



Dissecting the genetic basis of resistance to *Soil-borne cereal mosaic virus* (SBCMV) in durum wheat by bi-parental mapping and GWAS

Martina Bruschi¹ · Matteo Bozzoli¹ · Claudio Ratti¹ · Giuseppe Sciarra¹ · Ellen Goudemand² · Pierre Devaux² · Danara Ormanbekova¹ · Cristian Forestan¹ · Simona Corneti¹ · Sandra Stefanelli¹ · Sara Castelletti¹ · Elena Fusari¹ · Jad B Novi¹ · Elisabetta Frascaroli¹ · Silvio Salvi¹ · Dragan Perovic³ · Agata Gadaleta⁴ · Concepcion Rubies-Autonell¹ · Maria Corinna Sanguineti¹ · Roberto Tuberosa¹ · Marco Maccaferri¹

Received: 14 August 2023 / Accepted: 4 August 2024 / Published online: 2 September 2024
© The Author(s) 2024

Abstract

Soil-borne cereal mosaic virus (SBCMV), the causative agent of wheat mosaic, is a *Furovirus* challenging wheat production all over Europe. Differently from bread wheat, durum wheat shows greater susceptibility and stronger yield penalties, so identification and genetic characterization of resistance sources are major targets for durum genetics and breeding. The *Sbm1* locus providing high level of resistance to SBCMV was mapped in bread wheat to the 5DL chromosome arm (Bass in Genome 49:1140–1148, 2006). This excluded the direct use of *Sbm1* for durum wheat improvement. Only one major QTL has been mapped in durum wheat, namely *QSbm.ubo-2B*, on the 2BS chromosome region coincident with *Sbm2*, already known in bread wheat as reported (Bayles in HGCA Project Report, 2007). Therefore, *QSbm.ubo-2B* = *Sbm2* is considered a pillar for growing durum in SBCMV-affected areas. Herein, we report the fine mapping of *Sbm2* based on bi-parental mapping and GWAS, using the Infinium 90 K SNP array and high-throughput KASP®. Fine mapping pointed out a critical haploblock of 3.2 Mb defined by concatenated SNPs successfully converted to high-throughput KASP® markers coded as *KUBO*. The combination of *KUBO-27*, *wPt-2106-ASO/HRM*, *KUBO-29*, and *KUBO-1* allows unequivocal tracing of the *Sbm2*-resistant haplotype. The interval harbors 52 high- and 41 low-confidence genes, encoding 17 cytochrome p450, three receptor kinases, two defensins, and three *NBS-LRR* genes. These results pave the way for *Sbm2* positional cloning. Importantly, the development of *Sbm2* haplotype tagging KASP® provides a valuable case study for improving efficacy of the European variety testing system and, ultimately, the decision-making process related to varietal characterization and choice.

Communicated by Hermann Buerstmayr.

Martina Bruschi and Matteo Bozzoli have contributed equally to this work.

✉ Marco Maccaferri
marco.maccaferri@unibo.it

¹ Department of Agricultural and Food Sciences (DISTAL), Alma Mater Studiorum – Università di Bologna, 40127 Bologna, Italy

² S.A.S. Florimond-Desprez Veuve and Fils, BP41, 59242 Cappelle-en-Pévèle, France

³ Federal Research Centre for Cultivated Plants, Institute for Resistance Research and Stress Tolerance, Julius Kühn-Institut (JKI), Erwin-Baur-Str. 27, 06484 Quedlinburg, Germany

⁴ Department of Soil, Plant and Food Science (Di.S.S.P.A.), University of Bari ‘Aldo Moro’, 70126 Bari, Italy

Abbreviations

BLUES	Best linear unbiased estimates
Cc	Cultivar Ciccio
Cl	Cultivar Claudio
Cv	Cultivar
DArT®	Diversity array technology marker
ELISA	Enzyme-linked immunosorbent assay.
GO	Gene ontology
GS	Genetic similarity
GWAS	Genome-wide association study
HC	Genes annotated at high confidence
IWGSC	International Wheat Genome Sequencing Consortium
IBS	Identity by state
IBD	Identity by descent
K	Kinship
KASP®	Kompetitive allele-specific PCR
KUBO	KASP® marker developed by DISTAL laboratory at University of Bologna

LC	Genes annotated at low confidence
MAS	Marker-assisted selection
MLM	Mixed linear model
Mr	Cultivar Meridiano
NJ	Neighbor joining tree
PVE	Proportion of variance explained
QTL	Quantitative trait locus
RIL	Recombinant inbred line mapping population
SBCMV	Soil-borne cereal mosaic virus
Sbm1	Soil-borne mosaic resistance locus 1.
Sbm2	Soil-borne mosaic resistance locus 2
SBWMV	Soil-borne wheat mosaic virus
SEV	Disease severity
SNP	Single-nucleotide polymorphism
Sv	Cultivar Svevo

Introduction

Soil-borne cereal mosaic virus (SBCMV) is a furovirus affecting the production of both tetraploid durum (*Triticum turgidum* var. *durum* Desf.) and hexaploid bread wheat (*Triticum aestivum* L.) across Europe, especially in France, Italy, Germany, Poland, and Denmark (Budge et al. 2008; Kanyuka et al. 2004; Maccaferri et al. 2011a; Rubies-Autonell et al. 2003). SBCMV was initially considered the European strain of *soil-borne wheat mosaic virus* (SBWMV), a member of the genus *Furovirus* which affects mainly wheat crops in U.S., and later recognized by the International Committee on Taxonomy of Viruses as a separate species (Kanyuka et al. 2004).

Fields heavily infected with SBCMV have been shown to suffer a yield reduction of ca. 40–50% in susceptible winter wheat cultivars in the UK and up to 70% in highly susceptible durum wheat in Italy (Clover et al. 1999; Kühne 2009; Vallega and Rubies Autonell 1985; Vallega et al. 2003).

The name “soil-borne” is due to the infection mechanism which takes advantage of the plasmodiophorid vector *Polymyxa graminis*, a soil inhabitant fungal-like endoparasitic protist slime mold, endoparasite of wheat. As a vector, *P. graminis* produces resting spores that are easily spread mainly by water, wind, and machinery. SBCMV nests inside the resting spores of the vector during periods of adverse environmental conditions and long rotations of non-host crops, persisting and accumulating in soil for decades. When conditions reverse to favorable for disease development (e.g., presence of susceptible hosts and high soil moisture, particularly in the fall, when extended periods of high soil moisture favor the movement of *Polymyxa graminis*), the virus is released from the resting spores and, once vectored through zoospores, enters root hair cells (Kühne 2009). From there, the virus can reach the aerial part of the plant and infect the leaf tissue, causing symptoms ranging from mild mottling to

severe leaf mosaic. Moreover, vigor and plant height are also affected, with susceptible cultivars showing stunting (Kim et al. 2017; Kühne 2009; Perovic et al. 2009; Roberts 2014).

Due to the nature and persistence of the natural inoculum, agronomic practices like crop rotation or delayed sowing are mostly ineffective once the soil is infected. Chemical defense is neither effective nor environmentally friendly. As a consequence, the only possible option is the development and deployment of resistant cultivars (Kanyuka et al. 2003; Ordon et al. 2009).

The first and most effective SBCMV resistance locus identified in the bread wheat germplasm was *Sbm1*, located on chromosome 5DL (Bass et al. 2006). Resistance was detected in bread wheat cv. Cadenza by assessing a doubled haploid population obtained from the cross Avalon (susceptible cultivar) × Cadenza (resistant cultivar). Resistance, mainly of “translocation type” preventing the infection to spread from the roots to the stem and leaves (Kanyuka et al. 2004; Ordon et al. 2009), was found to be controlled by a major locus named *Sbm1* mapped to the long arm of chromosome 5D tagged by microsatellite *Xbarc110* and *Xwmc765* in a region of approximately 17 cM (Bass et al. 2006). Due to the relevance of *Sbm1* for hexaploid wheat breeding, it has been recently fine mapped (Liu et al. 2020) and confined in the 546,651,779–547,273,461-bp interval of Chinese Spring RefSeq v1.0 pseudomolecule (International Wheat Genome Sequencing Consortium (IWGSC), 2018), harboring 17 candidate genes, with *pto-interacting protein 1* (*PTI1*, TraesCS5D01G531200) identified as the most promising one. After Bass et al. (2006), another major QTL named *Sbm2* was mapped in the same Avalon × Cadenza population by Bayles et al. (2007) in chromosome arm 2BS. *Sbm2* acts in a distinct and complementary way to *Sbm1*, suggesting the possibility to pyramid these two sources of resistance in bread wheat. Taken together, these two major *Sbm* loci enable breeding for bread wheat varieties with valuable SBCMV resistance.

Conversely, durum wheat lacks the D genome; thus, *Sbm1* resistance locus is not directly exploitable for durum wheat breeding. The potentially higher impact of SBCMV on grain yield loss and the lack of known sources of resistance in durum wheat as compared to hexaploid wheat prompted the search for valuable sources of resistance in the durum wheat germplasm. Using a RIL population from Meridiano (resistant) × Claudio (moderately susceptible), Maccaferri et al. (2011a) identified *QSbm.ubo-2B*, a major QTL for SBCMV resistance in durum wheat located in the distal telomeric region of chromosome 2BS, close to the *wPt-2106* DArT® marker and three SSR markers (*Xwmc661-Xgwm210-Xbarc35*). The large proportion of variance explained (PVE) by this locus (55.8–68.0% for ELISA and up to 91.6% for symptom severity) suggested *QSbm.ubo-2B* as a potentially major target for durum

wheat breeding, marker-assisted selection, and QTL cloning. *QSBm.ubo-2B* was confirmed in a durum wheat panel (Maccaferri et al. 2011a) and in a second mapping population, Simeto (susceptible) × Levante (resistant) with a PVE of 60–70% for symptom severity. Meta-QTL analysis mapped *QSBm.ubo-2BS* on a two cM-wide CI centered to *wPt-2106*, proximal to *Xwmc661-Xgwm210-Xbarc35*, which allowed for the first molecular assay to trace the QTL resistant allele. Based on the tetraploid wheat consensus map (Maccaferri et al. 2015) and meta-QTL analysis, *QSBm.ubo-2BS* (Maccaferri et al. 2012) was mapped in a region coincident with the *Sbm2* locus in bread wheat (*QSBm.ubo-2BS = Sbm2*). This makes *QSBm.ubo-2BS = Sbm2* a valuable target for basic and applied research aimed at its fine mapping and positional cloning.

Currently, the use of single-nucleotide polymorphisms (SNPs, Yuan et al. 2014) has become more efficient, cheaper, and robust than other types of markers and related sequencing techniques in both bread and durum wheat, thanks to the development of species-specific SNP arrays like the Illumina Infinium *iSelect* 90 K wheat SNP array (Cavanagh et al. 2013; Wang et al. 2014). Maccaferri et al. (2015) developed a consensus map for tetraploid wheat harboring 30,144 markers, with high-density gene-derived SNPs as useful anchor points to isolate target loci. SNPs abundance in the wheat genome, together with the development of high-throughput fluorescence-based marker assays, like KASP® (Kompetitive Allele Specific PCR, Semang et al. 2014), makes them ideal for fine mapping and MAS breeding purposes.

The publication of the durum wheat reference cultivar Svevo (Maccaferri et al. 2019) allowed us to easily switch from the genetic to the physical map, accelerating both fine mapping and candidate gene identification. Genome-wide association studies (GWAS) coupled with dense SNPs arrays enable efficient exploitation of large phenotypic datasets to search for and fine map QTLs (Korte and Farlow 2013).

The present work aimed at fine mapping *QSBm.ubo-2BS* using a combination of mapping populations and GWAS. The development of KASP® markers was pursued with the double objective of providing genic markers anchored to the durum wheat Svevo genome assembly suitable for fine mapping of *Sbm2* and, at the same time, assays ready to implement MAS. These results were integrated with candidate genes, identified within the QTL region projected upon the durum wheat cv. Svevo, hexaploid wheat cv. Chinese Spring, and the wild emmer wheat Zavitan (*Triticum turgidum* ssp. *dicoccoides*, Körn. Ex Asch. & Graebn.) reference genomes. Moreover, the results of an extensive survey of the *QSBm.ubo-2BS = Sbm2*-resistant/susceptible haplotype distribution in durum wheat panels grown in Europe and worldwide are reported.

Materials and methods

Plant materials

Mapping populations and cultivated germplasm panels

A recombinant inbred line (RIL) population developed by the University of Bari (Gadaleta et al. 2009) composed of 3,075 lines obtained from cv. Svevo (hereafter referred to as Sv, resistant) and Ciccio (hereafter referred to as Cc, susceptible) was used for *Sbm2* fine mapping. Svevo harbors the *Xwmc661-Xgwm210-Xbarc35-wPt-2106* resistance tagging haplotype at *Sbm2* while Ciccio carries the susceptible haplotype, similar to Simeto (Maccaferri et al. 2011a). Cultivar Svevo was sequenced, assembled, and annotated at gold-standard level (Maccaferri et al. 2019). Ciccio is adapted to Southern Mediterranean conditions and is a derivative of Valnova founder, highly susceptible to SBCMV. A second population includes 181 RILs developed by Produttori Sementi Bologna SpA from a cross between cv. Meridiano (hereafter referred to as Mr, resistant) and Claudio (hereafter referred to as Cl, susceptible). The Mr × Cl RIL population was described in Maccaferri et al. (2011a, 2012).

The cultivated durum wheat panels herein considered provide a representation of the cultivated germplasm worldwide spanning a breeding period from 1914/1920 to 2020. A total of 549 varieties/breeding lines included in this study. A first panel, referred as the "UNIBO Durum Panel" (Maccaferri et al. 2011b), was used to perform GWAS and assess linkage disequilibrium extent in the *Sbm2* region and the repeatability and information content of diagnostic Illumina SNP markers converted to fluorescent-based high-throughput KASP® markers. The UNIBO Durum Panel included 258 breeding lines and cultivars released from 1970 to early 2000 from Italy, Spain, CIMMYT, ICARDA, North America, Southwestern USA, France, Austria, and Australia (Maccaferri et al. 2011b). This panel was field phenotyped for SBCMV response over three years. A wider panel including 291 more recent and diverse European varieties was assembled and analyzed in the framework of the Horizon 2020 project "InnoVar, next-generation variety testing for improved cropping on European farmland" and other sources. The complete panel was used to investigate the distribution of SBCMV-resistant and SBCMV-susceptible haplotypes worldwide and across breeding decades.

Molecular analysis

Enrichment of the *Sbm2* region with KASP® markers

Based on the consensus map of Maccaferri et al. (2015) and Svevo genome sequence, KASP® assays were developed from 16 Illumina SNPs previously mapped in the

Sbm2 region. Features of these SNPs are reported in Table 1. Primers were designed using PolyMarker (<http://Polymarker.tgac.ac.uk>) followed by manual curation. Probe sequences tagged with fluorescent dyes FAM and HEX were added to the 5'-end of A and B primers, respectively. KASP® marker development protocol is described in Method S1.

