Alma Mater Studiorum – Università di Bologna

DOTTORATO DI RICERCA IN

SCIENZE BIOTECNOLOGICHE E FARMACEUTICHE

Ciclo XXX

Settore Concorsuale: 07/E1

Settore Scientifico Disciplinare: AGR/07

Fine mapping of *QSbm.ubo-2BS*, a major QTL for resistance to Soil-Borne Cereal Mosaic Virus (SBCMV) in durum wheat

Presentata da: Dott.ssa Martina Bruschi

Coordinatore Dottorato

Chiar.mo Prof. S. Mario Spampinato

Supervisore

Chiar.mo Prof. Roberto Tuberosa

Esame finale anno 2018

Table of contents

1. Abstract

QSbm.ubo-2BS, a major QTL controlling the response to Soil-Borne Cereal Mosaic Virus in durum wheat, maps within a 2 cM-wide interval in the distal region of chromosome arm 2BS.

The consensus map developed by Maccaferri et al. (2015) allowed to enrich *QSbm.ubo-2BS* interval with SNPs from the 90K Illumina array common to several linkage maps. Eleven SNPs were successfully converted to KASP markers, which provided fluorescent high-throughput assays spanning the QTL region. Marker-assisted selection was performed on ~3,000 RILs from the Svevo (resistant) x Ciccio (medium-susceptible) population with two KASP markers flanking the QTL interval, *KUBO 9* and *KUBO 13.*

MAS identified a total of 521 lines recombinant between the two markers. In particular, 320 and 189 of these lines were characterized for SBCMV response in a field nursery under severe and uniform SBCMV infection in 2016 and 2017, respectively. The lines were scored for symptom severity on a 0 to 5 scale, where 0 = very resistant and 5 = very susceptible.

The same lines were genotyped with six KASP markers distributed along the QTL interval (*KUBO 1*, *KUBO 3*, *KUBO 27*, *KUBO 29*, *KUBO 40* and *KUBO 41*) and with the DArT marker *wPt-2106.*

The fine mapping conducted on Svevo x Ciccio allowed to locate the most probable *QSbm.ubo-2BS* support interval to 0.2 cM between the markers *KUBO 27* and *KUBO 41.* Single-marker analysis showed that all markers in the interval are strongly associated with the QTL. Markers order was confirmed through genotyping of Meridiano x Claudio RILs with recombinant events in the interested interval.

QSbm.ubo-2BS support interval corresponds to 3.2 Mb on the physical map of Svevo genome, a region that harbors 93 genes, including defensins, proteins belonging to NBS-LRR family and NBS-LRR-like resistance proteins. This work provides the foundation for the positional cloning of *QSbm.ubo-2BS.*

2. Introduction

2.1. Durum wheat: from the field to the table

2.1.1. Origin and economic value

Durum wheat (tribe Triticeae, subtribe Triticinae, genus *Triticum*, species *Triticum turgidum* subsp. *durum*) is an annual grass member of the Poaceae or Graminaceae family. Native to the Mediterranean region and southwest Asia, it is one of several species of cultivated wheat, now grown in temperate climates worldwide for its cereal grain.

FAO estimated the global wheat production in 2014 of 729 million tons, on a surface of 220 million ha (FAOSTAT, 2017).

Durum wheat represents only 7-8% of total wheat production (USDA, 2017), but it has a high economic relevance, since it is one of the top two cereal crops grown in the world for human consumption, along with rice (*Oryza sativa*). Durum wheat is mostly cultivated in Mediterranean countries, North America and Australia. Italy and Canada are the main producers, with an estimate of 3.9 and 5.2 million tons produced in 2014 respectively.

Durum wheat was developed by artificial selection of the domesticated emmer wheat (*Triticum turgidum* subsp. *dicoccum*) strains formerly grown in Central Europe and Near East around 7000 BC. It has larger, harder grains, a higher protein content and lower gluten content than bread wheat (*Triticum aestivum*) varieties. Durum wheat varieties are hardy and drought-resistant, often grown in dry regions, and are mostly planted in the spring, in contrast to many bread wheat varieties, known as winter wheat because they are planted in the fall and allowed to overwinter.

Wheat is high in carbohydrates, protein and vitamins B and E; durum wheat is slightly lower than bread wheat in nutrient content. Because of its low gluten content, durum wheat is not used for bread and baked goods, but it is employed to produce pasta, semolina, the Middle Eastern bulgur and North African couscous.

2.1.2. Durum wheat and the Italian pasta industry

Pasta is the most traditional Italian product and a mainstay of the Italian diet. Pasta products, largely consumed all over the world, are traditionally manufactured from durum wheat semolina, a coarse-ground flour, known to be the best raw material suitable for pasta production (Petitot et al., 2010).

With more than 3 million tons produced in 2015 and an annual per capita consumption of 25 kg, Italy is the first pasta producer worldwide (UN.A.F.P.A., 2015). The country holds the leadership thanks to 150 manufacturing plants active on its territory, which employ 8,300 workers. Pasta is consumed in many other countries including the United States, Brazil, Russia, and Germany (Carloni et al., 2017).

The Italian national demand, now in its maturity phase, is subject to a slow increase due to a 100% of penetration rate in the Italian families and a gradual replacement of pasta with other goods.

The export of pasta, instead, has significantly increased in the last 15 years, reaching 2 million tons in 2014. It represents the 7% of the national agricultural export value (ISMEA, 2015).

The pasta industry has undergone a deep change in the last 40 years: while in the ‘60s it was mainly represented by small local businesses, the Italian pasta market is now controlled by few companies, which entered the market pursuing a strategy of geographical and product diversification.

The leader company of the pasta market is Barilla. The brand has been on the national market for 125 years and on the international one for 40 years. Its core business is the semolina pasta.

Barilla stands out in the world panorama also for its investments in new technologies: the highly automated pasta production plant in Parma is the biggest dry pasta producer in the world, producing 330,000 tons of pasta per year, 65% of which is exported. Barilla exports about one third of its production to more than 100 countries, a share likely to increase with the advancement of international trading activities.

Process and product innovation are key factors in Barilla’s success: they are the results of investments in research and development promoted by the company also thanks to public subsidies,

aiming to conceive new and healthier kinds of pasta, in addition to new wheat varieties like Svevo (Magnatti, 2007).

2.1.3. Durum wheat products other than pasta

Bulgur

Bulgur has been known in Central Asia, the Middle East, Balkans and Turkey for a long time. It is commercially manufactured from durum wheat through boiling in water until whole grain gelatinizes, drying under sunlight or in drying-towers about 12% moisture content, slightly debranning, breaking with mill to different particle sizes, sifting and classifying.

More than 1 million tons of bulgur per year are produced in Turkey alone: it represents an excellent food source due to its low cost, storability (long shelf life), ease of preparation and high nutritional value. Moreover, it resists mold contamination and attack by insects and mites and it is stable in hot and humid environments. It is also stored for military and human nutrition purposes in some countries, because of its resistance to absorption of radiation. Because of the properties listed above it is used for food aid in famine regions by the World Food Programme.

Bulgur is a traditional and functional wheat product that is available as ready and half-ready to eat; it is used to prepare more than 250 meals such as pilaf, soup, bulgur balls and salads. Currently, the demand for functional food, defined as food containing good bioactive components for health, is increasing all over the world in direct response to health consciousness and health-care costs. Bulgur can be evaluated as a functional food because of its nutritional components, especially with respect to B vitamins, minerals, dietary fiber and its low glycemic index (Bayram et al., 2004; Bayram and Öner, 2006; Erbaş et al., 2016).

Couscous

Originating from North Africa, couscous has become very popular in many other parts of the world.

Its popularity is due to the great taste and good nutritional benefits: couscous contains only 100–120 calories per half-cup serving and includes complex carbohydrates, vitamin B and minerals (Çelik et al., 2004).

The traditional method for making couscous consists in mixing water with durum wheat semolina, rubbing the mixture between palms of the hands to form small irregularly shaped granules, screening the granules to proper size, steam precooking the granules and finally sun-drying those of the proper size. Sun-dried couscous has a long shelf life (Donnely et al., 1994). Today, in many countries including Turkey, couscous is made mechanically using extrusion technology (Çelik et al., 2004).

Couscous is the national dish of Morocco, where it is generally served with vegetables (carrots, potatoes, turnips) cooked in a spicy or mild broth or stew, and some meat (generally chicken, lamb or mutton). Besides being part of North Africa traditions, couscous is progressing towards becoming a world-renowned food as the working population increases: packaged sets containing a box of quick-preparation couscous and a can of vegetables and, generally, meat are sold in French, Spanish, Italian, and Portuguese grocery stores and supermarkets. In North America, Australia and the United Kingdom, couscous is available most commonly as either plain or pre-flavoured, quick preparation boxes. This instant, pre-steamed couscous takes less time to prepare than regular couscous, most dried pasta, or dried grains (such as rice; Coskun, 2013).

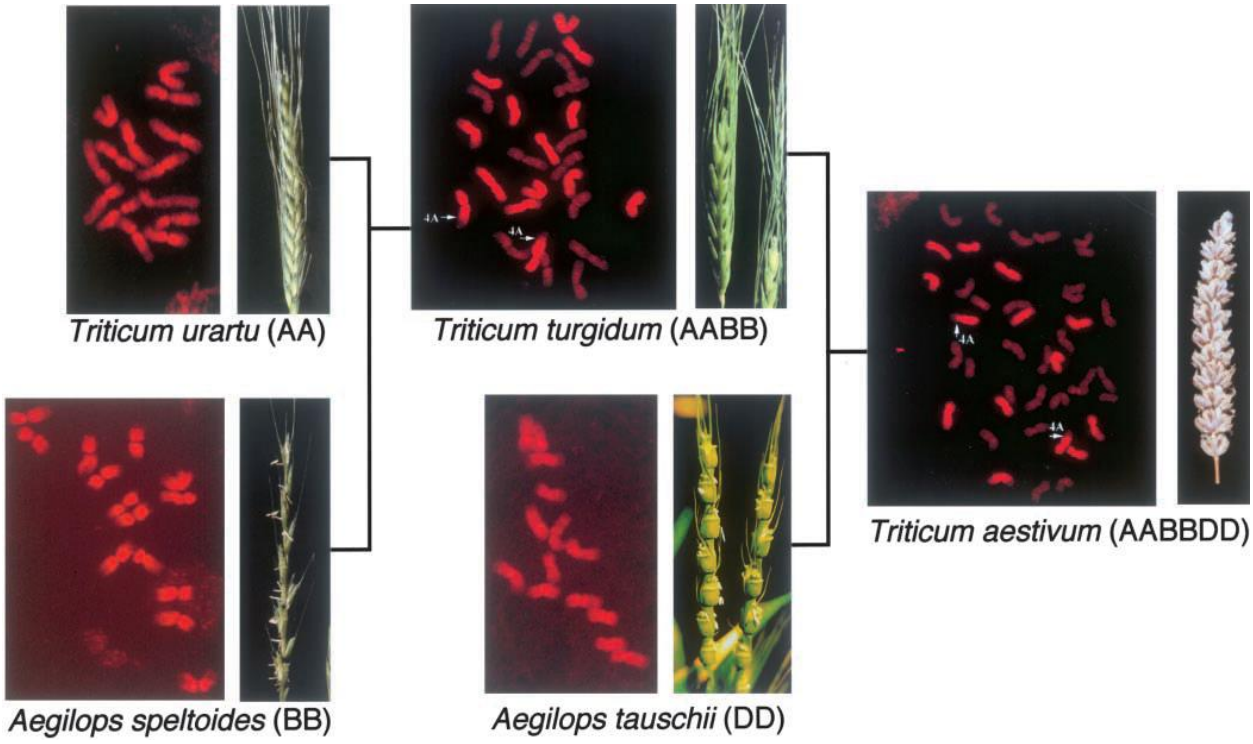
2.2. Genetics of durum wheat

Durum wheat has an allotetraploid genome (2n = 4x = 28 chromosomes, where x = 7), composed by A and B genomes (genomics formula = AABB). The donor of the A genome is *Triticum urartu* (AA genome, 2n = 2x = 14 chromosomes), an ancestor of the wild wheat. Because of the high conservation maintained during evolution, the A genome is considered the “pivot genome”, common to all wheat species. Origins of the B genome are not completely defined yet (Talbert et al., 1995; Zohary and Feldman, 1962). It has been theorized that B genome evolved from the S

genome belonging to an *Aegilops* species part of the *Sitopsis* section (van Slageren, 1994), and similar to the present *Aegilops speltoides* (Sarkar and Stebbins, 1956; **Figure 1**).

Wheat is the youngest polyploid species among the agricultural crops and the first one to be domesticated (Gill et al., 2004). It coevolved about 55-75 million years ago from an ancestor common also to rice and maize (Kellogg, 2001), though it differs greatly in genome size from these other two cereals. Bread wheat (*Triticum aestivum*, 2n = 6x = 42 chromosomes), which differs from durum wheat for the presence of the D genome in addition to A and B genomes (genomic formula = AABBDD), has a genome of 16 Gb (the largest among agricultural crops), ~8-fold larger than that of maize and 40-fold larger than that of rice (Arumuganathan and Earle, 1991). Durum wheat genome, whose sequencing has recently been completed by the International Durum Wheat Sequencing Consortium (Maccaferri et al., unpublished) measures ~12 Gb.

Figure 1. The evolution of wheat genomes (figure from Gill et al., 2004). Durum wheat (*Triticum turgidum* subsp. *durum*) derives from *Triticum urartu* and *Aegilops speltoides*, respectively donors of the A and B genomes. Bread wheat (*Triticum aestivum*) evolved from tetraploid wheat through its natural hybridization with *Aegilops tauschii*, donor of the D genome.



The progenitor of cultivated wheat, the diploid *Triticum monococcum* (AA), originated from the area of southeast Iran, northeast Iraq and southeast of Turkey. Between 15,000 and 10,000 B.C., inhabitant of these lands started harvesting and cultivating this kind of wheat, characterized by a fragile spike and caryopsis-adherent glumes. Starting from 9,000 years ago, selection operated by humans led to the isolation of ecotypes with consistent rachis, a characteristic that simplified the harvesting procedure. Domestication of tetraploid wheats, *Triticum turgidum* subsp. *dicoccum* (AABB) and *Triticum timophevii* (AAGG), took place at the same time in the area of the Fertile Crescent, Armenia and South Caucasus. Durum wheat genome presents homologies with both ancestors *Triticum turgidum* subsp. *dicoccum* (emmer wheat) and *Triticum turgidum* subsp. *dicoccoides* (wild emmer wheat). These two species are especially interesting for reintroducing ancestral useful alleles in cultivated wheat (Bianchi et al., 1989).

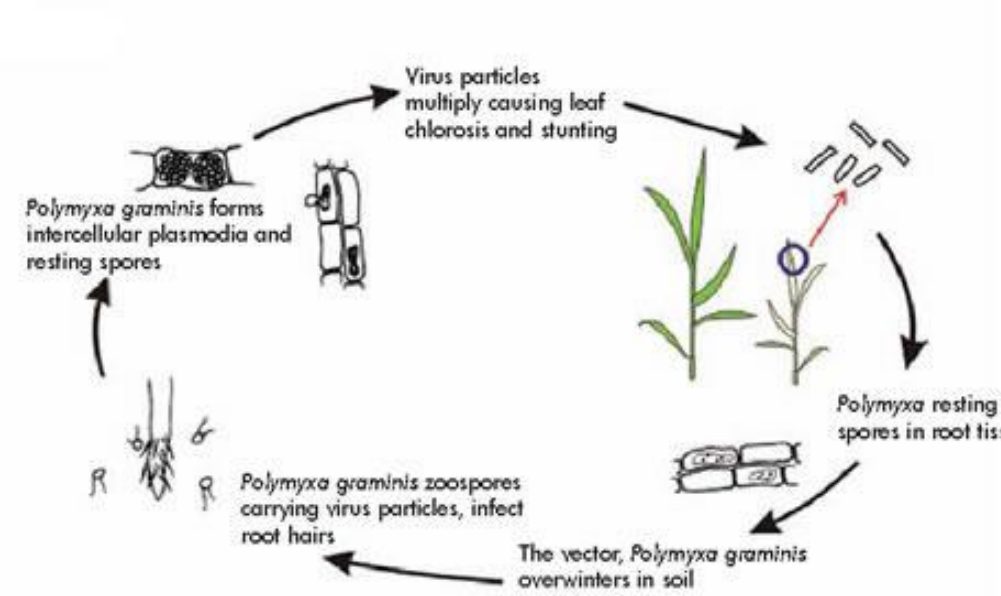
2.3. Soil-borne cereal mosaic virus

2.3.1. Mechanisms of the disease

Soil-borne cereal mosaic virus (SBCMV) is a furovirus that greatly affects the production of both durum and common wheat all over Europe (Budge et al., 2008; Rubies-Autonell et al., 2003). For a long time, and also presently by some authors, it has been considered the European strain of soil-borne wheat mosaic virus (SBWMV), which is also member of the genus Furovirus, but affects mainly wheat crops in U.S.. Recently, the International Committee on Taxonomy of Viruses (ICTV) has recognized SBCMV as a separate species than SBWMV (Kanyuka et al., 2004). The name “soil-borne” is due to the mechanism of infection, which takes advantage of the soil inhabitant plasmodiophorid *Polymyxa graminis*, used as a vector by the virus. The virus nests inside the resting spores of the vector where finds protection during periods of adverse environmental conditions and during long rotations of non-host crops, persisting in the soil also for decades. When the conditions return favorable (susceptible hosts and high soil moisture), the virus is released from the resting spores and, vectored through zoospores, enters plant’s root hair cells (Kühne, 2009).

From there, the virus is able to reach the aerial part of the plant and to infect the leaf tissue (**Figure 2**).

Figure 2. Infection mechanism of Soil-Borne Cereal Mosaic Virus through its vector *Polymyxa graminis*, a soil inhabitant plasmodiophorid.



SBCMV can cause yield losses of up to 70% in durum wheat, as reported by Vallega and Rubies Autonell (1985). Virus symptoms, which generally appear in early spring, range from a mild mottling to severe leaf mosaic. Also vigor and plant height are affected (stunting) (Kühne, 2009).

Both Vallega et al. (2003) and Maccaferri et al. (2011) reported a significant correlation between disease severity in durum wheat and yield and its components, such as thousand kernel weight.

Due to the nature of the infection, which once has infested a field can persist there indefinitely, agronomic practices like for example crop rotation or delayed sowing are not useful. Chemical defense is either not effective or not environment-friendly. Thus, the only possible solution is the development and employment of resistant cultivars (Ordon et al., 2009).

2.3.2. Genetic basis of resistance towards SBCMV in wheat

Due to the importance that the development of genotypes resistant to SBCMV has acquired as the only solution to this threat, genetics bases of SBCMV resistance have been thoroughly studied both in bread and durum wheat.

Resistance towards SBCMV was studied in the bread wheat UK cultivar Cadenza by Kanyuka et al. (2004). They demonstrated that a single locus, *Sbm1*, is responsible for SBCMV resistance in this cultivar. Resistance deriving from *Sbm1* is inherited in a monogenic manner and consists in a mechanism called "translocation resistance" that prevents the infection to spread from the roots to the stem and leaves. Bass et al. (2006) were able to locate *Sbm1* in the distal portion of the long arm of chromosome 5D, flanked by a set of SSR markers useful for both MAS and fine genetic mapping studies. The same kind of resistance, though not related by pedigree to the cultivar Cadenza, was found in the French cultivar Tremie and in the UK cultivar Claire (both hexaploid wheat cultivars) by Perovic et al. (2009). This resistance was located on chromosome 5DL as well and a SSR marker diagnostic for this locus was developed, which was found to co-segregate with *Sbm1* in Cadenza.

