094/PHYTO-03-16-0131-RVW

McIntosh, R. A., Dubcovsky, J., Rogers, W. J., Morris, C., Appels, R., & Xia, X. C. (2013). CATALOGUE OF GENE SYMBOLS FOR WHEAT: 2013-2014 SUPPLEMENT. 31.

McIntosh, R.A., Hart, G.E., Devos, K.M., Gale, M.D., and Rogers, W.J. 1998. Catalogue of gene symbols for wheat. In Proceedings of the 9th International Wheat Genetics Symposium, held 2–7 August 1998,

Saskatoon, Saskatchewan, Canada. Edited by A.E. Slinkard. University Extension Press, University of Saskatchewan, Saskatoon. Vol. 5. pp. 139–142.

McIntosh, Robert, et al. "Wheat stripe rust resistance gene Yr24/Yr26: A retrospective review." *The Crop Journal* 6.4 (2018): 321-329.

Motzo, R., Giunta, F., & Fois, S. (2001). Evoluzione varietale e qualità in frumento duro dalle vecchie popolazioni alle attuali cultivar. 173.

Mu, Jingmei, et al. "Genome-wide association study and gene specific markers identified 51 genes or QTL for resistance to stripe rust in US winter wheat cultivars and breeding lines." *Frontiers in plant science* 11 (2020): 998.

Naz, Ali Ahmad, et al. "Advanced backcross quantitative trait locus analysis in winter wheat: Dissection of stripe rust seedling resistance and identification of favorable exotic alleles originated from a primary hexaploid wheat (Triticum turgidum ssp. dicoccoides× Aegilops tauschii)." *Molecular Breeding* 30 (2012): 1219-1229.

Neddaf, H. M., Aouini, L., Bouznad, Z., & Kema, G. H. J. (2017). Equal distribution of mating type alleles and the presence of strobilurin resistance in Algerian Zymoseptoria tritici field populations. Plant Disease, 101(4), 544–549. https://doi.org/10.1094/PDIS-03-16-0298-RE

Parlevliet, J.E. & Zadoks, J.C. Euphytica (1977) 26: 5. https://doi-org.ezproxy.unibo.it/10.1007/BF00032062

Poland, Jesse A., et al. "Shades of gray: the world of quantitative disease resistance." Trends in plant science 14.1 (2009): 21-29.

Ponomarenko, Alisa, Stephen B. Goodwin, and Gert HJ Kema. "Septoria tritici blotch (STB) of wheat." Septoria tritici blotch (STB) of wheat (2011).

Pretorius, Z. A., et al. "First report of a Puccinia graminis f. sp. tritici race virulent to the Sr24 and Sr31 wheat stem rust resistance genes in South Africa." Plant disease 94.6 (2010): 784-784.

Quaedvlieg, W., et al. "Zymoseptoria gen. nov.: a new genus to accommodate Septoria-like species occurring on graminicolous hosts." Persoonia: Molecular Phylogeny and Evolution of Fungi 26 (2011): 57.

Rafalski, J. Antoni. "Association genetics in crop improvement." Current opinion in plant biology 13.2 (2010): 174-180.

Randhawa, M., Bansal, U., Valárik, M., Klocová, B., Doležel, J., & Bariana, H. (2014). Molecular mapping of stripe rust resistance gene Yr51 in chromosome 4AL of wheat. Theoretical and Applied Genetics, 127(2), 317–324. https://doi.org/10.1007/s00122-013-2220-8.

Richard D., 2017. "2017 Outlook of the U.S. and World Wheat Industries, 2017-2026," Agribusiness & Applied Economics Report 262199, North Dakota State University, Department of Agribusiness and Applied Economics.

Rosewarne, G. M., et al. "Analysis of leaf and stripe rust severities reveals pathotype changes and multiple minor QTLs associated with resistance in an Avocet× Pastor wheat population." *Theoretical and applied genetics* 124.7 (2012): 1283-1294.

Salamini, F., Özkan, H., Brandolini, A., Schäfer-Pregl, R., & Martin, W. (2002). Genetics and geography of wild cereal domestication in the near east. Nature Reviews Genetics, 3(6), 429–441. https://doi.org/10.1038/nrg817

Semagn, Kassa, et al. "Single nucleotide polymorphism genotyping using Kompetitive Allele Specific PCR (KASP®): overview of the technology and its application in crop improvement." Molecular breeding 33.1 (2014): 1-14.

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

Siah, A., Elbekali, A. Y., Ramdani, A., Reignault, P., Torriani, S. F. F., Brunner, P. C., & Halama, P. (2014). QoI resistance and mitochondrial genetic structure of Zymoseptoria tritici in Morocco. Plant Disease, 98(8), 1138–1144. https://doi.org/10.1094/PDIS-10-13-1057-RE

Singh RP, Hodson DP, Huerta-Espino J, Jin Y, Bhavani S, Njau P, Herrera-Foessel S, Singh PK, Singh S, Govindan V. The emergence of Ug99 races of the stem rust fungus is a threat to world wheat production. Annu Rev Phytopathol. 2011;49:465-81. doi: 10.1146/annurev-phyto-072910-095423. Review. PubMed PMID: 21568701.

Singh, R. P., et al. "Identification and mapping of gene Yr31 for resistance to stripe rust in Triticum aestivum cultivar Pastor." *Proceedings of the 10th International wheat genetics symposium*. Vol. 1. Rome, Italy: Instituto Sperimentale per la Cerealicoltura, 2003.

Sørensen, C. K., Labouriau, R., & Hovmøller, M. S. (2017). Temporal and Spatial Variability of Fungal Structures and Host Responses in an Incompatible Rust–Wheat Interaction. Frontiers in Plant Science, 8. https://doi.org/10.3389/fpls.2017.00484.

Spielmeyer, W., et al. "NBS-LRR sequence family is associated with leaf and stripe rust resistance on the end of homoeologous chromosome group 1S of wheat." *Theoretical and applied genetics* 101.7 (2000): 1139-1144.

Tuberosa, R., Graner, A., & Frison, E. (2014). Genomics of plant genetic resources: Volume 1. Managing, sequencing and mining genetic resources. Genomics of Plant Genetic Resources: Volume 1. Managing, Sequencing and Mining Genetic Resources, 1–710. https://doi.org/10.1007/978-94-007-7572-5

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, M. N., Wan, A. M., & Chen, X. M. (2015). Barberry as Alternate Host Is Important for Puccinia graminis f. Sp. Tritici But Not for Puccinia striiformis f. Sp. Tritici in the U.S. Pacific Northwest. Plant Disease, 99(11), 1507–1516. https://doi.org/10.1094/PDIS-12-14-1279-RE.

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.

Wellings, Colin R. (2011). Global status of stripe rust: A review of historical and current threats. Euphytica, 179(1), 129–141. https://doi.org/10.1007/s10681-011-0360-y.

Wu, J. H., Wang, Q. L., Chen, X. M., Wang, M. J., Mu, J. M., Lv, X. N., … Kang, Z. S. (2016). Stripe rust resistance in wheat breeding lines developed for central Shaanxi, an overwintering region for Puccinia striiformis f. Sp. Tritici in China. Canadian Journal of Plant Pathology, 38(3), 317–324. https://doi.org/10.1080/07060661.2016.1206039.

Xu, L. S., Wang, M. N., Cheng, P., Kang, Z. S., Hulbert, S. H., & Chen, X. M. (2013). Molecular mapping of Yr53, a new gene for stripe rust resistance in durum wheat accession PI 480148 and its transfer to common wheat. Theoretical and Applied Genetics, 126(2), 523–533. https://doi.org/10.1007/s00122-012-1998-0.

Yang, N., McDonald, M.C., Solomon, P.S. et al. Theor Appl Genet (2018) 131: 2765. https://doi-org.ezproxy.unibo.it/10.1007/s00122-018-3189-0

You, Qian, et al. "Development and applications of a high throughput genotyping tool for polyploid crops: single nucleotide polymorphism (SNP) array." Frontiers in plant science 9 (2018): 104.

Zhang, Chaozhong, et al. "An ancestral NB-LRR with duplicated 3' UTRs confers stripe rust resistance in wheat and barley." *Nature communications* 10.1 (2019): 1-12.

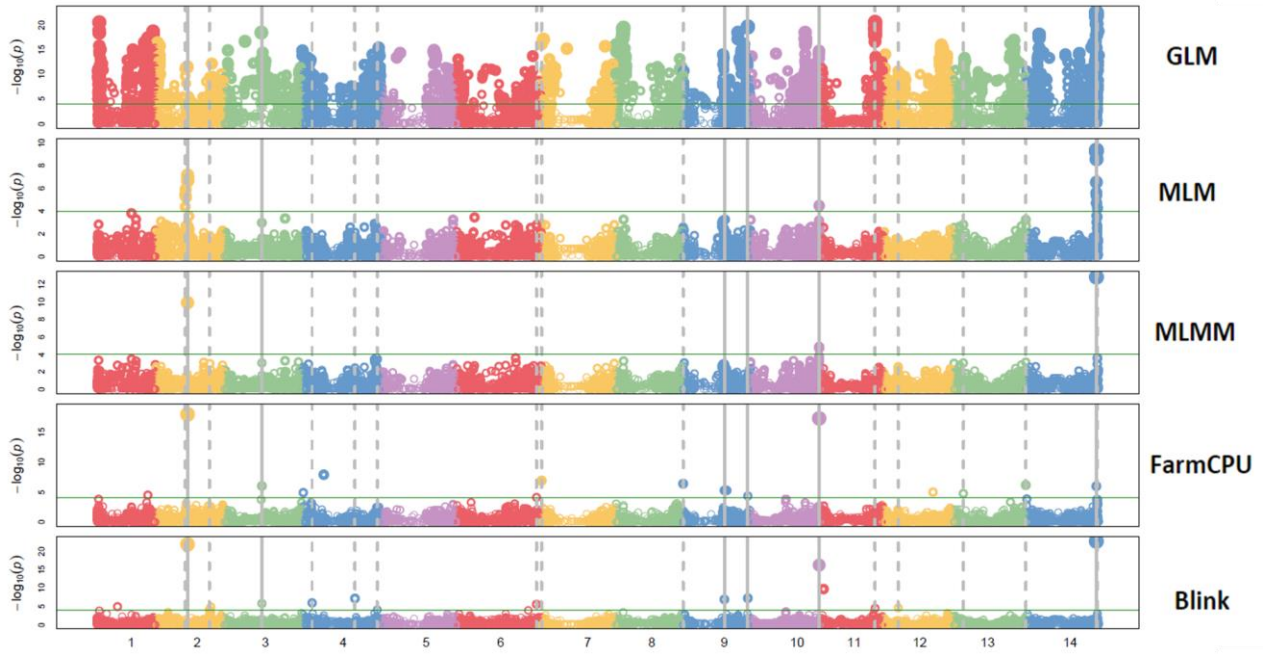
Zheng, S., Li, Y., Lu, L., Liu, Z., Zhang, C., Ao, D., … Zhang, L. (2017). Evaluating the contribution of Yr genes to stripe rust resistance breeding through marker-assisted detection in wheat. Euphytica, 213(2), 50. https://doi.org/10.1007/s10681-016-1828-6.