Table 1 Summary information with regard to the Illumina Infinium *iSelect* 90 K wheat SNPs chosen to develop polymorphic KASP® markers in the *Sbm2* confidence interval. Illumina 90 K wheat SNPs ID code, SNP marker name, newly developed KASP® marker name,

genetic position in the durum wheat consensus map, physical location in durum wheat cv. Svevo genome, and polymorphism in durum / tetraploid recombinant inbred lines. Markers are listed from the distal -5' to the proximal -3'- 2BS chromosomal region

SNP marker ID code	SNP marker name	KASP® marker	Consensus map position (cM) ^a	Svevo genome assembly position (bp) ^b	Polymorphisms in RIL maps ^c
IWB61884	RAC875_rep_c109471_154	KUBO-015*	7.9	9,766,826	Lt×MG_5323, Cs×Ld, Mr×Cl, Mh×Cr, Sv×Cc
IWB73347	Tdurum_contig76118_145	KUBO-012*	8.4	9,986,862	Lt×MG_5323, Cs×Ld, Mr×Cl, Mh×Cr, Sv×Cc, Sv×Zv
IWB11421	BS00085748_51	KUBO-026*	11.6	13,011,332	Lt×MG_5323, Cs×Ld, Mr×Cl, Sm×Lv, Mh×Cr, Sv×Cc, Sv×Zv
IWB23029	Excalibur_c1787_1037	KUBO-029*	11.6	15,643,691	Cs×Ld, Mr×Cl, Sm×Lv, Sv×Cc
IWB42660	Kukri_c22513_1780	KUBO-032	11.6	13,011,658	Lt×MG_5323, Cs×Ld, Mr×Cl, Sm×Lv, Mh×Cr, Sv×Cc, Sv×Zv
IWB28973	Excalibur_c8093_82	KUBO-001*	12.2	16,235,907	Cs×Ld, Mr×Cl, Mh×Cr, Sv×Cc
IWB45152	Kukri_c41556_619	KUBO-002	12.2	15,713,309	Cs×Ld, Mr×Cl, Mh×Cr, Sv×Cc
IWB41644	Kukri_c16758_443	KUBO-035	12.2	16,183,711	Sm×MC, Cs×Ld, Mr×Cl, Mh×Cr, Sv×Cc
IWB8328	BS00043055_51	KUBO-003*	12.3	15,805,908	Lt×MG_5323, Cl×Ld, Mr×Cl, Sm×Lv, Mh×Cr, Sv×Cc
IWB29097	Excalibur_c841_609	KUBO-040*	12.3	16,185,445	Sm×MC, Cs×Ld, Mr×Cl, Mh×Cr, Sv×Cc
IWB35524	IAAV8700	KUBO-041*	12.3	16,184,325	Sm×MC, Cs×Ld, Mr×Cl, Mh×Cr, Sv×Cc
IWB23330	Excalibur_c19499_948	KUBO-004	12.4	16,957,139	Cs×Ld, Mr×Cl, Sv×Cc
IWB24939	Excalibur_c30167_531	KUBO-005*	12.4	16,954,167	Sm×MC, Cs×Ld, Mr×Cl, Mh×Cr, Sv×Cc
IWB6204	BS00010318_51	KUBO-006*	19.0	24,308,711	Mr×Cl, Sm×Lv, Mh×Cr, Sv×Cc, Sv×Zv
IWB10512	BS00070900_51	KUBO-008*	19.0	24,313,744	Mr×Cl, Sm×Lv, Mh×Cr, Sv×Cc, Sv×Zv
IWB8390	BS00045163_51	KUBO-010	19.3	24,798,904	Lt×MG_5323, Cs×Ld, Sm×Lv, Mh×Cr

cM, (=centiMorgans) are reported as cumulative genetic distances on chromosome 2B according to the consensus map from Maccaferri et al. (2015)

^aGenetic map distances estimated based on the tetraploid wheat consensus map published in Maccaferri et al. (2015)

^bPhysical order on chromosome 2BS from distal (5') to proximal (3') side based on the Svevo genome RefSeq genome assembly published in Maccaferri et al. (2019)

^cLt=Latino, MG_5323=MG5323, Cs=Colosseo, Ld=Lloyd, Mr=Meridiano, Cl=Claudio, Mh=Mohawk, Cr=Cocorit69, Sv=Svevo, Cc=Ciccio, Zv=Zavitan, Sm=Simeto, Lv=Levante, MC=Molise Colli. The corresponding genetic maps have been reported in Maccaferri et al. (2015)

*: codominant KASP® marker

Genome-wide association study on the UNIBO Durum Panel

The UNIBO Durum Panel was phenotyped in a highly infected field nursery located in Cadriano, Bologna, Italy (44°35'N 11°27'E). The propagule soil content in the field nursery was progressively increased, maintained high, and evenly spread in the nursery soil by repeatedly adopting a wheat-to-wheat rotation with the highly susceptible durum cv. *Grazia*. This ensured spreading of virulent *P. graminis* spores as clusters of resting spores in root tissue debris or infective propagules directly released in soil and evenly distributed throughout the field.

Phenotypic data were obtained from unreplicated, augmented design field trials over three years (2005, 2007, 2010). Cvs. Svevo, Ciccio, Meridiano, and Claudio were used as common checks across blocks (five blocks per experimental field). Broad-sense heritability was calculated with the R package *heritability* (Kruijer et al. 2015).

Recorded data consisted of visual symptom severity/scores (SEV) in 2005 (single date at end of tillering/first node stage, according to Zadoks et al. (1974), 2007 (four dates), and 2010 (single date) and ELISA values (ELISA, indicating virus concentration in the leaves) in 2007 (two dates) and 2010 (single date). SEV was recorded according to a 0–5 scale modified from Vallega and Rubies Autonell (1985) where: 0 = no symptoms, 0.1–1.5 = 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. ELISA assay was performed as in Maccaferri et al. (2012).

Single-plot data were spatially adjusted according to a moving average based on the double adjacent plots. One representative score per year was chosen for further analysis based on correlation among dates, average infection level, and cultivar discrimination power. Dates considered were April 22, 2005, March 6, 2007, and April 7, 2010, for SEV, March 14, 2007, and April 7, 2010, for ELISA. Phenotyping was carried out only for years with nurseries showing an infection level of SEV ≥ 3.00 in susceptible parents and checks throughout the field nursery. Overall infection levels were also concomitantly monitored by ELISA (ELISA ≥ 0.70). Consequently, 2006, 2008, and 2009 nurseries were not considered for further analysis.

Adjusted visual scores and ELISA values were used to calculate best linear unbiased estimates (BLUEs), based on a mixed linear model including genotypes as fixed and all other factors as random (Bates et al. 2014).

GWAS was performed in Tassel 5 (Bradbury et al. 2007) based on genotypic data from Illumina 90 K SNPs, newly developed KASP® markers, *wPt-2106* and visual score and ELISA adjusted BLUEs. Genotypic data were converted to *hapmap* format, keeping molecular markers with MAF $\geq 5\%$, missing data $\leq 50\%$ and by removing genotypes with missing

data $\geq 50\%$. Missing SNPs were subsequently imputed with Beagle v5.4 (Browning and Browning 2022).

A kinship matrix was calculated with all non-redundant SNP markers identified in Haploview 4.2 (Gabriel et al. 2002) with the tagger function set to $r^2 = 1.0$. Kinship based on identity-by-state (IBS) similarities among accessions was calculated in TASSEL. Each marker was evaluated for association with phenotypic data using a mixed linear model (MLM; Bradbury et al. 2007) which included the kinship matrix (MLM + K). A second round of analysis was carried out using the marker most associated with *Sbm2* region as covariate to highlight additional QTL peaks other than the major one on chromosome 2B. *IWB6584* was chosen as the *Sbm2* tagging marker to be used as covariate.

Fine mapping of the *Sbm2* region in bi-parental recombinant inbred lines

Aiming to detect new informative recombinant lines in the *Sbm2* confidence region, the complete Svevo \times Ciccio RIL set (Sv \times Cc RILs, 3,075 lines) was genotyped with the two KASP® markers flanking the *QSbm.ubo-2BS* region, *KUBO-13* and *KUBO-9*, located at the distal (5') and proximal (3') sides of the QTL, respectively. The Sv \times Cc RILs selected through MAS were sown in two seasons (2015/2016 and 2016/2017) in a field experiment aimed at assessing their SBCMV response in Cadriano (Bologna, Italy, 44°35'N 11°27'E). Further genotyping was performed with seven markers located within the *KUBO-13/KUBO-9* interval (i.e., *KUBO-1*, *KUBO-3*, *KUBO-27*, *KUBO-29*, *KUBO-40*, *KUBO-41*, and *wPt-2106*, reported based on their distal-to-proximal chromosomal order) with the aim of precisely tracing the recombination events.

An additional set of 30 RILs from the Meridiano \times Claudio RIL population (*Mr* \times *Cl*, Maccaferri et al. 2011a) with recombinations between *KUBO-3* and *KUBO-9* were selected for detailed *KUBO* genotyping. For each population, genotypic data were used to construct a genetic map was assembled using JoinMap v. 4 (Van Ooijen 2006). Marker order and mapping distances between markers were calculated using the maximum likelihood algorithm and Haldane's mapping function. The two maps were compared to check for the order of KASP® markers in the region flanked by *KUBO-13* and *KUBO-9*. Once marker order was determined, genotypic and phenotypic data were integrated to define the restricted interval where *QSbm.ubo-2BS* was most probably located. Mapping and association analysis were carried out in Windows QTL Cartographer v. 2.5 by single-marker analysis (Wang et al. 2012). Interval mapping (IM) analysis was also carried out to calculate the LOD score in the interval between *KUBO-13* and *KUBO-9*. The threshold value was set at LOD 3.0 (rounded) via permutations. Walk speed was set at 0.5 cM.

Comparison of marker order in the *Sbm2* region among genomes

The marker order between *IWB73347* (= *KUBO-13*, position: 9,986,862 bp) and *IWB10512* (= *KUBO-9* position: 24,313,794 bp) was compared between Svevo, Chinese Spring (CS, IWGSC RefSeq V2.1; Zhu et al. 2021) and

Zavitan (WEW_v2.0; Zhu et al. 2019) (Figs. 1, 2, S5, S6, and S7).

Marker physical positions were determined by BLAST analysis while co-linearity among intervals was compared using *MUMmer4* (Marçais et al. 2018) and manual curation as described in Method S3.

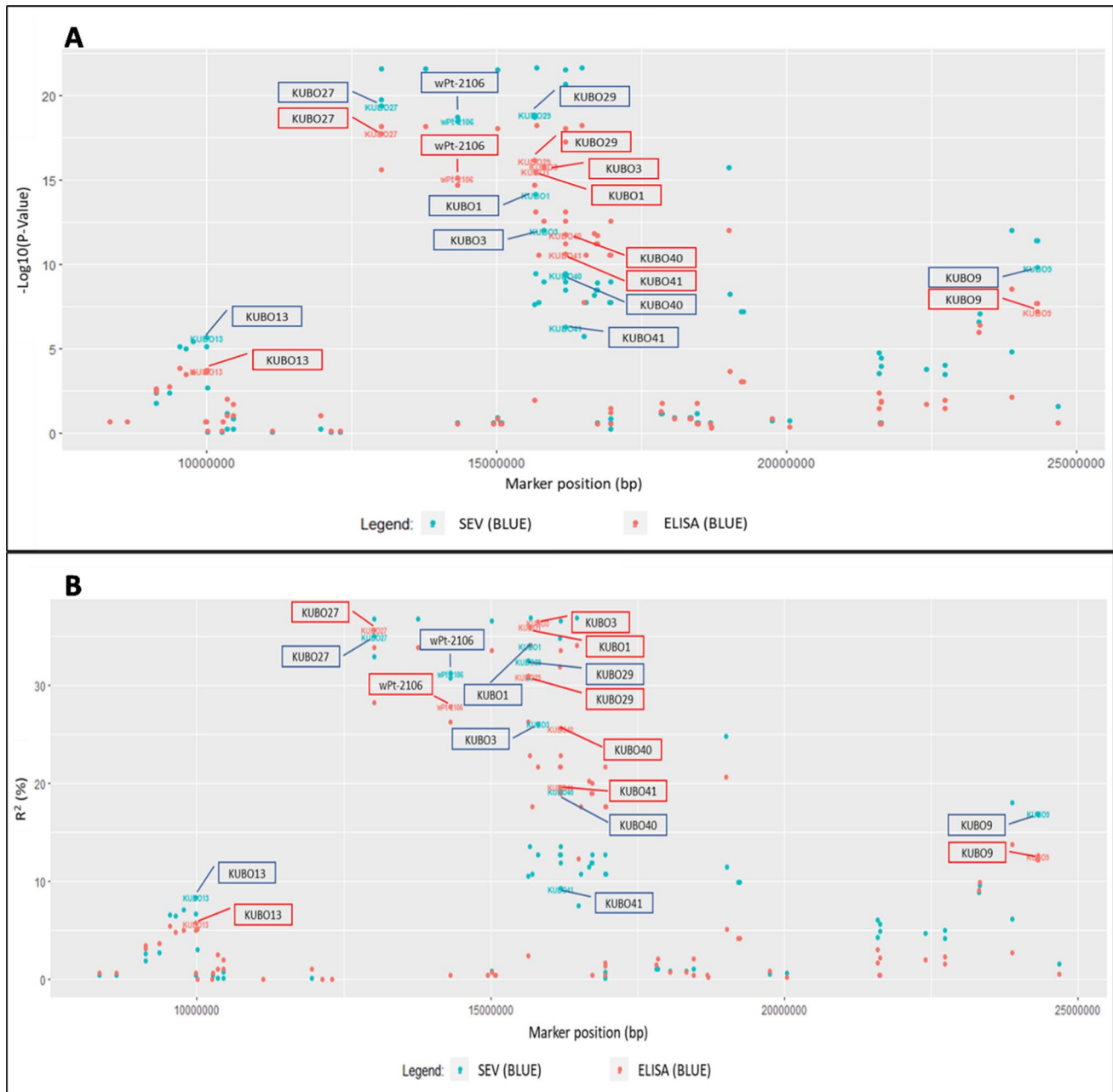


Fig. 1 SBCMV response GWAS association results for markers in the *Sbm2* region on chromosome 2BS. Minus $\log_{10}(P)$ panel A) and R^2 (% panel B) GWAS analysis results based on the UNIBO Durum Panel performed with visual score (SEV, shown in dark blue) and ELISA (shown in pale red) BLUEs of field data collected in 2005

(SEV), 2007 (SEV and ELISA), and 2010 (SEV and ELISA) in Cadriano (Bologna, Italy). Marker positions are from the Svevo reference genome assembly (Maccaferri et al. 2019). Newly developed KASP® *KUBO* is shown in the picture

