Usefulness of a Multiparent Advanced Generation Intercross Population With a Greatly Reduced Mating Design for Genetic Studies in Winter Wheat

Melanie Stadlmeier1,2*, Lorenz Hartl1 and Volker Mohler1,2

1 Bavarian State Research Center for Agriculture, Institute for Crop Science and Plant Breeding, Freising, Germany, 2 TUM School of Life Sciences Weihenstephan, Technical University of Munich, Freising, Germany

Multiparent advanced generation intercross (MAGIC) populations were recently developed to allow the high-resolution mapping of quantitative traits. We present a genetic linkage map of an elite but highly diverse eight-founder MAGIC population in common wheat (Triticum aestivum L.). Our MAGIC population is composed of 394 F6:8 recombinant inbred lines lacking significant signatures of population structure. The linkage map included 5435 SNP markers distributed over 2804 loci and spanning 5230 cM. The analysis of population parameters, including genetic structure, kinship, founder probabilities, and linkage disequilibrium and congruency to other maps indicated appropriate construction of both the population and the genetic map. It was shown that eight-founder MAGIC populations exhibit a greater number of loci and higher recombination rates, especially in the pericentromeric regions, compared to four-founder MAGIC, and biparental populations. In addition, our greatly simplified eight-parental MAGIC mating design with an additional eight-way intercross step was found to be equivalent to a MAGIC design with all 210 possible four-way crosses regarding the levels of missing founder assignments and the number of recombination events. Furthermore, the MAGIC population captured 71.7% of the allelic diversity available in the German wheat breeding gene pool. As a proof of principle, we demonstrated the application of the resource for quantitative trait loci mapping analyzing seedling resistance to powdery mildew. As wheat is a crop with many breeding objectives, this resource will allow scientists and breeders to carry out genetic studies for a wide range of breeder-relevant parameters in a single genetic background and reveal possible interactions between traits of economic importance.

Keywords: multiparental population, SNP, RIL, linkage map, recombination, wheat

INTRODUCTION

In plant breeding, the detection of quantitative trait loci (QTL) is no longer limited by the availability of genetic marker information and genotyping throughput (Mammadov et al., 2012; Chen et al., 2014; He et al., 2014; Unterseer et al., 2014; Cui et al., 2017), but rather by the genetic material employed (Flint-Garcia et al., 2003; Zhu et al., 2008; Asimit and Zeggini, 2010; Gibson, 2012). In an attempt to counteract this fact, nested association mapping (NAM) and multiparent advanced generation intercross (MAGIC) populations were
established. In crops, Mackay and Powell (2007) and Cavanagh et al. (2008) first discussed the latter population type. In MAGIC designs, multiple inbred founders are intercrossed several times in a well-defined order to combine the genetic material of all the founders in a single line (Cavanagh et al., 2008). This leads to highly diverse genotypes each with a unique mosaic of founder alleles. The higher number of parents and recombination events of a MAGIC population are clear advantages compared to a classical biparental population, while for both designs pedigree and genetic structure are well known. In association mapping (AM) panels, genetic diversity and recombination rates are higher than in MAGIC designs as these panels take advantage of a collection of diverse breeding lines. However, the main factor limiting AM studies is confounding due to population structure based on sampling effects (Flint-Garcia et al., 2003; Caldwell et al., 2006; Vilhjálmsson and Nordborg, 2013) which enhances the risk of detecting false positives (Ewens and Spielman, 2001; Dickson et al., 2010; Korte and Farlow, 2013). Therefore, MAGIC populations represent an intermediate to biparental crosses and diversity panels concerning substructure, allele diversity, the number of traits that can be investigated, resolution, and power (Rakshit et al., 2012; Pascual et al., 2016). To date, MAGIC populations were established in a wide range of various crops including rice (Bandillo et al., 2013), tomato (Pascual et al., 2015), fava bean (Sallam and Martinsch, 2015), maize (Dell’Acqua et al., 2015), barley (Sannemann et al., 2015), and sorghum (Ongom and Ejeta, 2018). In wheat, the first described MAGIC population was based on four founder genotypes (Huang et al., 2012). Further wheat MAGIC resources were generated using eight parents by Mackay et al. (2014) and Sannemann et al. (2018). A 16 founder wheat MAGIC population was established in the UK including elite and historical varieties (Fradgley et al., 2013). It is expected that a higher number of parents and initial crosses will result in a better dissection of complex traits (Huang et al., 2012). However, this implies more time and higher costs for population creation. So far, no study has been published investigating a MAGIC population design that involved a greatly reduced number of overall crossings.

Through intensive genetic analysis, we evaluated the impact of a MAGIC mating design with a greatly reduced number of overall crossings. First, we explored the genetic structure of our Bavarian MAGIC wheat population (BMWpop) for the validation of the crossing procedure. Second, we constructed a genetic linkage map, which we evaluated through population parameters and congruency with other maps. For comparative analysis, we used two other wheat MAGIC maps (Cavanagh et al., 2013; Gardner et al., 2016), one biparental map (Geyer et al., 2017), and the IWGSC RefSeq v1.0 (International Wheat Genome Sequencing Consortium, 2018). Recombination fraction, linkage disequilibrium, and founder probabilities were used to address population analysis parameters. Third, as a proof of principle, we demonstrated the application of the resource for QTL mapping using the trait seedling resistance to powdery mildew. The described genetic material will be shared with the scientific community upon request. It will enable scientists and breeders to carry out genetic studies for a wide range of economically relevant winter wheat breeding traits in a single genetic background and to reveal possible interactions.