Zhu, Xiaoxu, et al. "TaMYB29: A novel R2R3-MYB transcription factor involved in wheat defense against stripe rust." *Frontiers in Plant Science* 12 (2021).

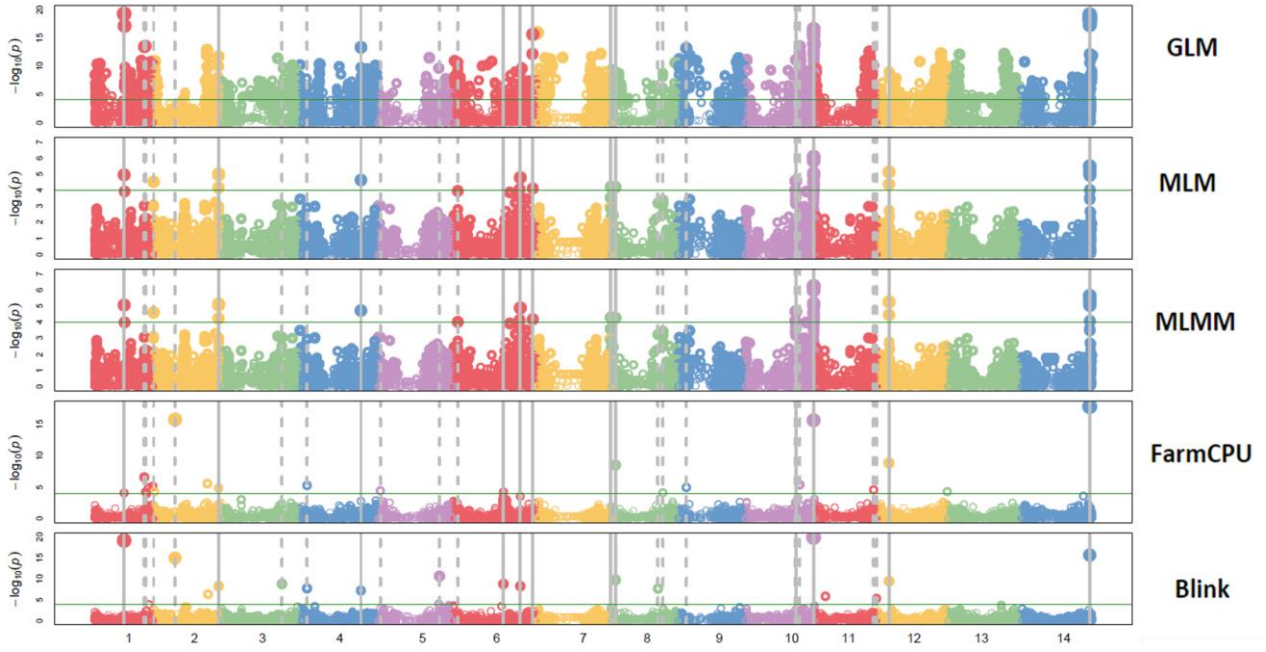
Žilić, S., Barać, M., Pešić, M., Dodig, D., & Ignjatović-Micić, D. (2011). Characterization of proteins from grain of different bread and durum wheat genotypes. International Journal of Molecular Sciences, 12(9), 5878–5894. https://doi.org/10.3390/ijms12095878

Zipfel, C. (2014). Plant pattern-recognition receptors. Trends in Immunology, 35(7), 345–351. https://doi.org/10.1016/j.it.2014.05.004.

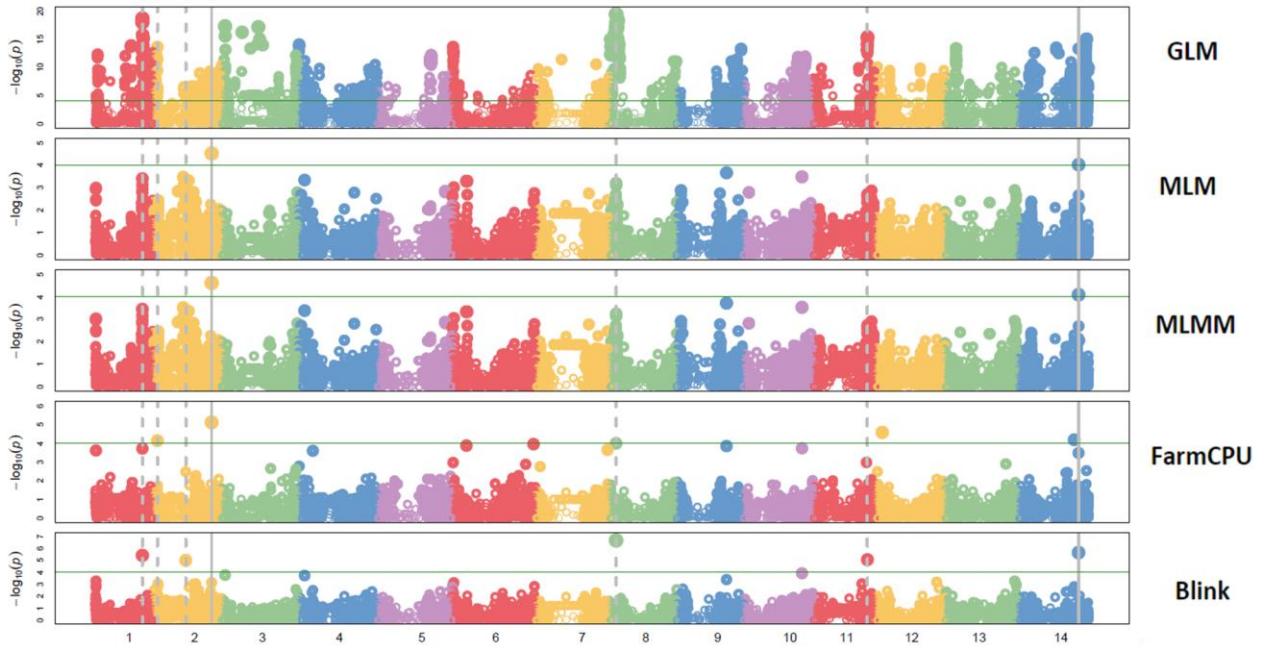
Supplementary Materials



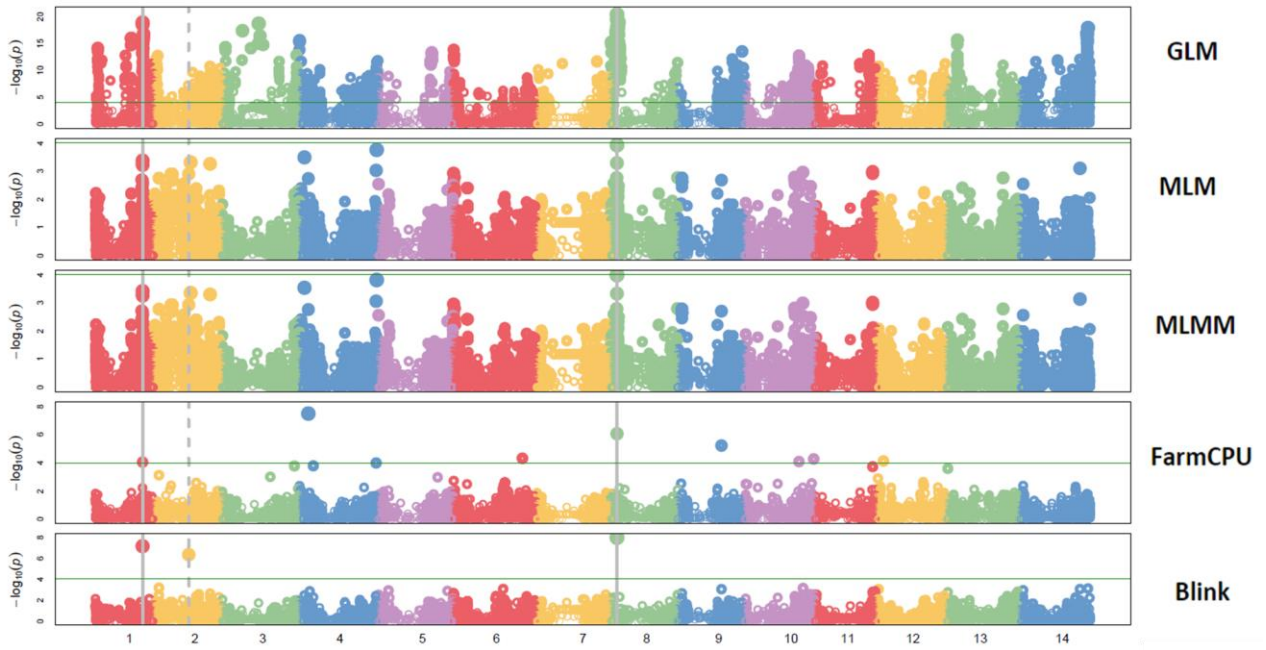
SUPPLEMENTARY FIGURE 1: MANHATTAN PLOT WITH ALL MODELS (GLM, MLM, MLMM, FARMCPU AND BLINK) FOR ARGENTINA IT



