Identification of a new gene *Rf9* and unravelling the genetic complexity for controlling fertility restoration in hybrid wheat

Fahimeh Shahinnia, Manuel Geyer, Annette Block, Volker Mohler, Lorenz Hartl*

Institute for Crop Science and Plant Breeding, Bavarian State Research Centre for Agriculture, Am Gereuth 8, 85354 Freising, Germany

* **Corresponding author:**

Dr. Lorenz Hartl

E-Mail: lorenz.hartl@lfl.bayern.de

Abstract

Wheat (*Triticum aestivum* L.*)* is a self-pollinating crop whose hybrids offer the potential to provide a major boost in yield. Male sterility induced by the cytoplasm of *Triticum timopheevii* is a powerful method for hybrid seed production. Hybrids produced by this method are often partially sterile and full fertility restoration is crucial for wheat production using hybrid cultivars. To identify genetic loci controlling fertility restoration in wheat, we produced two CMS-based backcross (BC_1) mapping populations. The restorer lines Gerek 79 and 71R1203 were used to pollinate the male-sterile winter wheat line CMS-Sperber. Seed set and numbers of sterile spikelets per spike were evaluated in 340 and 206 individuals of the populations derived from Gerek 79 and 71R1203, respectively. Genetic maps were constructed using 930 and 994 SNPs, spanning 2,160 and 2,328 cM over 21 linkage groups in the two populations, respectively. Twelve quantitative trait loci (QTL) controlled fertility restoration in both BC_1 populations, including a novel *restorer-of-fertility* (*Rf*) locus flanked by the single nucleotide polymorphism (SNP) markers *IWB72413* and *IWB1550* on chromosome 6AS. The locus was mapped as a qualitative trait in the BC_1 Gerek 79 population and was designated *Rf9*. Ninety-three putative candidate genes were predicted for the QTL region on chromosome 6AS. Among them were genes encoding tetratricopeptide and pentatricopeptide repeat-containing proteins in rice known to be associated with fertility restoration. This finding is a promising step to better understand the functions of genes for improving hybrid wheat.

Keywords: Cytoplasmic male sterility, seed set, sterile spikelet, *Triticum aestivum*, *Triticum timopheevii*

Introduction

Since the discovery of male sterility and restoration systems in the 1960s, hybrid wheat triggered attention due to its potential for improved grain and straw productivity and yield stability particularly under harsh and marginal environments (Longin et al. 2012; Whitford et al. 2013). The major gains of hybrid versus line varieties are improved trait values due to heterosis (Castillo et al. 2014). Heterosis or hybrid vigor is the phenomenon by which the progeny derived from a cross of two inbred lines outperform the parent lines and has provided considerable economic benefits in worldwide crop production (Chen and Liu 2014). Heterosis in winter wheat has been reported to provide uniform plant establishment and tolerance against frost, lodging, and diseases such as leaf rust, stripe rust, *Septoria tritici* blotch and powdery mildew (Gupta et al. 2019).

To harness yield gains associated with hybrid vigor, the cytoplasmic male sterility (CMS) system provides a cost-effective tool for efficient hybrid seed production (Chen and Liu 2014). CMS in plants is based on rearrangements of mitochondrial DNA that lead to chimaeric genes and a condition under which a plant is unable to produce fertile pollen (Whitford et al. 2013). CMS evades the need for manual removal of anthers, thus facilitating a technology to produce unlimited numbers for hybrid plants. It has been successfully used in crops such as rye, rice, maize and sunflower (Castillo et al. 2015). In wheat, the outcome of a hybridization event involving *T. timopheevii* as the female parent and *Triticum aestivum* (Bohra et al. 2016) causes CMS. The use of *T. timopheevii* cytoplasm in bread and durum wheat creates male sterility, whereas female fertility is not impaired (Lukaszewski 2017). CMS-based hybrid production uses three different breeding lines, a CMS line, a maintainer line, and a restorer line (Chen and Liu 2014). The CMS line is used as the female parent with at least one CMS-causing gene in the *T. timopheevii*-derived cytoplasm and lacking functional nuclear-encoded *Restorer-of-fertility* (*Rf*) genes (Schnable and Wise 1998). The maintainer line serves as the male parent in crosses for the propagation and maintenance of the CMS line, with the same nuclear genome as the CMS line, but a normal fertile *T. aestivum* cytoplasm. The restorer line retains (a) functional *Rf* gene(s) and acts as the male parent to cross with the CMS line to produce the F_1 hybrid seeds. In F_1 plants, the *Rf* gene(s) restore(s) male fertility, and the combination of the nuclear genomes from the CMS line and the restorer line produces hybrid vigor (Chen and Liu 2014). For commercial hybrid seed production, a male-sterile

line has to be crossed with a line carrying dominant restorer alleles and suitable pollinator qualities (Whitford et al. 2013). While fertility restoration is a crucial trait in hybrid breeding, hybrids produced using this method are often partially sterile due to the complex interaction between mitochondrial and nuclear genes controlling male specificity, occurrence, and restoration of fertility (Chen and Liu 2014). Therefore, incomplete fertility restoration poses a major bottleneck for hybrid wheat breeding, as it impairs the effect of heterosis on grain yield and the uniformity or quality of end-use products (Keydel 1973; Geyer et al. 2018).