**MATERIALS AND METHODS**

**Plant Material and Population Development**

Eight winter wheat lines ‘Event’ (A), ‘Format’ (B), ‘BAYP4535’ (C), ‘Potenzial’ (D), ‘Ambition’ (E), ‘Bussard’ (F), ‘Firl3565’ (G), and ‘Julius’ (H) were selected as founders of an eight-way MAGIC population based on the following criteria: (i) variation for disease resistance, quality, and agronomic traits, (ii) derivation from diverse breeding programs, and (iii) importance within the respective quality group (Table 1). All founders originated from wheat breeders in Germany except for the variety ‘Ambition’, which emerged from the Nordic Seed (Denmark) breeding program. To create F1 seed, one ear per plant was used in four two-way crosses (AB, CD, EF, and GH) (Figure 1). Out of each of the four two-way crosses, one randomly selected F1 seed was raised and two ears per plant were further mated to four-way crosses (ABCD and EFGH). Out of 32 available four-way crosses, sixteen independent plants were selected and crossed, to obtain eight F1 populations that involved four reciprocal cross combinations (ABCD, EF, and GH) (Figure 1). The four-way crosses were further hybridized with the four reciprocal eight-way crosses for establishing 16 eight-way intercross combinations (ABCD, EF, and GH). The term ‘MAGIC group 1–16’, as used hereinafter, refers to these 16 subpopulations. The F1 seeds of BMWpop were progressed to the F6 generation via single seed descent followed by two generations of bulk propagation in the field. The crossing procedure started in June 2009, and was completed in December 2010 followed by 5 years it took to self the plants to F6 recombinant inbred lines (RILs).

**DNA Extraction and Genotyping**

Ten individual primary leaves of each one of the 394 F6 RILs and all eight founders were harvested, pooled and freeze-dried. Genomic DNA was extracted according to the procedure described by Plaschke et al. (1995). All lines including the parents repeated twice were genotyped using the 15K + 5K Infinium® iSelect® array containing 17267 single nucleotide polymorphism (SNP) markers provided as a service by the company TraitGenetics (Gatersleben, Germany). The array combines markers from the 90K iSelect array (Wang et al., 2014) and the 820K Axiom array (Winfield et al., 2016). The population was further genotyped with a functional PCR marker for powdery mildew resistance allele Pm3a (Tommasini et al., 2006) since the parent ‘BAYP4535’ is known to carry that gene.

**Marker Filtering**

The filtering of the genotypic data was carried out with the synbreed package V0.12-6 (Wimmer et al., 2012) in R (R Core Team, 2017). First, all markers were excluded which were monomorphic for the population and for the founders. Further
TABLE 1 | Classification of founder lines of the eight-parent Bavarian MAGIC winter wheat population.

| Founder line | Quality | GPC | Yield | Resistance | Breeder | Registration period |
|--------------|---------|-----|-------|------------|---------|---------------------|
|              | Group^c | GPC | Yield | PM^f | LR^f | STB^f |
| Event^a      | E       | 6   | 7/5   | 3    | 4    | 6   | Saatzucht Josef Breun | 2009–2015 |
| Format^a     | A       | 8   | 6/6   | 5    | 5    | 4   | Saatzucht Schweiger    | 2007–2013 |
| BAYP4535^b   | C       | n.d.| 7/7   | 2    | 2    | 4   | LfL                   | Not released |
| Potenzial^a   | A       | 5   | 7/7   | 2    | 4    | 5   | Deutsche Saatveredelung | 2006–present |
| Ambition^a,b  | C       | 4   | 6/7   | 2    | 5    | 4   | Nordic Seed            | 2003–2012 |
| Bussard^a     | E       | 8   | 2/3   | 4    | 7    | 6   | KWS Lochow             | 1990–present |
| Firl3565^b    | A       | 4   | 6/7   | 4    | 5    | 4   | Saatzucht Firlbeck     | Not released |
| Julius^a      | B       | 4   | 8/8   | 3    | 3    | 3   | KWS Lochow             | 2008–present |

^aBased on the descriptive variety list of the Federal Plant Variety Office 2009 (www.bundessortenamt.de).
^bBased on the assessment of the Bavarian Research Center for Agriculture (LfL).
^cQuality group (bread-making): E (excellent), A (high), B (medium), and C (low/other use).
^dGrain protein content: 1 (low) – 9 (high).
^eGrain yield without growth regulator and fungicide treatment/at high production intensity: 1 (low) – 9 (high).
^fPM (Powdery mildew), LR (Leaf rust), STB (Septoria tritici blotch); rating score: 1 (no disease) – 9 (severe disease).

n.d., not determined.