Another major locus controlling SBCMV resistance in bread wheat, *Sbm2*, was mapped by Bayles et al. (2007) in the short arm of chromosome 2B and was shown to act in distinct and complementary way to *Sbm1*.

SBCMV resistance in durum wheat was extensively studied by Maccaferri et al. (2012, 2011) by using a population of 181 recombinant inbred lines (RILs) obtained from Meridiano (resistant) x Claudio (moderately susceptible) Italian elite cultivars. The RILs were characterized for SBCMV response in the field under severe and uniform infection conditions during 2007 and 2008. A large portion of the variability for SBCMV response was explained by a major QTL, *QSbm.ubo-2B*, located in the distal telomeric region of chromosome 2BS, close to the DArT marker *wPt-2106* and tagged by the SSR triplet X*wmc661-Xgwm210-Xbarc35*, defined as "the resistance-tagging haplotype". Subsequently, a similar analysis was carried out using a population of 180 RILs obtained from Simeto (susceptible) x Levante (resistant), in which *QSbm.ubo-2B* was reported to

account for 60-70% of the phenotypic variation. The meta-QTL analysis performed on the data of Meridiano x Claudio and Simeto x Levante populations allowed to confine *QSbm.ubo-2B* to a 2 cM-wide interval, tagged by SSR markers functional to conduct marker-assisted selection in breeding programs. In both populations, several minor QTLs (11 in Meridiano x Claudio and 7 in Simeto x Levante) were found to influence the response towards SBCMV infection.

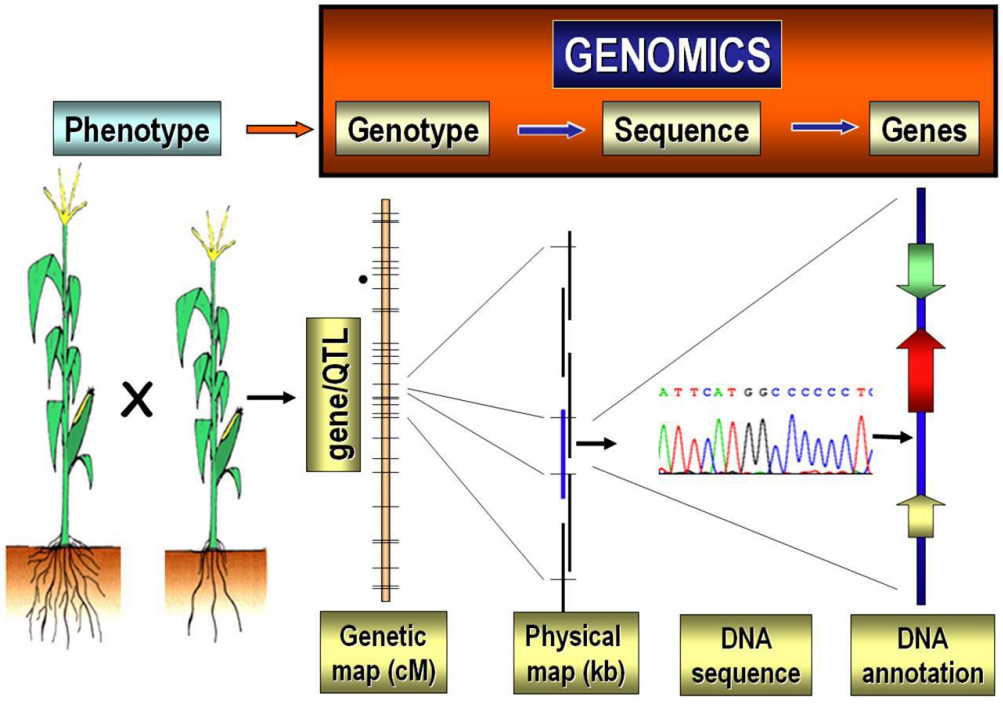
2.4. Fine mapping strategy

2.4.1. QTL fine mapping

QTLs (Quantitative Trait Loci) are defined as genetic loci where functionally different alleles segregate and cause significant effect on a quantitative trait (Salvi and Tuberosa, 2005). QTL mapping is composed by different steps. Initially, phenotypic and genotypic data (the latter obtained with molecular markers) about a segregating population are collected and statistical analysis are performed in order to detect loci where genotype and phenotype correlate. This strategy, which is now a standard procedure in quantitative genetics, allows only for a coarse mapping of QTLs, which are usually located in more or less wide intervals (~10-30 cM), and is thus considered a preliminary result that needs to be refined. QTL fine genetic mapping, which can be considered a second step in the QTL mapping strategy, consists in assigning the QTL to the shortest possible genetic interval, possibly at the sub-cM level. Fine mapping can be in turn subdivided in three steps: 1) production of a new segregant population, which possibly harbors a high amount of recombination events; 2) enrichment of the QTL region with new molecular markers; 3) comparison between the new support interval defined for the QTL on the genetic map and the same interval on the physical map and identification of candidate genes. The enrichment of the QTL region with molecular markers allows for a higher map resolution in the segregant population, necessary to further refine the QTL support interval. Once the genetic resolution is close to the cM level, the markers closest to the QTL can be used to anchor the genetic map to the physical map, i.e.

the genomic sequence, when available, or a BAC contig covering that region (QTL physical mapping). Sequence information available for the region of interest will then allow for the identification of the putative candidate genes. In this phase, gene prediction and annotation are performed through bioinformatics tools. QTL fine genetic mapping followed by QTL physical mapping and identification of candidate genes should eventually lead to the positional cloning of the QTL (Salvi and Tuberosa, 2005). **Figure 3** displays a scheme of the steps required for positional cloning.

Figure 3. Steps of the procedure that leads to positional cloning, from the acquisition of phenotypic and genotypic data to the individuation of candidate genes (figure from Graziani, 2012, modified from Varshney and Tuberosa, 2007).



2.4.2. Marker-assisted selection

QTL identification based only on phenotypic evaluation is not possible. A major progress in the characterization of QTLs was represented by the development of molecular markers in the '80s.

Markers linked to agronomically relevant genes (called gene "tagging") can be used as molecular tools for marker-assisted selection (MAS) in plant breeding (Collard et al., 2005).

In practice, MAS is based on genotypes rather than on phenotypes. This allows avoiding the phenotypic evaluation of plants, which can often be tedious, time-consuming, season and weather-dependent (Singh and Singh, 2015), especially if protocols for nurseries/greenhouse trials are not available. For these reasons, MAS can be a useful tool when applied to the fine mapping of QTLs. In this case, lines of a segregating population can be selected based on their genotype at markers flanking the region where the QTL is located. The goal is selecting lines that present a recombination between the two flanking markers i.e. that present a recombination in the confidence interval of the QTL. Selected lines are then further genotyped with markers mapping between the two flanking markers and phenotyped, in order to further narrowing the QTL support interval.

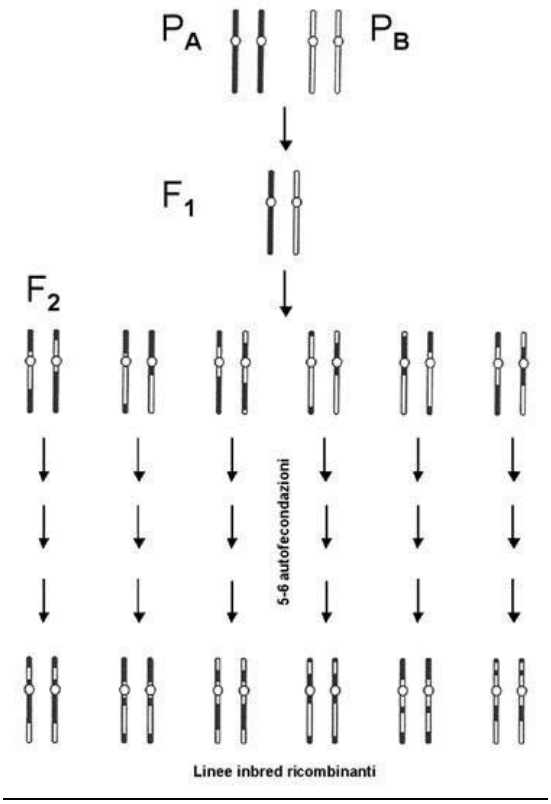
2.4.3. Recombinant Inbred Lines

A Recombinant Inbred Line (RIL) is produced from the cross of two homozygous parents (pure lines) followed by 5-6 generations of repeated selfing (**Figure 4**). In the F_6 generation, plants are almost completely homozygous and each RIL genome is a mosaic of the parental genomes (Broman, 2005).

RILs are usually derived from F_2 populations through a single seed descent (SSD) procedure. This technique consists in harvesting one seed from each plant starting from the F_2 generation. For each generation, seeds from all the plants are composited and planted to raise the next generation. This procedure is usually conducted for five or more generations. At the end, seeds from each plant are harvested separately and the number of RILs obtained is equal to the number of individual plants in the SSD population. At each generation of selfing, heterozygosity is reduced to one-half of that present in the previous generation, and homozygosity is consequently increased. The final RIL population consists of the two homozygotes for a locus at an expected ratio of 1:1, and a rather

small proportion of the heterozygote that depends on the number of generations up to which the SSD procedure was conducted (Singh and Singh, 2015).

Figure 4. Assembly of Recombinant Inbred Lines (figure from Barcaccia and Falcinelli, 2007). RILs are obtained through the cross of two pure lines (parent A and parent B), followed by 5-6 generations of repeated selfing. In the F_6 generation, RILs are almost completely homozygous.



RILs employed as genetic material in QTL fine mapping present many advantages. First, each strain needs to be genotyped only once. Moreover, RILs allow to obtain a high mapping resolution, since recombination events that accumulates in progressive generations are denser than those that occur in a single meiosis (Broman, 2005). Finally, seed of each line can be multiplied indefinitely and thus the line propagated eternally (immortalized), allowing for replicated trials in space (different locations) and time (different years).

2.4.4. SNP markers

Single-nucleotide polymorphisms (SNPs) are single base pairs variations among individuals of a species in a determinate site of their genome. Only nucleotide polymorphisms for which the least frequent allele is present at a frequency ≥ 1% are considered SNPs. A SNP locus can present up to four alleles (corresponding to the four DNA nucleotides); however, SNPs are usually scored as biallelic markers. SNPs abundance in plant genomes is about one SNP every 100-300 bp; their mutation rate is relatively low and they are easily detectable (Singh and Singh, 2015). Their frequency makes them suitable to be used in marker enrichment of QTL regions for fine mapping purpose. Moreover, SNPs are amenable to robotic management, i.e. they can be employed in array platforms and in various high-throughput genotyping technologies, like KASP (Kompetitive Allele Specific PCR) and HRM (High Resolution Melting). For these reasons, SNPs are presently the most widely utilized markers, extensively exploited in crops genetics (Maccaferri et al., 2015).

Recently, wheat-dedicated Illumina iSelect platforms containing 9,000 and 90,000 SNPs have been developed by Cavanagh et al. (2013) and Wang et al. (2014), respectively. These arrays present numerous advantages, like high call frequency, ease of use, high accuracy and repeatability, reduced computational requirements for downstream data processing and low error rate (Wang et al., 2014), which make them highly valuable genotyping tools. SNPs in 9K and 90K arrays are transcripts-derived, i.e. associated to genes. This represents a convenient feature when it comes to the identification of candidate genes, once a QTL region has been enriched with these SNPs. Both 9K and 90K SNPs arrays were used to construct consensus maps based on SNPs. Consensus maps are useful to overcome limitations represented by linkage maps based on independent crosses, in particular to extend mapped genome portions, increase local genetic resolution and validate marker order (Maccaferri et al., 2015; Salvi et al., 2009).

Maccaferri et al. (2015) developed a consensus map integrating genotyping datasets of 13 tetraploid independent biparental populations involving durum wheat, cultivated emmer and wild emmer genotypes. Nine of these populations were genotyped with the 90K and one with the 9K SNP

Illumina arrays. The final map spans 2,631 cM of all 14 durum wheat chromosomes. In addition to SNPs, the consensus map includes SSR and DArT markers, for a total of 30,144 markers, which to date make it the consensus map with the highest number of markers. Since most of these SSR and SNP markers were previously mapped in bread wheat, this map can be considered a bridge for exchanging and extrapolating genomic information between durum and bread wheat, representing a valuable tool for mapping in the A and B genomes of both species. Moreover, useful anchor points for positional cloning are provided by the high density of gene-derived SNPs.

2.4.5. KASP (Kompetitive Allele Specific PCR) technology

The KASP® genotyping system is a homogeneous, fluorescent, endpoint genotyping technology. It allows for the genotyping of individuals based on SNP variation. The chemistry reaction of KASP system, described hereafter, is based on the database information available on the website www.lgcgroup.com/products/kasp-genotyping-chemistry/.

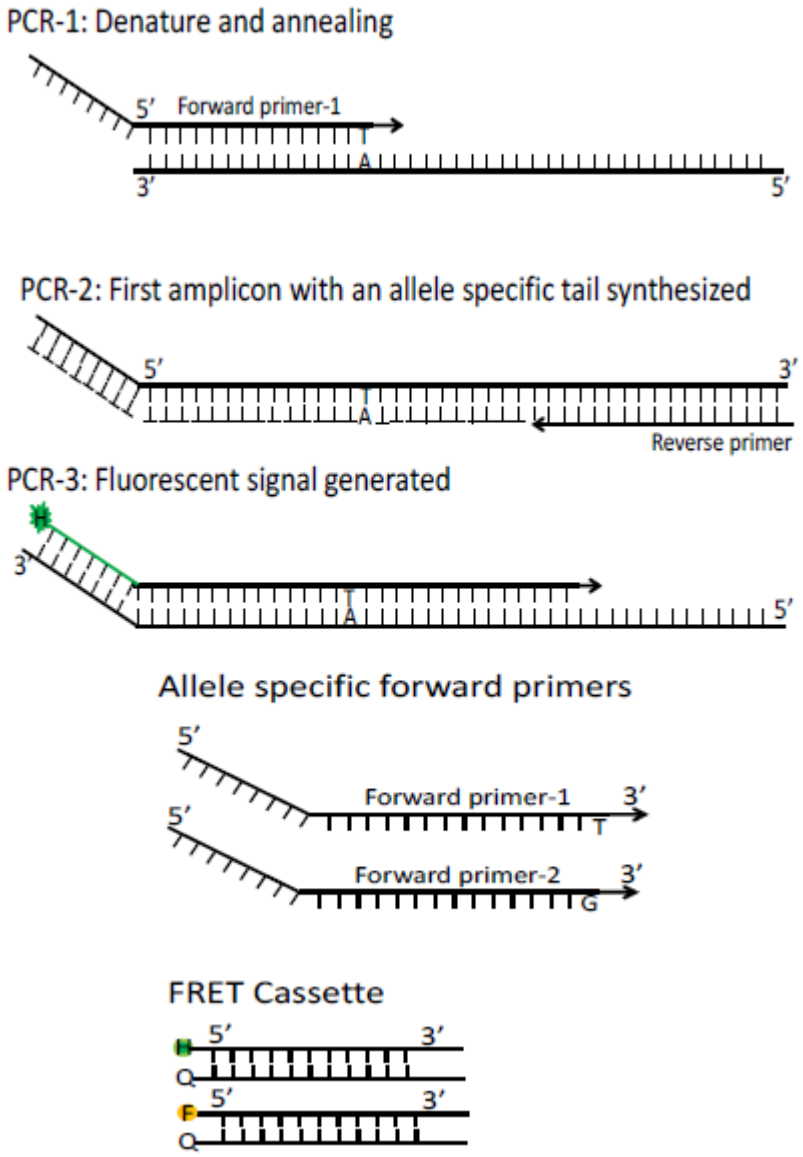
KASP PCR reaction is based on three components: the DNA template, containing the target sequence harboring the SNP, the primer mix and the KASP master mix, containing all other reagents necessary to PCR reaction. The primer mix is composed by three primers: two allele-specific primers A and B, one for each SNP allelic variant, and one common primer C. Primers A and B present unique tails at 5' ends, which are the same sequences as the nucleotides contained in the FRET (Fluorescence Resonance Energy Transfer) cassettes contained in the KASP master mix. KASP master mix contains the two FRET cassettes, the passive reference dye ROX, the Taq polymerase enzyme and $MgCl_2$, all reagents mixed in an optimized buffer. FRET cassettes contain oligonucleotide sequences marked with FAM and HEX fluorochromes respectively and are "universal", since they function with any primer containing the respective FAM/HEX probe sequence tail at 5'. In the KASP master mix, the proximity of reporter FAM/HEX and quencher in the FRET cassette prevents the emission of any fluorescence from the reporter (Yuan et al., 2014).

During the first PCR cycle, the two allele-specific primers compete for the same annealing site, but only one, specific for the target SNP variant contained in the DNA template, will be able to pair correctly, allowing for the extension. The tail at 5' does not anneal to the template and consequently is not copied, but it is incorporated in the amplification product. During the extension step of the second PCR round, the tail is copied and a complementary sequence is generated. During the third PCR round, the corresponding FRET oligonucleotide in the master mix pairs with the sequence complementary to the tail and is incorporated in the amplification product. When the FRET cassette binds to the tail, the 5'-3' exonuclease activity of the Taq DNA polymerase cleaves the FRET, separating the reporter from the quencher and allowing the fluorescence to be detected (Yuan et al., 2014; **Figure 5**). During the following rounds of PCR, many amplicons are generated, whose detected FAM/HEX fluorescence corresponds to the SNP variant enclosed in the DNA template. At the end of the PCR cycle, the temperature is decreased to inactivate the FRET oligonucleotides that were not incorporated in amplification products. The final PCR product will emit fluorescence corresponding to FAM or HEX if the sample is homozygous for one of the two alleles, corresponding to both FAM and HEX if the sample is heterozygous.

Once the KASP reaction is completed, the fluorescence of each sample is measured with a plate-reader connected with a software. The software visualizes the output as a cluster plot. In the graph, the fluorescence value of HEX for each sample is plotted on the x-axis, while the fluorescence value of FAM is plotted on the y-axis (it can be the opposite depending on the software). A sample homozygous for the allele reported on the x-axis (which is conventionally indicated by the software as "allele 1") will generate a high HEX signal and no FAM signal and will thus be represented as a dot positioned in the lower-right area of the plot. Likewise, a sample homozygous for the allele reported on the y-axis (indicated as "allele 2") will generate a high FAM signal and no HEX signal and will be represented as a dot in the upper-left area of the plot. Heterozygous samples will generate both HEX and FAM signals, each signal having an intensity equal to the half of that generated by the homozygous samples, and will thus be represented in the central area of the graph.

Since all samples having the same genotype will generate a similar signal, they will be represented in the same graph area, generating clusters. Based on cluster position, the software will assign to each sample a color and one of the following genotypes: “homozygous 1/1”, “homozygous 2/2”, “heterozygous 1/2” and “undetermined”. An undetermined genotype corresponds to a missing data and is represented on the graph as a black cross. The software generates an output Excel file with the list of samples and corresponding genotypes.

Figure 5. KASP PCR reaction (figure from Yuan et al., 2014). In the first PCR round (PCR-1), the allele-specific primer anneals to the DNA template. In the second PCR round (PCR-2), the allele-specific tail is copied and a complementary sequence is generated. In the third PCR round (PCR-3), the corresponding FRET oligonucleotide pairs with the complementary sequence and the fluorescent signal is generated. Following, the two allele-specific primers with the FAM/HEX sequence at their 5' ends and the two FRET cassettes contained in the KASP master mix are represented.



3. Objectives

Sustainable production of crop plants is one of the main challenges faced by mankind, not only for economic or environmental reasons, but also because of the strong need of finding solutions to the challenge of feeding a rapidly growing population, which is estimated to reach 10 billion by 2050, a particularly daunting issue in view of the dwindling availability of arable land per person (Kühne, 2009).