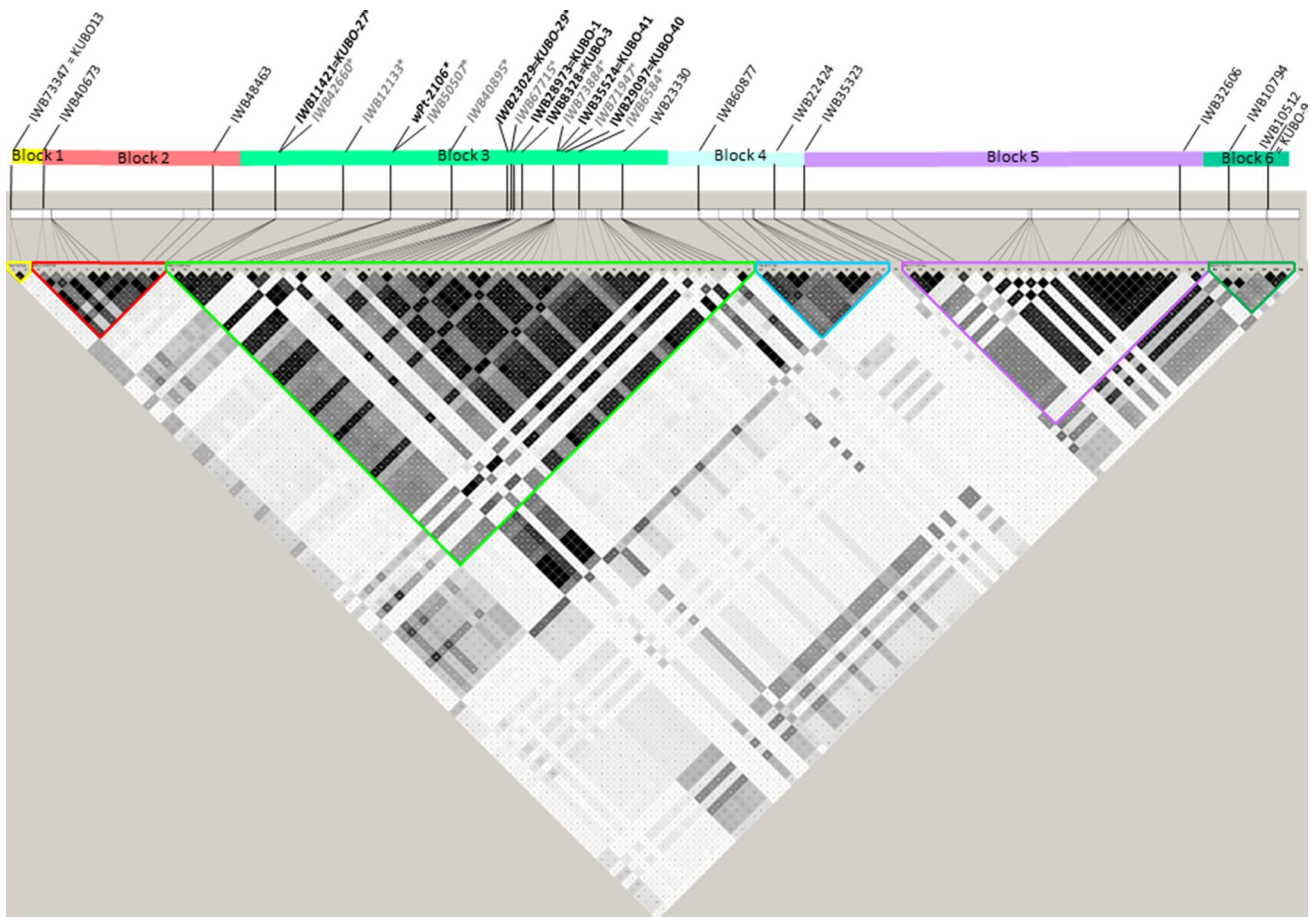


Fig. 2 Local linkage disequilibrium (LD) matrix based on pairwise r^2 values in the *KUBO13-KUBO9* chromosome 2BS interval harboring *Sbm2*. The r^2 matrix has been generated in *Haploview* (Barrett

et al. 2005). Color-labeled bars represent six adjacent linkage disequilibrium blocks (*hapblocks*) defined based on Gabriel et al. (2002) method

Projection of the *Sbm2* support interval into the Svevo physical map for candidate gene identification

The *KUBO-27* to *KUBO-40* interval (Figs. 1, 2, S5, S6, and S7) was considered to retrieve annotated genes from the Svevo genome reference assembly (Maccaferri et al. 2019). Gene functions were recovered using the functional annotation program *AHRD*, considering the longer transcript as the representative one. GO terms (GO Slim) in *R Bioconductor* (version 3.4) package *GOstats* (Falcon and Gentleman 2007; R Core Team 2017) were used to categorize genes according to their molecular functions and biological processes. Gene intervals were explored with *gProfiler* software (Raudvere et al. 2019). Functional profiling was computed extracting GO molecular functions and biological processes from all the genes of the interval.

Sbm2 haplotypes present in the UNIBO Durum Panel

Based on GWAS results obtained from the Durum Panel UNIBO, the region most associated with *Sbm2* was identified as coincident with the haplotype block containing markers *KUBO-27*, *KUBO-29*, *KUBO-1*, *KUBO-3*, *KUBO-41*, and *KUBO-40* (listed from distal to proximal region of chromosome 2BS, Figs. 1, 2, S5, S6, and S7). UNIBO Durum Panel SNP genotype calls in the *Sbm2* region in the *KUBO-13* to *KUBO-9* interval. This interval was used to inspect the local *Sbm2* specific SNP haplotype distribution patterns in cultivars/breeding lines based graphically on a two-color mode (Fig. 3). In the *Sbm2* region, the SNP alleles belonging to the resistant parents/founders Levante, Meridiano, and Neodur (showing identity by descent, IBD, in the region) were dark blue colored while the alternative alleles were pale red colored. Additionally, the local *Sbm2* haplotypes depicted

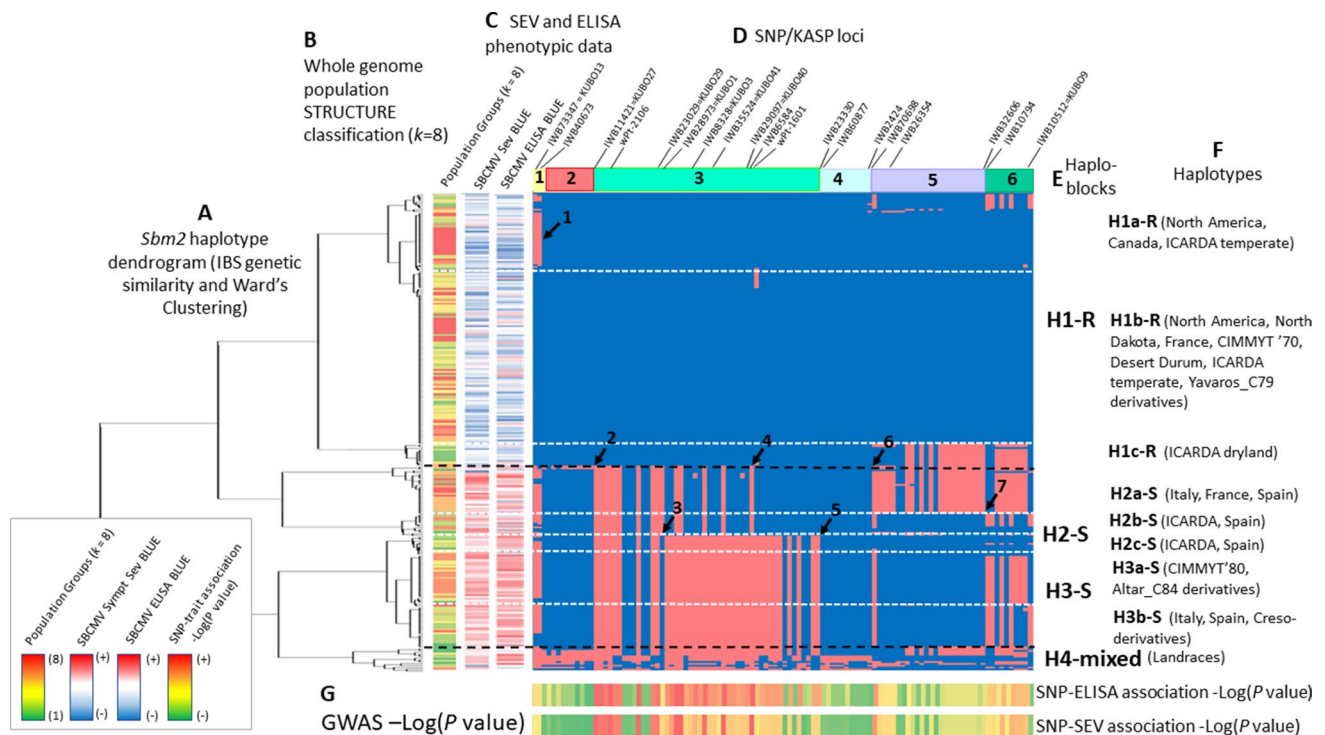


Fig. 3 UNIBO Durum Panel local haplotypes in the *KUBO13-KUBO9* chromosome 2BS interval harboring *Sbm2*. Haplotype-based UPGMA identity-by-state IBS genetic similarity was used to group durum panel genotypes and landraces (panel **A**). Each genotype was associated with population structure Q file at $Q=8$ (**B**) and with SEV and ELISA phenotypic data (**C**). Critical SNP and KUBO KASP® markers are shown in (**D**). Haploblocks detected in Haploview are shown in (**E**) according to Fig. 5. Nucleotide-based haplotypes were

represented based on a two-color scheme in panel (**F**). The SNP allele from the resistant parents (Meridiano, Levante, Neodur) was considered as the main leading allele and was colored dark blue while the alternative SNP allele was colored pale red. GWAS minus \log_{10} (P -value) of SNP-ELISA and SNP-SEV are shown as heat bars in panel (**G**). Critical recombination events are reported with black arrows and numbered from 1 to 7

in Fig. 3 were clustered based on local genetic similarities (simple matching genetic similarity algorithm, GS, also corresponding to identity by state, IBS) and Ward's clustering algorithm. Simple matching GS were computed in Tassel 5 and Ward clustering in *heatmap* R package as in Kolde and Kolde (2015). Linkage disequilibrium among local SNPs and linkage blocks were determined using Haploview (Barrett et al. 2005), defining haploblocks by the algorithm developed by Gabriel et al. (2002). Box plots of SEV and ELISA phenotypes were produced for each haplotype detected at this block.

Results

KASP® marker design and validation

Fine mapping of loci of breeding interest and downstream diagnostic (haplotype-based) marker-assisted selection require the development of locally dense molecular marker maps to saturate the confidence interval regions. Therefore, in the case of *Sbm2*, sequences of 16 Illumina Infinium 90 K

wheat SNPs mapped in the *Sbm2* region by multiple tetraploid mapping populations (Maccaferri et al. 2015) were deployed to develop high-throughput, fluorescence-based KASP® assays, subsequently used to genotype both RILs and breeding germplasm panels for fine mapping.

Nineteen out of the 35 tested assays (*KUBO-1*, *KUBO-3*, *KUBO-5*, *KUBO-6*, *KUBO-8*, *KUBO-9*, *KUBO-12*, *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*, reported based on map order, from the distal 5'-end to the proximal 3'-end region of the QTL on chromosome arm 2BS) proved to be functional, highly repeatable, and codominant. These are ideal for genotyping, resulting in a 54.2% SNP-to-KASP® conversion rate (Illumina SNP and KASP® sequences in Table 1 and Table S1a, respectively).

KUBO-13 and *KUBO-9*, positioned at the two borders of *Sbm2* region 2.5 cM upstream and 8.1 cM downstream of *wPt-2106*, respectively, were selected to identify recombinants in the Sv × Cc RILs (Table 1 and Fig. S5).

The DAiT® marker *wPt-2106*, close to the *Sbm2* peak LOD region, was sequenced and converted to two PCR-based assays, one based on allele-specific (ASO) primers

and detectable through horizontal gel electrophoresis assay, and the other based on the high-throughput high-resolution melting technique (Table S1b). Detailed information is reported in Method S1 while Fig. S1 shows PCR assay results for *wPt-2106* DArT marker converted to *wPt-2106* Allele-Specific Oligonucleotide (ASO), *wPt-2106* High-Resolution Melting (HRM, Wittwer et al. 2003).

The newly developed KASP® markers were used to genotype the UNIBO Durum Panel and on RIL populations and validated by comparison with the former Illumina 90 K SNP genotypic calls. KASP® diagnostic marker accuracy results are reported in Table S2 while their amplification plots are shown in Fig. S2.

GWAS based on the UNIBO Durum Panel

The UNIBO Durum Panel was used as an independent resource to fine map *Sbm2* and to validate the newly developed KASP® markers (genotyping information reported in Method S2). The UNIBO Durum Panel was field evaluated across three years in a dedicated nursery in Cadriano, Bologna, Italy for both visual symptom severity (SEV) and virus content in the leaves (ELISA), as reported in Table S3. SEV

and ELISA traits were highly related (Figure S3), with an *r* value of 0.69.

The UNIBO Durum Panel showed a population genetic structure including eight well-defined groups of varieties/breeding lines corresponding to the main breeding programs/pedigree worldwide, originated from well-known “founders” (Fig. 4a, Data_S1).

SBCMV response (both SEV and ELISA) appeared to be highly differentiated according to population structure, as shown in the box-plot phenotypic distributions based on population structure reported in Fig. 4b.

Germplasm of worldwide success as CIMMYT’60 (Jori C69 founder), CIMMYT’70 (Yavaros C79 founder), CIMMYT’80 (Altar C84 founder), the Italian (Valnova founder), and the North American and French germplasm (cvs. Kyle, Langdon, and Neodur, Vic, founders) are represented in the panel. CIMMYT’60, CIMMYT’70, and the North American/French groups showed a higher frequency of resistance as compared to other groups like the ICARDA temperate, while Italian varieties originated from Valnova and the CIMMYT’80 varieties originated from Altar C84 were all highly susceptible (Fig. 3). Additionally, resistance was observed in high frequency in old varieties derived from West Asian landraces (Haurani/Omrabi group, Fig. 3).

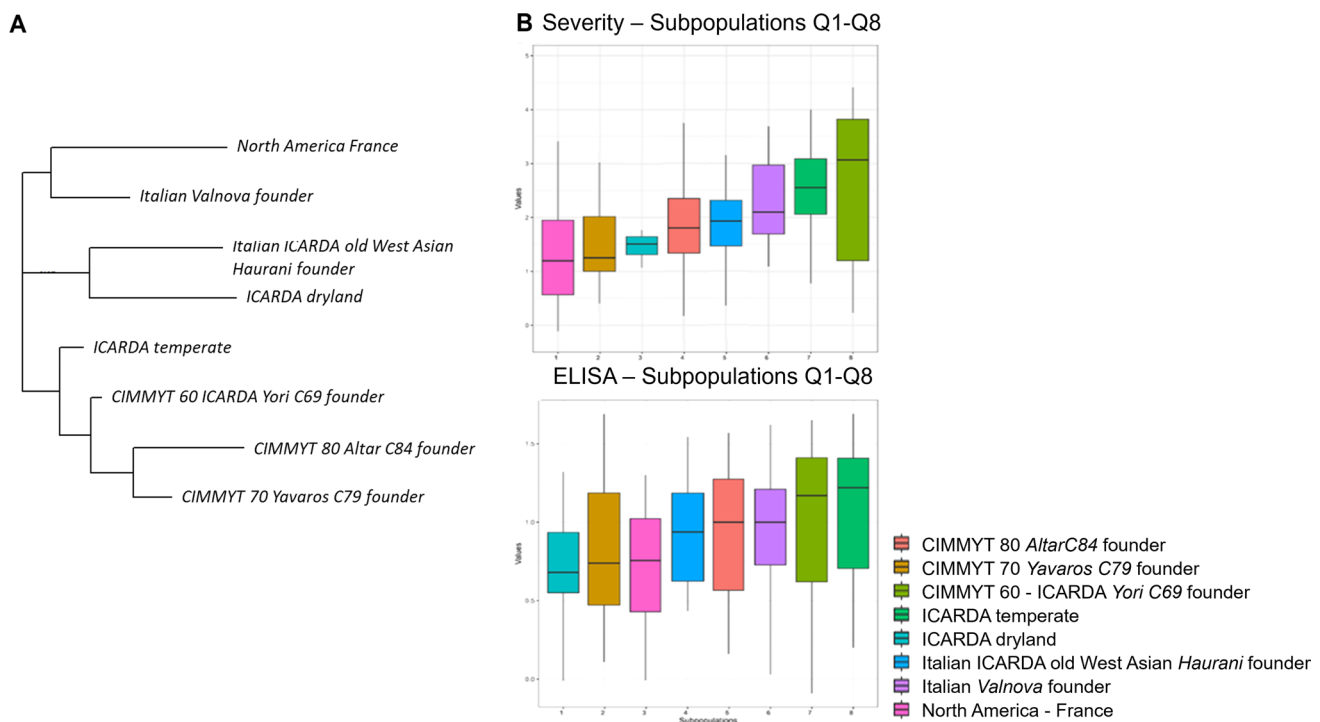


Fig. 4 Genetic population structure and SBCMV response phenotypic distributions of varieties and breeding lines included in the UNIBO Durum Panel. **A** neighbor joining tree of the main eight subpopulation groups detected in the UNIBO Durum Panel based on Illumina 90 K SNP. **B** Box-plot phenotypic distributions of UNIBO