Fertility restoration is a genetically complex process and is mainly controlled by the mitochondrial genome in interaction with *Rf* genes (Castillo et al. 2014). Besides, it is known that fertility restoration is influenced by environmental factors including photoperiod, water stress, light intensity and temperature (Castillo et al. 2015). The mitochondrial gene families produce proteins that share the common structural organization of similar repeated helical motifs and include half-tetratricopeptides, pentatricopeptide repeat (PPR) proteins, and mitochondrial transcription termination factors (mTERF) (Pan et al. 2019). Tetratricopeptide repeat (TPR) motifs occur in proteins from bacteria to humans and play important roles in various processes, including hormone regulation, protein kinase inhibition, protein folding, cell cycle and regulation of reproductive development (Yu et al. 2016). The effect of cytoplasmic male sterility can be weakened by inhibiting the accumulation of the CMS-conferring gene products through the function of a class of *Rf* genes, generally belonging to a large family of genes that encode organelle-targeted PPR proteins (Hu et al. 2012). mTERF genes are widely distributed in metazoans, plants and green alga. They regulate transcription, translation and DNA replication of mitochondrial genes in metazoans while regulating gene expression in chloroplasts and mitochondria in plants (Pan et al. 2019). In wheat, the presence of eight major genes (*Rf1*-*Rf8*) for timopheevii-based cytoplasmic male sterility is known and assigned to the chromosomes 1A, 7D, 1B, 6B, 6D, 5D, 7B and 2D, respectively (Mukai and Tsunewaki 1979; Sinha et al. 2013; Gupta et al. 2019). *Rf1* and *Rf3* are the most effective genes for achieving restoration in wheat (Geyer et al. 2016, 2018; Würschum et al. 2017). Previous studies have indicated that combinations of two or three major *Rf* genes and restorer genes with small effect or low penetrance (modifier loci) can modify the degree of fertility restoration (Ma et al. 1995; Ahmed et al. 2001; Zhou et al. 2005; Stojałowski et al. 2013; Whitford et al. 2013; Lukaszewski

2017). Consequently, attempts are made to pyramid multiple dominant or partially dominant alleles of the most favourable genes or quantitative trait loci (QTL) including those involved in epistatic interactions in order to achieve complete fertility restoration in hybrid wheat (Zhao et al. 2013; Jiang et al. 2017; Gupta et al. 2019). Understanding the genetic mechanisms underlying restoration of fertility and developing elite restorer lines are crucial to overcoming the intricate barriers in hybrid breeding programs. For this reason, our objective was to identify new genetic loci controlling fertility restoration that can be employed in hybrid wheat breeding. Here, we developed two CMS-based backcross (BC) mapping populations, which we used for QTL mapping and identification of candidate genes. Our study identified a new *Rf* gene (*Rf9*) and novel QTL for seed set on chromosomes 4AL, 5BL and 6AS and for number of sterile spikelets per spike on chromosomes 1DS, 2AL and 6AS. Putative candidate genes located in the target regions are discussed.

Materials and methods

Plant materials and population development

Two BC_1 mapping populations were developed using Gerek 79 and 71R1203 as fertility-restoring parental lines. The winter wheat cultivar Gerek 79 (PI 559560, pedigree: Mentana/Mayo-48//4-11/3/Yayla-305) originated in 1979 from the Transitional Zone Agricultural Research Institute, Anadolu ARI, Eskisehir, Turkey (http://wheatpedigree.net/sort/show/21747). The restoration capacity of Gerek 79 was found in initial screening experiments (unpublished) when pollinating CMS-Sperber with cultivars from various regions and testing the hybrids for self-fertility in the greenhouse at the Bavarian State Research Centre for Agriculture (LfL). The winter wheat restorer line 71R1203 (PI 473552, pedigree: NB542437/CI 13438//2*Burt/3/NB542437/2*CI13438) was developed in 1982 by USDA-ARS and Washington State University. The specific fertility restoration genes possessed by each of the two sources had not been previously determined, but 71R1203 was known to potentially carry restorer genes *Rf1* and *Rf2* that are present in NB542437 (Allan and Rubenthaler 1984). The variety Sperber (registered 1983) and CMS-Sperber are maintained at the LfL. Seeds of Gerek 79 were kindly provided by Prof. Friedrich Zeller (Technical University of Munich, Freising, Germany) and are available at the

Germplasm Resources Information Network (GRIN), U.S. National Plant Germplasm System. Seeds of line 71R1203 were obtained from National Small Grains Collection, US. Gerek 79 and 71R1203 were used as restorer lines in crosses with the male-sterile winter wheat line CMS-Sperber. The hybrids were then backcrossed with the maintainer line Sperber to develop the mapping populations CMS-Sperber/Gerek 79//Sperber (BC_1 Gerek 79) and CMS-Sperber/71R1203//Sperber (BC_1 71R1203).

Field trials and phenotyping

BC_1 Gerek 79 population was vernalised in a climate chamber at 6 °C for 8 weeks and planted in spring 2019 in an LfL field at Freising (48°24'12.64"N, 11°44'55.54"E), Germany. The BC_1 71R1203 population was sown in autumn 2018 in the field at KWS LOCHOW GMBH in Bergen (52°48'30.13"N, 9°57'49.46"E), Germany. To assess fertility restoration in the mapping populations, four emerging spikes from the main tillers of each plant were covered before anthesis using glassine bags. After ripening, the spikes were harvested and the seed set (as the restored fertility trait) and number of sterile spikelets per spike (as the non-restored fertility trait) were counted. The seed set of a plant was calculated as the number of kernels divided by the number of spikelets, averaged over all four bagged spikes per individual. Plants were considered fertile if they had at least one seed per spike and male sterile when no seed was produced. Observed ratios of fertile to sterile plants in each mapping population were tested against the expected segregation pattern using *chi-squared goodness*-of-*fit tests*. Statistical analyses including descriptive statistics, correlation and frequency distribution of the traits were conducted in SigmaPlot (Systat Software, San Jose, CA, USA).

Genotyping and linkage analysis

Genomic DNA of parental lines and BC_1 progenies was extracted from young leaf tissues following the procedure of Plaschke et al. (1995). Based on fertility restoration phenotypic data, the DNA of 273 and 184 individuals from BC_1 Gerek 79 and BC_1 71R1203, respectively, were selected for genotyping using a bead chip comprising 16,762 Single Nucleotide Polymorphism (SNP) markers selected from the 90K iSelect® array (Wang et al. 2014). SNP genotyping was done by KWS SAAT S & Co, KGaA, Einbeck, Germany. The raw SNP data were analysed as described by Geyer et al. (2018). Briefly, all

monomorphic SNPs and those with more than 10 % missing values and a minor allele frequency of less than 10 % were discarded from further analysis using the synbreed package V0.12-6 (Wimmer et al. 2012) in R (R Core Team, 2017). Linkage analysis was done using JoinMap® (Kyazma BV, Wageningen, The Netherlands). The Kosambi mapping function (Kosambi 1944) was used to convert recombination frequencies into centimorgans (cM).