Durum wheat, as one of the main cereal crops destined to human consumption, is at the center of this issue, particularly in Mediterranean countries where it provides the staple to tens of millions. Moreover, as previously explained, the durum value chain has a high economic value and is at the top list in the food market worldwide.

Soil-borne cereal mosaic virus, with its potential to curtail yield up to 70%, represents a serious threat for durum wheat production. Since the only solution to face this disease is the selection and cultivation of resistant cultivars, it is important to gain a good knowledge of the genetic basis of virus resistance in durum wheat.

In the study conducted by Maccaferri et al. (2012), *QSbm.ubo-2BS* was located in a ~2-cM-wide interval, based on the joint analysis of two distinct molecular and phenotypic datasets. The interested chromosomal region, defined as gene-rich and highly recombinogenic, was marked by a set of SSR markers, which were also useful to carry out MAS, in order to increase the SBCMV resistance, as evidence of the applicative implication of this kind of research.

In order to proceed towards the fine mapping of *QSbm.ubo-2BS*, it was first necessary to enrich the interested region with new molecular markers. Along this line, Maccaferri et al. (2015) were able to construct a high-density, SNP-based consensus map of tetraploid wheat, which for the first time integrated SSR, DArT and SNP markers in a single framework. The consensus map integrates data obtained from thirteen independent biparental populations: nine *T. durum* 'élite x élite', three *T. dicoccum* x *T. durum* and one *T. dicoccoides* x *T. durum*. These populations include the RILs Meridiano x Claudio and Simeto x Levante, i.e. the plant materials used in the previous studies

about *QSbm.ubo-2BS* (Maccaferri et al., 2011, 2012), and the Svevo x Ciccio RILs chosen for this study. Maccaferri et al. (2015) used the Illumina 90K SNP array (Wang et al., 2014) to enrich the individual linkage maps of SNP markers, allowing mapping several of them in the region were *QSbm.ubo-2BS* had been located and in its surroundings.

SNP markers presented many advantages for the present study, mainly the high density with which they are present in the wheat genome and the possibility to be converted in KASP markers, which are amenable to robotic management. Moreover, SNPs are now widely used within MAS pipelines of breeding programs and thus the shift from SSR to SNP markers represents a step toward a more practical implementation (Ramirez-Gonzalez et al., 2015a).

Based on these premises, the research herein described targeted the following objectives:

1) Development of KASP markers from SNPs mapped in the *QSbm.ubo-2BS* interval
2) Marker-assisted selection (MAS) with KASP markers flanking *QSbm.ubo-2BS* interval on the RIL population Svevo x Ciccio
3) Fine mapping of *QSbm.ubo-2BS* on the selected RILs by means of KASP markers mapping in the interval
4) Comparison with the physical map and identification of putative candidate genes responsible for resistance towards SBCMV.

4. Materials and methods

4.1. Plant materials

4.1.1. Svevo x Ciccio RILs

The first population employed in this study included 3,075 RILs developed from a cross between the durum wheat commercial cultivars Svevo (hereafter referred to as Sv) and Ciccio (hereafter referred to as Cc). The RIL population was developed by UNIBA (Gadaleta et al., 2009) by advancing random individual F_2 plants to the F_7 generation by single seed descent.

Sv, a Cimmyt (pedigree: rock/fg//stil/3/dur1/4/sapi/teal//hui)/Zenit line, was released in 1996 by Produttori Sementi Bologna SpA (now part of Syngenta) for the Barilla production chain. Sv is a spring, medium-high, very early heading variety selected for the high yellow index and the very high protein content. It also presents several characteristics suitable for production: medium-high potential yield, good test weight and gluten quality, good resistance to powdery mildew and medium resistance to low temperature, leaf rust, septoria and lodging. Sv is well adapted to the Mediterranean environment.

Sv was used in the Italian sequencing project of durum wheat, a collaboration that involved several Italian and international institutions and companies: CREA, CNR and University of Bologna (Italy), Crop Development Centre of the University of Saskatchewan (Canada), University of Tel Aviv (Israel), IPK Gatersleben (Germany), Montana State University (USA) and NRGene (Israel), a leading company in genome sequencing. The sequence information produced from this project can be useful in many research lines involving Sv as plant material.

Sv is also the parent of two additional RILs, Kofa x Sv and Sv x Zavitan, which, with Sv x Cc and other mapping populations, were used by Maccaferri et al. (2015) for the assembly of a high-density tetraploid wheat consensus map harbouring 30,144 markers.

Cc (pedigree: Appulo/Valnova//Valforte/Patrizio), a medium-size variety, presents medium-high protein content, medium yellow index, high hectoliter weight and gluten tenacity, medium-early

heading time, medium resistance to powdery mildew and cold and high resistance to brown rust and lodging.

The RIL population Sv x Cc chosen for this study is a cross between two varieties, which have respectively a resistant phenotype (Sv) and a medium-susceptible phenotype (Cc) towards SBCMV. Moreover, Sv was shown to harbour the SBCMV-resistance haplotype at *QSbm.ubo-2BS*, which on the other hand was not found in Cc (Maccaferri et al., 2011).

RILs are particularly suitable for this study, since the vast majority of lines are homozygous for either the resistant or the susceptible allele at each marker in the QTL interval and thus are useful for selecting the recombinants required to narrow the QTL interval in the fine mapping strategy. Moreover, the high number (3,075 in total) of lines in the Sv x Cc population increases the probability of finding recombinants. Finally, the high quantity of seed available for each line allowed us to perform replicated field experiments, which are useful to control environmental variation.

4.1.2. Meridiano x Claudio RILs

The second population employed in the study includes 181 RILs developed by Produttori Sementi Bologna SpA from a cross between two elite Italian cultivars, Meridiano (hereafter referred to as M) and Claudio (hereafter referred to as C), advanced until the $F_{7:8}$ generation by single seed descent.

M (pedigree: Simeto/WB881//Duilio/F21) is a medium-early heading cultivar, while C (pedigree: CIMMYT selection/Durango//ISI938/Grazia) is medium-late. Both varieties were released in Italy during 1999 and are widely cultivated across Europe. They have high grain yield potential across southern Europe (Maccaferri et al., 2011).

M x C RIL population had already been used by Maccaferri et al. (2011, 2012) to dissect the genetic basis of SBCMV resistance in durum wheat, since it represents a cross between a cultivar resistant (M) and a cultivar moderately susceptible (C) to the virosis.

4.2. Development of new molecular markers in the *QSbm.ubo-2BS* interval

4.2.1. Conversion of SNPs to KASP markers

Based on the consensus map developed by Maccaferri et al. (2015), KASP assays were developed from sixteen SNPs previously mapped in a 11.4 cM interval in the *QSbm.ubo-2BS* region. The interval spans up to 3.0 cM upstream and 8.4 cM downstream of *wPt-2106* (the marker most associated with the phenotype in the M x C RIL population, Maccaferri et al., 2011), a region proximal to the SSR triplet *Xwmc661-Xgwm210-Xbarc35*, and includes the DArT marker *wPt-1601*. Fifteen of these SNPs are polymorphic in M x C, fifteen in Sv x Cc and seven in Simeto x Levante**. Table 1** reports the main SNP characteristics.

Primers were designed by means of the software PolyMarker (http://Polymarker.tgac.ac.uk), an automated bioinformatics pipeline used to convert SNP markers into codominant genome-specific KASP assays (Ramirez-Gonzalez et al., 2015a). Each KASP assay is based on two allele-specific primers and a common primer. The allele-specific primers are designed on a varietal SNP, which is a single nucleotide variation between the sequences of different wheat lines on the same genome. The common primer is designed on a homoelogous SNP, namely a single nucleotide variation between the A and B genomes of tetraploid wheat. In fact, since the allele-specific primers are restricted in position, the common primer is used to incorporate the chromosome specific variants when possible (Ramirez-Gonzalez et al., 2015b). Designing primers with PolyMarker therefore allows for both the locus and genome specificity required in allopolyploid species such as wheat.

Table 1. SNPs used to develop KASP markers. SNPs information are reported as in the Supplemental Table S2 of Maccaferri et al. (2015). cM (= centiMorgans) are reported as cumulative genetic distances on chromosome 2B according to the consensus map. RILs maps are the linkage maps of individual populations on which the SNP resulted polymorphic. LT = Latino, CL = Claudio, LD = Lloyd, MR = Meridiano, MH = Mohawk, CR = Cocorit69, SV = Svevo, CC = Ciccio, ZV = Zavitan, SM = Simeto, LV = Levante, ML = Molise Colli.

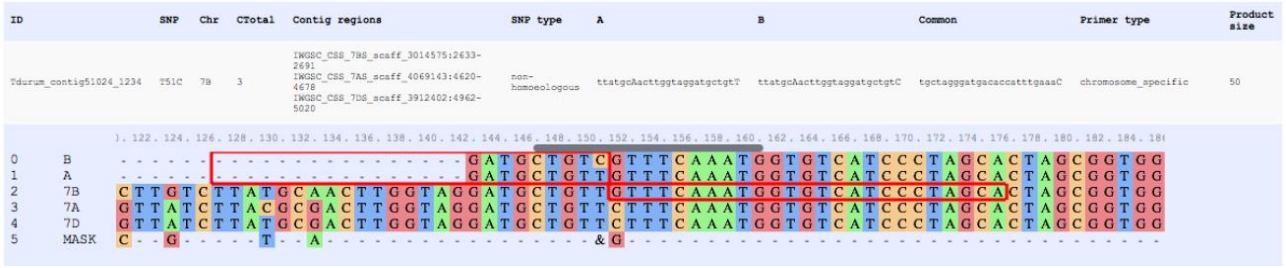
Marker ID code	**Marker name**	**cM**	**RIL maps**
IWB61884	RAC875_rep_c109471_154	7.9	LT×MG_5323, CL×LD, MR×CD, MH×CR, SV×CC
IWB73347	Tdurum_contig76118_145	8.4	LT×MG_5323, CL×LD, MR×CD, MH×CR, SV×CC, SV×ZV
IWB11421	BS00085748_51	11.6	LT×MG_5323, CL×LD, MR×CD, SM×LV, MH×CR, SV×CC, SV×ZV
IWB23029	Excalibur_c1787_1037	11.6	CL×LD, MR×CD, SM×LV, SV×CC
IWB42660	Kukri_c22513_1780	11.6	LT×MG_5323, CL×LD, MR×CD, SM×LV, MH×CR, SV×CC, SV×ZV
IWB28973	Excalibur_c8093_82	12.2	CL×LD, MR×CD, MH×CR, SV×CC
IWB45152	Kukri_c41556_619	12.2	CL×LD, MR×CD, MH×CR, SV×CC
IWB41644	Kukri_c16758_443	12.2	SM×ML, CL×LD, MR×CD, MH×CR, SV×CC
IWB8328	BS00043055_51	12.3	LT×MG_5323, CL×LD, MR×CD, SM×LV, MH×CR, SV×CC
IWB29097	Excalibur_c841_609	12.3	SM×ML, CL×LD, MR×CD, MH×CR, SV×CC
IWB35524	IAAV8700	12.3	SM×ML, CL×LD, MR×CD, MH×CR, SV×CC
IWB23330	Excalibur_c19499_948	12.4	CL×LD, MR×CD, SV×CC
IWB24939	Excalibur_c30167_531	12.4	SM×ML, CL×LD, MR×CD, MH×CR, SV×CC
IWB6204	BS00010318_51	19.0	MR×CD, SM×LV, MH×CR, SV×CC, SV×ZV
IWB10512	BS00070900_51	19.0	MR×CD, SM×LV, MH×CR, SV×CC, SV×ZV
IWB8390	BS00045163_51	19.3	LT×MG_5323, CL×LD, SM×LV, MH×CR

The handbooks “Designing genome specific primers” and “How to use PolyMarker” hosted on the website http://www.wheat-training.com were used to prepare an input .csv file containing the following information about the SNPs to be converted: the SNP name (or ID), the target chromosome and genome to amplify from and the sequence containing the varietal SNP allele in square brackets (example: [A/C]). For each SNP the information found in Table S2 of Maccaferri et al. (2015) was entered. Once the file has been uploaded on the PloyMarker website, the software generates a multiple alignment between the target SNP sequence and the IWGSC chromosome survey sequences (International Wheat Genome Sequencing Consortium, 2014) for each of the three wheat genomes and produces two output files. The first one, named “Primers”, contains the sequences of the designed primers: the A and B allele-specific primers incorporate the alternative varietal nucleotides at their 3’ ends; the common primer incorporates the homoeologue-specific or semi-specific base at the 3’ end (Ramirez-Gonzalez et al., 2015a). Each assay specifies whether the varietal SNP is homoeologous or non-homoeologous (i.e. monomorphic in both genomes of the same variety/cultivar). Non-homoelogous SNPs are preferred because they allow to discriminate between two varieties/cultivars, requiring that each of them has the same allelic variant in both genomes. The field “primer_type” describes the type of genome specificity achieved with the designed primer. “Specific” means that the common primer is genome specific to the target chromosome; “semi-specific” means that the common primer discriminates between the target genome and one, but not both, of the other two genomes (in case of tetraploid wheat this is relevant only if the common primer cannot discriminate between A and B genomes); “non-specific” means that no variation could be found between the target genome and non-target genome to design gene-specific common primers. The output file also reports on the size of the amplicon expected from using the primer pair and on the errors that might be encountered during the design process.

Whenever a chromosome-specific marker for a given SNP was unavailable, the second output file, named “Mask” (**Figure 6**) was used to manually design primers. This file contains the graphical details of the alignment performed by the software. The mask highlights the varietal SNP and

various chromosome-specific and semi-specific bases located near the varietal SNP, including the one incorporated in the common primer. Chromosome-specific bases (bases that were different between the chromosomes 2A and 2B) were used to manually design various common primers that were coupled with the same A and B allele-specific primers. In some cases, it was possible to obtain different assays (up to three) for the same SNP, changing only the common primer.

Figure 6. Example of output file "Mask" generated by the software PloyMarker (figure from Ramirez-Gonzalez et al., 2015b). The details of the assay are reported in the upper part of the figure: the SNP ID, its local position within the marker in the input, the target chromosome, the total number and list of contigs where the marker maps, the SNP type, the sequences of the two allele-specific primers A and B and of the common primer, the primer type and the product size. The lower part of the figure shows a schematic of the alignment performed by the software, organized in rows from 0 to 5: 0 - 1 sequence of the markers, with the designed allele-specific primers A and B enclosed by a red outline; 2 – 4 local alignment to the chromosomes where the marker maps with the reverse complement of the designed common primer enclosed by a red outline; 5 mask highlighting the feature of each base: the ampersand (&) indicates the position of the varietal non-homoeologous target SNP and the bases indicate candidate starting positions for specific or semi-specific common primers, included the one used to design the assay common primer.



4.2.2. Validation of KASP markers on parental genotypes

The developed KASP markers were further validated on the parental lines of the RIL populations used in this study (M x C and Sv x Cc). To test KASP assays, DNA samples of the four parental lines (M, Sv, C and Cc) were used at a concentration of 20 ng/µL. Stocks of parental lines were already available at the Plant Genetics Laboratory at UNIBO. Artificial heterozygous samples were produced by mixing the same amount of DNA of a resistant and of a susceptible parent at the same concentration (20 ng/µL in this case). A sample of artificial heterozygous composition was obtained by mixing M and C (MC-heterozygous) and a sample of artificial heterozygous composition by

mixing Sv and Cc (SvCc-heterozygous). Each KASP assay was tested on several replicates of the four parental lines and the two heterozygous samples.

PCR for KASP assays were set up in 96-wells optical plates in a total volume of 10 μL per reaction, containing 2 μL of genomic DNA at a concentration of 20 ng/μL, 5 μL of KASP master mix (2X), 0.14 μL of primer mix and 2.86 μL of Milli-Q water. Primer mix was composed by 12 μL of primer A, 12 μL of primer B, 30 μL of primer C (each primer at 100 μM) and 46 μL of Milli-Q water (total volume = 100 μL).

The KASP thermal cycling program used is reported in **Table 2**. PCR reaction protocol and thermal cycling program were used as proposed by the "KASP genotyping chemistry User guide and manual" available at the website https://www.lgcgroup.com/products/kasp-genotyping-chemistry (PCR protocol was adjusted for the quantity of DNA/water per reaction). Primer mix was prepared as reported in Yuan et al. (2014).

Table 2. Thermal cycling program used for KASP assays as proposed by the "KASP genotyping chemistry User guide and manual" available at the website https://www.lgcgroup.com/products/kasp-genotyping-chemistry.

94°C for 15 minutes	Hot-start activation
94°C for 20 seconds 61-55°C for 60 seconds (dropping 0.6°C per cycle)	10 cycles
94°C for 20 seconds 55°C for 60 seconds	26 cycles

Once the PCR thermal cycling program was completed, optical plates were read on an ABI 7500 instrument. This stage of the process consists in performing a thermal recycling program composed of three cycles as follows: denaturation at 94 °C for 20 sec followed by annealing/elongation at 57 °C for 60 sec. The recycling program can be repeated several times until three defined genotyping

clusters are outlined in the output graphical plot. Usually one to three recycling runs are needed, depending on the assay. However, not more than five runs were performed to avoid the detection of aspecific PCR products. At the end of each recycling run, the resulting cluster plot was visually inspected and the results checked for correct correspondence of the samples used as control. Results were then saved and exported in the form of Excel files containing the resulting genotype (homozygous for the resistant allele/homozygous for the susceptible allele/heterozygous/undetermined) for each investigated sample.

4.3. Marker-assisted selection on the Sv x Cc RIL population

A marker-assisted selection (MAS) was performed on the Sv x Cc RIL population with two KASP markers flanking the region were *QSbm.ubo-2BS* was located: *KUBO 13* and *KUBO 9*. The objective of the MAS was to select RILs with recombination between the two flanking markers. A first round of MAS was performed in 2015 on a subset of 1,935 RILs Sv x Cc and a second round of MAS in 2016 on a subset of 1,125 Sv x Cc RILs, for a total of 3,060 RILs analyzed.

4.3.1. DNA isolation

DNA was isolated from lyophilized leaf tissue of Sv x Cc RILs using a CTAB DNA isolation protocol for 96-well plates. Leaf tissue was distributed in plates (one RIL per tube) and ground with inert material by means of a tissue lyser mixer mill. Two tubes per plate were left empty to host the DNA of the controls (the two parental lines Sv and Cc), already available. 400 µL of warm CTAB extraction buffer (prepared as described in Doyle and Doyle, 1990) were added to each tube. Plates were inverted several times to mix the content and then incubated for 90 minutes at 65 °C with continuous rocking at 160 rpm speed. After adding 400 µL of chloroform to each tube, plates were inverted and centrifuged for 25 min at 2500 rpm speed. The top aqueous layer was poured off from each tube and transferred to a new plate. An equal volume of cold isopropanol was added to the

transferred aliquot and plates were inverted several times to mix the content. Plates were centrifuged for 25 min at 2800 rpm and the aqueous solution removed by inversion of the tubes. After adding 500 μL of ethanol 70%, plates were left at room temperature for 1 h. Plates were centrifuged for 10 min at 2800 rpm and the ethanol was removed by inversion. The tubes were left to dry overnight at room temperature. Finally, 200 μL of Milli-Q water were added to each tube to resuspend the DNA.

4.3.2. PCR reactions

The PCR reactions were performed on the isolated DNA of the RILs following the protocol already described in section **4.2.2**. One sample of each of the two parental lines (Sv and Cc) per plate was used as control.

4.4. Phenotyping of selected Sv x Cc RILs

4.4.1. Field trials

The Sv x Cc RILs selected through MAS with two markers (*KUBO 13* and *KUBO 9*) flanking the *QSbm.ubo-2BS* interval were sown in a field experiment aimed at measuring the severity of the SBCMV infection.