Durum Panel ELISA and SEV BLUES across years according to the eight subpopulation groups. The molecular-based genetic similarity matrix (=kinship matrix) was obtained using the identity-by-state (IBS) genetic similarity algorithm

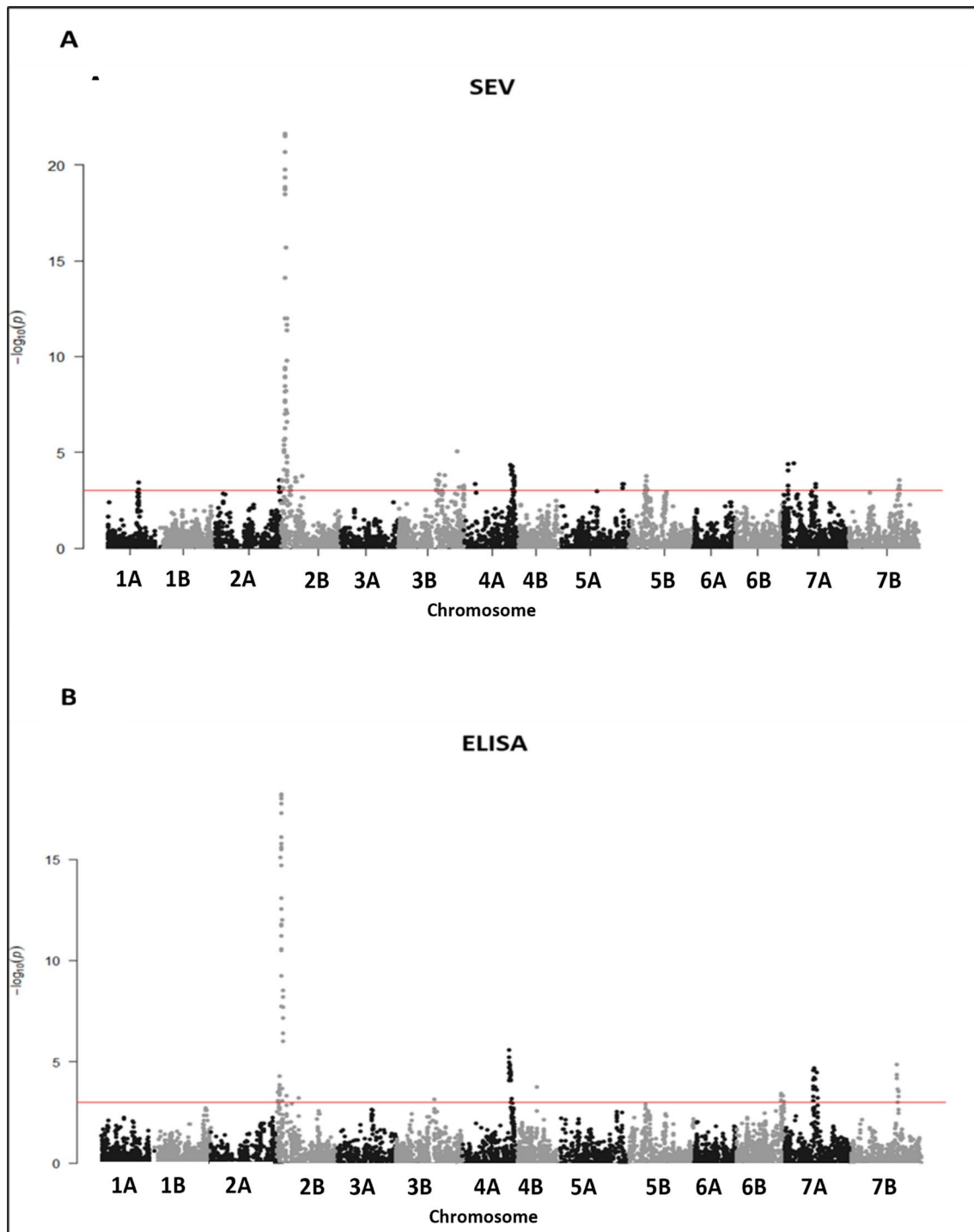


Fig. 5 Manhattan plots of GWAS association analysis results for SBCMV response performed on the UNIBO Durum Panel. **A** GWAS for symptom severity scores (SEV; years 2005, 2007 and 2010). **B**

GWAS for ELISA (years 2007 and 2010) data collected in Cadriano (Bologna, Italy). Red line represents the significance threshold set at $-\log_{10}(P\text{-value})=3$

Population structure accounted for 21.5 and 13.3% of phenotypic variation for SEV and ELISA, respectively, indicating that SBCMV-resistant and SBCMV-susceptible alleles at *Sbm2* were unevenly partitioned among

subpopulations/breeding groups. Figure 4b shows the NJ tree of the eight main subpopulations/breeding groups and the corresponding pattern of phenotypic distribution of SBCMV responses. Detailed population structure Q

membership values together with phenotypic and genotypic data are reported in supplemental Data_S1 and Data_S2.

GWAS showed a major peak locus on chromosome 2B, corresponding to *Q_{Sbm.ubo-2BS}=Sbm2* (Fig. 5). The LOD of this peak was 21.6 for visual score (SEV) and 18.2 for ELISA; R^2 (=PVE) was 36.9% for visual score (SEV) and 34.1% for ELISA, respectively (Fig. 1). The three peak markers (*IWB72375*, *IWB67715*, and *IWB6584*) from the two SEV and ELISA phenotypes were coincident. Detailed marker–trait association results for the *Sbm2* region are reported in supplementary material section (Data_S2). Notably, the association results for SEV and ELISA were highly consistent.

The region corresponding to the peak of significance on Svevo genome extended from *IWB11421* (*KUBO-27*) at 13,011,332 Mb to *IWB6584* at 16,464,631 Mb, very close to the previously mapped *wPt-1601* at 16,497,383 Mb (Fig. 1 and Table 2). In this interval of ca. 3.5 Mb, 14 out of 35 markers showed very high LOD values, clearly identifying the *Sbm2* hotspot region. The entire region is tagged by the series *KUBO-27*, *wPt-2106*, *KUBO-29*, *KUBO-3*, *KUBO-41*, and *KUBO-40* (reported from distal -5' to proximal -3' 2BS chromosomal region, details of association results reported in Table 3), thus identifying the corresponding haplotypes highlighted in Fig. 3. Among KASP® markers (Table 3), the most closely associated is *KUBO-27*. Overall, $-\log_{10}(P)$ ranged from 19.36 for *KUBO-27* from

Table 2 Peak markers in *Sbm2* chromosome region resulted from the GWAS analysis on the UNIBO Durum Panel

Marker	Position (cM) ^a	Position (bp) ^b	Strand (\pm) ^c	$-\log_{10}(P)$	R^2 (%)	Allele nucleotide	Allele frequency	Trait ^d
IWB72375	12.3	15,686,394	+	21.63	36.9	A/G	0.42/0.58	SEV_BLUEs
IWB67715	12.3	15,686,970	–	21.63	36.9	A/G	0.58/0.42	SEV_BLUEs
IWB6584	12.3	16,464,631	–	21.63	36.9	A/G	0.58/0.42	SEV_BLUEs
IWB72375	12.3	15,686,394	+	18.22	34.1	A/G	0.42/0.58	ELISA_BLUEs
IWB67715	12.3	15,686,970	–	18.22	34.1	A/G	0.58/0.42	ELISA_BLUEs
IWB6584	12.3	16,464,631	–	18.22	34.1	A/G	0.58/0.42	ELISA_BLUEs

^aGenetic map positions (cM) are reported as in the consensus map of Maccaferri et al. (2015)

^bPhysical map positions (bp) are reported as the start positions obtained by blasting the marker sequence on the durum wheat genome reference Svevo (Maccaferri et al. 2019)

^cStrand defines based on the direction of the marker sequence

^dBest linear unbiased estimates (BLUEs) were calculated based on LUEs of A) symptom severity scores (SEV) data from 2005, 2007, and 2010 (SEV_BLUEs) and of ELISA data collected in 2007 and 2010 (ELISA_BLUEs) in Cadriano (Bologna, Italy), spatially corrected

Table 3 Association results for KUBO KASP® markers and *wPt-2106* in GWAS performed on the UNIBO Durum Panel

Marker	Position (cM) ^a	Position (bp) ^{bs}	Strand (\pm) ^c	Resistant allele	Susceptible allele	Allele frequency	SEV ^d		ELISA ^d	
							$-\log_{10}(P)$	R^2 (%)	$-\log_{10}(P)$	R^2 (%)
KUBO-13	8.4	9,986,862	+	A	G	0.45/0.45	5.65	8.3	3.72	5.70
KUBO-27	11.6	13,011,332	+	A	G	0.56/0.39	19.36	35	17.76	35.6
wPt-2106	10.9	14,314,449	–	I	O	0.38/0.59	18.48	31.2	15.1	27.8
KUBO-29	11.6	15,643,691	+	G	T	0.57/0.38	18.84	32.4	16.12	30.9
KUBO-1	12.2	15,672,582	–	T	C	0.56/0.26	14.11	34	15.49	35.9
KUBO-3	12.3	15,805,908	+	T	C	0.26/0.57	11.99	26	15.78	36.4
KUBO-41	12.3	16,184,452	–	C	T	0.68/0.25	6.27	9.3	10.58	19.6
KUBO-40	12.3	16,185,445	+	C	T	0.58/0.25	9.35	19.2	11.75	25.6
KUBO-9	19.0	24,313,744	+	G	A	0.26/0.66	9.8	16.9	7.17	12.6

^aGenetic positions (cM) are reported as in the consensus map of Maccaferri et al. (2015)

^bPhysical positions (bp) are reported as the start positions obtained by blasting the marker sequence on the durum wheat genome reference Svevo (Maccaferri et al. 2019)

^cStrand defines the direction of the marker sequence

^dBest linear unbiased estimates (BLUEs) were calculated based on symptom severity scores (SEV) data from 2005, 2007 and 2010 (SEV_BLUEs) and ELISA data collected in 2007 and 2010 (ELISA_BLUEs) in Cadriano (Bologna, Italy), spatially corrected

visual score (SEV) to 3.72 for *KUBO-13* as from ELISA test. The highest PVE was obtained for *KUBO-3* (36.4%, ELISA, Tables 2 and 3), while the lowest was obtained for *KUBO-13* (5.7%, ELISA; Tables 2 and 3) *Sbm2* allelic distribution appeared to be highly related to population structure, as all the founders of groups associated with resistance including ICARDA dryland, CIMMYT'60, CIMMYT'70, and North American / French varieties carried the *Sbm2*-resistant allele, while founders of ICARDA temperate, Italian, and CIMMYT'80 groups carried the *Sbm2*-susceptible allele (Data_S2).

Several minor putative QTLs across the genome resulted from the GWAS (Table S4). Up to 20 putative minor QTLs at $-\log_{10}(P) > 3$ in GWAS were mapped, eight of which were common to both SEV and ELISA (Tables S4a and S4b, for SEV and ELISA, respectively).

Minor QTLs common to both SEV and ELISA were found on chromosomes 2A (*IWB71381*, 206.5 cM), 2B (two QTLs, at *IWB6584*, 12.3 cM, and *IWB48240*, 67.6 cM), 4A (two QTLs, at *IWB6276*, 147.4 cM and *IWB37657*, 156.9 cM), 5B (*IWB67284*, 53.8 cM), 6B (*wpt-9589*, 153.8 cM), 7A (*IWB72581*, 102.4 cM), and 7B (*IWB71789*, 146.2 cM).

After completing the GWAS, the *Q_{Sbm.ubo-2BS}*=*Sbm2*, the additional minor QTLs, and population structure *Q* values were modeled in GLM (Table S5). Major and minor QTL and population structure effects on SEV and ELISA were tested based on a multi-marker general linear model (*MM-GLM* + *Q*). The *p*-values and adjusted *R*² of the full *Sbm2* and minor QTL models are reported in Table S5. The population structure *Q* memberships accounted for 21.5 and 13.3% of SEV and ELISA blues values, respectively, a finding mainly explained by the allele distribution at *Sbm2*, whose resistant and susceptible alleles were differently distributed according to main population structure partitioning. Therefore, we modeled the effect of *Sbm2* together with the additional minor QTLs suggested by GWAS in a GLM excluding population structure as covariate. Notably, *Sbm2* accounted for 54.6 and 46.9% of the total phenotypic variation for SEV and ELISA, respectively. Minor QTLs were mostly confirmed by GLM for either SEV or ELISA, or both, except for *IWB48240_2B* and *IWB26062_4A*, which were not validated by the GLM. *Sbm2* × other QTL interactions were mostly non-significant. Importantly, the addition of minor QTL effects and their interactions raised the total variation explained by the global genetic effects up to 67.6 and 54.7% for SEV and ELISA, respectively (Table S5).

The distribution of resistant and susceptible alleles at *Sbm2* and minor QTLs in the UNIBO durum panel is detailed in Data_S3. Resistant alleles at minor QTLs were distributed across multiple varieties, from a minimum of four to a maximum of 18 alleles per variety. The linear model plot reports the cumulative number of resistant alleles across different genotypes in the X-axis, based on

the tag-QTL marker scores, whereas the Y-axis reports the BLUES of SEV and ELISA data. Based on this data, linear regression was computed to estimate the overall effect of minor QTL combinations on the phenotypic data (Data_S3).

The global effect of the susceptible-to-resistant ideal progressive substitution of alleles in germplasm is exemplified by the histograms reported in Data_S3, reported separately for varieties carrying either the resistant or susceptible allele at *Sbm2*.

In the 148 *Sbm2* resistant varieties, the cumulated substitution effect of minor QTLs was estimated (based on the regression reported in Data_S3) to improve SEV only marginally from 1.63 (four minor resistant alleles) to 0.92. However, the effect was more marked on the ELISA value that decreased from 0.92 (4 minor alleles) to 0.40 (18 alleles).

Minor QTL allelic substitution effect impacted more evidently on the group of 110 *Sbm2* susceptible varieties, as expected. In this group, the SEV response of varieties decreased from 3.81 (varieties cumulating four minor resistant alleles only) to 1.82 (varieties cumulating 16 cumulated minor resistant alleles) and, for ELISA, the average value decreased from 1.45 (four resistant alleles) to 0.94 (16 resistant alleles), thus providing a substantial improvement in resistance.