To determine whether Gerek 79 and 71R1203 carried *Rf3*, they were genotyped with SNP marker *IWB72107*, earlier shown to have a high potential for predicting *Rf3* (Geyer et al. 2016). The SNP-containing sequence for *IWB72107* was retrieved from The Triticeae Toolbox (T3, https://triticeaetoolbox.org/wheat/) and converted to a Kompetitive Allele Specific Polymerase (KASP) chain reaction marker assay (Table 1). Plants were genotyped according to the manufacturer's instructions (LGC Genomics, Hoddeson, UK). Each KASP reaction was prepared in a volume of 10 µL with 5 µL DNA and 5 µL of the genotyping master mix. Amplification was carried out using the CFX96 Touch™ Real-Time PCR SNP Detection System (Bio-Rad, CA, USA), starting with 15 min at 94 °C, followed by 40 cycles of PCR with 94 °C for 20 s and 65 °C for 1 min and 10 cycles of touch down PCR where the annealing temperature was gradually reduced by 0.8 °C per cycle. Endpoint analysis and allelic discrimination related to SNP calls were accomplished using the CFX96 Touch™software (Bio-Rad, CA, USA). The DNA of the restorer line Primepi was used as a reference control for *Rf3* (Geyer et al. 2016).

QTL mapping

To detect QTL controlling seed set and number of sterile spikelets per spike in BC_1 populations, composite interval mapping with a 5-cM window and a maximum of 10 marker cofactors per model was carried out using Windows QTL Cartographer version 2.5 (Wang et al, 2012). Tests were performed at 1-cM intervals, and cofactors were selected by forward-backward stepwise regression Model 6 (Wang et al. 2012; Shahinnia et al. 2009). Genome-wide, trait-specific threshold values ($\alpha = 0.05$) of the likelihood ratio test statistic for declaring the presence of a significant QTL was predicted by random sampling of phenotypic data and conducting 2000 permutations test (Churchill and Doerge 1994). The additive effect of an allelic substitution at each QTL and the phenotypic variation explained by a QTL

(R^2) conditioned by the composite interval mapping cofactors involved in the model was calculated at the most likely QTL position. The LOD peak of each significant QTL was reflected as the QTL location on the linkage map. To identify markers associated with trait variation located in the confidence interval of a target QTL, single marker analysis was performed using Wald statistics (Kenward et al. 1997; Shahinnia et al. 2016). QTL designation followed the recommended rules for wheat (https://wheat.pw.usda.gov/ggpages/wgc/98/Intro.htm). QTL nomenclature (*Qphenotype.lab-chromosome.Qnumber*) included 'lfl', 'Rf' and 'StS' representing 'Bayerische Landesanstalt für Landwirtschaft' (LfL), seed set and number of sterile spikelets per spike, respectively. The *Qnumber* after the chromosome designation refers to overlapping QTL identified on the same chromosome in the two BC_1 populations.

Physical mapping and identification of candidate genes

Candidate genes were identified to better understand the genetic basis of fertility restoration. Sequences and physical positions of SNP markers within the QTL regions were retrieved based on the entire bread wheat NRGene genome assembly version 1.0 (https://wheat-urgi.versailles.inra.fr/Seq-Repository/Assemblies) (Appels et al. 2018) using the Triticeae Toolbox (T3, https://triticeaetoolbox.org/wheat/). Candidate genes were identified using web application POPSEQ Ordered *Triticum aestivum* Gene Expression (POTAGE) (Suchecki et al. 2017). The obtained sequences for SNP markers were then used as a query for BLASTn search in POTAGE (http://crobiad.agwine.adelaide.edu.au/potage/). The search by sequence option allowed convenient access to high confidence predictions of genes within the CSS contigs and functional annotation with the respective Munich Information Centre for Protein Sequences (MIPS) (ftp://ftpmips.helmholtz-muenchen.de/plants/wheat/IWGSC/genePrediction_v2.1/), supplemented with rice annotations translated sequence similarities (RGAP version 7). *In silico* expression values for tissue series of the wheat spike, root, leaf, grain and stem organs at different developmental stages (Zadoks et al. 1974) was obtained through the WheatExp (https://wheat.pw.usda.gov/WheatExp/) (Pearce et al. 2015) and a bread wheat tissue series RNA-Seq data set (http://urgi.versailles.inra.fr/files/RNASeqWheat/) in POTAGE.

Fragments per kb per million reads (FPKM) were used to show the gene expression quantity, thus avoiding the influence of sequencing length and differences on expression values.

Results

Evaluation of fertility restoration

The values for seed set and number of sterile spikelets per spike in BC_1 Gerek 79 and BC_1 71R1203 are presented in Table 2. Whereas a 1:1 segregation ratio for fertile to sterile lines was observed in BC_1 Gerek 79 (174:166) and BC_1 71R1203 (100:106), the average seed set was higher in BC_1Gerek 79 (0.5) than in BC_1 71R1203 (0.3). Number of sterile spikelets per spike showed a negative correlation with seed set in both BC_1 Gerek 79 (r = -0.65) and BC_1 71R1203 (r = -0.87). Frequency distribution of the traits ranged between 0-2.5 and 0-2.1 for seed set (Fig. 1a, b), whereas a range between 5.5-29.5 and 3.7-23.0 was observed for number of sterile spikelets per spike (Fig. 1c, d) in BC_1 Gerek 79 and BC_1 71R1203, respectively.

Construction of genetic maps

Following filtration of 16,762 SNPs used for genotyping of BC_1 lines, the resulting genetic base maps consisted of 929 and 994 unique SNP loci, spanning 2,160 and 2,328 cM over 21 linkage groups in BC_1 Gerek 79 and BC_1 71R1203, respectively. The average distance (2.4 cM) between two unique loci was similar in both linkage maps (Tables S1 & S2).

Using the categorical fertility phenotypes (completely sterile or fertile), a new restorer locus was mapped as a qualitative (monogenically inherited) trait between SNP markers *IWB72413* (4.3 cM) and *IWB1550* (4.7 cM) in the subtelomeric region of chromosome 6AS in BC_1 Gerek 79 (Fig. 2a). The newly dissected locus was designated *Rf9* following the Catalogue of Gene Symbols (https://shigen.nig.ac.jp/wheat/komugi/genes/symbolClassListAction.do?geneClassificationId=68) for Restorers for Cytoplasmic Male Sterility in wheat. No restorer locus underlying the binary phenotype could be genetically mapped in BC_1 71R1203.