In total, 320 of 337 RILs selected in 2015 were sown in early November of the same year in two experimental fields (field 1-2016 and field 2-2016, see **Table 3**) in Cadriano (Bologna, Italy, 44°35' N 11°27' E). Experimental fields were characterized by a severe and uniform SBCMV infection, maintained by continuously growing the highly susceptible durum cultivar Grazia. In both fields, the experimental design was a randomized complete block design with two replicates. Each plot was composed by four 1.80 m-long rows. Sowing density was 300 seeds/m^2. Sv, Cc and Grazia were included in each block as control.

181 of 184 RILs selected in 2016 were sown in early November of the same year in Cadriano (Bologna), together with eight RILs selected and sown in 2015, which resulted informative from the genotypic and phenotypic analysis, for a total of 189 RILs. 177 of these RILs were sown in a field with the same characteristics described above (field 1-2017), but using a design with three replicates and only Sv and Cc as control. The 12 remaining RILs, for which less seed was available, were sown in a separate field (field 2-2017) in plots with two 1.50 m-long rows, at a density of 300 seeds /m^2. Also in this case a randomized complete block design with three replicates with Sv and Cc as control was adopted.

4.4.2. SBCMV symptom severity evaluation

The phenotypic evaluation of SBCMV response was performed by visually scoring each RIL for symptom severity (SS) on several dates from the appearance of the first symptoms on plant leaves until the infection showed his maximum intensity, corresponding to a period from the middle of March to the beginning of April. This period corresponded to plant growing stages from mid-end of tillering (Z25 in the Zadoks scale, Zadoks et al., 1974), to first (Z31) and second (Z32) node appearance. Scoring dates for each field are listed in **Table 3**. SS was recorded by means of a 0-5 scale modified from Vallega and Rubies Autonell (1985) where: 0-1.5 = no or slight symptoms, 1.51-2.5 = mild mottling and stunting, 2.51-3.5 = mottling and stunting, 3.51-4.5 = severe mottling and stunting, 4.51-5.0 = plants killed by virus.

Table 3. Scoring dates for SBCMV symptom severity evaluation in years 2016 and 2017. For each field, dates in which the evaluation was performed are marked with an "X".

	2016				
	March 14th	March 17th	March 30th		
Field 1	x		x		
Field 2		x	x		
	2017				
	March 17th	March 21st	March 23rd	March 28th	April 3rd
Field 1	x	x		x	x
Field 2			x	x	x

4.5. Genotyping of selected Sv x Cc RILs

RILs selected through MAS in 2015 and 2016 and sown in the experimental fields were genotyped with the following KASP markers: *KUBO 1*, *KUBO 3*, *KUBO 27*, *KUBO 29*, *KUBO 40* and *KUBO 41*. RILs were also genotyped with *KUBO 13* and *KUBO 9* to double-check the previous results and with the DArT marker *wPt-2106.*

4.5.1. DNA isolation

Seeds of selected RILs were planted in the greenhouse. Leaves were collected in Eppendorf tubes (one RIL per tube), lyophilized and grinded with inert material by means of a tissue lyser mixer mill. DNA was isolated by means of the kit "NucleoSpin® Plant II" by Macherey-Nagel. 400 µL of buffer PL1 (lysis buffer) were added to each tube. After vortexing the mixture, 10 µL of RNase A solution were added to remove RNA. Samples were mixed manually by shaking the tubes and then incubated for 10 min at 65 °C and continuous rocking at 550 rpm speed. Then they were centrifuged for 1 minute at 11,000 x g and the lysate was loaded onto a NucleoSpin® Filter column for clarification. Columns were centrifuged for 2 min at 11,000 x g and flow-through was collected.

Since a pellet was visible in the flow-through, the clear supernatant was transferred to a new 1.5 mL microcentrifuge tube. 450 µL of buffer PC were added to each tube to adjust DNA binding conditions. Samples were vortexed and then loaded onto a NucleoSpin®Plant II Column to allow the DNA to bind to the silica membrane. Columns were centrifuged for 1 min at 11,000 x g and flow-through was discarded. A first wash of the silica membrane was performed by adding 400 µL of buffer PW1 to each column. Columns were centrifuged for 1 min at 11,000 x g and flow-through was discarded. A second wash was performed by adding 700 µL of buffer PW2. Columns were centrifuged for 1 min at 11,000 x g and flow-through was discarded. Another 200 µL of buffer PW2 were added to the columns. A 2 minutes centrifugation at 11,000 x g was performed in order to remove the wash buffer and dry the silica membrane completely. Each column was then moved to a new 1.5 mL microcentrifuge tube and 50 uL of Buffer PE (preheated at 65 °C) were pipetted onto the membrane. Columns were incubated for 5 min at 65 °C and then centrifuged for 1 min at 11,000 x g to elute the DNA. This step was repeated with 50 µL of buffer PE and DNA was eluted into the same tube.

DNA was quantified by means of 1% agarose gel electrophoresis and diluted to a concentration of 20 ng/µL (working stock).

4.5.2. PCR reactions

The PCR protocol used for KASP markers was already described in section **4.2.2**. Two samples of each parental line (Sv and Cc) and two samples of artificial SvCc-heterozygous were included in each plate as control.

PCR reactions for *wPt-2106* were set up in 96-well plates, in a total volume of 20 µL per reaction composed as following: 5 µL of genomic DNA at 20 ng/µL, 4 µL of buffer (5X), 1.6 µL of $MgCl_2$, 0.16 µL of dNTPs, 0.05 µL of Taq Polymerase, 2.4 µL of primer mix and 6.79 µL of Milli-Q water. Primer mix was composed by 100 µL of primer forward Meridiano, 100 µL of primer forward Claudio and 100 µL of primer reverse (each primer at 100 µM) and 700 µL of Milli-Q water (total

volume = 1 mL). The thermal cycling program used was the following: 5 min at 94 °C, 45 sec at 94° C, 45 sec at 60 °C, 1 min at 72 °C, 10 min at 72 °C and ∞ at 12 °C for 35 cycles. PCR products were electrophoresed in 2.5% agarose gel and visualized after staining with ethidium bromide. Parental lines and artificial heterozygous samples were loaded on each gel as control.

4.6. Genotyping of M x C RILs with KASP markers

The 30 (out of 181) M x C RILs presenting recombination between the SNPs used to design the assays for *KUBO 13* and *KUBO 9* or immediately outside this interval (based on the M x C linkage map of Maccaferri et al., 2015) were selected and genotyped with KASP markers *KUBO 1, KUBO 3, KUBO 9, KUBO 13, KUBO 27, KUBO 29, KUBO 40* and *KUBO 41*. DNA of RILs was already available at the Plant Genetics Laboratory at UNIBO. PCR reactions were performed following the protocol previously described for KASP markers in section **4.2.2**. Two samples of each parental line (M and C) and two samples of artificial MC-heterozygous were included in each plate as control.

4.7. Data analysis

4.7.1. Phenotypic data analysis

The SS data collected for Sv x Cc RILs were corrected according to various methods: a moving average model based on genotypes means, a moving average model based on population means, a mixed model with columns and rows of the experimental layout considered as fixed factors and genotypes as random, and various combinations of these methods.

The moving average model was applied according to a two-step strategy, considering each plot of the experimental field layout as a position of coordinates X and Y, where X are rows and Y are columns: 1) the difference of the SS score from the genotypic or population mean was calculated for each XY position, 2) the mean of all values obtained for each XY position in the area including the XY position and all the plots surrounding it was subtracted from the original SS score of each

XY position. This strategy allowed estimating the effects of the micro-environmental variation on the observed SS scores.

ANOVA was performed for data calculated according to each model and the heritability was calculated as following: $h^2 = \sigma^2_G / (\sigma^2_G + \sigma^2_E / r)$, where r is the number of replicates, σ^2_G = (MS genotypes – MS $_{error}$) / r and σ^2_E = MS $_{error}$ (MS = mean square values).

The moving average model based on genotypes means showed the highest heritability and better maximized the difference between Sv and Cc compared to all other models. In the next step, BLUEs (Best Linear Unbiased Estimators) of data corrected according to the moving average of genotypic effects were calculated considering the date of phenotyping as random effect factors and the genotypes as fixed-effect factors. BLUEs of the eight RILs that were sown in both 2016 and 2017 field trials were estimated considering both years data. Data correction and calculation of BLUEs were performed using the statistical software R and the package "lme4" (Bates et al., 2014; R Core Team, 2017).

4.7.2. Genotypic data analysis

Genotypic data of Sv x Cc and M x C RIL populations were used to construct a genetic map for each population using the software JoinMap v. 4 (Van Ooijen, 2006).

In the case of Sv x Cc RILs, all RILs non recombinant between the flanking markers *KUBO 13* and *KUBO 9* were assumed to have either a haplotype completely resistant (Sv allele at all markers) or completely susceptible (Cc allele at all markers) in ratio 1:1. The presence of a 2% of double recombinants was assumed between the two flanking markers. These assumptions allowed to calculate map distances based on the 3,060 RILs of Sv x Cc population. To perform the analysis, heterozygous genotypes were considered as missing data.

In the case of M x C RILs, genotypic data for KASP markers were integrated with those already available for SNP, SSR and DArT markers in the *QSbm.ubo-2BS* region of the M x C linkage map of Maccaferri et al. (2015) and markers order was recalculated based on the integrated data.

Markers order and genetic distances between markers were calculated using the Maximum Likelihood algorithm, which bases the calculation of genetics distances on the Haldane's function. The two maps were compared to check the order of KASP markers in the region between *KUBO 13* and *KUBO 9*.

Once markers order was determined, genotypic data were integrated with phenotypic data to evaluate the presence of informative recombination events between markers in the *QSbm.ubo-2BS* interval and to define the interval where the QTL is most probably confined.

4.8. Single-marker analysis

Phenotypic and genotypic data of the Sv x Cc RIL population were used to conduct a single-marker analysis by means of the software Windows QTL Cartographer (Wang et al., 2012) to investigate the association between each marker and the SS phenotypic trait. In the following step, interval mapping (IM) analysis was carried out with the same software in order to calculate the LOD score at a walking step of 0.5 cM in the interval between *KUBO 13* and *KUBO 9*.

4.9. Physical region and candidate genes

The interval between markers *KUBO 27* and *KUBO 40 – KUBO 41* was compared with the same region on the physical map of the Svevo genome, recently sequenced by the International Durum Wheat Sequencing Consortium (Maccaferri et al., unpublished), in order to determine the length of the interval in base pairs. The number of genes included in the interval was determined and, for each of them, the longer transcript was considered as representative of all transcripts. All the information about gene functions was recovered using the functional annotation program AHRD and the number of genes corresponding to each function description was counted. Lastly, GO terms (GO Slim) in R Bioconductor (version 3.4) package GOstats (Falcon and Gentleman, 2007; R Core

Team, 2017) was used to categorize the genes in the region according to their molecular functions and biological processes.

5. Results

5.1. KASP markers developed in *QSbm.ubo-2BS* region

The high-density consensus map developed by Maccaferri et al. (2015) was instrumental for this study, since it allowed to develop KASP markers from SNPs found polymorphic in several maps. Accordingly, SNPs were chosen based both on their position in the QTL interval and on the maps in which they were found polymorphic. Markers polymorphic in the three RIL populations M x C, Simeto x Levante and Sv x Cc were preferred.

In total, 35 KASP assays were developed from 16 SNPs mapped in the consensus map. Designed primers for each KASP assay are listed in **Table 4**. The diagnostic allelic primers designed by PolyMarker do not contain probe sequences for fluorescent dyes (FAM and HEX) used in KASP. These were added to the 5' end of the primers. In all cases, the probe sequence for FAM (gaaggtgaccaagttcatgct) and for HEX (gaaggtcggagtcaacggatt) were added to A and B primers, respectively.

Table 4. Primers designed for KASP assays. SNP ID = ID code of the SNP on which the KASP assay was developed. KASP marker = name of the developed KASP assay. Each assay is composed by three primers: two allele-specific primers A and B and a common primer C. Allele-specific primers A and B present the probe sequence respectively for FAM (5' GAAGGTGACCAAGTTCATGCT 3') and HEX (5' GAAGGTCGGAGTCAACGGATT 3') at 5' end. For each developed assay, information about its codominance is reported.

SNP ID	KASP marker	Primer allele-specific A	Primer allele-specific B	Primer common C	Codominant
IWB28973	KUBO1	5'FAM-cgtcttgttggtcagctgT	5'HEX-cgtcttgttggtcagctgC	gcagctgtgacacgaacatta	yes
IWB28973	KUBO38	5'FAM-ccgtcttgttggtcagctgt	5'HEX-ccgtcttgttggtcagctgC	gcagagcagctgtgacacg	yes
IWB8328	KUBO3	5'FAM-tgaagtggaggaactagaagcT	5'HEX-tgaagtggaggaactagaagcC	gagaaggtgctgacatcattttc	yes
IWB24939	KUBO5	5'FAM-caaaatctggggcatgctcA	5'HEX-caaaatctggggcatgctcG	agagctagggtggtttcagg	yes
IWB6204	KUBO6	5'FAM-gacaaagttggtggtaaattctcttaC	5'HEX-ggacaaagttggtggtaaattctcttaT	gcaagcatgccttccaccaccaa	yes
IWB10512	KUBO8	5'-FAMatcagaatatgtacaattattatgccaagaT	5'HEX-cagaatatgtacaattattatgccaagaC	ggcggaaattctcttccttttccttattta	yes
IWB10512	KUBO9	5'FAM-aggatgcgttgtttgccaactctA	5'HEX-ggatgcgttgtttgccaactctG	gccatccaccaccaccataagtta	yes
IWB73347	KUBO12X	5'-FAMggctgacagggacattgcT	5'HEX-ggctgacagggacattgcC	ctgtaaacggactctcgaca	yes
IWB73347	KUBO13	as KUBO12X_A	as KUBO12X_B	gtgcatttcaaaattcattttctc	yes
IWB73347	KUBO14	5'FAM-gcatctctgacattctcagaacA	5'HEX-gcatctctgacattctcagaacG	cggggacggtttcagg	yes
IWB61884	KUBO15	5'FAM-aaggagtgaaaccaagcgtaT	5'HEX-aaggagtgaaaccaagcgtaC	caccaactctgacactggc	yes
IWB61884	KUBO16	as KUBO15_A	as KUBO15_B	tgacatcacccaccttatcgt	yes
IWB61884	KUBO17	5'-FAM-ccccaagcgctgacG	5'HEX-ccccaagcgctgacA	ggagcatggtatgacagacatc	yes
IWB11421	KUBO26	5'FAM-taaacaatcactcgctaataagcaaC	5'HEX-taaacaatcactcgctaataagcaaT	aaagtaggtgcacagcctgag	yes
IWB11421	KUBO27	as KUBO26_A	as KUBO26_B	gccatatgctgtgactgctg	yes
IWB23029	KUBO29	5'FAM-gctggaggaactagaagctgaG	5'HEX-gctggaggaactagaagctgaT	aacagagcattgcaaaaccta	yes
IWB23029	KUBO31	as KUBO29_A	as KUBO29_B	ggagtcagttggtgcattcg	yes
IWB29097	KUBO40	5'FAM-cctTtTaTcccaggaA	5'HEX-GaTGCTcctTtTaTcccaggaG	tcgaatcatgaggacctggg	yes
IWB35524	KUBO41	5'FAM-tgtGcattgtgaaggtgatttaaA	5'HEX-tgtGcattgtgaaggtgatttaaG	acttgcagacgaaccatgtag	yes
IWB23330	KUBO4	5'FAM-tgcacccctttgttctttgaaT	5'HEX-tgcacccctttgttctttgaaC	aaacaagacaagataatgcaagtct	no
IWB10512	KUBO7	5'FAM-aggatgcgttgtttgccaactctA	5'HEX-ggatgcgttgtttgccaactctG	gccatccaccaccaccataagtta	no
IWB8390	KUBO10	5'FAM-cctctgcaacgccgccgC	5'HEX-cctctgcaacgccgccgT	tatgcgggtcggcgatgacgtt	no
IWB8390	KUBO11	5'FAM-agagtcaaggaactcccacG	5'HEX-agagtcaaggaactcccacA	gtgtgaatgagagaatctttgacg	no
IWB8390	KUBO12	5'FAM-ccgatactaatcctaggaattacgG	5'HEX-ccgatactaatcctaggaattacgA	aggaagtcatgagtccttggta	no
IWB11421	KUBO28	as KUBO26_A	as KUBO26_B	tttctgcacctgccatatgc	no
IWB23029	KUBO30	as KUBO29_A	as KUBO29_B	cgtcgagaagaatttgacagaa	no
IWB42660	KUBO32	5'FAM-ttggcctgtaaaaggctatcG	5'HEX-ttggcctgtaaaaggctatcA	gatgagcataatctacttcctgatg	no
IWB42660	KUBO33	as KUBO32_A	as KUBO32_B	tcctgatgccctagagagtgac	no
IWB42660	KUBO34	as KUBO32_A	as KUBO32_B	ctatgttctcaatagtcatgaggc	no
IWB41644	KUBO35	5'FAM-aagtgcaggcatgtgttgagataA	5'HEX-aagtgcaggcatgtgttgagataC	cggcgacattactgtcttaacc	no
IWB41644	KUBO36	as KUBO35_A	as KUBO35_B	ggctggagcttgactgagaat	no
IWB41644	KUBO37	as KUBO35_A	as KUBO35_B	ccagccctgctcctgatc	no
IWB28973	KUBO39	5'FAM-gaggcgtcgtgctgcctA	5'HEX-gaggcgtcgtgctgcctG	tcgcaccgcatgtatggt	no
IWB45152	KUBO2	5'FAM-acaaacgaagaacaacacggT	5'HEX-acaaacgaagaacaacacggC	tcgccctcccctttcctg	no
IWB45152	KUBO42	as KUBO2_A	as KUBO2_B	ccctttcctggtgacagtga	no

Among the 35 assays that were developed, 19 (*KUBO 1, KUBO 3, KUBO 5, KUBO 6, KUBO 8, KUBO 9, KUBO 12X, KUBO 13, KUBO 14, KUBO 15, KUBO 16, KUBO 17, KUBO 26, KUBO 27, KUBO 29, KUBO 31, KUBO 38, KUBO 40* and *KUBO 41*) were found to be codominant. Examples of output cluster plots for KASP assays are reported in **Figure 7.** A KASP assay is considered efficient when it detects and reveals the polymorphism enclosed in the starting SNP marker, that is when the assay clearly discriminates the two parental and the heterozygous alleles. Good KASP markers produce an output cluster plot with three well-distinct clusters: samples homozygous for the resistant allele (M and Sv), samples homozygous for the susceptible allele (C and Cc) and heterozygous samples. An example of such assays is the output plot of *KUBO 40* presented in **Figure 7A**. KASP assays that result in plots where samples look scattered across the scheme and are not grouped into clusters, as is the plot for *KUBO 42* represented in **Figure 7B**, are not considered suitable. In this case, samples homozygous for the resistant allele are not discernable from samples homozygous for the susceptible allele and thus the assay is not useful to conduct genotyping on the population. Another kind of plot outcome that can result from KASP assays is the dominant marker profile. In this case, samples are grouped in two clusters (homozygous for the resistant allele and homozygous for the susceptible allele) and heterozygous genotypes are not discernable from these two groups. Although this kind of assay can be used for genotyping, it could lead to a mistake in the evaluation of recombinant events, for example in case of RILs with residual heterozygosity, hence not fixed at some markers in the interested interval.