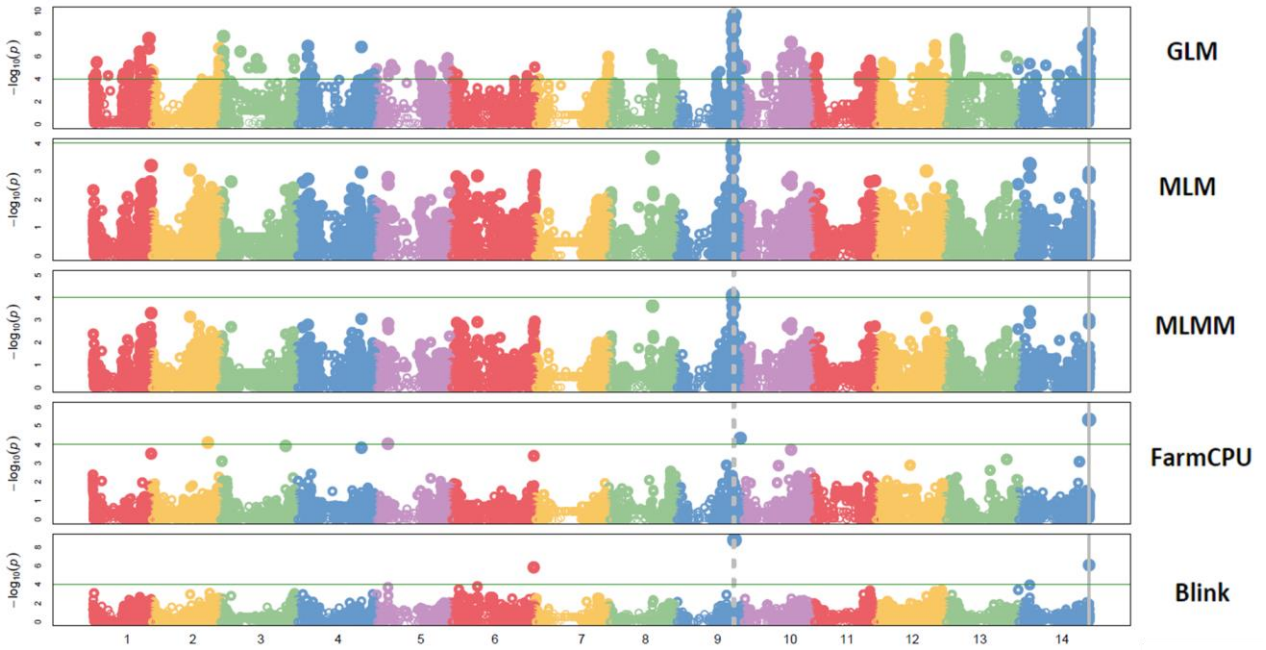
SUPPLEMENTARY FIGURE 2: MANHATTAN PLOT WITH ALL MODELS (GLM, MLM, MLMM, FARMCPU AND BLINK) FOR ARGENTINA SEV



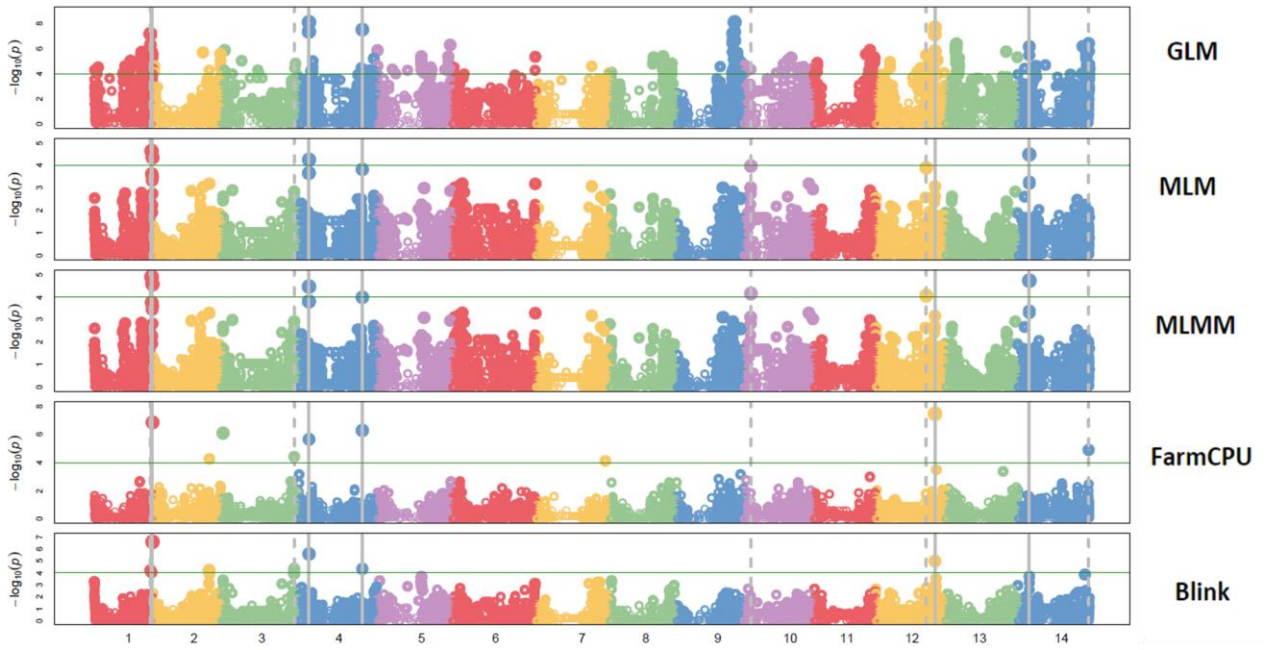
SUPPLEMENTARY FIGURE 3: MANHATTAN PLOT WITH ALL MODELS (GLM, MLM, MLMM, FARMCPU AND BLINK) FOR BAUCINA IT



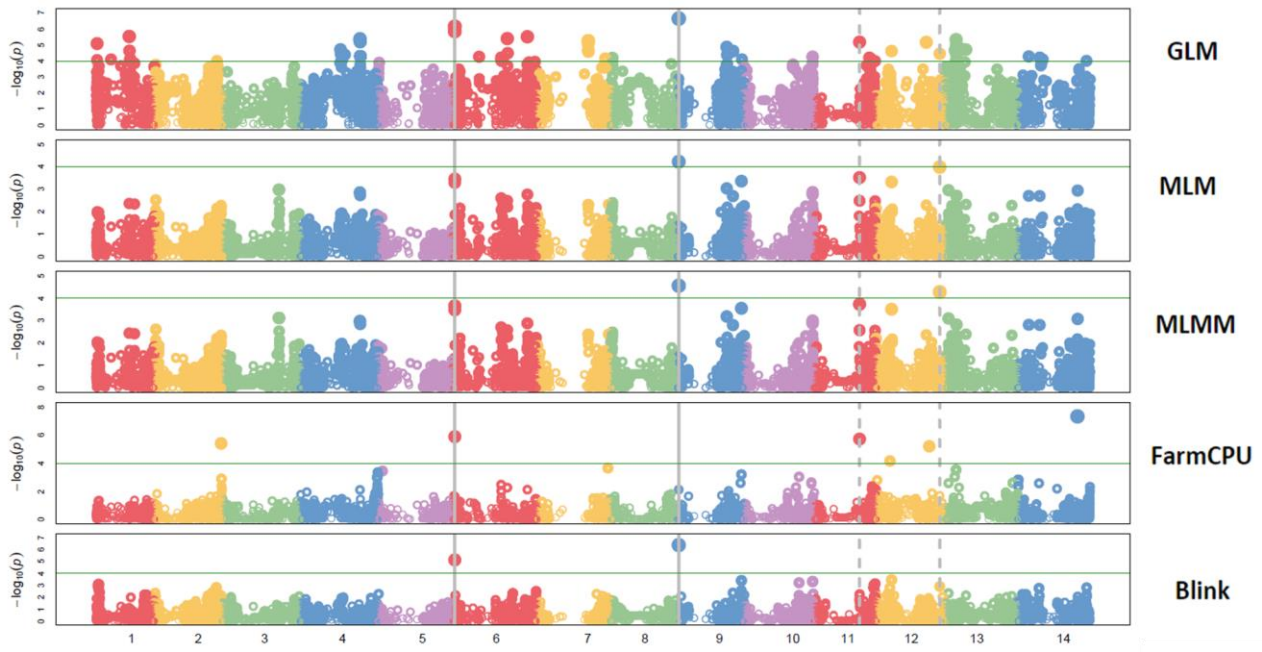
SUPPLEMENTARY FIGURE 4: MANHATTAN PLOT WITH ALL MODELS (GLM, MLM, MLMM, FARMCPU AND BLINK) FOR BAUCINA SEV