markers were removed based on the following conditions: (i) missing data ≥ 5%, (ii) minor allele frequency ≤ 5%, (iii) duplicated segregation pattern (marker with the lowest number of missing data was kept), (iv) missing and heterozygous data in founders and (v) different allele calls between the double-genotyped founders. Markers exhibiting segregation distortion were left in the data set, as segregation distortion does not affect recombination fractions (Kjær et al., 1995).
Population Validation
The genotypic validation of BMWpop was based on a subset of 6717 markers (selection procedure see section ‘2.2 Marker Filtering’). The population structure was investigated with a principle coordinates analysis (PCoA) based on Roger’s distance matrix in R/ape V5.0 (Paradis et al., 2004). The kinship matrix K was calculated with the R package Genomic Association and Prediction Integrated Tool (GAPIT Tool 3.0) (Lipka et al., 2012) using the method of VanRaden (2008). Furthermore, we examined the proportion of heterozygote allele calls i.e., the combination of the heterozygosity of each plant and the segregation of the allele within the ten bulked plants of the F_{6:8} lines. The deviation of the observed segregation from the expected ratio of each marker was analyzed after setting all heterozygous data to missing. The multipoint probability of each locus that the observed genotype was inherited from one of the eight founders was calculated based on haplotype structure using the function ‘mpprob’ implemented in R/mpMap V2.0.2 (Huang and George, 2011). The threshold to determine the parental origin of an allele was set to 0.7. This threshold allowed us to compare the proportion of missing founder probabilities and the number of recombination events with other studies. The number of recombination events in the RILs of BMWpop was compared to the recombination events in the eight-founder MAGIC population NIAB 2015 (Gardner et al., 2016), the four-founder MAGIC design 9kMAGIC (Cavanagh et al., 2013), and the biparental backcross population L19 BC1 (Geyer et al., 2017). Information about the population size and type, the generation of the lines, and the genotyping platform used of the populations mentioned above is given in Table 2. Recombination events of RILs in all MAGIC populations were recognized by the change of the origin of parental alleles along the chromosome, whereas in the L19 BC1 lines the change of allele calls from heterozygote to homozygote along the chromosome and vice versa was assessed. Genome-wide recombination counts were available for 9kMAGIC (Huang et al., 2012), whereas for NIAB 2015 data were only published for chromosome 3B. The allelic diversity captured in BMWpop was compared to a panel of 524 common wheat breeding lines provided by six breeding companies (Geyer et al., 2016). This genetic library represents well the germplasm of the German wheat breeding pool and was genotyped with the Illumina® Infinium® 15 k SNP array representing a subset of markers of the 15K + 5K Infinium® iSelect® array.

Map Construction
The construction of the genetic map was based on the two R packages mpMap V2.0.2 (Huang and George, 2011) and mpMap2 V0.0.3.1,2. Unless otherwise stated, the functions used were from the mpMap package. Heterozygous calls were set to missing. A recombination fraction (rf) matrix between all pairs of markers was calculated using the function ‘estmateRF’ in the package mpMap2 at default settings. The estimation of the rf matrix was based on identical by descent probabilities given in Broman (2005). Markers were grouped to 200 linkage groups using ‘mpgroup’ function. The linkage groups were assigned to chromosomes based on the information of published maps (Cavanagh et al., 2013; Wang et al., 2014; Marcotuli et al., 2017). Within linkage groups, markers were ordered in two steps using two-point ordering: First, an overall path order was constructed to minimize total map length using the ‘mporder’ function; second, fine ordering was performed based on the overall path order using the ‘orderCross’ function in mpMap2. All linkage groups belonging to the same chromosome were merged based on the consensus map of Cavanagh et al. (2013). Finally, the validity of the marker order over the whole chromosome was examined using the R/qtl ripple function implemented in the ‘mporder’ function. Genetic map distances were calculated using Haldane mapping function using ‘computemap’.

Map Validation
Chromosome length and number of markers and loci were compared to the wheat genetic maps of NIAB 2015, 9kMAGIC, and L19 BC1 populations. The congruency of marker order was assessed in relation to the three above-mentioned genetic maps and to the physical marker positions of IWGSC RefSeq v1.0 (International Wheat Genome Sequencing Consortium, 2018) employing a set of markers common between BMWpop and the respective other population.

The rf heatmaps (R/lattice V0.20-35; Sarkar, 2008) were used to visualize marker order. Linkage disequilibrium (LD) was estimated as the squared correlation coefficient between markers using R/genetics V1.3.8.1 (Warnes et al., 2013). The extent of LD decay was analyzed similar to Breseghello and Sorrells (2006); a population-specific r^2 value was determined at the 95th percentile of the square root transformed r^2 value distribution for unlinked markers. r^2 values above this threshold were expected to be caused by genetic linkage. A least square regression (loess) curve was fitted to the LD decay estimation using a smoothing span parameter of 0.10 (R/stats V3.4.3, R Core Team, 2017).

QTL Mapping
Two independent trials were conducted for powdery mildew (PM) seedling resistance in the glasshouse under controlled conditions with a mean relative humidity of 67%, a temperature of 18°C, and 16 h of light per day. For each experiment, the population and the eight founders were sown in multi pot plates in a randomized complete block design with two replicates each. After the primary leaf emerged up to a length of 5 cm, the plants were inoculated with a B. graminis tritici isolate mixture by evenly shaking conidia of heavily infected wheat seedlings above the trial. The isolate mixture consisted of two isolates with virulence to the resistance genes Pm2, Pm4b, Pm5a, Pm6, and Pm8 (Mohler et al., 2013). Disease reaction was scored from 1 (no symptoms) to 9 (susceptible) 2 weeks after inoculation. The assessment was based on ten plants per genotype and replicate.

Phenotypic data were adjusted in R/lme4 V1.1-14 (Bates et al., 2014) based on the following model:

\[ y_{ijk} = \mu + g_i + l_j + g_i l_j + r_{kj} + e_{ijk} \] (1)
TABLE 2 | Description of BMWpop and reference populations and linkage map summary statistics.