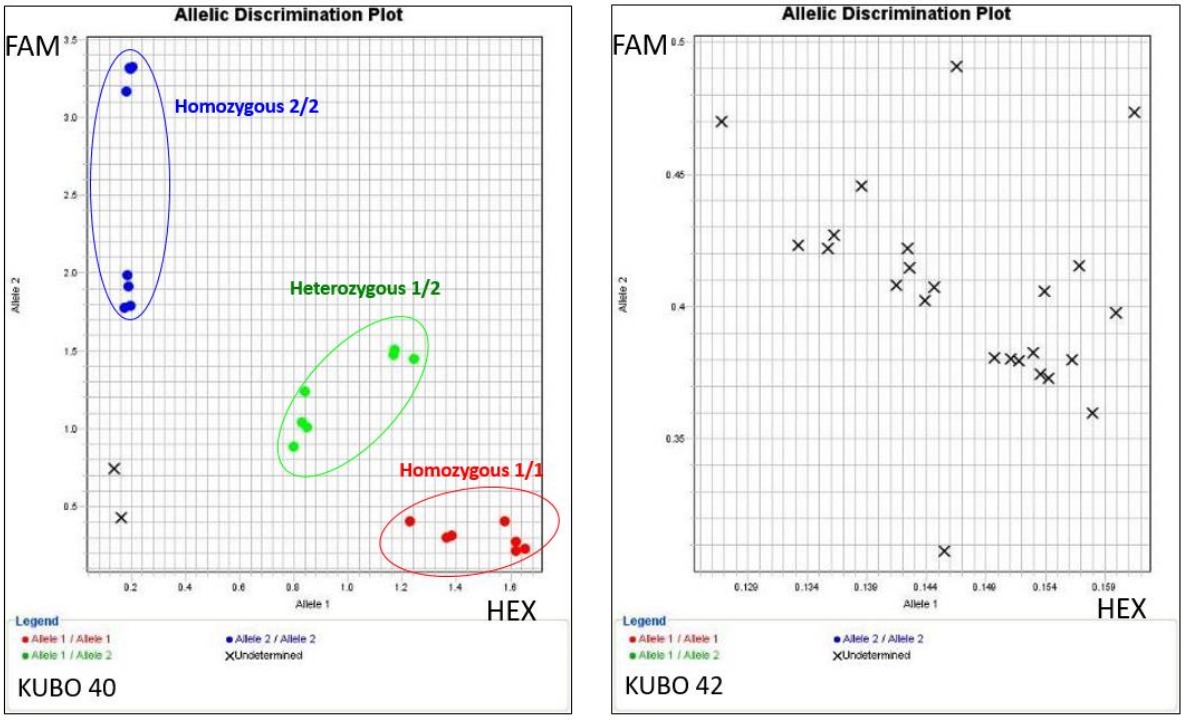
In total, 11 of the starting 16 SNPs were successfully converted in codominant KASP assays, obtaining a conversion efficiency from SNP to KASP of 69%. When more than one codominant KASP assay per SNP was available, the one with the best resolving power was chosen, i.e. the one that better clustered the different genotypes.

KUBO 13 and *KUBO 9* were selected to perform MAS on Sv x Cc RILs since they were suitable when validated on parental genotypes and for their position flanking the QTL interval. The SNPs on which the two KASP assays were developed are located 2.5 cM upstream and 8.1 cM downstream

of *wPt-2106*, respectively (based on the consensus map) and flank the ~2 cM interval where *QSbm.ubo-2BS* was confined in previous studies. Both KASP assays resulted polymorphic between the resistant and the susceptible allele when validated on the four parental lines M, Sv, C and Cc. Both markers are codominant and thus able to detect the eventual residual heterozygosity in the RILs. *KUBO 9* was further validated on the M x C RILs and mapped in the same position as the SNP from which it was developed.

Figure 7. Examples of output cluster plots for KASP assays. A) *KUBO 40* represents a codominant marker with a good resolution output cluster plot. In this case, samples homozigous for allele 1 represent the resistant parents (Sv and M), samples homozigous for allele 2 represent the susceptible parents (Cc and C) and heterozygous samples (allele 1/allele 2) represent the artificial heterozigous. B) *KUBO 42* represents a non-informative KASP assay. Samples are scattered all over the plot and not clusterized into genotype groups.

A) **B)**



After performing MAS, six codominant KASP markers with good resolving power were selected to genotype recombinant Sv x Cc RILs: *KUBO 1*, *KUBO 3*, *KUBO 27*, *KUBO 29*, *KUBO 40* and *KUBO 41*. KASP assays for these markers were developed from SNPs mapping up to 2.5 cM upstream and 1.4 cM downstream of *wPt-2106* and between SNPs used to develop the flanking markers *KUBO 13* and *KUBO 9* (based on the consensus map). Genotyping with KASP markers mapping inside the interval was performed in order to search for recombination events that allow to fine map *QSbm.ubo-2B*.

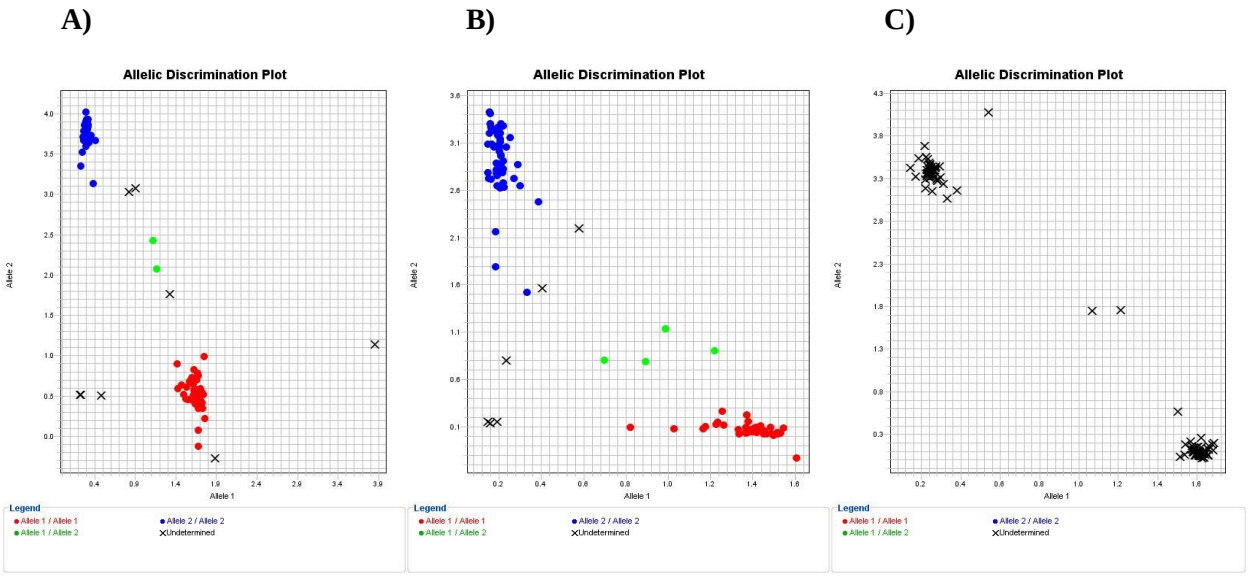
5.2. Marker-assisted selection on the Sv x Cc RIL population

In total 337 RILs were selected through MAS performed in 2015. 160 RILs were Sv at *KUBO 13* and Cc at *KUBO 9*, 157 were Cc at *KUBO 13* and Sv at *KUBO 9*, 2 lines were heterozygous at one of the two markers. The leftover 18 lines were double recombinants (Sv-Cc-Sv or Cc-Sv-Cc) selected through genotyping a part of the RILs with *wPt-2106* in addition to *KUBO 13* and *KUBO 9*.

In total 184 RILs were selected through MAS performed in 2016. 94 RILs were Sv at *KUBO 13* and Cc at *KUBO 9* and 90 RILs presented the opposite genotype.

Examples of KASP cluster plots of Sv x Cc RILs genotyped with *KUBO 13* and *KUBO 9* are reported in **Figure 8**. In some cases, the software was unable to assign a color code to the analyzed samples and these were marked as "undetermined" (**Figure 8C**). This was probably due to overheating of the plate, since plates should be read at a temperature below 40 °C. If samples were divided in three tight clusters, clearly referable to as homozygous for allele 1, homozygous for allele 2 and heterozygous and concordant with controls, then color codes (and consequently genotypes) were assigned manually before exporting the results.

Figure 8. Examples of cluster plots obtained genotyping the population Sv x Cc with *KUBO 13* and *KUBO 9*. A) Cluster plot of *KUBO 9* on 94 Sv x Cc RILs and two parental samples (Sv and Cc). B) and C) Cluster plot of *KUBO 13* on 94 Sv x Cc RILs and two parental samples (Sv and Cc).



5.3. Phenotyping of Sv x Cc RILs selected through MAS

RILs selected through MAS were sown in experimental fields infected with SBCMV. 17 RILs selected in 2015 and 3 RILs selected in 2016 were not sown for lack of available seed.

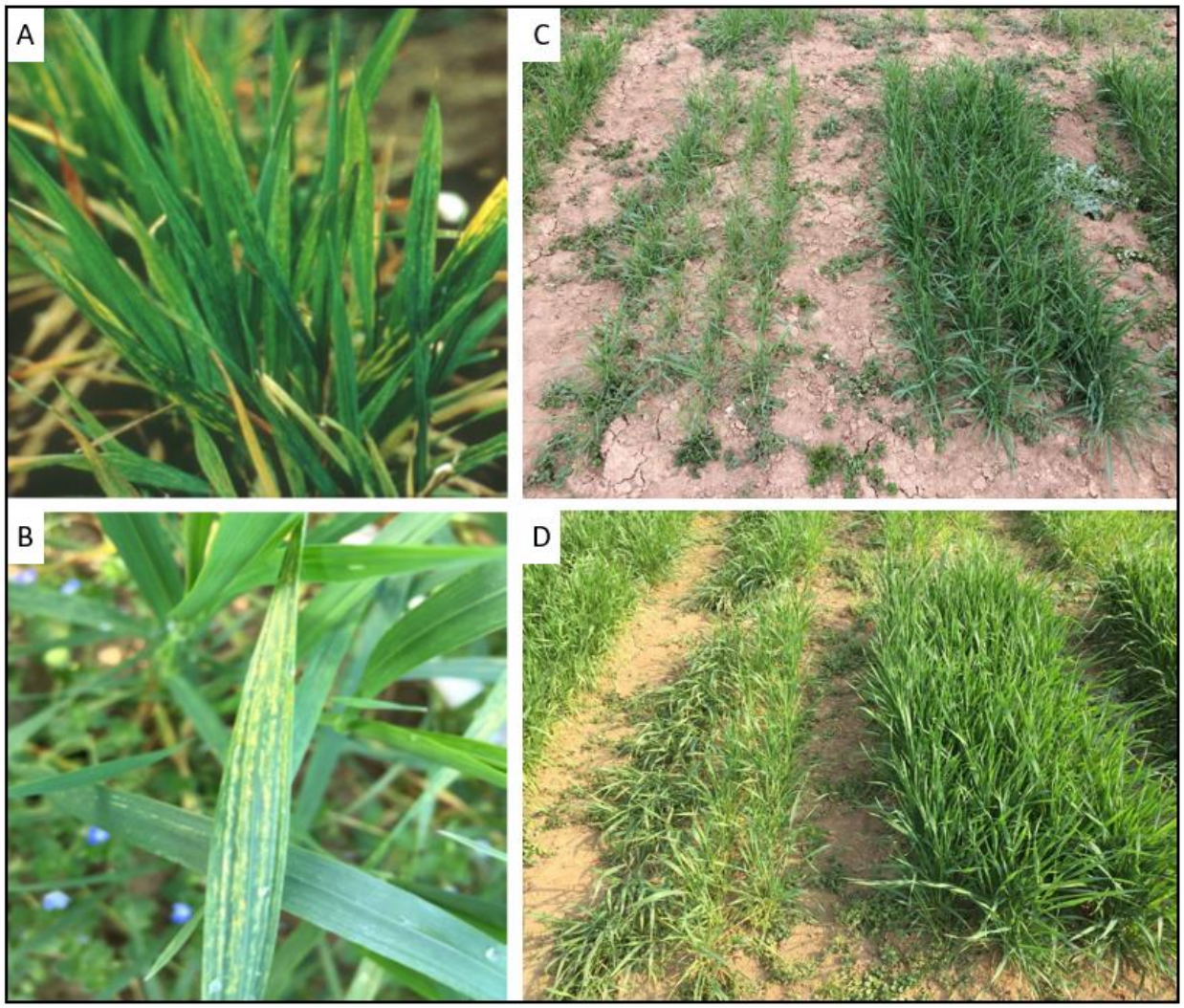
SS score for 320 RILs selected in 2015 and 181 RILs selected in 2016 was collected, with a total of 501 RILs scored. Of the 320 RILs selected in 2015 and scored in 2016, eight were informative from preliminary analysis conducted on that first dataset and were thus re-sown and scored again in 2017.

Examples of symptoms caused by SBCMV observed during the scoring activity are reported in **Figure 9**.

SS data collected in 2016 and 2017 were corrected according to the moving average of genotypic effects model, which showed the highest heritability (93.3%) and the maximum difference between parental lines compared to all other models. Statistically significant differences ($p < 0.0001$)

resulted from the ANOVA performed on the corrected data among genotypes and scoring dates as well as a significant genotype x scoring dates interaction.

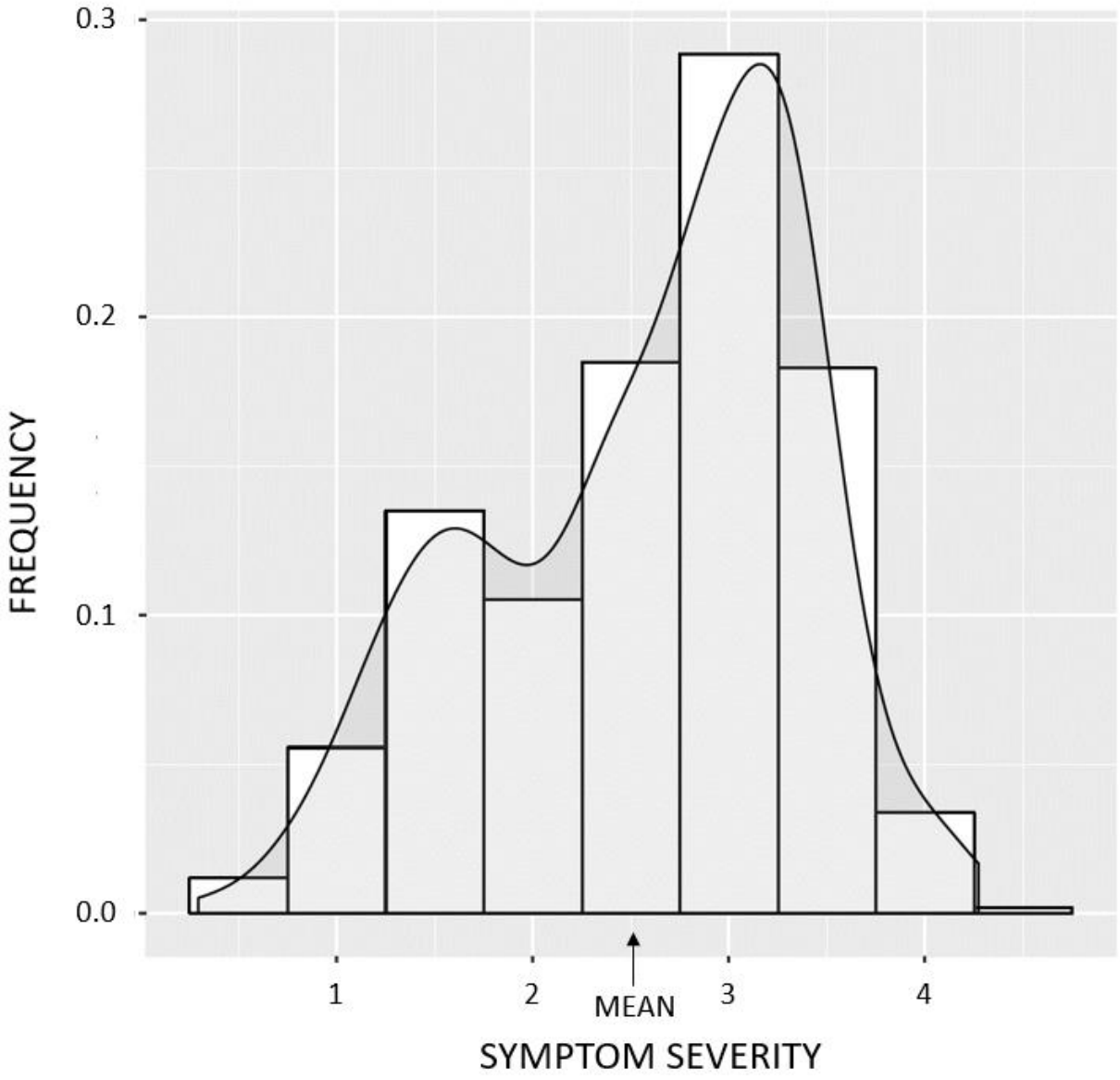
Figure 9. Symptoms caused by SBCMV infection in durum wheat. A) and B) Chlorotic mottling caused by SBCMV infection. C) and D) The left plot (susceptible) shows stunting compared to the right plot (resistant), which shows a regular growth.



Frequency distribution of BLUEs obtained from corrected SS data are reported in **Figure 10**. Data showed a mean value of 2.58 with extreme values of 0.29 and 4.27. The distribution histogram shows two peaks, one centered on a SS score of 1.5 (medium resistant) and one on a SS score of 3.0 (medium susceptible). The frequency distribution reflects the segregation of the resistance trait in

the population, with ~41% of individuals presenting a SS score <2.51 (resistant) and ~59% of individuals presenting a SS score ≥2.51 (susceptible).

Figure 10. Frequency distribution of BLUEs obtained from symptom severity scores data collected in 2016 and 2017 for Sv x Cc RILs and corrected acording to the moving average of genotypic effects model.



5.4. Genotyping of RILs Sv x Cc and M x C and maps construction

The 501 Sv x Cc RILs (320 selected in 2015 and 181 selected in 2016) were genotyped with six KASP markers mapping inside the interval (*KUBO 1*, *KUBO 3*, *KUBO 27*, *KUBO 29*, *KUBO 40* and *KUBO 41*), with the two KASP markers flanking the interval (*KUBO 13* and *KUBO 9*) to double-check previous results and with *wPt-2106*. Genotyping results are reported in **Table 5**. "A" and "B" indicate a RIL homozygous for the resistant (Sv) and the susceptible allele (Cc), respectively, while "H" indicates a heterozygous RIL and "m" indicates a missing data. In total, 227 RILs presented the Sv (resistant) haplotype in the whole interval except for one of the two flanking markers and 199 presented the Cc (susceptible) haplotype in the whole interval except for one of the two flanking markers. Only seven RILs showed recombination inside the target interval and where thus informative for fine mapping. The remaining 68 RILs showed either the Sv/Cc allele at all markers, a double recombination between the flanking markers *KUBO 13* and *KUBO 9* or presented residual heterozygosity or were missing (data not reported).

Based on the information already available from the M x C linkage map, 30 M x C RILs were selected for the presence of recombination events between the SNPs used to design *KUBO 13* and *KUBO 9* or immediately outside this interval. Genotyping profiles of KASP markers were integrated with profiles of SNPs already available from the M x C linkage map and markers order was recalculated with JoinMap v.4. Results for the 18 informative RILs are reported in **Table 6** where "A" and "B" indicate homozygous RILs for the resistant (M) and the susceptible (C) allele respectively.

Genotyping results of Sv x Cc and M x C RILs were used to construct a genetic map for each population using the software JoinMap v.4 (**Figure 11**). As expected, in the M x C map each KASP marker maps at the same genetic distance of the SNP from which the respective KASP assay was developed.