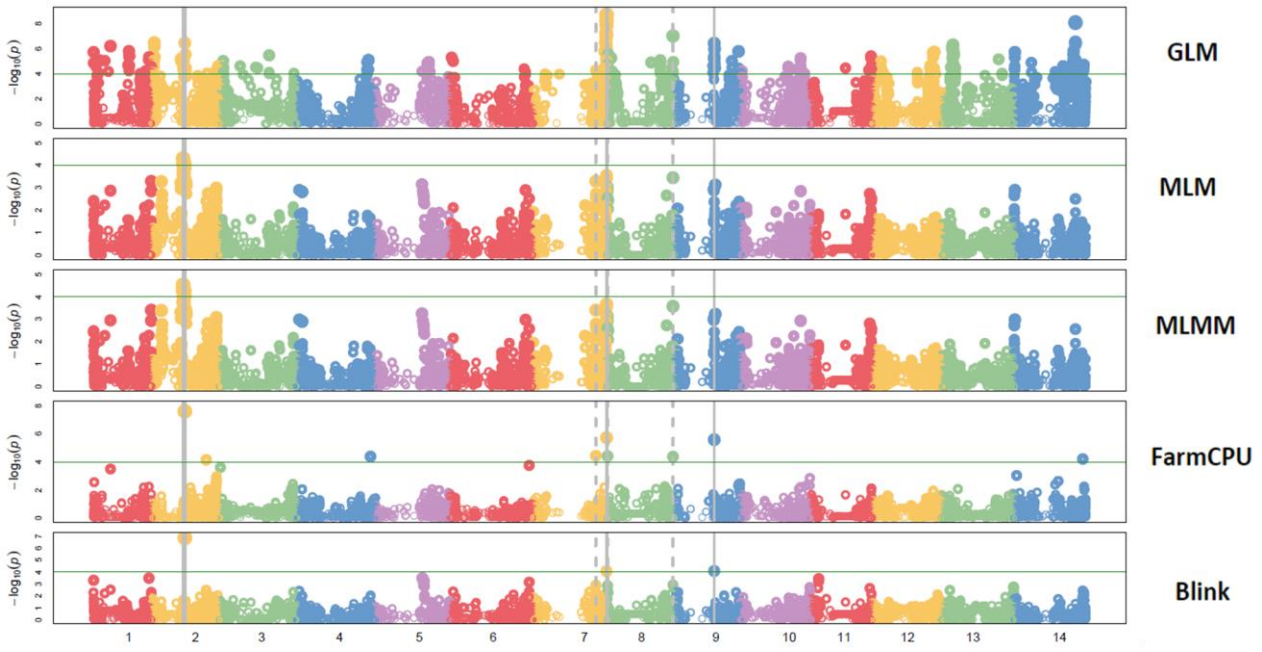
SUPPLEMENTARY FIGURE 5: MANHATTAN PLOT WITH ALL MODELS (GLM, MLM, MLMM, FARMCPU AND BLINK) FOR EGYPT IT



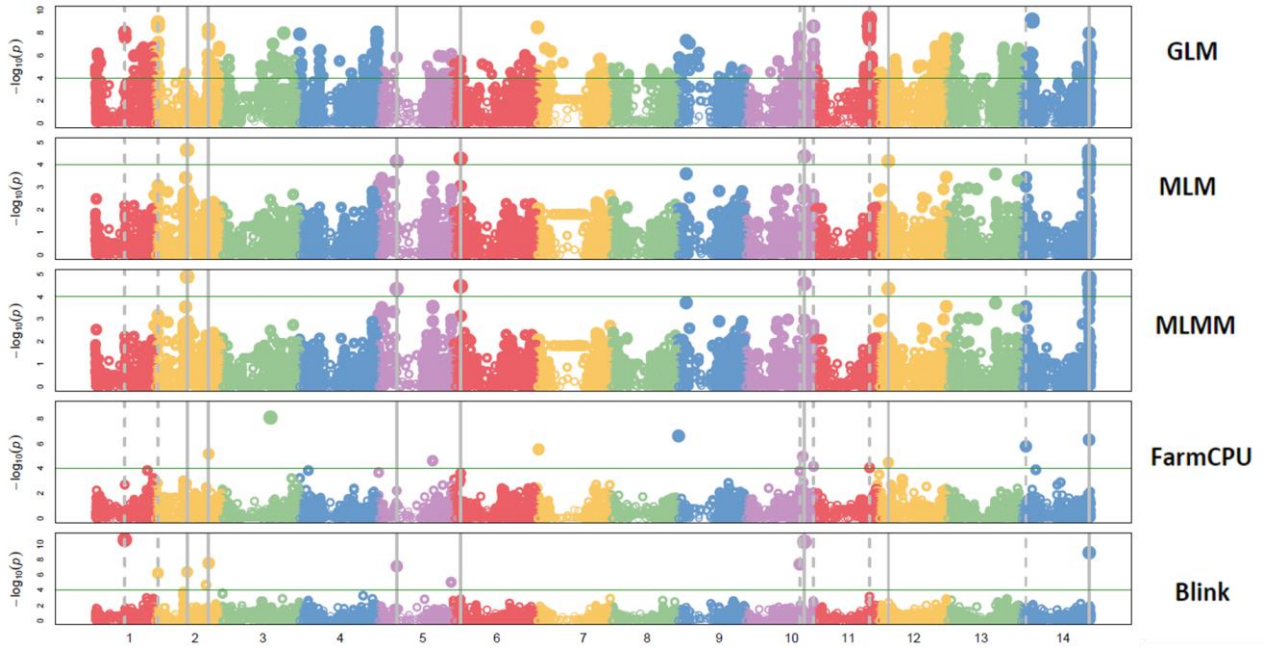
SUPPLEMENTARY FIGURE 6: MANHATTAN PLOT WITH ALL MODELS (GLM, MLM, MLMM, FARMCPU AND BLINK) FOR EGYPT SEV



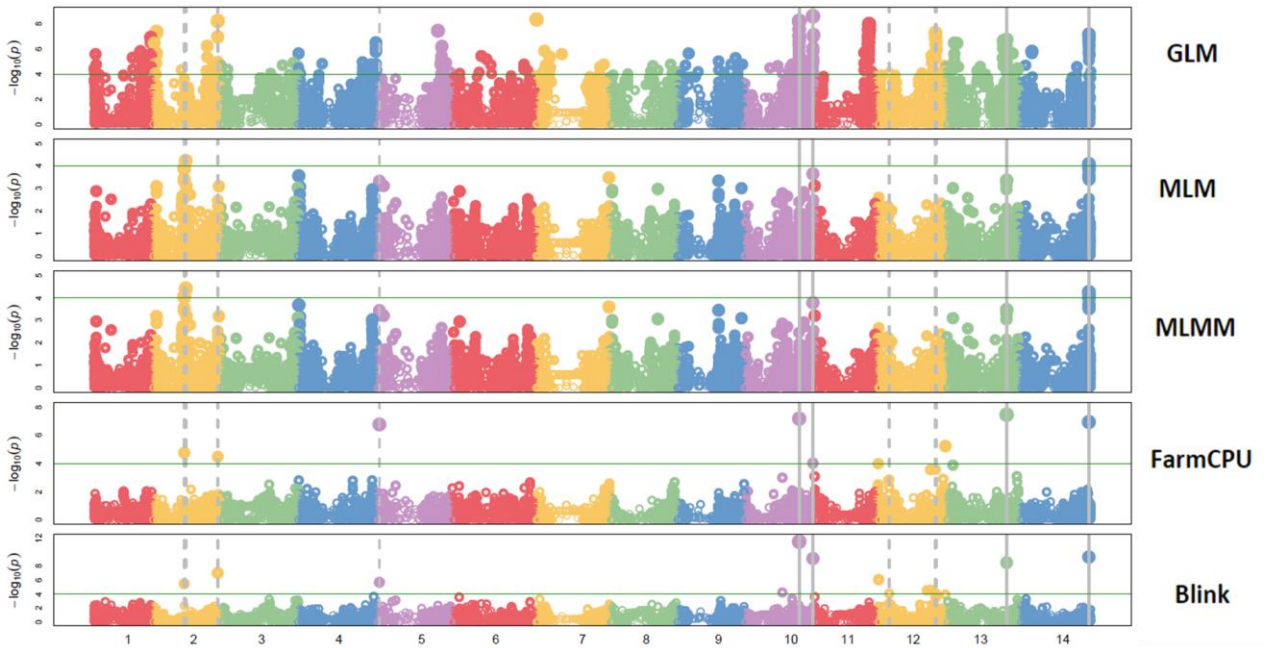
SUPPLEMENTARY FIGURE 7: MANHATTAN PLOT WITH ALL MODELS (GLM, MLM, MLMM, FARMCPU AND BLINK) FOR FOGGIA IT



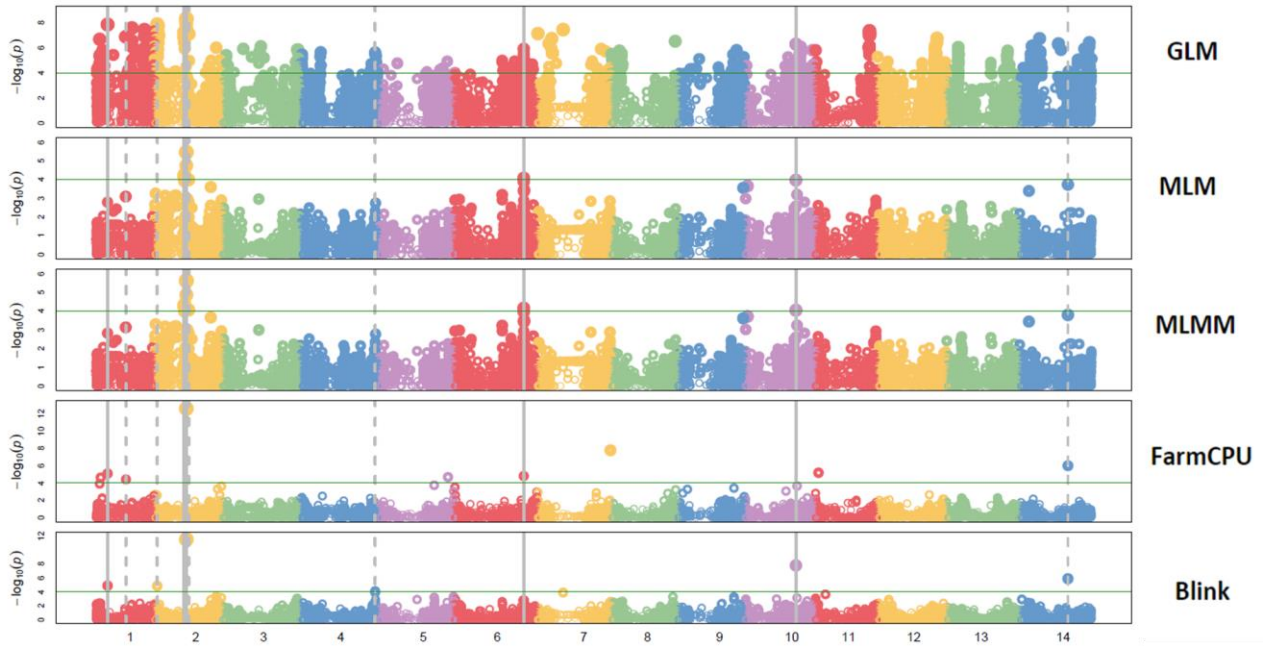
SUPPLEMENTARY FIGURE 8: MANHATTAN PLOT WITH ALL MODELS (GLM, MLM, MLMM, FARMCPU AND BLINK) FOR FOGGIA SEV