| Population type | BMWpop | NIAB 2015 | 9kMAGIC | L19 BC1 |
|-----------------|--------|-----------|---------|----------|
| Generation      | F6:8   | F4:5      | F6      | BC1      |
| Population size | 394    | 643       | 1440    | 230      |
| SNP array       | 20k    | 80k       | 9k      | 20k      |
| Total marker number | 5436 | 18601     | 4300    | 2100     |
| Marker number A genome | 2185 | 6832      | 2055    | 967      |
| Marker number B genome | 2603 | 9388      | 1944    | 906      |
| Marker number D genome | 648  | 2381      | 301     | 227      |
| Min number/chromosome | 35   | 80        | 14      | 9        |
| Max number/chromosome | 483  | 2327      | 451     | 231      |
| Total chromosome length | 5230 | 5405      | 3522    | 2823     |
| Chromosome length A genome | 2012 | 2009      | 1420    | 1283     |
| Chromosome length B genome | 2070 | 2108      | 1405    | 1109     |
| Chromosome length D genome | 1149 | 1287      | 697     | 579      |
| Min chromosome length | 87   | 126       | 44      | 30       |
| Max chromosome length | 390  | 386       | 300     | 219      |
| Total loci number | 2804 | 4578      | 1813    | 1920     |
| Loci A genome    | 1126  | 1892      | 798     | 874      |
| Loci B genome    | 1327  | 2160      | 878     | 838      |
| Loci D genome    | 351   | 526       | 137     | 208      |
| Min loci/chromosome | 25   | 39        | 11      | 9        |
| Max loci/chromosome | 248  | 452       | 202     | 209      |
| Mean recombination event/RIL | 73  |            | 37      | 53       |
| Range recombination event/RIL | 49–100 | 1–80 | 23–126 |

where \( y_{ijk} \) is the trait observation, \( \mu \) is the overall mean, \( g_i \) is the fixed effect of genotype \( i \), \( l_j \) is the random effect of the trial \( j \), \( g_l_{ij} \) is the random interaction effect of genotype \( i \) with trial \( j \), \( r_{kj} \) is the random effect of replication \( k \) nested within trial \( j \), and \( e_{ijk} \) is the random residual error. To obtain variance components, the genotype was fitted as random. Heritability was estimated on a progeny mean basis according to Hallauer and Miranda (1981).

The QTL mapping was carried out using the mpMap package V2.0.2. Simple interval mapping was based on founder probabilities computed as mentioned above (2.4 Population validation). The function ‘mpIM’ was used for mapping. QTL were detected at a genome-wide significance threshold of \( \alpha < 0.001 \). The threshold was derived from an empirically null distribution with 1000 simulation runs similar to Churchill and Doerge (1994). All detected QTL of the base model were simultaneously fitted in a full model using the function ‘fit’. From this model fit, only QTL were kept with a p-value < 0.05 and the full model was fitted again to obtain additive founder effects relative to Julius and the phenotypic variance explained \( (R^2) \) by individual QTL. The QTL support interval (S.I.) was defined as the map distance in cM surrounding a QTL peak at a -log10(p) drop of \( \pm 1.0 \). The designation of QTL followed the recommended rules for gene symbolization in wheat (McIntosh et al., 2013).

Data and Material Availability
All raw genotypic data and the complete pedigree traceable to two-way crosses are publicly available at http://doi.org/10.14459/2018mp1435172. Seed of all 394 RILs of BMWpop and the eight founders are available for non-commercial use upon request from the Bavarian State Research Center for Agriculture (Freising, Germany).

RESULTS

Construction of the MAGIC Population
The crossing design of the MAGIC population was greatly simplified compared to the maximum number of possible crosses (Mackay et al., 2014); however, it involved an additional eight-way intercross step. Any developed line can be sourced to one of the 16 eight-way intercrosses (MAGIC group 1-16). The RILs built up MAGIC groups of 13 to 58 lines with an average size of 32 lines. Altogether 972 seeds, each representing a unique genotype, were generated during MAGIC crossing procedure. Five hundred sixteen lines were advanced to \( F_{6:8} \) generation and a number of 394 lines were selected based on sufficient seed availability for field trials and suitability for experiments in the agricultural environment.
Genotypic Data Analysis

Of the 17,267 SNP markers available on the 15K + 5K Infinium® iSelect® array, 11,426 markers (66.2%) were segregating in the population. After marker filtering 6,738 SNPs (58.8%) remained. Since founder lines were genotyped twice, a genotyping error rate of 0.1% could be observed. Twenty-two markers were found to be polymorphic in the population but not in the founders. These markers were also excluded from the data set, and a total number of 6,716 codominant SNP markers and one dominant PCR marker remained for further analysis. The percentage of missing marker data per line was in the range of 0.0–9.5%, with an overall mean of 0.3% (Figure 2A). The four lines 'BMW2329', 'BMW2429', 'BMW2151', and 'BMW2120' showed the highest proportion of missing allele calls (Figure 2A). The mean proportion of heterozygote allele calls of the bulk of ten plants per MAGIC RIL was 0.8%. RIL 'BMW2275' showed an unexpected high level of heterozygote allele calls of 19.3%, but also three of the four above-mentioned lines had higher proportions of heterozygote calls (Figure 2A). The deviation of the observed segregation from the expected ratio followed a normal distribution with a mean of -0.01 (Figure 2B). The range of the deviation was from -0.32 to 0.26. A total of 71.7% of markers segregating in a diversity panel described by Geyer et al. (2016) were also polymorphic in BMWpop.