Maps were compared to check the order of KASP markers, which coincided in both maps. However, in the M x C map *KUBO 27* is located in the recombination bin including also *KUBO 29*

and *wPt-2106*, while in Sv x Cc it is located in a different bin. Moreover, *KUBO 1* and *KUBO 3* in Sv x Cc map are in the same bin as *KUBO 29* and *wPt-2106* and in a different one in respect to *KUBO 40* and *KUBO 41*, while in M x C *KUBO 1* and *KUBO 3* are separated from *KUBO 29* and *wPt-2106*, but not from *KUBO 40* and *KUBO 41*.

Another important difference between the two maps pertains to the genetic distances. As an example, the interval between *KUBO 27* and *KUBO 40* shows seven recombination events on a total of 3,060 lines in the Sv x Cc population and three recombination events on a total of 181 lines in the M x C population. This results in a genetic distance between the two markers of 0.2 cM in Sv x Cc and 0.9 cM in M x C.

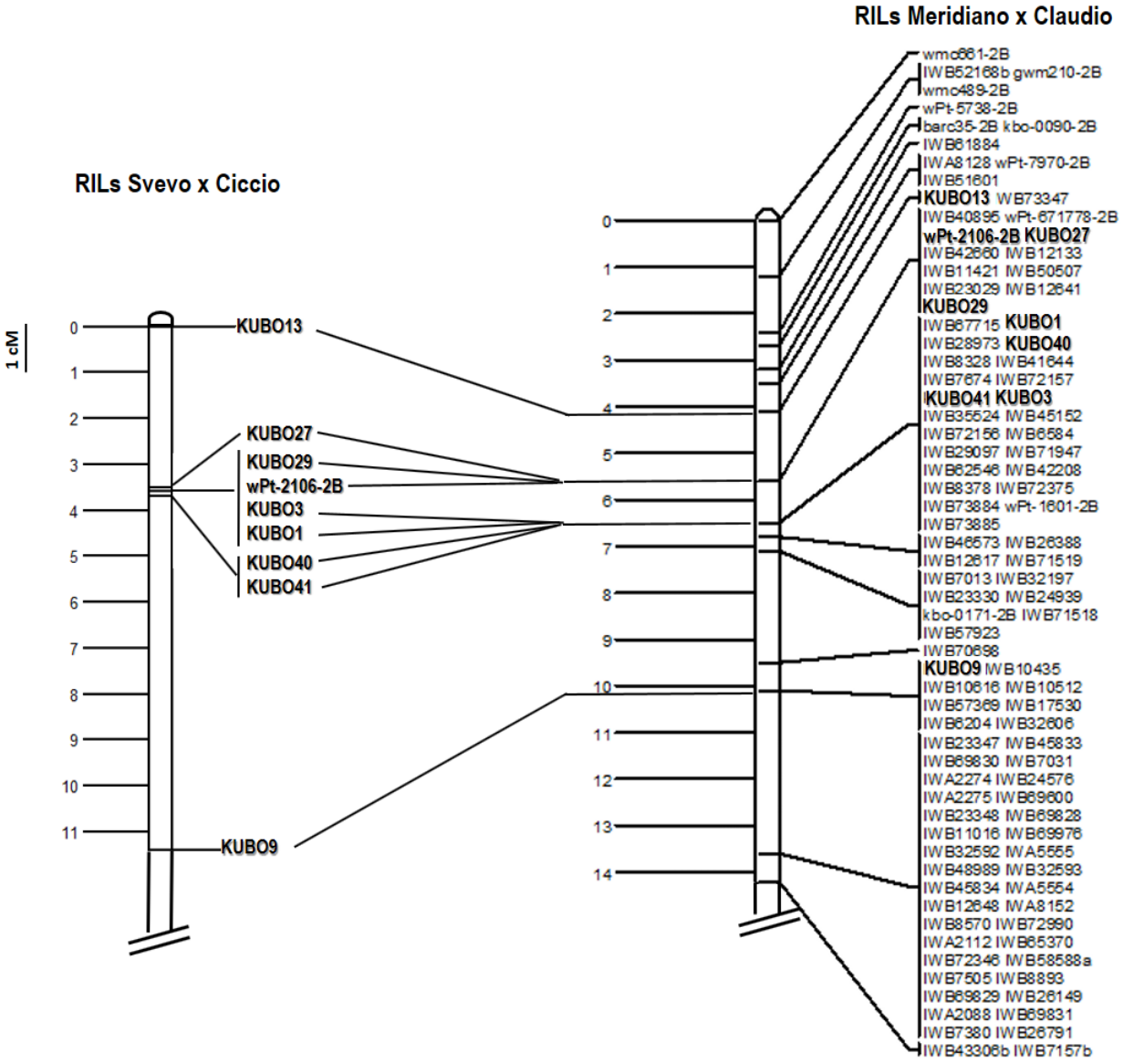
Table 5. Genotyping results of Sv x Cc RILs selected by MAS. A = Sv allele = resistant, B = Cc allele = susceptible, H =heterozygous, m = missing data. Genetic distances (cM) were calculated with JoinMap v. 4. N° of RILs = number of individuals presenting that specific haplotype. Symptom Severity BLUEs were calculated as the mean of BLUEs of all individuals presenting that specific haplotype. Red square indicates the interval where most probably *QSbm.ubo-2BS* is located.

Markers	**cM**	140	3	59	1	1	2	162	65	**n° of RILs**
KUBO13	0	B	B	A	A	A	A	A	B	
KUBO27	3.5	B	B	B	A	A	A	A	A	
wPt-2106-2B	3.6	B	B	B	B	A	A	A	A	
KUBO29	3.6	B	B	B	B	A	A	A	A	
KUBO3	3.6	B	m	B	B	H	A	A	A	
KUBO1	3.6	B	m	B	B	m	A	A	A	
KUBO40	3.7	B	A	B	B	B	B	A	A	
KUBO41	3.7	B	A	B	B	B	B	A	A	
KUBO9	11.4	A	A	B	B	B	B	B	A	
		3.20	3.25	3.24	3.34	2.56	1.47	2.01	1.90	**BLUEs**

Table 6. Genotypic profiles of M x C RILs recombinant between *KUBO 13* and *KUBO 9*. A = M allele = resistant, B = C allele = susceptible. Genetic distances (cM) were calculated with JoinMap v. 4. N° of RILs = number of individuals presenting that specific haplotype. Symptom Severity BLUEs were calculated as the mean of BLUEs of all individuals presenting that specific haplotype, based on phenotypic data of Maccaferri et al. (2011) adjusted according to the method used to record Sv x Cc symptom severity scores for ease of comparison. Red square indicates the interval where most probably *QSbm.ubo-2BS* is located.

Markers	cM	1	1	2	1	3	1	5	1	1	2	n° of RILs
IWB73347	4.1	A	B	B	B	B	B	A	A	A	A	
KUBO13	**4.1**	A	B	B	B	B	B	A	A	A	A	
KUBO27	**5.6**	B	B	B	B	A	A	A	A	A	A	
IWB11421	5.6	B	B	B	B	A	A	A	A	A	A	
IWB12641	5.6	B	B	B	B	A	A	A	A	A	A	
IWB40895	5.6	B	B	B	B	A	A	A	A	A	A	
IWB50507	5.6	B	B	B	B	A	A	A	A	A	A	
wPt-671778-2B	5.6	B	B	B	B	A	A	A	A	A	A	
KUBO29	**5.6**	B	B	B	B	A	A	A	A	A	A	
wPt-2106-2B	**5.6**	B	B	B	B	A	A	A	A	A	A	
IWB42660	5.6	B	B	B	B	A	A	A	A	A	A	
IWB12133	5.6	B	B	B	B	A	A	A	A	A	A	
IWB23029	5.6	B	B	B	B	A	A	A	A	A	A	
KUBO3	**6.5**	B	B	B	A	A	A	A	A	A	B	
IWB67715	6.5	B	B	B	A	A	A	A	A	A	B	
IWB72157	6.5	B	B	B	A	A	A	A	A	A	B	
IWB28973	6.5	B	B	B	A	A	A	A	A	A	B	
IWB8378	6.5	B	B	B	A	A	A	A	A	A	B	
IWB45152	6.5	B	B	B	A	A	A	A	A	A	B	
IWB73884	6.5	B	B	B	A	A	A	A	A	A	B	
IWB41644	6.5	B	B	B	A	A	A	A	A	A	B	
IWB6584	6.5	B	B	B	A	A	A	A	A	A	B	
IWB73885	6.5	B	B	B	A	A	A	A	A	A	B	
IWB7674	6.5	B	B	B	A	A	A	A	A	A	B	
IWB62546	6.5	B	B	B	A	A	A	A	A	A	B	
IWB29097	6.5	B	B	B	A	A	A	A	A	A	B	
IWB42208	6.5	B	B	B	A	A	A	A	A	A	B	
IWB72156	6.5	B	B	B	A	A	A	A	A	A	B	
IWB35524	6.5	B	B	B	A	A	A	A	A	A	B	
IWB72375	6.5	B	B	B	A	A	A	A	A	A	B	
IWB71947	6.5	B	B	B	A	A	A	A	A	A	B	
wPt-1601-2B	6.5	B	B	B	A	A	A	A	A	A	B	
IWB8328	6.5	B	B	B	A	A	A	A	A	A	B	
KUBO1	**6.5**	B	B	B	A	A	A	A	A	A	B	
KUBO40	**6.5**	B	B	B	A	A	A	A	A	A	B	
KUBO41	**6.5**	B	B	B	A	A	A	A	A	A	B	
IWB46573	6.8	B	B	B	A	A	A	A	A	B	B	
IWB71519	6.8	B	B	B	A	A	A	A	A	B	B	
IWB12617	6.8	B	B	B	A	A	A	A	A	B	B	
IWB26388	6.8	B	B	B	A	A	A	A	A	B	B	
IWB23330	7.1	B	B	B	A	A	A	A	B	B	B	
IWB57923	7.1	B	B	B	A	A	A	A	B	B	B	
IWB7013	7.1	B	B	B	A	A	A	A	B	B	B	
IWB24939	7.1	B	B	B	A	A	A	A	B	B	B	
IWB32197	7.1	B	B	B	A	A	A	A	B	B	B	
kbo-0171-2B	7.1	B	B	B	A	A	A	A	B	B	B	
IWB71518	7.1	B	B	B	A	A	A	A	B	B	B	
IWB70698	9.5	B	B	A	A	A	A	B	B	B	B	
IWB32606	10.1	B	A	A	A	A	B	B	B	B	B	
IWB57369	10.1	B	A	A	A	A	B	B	B	B	B	
IWB10616	10.1	B	A	A	A	A	B	B	B	B	B	
IWB10435	10.1	B	A	A	A	A	B	B	B	B	B	
IWB10512	10.1	B	A	A	A	A	B	B	B	B	B	
IWB6204	10.1	B	A	A	A	A	B	B	B	B	B	
IWB17530	10.1	B	A	A	A	A	B	B	B	B	B	
KUBO9	**10.1**	B	A	A	A	A	B	B	B	B	B	
		3.13	3.57	2.98	3.34	0.28	0.40	0.26	0.25	0.17	0.10	**BLUEs**

Figure 11. Genetic maps obtained for the two RIL populations Sv x Cc and M x C with the software JoinMap. In case of RILs M x C, genotypic data on KASP markers were integrated with those of SNP, SSR and DarT markers already available in the region.



5.5. Analysis of recombination events and fine mapping

In **Table 5** and **Table 6** genotyping results obtained for Sv x Cc and M x C populations are integrated with phenotyping results. Phenotypic SS scores are reported as the mean of BLUEs of all lines with a specific genotypic profile and are represented in a color scale from red (susceptible) to green (resistant).

Sv x Cc genotypic data (**Table 5**) indicate that only seven RILs present recombination at KASP markers developed in the interval between *KUBO 13* and *KUBO 9* and are thus informative to fine map *QSbm.ubo-2BS*. Specifically, one line presents recombination between *KUBO 27* and the block of markers *wPt-2106 - KUBO 29 - KUBO 3 - KUBO 1*, two lines present recombination between the block *wPt-2106 - KUBO 29 - KUBO 3 - KUBO 1* and the block *KUBO 40 - KUBO 41* and four lines present recombination between *wPt-2106 - KUBO 29* and *KUBO 40 - KUBO 41*. Regarding the latter four lines, it was not possible to determine the genotype at marker *KUBO 3* for three of them and at marker *KUBO 1* for all of them. One line resulted heterozygous at marker *KUBO 3*.

Based on the comparison of the genotypic and phenotypic results, some of the markers appear located outside the putative support interval where most probably *QSbm.ubo-2BS* is located. Considering lines recombinant only between *KUBO 13* and *KUBO 9* it is possible to exclude the two flanking markers. *KUBO 13* can be excluded because 59 lines with genotype "A" (resistant) at this marker showed a susceptible phenotype (mean BLUE = 3.24) and 65 lines with genotype "B" (susceptible) showed a resistant phenotype (mean BLUE = 1.9). Additionally, *KUBO 9* can also be excluded because 140 lines with genotype "A" (resistant) showed a susceptible phenotype (mean BLUE = 3.2) and 162 lines with genotype "B" (susceptible) showed a resistant phenotype (mean BLUE = 2.01). Moreover, *KUBO 27* can also be excluded from the support interval based on the line that presents recombination between this marker and the block of contiguous markers *wPt-2106 - KUBO 29 - KUBO 3 - KUBO 1*: in fact, this line presents genotype "A" (resistant) at *KUBO 13* as well as *KUBO 27* and genotype "B" (susceptible) at all other markers and a susceptible phenotype (BLUE = 3.34). Lastly, the block of markers *KUBO 40 - KUBO 41* can be excluded based on three

lines with recombination between *KUBO 40 - KUBO 41* and *wPt-2106 - KUBO 29* and two lines with recombination between *KUBO 40 - KUBO 41* and the block *wPt-2106 - KUBO 29 -KUBO 3 - KUBO 1*. The first three lines present genotype "A" (resistant) at *KUBO 40 - KUBO 41*, genotype "B" (susceptible) at *wPt-2106 - KUBO 29* and a susceptible phenotype (mean BLUE = 3.25) and the other two lines present genotype "B" (susceptible) at *KUBO 40 - KUBO 41*, genotype "A" (resistant) at all other markers and a resistant phenotype (mean BLUE = 1.47). The remaining line, which presents a recombination between *KUBO40 - KUBO41* and *wPt-2106 - KUBO 29*, presents an intermediate phenotype (BLUE = 2.56) and is thus not useful to exclude a portion of markers from the support interval. Based on the above considerations, the fine mapping based on Sv x Cc RILs locates the support interval of *QSbm.ubo-2BS* in 0.2 cM between the markers *KUBO 27* and *KUBO 40 - KUBO 41*.

Genotyping results regarding M x C population integrated with phenotyping results already available at UNIBO (Maccaferri et al., 2011) are reported in **Table 6**. The lines with recombination between *KUBO 9* and *KUBO 13* are 18. The comparison of these results with those obtained for Sv x Cc RILs, indicates that markers *KUBO 13, KUBO 9, KUBO 40* and *KUBO 41* are not in the support interval, but in this case there are not recombination events that exclude *KUBO 27*, which is in the same recombination bin as *KUBO 29* and *wPt-2106*. Nevertheless, markers *KUBO 3* and *KUBO 1* (which map at the same distance as *KUBO 40* and *KUBO 41*) can be excluded because of three informative recombination events. In fact, one line has genotype "B" (susceptible) at *KUBO 27 - KUBO 29 - wPt-2106*, genotype "A" (resistant) at *KUBO 3 - KUBO 1 - KUBO 40 - KUBO 41* and a susceptible phenotype (BLUE = 3.34) and two lines show opposite results (mean BLUE = 0.10). In this case, the support interval is confined between *KUBO 13* and *KUBO 3 - KUBO 1 - KUBO 40 - KUBO 41*, which corresponds to a 2.4 cM interval on the M x C map.

5.6. Single-marker analysis

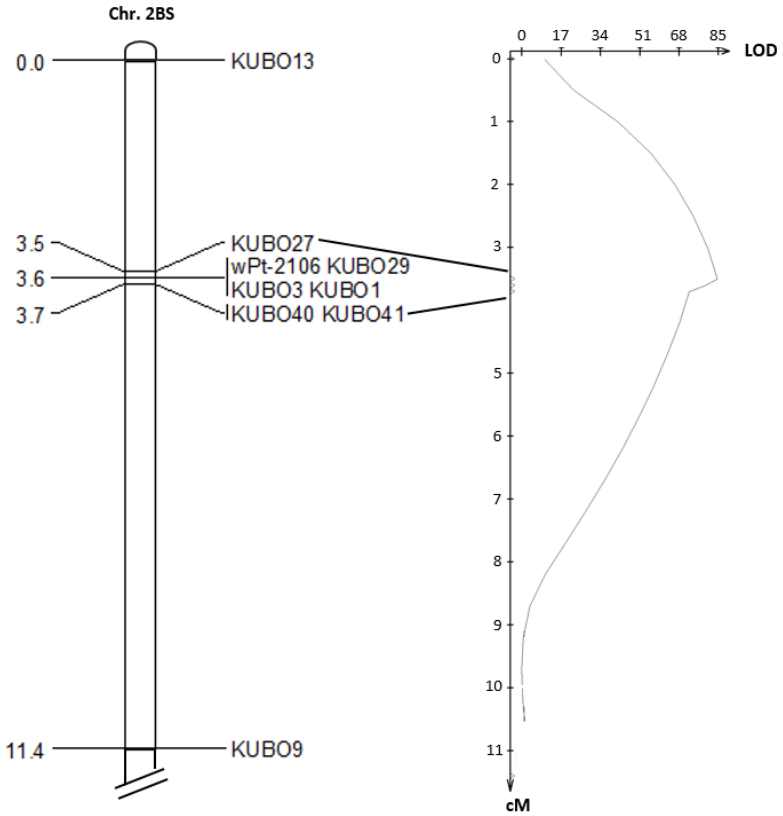
The single-marker analysis for KASP markers was based on the software Windows QTL Cartographer. The analysis showed a strong association for all markers (**Table 7**). The marker most associated with the phenotype was *wPt-2106*, a result consistent with those of Maccaferri et al., (2012, 2011). Markers in the same recombination bin as *wPt-2106* are more closely associated as compared to other markers, except for *KUBO 27,* which is in fact closer to the block of *wPt-2106* than to *KUBO 40* and *KUBO 41*. As expected, *KUBO 13* and *KUBO 9*, are the least associated markers.

Subsequently, the interval mapping (IM) analysis, which is an extension of single-marker analysis, was utilized to further refine the results. This analysis tests the probability of QTL presence versus that of QTL absence for each map interval included between two adjacent markers. It is based on the statistic method of the LOD (Logarithm of ODds) value, which considers the QTL position, the mean of genotypes for each QTL allele and the phenotypic variation. The higher the LOD, the greater the evidence for a QTL. The LOD score is calculated at each increment (walking step) in the interval. The IM analysis was based on a walking step of 0.5 cM, the smallest walking step allowed by the software. Results of IM analysis are reported in **Figure 12**. As expected, flanking markers *KUBO 13* and *KUBO 9* showed a LOD much lower (10 and 6 respectively) in respect to all other markers. The LOD increases as it gets closer to the interval between *KUBO 27* and *KUBO 40-KUBO 41*, with *wPt-2106* showing the highest LOD value (89).

Table 7. Results of the single-marker analysis on KASP markers and *wPt-2106*. F statistic is a measure of the association between marker and phenotypic trait (the higher the F, the higher the association). ****: F values significant at the 0.01% probability level.