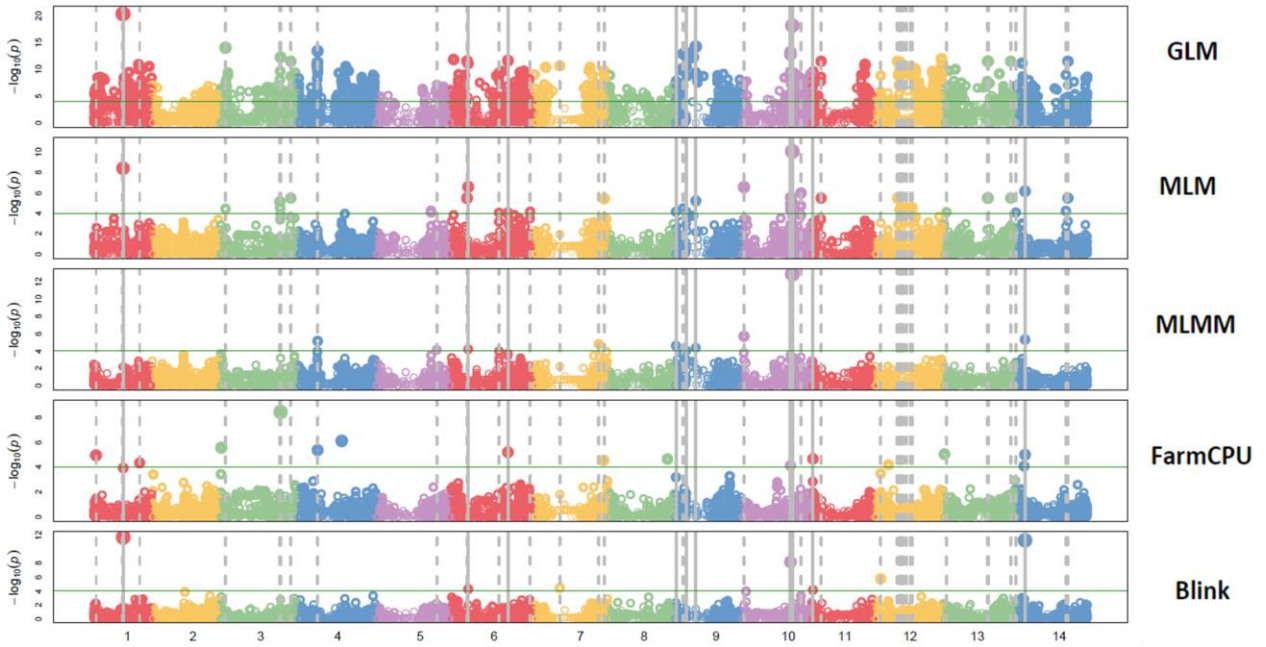
SUPPLEMENTARY FIGURE 9: MANHATTAN PLOT WITH ALL MODELS (GLM, MLM, MLMM, FARMCPU AND BLINK) FOR GROSSETO 2019 IT



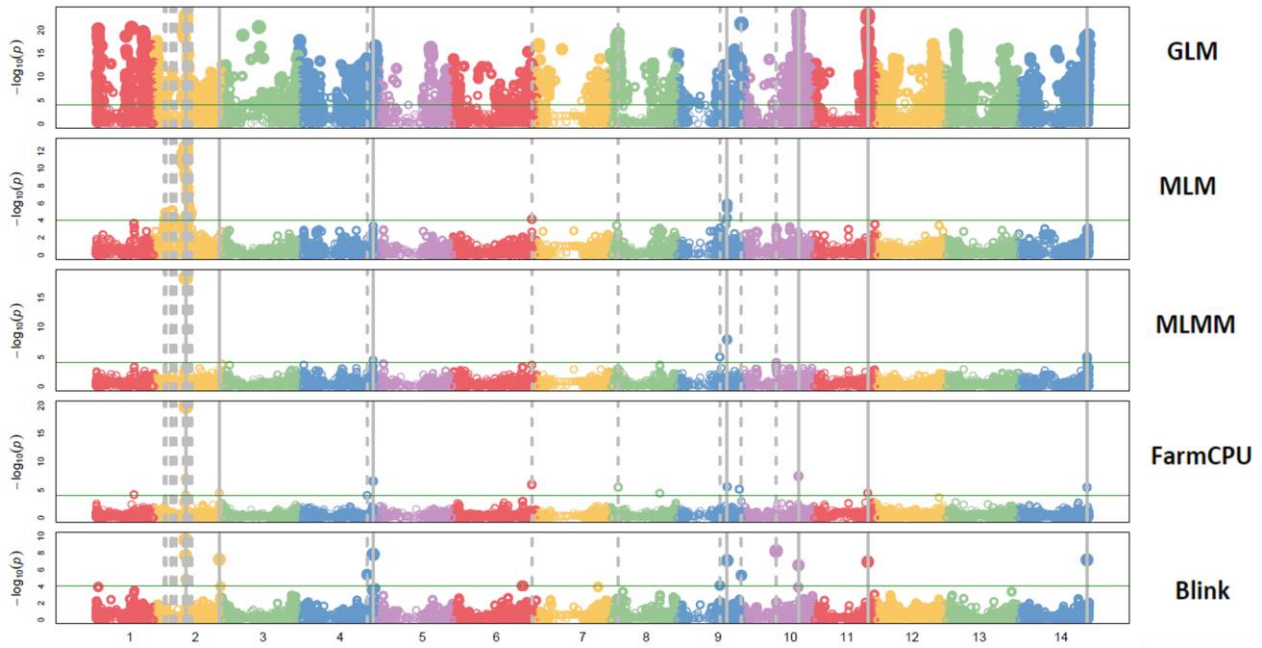
SUPPLEMENTARY FIGURE 10: MANHATTAN PLOT WITH ALL MODELS (GLM, MLM, MLMM, FARMCPU AND BLINK) FOR GROSSETO 2019 SEV



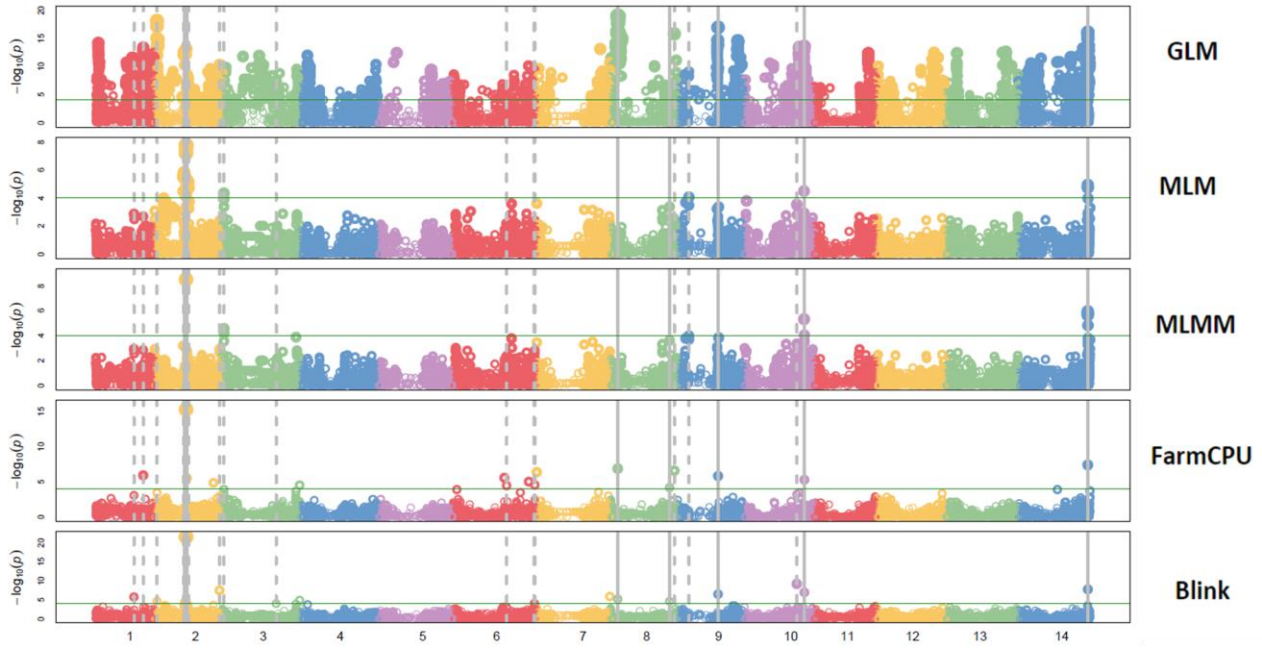
SUPPLEMENTARY FIGURE 11: MANHATTAN PLOT WITH ALL MODELS (GLM, MLM, MLMM, FARMCPU AND BLINK) FOR GROSSETO 2020 IT