Structure of the MAGIC Population

The first two principle coordinates of the PCoA accounted for 2.7 and 2.3% of the molecular variation, respectively (Figure 3). Individuals of MAGIC groups with the same paternal crossing partner in the 8-way × 8-way cross showed a weak clustering. The kinship showed that all lines were relatives and that there was no kinship structure present although lines built up some more related groups (Supplementary Figure S1). Thus, analyses revealed a mild population structure. Average congruency of the marker pattern between all lines was 62.8%. The range of the similarity of the marker pattern of two lines was from 52.2 to 97.6%. The genome-wide founder contributions in BMWpop were analyzed with haplotype probabilities using a threshold of 0.7. Based on the threshold setting, different proportions of founder probabilities per chromosome remained unknown (Figure 4). The average amount of missing founder contribution over the whole genome was 31.6%, ranging from 16.0% on chromosome 6B to 70.3% on chromosome 4D with average values of 26.0, 28.2, and 40.5% for the A, B, and D genomes, respectively. Generally, all eight parental genomes contributed to all chromosomes except for chromosome 4D. Based on an eight-founder crossing scheme as used in this study, each founder genome is expected to be represented by 12.5% in the population. In BMWpop, the average founder contribution varied from 7.0 to 9.6%. Generally, for the 18 chromosomes with higher explained contributions (Figure 4), the probability explained by the eight founders was quite equally distributed except for chromosome 2B. Here, the parent 'Firl3565' was underrepresented with 0.1%, whereas 'BAYP4535' was overrepresented with 17.8%.

Genetic Map Construction

In total, 5,436 markers were assigned to the 21 wheat chromosomes. Summary statistics are shown in Table 2. The complete genetic map of BMWpop is presented in Supplementary Table S1. The majority of markers of 40.2 and 47.9% were mapped to the A and B genomes, respectively. Only 11.9% of markers were located on the D genome. The number
of markers per chromosome varied between 35 on chromosome 4D and 483 on chromosome 5B. The total linkage map spanned 5230 cM. The A and B genomes contributed equally to the map length with 38.4 and 39.6%, respectively, whereas the D genome had only a share of 22.0%. The individual chromosome length ranged from 87.0 cM for chromosome 4D to 389.6 cM.
corresponds to approximately half of the length of the wheat consensus map (Somers et al., 2004) with a total length of ~2500 cM (Huang et al., 2012). BMWpop, 9kMAGIC, and L19 BC1 showed 71.6, 48.1, and 103.9% of the simulated average recombination number, respectively.

**Linkage Map Validation**

The overall heatmap of the recombination fraction matrix (Supplementary Figure S4) showed a clear diagonal. The analysis of individual chromosomes reflected the signatures for correctly mapped A, B, and D genomes (Supplementary Figure S5; Gardner et al., 2016). The A and B genome chromosomes showed a distinct diagonal with small-sized tightly linked blocks (Supplementary Figure S5, Chromosome 5, panel A). In some cases, the markers around the centromeric region were all closely linked leading to bigger-sized blocks (Supplementary Figure S5, Chromosome 7, panel A). An extreme case of this was seen for chromosome 6B (Supplementary Figure S5, Chromosome 6, panel B). This result confirmed the already described truncated map length around the centromeric region of this chromosome (see “Map Comparison”). The rf heatmaps of the D genome chromosomes (Supplementary Figure S5, Chromosome 7, panel C) showed a limited number of blocks.

Plots of LD showed a clear pattern of intra-chromosomal LD decay in BMWpop for the genome and for the individual chromosomes (Supplementary Figures S6, S7). The mean LD for the genome decreased to $r^2 < 0.2$ within 9.3 cM. Individual chromosome LD decay to $r^2 < 0.2$ ranged from 3.5 cM (chromosome 6D) to 18.5 cM (chromosome 7D). To consider LD between loci due to genetic linkage the $r^2$ critical population-specific threshold was 0.017. Thus for the genome, markers were assumed to be genetically linked within a distance of 68.6 cM.

**Powdery Mildew QTL**

The repeatability of PM trials was 81.3 and 81.9% with a correlation between the trials of 0.8 (Supplementary Table S2). The residuals of the adjusted means followed a normal distribution (Supplementary Figure S8A), a prerequisite for appropriate QTL mapping. The progeny mean was not significantly different ($p < 0.05$) from the parental mean (Supplementary Table S2). The adjusted means (Supplementary Table S3) did not follow a normal distribution but rather showed a bimodal distribution (Supplementary Figure S8B). The heritability estimate was high with 93.0% (Supplementary Table S2). Five QTL were detected in simple interval mapping explaining 72.5% of the total phenotypic variance (Table 3). The individual $R^2$ values ranged from 4.5 to 34.1%. The support interval for all QTL was between 0 and 9 cM. The strongest QTL $Qpm.jfl-1A$ explained 34.1% of the phenotypic variance and coincided with the functional marker for the $Pm3a$ gene. The resistance allele was inherited from ‘BAYP4535’, which is the only founder carrying this gene, and reduced disease severity by 3.6 grades on a scale of 1–9. The QTL on chromosomes 6B and 7A had also high -log(p) values and accounted for 17.4 and
18.3%, respectively, of the variation for PM. The alleles reducing disease severity at QPm.lfl-6B and QPm.lfl-7A were contributed by 'Event' and 'Ambition', respectively. The QTL QPm.lfl-1B explained 4.5% of the phenotypic variance with all founder effects being negative relative to the parent Julius, whereas four parental lines decreased disease severity for QPm.lfl-4A.

**DISCUSSION**

A full eight-founder MAGIC crossing design is built up of 28 two-way crosses, 210 four-way crosses, and 315 eight-way crosses (Mackay et al., 2014). Starting with just four two-way crosses, 32 four-way crosses and eight eight-way crosses our mating design was greatly reduced but it involved sixteen additional eight-way intercrosses to compensate for the low number of overall crosses. We evaluated possible effects of this crossing design on population structure, number of recombination events, genetic map construction, founder probabilities, and QTL mapping.