Identification of QTL in BC_1 Gerek 79 and BC_1 71R1203

Composite interval mapping in BC_1 Gerek 79 (Table 3) detected two QTL for seed set on chromosomes 6AS (*QRf.lfl-6AS.1*) and 4AL (*QRf.lfl-4AL*) that explained 18 % and 14 % of the phenotypic variation, respectively. At both loci, the parental line Sperber contributed with negative additive effects indicating that the Gerek 79 alleles increased seed set values. In this population, three QTL for number of sterile spikelets per spike were identified on chromosomes 6AS (*QStS.lfl-6AS.1*), 6BS (*QStS.lfl-6BS*) and 2AL (*QStS.lfl-2AL*). Of these, *QStS.lfl-6AS.1*, located close to *IWB72428*, showed the highest LOD score (46.3) and explained 53 % of the total phenotypic variation for number of sterile spikelets per spike. The QTL allele that increased the number of sterile spikelets was inherited from the parental line Sperber (Table 3).

Four QTL for seed set in BC_1 71R1203 population (Table 3) were identified on chromosomes 1AS (*QRf.lfl-1AS*), 1BS (*QRf.lfl-1BS*), 5BL (*QRf.lfl-5BL*), and 6AS (*QRf.lfl-6AS.2*), explaining 12, 7, 5 and 11 % of the total phenotypic variation, respectively. A higher seed set was conferred by the 71R1203 allele at all loci. The most significant QTL for number of sterile spikelets per spike with a LOD score of 26.3 was detected on chromosome 1AS (*QStS.lfl-1AS*) near to *IWB7436* with a positive additive effect derived from Sperber. This QTL together with two QTL on chromosomes 1DS (*QStS.lfl-1DS*) and 6AS (*QStS.lfl-6AS.2*) explained 46, 4 and 12 %, respectively, of the total phenotypic variation for the trait (Table 3).

A QTL hotspot for seed set and number of sterile spikelets per spike was deciphered in the nearly 20 cM spanning interval of the common SNP markers *IWB66015-IWB26040* on chromosome 6AS in both mapping populations (Fig. 2a, b). In BC_1 Gerek 79, the same QTL were identified in the interval between *IWB66015* (1.8 cM) and *CAP11_c1594* (14.2 cM) for controlling seed set (*QRf.lfl-6AS.1*) and number of sterile spikelets per spike (*QStS.lfl-6AS.1*) with the opposite allelic effect of Sperber (-0.34 and 8.59, respectively) (Table 3). Remarkably, the peak of both QTL harboured *Rf9*, located 4.4 cM proximal to the subtelomeric region in BC_1 Gerek 79 (Fig. 2a). The magnitudes and directions of allelic effects at *Rf9* and SNP loci *IWB72413* (4.3 cM) and *IWB1550* (4.7 cM) showed a highly significant effect for seed set (Fig. 3a) and number of sterile spikelets per spike (Fig. 3b), with the favourable allele derived from Gerek 79. In BC_1 71R1203, the QTL for seed set and number of sterile spikelets per spike also

overlapped the *IWB66015-IWB26040* interval on chromosome 6AS (Fig. 2b) and revealed negative (-0.21 for *QRf.lfl-6AS.2*) and positive (2.73 for *QStS.lfl-6AS.2*) allelic effects of Sperber for seed set and number of sterile spikelets per spike, respectively.

Candidate genes associated with co-located QTL

To further investigate the chromosomal region associated with co-located QTL controlling the traits on chromosome 6AS, SNP markers *IWB66015* (1.8 cM) and *CAP11_c1594* (14.2 cM) flanking *QRf.lfl-6AS.1* and *QStS.lfl-6AS.1* were used to search for putative candidate genes (Table S3). The search resulted in 93 candidate genes physically located in this interval based on rice annotation translated sequence similarities (RGAP version 7). Among those, 17 genes belonged to the mitochondrial transcription termination factor (mTERF) family. Genes *Traes_6AS_60B99C6F3* encoded for proteins of the tetratricopeptide repeat (TPR) family, whereas *Traes_6AS_97F4D73A5* and *Traes_6AS_A5F88DE9A* (*Rf1*) encoded pentatricopeptide repeat (PPR) proteins. *In silico* expression analysis of candidate genes (Table S3) showed a wide range of expression for the genes in different organs and at three developmental stages. The gene *Traes_6AS_60B99C6F3* was highly expressed (0.98 FPKM) at Zadoks 39 (flag leaf ligule visible) only in spikes of wheat (Table S3). The expression (Fig. 4) of *Traes_6AS_97F4D73A5* in spikes was higher (1.16 FPKM) at Zadoks 32 (second node visible) whereas *Traes_6AS_A5F88DE9A* (*Rf1*) was highly expressed (1.99 FPKM) in spikes at Zadoks 65 (full flowering: 50 % of anthers matured). Their expression patterns indicated that they could have biological roles in spike and, more likely spikelet, development in wheat.

Discussion

Yield gains associated with heterosis in wheat have been achieved at a slower pace than in other major crops such as maize and rice. The lack of an efficient system for producing hybrid seed is the major bottleneck impairing the competitiveness of hybrid breeding over line-breeding. At the breeding level, it hinders efficient development of large numbers of test crosses required for development of heterotic pools. The CMS hybridisation system based on sterility induced by the cytoplasm of *T. timopheevii* has been proven as a potentially efficient strategy for hybrid seed production. However, further genes and

QTL controlling fertility restoration have to be identified and tagged with breeder-friendly molecular markers for a more efficient breeding process. In wheat, this was severely constrained by its large complex genome of about 17 Gb, until the recent release of a reference genome sequence (Appels et al. 2018). QTL for general fertility restoration in common wheat have been identified on chromosomes 1BS, 2AL, 2BS, 4BS and 6AS (Ahmed et al. 2001), 1BS, 5A and 7D (Zhou et al. 2005), 2DS (Dou et al. 2009), 1AS, 1BS and 6BS (Geyer at al. 2018) and 2DS, 4BS and 7AL (Yuan et al. 2020). In the present study, we performed QTL analyses in two populations segregating for restorer lines discovered in Gerek 79 and 71R1203 and identified a novel restorer locus on chromosome 6AS (*Rf9*) together with QTL on chromosomes 1AS, 1BS, 1DS, 2AL, 4AL, 5BL and 6BS.