Fine mapping of *Sbm2* based on the Svevo × Ciccio and Meridiano × Claudio RIL populations

Genetic maps constructed for each population are presented in Fig. S4. KASP® order was coincident in the two maps. In the Mr × Cl map, *KUBO-27* was in a recombination bin including *KUBO-29* and *wPt-2106*, while in the Sv × Cc map it was separated by a recombination event. KASP® and *wPt-2106* genotypes of the Sv × Cc RILs are reported in Table S5a. Crossing overs were not evenly distributed in the target interval. Only seven RILs showed recombination inside the target interval and were therefore informative for fine mapping (Table S5b). In total, 433 recombinant Sv × Cc RILs in the *KUBO-13* to *KUBO-9* interval were selected and scored for SBCMV resistance in 2016 and 2017. BLUES frequency distribution obtained from corrected SEV data is reported in Figure S6. Data presented a mean value of 2.58 with extreme values of 0.29 and 4.27. The distribution histogram showed two peaks, one centered on a SEV score of 1.53 (medium resistant) and one on a SEV score of 3.16 (medium susceptible), with heritability value of 0.83.

The joint genotypic and phenotypic analysis of the Sv × Cc recombinant lines allowed exclusion of the two distal regions marked by *KUBO-13* and *KUBO-9* flanking markers. Sixty-five resistant lines showed that the QTL was most probably downstream of *KUBO-13* and 140 susceptible lines showed that it was upstream of *KUBO-9* (Table S6a). Three recombinant susceptible lines (#652, #2077, #3089)

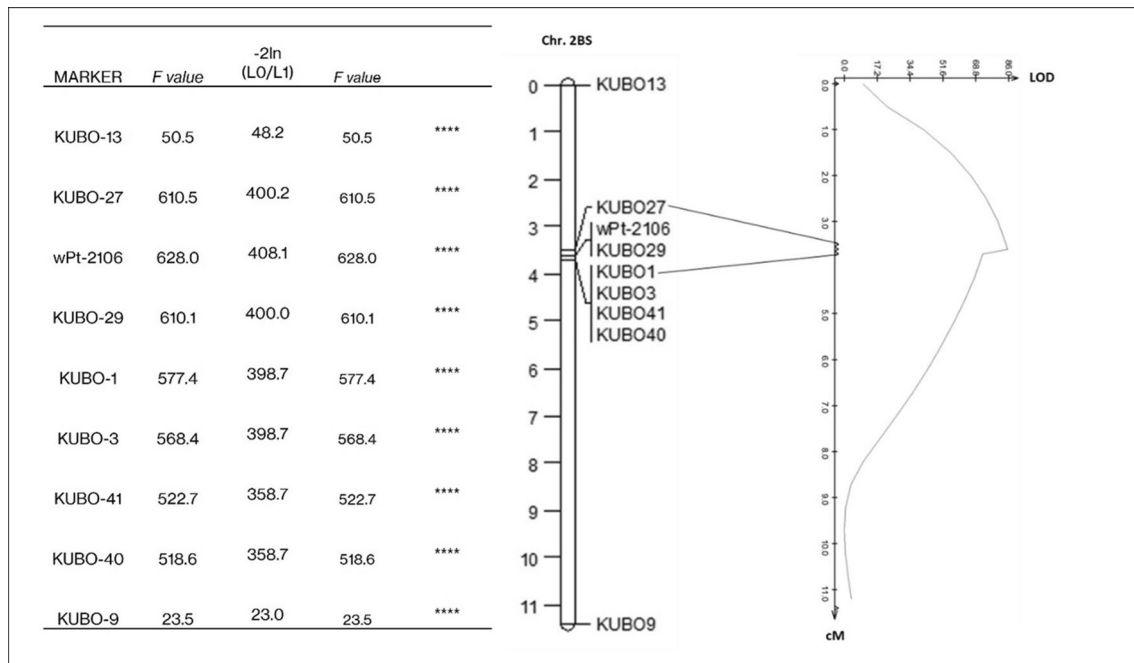


Fig. 6 Single-marker QTL analysis results for *KUBO* KASP® markers and *wPt-2106* mapped in *Sv* × *Cc* population. Interval mapping was based on SEV trait data collected in 2016 (3 dates) and 2017 (5 dates) in Cadriano (Bologna, Italy) for *Sv* × *Cc* RILs

together with three recombinant resistant lines (#1450, #126, #1166) showed that the QTL is most probably upstream of *KUBO-1* located at 13,011,332 bp in Svevo (Table S6a). However, only one additional recombinant line (#139) provided evidence for the QTL to be downstream of *KUBO-27* (Table S6a). Therefore, *Sv* × *Cc* RILs positioned the supporting interval of *Sbm2* in 0.2 cM between *KUBO-27* and *KUBO-1*, corresponding to 2.661 Mb.

As to the *Mr* × *Cl* RIL population, beside the confirmation of results from the *Sv* × *Cc* populations, 18 RILs were informative for fine mapping (Table S6b) and confirmed the interval downstream of *KUBO-13* (supported by five RILs) and upstream of *KUBO-1* (supported by two recombinant lines) as the most probable for harboring *Sbm2*. Single-marker analysis (Fig. 6) showed that *wPt-2106* was the marker most associated with the phenotype. Consistently with GWAS, *KUBO-27* was the most associated with the phenotype among all newly developed KASP® markers. The result of the Interval Mapping analysis is reported in Fig. 6. As expected, the LOD value peaked in the region identified by *KUBO-27*, *wPt-2106*, *KUBO-29*, and *KUBO-1*, with a peak LOD value of 86.

Comparison among genomes for marker and gene order in *Sbm2* region

A comparison of marker order in *Sbm2* region among the three wheat genomes (*cvs.* Zavitan, Svevo, and Chinese

Spring) is shown in Figure S6. Order inconsistency among physical maps of the three genomes was observed for 26 markers, mostly located in a few regions that showed rearrangements among assemblies, as in the case of a group of eight consecutive markers (*IWB71947*, *IWB35524*, *IWB72156*, *IWB62546*, *IWB41644*, *IWB73884*, *IWB29097*, and *IWB72157*) that identified an inversion of ca. 2 kb in Svevo. The three wheat genomes were therefore compared for gene collinearity in the interval *IWB73347* (*KUBO-13*, position: 9,986,862 bp) and *IWB10512* (*KUBO-9*, position: 24,313,794 bp), including *Sbm2*. This analysis further confirmed the local inversion spanning the region tagged by *IWB45885*, *IWB4596*, *IWA7545*, *IWB26871*, *IWB26232*, *IWA3589*, and *IWB26233* (Fig. S7). No inversions were detected between Zavitan and Chinese Spring assemblies, indicating that the local inversions in Svevo genome could be an artifact of assembly. Additional analysis using long-read sequencing data is required to confirm this observation and its potential effects on differential gene expression.

Sbm2 molecular haplotypes in cultivated durum germplasm

With the aim to define the linkage disequilibrium and haplotype patterns in the *Sbm2* region, SNP data from the UNIBO durum panel were subjected to local disequilibrium analysis using the linkage block method implemented in the Haploview software. Considering the extended region between

KUBO-13 and *KUBO-9*, six main linkage blocks of various Mb size were identified (Fig. 2). The target interval on chromosome 2B is comprised between 9,986,912 and 24,313,794 bp. Based on the block partitioning, both fine mapping and GWAS results indicate that *hapblock-3* harbors the QTL hotspot.

SNPs were partitioned into six haplotype blocks (*hapblocks*) as in Fig. 2 and their $-\log_{10}(P)$ of association with SBCMV response are reported as horizontal heat bars in the lower side of the figure (F-G panes).

Detailed local SNP-based haplotypes present in the UNIBO Durum Panel varieties for the *KUBO-13* to *KUBO-9* region are reported in Fig. 3 and in Data_S2. In Fig. 3, varieties were clustered based on the identity-by-state (IBS) genetic distance matrix and Ward's clustering obtained using SNPs from the *Sbm2* region only (Fig. 3, A pane), thus reflecting the genetic relationships among varieties based on *Sbm2* haplotypes only. In Fig. 3, clustered varietal haplotypes are shown side by side with their respective whole-genome population structure Q values (eight Q sub-populations, reported as heat bar), SEV- and ELISA-BLUES values, all reported as vertical heat bars aligned from left to right (Fig. 3, B and C panes). Haplotypes of varieties/breeding lines at *Sbm2* are shown as horizontal bars adopting a two-color scheme where the SNP allele from the resistant Meridiano, Levante, and Neodur parents was considered as the main reference allele and hence is dark blue colored, while the alternative SNP allele, carried by susceptible parents, is pale red colored (Fig. 3, D, E, and F panes).

Considering the region as a whole, two main haplotype groups (H1 and H2) were identified and were clearly associated with resistance (H1-R) and susceptibility (H2-S and H3-S), respectively. Within resistant and susceptible haplotype groups, sub-haplotypes were defined based on haplotype differentiation, termination, and/or recombination events among haplotypes (indicated as black arrows, numbered from 1 to 6). Additional haplotypes were observed in a group of a few landraces only with a much more complex haplotype structure, mostly associated with a susceptible response (bottom side of the figure). R- and S-haplotype groups showed population structure effects, with H1-R being more frequent in North American, French, CIMMYT'70, Desert Durum, and some ICARDA temperate varieties. Conversely, H2-S was more represented in ICARDA dryland, Italian, French, and Spanish varieties derivatives of Valnova/Simeto while H3-S was specifically found in highly susceptible varieties derived from the CIMMYT'80 Altar84 founder.

SNPs differentiating the H1-R group from the two H2-S and H3-S groups can already be identified, from left to right, in *hapblock-1* ($IWB73347=KUBO13$, A/G, indicated by black arrow 1) and *hapblock-2* (including all SNPs of the block, reported in Data_S2). However, these SNPs are highly recombined with the phenotype and have low discrimination power,

indicating that *hapblock-1* and *-2* are distant from the causal sequence. For instance, at $IWB73347=KUBO-13$, the G variant (in pale red) is present in relatively high frequency in both R and S varieties. Similarly, all SNPs included in *hapblock-2* showed variants common to both R and S varieties and were therefore unable to discriminate between resistant and susceptible varieties, while the rare variants discriminated between modern and ancient germplasm only.

Hapblock-3 starting in correspondence of black arrow 2 in Fig. 3 ($IWB11421=KUBO-27$) and ending with SNP $IWB23330$ showed the highest density of SNPs with the highest association to SBCMV response and the highest resolving power for H1-R and H2-S haplotypes ($IWB11421=KUBO-27$, $IWB42660$, $IWB12133$, $wPt-2106$, $IWB50507$, $IWB40895$, $IWB23029=KUBO-29$, $IWB72375$, $IWB67715$, $IWB73884$, $IWB71947$, and $IWB6584$), three of which were transformed into breeder friendly markers suitable for marker-assisted selection (*KUBO-27*, *wPt-2106*, and *KUBO-29*).

Within *hapblock-3*, a main H1-R haplotype associated with resistance and two haplotypes associated with susceptibility, H2-S and H3-S, can be clearly distinguished. Interestingly, they showed a core haplotype region common to both S haplotypes and highly differentiated from H1-R, comprised between $IWB11421=KUBO-27$ (black arrow 2 in Fig. 3) and $IWB28973=KUBO-1$ (black arrow 3). From *KUBO-1* downwards, still in *hapblock-3*, the two S haplotypes start to differentiate as to each other and differentiate from H1-R for a wide segment. Therefore, based on GWAS results, the *KUBO-27* to *KUBO-1* interval is the main candidate to harbor the *Sbm2* causal sequence.

On the proximal side, *hapblock-4*, *-5*, and *-6* all showed SNPs weakly associated with SBCMV, with resolving power considerably lower than those in *hapblock-3*, clearly indicating *hapblock-3* as the most relevant region for *Sbm2*. Interestingly, H1-R haplotype included a group of resistant cultivars with sub-haplotype H1c-R (mainly ICARDA dryland varieties) showing a haplotype clearly recombined with H2-S particularly in *hapblock-5* and *-6*. Overall, these results indicated *hapblock-3*, harboring the QTL hotspot. The phenotypic effect on SEV and ELISA of the three main haplotypes present in *hapblock-3* and tagged by the multi-marker *KUBO-27*, *wPt-2106*, *KUBO-29*, *KUBO-1*, *KUBO-3*, *KUBO-41*, and *KUBO-40* series is reported as box-plot in Fig. 7.

Overall, these observations are consistent with the results obtained from the fine mapping experiment with the biparental populations.

Haplotype distribution in a panel of worldwide durum wheat varieties across decades

To broaden the survey regarding *Sbm2* haplotypes in cultivated durum wheat, a larger panel of 549 varieties was

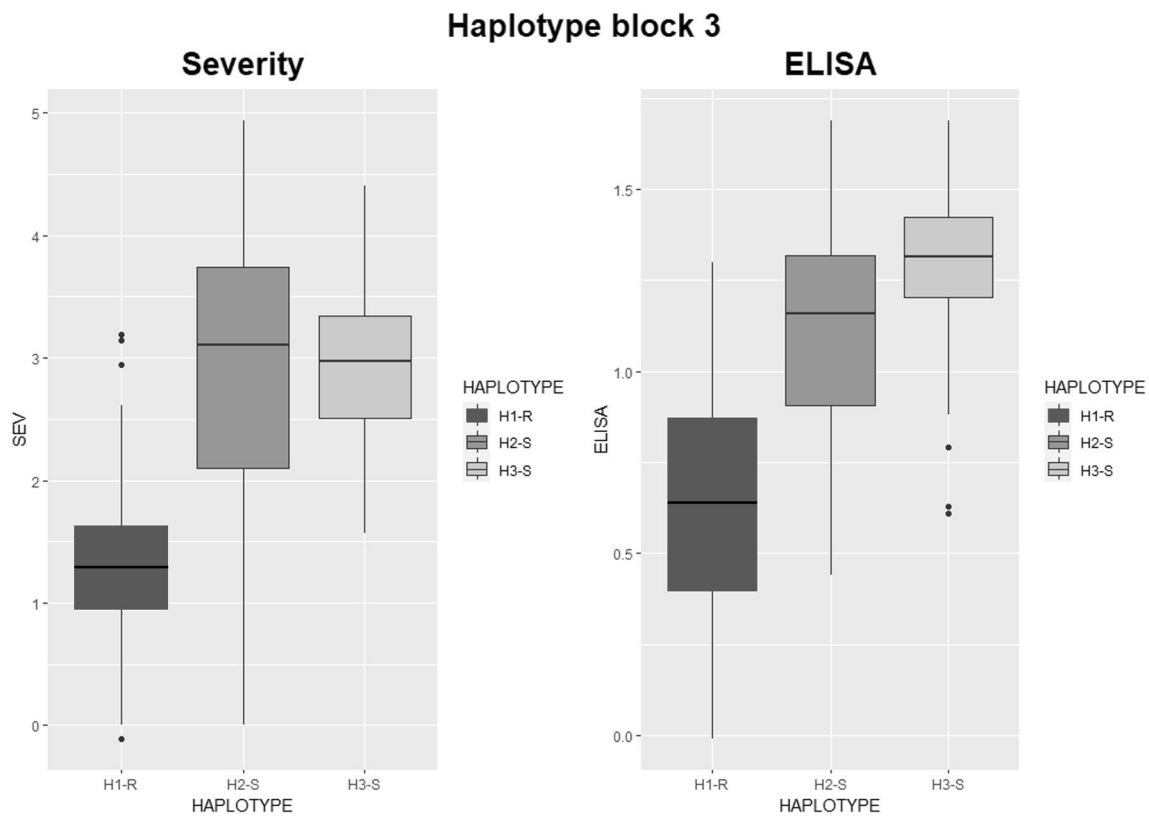


Fig. 7 Distribution of phenotypic values (y-axis) for UNIBO Durum Panel varieties and breeding lines, based on the main three haplotypes detected in *hapblock3* (x-axis) for SEV (panel A) and ELISA (panel B) phenotypic values (BLUES). Phenotypic data were collected in

2005 (SEV), 2007 (SEV and ELISA) and 2010 (SEV and ELISA) in Cadriano (Bologna, Italy) and spatially corrected. Haplotype one is associated with resistance (H1-R) and haplotype two (H2-S) and three (H3-S) are associated with susceptibility

investigated and genotyped with diagnostic markers *KUBO-27*, *KUBO-29*, and *KUBO-1*. Based on the genotyping results, haplotypes were classified in categories: resistant haplotype, susceptible haplotype, recombinant haplotype of type-1, type-2, type-3, double recombinant haplotype, and admixed. The proportion of accessions represented in each haplotype category is reported in Figure S8. The resistant haplotype was the most frequent, being detected in 53% of the accessions, followed by the susceptible haplotype detected in 31% of accessions. The three categories of recombinant haplotype type-1, type-2, and type-3 represented in total 11%, while accessions with a double recombinant haplotype were less than 1%.