Our greatly simplified MAGIC design and the small population size of 394 RILs did not cause a considerable structure in the population as indicated by population structure and kinship analyses (Figure 1 and Supplementary Figure S1). This notwithstanding, we observed a negligible impact of the male eight-way intercrossing partner. Possible biased maternal effects (Donohue, 2009; Roach and Wulff, 1987) cannot be ruled out as in our simplified crossing design the parent Event was the ultimate female parent of all 394 RILs. Another parameter that could have been affected is the number of recombination events. Based on a simulation model, the number of recombination events per RIL in a population is expected to increase with the number of founders and the number of crosses involved (Huang et al., 2012; Ladejobi et al., 2016). The average number of 73 recombination events per line in BMWpop compared to 37 in a four-founder MAGIC population (Table 2) supported the assumption of the simulation model. Despite this, the experimentally derived number of recombination events per line in BMWpop was 1.6 and 1.8 recombination events per line per Morgan was detected on chromosome 3B for NIAB 2015 (Gardner et al., 2016) and BMWpop, respectively. Despite this observation, which is limited to a single chromosome, it could be still possible that in a simplified MAGIC mating design a lower number of unique recombinations is available. A simplified MAGIC design could also have led to lower accuracy in estimation of the rf matrix, which in turn would have adversely affected genetic map construction. Nevertheless, of 6717 markers passing quality control, 5436 markers were mapped to the 21 wheat chromosomes. The rf heatmaps (Supplementary Figures S4, S5), and the agreement of BMWpop map to other genetic maps and the IWGSC RefSeq v1.0 (Supplementary Figure S2) supported the appropriateness of marker ordering. Of the 1281 unmapped markers, the largest share of 49.3 and 36.0% was rejected due to too many missing values (> 20%) in the rf matrix and marker ordering, respectively. The trivialization of both criteria resulted in incorrect marker order and artificially increased map length. Therefore, to ensure accurate mapping in BMWpop, strict quality control of the rf matrix and extensive manual curation during the mapping process was carried out. Similar conclusions were reported by Gardner et al. (2016) who used an almost fully realized mating design for MAGIC population creation. A valid genetic map and a sufficient number of recombinations are prerequisites for the determination of the parental origin of an allele. Increasing the number of crosses and progenies in a MAGIC population would increase the possible number of recombination events and could reduce the amount of missing founder assignments. However, the genome-wide proportion of missing haplotype data of our greatly reduced mating design compared to the mating design of NIAB 2015 with 643 RILs was similar with 17.3 and 18.6%, respectively. Therefore, it appears that a higher number of crosses will not enhance the proportion of alleles assigned to parental origin. Since founder assignments are based on haplotypes, increasing the number of unique founder haplotypes would considerably improve founder assignment. This aim may be achieved for example by using a higher number of SNP markers, including multiallelic marker systems, or new approaches to construct the haplotypes (Gabriel et al., 2002; Pook et al., 2018). Additionally, to keep the error rate low the threshold to determine the parental origin of an allele needs to be set in the upper range strengthening missing founder assignments. Beside technical issues, there could be also genetic reasons for missing founder probabilities such as...
and breeders to carry out genetic studies for a wide range of breeder-relevant parameters in a single genetic background.

**AUTHOR CONTRIBUTIONS**

VM and LH conceived the study. MS conducted the experiments and analyzed the data. VM supervised the research. MS and VM drafted the manuscript. All authors read, edited, and approved the final manuscript.

**FUNDING**

This study was part of the ERA-CAPS project ‘An effector- and genomics-assisted pipeline for necrotrophic pathogen resistance breeding in wheat (EfectaWheat)’ funded by the German Research Foundation (DFG; Project No. 263641700).

**ACKNOWLEDGMENTS**

We deeply thank Sabine Schmidt and the working group Wheat and Oat Breeding of the Bavarian State Research Center for Agriculture for their excellent and accurate technical effort during MAGIC population construction. The authors also gratefully thank Petra Greim of the Bavarian State Research Center for Agriculture for her excellent assistance in the lab. We express deep gratitude to Chris-Carolin Schön and the staff of Plant Breeding, Technical University of Munich (TUM) for the outstanding opportunity of using computing power and office space. We thank Rohan Shah for helpful advice for the use of ‘mpMap2’ package. We thank the International Wheat Genome Sequencing Consortium (http://www.wheatgenome.org/) for allowing access to the IWGSC RefSeq v1.0 and we gratefully acknowledge Martin Ganal and Jörg Plieske of TraitGenetics for providing unpublished assignments of 15K + 5K Infinium® iSelect® array markers to physical positions based on the IWGSC RefSeq v1.0. The authors thank Beatrice Corsi, Keith Gardner, and James Cockram for the QTL mapping course on multiparental populations at NIAB. We thank the EfectaWheat Consortium for giving us all the time needed to carry out this study.

**SUPPLEMENTARY MATERIAL**

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2018.01825/full#supplementary-material

**REFERENCES**

Asimit, J., and Zeggini, E. (2010). Rare variant association analysis methods for complex traits. *Annu. Rev. Genet.* 44, 293–308. doi: 10.1146/annurev-genet-102209-163421