Gerek 79 was the most widely grown winter wheat cultivar in Turkey in the 1990s. It is a tall wheat derived from a Turkish landrace 'Yala' which has resistance to common bunt caused by *Tilletia caries* and *T. foetida* (Bonjean et al. 2001). 71R1203, developed from a cross of winter wheat 'NB542437' was one the best restorer lines identified in field trials conducted in Eastern Washington during 1974-1978 (Allan and Rubenthaler 1984). In BC_1 populations derived from Gerek 79 and 71R1203, we found 12 QTL for seed set and number of sterile spikelets per spike. Four of these QTL (*QRf.lfl-6AS.1*, *QStS.lfl-6AS.1*, *QRf.lfl-6AS.2*, *QStS.lfl-6AS.2*) overlapped at the distal end of chromosome 6AS. Ahmed et al. (2001) detected a minor QTL for seed fertility on chromosome 6AS, but located much closer to the centromere using recombinant inbred lines derived from a cross of *T. aestivum* cv. 'Chinese Spring' and *T. spelta* var. *duhamelianum*.

It seems the new restorer locus *Rf9* that coincided with the QTL peaks of *QRf.lfl-6AS.1* and *QStS.lfl-6AS.1* in the subtelomeric region of chromosome 6AS (Fig. 2a) is new with no *Rf* gene previously reported in the region. Putatively annotated high confidence genes located between the flanking markers of the gene *Rf9* showed possible associations with candidate genes coding for TPR, PPR and mTERF proteins with higher expression patterns in the spikes (Table S3). A T-DNA insertion mutant OsAPC6 related to a TPR containing protein in rice caused abnormal mitotic divisions during megagametogenesis and led to reduced plant height and up to 45 % decreased seed set (Awasthi et al. 2012). The involvement of TPR-containing proteins in hybrid sterility was reported by Yu et al. (2016), who showed that down-regulation of ORF3 encoding a TPR domain-containing protein restored spikelet fertility in hybrid rice

by sporogametophyte allelic interaction. PPR genes have been cloned and well-characterised as essential components for fertility restoration in rice (Hu et al. 2012) and sorghum (Tang et al. 1996). Members of the mitochondrial transcription termination factor (mTERF) family are also involved in fertility restoration in cereals. Pan et al. (2019) identified the candidate gene *Zea mays small kernel 3 (Zmsmk3)*, which contained two mTERF motifs and was required for the intron splicing of mitochondrial nad4 and nad1 and kernel development. Genome-wide association studies in a multiparental mapping population in hybrid barley detected two mTERF proteins linked to the restorer locus *Rfm3* on the short arm of chromosome 6H (Bernhard et al. 2019). Functional association of the wheat *Rf9* locus and these other candidate genes deserves attention in future studies.

In BC_1 71R1203, a genome-wide QTL scan revealed two overlapping major QTL, *QRf.lfl-1AS* and *QStS.lfl-1AS*, within the interval *IWB7436* (1.6 cM) - *IWB31602* (22.9 cM) on chromosome 1AS (Fig. 2b). The identified QTL probably corresponded to *Rf1* in line R3, the first reported restorer of *T. timopheevii* CMS (Livers 1964). It was later located on chromosome 1A by monosomic analysis (Yen et al. 1969). Geyer et al. (2018) detected a QTL for seed set in association with fertility restoration *Rf1* locus in a BC_1 population derived from CMS-Sperber and line R3 restorer.

Composite interval mapping also detected minor QTL *QRf.lfl-4AL, QStS.lfl-2AL and QStS.lfl-6BS* in BC_1 Gerek 79 and *QRf.lfl-1BS, QRf.lfl-5BL* and *QStS.lfl-1DS* in BC_1 71R1203 (Table 3). Except for two QTL on chromosomes 1BS and 6BS, previously identified as minor or modifier loci for major loci *Rf3* and *Rf4* (Ahmed et al. 2001; Zhou et al. 2005; Geyer et al. 2018), we detected new QTL on chromosomes 1DS, 2AL, 4AL and 5BL (Table 3). A KASP assay developed based on the SNP *IWB72107* was used to determine whether QTL *QRf.lfl-1BS* could be associated with *Rf3* located on chromosome 1BS (Geyer et al. 2018). No genetic polymorphism at this SNP locus was found between the parental lines when compared to the marker genotype of Primepi as a reference for the *Rf3* allele (data are not shown), indicating that *Rf3* was not present in BC_1 Gerek 79 or BC_1 71R1203. A comparison of flanking SNP markers and genetic distances also showed that none of these minor QTL was linked to a region associated with a major *Rf* gene detected previously (Ma et al. 1995; Ahmed et al. 2001; Zhou et al. 2005; Stojałowski et al. 2013; Whitford et al. 2013; Lukaszewski 2017).

Both BC_1 populations exhibited moderate restoration potential leading to incomplete fertility in most of the individuals. Nevertheless, genetic components of restorer parental lines could improve fertility restoration when combined with other restorer loci. Robertson and Curtis (1967) identified modifier genes located on chromosomes 1B, 2A, 3D, 6A and 6B in Chinese Spring wheat that prevented full restoration, even in the presence of *Rf1* and *Rf2*. Geyer et al. (2018) reported a significant effect of modifier loci located on chromosomes 1BS and 6BL on the penetrance of *Rf1*, however; such interactions between major genes and modifier loci were proposed in several studies as essential for full fertility restoration (Ahmed et al. 2001, Zhou et al. 2005; Stojałowski et al. 2013; Würschum et al. 2017). Besides the need to modify the cleistogamous behavior of wheat to a more outcrossing system, the production of commercial hybrid seed will require a stable and efficient hybridization system comprising an easily pollinated female parent and a male parent with ability to completely rerstore fertility under most field conditions. Selection for male characteristics such as plant height, anther size, pollen viability and longevity, general and specific combining ability with the female parent(s) and genetic diversity is important for exploring the best pollinators. However; the emphasis should be on the integration of genetic factors that control fertility restoration comprising QTL and *Rf* genes for developing males through repeated backcrossing. Selection of males segregating for fertility restoration loci in *T. timopheevii* cytoplasm could be performed by test crossing and evaluation of the plants particularly under conditions that might inhibit the expression of fertility restoration such as heat and drought (Bonjean et al. 2001). To reach an acceptable level of hybrid fertility, a combination of two to three genes should be sufficient; therefore restorer lines with effective major QTL and genes such as *Rf1*, *Rf3*, and *Rf9* could be tested for maximum fertility restoration. Minor QTL for fertility restoration could be used for improving the female lines. Further, the development of functional and predictive molecular markers for tracking fertility restoration loci ultimately increases the efficiency of breeding processes and facilitates production of restorer parents. The results of our study could be used for developing new genetic resources of restorer lines and marker-assisted breeding in hybrid wheat. Also, it provides new insights into the genetic mechanisms controlling fertility restoration in wheat with the cytoplasm of *T. timopheevii* and lays the groundwork for gene characterization, cloning and manipulation in future fundamental studies.