The distribution of resistant, susceptible, and recombinant haplotypes based on varietal origin is reported in Fig. 8. For all origins, varieties with the resistant haplotype represented more than 45% of the total, with Portugal–Spain having the minor proportion of resistant accessions (46%). The highest proportion of resistant varieties was observed for the Canada–USA (81% resistant varieties) and the lowest for CIMMYT–Mexico (49% resistant varieties). In Europe, the resistant haplotype was mostly represented in France (51%), even if the proportion was very similar to that observed for

the other European countries, ranging from 46 to 51%. ICARDA materials showed the presence of the resistant haplotype in relatively high frequency (72%). The susceptible haplotype was poorly represented in Canada–USA (14%) and ICARDA (25%), while ranged from 35 to 44% for all the other locations.

Time course haplotype distribution trends were inspected from 1955 to 2020 (Figure S9). Notably, the registration of resistant varieties in North America underwent an outburst in the 1995–2000 interval. Additionally, in Italy and France the release of varieties with the resistant haplotype increased starting from 2000. As to Portugal–Spain, the release of varieties is more evenly distributed throughout the years, with a major release of resistant varieties in the interval 1995–2000. Release of varieties from ICARDA was concentrated in the first half of the considered period, from 1985 to 2000, and most of the novel varieties presented the resistant haplotype (Fig. 8).

Candidate gene analysis at *Sbm2*

The *hapblock-3* hotspot interval between *KUBO-27* and *KUBO-1* up to *KUBO-40/IWB6584/wPt-1601* corresponded

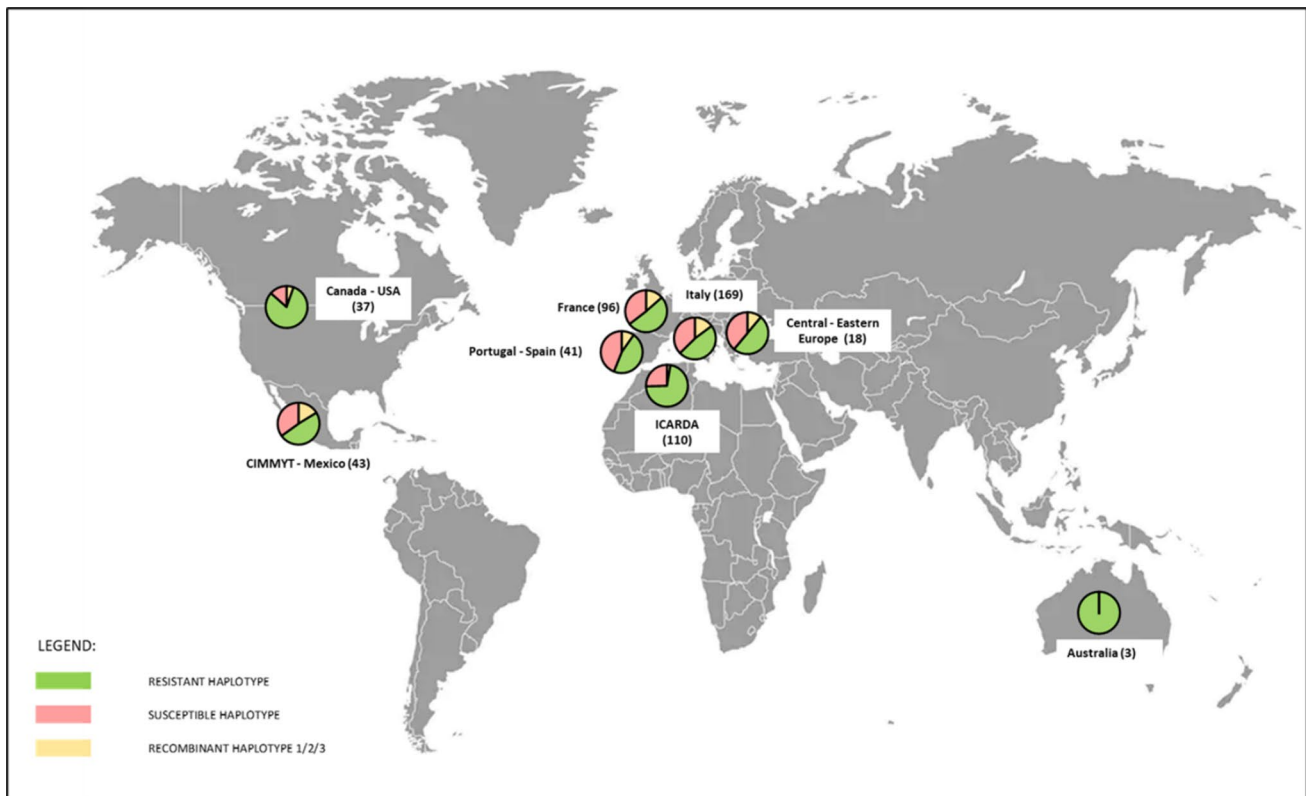


Fig. 8 Proportion of SBCMV-resistant, SBCMV-susceptible, and recombinant haplotypes in durum wheat varieties based on their origin. The number in brackets refers to the total number of accessions.

The haplotype refers to three markers (*KUBO-27*, *KUBO-29* and *KUBO-1*) and includes the *Sbm2* support interval

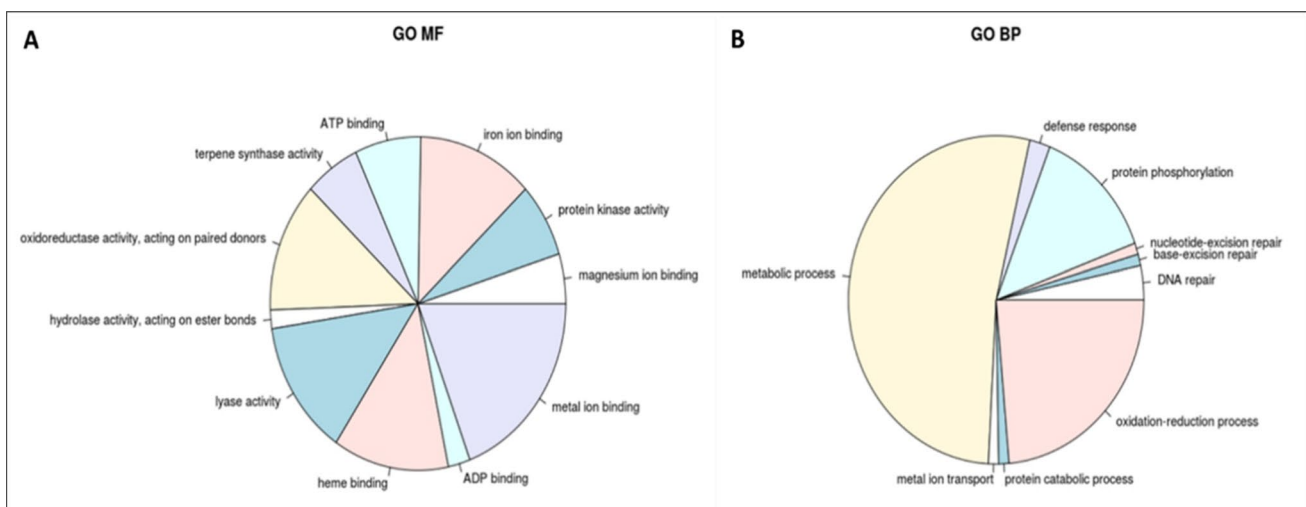


Fig. 9 Molecular functions (A) and biological processes (B) of genes present in the support interval of *Sbm2* based on Gene Ontology

to 3.2 Mb on the Svevo genome (Maccaferri et al. 2019), a region harboring a total of 93 genes (Table S8), of which 52 were high-confidence (HC) and 41 low-confidence (LC) genes. Genes present in this region appeared to be involved mainly in metabolic, oxidation–reduction and protein

phosphorylation processes (Fig. 9). Part of those genes are involved in plant defense response, hence representing candidate genes putatively involved in SBCMV resistance. Such genes belong to defensins, disease resistance protein (NBS-LRR class) family and NBS-LRR-like resistance proteins

(Table S7). In the fine-mapped interval, we observed: 17 cytochrome p450, three receptor kinases, two defensins, one high-confidence and one low-confidence, and three NBS-LRRs in addition to further genes involved in secondary metabolic pathways and transporter activity. Defensins class included an HC gene and a LC gene. As regards to the NBS-LRR genes (HC), two of them encoded for NBS-LRR resistance proteins, while one was considered an NBS-LRR like resistance protein.

Based on the gene ontology enrichment (GO), the molecular functions ascribed to mono-oxygenase activity, terpene biosynthesis, and ion binding are highly represented in the *Sbm2* confidence interval between markers *KUBO-27* and *KUBO-40*, based on the logarithm of adjusted *p*-value (Fig. 9). With regard to the biological processes, general metabolic pathways are highly frequent such as terpenoid and isoprenoid biosynthetic processes and protein phosphorylation (Figure S10), classes known to be involved in the plant–pathogen interaction.

The genes included in the interval between *KUBO-27* and *KUBO-40* were also explored for KEGG terms using *Blast-Koala* bioinformatic tool (Kanehisa et al. 2016). As to the organism system attribution, the main represented pathway is "plant–pathogen interaction," which reports the metabolic pathways involved in the recognition between a host organism and a virulent pathogen. As shown in Fig. S11, the gene *NBS-LRR RPM1 (TRITD2Bv1G007240)* appears to be involved in this pathway as responsible for the hypersensitive response (defense response) as a consequence of the recognition between the NBS-LRR RPM1 product and the pathogen avirulent protein (avr).

Additionally, the gene content in the *Sbm2* confidence interval was compared to the candidate genes identified in the fine-mapped *Sbm1* region on chromosome 5DL of bread wheat by Liu et al. (2020). The *Sbm1* interval was positioned between 546,651,779 bp and 547,273,461 bp, harboring 17 candidate genes from CS RefSeq v1.0 pseudomolecule (IWGSC, 2018). The candidate gene functions belong to Ser/Thr kinase proteins, phosphorylation proteins for signal transduction, membrane transporter binding proteins, transcription factor proteins, and secondary metabolism proteins. Functions such as secondary metabolism, signal transduction by protein phosphorylation, and in particular Ser/Thr proteins and transcription factors are shared with gene functions reported for *Sbm2* (Table S8). However, Liu and colleagues (2020) identified as the most promising candidate a pto-interacting protein 1 (PTI1) (*TraesCS5D01G531200*) based on RNA-Seq experimental data. This specific gene function was not identified among the *Sbm2* candidate genes which, on the other hand, represented other promising candidates. With the aim to find ortholog and/or paralog genes shared between *Sbm1* and *Sbm2* durum and bread wheat intervals, the ENSEMBL plant database

where orthologs and paralogs of each annotated gene are listed based on alignment identity, common domains, and molecular function was surveyed. The comparison between *Sbm2* and bread wheat CS RefSeq v1.0 genome was also performed, retrieving orthologs in CS reference genome and paralogs in Svevo durum wheat genome. The results of the ENSEMBLE search are reported in Data_S4 and Data_S5. As to *Sbm1* 5D genes, no orthologs/paralogs were identified in the *Sbm2* region of chromosome 2B. Most of the *Sbm1* 5D orthologs mapped to 4A and 5B chromosomes, as expected based on the wheat 4AL/5AL translocation (Devos et al. 1995), and not on chromosome 2 group. On the other hand, mapping of *Sbm2* orthologs and paralogs, showed absence of the corresponding orthologs/paralogs in *Sbm1* 5D in Chinese Spring and Svevo genomes.

Discussion

Fine mapping of *Sbm2*

This work was based on the previous knowledge gathered on the genetic basis of SBCMV resistance in durum wheat and was aimed at fine mapping the unique major QTL responsible for this mechanism known so far in durum (Maccaferri et al. 2012). The fine mapping strategy took advantage of the combination of both production of novel recombinants obtained by artificial crossing and use of GWAS in a panel of varieties cultivated worldwide.

Based on these results, the most informative Illumina SNPs mapped in the tetraploid wheat consensus map and in the Svevo genome assembly (Maccaferri et al. 2015 and 2019) were successfully converted to KASP® assays paving the way for haplotype tracing, germplasm characterization, and MAS (Koebner and Summers 2003; Maccaferri et al. 2022). SNP markers proved very effective for MAS when coupled with a high-throughput technology which allows greater flexibility as compared to the original SNP array technology (Shavrukov 2016). The 69% SNP-to-KASP conversion rate allowed for an effective exploitation of SNP markers singled out with the Illumina SNP array. The probability of success in converting an SNP to a KASP® marker, however, varied greatly according to each SNP. For some SNPs, the design of suitable codominant KASP® assays was particularly difficult or impossible, as already reported by Makhoul et al. (2020).

To fine map *Sbm2*, a combination of GWAS and linkage mapping in progenies derived from two different populations, namely the Sv × Cc and the Mr × Cl RILs, was used (Salvi and Tuberosa 2005). Recombinants were identified with two flanking KASP® in a 2.0 cM interval. RIL observations confirmed what was reported by Maccaferri et al. (2011a; 2012) on *Sbm2* as a major locus for SBCMV

resistance in durum wheat. The divergence of the phenotypic response from a normal distribution toward a bimodal distribution supports this hypothesis. However, the absence of a clear bimodality suggests the influence of additional QTLs with minor effects, an expected outcome of QTLome dissection (Salvi and Tuberosa 2015).

The spatially adjusted model applied to phenotypic data analysis effectively allowed for accounting micro-environmental variation and uneven incidence of SBCMV infection across the field, scoring/sampling date, and seasons. It is well known that SBCMV infection effects are strongly influenced by environmental factors such as temperatures and soil humidity that impact both vector and virus spreading, hence causing spatial patchiness of infection incidence (Bayles et al. 2007; Liu et al. 2014). Moreover, effects of cold weather on plant growth can worsen virus symptoms. For these reasons, implementing growth chamber experiments under controlled conditions is highly desirable to standardize infection effects and to increase heritability.

The frequency of informative recombination events between *KUBO-13* and *KUBO-9* obtained from the two mapping populations was highly heterogeneous and considerably lower in $Sv \times Cc$ as compared to $Mr \times Cl$. This was possibly due to local cross-specific suppression of recombination (Akhunov et al. 2003; Conley et al. 2004), supporting the need for diverse genetic materials. Maccaferri et al. (2019) reported that the chr. 2BS telomere of the Svevo genome is highly recombinogenic, with an average physical-to-genetic ratio of 1.8 Mb/cM. For the interval between *KUBO-13* and *KUBO-9*, our results demonstrated a physical-to-genetic ratio of 1.3 and 2.5 Mb/cM for $Sv \times Cc$ and $Mr \times Cl$ RILs, respectively. However, in correspondence of the fine-mapped interval of *Q_{Sbm.ubo-2BS}* interval between *KUBO-27* and *KUBO-1*, the ratio was 3.1 Mb/cM for $Mr \times Cl$ and 13.5 Mb/cM for $Sv \times Cc$.