Bandillo, N., Raghavan, C., Muyco, P., Sevilla, M. A. L., Lobina, I. T., Dilla-Ermita, C., et al. (2013). Multi-parent advanced generation inter-cross (magic) populations in rice: progress and potential for genetics research and breeding. *Rice* 6:11. doi: 10.1186/1939-8433-6-11
Bates, D., Mächler, M., Bolker, B., and Walker, S. (2014). Fitting Linear Mixed-Effects Models Using lme4. Available at: http://arxiv.org/abs/1406.5823
Bresghello, F., and Sorrells, M. E. (2006). Association mapping of kernel size and milling quality in wheat (Triticum aestivum L.) cultivars. Genetics 172, 1165–1177. doi: 10.1534/genetics.105.054386
Broman, K. W. (2005). The genomes of recombinant inbred lines. Genetics 169, 1133–1146. doi: 10.1534/genetics.104.035212
Calderwood, K. S., Russell, J., Langridge, P., and Powell, W. (2006). Extreme population-dependent linkage disequilibrium detected in an inbreeding plant species. Hordeum vulgare. Genetics 172, 557–567. doi: 10.1534/genetics.104.038489
Cavanagh, C., Morell, M., Mackay, I., and Powell, W. (2008). From mutations to magic: resources for gene discovery, validation and delivery in crop plants. Curr. Opin. Plant Biol. 11, 215–221. doi: 10.1016/j.pbi.2008.01.002
Cavanagh, C. R., Chao, S., Wang, S., Huang, B. E., Stephen, S., Kiani, S., et al. (2013). Genome-wide comparative diversity uncovers multiple targets of selection for improvement in hexaploid wheat landraces and cultivars. Proc. Natl. Acad. Sci. U. S. A. 110, 8057–8062. doi: 10.1073/pnas.1217113110
Chen, H., Xie, W., He, H., Yu, H., Chen, W., Li, J., et al. (2014). A high-density SNP genotyping array for rice biology and molecular breeding. Mol. Plant 7, 541–553. doi: 10.1093/mps/sts135
Churchill, G. A., and Doerge, R. W. (1994). Empirical threshold values for quantitative trait mapping. Genetics 138, 963–971.
R Core Team, (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available at: https://www.R-project.org/
Cui, F., Zhang, N., Fan, X., Zhang, W., Zhao, C., Yang, L., et al. (2017). Utilization of a Wheat660K SNP array-derived high-density genetic map for high-resolution mapping of a major QTL for kernel number. Sci. Rep. 7:3788. doi: 10.1038/s41598-017-04028-6
Dell’Acqua, M., Gatti, D. M., Pea, G., Cuttonaro, F., Coppens, F., Magris, G., et al. (2015). Genetic properties of the magic maize population: a new platform for high definition QTL mapping in Zea mays. Genome Biol. 16:167. doi: 10.1186/s13059-015-0716-z
Dickson, S. P., Wang, K., Krantz, I., Hakonarson, J., and Goldstein, D. B. (2010). Rare variants create synthetic genome-wide associations. PLoS Biol. 8:e1000294. doi: 10.1371/journal.pbio.1000294
Donohue, K. (2009). Completing the cycle: maternal effects as the missing link of fertility restoration controlled by Rf1 in common wheat (Triticum aestivum L.) cultivars. Theor. Popul. Biol. 77, 102–107. doi: 10.1016/j.tpb.2009.09.005
Ewens, W. J., and Spielman, R. S. (2001). Locating genes by linkage and association. Nat. Rev. Genet. 2, 33–43. doi: 10.1038/35048900
Flint-Garcia, S. A., Thornsberry, J. M., and Buckler, E. S. (2003). Structure of recombination in maize. Theor. Popul. Biol. 64, 1059–1074. doi: 10.1016/S0040-5809(03)00058-5
Fradgley, N., Bentley, A. R., Gardner, K. A., Howell, P. J., Mackay, I., Scott, M., et al. (2017). Development and QTL mapping in a 16 founder wheat magic population: a new platform for high definition QTL mapping. Theor. Popul. Biol. 114, 215–221. doi: 10.1016/j.tpb.2016.01.002
He, J., Zhao, X., Laroche, A., Lu, Z.-X., Liu, H., and Li, Z. (2014). Genotyping-by-sequencing (GBS), an ultimate marker-assisted selection (MAS) tool to accelerate plant breeding. Front. Plant Sci. 5:484. doi: 10.3389/fpls.2014.00484
Huang, B. E., and George, A. W. (2011). R/mpmap: a computational platform for the genetic analysis of multiparent recombinant inbred lines. Bioinformatics 27, 727–729. doi: 10.1093/bioinformatics/btq119
Kjær, B., Jensen, J., and Giese, H. (1995). Quantitative trait loci for heading date and straw characters in barley. Genome 38, 1098–1104. doi: 10.1139/g95–146
Korte, A., and Farlow, A. (2013). The advantages and limitations of trait analysis with GWAS: a review. Plant Methods 9:29. doi: 10.1186/1746-4811-9-29
Ladejobi, O., Elderfield, J., Gardner, K. A., Gaynor, R. C., Hickey, J., Hibberd, J. M., et al. (2016). Maximizing the potential of multi-parental crop populations. Appl. Transl. Genom. 11, 9–17. doi: 10.1016/J.ATG.2016.10.002
Lipka, A. E., Tian, F., Wang, Q., Peiffer, J., Li, M., Bradbury, P. J., et al. (2012). GAIT: genome association and prediction integrated tool. Bioinformatics 28, 2397–2399. doi: 10.1093/bioinformatics/bts444
Mackay, I., and Powell, W. (2007). Methods for linkage disequilibrium mapping in crops. Trends Plant Sci. 12, 57–63. doi: 10.1016/JTPLANTS.2006.12.001
MamoDav, J., Aggarwal, R., Buyarapu, R., and Kumpatla, S. (2012). SNP markers and their impact on plant breeding. Int. J. Plant Genom. 2012, 1–11. doi: 10.1155/2012/728398
Mcintosh, R. A., Yamazaki, Y., Dubcovsky, J., Rogers, J., Morris, C., Appels, R., et al. (2013). Catalogue of Gene Symbols for Wheat. 12th International Wheat Genetic Symposium. Minato: Yokohama.
Mohler, V., Bauer, C., Schweizer, G., Kempf, H., and Harl, L. (2013). PM50: a new powdery mildew resistance gene in common wheat derived from cultivated emmer. J. Appl. Genet. 54, 259–263. doi: 10.1007/s13353-013-0158-9
Ongsombut, P. O., and Ejeta, G. (2018). Mating design and genetic structure of a multi-parent advanced generation intercross magic population. Cereal Res. Commun. 54, 135–139. doi: 10.1007/s12041-012-04028-6
Ongsomphot, P., and Eletas, M. K., et al. (2012). A multiparent advanced generation inter-cross population of common and durum wheat. J. Appl. Genet. 53, 357–374. doi: 10.1007/s13353-012-0158-9
Pascual, L., Albert, E., Sauvage, C., Duangjit, J., Bouchet, J. P., Bitton, F., et al. (2016). Dissecting quantitative trait variation in the resequencing era: mating design and genetic structure of a multi-parent advanced generation intercross magic population of sorghum (Sorghum bicolor (L.) Moench). G3 8, 331–341. doi: 10.1534/g3.117.300248
Paradis, E., Claude, J., and Strimmer, K. (2004). APE: analyses of phylogenetics and evolution in R language. Bioinformatics 20, 289–290. doi: 10.1093/bioinformatics/btg142
Pascual, L., Albert, E., Sauvage, C., Duangjit, J., Bouchet, J. P., Bitton, F., et al. (2016). Dissecting quantitative trait variation in the resequencing era: complementarity of bi-parental, multi-parental and association panels. Plant Sci. 242, 120–130. doi: 10.1016/JPLANTS.2015.06.017
Paradis, E., Desplat, N., Huang, B. E., Desgroux, A., Bruguier, L., Bouchet, J. P., et al. (2015). Potential of a tomato magic population to decipher the genetic control of quantitative traits and detect causal variants in the resequencing era. Plant Biotechnol. J. 13, 565–577. doi: 10.1111/pbi.12282
Plaschke, J., Galal, M. W., and Röder, M. S. (1995). Detection of genetic diversity in closely related bread wheat using microsatellite markers. Theor. Appl. Genet. 91, 1001–1007. doi: 10.1007/BF00223912
Roach, D. A., and Wulff, R. D. (1987). Maternal Effects in Plants. *Annu. Rev. Ecol. Syst.* 18, 209–235. doi: 10.1146/annurev.es.18.110187.001233