References

Ahmed TA, Tsujimoto H, Sasakuma T (2001) QTL analysis of fertility-restoration against cytoplasmic male sterility in wheat. Genes Genet Syst 76:33–38. doi: 10.1266/ggs.76.33

Allan RE and Rubenthaler GL (1984) Registration of male fertility restoration wheat germplasm. Crop Sci 24(6):1220.

Appels R, Eversole K, Feuillet C, et al (2018) Shifting the limits in wheat research and breeding using a fully annotated reference genome. Science (80) 361. doi: 10.1126/science.aar7191

Awasthi A, Paul P, Kumar S, et al (2012) Abnormal endosperm development causes female sterility in rice insertional mutant OsAPC6. Plant Sci 183:167–174. doi: 10.1016/j.plantsci.2011.08.007

Bahl P, Maan S (1973) Chromosomal location of male fertility restoring genes in six lines of common wheat. Crop Sci 317–320

Bernhard T, Koch M, Snowdon RJ, et al (2019) Undesired fertility restoration in msm1 barley associates with two mTERF genes. Theor Appl Genet 132:1335–1350. doi: 10.1007/s00122-019-03281-9

Bohra A, Jha UC, Adhimoolam P, et al (2016) Cytoplasmic male sterility (CMS) in hybrid breeding in field crops. Plant Cell Rep 35:967–993. doi: 10.1007/s00299-016-1949-3

Bonjean AP, Angus WJ, van Ginkel M (2001) The world wheat book: A history of wheat breeding. Tec & Doc intercept Ltd. Lavoisier, Paris.

Castillo A, Atienza SG, Martin AC (2014) Fertility of CMS wheat is restored by two Rf loci located on a recombined acrocentric chromosome. J Exp Bot 65:6667–6677. doi: 10.1093/jxb/eru388

Castillo A, Rodríguez-Suárez C, Martín AC, Pistón F (2015) Contribution of chromosomes 1HchS and 6HchS to fertility restoration in the wheat msH1 CMS system under different environmental conditions. PLoS One 10:1–11. doi: 10.1371/journal.pone.0121479

Chen L, Liu YG (2014) Male sterility and fertility restoration in Crops. Annu Rev Plant Biol 65:579–606. doi: 10.1146/annurev-arplant-050213-040119

Churchill G, Doerge RW (1994) Empirical threshold values for quantitative trait mapping. Genetics 138:963–971. doi: 10.1534/genetics.107.080101

Dou B, Hou B, Xu H, et al (2009) Efficient mapping of a female sterile gene in wheat (Triticum aestivum L.). Genet Res 91:337–343. doi: 10.1017/S0016672309990218

Geyer M, Albrecht T, Hartl L, Mohler V (2018) Exploring the genetics of fertility restoration controlled by Rf1 in common wheat (Triticum aestivum L.) using high-density linkage maps. Mol Genet Genomics 293:451–462. doi: 10.1007/s00438-017-1396-z

Geyer M, Bund A, Albrecht T, et al (2016) Distribution of the fertility-restoring gene Rf3 in common and spelt wheat determined by an informative SNP marker. Mol Breed 36: 167–177. doi: 10.1007/s11032-016-0592-6

Gupta PK, Balyan HS, Gahlaut V, et al (2019) Hybrid wheat: past, present and future. Theor Appl Genet 132:2463–2483. doi: 10.1007/s00122-019-03397-y

Hu J, Wang K, Huang W, et al (2012) The rice pentatricopeptide repeat protein RF5 restores fertility in Hong-Lian cytoplasmic male-sterile lines via a complex with the glycine-rich protein GRP162. Plant Cell 24:109–122. doi: 10.1105/tpc.111.093211

Jiang Y, Schmidt RH, Zhao Y, Reif JC (2017) A quantitative genetic framework highlights the role of epistatic effects for grain-yield heterosis in bread wheat. Nat Genet 49(12):1741–1746. doi:10.1038/ng.3974

Kenward MG, Roger JH. (1997) Small sample inference for fixed effects from restricted maximum likelihood. Biometrics 53:983–97.

Keydel F (1973) Die Restoration der Fertilität in Weizenhybriden. Bayer Landwirtsch Jahrb 50:424–430

Kosambi DD (1944) The estimation of map distances from recombination values. Ann Eugen 12:172–175

Livers RW (1964) Fertility restoration and its inheritance in cytoplasmic male-sterile wheat. Science 144:420.