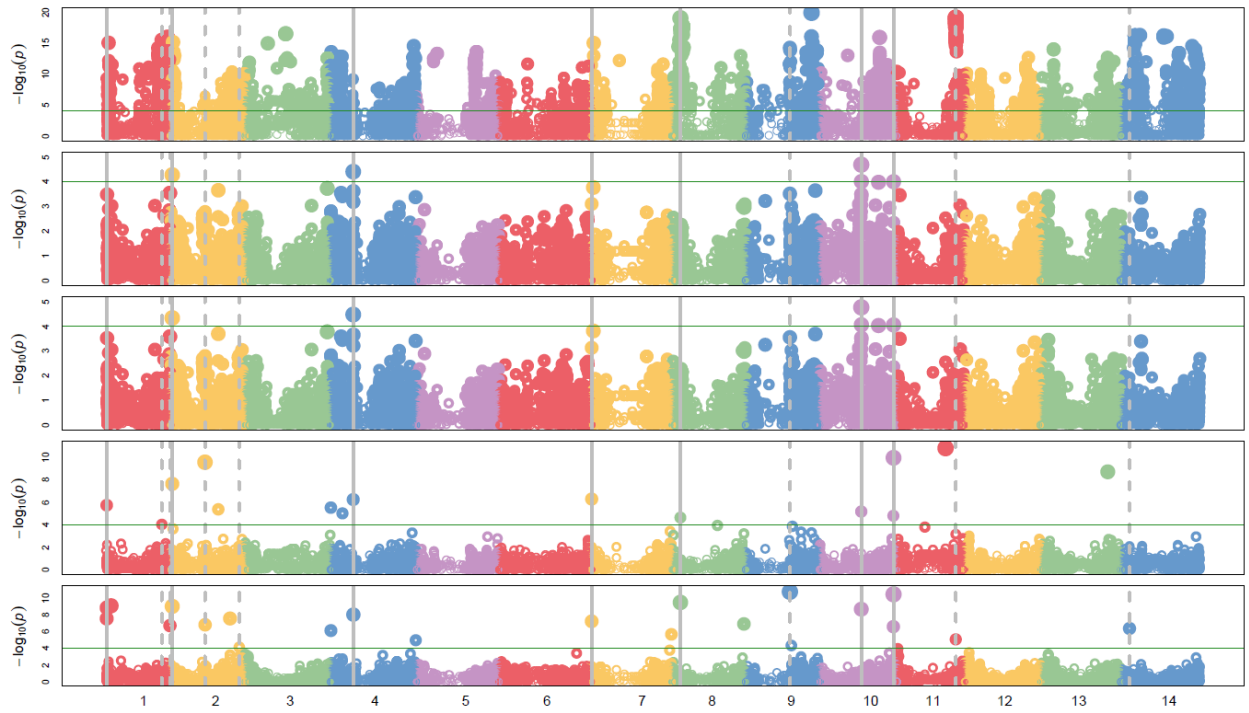
SUPPLEMENTARY FIGURE 12: MANHATTAN PLOT WITH ALL MODELS (GLM, MLM, MLMM, FARMCPU AND BLINK) FOR GROSSETO 2020 SEV



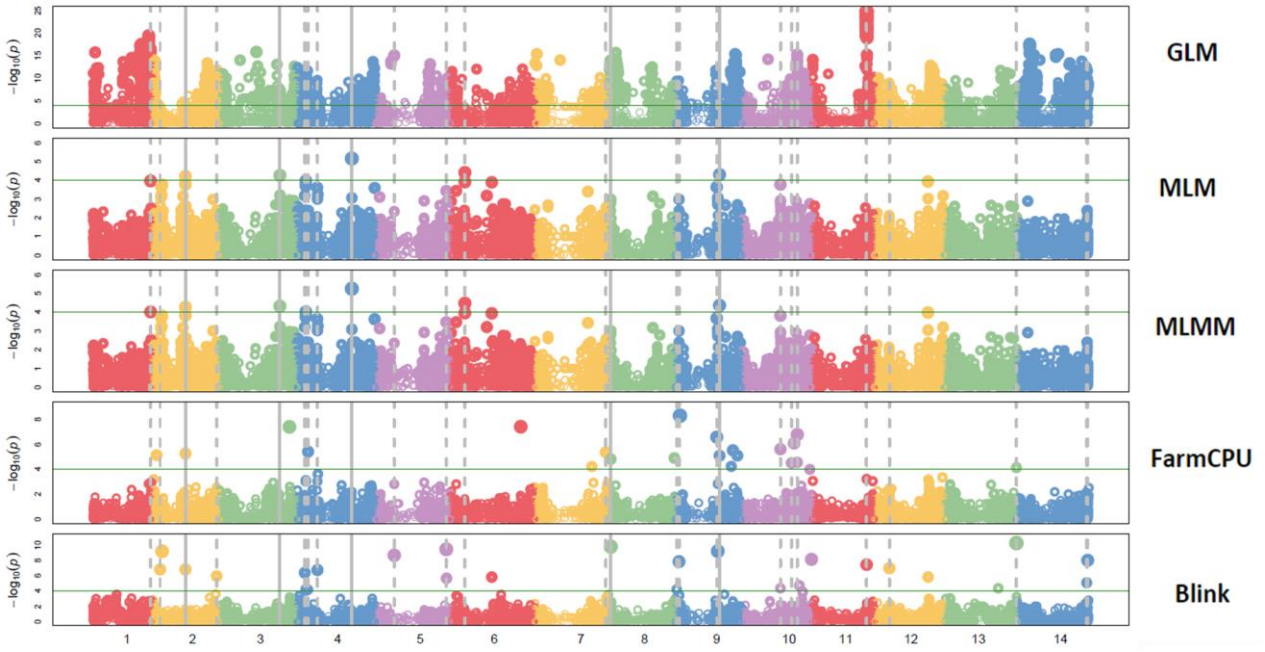
SUPPLEMENTARY FIGURE 13: MANHATTAN PLOT WITH ALL MODELS (GLM, MLM, MLMM, FARMCPU AND BLINK) FOR GROSSETO 2021 IT



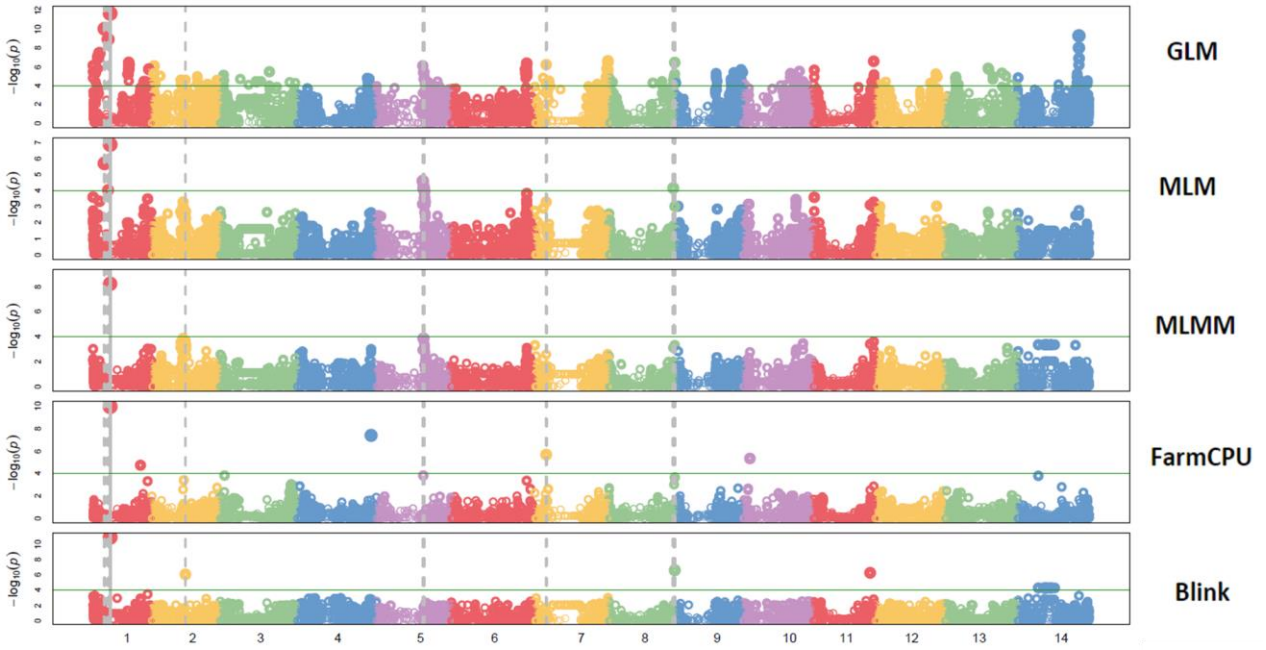
SUPPLEMENTARY FIGURE 14: MANHATTAN PLOT WITH ALL MODELS (GLM, MLM, MLMM, FARMCPU AND BLINK) FOR GROSSETO 2021 SEV