The complementary GWAS analysis conducted on modern cultivars from around the world allowed mapping *Sbm2* to the same short interval obtained from the RIL fine mapping experiments, hence providing an independent validation of results, while enabling the evaluation of allelic distribution and diagnostic power of significant SNPs in cultivated germplasm. All markers were validated and the assessment of their diagnostic power highlighted *wPt-2106_ASO*, *wPt-2106_HRM*, *KUBO-27*, *KUBO-29*, *KUBO-41*, and *KUBO-3* as the best markers for MAS.

The potential to predict the phenotype at the causal locus among genotypes belonging to different breeding origins/lineages is important for wide application of MAS in diverse genetic backgrounds and contributes to define the breeding value of a marker (Terracciano et al. 2013). A drawback of SNP markers, as compared to SSR markers, is their biallelism, as SNPs fit with the infinite site and allele mutation model (Kimura and Crow 1964; Kimura 1969). This strongly

limits the chance of a single SNP marker to consistently match with the causative mutations and haplotypes at the QTL. Notably, the availability of multiple SNP markers tightly associated with the target locus allows for more accurate haplotype-based screening, a more effective approach for increasing the diagnostic power of MAS (Meuwissen et al. 2001; Mucha and Wierzbicki 2012) and a more effective breeding by design approach (Peleman and Voort 2003).

In this study, several concatenated SNPs highly diagnostic of the allelic state at *Sbm2* have been identified and some of them were successfully transformed into high-throughput KASP® or HRM markers, in sufficient number to perform haplotype-based characterization. This enables to trace identity by descent of *Sbm2* alleles at high level of confidence. The combination of *KUBO-27*, *wPt-2106*, *KUBO-29*, and *KUBO-41* provides the best multi-marker haplotype assay available to ascertain the presence of the resistant-*Sbm2* allele for germplasm characterization, including diverse breeder's materials, segregating individuals, and lines from artificial crosses, while characterizing selected breeding lines as possible candidate for varietal registration.

Until now, *Sbm2* has been considered the only major QTL to provide SBCMV resistance in durum wheat (Bayles et al. 2007; Maccaferri et al. 2012). If additional major sources of resistance are found in *Triticum turgidum* and/or other tetraploid wheat germplasm, or in chromosomes A and/or B of bread wheat, it would be useful to pyramid them into new varieties to increase both virus resistance and durability; hence, the KASP® markers developed in this study are instrumental.

Recently, the *Jmv1* locus for resistance against the *Japanese soil-borne wheat mosaic virus*, a Furovirus also transmitted by *Polymyxa graminis*, has been identified in chromosome 2H of barley (Okada et al. 2020 and 2023). A joint comparative analysis of the two loci in Triticeae would increase the chance to better understand the causal locus and mechanism of resistance.

Candidate genes in *Sbm2* interval

The Svevo genome enabled the identification of the gene(s) responsible for *Cdu-B1*, the major locus underlying the cadmium concentration QTLome in grain, mainly based on candidate gene analysis and translational genomics, previously hindered by the low recombination rate in the target region (Maccaferri et al. 2019). A similar approach could be used to dissect the *Sbm2* region. Other solutions are represented by alternative cloning approaches which do not rely on positional fine mapping, like the MutRenSeq method proposed by Steuernagel et al. (2016).

A glance at the list of candidate genes proposed in this study identified three gene categories as most promising for underlying SBCMV resistance: defensins, disease resistance

protein (NBS-LRR class) family, and NBS-LRR-like resistance proteins. Plant defensins are a family of folded antimicrobial peptides thought to be effective against fungal pathogen (Thomma et al. 2002), while NBS-LRR resistance proteins represent the largest class of plant resistance genes (*R* genes) and *R* genes are involved in resistance against diverse pathogens, including viruses (Calil and Fontes 2017).

Specifically to SBCMV, Vallega et al. (2006) proposed a hypothesis about a generalized resistance to virus accumulation in durum wheat, i.e., a mechanism controlling virion accumulation in plants by limiting the spread of the virus particles from the roots to the shoot. Kanyuka et al. (2004), Ordon et al. (2009), and Perovic et al. (2009) reported a potential *Sbm1* mechanism defined as “translocation resistance,” which prevents the virus to spread from the roots to the stem and leaves in SBCMV-resistant hexaploid cultivars. However, a more complete explanation about the mechanism of SBCMV defense in wheat has not yet been sufficiently defined. Thus, it would be inappropriate to narrow down the search of candidate genes to the *R* gene category only. Singh and Sharma (2015) reported that several terpenoids have a role in plant defense, acting as phytoalexins, low molecular weight compounds, known to exhibit antimicrobial properties in rice (Prisic et al. 2004) and deterrent for insects and herbivores in corn, lima bean, poplar, and cotton (Rodriguez-Saona and Crafts-Brandner 2003; Arimura et al. 2004; Mithofer 2005; Schnee et al. 2006). Ten genes involved in terpene synthase activity were detected in the *Sbm2* region, representing the second most frequent gene category. Another uncertainty lies in disentangling the direct resistance against the virus from the indirect resistance against the *Polymyxa* vector, particularly because *P. graminis* is not well known as *Polymyxa betae*. Additionally, with the objective of providing an overall view of the entire list of genes present in the confidence interval, we did not exclude low-confidence genes. High-/low-confidence classification based on a single reference genome is not a valid criterion to discriminate putative candidate genes. Jupe et al. (2012), in a survey of NBS-LRR genes in the potato genome, showed that many of these genes previously annotated as partial/low confidence by the Consortium (2011) were actually functional.

Herein, an overall view on the list of genes presents in the confidence interval of the Svevo durum wheat reference genome and resistance donor was outlined, providing a starting point for a deeper and detailed scan of functions and biological processes.

Conclusions

In this study, SNP markers were used to dissect the genetic basis of SBCMV resistance in durum wheat. High-density SNP arrays allowed for marker enrichment of *Q_{Sbm}*.

ubo-2BS = Sbm2 region, while simple, cheap, and robust SNP-specific assays are required for targeted high-throughput screening such as fast evaluation of large segregant populations and efficient marker-assisted selection. Among the available technical solutions, KASP® technology is currently the most used worldwide. Fine mapping of *Sbm2* was achieved, with *Sbm2* now delimited to a 0.2-cM interval including *wPt-2106* and flanked by *KUBO-27* and *KUBO-40-KUBO-41*, overlapping the target QTL.

This study provides a series of markers organized into haplotypes closely associated and overlapping to the target QTL, hence highly diagnostic for resistance. Such markers are instrumental for breeding programs to select SBCMV-resistant genotypes in durum wheat, with the aim of increasing the frequency of *Sbm2*-resistant haplotype. This work paves the way for widening the investigation to bread wheat germplasm, where *Sbm2* has also been described.

Additionally, the effective development of *Sbm2* haplotype tagging KASP® presented herein can be considered as a case study to innovate plant variety testing in order to improve the efficacy and accuracy of European variety tests as well as the assessment of the value for cultivation and use and the decision-making process.

The recent availability of the genome assembly of the parental line and germplasm founder Svevo durum wheat allowed us to compare the genetic and physical maps at the target region and identify putative candidate genes. Therefore, this study provides the basis for further refining the *Sbm2* interval and paves the way for the positional cloning of this locus.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s00122-024-04709-7>.

Author Contribution Statement MM, RT, MCS, CR, CR-A, DP, EG, and PD conceived the experiment and supervised the work. MM, SC (Simona Corneti), AG, PD, EG developed and/or provided the genetic materials. MB (Matteo Bozzoli) and MM selected varieties for the germplasm panels. MM, GS, CR, CR-A, SS, MCS, RT designed the phenotyping and genotyping methodology. SC (Simona Corneti), SC (Sara Castelletti), MB (Martina Bruschi), MB (Matteo Bozzoli), JBN, MM, DP, GS performed the phenotypic evaluation and genotyping analyses. MB (Martina Bruschi), MB (Matteo Bozzoli), EF (Elisabetta Frascaroli), and GS performed statistical analyses. DO and MB performed bioinformatics analyses. MB (Martina Bruschi), MB (Matteo Bozzoli), JBN, SC (Sara Castelletti), EF (Elena Fusari), and SC (Simona Corneti) performed DNA extraction and laboratory analysis including KASP® development. MB (Martina Bruschi), MM, GS, and DO analyzed the data and interpreted the results. CR and CR-A provided ELISA data. MB (Martina Bruschi), MB (Matteo Bozzoli), MM, RT wrote the manuscript. MM and RT reviewed and edited the manuscript. All authors have read and approved the manuscript.

Funding Open access funding provided by Alma Mater Studiorum - Università di Bologna within the CRUI-CARE Agreement. This research study was conducted with the financial support of the French project “FSOV 2016 C MOSAIC RESISTANCE IN DURUM,” the Italian Cluster AGRIFOOD SOSTENIBILITA’ DELLA FILIERA

AGROALIMENTARE (SO.FI.A) project, the support of the Interdepartmental Centre for Industrial Agrofood Research—CIRI Agrofood, and the European Union's Horizon 2020 research and innovation program project "INNOVAR—Next generation variety testing for improved cropping on European farmland," under grant agreement No 818144.

Data availability The datasets generated in the current study are available as Supplementary Information (SI). Additional datasets are available from the corresponding author on request.

Declarations

Conflict of interests The authors have no relevant financial or non-financial interests to disclose.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Akhunov ED, Goodyear AW, Geng S, Qi L-L, Echalié B, Gill BS, Miftahudin GJP, Lazo G, Chao S, Anderson OD, Linkiewicz AM, Dubcovsky J, Rota ML, Sorrells ME, Zhang D, Nguyen HT, Kalavacharla V, Hossain K, Kianian SF, Peng J, Lapitan NLV, Gonzalez-Hernandez JL, Anderson JA, Choi D-W, Close TJ, Dilibirli M, Gill KS, Walker-Simmons MK, Steber C, McGuire PE, Qualset CO, 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>
- Arimura G, Huber DPW, Bohlmann J (2004) Forest tent caterpillars (*Malacosoma disstria*) induce local and systemic diurnal emissions of terpenoid volatiles in hybrid poplar (*Populus trichocarpa* × *deltoides*): cDNA cloning, functional characterization, and patterns of gene expression of (–) germacrene D synthase, PtdTPS1. *Plant J* 37:603–616. <https://doi.org/10.1111/j.1365-3113X.2003.01987.x>
- Barrett JC, Fry B, Maller J, Daly MJ (2005) Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 21:263–265. <https://doi.org/10.1093/bioinformatics/bth457>
- Bass C, Hendley R, Adams MJ, Hammond-Kosack KE, 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.* 2014.2016
- 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 Project Report
- Bradbury PJ, Zhang Z, Kroon DE, Casstevens TM, Ramdoss Y, Buckler ES (2007) TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics* 23:2633–2635. <https://doi.org/10.1093/bioinformatics/btm308>
- Browning BL, Browning SR (2022) Genotype error biases trio-based estimates of haplotype phase accuracy. *Am J Hum Genet* 109(6):1016–1025
- Budge GE, Ratti C, Rubies-Autonell C, Lockley D, Bonnefoy M, Vallega V, Pietravalle S, Henry CM (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>
- Calil IP, Fontes EPB (2017) Plant immunity against viruses: antiviral immune receptors in focus. *Ann Bot* 119:711–723. <https://doi.org/10.1093/aob/mcw200>
- Cavanagh CR, Chao S, Wang S, Huang BE, Stephen S, Kiani S, Forrest K, Saintenac C, Brown-Guedira GL, Akhunova A, See D, Bai G, Pumphrey M, Tomar L, Wong D, Kong S, Reynolds M, da Silva ML, Bockelman H, Talbert L, Anderson JA, Dreisigacker S, Baenziger S, Carter A, Korzun V, Morrell PL, Dubcovsky J, Morell MK, Sorrells ME, Hayden MJ, 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>
- Clover GRG, Wright DM, Henry CM (1999) Occurrence of soil-borne wheat mosaic virus in the UK. In: *Proceedings Fourth Symposium International Work. Group Plant Viruses Fungal Vectors Asilomar Conference Cent. Monterey California USA 5–8 Oct 1999*, pp 105–108
- Conley EJ, Nduati V, Gonzalez-Hernandez JL, Mesfin A, Trudeau-Spanjers M, Chao S, Lazo GR, Hummel DD, Anderson OD, Qi LL, Gill BS, Echalié B, Linkiewicz AM, Dubcovsky J, Akhunov ED, Dvořák J, Peng JH, Lapitan NLV, Pathan MS, Nguyen HT, Ma X-F, Miftahudin GJP, Greene RA, Sorrells ME, Hossain KG, Kalavacharla V, Kianian SF, Sidhu D, Dilibirli M, Gill KS, Choi DW, Fenton RD, Close TJ, McGuire PE, Qualset CO, Anderson JA (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 TPGS (2011) Genome sequence and analysis of the tuber crop potato. *Nature* 475:189. <https://doi.org/10.1038/nature10158>
- Devos KM, Dubcovsky J, Dvořák J, Chinoy CN, Gale MD (1995) Structural evolution of wheat chromosomes 4A, 5A, and 7B and its impact on recombination. *Theor Appl Genet* 91:282–288. <https://doi.org/10.1007/BF00220890>
- 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>
- Gabriel SB, Schaffner SF, Nguyen H et al (2002) The structure of haplotype blocks in the human genome. *Science*. 296(5576):2225–2229. <https://doi.org/10.1126/science.1069424>
- Gadaleta A, Giancaspro A, Giove SL, 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* 118:1015–1025. <https://doi.org/10.1007/s00122-008-0958-1>
- 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>
- International Wheat Genome Sequencing Consortium IWGS, Eversole K, Stein N, Feuillet C, Keller B, Rogers J, Pozniak CJ, Choulet F, Distelfeld A, Poland J, Ronen G, Sharpe AG, Barad O, Baruch K, Keeble-Gagnère G, Mascher M, Ben-Zvi G, Joselin A-A, Himmelbach A, Balfourier F, Gutierrez-Gonzalez J, Hayden M, Koh C, Muehlbauer G, Pasam RK, Paux E, Rigault