Longin CFH, Mühleisen J, Maurer HP, et al (2012) Hybrid breeding in autogamous cereals. Theor Appl Genet 125:1087–1096. doi: 10.1007/s00122-012-1967-7

Lukaszewski AJ (2017) Chromosomes 1BS and 1RS for control of male fertility in wheats and triticales with cytoplasms of Aegilops kotschyi, Ae. mutica and Ae. uniaristata. Theor Appl Genet 130:2521–2526. doi: 10.1007/s00122-017-2973-6

Ma ZQ, Sorrells ME, Zhao YH (1995) Inheritance and chromosomal locations of male fertility restoring gene transferred from Aegilops umbellulata Zhuk. to Triticum aestivum L. Mol Gen Genet 247:351–357. doi: 10.1007/BF00293203

Mukai Y, Tsunewaki K (1979) Basic studies on hybrid wheat breeding. Theor Appl Genet 54:153–160. doi: 10.1007/bf00263045

Pan Z, Ren X, Zhao H, et al (2019) A mitochondrial transcription termination factor, ZmSmk3, is required for nad1 Intron4 and nad4 Intron1 splicing and kernel development in maize. G3 Genes, Genomes, Genet 9:2677–2686. doi: 10.1534/g3.119.400265

Pearce S, Vazquez-Gross H, Herin SY, et al (2015) WheatExp: An RNA-seq expression database for polyploid wheat. BMC Plant Biol 15:1–8. doi: 10.1186/s12870-015-0692-1

Plaschke J, Ganal MW, Röder MS (1995) Detection of genetic diversity in closely related bread wheat using microsatellite markers. Theor Appl Genet 91:1001–1007

R Core Team, (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Robertson LD, Curtis BC (1967) Monosomic analysis of fertility restoration in common wheat (Triticum aestivum L.). Crop Sci 7:493–495.

Schnable PS, Wise RP (1998) The molecular basis of cytoplasmic male sterility and fertility restoration. Trends Plant Sci 3:175–180. doi: 10.1016/S1360-1385(98)01235-7

Shahinnia F, LeRoy J, Laborde B, et al (2016) Genetic association of stomatal traits and yield in wheat grown in low rainfall environments. BMC Plant Biol. doi: 10.1186/s12870-016-0838-9

Shahinnia F, Sayed-Tabatabaei BE, Sato K, et al (2009) Mapping of QTL for intermedium spike on barley chromosome 4H using EST-based markers. Breed Sci 59:383–390. doi: 10.1270/jsbbs.59.383

Sinha P, Tomar SMS, Vinod, et al (2013) Genetic analysis and molecular mapping of a new fertility restorer gene Rf8 for Triticum timopheevi cytoplasm in wheat (Triticum aestivum L.) using SSR markers. Genetica 141:431–441. doi: 10.1007/s10709-013-9742-5

Stojałowski S, Bobrowska A, Hanek M, Myśków B (2013) The importance of chromosomes from the sixth homeologic group in the restoration of male fertility in winter triticale with Triticum timopheevii cytoplasm. J Appl Genet 54:179–184. doi: 10.1007/s13353-013-0144-2

Suchecki R, Watson-Haigh NS, Baumann U (2017) POTAGE: A Visualisation Tool for Speeding up Gene Discovery in Wheat. Sci Rep 7:1–8. doi: 10.1038/s41598-017-14591-7

Tahir CM, Tsunewaki K (1969) Monosomic analysis of Triticum spelta var duhameliamum, a fertility-restorer for T. timopheevi cytoplasm. Jpn J Genet 44:1–9

Tang HV, Pring DR, Shaw LC, et al (1996) Transcript processing internal to a mitochondrial open reading frame is correlated with fertility restoration in male-sterile sorghum. Plant J 10:123–133

Wang S, Basten CJ, Zeng ZB (2012) Windows QTL Cartographer 2.5. Department of Statistics, North Carolina State University, Raleigh, NC.

Wang S, Wong D, Forrest K, et al (2014) Characterization of polyploid wheat genomic diversity using a high-density 90 000 single nucleotide polymorphism array. Plant Biotechnol J 12:787–796. doi: 10.1111/pbi.12183

Whitford R, Fleury D, Reif JC, et al (2013) Hybrid breeding in wheat: Technologies to improve hybrid wheat seed production. J Exp Bot 64:5411–5428. doi: 10.1093/jxb/ert333

Wimmer V, Albrecht T, Auinger H, et al (2012) synbreed : a framework for the analysis of genomic prediction data using R. Bioinformatics 28:2086–2087. doi: 10.1093/bioinformatics/bts335

Würschum T, Leiser WL, Weissmann S, Maurer HP (2017) Genetic architecture of male fertility restoration of Triticum timopheevii cytoplasm and fine-mapping of the major restorer locus Rf3 on chromosome 1B. Theor Appl Genet 130:1253–1266. doi: 10.1007/s00122-017-2885-5

Yen FS, Evans LE, Larter EH (1969) Monosomic analysis of fertility restoration in three restorer lines of wheat. Can J Genet Cytol 11:531–546

Yu Y, Zhao Z, Shi Y, et al (2016) Hybrid Sterility in Rice (Oryza sativa L.) Involves the. Genetics 203:1439–1451. doi: 10.1534/genetics.115.183848

Yuan SH, Bai JF, Guo HY, et al (2020) QTL mapping of male sterility-related traits in a photoperiod and temperature-sensitive genic male sterile wheat line BS366. Plant Breed 1–10. doi: 10.1111/pbr.12811

Zadoks JC, Chang, TT, Konzak CF (1974) A decimal code for the growth stages of cereals. Weed Res 14:415–421.

Zhao Y, Zeng J, Fernando R, Reif JC (2013) Genomic prediction of hybrid wheat performance. Crop Sci 53:802–810. doi: 10.2135/cropsci2012.08.0463

Zhou W, Kolb FL, Domier LL, Wang S (2005) SSR markers associated with fertility restoration genes against Triticum timopheevii cytoplasm in Triticum aestivum. Euphytica 141:33–40. doi: 10.1007/s10681-005-5067-5

Acknowledgments

The authors would like to thank Sabine Schmidt, Petra Greim, Mahira Duran and the working group Wheat and Oat Breeding Research of the Bavarian State Research Center for Agriculture for excellent technical assistance. We are thankful to Dr. Bianca Büttner, Dr. Günther Schweizer and Dr. Theresa Albrecht for helpful discussions. We gratefully acknowledge valuable support and cooperation from Dr. Julia Rudloff and Dr. Erhard Ebmeyer, KWS LOCHO, Germany in this project, especially for phenotyping the BC1 71R1203 population. We appreciate Prof. Robert McIntosh for providing critical information on the gene symbols and reviewing the manuscript. The project was supported by funds of the Federal Ministry of Food and Agriculture (BMEL) based on a decision of the Parliament of the Federal Republic of Germany via the Federal Office for Agriculture and Food (BLE) under the innovation support programme.

Conflict of interest

The authors declare that they have no conflicts of interest.

Ethical approval

This article does not contain any studies with human participants or animals performed by any of the authors.