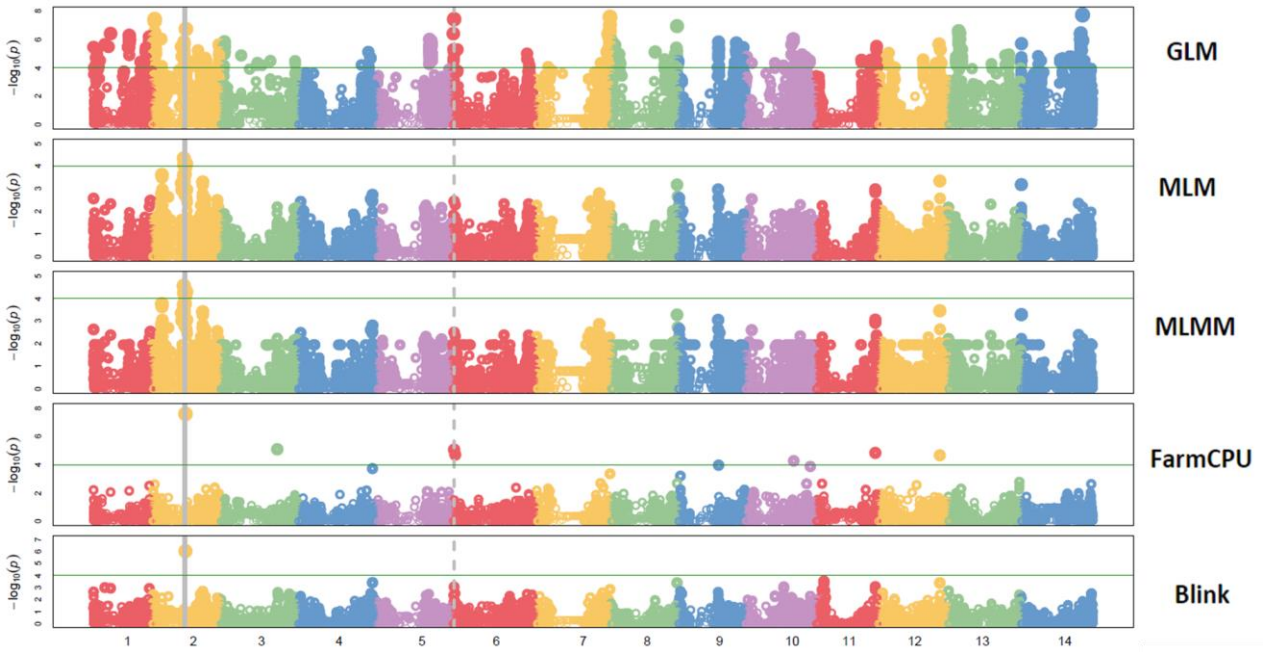
SUPPLEMENTARY FIGURE 15: MANHATTAN PLOT WITH ALL MODELS (GLM, MLM, MLMM, FARMCPU AND BLINK) FOR LEBANON IT



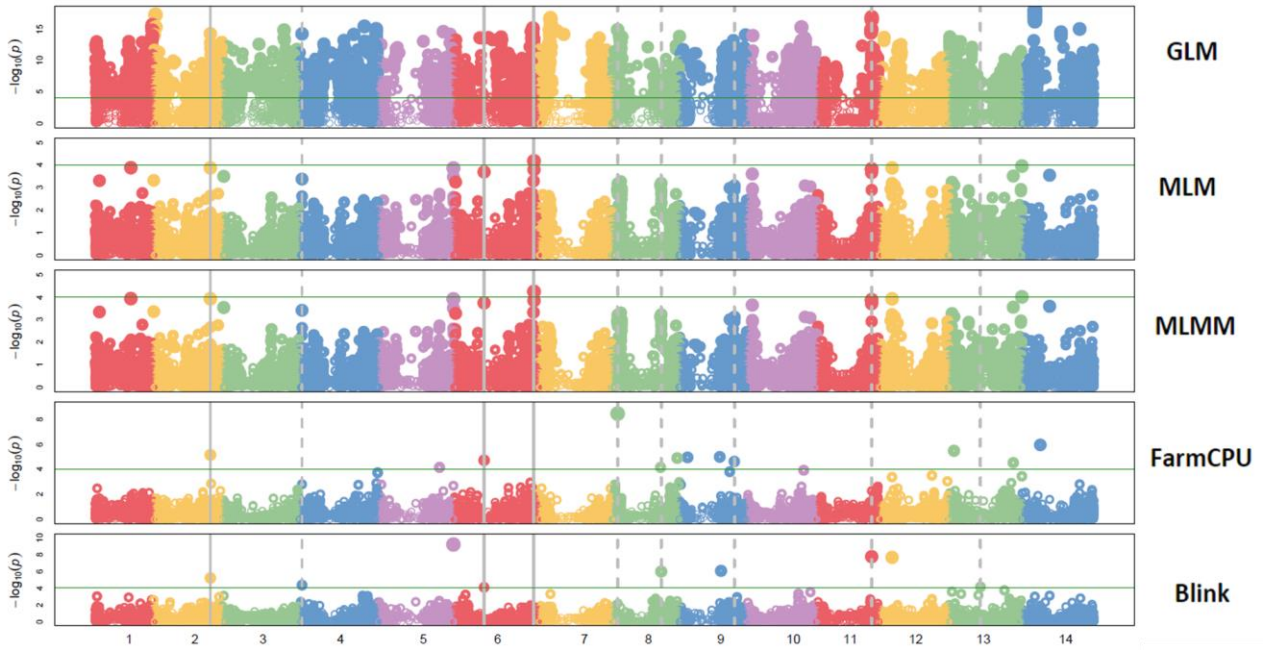
SUPPLEMENTARY FIGURE 16: MANHATTAN PLOT WITH ALL MODELS (GLM, MLM, MLMM, FARMCPU AND BLINK) FOR LEBANON SEV



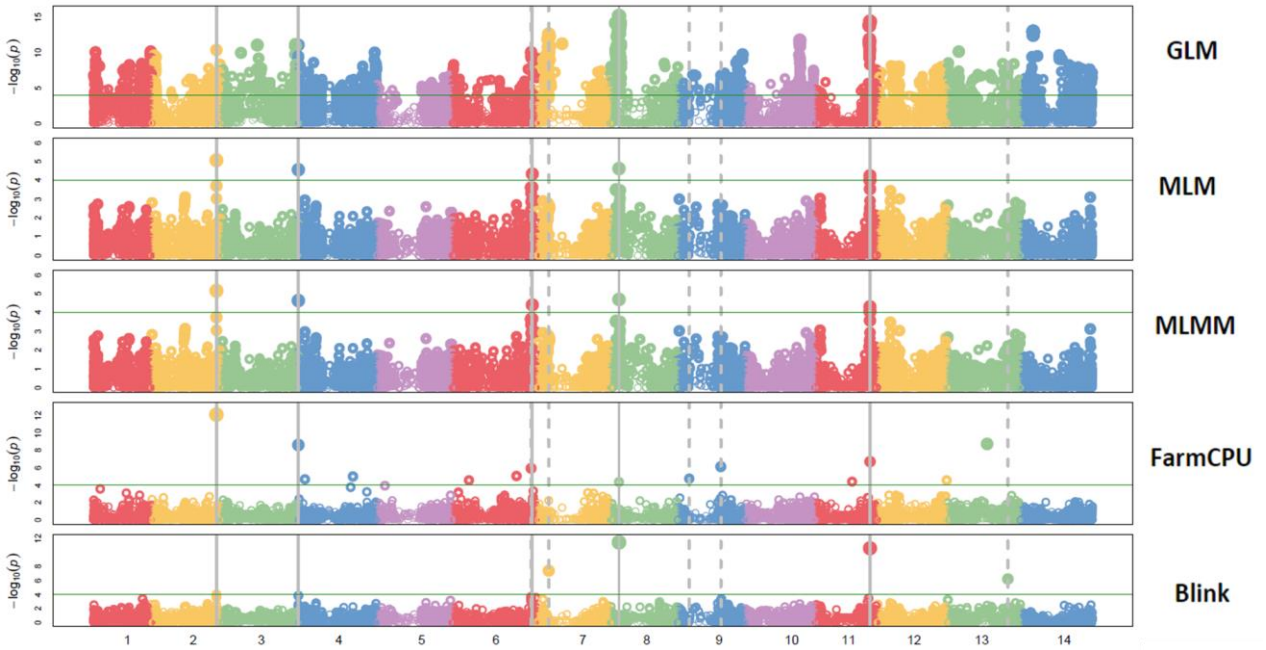
SUPPLEMENTARY FIGURE 17: MANHATTAN PLOT WITH ALL MODELS (GLM, MLM, MLMM, FARMCPU AND BLINK) FOR MOROCCO IT



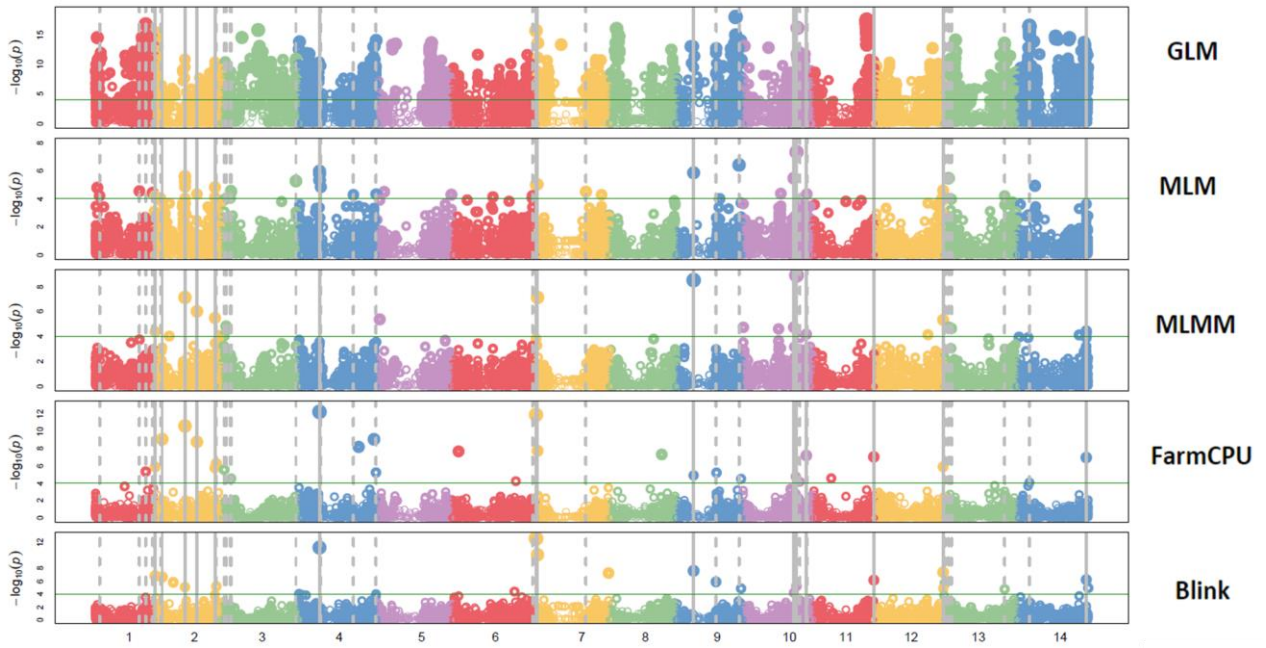
SUPPLEMENTARY FIGURE 18: MANHATTAN PLOT WITH ALL MODELS (GLM, MLM, MLMM, FARMCPU AND BLINK) FOR MOROCCO SEV