- P, Tibbits J, Tiwari V, Spannagl M, Lang D, Gundlach H, Haberer G, Mayer KFX, Ormanbekova D, Prade V, Šimková H, Wicker T, Swarbreck D, Rimbart H, Felder M, Guilhot N, Kaithakottil G, Keilwagen J, Leroy P, Lux T, Twardziok S, Venturini L, Juhász A, Abrouk M, Fischer I, Uauy C, Borrill P, Ramirez-Gonzalez RH, Arnaud D, Chalabi S, Chalhou B, Cory A, Datla R, Davey MW, Jacobs J, Robinson SJ, Steuernagel B, van Ex F, Wulff BBH, Benhamed M, Bendahmane A, Concia L, Latrasse D, Bartoš J, Bellec A, Berges H, Doležel J, Frenkel Z, Gill B, Korol A, Letellier T, Olsen O-A, Singh K, Valárik M, van der Vossen E, Vautrin S, Weining S, Fahima T, Glikson V, Raats D, Čížalíková J, Toegelová H, Vrána J, Sourdille P, Darrier B, Barabaschi D, Cattivelli L, Hernandez P, Galvez S, Budak H, Jones JDG, Witek K, Yu G, Small I, Melonek J, Zhou R, Belova T, Kanyuka K, King R, Nilsen K, Walkowiak S, Cuthbert R, Knox R, Wiebe K, Xiang D, Rohde A, Golds T, Čížková J, Akpinar BA, Biyiklioglu S, Gao L, N'Daiye A, Kubaláková M, Šafář J, Alfama F, Adam-Blondon A-F, Flores R, Guerche C, Loaec M, Quesneville H, Condie J, Ens J, Maclachlan R, Tan Y, Alberti A, Aury J-M, Barbe V, Couloux A, Cruaud C, Labadie K, Mangenot S, Wincker P, Kaur G, Luo M, Sehgal S, Chhuneja P, Gupta OP, Jindal S, Kaur P, Malik P, Sharma P, Yadav B, Singh NK, Khurana JP, Chaudhary C, Khurana P, Kumar V, Mahato A, Mathur S, Sevanthi A, Sharma N, Tomar RS, Holušová K, Plíhal O, Clark MD, Heavens D, Kettleborough G, Wright J, Balcárková B, Hu Y, Salina E, Ravin N, Skryabin K, Beletsky A, Kadnikov V, Mardanov A, Nesterov M, Rakitin A, Sergeeva E, Handa H, Kanamori H, Katagiri S, Kobayashi F, Nasuda S, Tanaka T, Wu J, Cattonaro F, Jiumeng M, Kugler K, Pfeifer M, Sandve S, Xun X, Zhan B, Batley J, Bayer PE, Edwards D, Hayashi S, Tulpová Z, Visendi P, Cui L, Du X, Feng K, Nie X, Tong W, Wang L (2018) Shifting the limits in wheat research and breeding using a fully annotated reference genome. *Science* 361:eaar7191. <https://doi.org/10.1126/science.aar7191>
- Jupe F, Pritchard L, Etherington GJ, MacKenzie K, Cock PJ, Wright F, Sharma SK, Bolser D, Bryan GJ, Jones JD, 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>
- Kanehisa M, Sato Y, Morishima K (2016) BlastKOALA and GhostKOALA: KEGG tools for functional characterization of genome and metagenome sequences. *J Mol Biol* 428:726–731
- Kanyuka K, Ward E, Adams MJ (2003) *Polymyxa graminis* and the cereal viruses it transmits: a research challenge. *Mol Plant Pathol* 4:393–406. <https://doi.org/10.1046/j.1364-3703.2003.00177.x>
- Kanyuka K, Lovell DJ, Mitrofanova OP, Hammond-Kosack K, Adams MJ (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>
- Kim B, Jeong J, Yoon T (2017) Analysis of soil-borne wheat mosaic virus and soil-borne cereal mosaic virus using datamining. *Int J Pharma Med Biol Sci*. 6:11–15. <https://doi.org/10.18178/ijpmb.6.1.11-15>
- Kimura M (1969) The number of heterozygous nucleotide sites maintained in a finite population due to steady flux of mutations. *Genetics* 61:893–903. <https://doi.org/10.1093/genetics/61.4.893>
- Kimura M, Crow J (1964) The number of alleles that can be maintained in a finite population. *Genetics* 49:725–738. <https://doi.org/10.1093/genetics/49.4.725>
- Koebner RMD, Summers RW (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](https://doi.org/10.1016/S0167-7799(02)00036-7)
- Kolde R, Kolde MR (2015) Package ‘pheatmap’. R package 1(7):790
- Korte A, Farlow A (2013) The advantages and limitations of trait analysis with GWAS: a review. *Plant Methods* 9:29. <https://doi.org/10.1186/1746-4811-9-29>
- Kruijer W, Boer MP, Malosetti M, Flood PJ, Engel B, Kooke R, Keurentjes JJB, van Eeuwijk F (2015) Marker-based estimation of heritability in immortal populations. *Genetics* 199:379–398. <https://doi.org/10.1534/genetics.114.167916>
- Kühne T (2009) Soil-borne viruses affecting cereals: known for long but still a threat. *Virus Res* 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. *Theor Appl Genet* 127:1039–1047. <https://doi.org/10.1007/s00122-014-2277-z>
- Liu S, Bai G, Lin M, Luo M, Zhang D, Jin F, Tian B, Trick HN, Yan L (2020) Identification of candidate chromosome region of Sbwml1 for Soil-borne wheat mosaic virus resistance in wheat. *Sci Rep* 10:8119. <https://doi.org/10.1038/s41598-020-64993-3>
- Maccaferri M, Ratti C, Rubies-Autonell C, Vallega V, Demontis A, Stefanelli S, Tuberosa R, Sanguineti MC (2011a) Resistance to Soil-borne cereal mosaic virus in durum wheat is controlled by a major QTL on chromosome arm 2BS and minor loci. *Theor Appl Genet* 123:527–544. <https://doi.org/10.1007/s00122-011-1605-9>
- Maccaferri M, Sanguineti MC, Demontis A, El-Ahmed A, Garcia del Moral L, Maalouf F, Nachit M, Nserallah N, Ouabou H, Rhouma S, Royo C, Villegas D, Tuberosa R (2011b) Association mapping in durum wheat grown across a broad range of water regimes. *J Exp Bot* 62:409–438. <https://doi.org/10.1093/jxb/erq287>
- Maccaferri M, Francia R, Ratti C, Rubies-Autonell C, Colalongo C, Ferrazzano G, Tuberosa R, Sanguineti MC (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, Ricci A, Salvi S, Milner SG, Noli E, Martelli PL, Casadio R, Akhunov E, Scalabrini S, Vendramin V, Ammar K, Blanco A, Desiderio F, Distelfeld A, Dubcovsky J, Fahima T, Faris J, Korol A, Massi A, Mastrangelo AM, Morgante M, Pozniak C, N'Daiye 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>
- Maccaferri M, Harris NS, Twardziok SO, Pasam RK, Gundlach H, Spannagl M, Ormanbekova D, Lux T, Prade VM, Milner SG, Himmelbach A, Mascher M, Bagnaresi P, Faccioli P, Cozzi P, Lauria M, Lazzari B, Stella A, Manconi A, Gnocchi M, Moscatelli M, Avni R, Deek J, Biyiklioglu S, Frascaroli E, Corneti S, Salvi S, Sonnante G, Desiderio F, Marè C, Crosatti C, Mica E, Özkan H, Kilian B, De Vita P, Marone D, Joukhadar R, Mazzucotelli E, Nigro D, Gadaleta A, Chao S, Faris JD, Melo ATO, Pumphrey M, Pecchioni N, Milanese L, Wiebe K, Ens J, MacLachlan RP, Clarke JM, Sharpe AG, Koh CS, Liang KYH, Taylor GJ, Knox R, Budak H, Mastrangelo AM, Xu SS, Stein N, Hale I, Distelfeld A, Hayden MJ, Tuberosa R, Walkowiak S, Mayer KFX, Ceriotti A, Pozniak CJ, Cattivelli L (2019) Durum wheat genome highlights past domestication signatures and future improvement targets. *Nat Genet* 51:885–895. <https://doi.org/10.1038/s41588-019-0381-3>
- Maccaferri M, Bruschi M, Tuberosa R (2022) Sequence-based marker assisted selection in wheat. In: Reynolds MP, Braun HJ (eds) *Wheat improvement: food security in a changing climate*. Springer International Publishing, Cham, pp 513–553. <https://doi.org/10.1007/978-3-030-90673-3>

- Makhouf M, Rambla C, Voss-Fels KP, Hickey LT, Snowdon RJ, Obermeier C (2020) Overcoming polyploidy pitfalls: a user guide for effective SNP conversion into KASP markers in wheat. *Theor Appl Genet* 133:2413–2430. <https://doi.org/10.1007/s00122-020-03608-x>
- Marçais G, Delcher AL, Phillippy AM, Coston R, Salzberg SL, Zimin A (2018) MUMmer4: a fast and versatile genome alignment system. *PLoS Comput Biol* 14:
- Meuwissen THE, Hayes BJ, Goddard ME (2001) Prediction of total genetic value using genome-wide dense marker maps. *Genetics* 157:1819–1829
- Mithofer A (2005) Effects of feeding *Spodoptera littoralis* on lima bean leaves. II. Continuous mechanical wounding resembling insect feeding is sufficient to elicit herbivory-related volatile emission. *Plant Physiol* 137:1160–1168. <https://doi.org/10.1104/pp.104.054460>
- 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>
- Okada K, Kato T, Oikawa T, Komatsuda T, Namai K (2020) A genetic analysis of the resistance in barley to Soil-borne wheat mosaic virus. *Breed Sci* 70:617–622. <https://doi.org/10.1270/jsbbs.20071>
- Okada K, Xu W, Mishina K, Oono Y, Kato T, Namai K, Komatsuda T (2023) Genetic resistance in barley against Japanese soil-borne wheat mosaic virus functions in the roots. *Front Plant Sci* 14:1149752. <https://doi.org/10.3389/fpls.2023.1149752>
- 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>
- Peleman JD, Van der Voort JR (2003) Breeding by design. *Trends Plant Sci* 8(7):330–334. [https://doi.org/10.1016/S1360-1385\(03\)00134-1](https://doi.org/10.1016/S1360-1385(03)00134-1)
- Perovic D, Förster J, Devaux P, Hariri D, Guillieroux 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>
- Prisic S, Xu M, Wilderman PR, Peters RJ (2004) Rice contains two disparate ent-Copalyl Diphosphate synthases with distinct metabolic functions. *Plant Physiol* 136:4228–4236. <https://doi.org/10.1104/pp.104.050567>
- R Core Team (2017) R: a language and environment for statistical computing
- Raudvere U, Kolberg L, Kuzmin I, Arak T, Adler P, Peterson H, Vilo J (2019) g: Profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update). *Nucleic Acids Res*. 47:91–198. <https://doi.org/10.1093/nar/gkz369>
- Roberts AG (2014) Plant viruses: soil-borne. ELS John Wiley Sons Ltd. Chichester. <https://doi.org/10.1002/9780470015902.a0000761.pub3>
- Rodriguez-Saona C, Crafts-Brandner SJ (2003) Volatile emissions triggered by multiple herbivore damage: beet armyworm and whitefly feeding on cotton plants. *J Chem Ecol* 29:2539–2550. <https://doi.org/10.1023/A:1026314102866>
- 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, 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>
- Salvi S, Tuberosa R (2015) The crop QTLome comes of age. *Curr Opin Biotechnol* 32:179–185. <https://doi.org/10.1016/j.copbio.2015.01.001>
- Schnee C, Kollner TG, Held M, Turlings TCJ, Gershenzon J, Degenhardt J (2006) The products of a single maize sesquiterpene synthase form a volatile defense signal that attracts natural enemies of maize herbivores. *Proc Natl Acad Sci* 103:1129–1134. <https://doi.org/10.1073/pnas.0508027103>
- Semang K, Babu R, Hearne S, Olsen S (2014) Single nucleotide polymorphism genotyping using Kompetitive Allele Specific PCR (KASP): overview of the technology and its application in crop improvement. *Mol Breeding* 33:1–14. <https://doi.org/10.1007/s11032-013-9917-x>
- Shavrukov Y (2016) Comparison of SNP and CAPS markers application in genetic research in wheat and barley. *BMC Plant Biol*. <https://doi.org/10.1186/s12870-015-0689-9>
- Singh B, Sharma RA (2015) Plant terpenes: defense responses, phylogenetic analysis, regulation and clinical applications. *Biotech* 5:129–151. <https://doi.org/10.1007/s13205-014-0220-2>
- Steuernagel B, Periyannan SK, Hernández-Pinzón I, Witek K, Rouse MN, Yu G, Hatta A, Ayliffe M, Bariana H, Jones JDG, Lagudah ES, Wulff BBH (2016) Rapid cloning of disease-resistance genes in plants using mutagenesis and sequence capture. *Nat Biotechnol* 34:652–655. <https://doi.org/10.1038/nbt.3543>
- Terracciano I, Maccaferri M, Bassi F, Mantovani P, Sanguineti MC, 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 BP, Cammue BP, Thevissen K (2002) Plant defensins. *Planta* 216:193–202. <https://doi.org/10.1007/s00425-002-0902-6>
- Vallega V, Rubies Autonell C (1985) Reactions of Italian *Triticum durum* cultivars to soilborne wheat mosaic. *Plant Dis* 69:64–66
- Vallega V, Ratti C, Rubies-Autonell C (2003) Reaction of durum wheat cultivars to mixed SBWMV and WSSMV infection in Central Italy. *Phytopathol Mediterr* 42(2):177–182
- Vallega V, Rubies-Autonell C, Ratti C (2006) Resistance to accumulation of Soil-borne cereal mosaic virus in eight cultivars of *Triticum durum* Desf. *Parasitica* 62:79–96
- Van Ooijen JW (2006) JoinMap 4. software for the calculation of genetic linkage maps in experimental populations Kyazma BV, Wageningen, Netherlands Voorrips RE MapChart: software for the graphical presentation of linkage maps and QTLs. *Theor Appl Genet* 124:323332
- Wang S, Basten CJ, Zeng Z -B (2012) Windows QTL Cartographer 2.5. Department of Statistics, North Carolina State University, Raleigh, NC. (<http://statgen.ncsu.edu/qtlcart/WQTLCart.htm>)
- Wang S, Wong D, Forrest K, Allen A, Chao S, Huang BE, Maccaferri M, Salvi S, Milner SG, Cattivelli L, Mastrangelo AM, Whan A, Stephen S, Barker G, Wieseke R, Plieske J, Lillemo M, Mather D, Appels R, Dolferus R, Brown-Guedira G, Korol A, Akhunova AR, Feuillet C, Salse J, Morgante M, Pozniak C, Luo M-C, Dvorak J, Morell M, Dubcovsky J, Ganai M, Tuberosa R, Lawley C, Mikoulitch I, Cavanagh C, Edwards KJ, 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>
- Wittwer CT, Reed GH, Gundry CN, Vandersteen JG, Pryor RJ (2003) High-resolution genotyping by amplicon melting analysis using LCGreen. *Clin Chem* 49(6 Pt 1):853–860. <https://doi.org/10.1373/49.6.853>

- 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 JC, Chang TT, Konzak CF (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>
- Zhu T, Wang L, Rodriguez JC, Deal KR, Avni R, Distelfeld A, McGuire PE, Dvorak J, Luo M-C (2019) Improved genome sequence of Wild Emmer Wheat Zavitan with the aid of optical maps. *G3 Genes Genomes Genet* 9:619–624. <https://doi.org/10.1534/g3.118.200902>
- Zhu T, Wang L, Rimbert H, Rodriguez JC, Deal KR, Oliveira RD, Choulet F, Keeble-Gagnère G, Tibbits J, Rogers J, Eversole K, Appels R, Gu YQ, Mascher M, Dvorak J, Luo M-C (2021) Optical maps refine the bread wheat *Triticum aestivum* cv. Chin Spring Genome Assem *Plant J* 107:303–314. <https://doi.org/10.1111/tpj.15289>

Cited websites:

LGC Group: <https://www.lgcgroup.com/products/kasp-genotyping-chemistry/> Polymarker: <http://Polymarker.tgac.ac.uk>. Windows QTL cartographer: Wang S., C. J. Basten, and Z.-B. Zeng (2012). Windows QTL Cartographer 2.5. Department of Statistics, North Carolina State University, Raleigh, NC. (<http://statgen.ncsu.edu/qtldcart/WQTLCart.htm>) The Durum Wheat Genome Browser (<https://d-gbrowse.interomics.eu/>) The Durum Wheat Genome Database (<https://d-data.interomics.eu/>)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.