List of Tables

Table 1 Primer sequences designed for *KASP_IWB72107* marker associated with the gene *Rf3*.

Table 2 Descriptive statistics for seed set and number of sterile spikelets per spike in BC_1 Gerek 79 and BC_1 71R1203. Segregation of the number of fertile *vs.* sterile plants was compared to a 1:1 ratio using the χ^2 tests.

Table 3 Chromosomal locations, map intervals, flanking markers, LOD scores, additive effects (A.E.) of Sperber and percentage of explained variance by QTL detected for seed set (*QRf*) and number of sterile spikelets per spike (*QStS*) in BC_1 Gerek 79 and BC_1 71R1203.

Supplementary Tables

Table S1 Position of the markers (cM) in 21 linkage groups of wheat in BC_1 Gerek 79 mapping population.

Table S2 Position of the markers (cM) in 21 linkage groups of wheat in BC_1 71R1203 mapping population.

Table S3 List of the candidate genes associated with co-located QTL *Q1Rf.lfl-6AS* and *Q1StS.lfl-6AS* including physical position, contig ID, MIPS annotation hit ID and description, rice annotation hit ID and description and expression analysis in wheat organs at different developmental stages according to Zadoks et al. (1974).

List of Figures

Fig. 1 Frequency distribution of phenotypes for (a, b) seed set and (c, d) number of sterile spikelets per spike in BC_1 Gerek 79 and BC_1 71R1203, respectively.

Fig. 2 Genetic position of the gene *Rf9* and co-location of QTL detected for seed (*QRf*, dark bars) and number of sterile spikelets per spike (*QStS*, light bars) on chromosome 6AS in (a) BC_1 Gerek 79 and (b) BC_1 71R1203. The peak of each QTL is shown with arrows. Correspondence interval of *IWB66015-IWB26040* common markers in two maps is shown with dot lines. A partial linkage map of the short arm of chromosome 6A is presented.

Fig. 3 Comparison of allele effects for *Rf9* locus and two flanking SNP loci (*IWB72413* and *IWB1550*) at the peak of *QRf.lfl-6AS.1* and *QStS.lfl-6AS.1* on (a) seed set and (b) number of sterile spikelets per spike in BC_1 Gerek 79. Dotted bars refer to the alleles of CMS-Sperber (AA) *vs*. restorer line Gerek 79 (Aa). T-tests based on the mean value of each trait showed highly significant differences ($p<0.001$) between the allele effects of CMS-Sperber compared to Gerek 79 parental lines for all loci.

Fig. 4 *In silico* expression analysis of the candidate genes *Traes_6AS_60B99C6F3* and *Traes_6AS_97F4D73A5* and *Traes_6AS_A5F88DE9A* (*Rf1*) encoded for tetratricopeptide repeat (TPR) and pentatricopeptide repeat (PPR) containing proteins, respectively in different wheat plant organs at different developmental stages according to Zadoks et al. (1974).

Table 1 Primer sequences designed for *KASP_IWB72107* marker associated with the gene *Rf3*.

Primer		**Sequences (5′-3′)**
IWB72107	Fam-labelled	GAAGGTGACCAAGTTCATGCTCGGCGTCCCAGTCCGCGT
	Hex-labelled	GAAGGTCGGAGTCAACGGATTGGCGTCCCAGTCCGCGC
	Common	CGACGACGAGGATGCCGAGTTT

Table 2 Descriptive statistics for seed set and number of sterile spikelets per spike in BC_1 Gerek 79 and BC_1 71R1203. Segregation of the number of fertile *vs.* sterile plants was compared to a 1:1 ratio using the χ^2 tests.

Trait/ Population	**No. lines**	**Mean**	**Max**	**Min**	**SE**	**No. fertile:sterile**	**χ^2 test (p< 0.01)**
BC_1 Gerek 79	340					174:166	0.76^{ns}
Seed set		0.5	2.5	0	0.03		
Number of sterile spikelets per spike		19.1	29.5	5.5	0.32		
BC_1 71R1203	206					100:106	0.77^{ns}
Seed set		0.3	2.1	0	0.02		
Number of sterile spikelets per spike		16.8	23	3.7	0.31		

ns= not significant

Table 3 Chromosomal locations, map intervals, flanking markers, LOD scores, additive effects (A.E.) of Sperber and percentage of explained variance by QTL detected for seed set (*QRf*) and number of sterile spikelets per spike (*QStS*) in BC_1 Gerek 79 and BC_1 71R1203.

Trait/ Population	QTL	Chromosome	Interval (cM)	Flanking markers	LOD	R^2 (%)	A.E.
BC_1 Gerek 79							
Seed set	*QRf.lfl-4AL*	4AL	48.6-50.7	*BS00062059-IWB74057*	3.1	14	-0.46
	QRf.lfl-6AS.1	6AS	1.8-14.2	*IWB66015-CAP11_c1594*	6.2	18	-0.34
Number of sterile spikelets per spike	*QStS.lfl-2AL*	2AL	46.7-47.3	*IWB12381-IWB53795*	3.1	9	2.07
	QStS.lfl-6AS.1	6AS	1.8-14.2	*IWB66015-CAP11_c1594*	46.3	53	8.59
	QStS.lfl-6BS	6BS	0-34	*IWA921-IWB12660*	14	21	6.35
BC_1 71R1203							
Seed set	*QRf.lfl-1AS*	1AS	1.6-20.7	*IWB7436-IWB28549*	7.2	12	-0.26
	QRf.lfl-1BS	1BS	0-10.2	*IWB69597-IWB12475*	4.4	7	-0.21
	QRf.lfl-5BL	5BL	133.1-147.1	*IWB33310-IWB26869*	3.6	5	-0.19
	QRf.lfl-6AS.2	6AS	9.0-23.5	*IWB3722-IWB47595*	5.9	11	-0.21
Number of sterile spikelets per spike	*QStS.lfl-1AS*	1AS	1.6-22.9	*IWB7436-IWB31602*	26.3	46	6.11
	QStS.lfl-1DS	1DS	2.5-21.8	*IWB11524-IWB59019*	3.6	4	2.29
	QStS.lfl-6AS.2	6AS	2.5-20.8	*IWB33595-IWB73288*	5.8	12	2.73