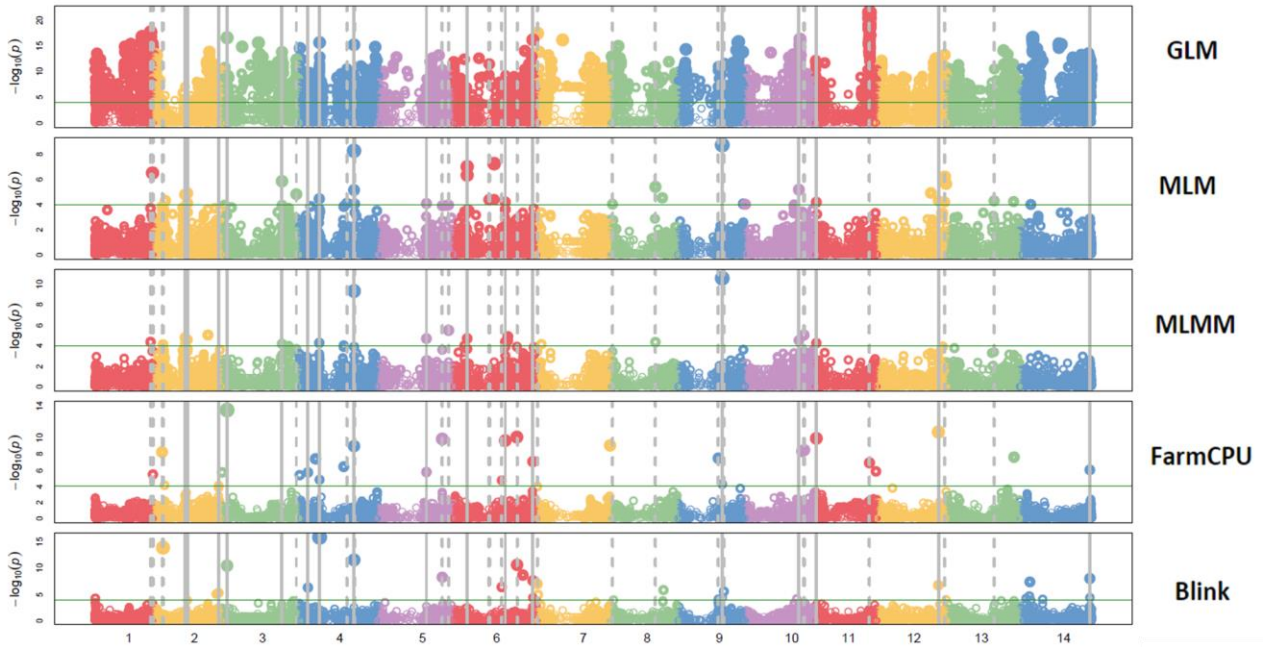
SUPPLEMENTARY FIGURE 19: MANHATTAN PLOT WITH ALL MODELS (GLM, MLM, MLMM, FARMCPU AND BLINK) FOR TURKEY IT



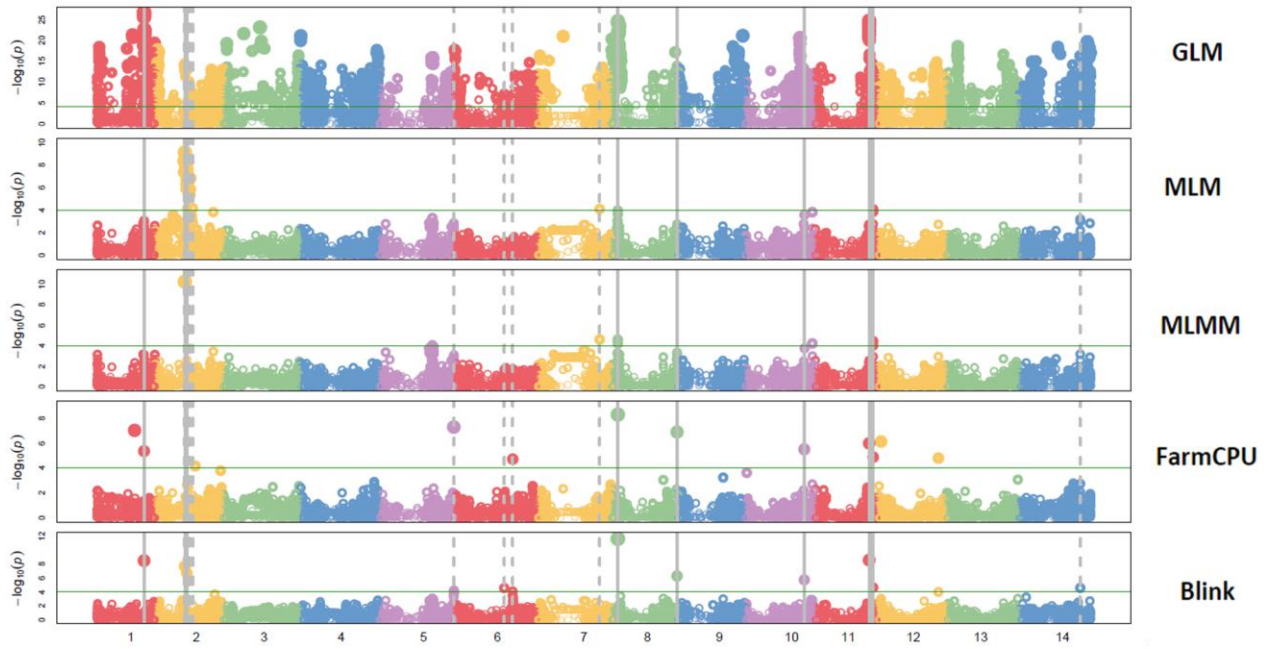
SUPPLEMENTARY FIGURE 20: MANHATTAN PLOT WITH ALL MODELS (GLM, MLM, MLMM, FARMCPU AND BLINK) FOR TURKEY SEV



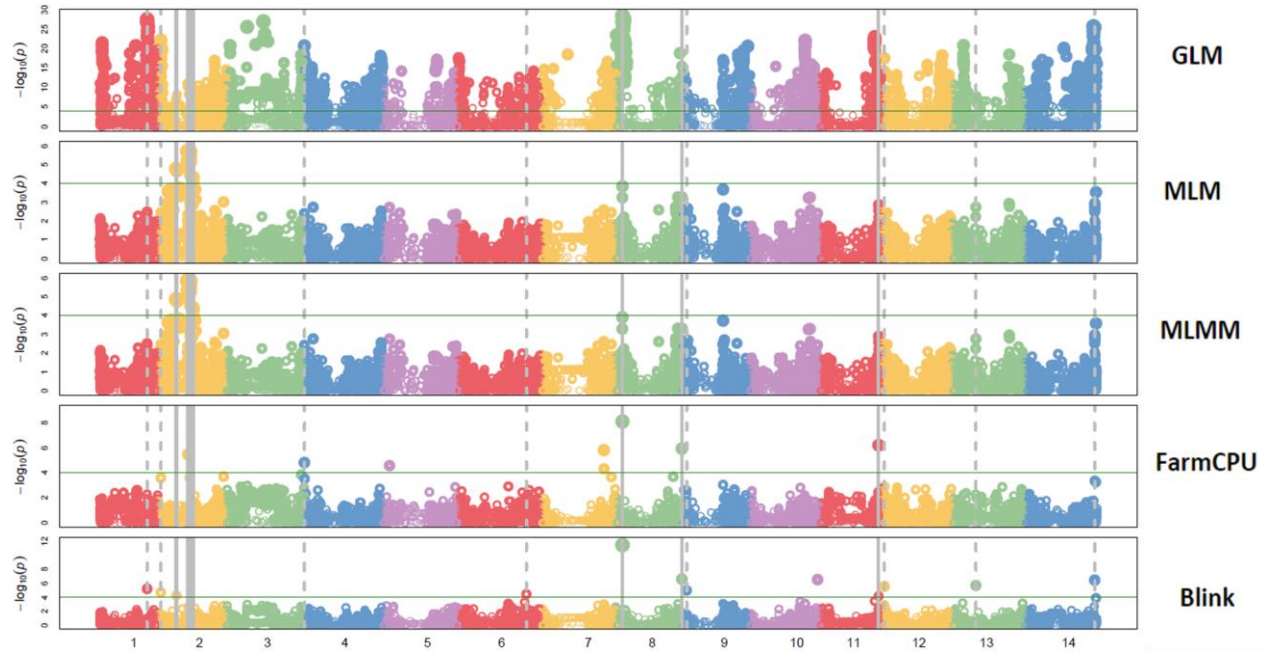
SUPPLEMENTARY FIGURE 21: MANHATTAN PLOT WITH ALL MODELS (GLM, MLM, MLMM, FARMCPU AND BLINK) FOR FOUR ENVIRONMENTS IT



SUPPLEMENTARY FIGURE 22: MANHATTAN PLOT WITH ALL MODELS (GLM, MLM, MLMM, FARMCPU AND BLINK) FOR FOUR ENVIRONMENTS SEV



SUPPLEMENTARY FIGURE 23: MANHATTAN PLOT WITH ALL MODELS (GLM, MLM, MLMM, FARMCPU AND BLINK) FOR THREE ENVIRONMENTS IT



SUPPLEMENTARY FIGURE 24: MANHATTAN PLOT WITH ALL MODELS (GLM, MLM, MLMM, FARMCPU AND BLINK) FOR THREE ENVIRONMENTS SEV