MARKER	F	pr(F)
KUBO 13	50.470	0.000****
KUBO 27	610.844	0.000****
wPt-2106-2B	628.362	0.000****
KUBO 29	610.382	0.000****
KUBO 3	607.449	0.000****
KUBO 1	607.449	0.000****
KUBO 40	522.469	0.000****
KUBO 41	522.469	0.000****
KUBO 9	23.486	0.000****

Figure 12. Map of chromsome 2BS with KASP markers (genetic distances based on the population Sv x Cc) and graph with LOD scores plotted in the interval between *KUBO 13* and *KUBO 9*. A high LOD value on the graph indicates a good QTL candidate.



5.7. Physical region and candidate genes

The support interval included between markers *KUBO 27* and *KUBO 41* corresponds to 3.2 Mb on the physical map of the Sv genome. The number of genes present in this region is 93. Description and number of genes corresponding to each description are reported in **Table 8**. The cytochrome P450 is the most represented category, followed by the chloroplastic myrcene synthase. In fact, **Figure 14** indicates that the biological process mostly represented in this list of genes is the metabolic one. Categories that are most useful to define a list of candidate genes responsible for resistance towards SBCMV are three: “defensins”, “NBS-LRR-like resistance proteins” and “disease resistance protein (NBS-LRR class) family”, which correspond to the biological process “defense response”.

Table 8. Description and number of genes present in the support interval of *QSbm.ubo-2BS*, based on the physical map of the Sv genome.

GENE DESCRIPTION	N°
Cytochrome P450	21
Myrcene synthase, chloroplastic	10
Glutamyl-tRNA (Gln) amidotransferase subunit A	9
Receptor kinase 1	7
Serpin-like protein	4
Receptor-like protein kinase	3
Defensin	2
Inner capsid protein lambda-1	2
NBS-LRR-like resistance protein	2
(DL)-glycerol-3-phosphatase 2	1
2-succinyl-5-enolpyruvyl-6-hydroxy-3-cyclohexene-1-carboxylate synthase	1
ABC-2 type transporter family protein	1
ARM repeat superfamily protein	1
ATP synthase subunit b 2	1
ATP-dependent Clp protease adapter protein ClpS	1
Basic helix-loop-helix (bHLH) DNA-binding superfamily protein	1
Carboxyl methyltransferase	1
CDT1-like protein a, chloroplastic	1
Disease resistance protein (NBS-LRR class) family	1
Formamidopyrimidine-DNA glycosylase	1
GDSL esterase/lipase	1
Glycosyltransferase	1
Jasmonate-induced protein	1
Kinase family protein	1
Lysine ketoglutarate reductase trans-splicing-like protein (DUF707)	1
Myosin regulatory light chain 2, skeletal muscle isoform	1
NAD(P)-binding Rossmann-fold superfamily protein	1
Omega-6 fatty acid desaturase, endoplasmic reticulum isozyme 2	1
Peptide methionine sulfoxide reductase MsrA/MsrB	1
Piezo-type mechanosensitive ion channel component	1
PQ-loop repeat family protein / transmembrane family protein	1
Protein kinase superfamily protein	1
Protein kinase, putative	1
Protein MID1-COMPLEMENTING ACTIVITY 1	1
Protein translocase subunit SecA 1	1
Protein UPSTREAM OF FLC	1
Receptor kinase-like protein	1
Receptor-like kinase	1
RING/U-box superfamily protein	1
Terpene synthase	1
TSL-kinase interacting protein 1	1
Zinc/iron-chelating domain protein	1
TOTAL	93

Figure 13. Gene ontology – molecular functions of genes present in the support interval of *QSbm.ubo-2BS.*

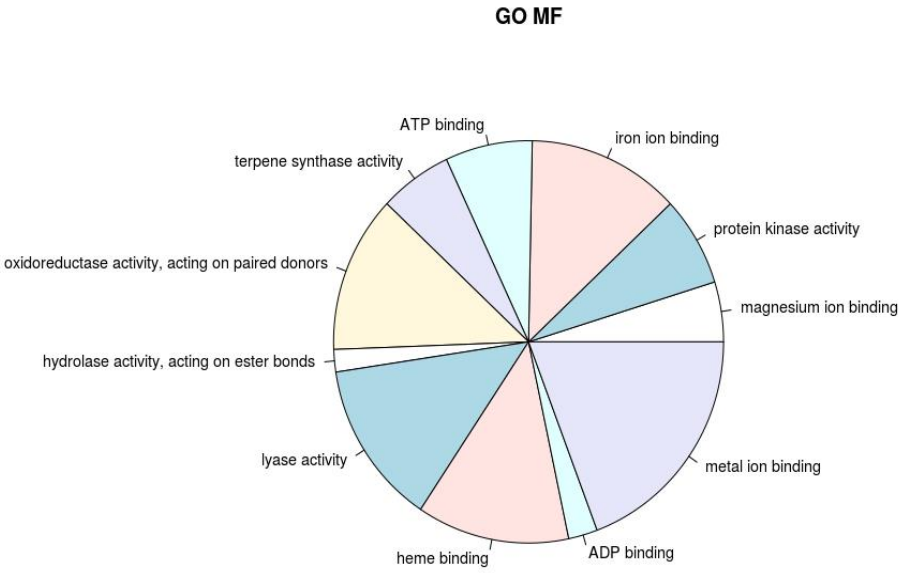
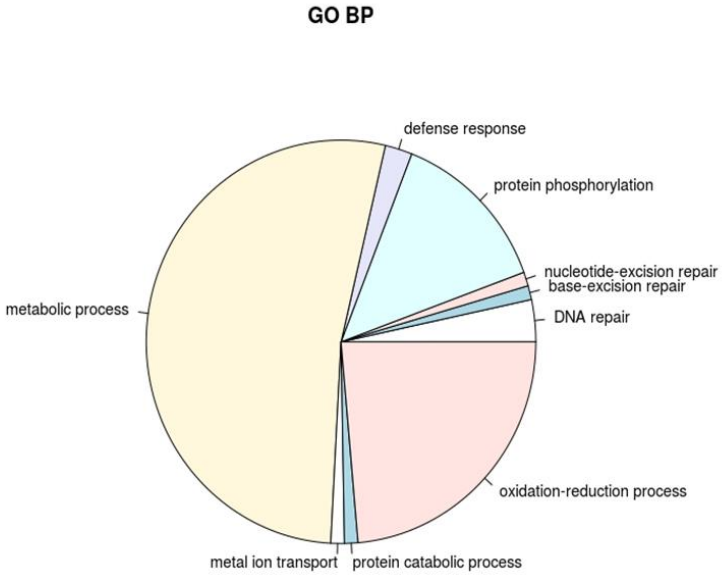


Figure 14. Gene ontology – biological processes of genes present in the support interval of *QSbm.ubo-2BS.*



6. Discussion

6.1. Exploiting SNPs through KASP technology

In this study, two powerful tools were combined with the aim of fine mapping an important QTL responsible for the resistance towards SBCMV in durum wheat. The first one is represented by the SNP-based consensus map of Maccaferri et al. (2015). Previous studies conducted to dissect the genetic basis of SBCMV resistance in durum wheat were based on SSR and DArT markers (Maccaferri et al., 2011, 2012), which were massively exploited in the past years as "second generations markers" (Singh and Singh, 2015). SNP markers provide many advantages in respect to SSR and DArT markers, mainly due to their high density in wheat genome and the possibility to be coupled with high-throughput genotyping technologies, like KASP and HRM. Lately, exploiting SNPs has been made possible thanks to the development of arrays like the Illumina 9K (Cavanagh et al., 2013) and the Illumina 90K (Wang et al., 2014). The consensus map of Maccaferri et al. (2015) provided a large set of SNPs common to several tetraploid wheat biparental populations. Through this instrument, SNPs were selected based both on their position in the region of interest on chromosome 2B and on their polymorphism information.

The second tool is represented by KASP technology. KASP markers have many advantages, such as the availability of a free, user-friendly and well-implemented platform for primers design like PolyMarker and a high-throughput lab protocol. Overall, KASP technology allowed for an effective and convenient exploiting of SNPs markers. The 69% SNP to KASP conversion efficiency obtained resulted satisfactory. The probability of success in converting a SNP in KASP, however, varies greatly according to each SNP. For some SNPs, the design of suitable codominant KASP assays was particularly difficult, most likely because the alignment performed by PolyMarker is based upon Chinese Spring survey sequence scaffolds (International Wheat Genome Sequencing Consortium, 2014) and, as such, does not guarantee the sequence identity with the genotypes used

in this study. For this reason, it is always necessary to validate the developed assays on parental lines of the target population.

Once the assays are designed, their validation is very straightforward and fast. More than one assay can be tested in the same optical 96-wells plate, since the software automatically separates the different SNP assays in different output cluster plots, which allows to save time in plate preparation, run and reading as well as lab consumables. All PCR components are already included in the KASP master mix provided by LGC, to which only DNA, primer mix and water need to be added, hence allowing to speed-up the entire analysis. FRET cassettes present in the master mix are universal and compatible with every kind of assay without the need of marked primers. PCRs can be run on almost every kind of thermal cycler and directly read with a real-time PCR machine, hence allowing to skip the agarose gel electrophoresis step. The output is exportable as an Excel file containing the genotyping results without the need of manually scoring the samples one by one. All these advantages make KASP technology amenable to robotic management and suitable to genotype large populations, like the ~3,000 Sv x Cc RILs, in a short time.

KASP technology is also considerably advantageous in terms of costs, compared to other high-throughput genotyping platforms. The cost for a single data point analysis, for example, is lower for KASP technology than for HRM technology (Terracciano et al., 2013). Yuan et al. (2014) compared KASP and TaqMan® technologies and concluded that the cost of KASP master mix per data point was higher than that of TaqMan®. KASP master mix is produced only by LGC and this does not allow any cost flexibility. On the other end, the universality of FRET system included in the KASP master mix allows its use with primers synthesized by any company. Moreover, as already mentioned above, KASP technology is adaptable to almost any lab equipment in terms of thermal cyclers, optical plates and plate-readers, thus favoring cost reduction.

6.2. Fine mapping of *QSbm.ubo-2BS*

To fine map *QSbm.ubo-2BS*, the 3000 Sv x Cc RILs were fingerprinted with the two KASP markers flanking the QTL region. Genotyping and phenotyping data obtained on the selected RILs reflect the segregation of the resistance trait in the population, confirming the data obtained in the previous studies of Maccaferri et al. (2011, 2012) about the resistance towards SBCMV in durum wheat being controlled mainly by a major genetic locus, which is inherited as a simple Mendelian trait.

The correction model that was applied to phenotypic data allowed for the adjustment of the micro-environmental variations and the uneven incidence of SBCMV infection across the field, phenotyping dates and seasons. In bread wheat, it has been shown that the effects of the infection are strongly influenced by environmental factors such as temperatures and soil humidity that impact both vector and virus spreading and rooting, hence causing spatial patchiness of infection incidence (Bayles et al., 2007; Liu et al., 2014). Moreover, effects of cold damages on plant growth can worsen the virus symptoms. For these reasons, there is a strong necessity of implementing growth chamber experiments under controlled conditions that allow for the standardization of infection effects on the various genotypes.

The fine mapping study performed on the Sv x Cc population identified the *QSbm.ubo-2BS* support interval in 0.2 cM between the markers *KUBO 27* and *KUBO 40-KUBO 41*. However, comparison with the analysis carried out on the M x C RILs suggests that *KUBO 3* and *KUBO 1* can also be excluded from the support interval. Indeed, in Sv x Cc four recombination events were identified involving exactly these two markers, but they could not be excluded since the KASP markers data on the relevant RILs presenting these recombination events were inconsistent. This could be due to the fact that these RILs were not yet fixed at these markers or to a KASP assays not sufficiently efficient with these lines. Further investigation on two backcross populations obtained crossing M with the susceptible RILs M x C 110 and M x C 250, respectively, and on three biparental populations obtained crossing a resistant and a susceptible cultivar, thanks to the collaboration with the French breeding company Florimond Desprez, will help clarifying this point. If the exclusion of

KUBO 3 and *KUBO 1* will be confirmed, then also the DArT marker *wPt-1601* can be excluded from the support interval, which maps in the same recombination bin as *KUBO 3* and *KUBO 1* in M x C and was found to tag *QSbm.ubo-2BS* by Maccaferri et al. (2012). This would allow for an additional narrowing of the interval defined for the QTL.

Another couple of considerations are due on the matter of fine mapping with Sv x Cc RILs. The informative recombinants obtained from this cross are definitely few in respect to those obtained with the cross M x C, considering the number of starting lines available for each population, possibly due to the presence of some suppressor of recombination in a region which is considered to be highly recombinogenic because of its position at the telomere (Akhunov et al., 2003; Conley et al., 2004). The presence of one line showing recombination in the support interval and an intermediate phenotypic value (BLUE = 2.56) reflects the influence of additional QTLs with minor effects, which accounts for the presence of medium-resistant RILs (Maccaferri et al., 2011). The Sv x Cc RILs showing symptoms corresponding to an intermediate level (approximately 2.5) were frequent during field scoring, especially in the second season and in certain areas of the field, accounting for the influence of minor QTLs and of the environment.

Single-marker analysis confirmed that all KASP markers mapping in the interval between *KUBO 13* and *KUBO 9* are highly associated with resistance. However, in order to be exploited in breeding programs aimed at selecting genotypes for resistance to SBCMV, they still need to be validated by GWAS on accession panels. In fact, the capacity to predict the phenotype at the causal locus among genotypes belonging to the breeding germplasm is a relevant aspect for MAS applications, and defines the breeding value of a marker (Terracciano et al., 2013). The advantage of these markers is that they are all codominant and closely linked to the target QTL at the sub-cM level and are thus considered valid tools for molecular breeding (Koebner and Summers, 2003). Moreover, the availability of multiple markers associated to the target locus would favor an accurate haplotype-based screening, an effective approach for the accurate prediction of the susceptible/resistance status of a breeding line (Meuwissen et al., 2001; Mucha and Wierzbicki, 2012). Markers

codominance allowed detecting residual heterozygosity in RILs, which can be exploited in developing Heterogeneous Inbred Families (HIFs), which in turn can be used to develop Near-Isogenic Lines (NILs) with contrasting alleles in the QTL region.

Interval mapping analysis resulted in an extremely high LOD score (between 79.9 and 89.0) for all markers mapped between *KUBO 13* and *KUBO* 9, confirming the evidence of a QTL in the interval. *QSbm.ubo-2BS* has been considered until now the only major QTL responsible for SBCMV resistance in durum wheat, coincident with *Sbm2* in bread wheat (Bayles et al., 2007; Maccaferri et al., 2012). If other major sources of resistance will be found in durum wheat, or in chromosomes A or B of bread wheat, it would be useful to pyramid them into single resistant varieties in order to prevent the selection of new virus strains (Kühne, 2009; Ramirez-Gonzalez et al., 2015a; Terracciano et al., 2013). Moreover, Maccaferri et al. (2012) suggested the possibility of pyramiding sources of resistance arising from minor QTLs. Markers associated with the resistance represent a useful tool for this purpose. Finally, genomic selection (Cabrera-Bosquet et al., 2012; Heffner et al., 2009) could also be used to enhance virus resistance based on the portion of variability due to minor QTLs that is not possible to map due to their small effects.

6.3. Candidate genes in *QSbm.ubo-2BS* interval

By taking a glance at the list of candidate genes proposed as result of this study, three categories seem the most promising to host possible candidate genes responsible of resistance versus SBCMV: defensins, disease resistance protein (NBS-LRR class) family and NBS-LRR-like resistance proteins. Plant defensis are a family of folded antimicrobial peptides that have representatives in vertebrates and invertebrates as well (Broekaert et al., 1995). Plant defensins are thought to be effective especially against fungal pathogen (Thomma et al., 2002). NBS-LRR resistance proteins represent the largest class of plant R genes (resistance genes). R genes are involved in resistance against diverse pathogens, including viruses. In the interval of interest, only one gene belongs to the NBS-LRR resistance protein family and only two are NBS-LRR-like resistance proteins. All these

genes are classified as high-confidence genes, that means coding sequences for which a start and a stop codon have been annotated and for which a representative hit to reference protein sequence is available. Nevertheless, high-confidence/low-confidence classification is not a valid criterion to discriminate putative candidate genes. Jupe et al. (2012), in studies regarding NBS-LRR genes identified in the potato genome, showed that many of these genes that were annotated as partial by the Consortium (2011) were actually full-length and that missing sequences resided in the flanking regions. Therefore, low-confidence genes cannot be excluded from the list of candidates. Neither genes that do not belong to these three categories can be excluded. In fact, although the mechanism of interaction between the virus and his host is not yet completely clear, it is known that, at least for hexaploid wheat, SBCMV-resistant cultivars exhibit what is defined as "translocation resistance", which prevents the virus to spread from the roots to the stem and leaves (Perovic et al., 2009). The responsible of such mechanism could be theoretically any kind of gene, with any kind of function. For this reason, a detailed scan of the entire list is required in order to identify the putative responsible.

7. Conclusions

In this study, SNP markers were used for the first time to dissect the genetics basis of SBCMV resistance in durum wheat. Their high density in wheat genome favored markers enrichment of *QSbm.ubo-2BS* region, while KASP technology allowed for a convenient exploitation of this resource. Considering the wide use of SNP markers in current research, their use in this work represents a further step towards a practical implementation of the results, in respect to the previous use of SSR and DArT markers.

The development of KASP assays allowed for a fast evaluation of a large segregant population like Sv x Cc and the fine mapping of *QSbm.ubo-2BS*, which was located in a 0.2 cM interval that includes the marker *wPt-2106* and is flanked by *KUBO 27* and the co-mapping markers *KUBO 40* and *KUBO 41*. The markers order found in Sv x Cc was confirmed in M x C.

The study provides a series of markers strongly associated to the target QTL. Such markers, once validated, could be included in breeding programs and used to select SBCMV resistant genotypes. Since the consensus map of Maccaferri et al. (2015) was conceived as a bridge to integrate genetics of durum and bread wheat, KASP assays developed from SNPs harbored in this map could be used also in bread wheat, assumed that the assay is polymorphic in the target parental lines.

Moreover, the study lays down the basis for a further refinement of the interval by means of other segregant populations that can be genotyped with the markers herein developed.

Finally, the employment of the parental line Sv allowed to compare the resulting genetic map with the physical map of the recently sequenced Sv genome and to delineate a first list of candidate genes.

8. References

Akhunov, E.D., Goodyear, A.W., Geng, S., Qi, L.-L., Echalier, B., Gill, B.S., Miftahudin, Gustafson, J.P., Lazo, G., Chao, S., Anderson, O.D., Linkiewicz, A.M., Dubcovsky, J., Rota, M.L., Sorrells, M.E., Zhang, D., Nguyen, H.T., Kalavacharla, V., Hossain, K., Kianian, S.F., Peng, J., Lapitan, N.L.V., Gonzalez-Hernandez, J.L., Anderson, J.A., Choi, D.-W., Close, T.J., Dilbirligi, M., Gill, K.S., Walker-Simmons, M.K., Steber, C., McGuire, P.E., Qualset, C.O., Dvorak, J., 2003. The Organization and Rate of Evolution of Wheat Genomes Are Correlated With Recombination Rates Along Chromosome Arms. Genome Res. 13, 753–763. https://doi.org/10.1101/gr.808603

Arumuganathan, K., Earle, E.D., 1991. Nuclear DNA content of some important plant species. Plant Mol. Biol. Report. 9, 208–218. https://doi.org/10.1007/BF02672069

Barcaccia, G., Falcinelli, M., 2007. Genetica e genomica. Liguori.

Bass, C., Hendley, R., Adams, M.J., Hammond-Kosack, K.E., Kanyuka, K., 2006. The *Sbm1* locus conferring resistance to Soil-borne cereal mosaic virus maps to a gene-rich region on 5DL in wheat. Genome 49, 1140–1148. https://doi.org/10.1139/g06-064

Bates, D., Mächler, M., Bolker, B., Walker, S., 2014. Fitting Linear Mixed-Effects Models using lme4. ArXiv14065823 Stat.

Bayles, R., O'Sullivan, D., Lea, V., Freeman, S., Budge, G., Walsh, K., Henry, C., 2007. Controlling Soil-borne cereal mosaic virus in the UK by developing resistant wheat cultivars. HGCA Proj. Rep.

Bayram, M., Öner, M.D., 2006. Determination of applicability and effects of colour sorting system in bulgur production line. J. Food Eng. 74, 232–239. https://doi.org/10.1016/j.jfoodeng.2005.03.016

Bayram, M., Öner, M.D., Kaya, A., 2004. Influence of soaking on the dimensions and colour of soybean for bulgur production. J. Food Eng. 61, 331–339. https://doi.org/10.1016/S0260-8774(03)00137-7

Bianchi, A., Lorenzoni, C., Salamini, F., 1989. Genetica dei cereali. Edagricole.

Broekaert, W.F., Terras, F.R., Cammue, B.P., Osborn, R.W., 1995. Plant defensins: novel antimicrobial peptides as components of the host defense system. Plant Physiol. 108, 1353–1358.

Broman, K.W., 2005. The Genomes of Recombinant Inbred Lines. Genetics 169, 1133–1146. https://doi.org/10.1534/genetics.104.035212

Budge, G.E., Ratti, C., Rubies-Autonell, C., Lockley, D., Bonnefoy, M., Vallega, V., Pietravalle, S., Henry, C.M., 2008. Response of UK winter wheat cultivars to Soil-borne cereal mosaic and Wheat spindle streak mosaic viruses across Europe. Eur. J. Plant Pathol. 120, 259–272. https://doi.org/10.1007/s10658-007-9214-9

Cabrera-Bosquet, L., Crossa, J., von Zitzewitz, J., Serret, M.D., Luis Araus, J., 2012. High-throughput Phenotyping and Genomic Selection: The Frontiers of Crop Breeding Converge. J. Integr. Plant Biol. 54, 312–320. https://doi.org/10.1111/j.1744-7909.2012.01116.x

Carloni, E., Amagliani, G., Omiccioli, E., Ceppetelli, V., Del Mastro, M., Rotundo, L., Brandi, G., Magnani, M., 2017. Validation and application of a quantitative real-time PCR assay to detect common wheat adulteration of durum wheat for pasta production. Food Chem. 224, 86–91. https://doi.org/10.1016/j.foodchem.2016.12.053

Cavanagh, C.R., Chao, S., Wang, S., Huang, B.E., Stephen, S., Kiani, S., Forrest, K., Saintenac, C., Brown-Guedira, G.L., Akhunova, A., See, D., Bai, G., Pumphrey, M., Tomar, L., Wong, D., Kong, S., Reynolds, M., Silva, M.L. da, Bockelman, H., Talbert, L., Anderson, J.A., Dreisigacker, S., Baenziger, S., Carter, A., Korzun, V., Morrell, P.L., Dubcovsky, J., Morell, M.K., Sorrells, M.E., Hayden, M.J., Akhunov, E., 2013. Genome-wide comparative diversity uncovers multiple targets of selection for improvement in hexaploid wheat

landraces and cultivars. Proc. Natl. Acad. Sci. 110, 8057–8062. https://doi.org/10.1073/pnas.1217133110

Çelik, İ., Işık, F., Gürsoy, O., 2004. Couscous, a traditional Turkish food product: production method and some applications for enrichment of nutritional value. Int. J. Food Sci. Technol. 39, 263–269. https://doi.org/10.1111/j.1365-2621.2004.00780.x

Collard, B.C.Y., Jahufer, M.Z.Z., Brouwer, J.B., Pang, E.C.K., 2005. An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop improvement: The basic concepts. Euphytica 142, 169–196. https://doi.org/10.1007/s10681-005-1681-5

Conley, E.J., Nduati, V., Gonzalez-Hernandez, J.L., Mesfin, A., Trudeau-Spanjers, M., Chao, S., Lazo, G.R., Hummel, D.D., Anderson, O.D., Qi, L.L., Gill, B.S., Echalier, B., Linkiewicz, A.M., Dubcovsky, J., Akhunov, E.D., Dvořák, J., Peng, J.H., Lapitan, N.L.V., Pathan, M.S., Nguyen, H.T., Ma, X.-F., Miftahudin, Gustafson, J.P., Greene, R.A., Sorrells, M.E., Hossain, K.G., Kalavacharla, V., Kianian, S.F., Sidhu, D., Dilbirligi, M., Gill, K.S., Choi, D.W., Fenton, R.D., Close, T.J., McGuire, P.E., Qualset, C.O., Anderson, J.A., 2004. A 2600-Locus Chromosome Bin Map of Wheat Homoeologous Group 2 Reveals Interstitial Gene-Rich Islands and Colinearity With Rice. Genetics 168, 625–637. https://doi.org/10.1534/genetics.104.034801

Consortium, T.P.G.S., 2011. Genome sequence and analysis of the tuber crop potato. Nature 475, 189. https://doi.org/10.1038/nature10158

Coskun, F., 2013. Production of couscous using the traditional method in Turkey and couscous in the world. Afr. J. Agric. Res. 8, 2609–2615. https://doi.org/10.5897/AJAR12.2195

Donnely, B.J., Debbouz, A., Hagen, K.R., Foundation, N.D.S.U.R., 1994. Extrusion method for making a couscous food product.

Doyle, Doyle, 1990. Isolation of plant DNA from fresh tissue. Focus 12, 13–15.

Erbaş, M., Aykın, E., Arslan, S., Durak, A.N., 2016. Adsorption behaviour of bulgur. Food Chem., Special issue from the 8th International Conference on Water in Food 195, 87–90. https://doi.org/10.1016/j.foodchem.2015.06.050

Falcon, S., Gentleman, R., 2007. Using GOstats to test gene lists for GO term association. Bioinformatics 23, 257–258. https://doi.org/10.1093/bioinformatics/btl567

Gadaleta, A., Giancaspro, A., Giove, S.L., Zacheo, S., Mangini, G., Simeone, R., Signorile, A., Blanco, A., 2009. Genetic and physical mapping of new EST-derived SSRs on the A and B genome chromosomes of wheat. Theor. Appl. Genet. 118, 1015. https://doi.org/10.1007/s00122-008-0958-1

Gill, B.S., Appels, R., Botha-Oberholster, A.-M., Buell, C.R., Bennetzen, J.L., Chalhoub, B., Chumley, F., Dvořák, J., Iwanaga, M., Keller, B., Li, W., McCombie, W.R., Ogihara, Y., Quetier, F., Sasaki, T., 2004. A Workshop Report on Wheat Genome Sequencing: International Genome Research on Wheat Consortium. Genetics 168, 1087–1096. https://doi.org/10.1534/genetics.104.034769

Graziani, M., 2012. Mappaggio fine di un QTL principale per la resa in granella sul cromosoma 3B di frumento duro (Tesi di dottorato). DOI 10.6092/unibo/amsdottorato/4283.

Heffner, E.L., Sorrells, M.E., Jannink, J.-L., 2009. Genomic Selection for Crop Improvement. Crop Sci. 49, 1–12. https://doi.org/10.2135/cropsci2008.08.0512

International Wheat Genome Sequencing Consortium (IWGSC), 2014. A chromosome-based draft sequence of the hexaploid bread wheat (*Triticum aestivum*) genome. Science 345, 1251788. https://doi.org/10.1126/science.1251788

Jupe, F., Pritchard, L., Etherington, G.J., MacKenzie, K., Cock, P.J., Wright, F., Sharma, S.K., Bolser, D., Bryan, G.J., Jones, J.D., Hein, I., 2012. Identification and localisation of the NB-LRR gene family within the potato genome. BMC Genomics 13, 75. https://doi.org/10.1186/1471-2164-13-75

Kanyuka, K., Lovell, D.J., Mitrofanova, O.P., Hammond-Kosack, K., Adams, M.J., 2004. A controlled environment test for resistance to Soil-borne cereal mosaic virus (SBCMV) and its use to determine the mode of inheritance of resistance in wheat cv. Cadenza and for

screening *Triticum monococcum* genotypes for sources of SBCMV resistance. Plant Pathol. 53, 154–160. https://doi.org/10.1111/j.0032-0862.2004.01000.x

Kellogg, E.A., 2001. Evolutionary history of the grasses. Plant Physiol. Rockv. 125, 1198–205.

Koebner, R.M.D., Summers, R.W., 2003. 21st century wheat breeding: plot selection or plate detection? Trends Biotechnol. 21, 59–63. https://doi.org/10.1016/S0167-7799(02)00036-7

Kühne, T., 2009. Soil-borne viruses affecting cereals—Known for long but still a threat. Virus Res., Plant Virus Epidemiology: Controlling epidemics of emerging and established plant viruses - the way forward 141, 174–183. https://doi.org/10.1016/j.virusres.2008.05.019

Liu, S., Yang, X., Zhang, D., Bai, G., Chao, S., Bockus, W., 2014. Genome-wide association analysis identified SNPs closely linked to a gene resistant to Soil-borne wheat mosaic virus. TAG Theor. Appl. Genet. Theor. Angew. Genet. 127, 1039–1047. https://doi.org/10.1007/s00122-014-2277-z

Maccaferri, M., Francia, R., Ratti, C., Rubies-Autonell, C., Colalongo, C., Ferrazzano, G., Tuberosa, R., Sanguineti, M.C., 2012. Genetic analysis of Soil-Borne Cereal Mosaic Virus response in durum wheat: evidence for the role of the major quantitative trait locus QSbm.ubo-2BS and of minor quantitative trait loci. Mol. Breed. 29, 973–988. https://doi.org/10.1007/s11032-011-9673-8

Maccaferri, M., Ratti, C., Rubies-Autonell, C., Vallega, V., Demontis, A., Stefanelli, S., Tuberosa, R., Sanguineti, M.C., 2011. Resistance to Soil-borne cereal mosaic virus in durum wheat is controlled by a major QTL on chromosome arm 2BS and minor loci. TAG Theor. Appl. Genet. Theor. Angew. Genet. 123, 527–544. https://doi.org/10.1007/s00122-011-1605-9

Maccaferri, M., Ricci, A., Salvi, S., Milner, S.G., Noli, E., Martelli, P.L., Casadio, R., Akhunov, E., Scalabrin, S., Vendramin, V., Ammar, K., Blanco, A., Desiderio, F., Distelfeld, A., Dubcovsky, J., Fahima, T., Faris, J., Korol, A., Massi, A., Mastrangelo, A.M., Morgante, M., Pozniak, C., N'Diaye, A., Xu, S., Tuberosa, R., 2015. A high-density, SNP-based consensus map of tetraploid wheat as a bridge to integrate durum and bread wheat genomics and breeding. Plant Biotechnol. J. 13, 648–663. https://doi.org/10.1111/pbi.12288

Magnatti, P., 2007. Caratteristiche, evoluzione e prospettive del settore delle paste alimentari. L'industria. https://doi.org/10.1430/24253

Meuwissen, T.H.E., Hayes, B.J., Goddard, M.E., 2001. Prediction of Total Genetic Value Using Genome-Wide Dense Marker Maps. Genetics 157, 1819–1829.

Mucha, A., Wierzbicki, H., 2012. Linear models for breeding values prediction in haplotype-assisted selection - an analysis of QTL-MAS Workshop 2011 Data. BMC Proc. 6, S11. https://doi.org/10.1186/1753-6561-6-S2-S11

Ordon, F., Habekuss, A., Kastirr, U., Rabenstein, F., Kühne, T., 2009. Virus Resistance in Cereals: Sources of Resistance, Genetics and Breeding. J. Phytopathol. 157, 535–545. https://doi.org/10.1111/j.1439-0434.2009.01540.x

Perovic, D., Förster, J., Devaux, P., Hariri, D., Guilleroux, M., Kanyuka, K., Lyons, R., Weyen, J., Feuerhelm, D., Kastirr, U., Sourdille, P., Röder, M., Ordon, F., 2009. Mapping and diagnostic marker development for Soil-borne cereal mosaic virus resistance in bread wheat. Mol. Breed. 23, 641–653. https://doi.org/10.1007/s11032-009-9262-2

Petitot, M., Boyer, L., Minier, C., Micard, V., 2010. Fortification of pasta with split pea and faba bean flours: Pasta processing and quality evaluation. Food Res. Int., Molecular, Functional and Processing Characteristics of Whole Pulses and Pulse Fractions and their Emerging Food and Nutraceutical Applications 43, 634–641. https://doi.org/10.1016/j.foodres.2009.07.020

R Core Team, 2017. R: A Language and Environment for Statistical Computing.

Ramirez-Gonzalez, R.H., Segovia, V., Bird, N., Fenwick, P., Holdgate, S., Berry, S., Jack, P., Caccamo, M., Uauy, C., 2015a. RNA-Seq bulked segregant analysis enables the identification of high-resolution genetic markers for breeding in hexaploid wheat. Plant Biotechnol. J. 13, 613–624. https://doi.org/10.1111/pbi.12281

Ramirez-Gonzalez, R.H., Uauy, C., Caccamo, M., 2015b. PolyMarker: A fast polyploid primer design pipeline. Bioinformatics 31, 2038–2039. https://doi.org/10.1093/bioinformatics/btv069

Rubies-Autonell, C., Vallega, V., Ratti, C., 2003. Reactions of cultivars of common wheat (*Triticum aestivum L.*) to Soilborne wheat mosaic virus in northern Italy / Reaktionen von Weizensorten auf den Befall durch SBWMV in Norditalien. Z. Für Pflanzenkrankh. Pflanzenschutz J. Plant Dis. Prot. 110, 332–336.

Salvi, S., Castelletti, S., Tuberosa, R., 2009. An updated consensus map for flowering time qtls in maize.

Salvi, S., Tuberosa, R., 2005. To clone or not to clone plant QTLs: present and future challenges. Trends Plant Sci. 10, 297–304. https://doi.org/10.1016/j.tplants.2005.04.008

Sarkar, P., Stebbins, G.L., 1956. Morphological Evidence Concerning the Origin of the B Genome in Wheat. Am. J. Bot. 43, 297–304. https://doi.org/10.2307/2438947

Singh, B.D., Singh, A.K., 2015. Marker-Assisted Plant Breeding: Principles and Practices. Springer.

Talbert, L.E., Blake, N.K., Storlie, E.W., Lavin, M., 1995. Variability in wheat based on low-copy DNA sequence comparisons. Genome 38, 951–957. https://doi.org/10.1139/g95-125

Terracciano, I., Maccaferri, M., Bassi, F., Mantovani, P., Sanguineti, M.C., Salvi, S., Šimková, H., Doležel, J., Massi, A., Ammar, K., Kolmer, J., Tuberosa, R., 2013. Development of COS-SNP and HRM markers for high-throughput and reliable haplotype-based detection of Lr14a in durum wheat (*Triticum durum* Desf.). Theor. Appl. Genet. 126, 1077–1101. https://doi.org/10.1007/s00122-012-2038-9

Thomma, B.P., Cammue, B.P., Thevissen, K., 2002. Plant defensins. Planta 216, 193–202. https://doi.org/10.1007/s00425-002-0902-6

Vallega, V. (Istituto S. per la C., Rubies-Autonell, C., Ratti, C. (Bologna U. (Italy) D. di S. e T.A., 2003. Reaction of durum wheat cultivars to mixed SBWMV and WSSMV infection in Central Italy [Triticum durum Desf. - Latium]. Phytopathol. Mediterr. Italy.

Vallega, V., Rubies Autonell, C., 1985. Reactions of Italian *Triticum durum* cultivars to soilborne wheat mosaic. Plant Dis. 69, 64–66.

Van Ooijen, J.W., 2006. JoinMap 4. Softw. Calc. Genet. Link. Maps Exp. Popul. Kyazma BV Wagening. Neth. 33.

van Slageren, M.W., 1994. Wild wheats: a monograph of *Aegilops L.* and *Amblyopyrum* (Jaub. & Spach) Eig (Poaceae).

Varshney, R.K., Tuberosa, R., 2007. Genomics-Assisted Crop Improvement: An Overview, in: Genomics-Assisted Crop Improvement. Springer, Dordrecht, pp. 1–12. https://doi.org/10.1007/978-1-4020-6295-7_1

Wang, S., Basten, C.J., Zeng, Z.-B., 2012. Windows QTL Cartographer 2.5.

Wang, S., Wong, D., Forrest, K., Allen, A., Chao, S., Huang, B.E., Maccaferri, M., Salvi, S., Milner, S.G., Cattivelli, L., Mastrangelo, A.M., Whan, A., Stephen, S., Barker, G., Wieseke, R., Plieske, J., International Wheat Genome Sequencing Consortium, Lillemo, M., Mather, D., Appels, R., Dolferus, R., Brown-Guedira, G., Korol, A., Akhunova, A.R., Feuillet, C., Salse, J., Morgante, M., Pozniak, C., Luo, M.-C., Dvorak, J., Morell, M., Dubcovsky, J., Ganal, M., Tuberosa, R., Lawley, C., Mikoulitch, I., Cavanagh, C., Edwards, K.J., Hayden, M., Akhunov, E., 2014. Characterization of polyploid wheat genomic diversity using a high-density 90 000 single nucleotide polymorphism array. Plant Biotechnol. J. 12, 787–796. https://doi.org/10.1111/pbi.12183

Yuan, J., Wen, Z., Gu, C., Wang, D., 2014. Introduction of High Throughput and Cost Effective SNP Genotyping Platforms in Soybean. Plant Genet. Genomics Biotechnol. 2, 90–94. https://doi.org/10.5147/pggb.v2i1.155

Zadoks, J.C., Chang, T.T., Konzak, C.F., 1974. A decimal code for the growth stages of cereals. Weed Res. 14, 415–421. https://doi.org/10.1111/j.1365-3180.1974.tb01084.x

Zohary, D., Feldman, M., 1962. Hybridization Between Amphidiploids and the Evolution of Polyploids in the Wheat (*aegilops-Triticum*) Group. Evolution 16, 44–61. https://doi.org/10.1111/j.1558-5646.1962.tb03197.x

Cited websites:

Food and Agriculture Organization of the United Nations (FAO): http://www.fao.org
Ismeamercati: http://www.ismeamercati.it/analisi-e-studio-filiere-agroalimentari
LGC Group: https://www.lgcgroup.com/products/kasp-genotyping-chemistry/
Miller Magazine: http://www.millermagazine.com/english
Polymarker: http://Polymarker.tgac.ac.uk
Union des Associations de Fabricants de Pates Alimentaires de l'U.E. (UN.A.F.P.A.): http://www.pasta-unafpa.org/
U.S. Department of Agriculture (USDA): https://www.ers.usda.gov/data-products/wheat-data.aspx
Wheat training: http://www.wheat-training.com

9. Acknowledgements

Part of this research study was conducted with the support of the project “FSOV 2016 C MOSAIC RESISTANCE IN DURUM”.

I would like to thank my supervisor Professor Roberto Tuberosa for giving me the chance to work in his research group during my PhD program.

I would like to acknowledge who participated with me in this research study: Dr. Marco Maccaferri, Dr. Giuseppe Sciara, Dr. Danara Ormanbekova, Dr. Simona Corneti, Dr. Sandra Stefanelli and Dr. Elena Fusari.

Moreover, I would like to thank all the colleagues at DipSA: Dr. Giuseppe Condorelli, Dr. Eder Groli, Dr. Linda Zamariola, Dr. Silvia Giuliani, Dr. Riccardo Bovina, Dr. Fabio Valli and Prof. Silvio Salvi for help, advices, and company during these three years.

Finally, I would like to thank my husband Giulio for moral support and practical help.