Sallam, A., and Marsch, R. (2015). Association mapping for frost tolerance using multi-parent advanced generation inter-cross (magic) population in faba bean (*Vicia faba* L.). *Genetica* 143, 501–514. doi: 10.1007/s10709-015-9848-z

Sannemann, W., Huang, B. E., Mathew, B., and Léon, J. (2015). Multi-parent advanced generation inter-cross in barley: high-resolution quantitative trait locus mapping for flowering time as a proof of concept. *Mol. Breed.* 35:86. doi: 10.1007/s11032-015-0284-7

Sannemann, W., Lisker, A., Maurer, A., Léon, J., Kazman, E., Cöster, H., et al. (2018). Adaptive selection of founder segments and epistatic control of plant height in the MAGIC winter wheat population WM-800. *BMC Genomics* 19:459. doi: 10.1186/s12864-018-4915-3

Sarkar, D. (2008). *Lattice: Multivariate Data Visualization with R.* New York, NY: Springer. doi: 10.1007/978-0-387-75969-2

Somers, D. J., Isaac, P., and Edwards, K. (2004). A high-density microsatellite consensus map for bread wheat (*Triticum aestivum* L.). *Theor. Appl. Genet.* 109, 1105–1114. doi: 10.1007/s00122-004-1740-7

Tommasini, L., Yahiaoui, N., Srichumpa, P., and Keller, B. (2006). Development of functional markers specific for seven Pm3 resistance alleles and their validation in the bread wheat gene pool. *Theor. Appl. Genet.* 114, 165–175. doi: 10.1007/s00122-006-0420-1

Unterseer, S., Bauer, E., Haberer, G., Seidel, M., Knaak, C., Ouzounova, M., et al. (2014). A powerful tool for genome analysis in maize: development and evaluation of the high density 600 k SNP genotyping array. *BMC Genomics* 15:823. doi: 10.1186/1471-2164-15-823

VanRaden, P. M. (2008). Efficient methods to compute genomic predictions. *J. Dairy Sci.* 91, 4414–4423. doi: 10.3168/jds.2007-0980

Vilhjálmsson, B. J., and Nordborg, M. (2013). The nature of confounding in genome-wide association studies. *Nat. Rev. Genet.* 14, 1–2. doi: 10.1038/nrg3382

Wang, S., Wong, D., Forrest, K., Allen, A., Chao, S., Huang, B. E., 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

Warnes, G., with contributions from Gorjanc, G., Leisch, F., and Man, M. (2013). *Genetics: Population Genetics. R Package Version 1.3.8.1.* Available at: https://cran.r-project.org/package=genetics

Wimmer, V., Albrecht, T., Auinger, H.-J., and Schön, C.-C. (2012). Synbreed: a framework for the analysis of genomic prediction data using R. *Bioinformatics* 28, 2086–2087. doi: 10.1093/bioinformatics/bts335

Winfield, M. O., Allen, A. M., Burridge, A. J., Barker, G. L. A., Benbow, H. R., Wilkinson, F. A., et al. (2016). High-density SNP genotyping array for hexaploid wheat and its secondary and tertiary gene pool. *Plant Biotechnol. J.* 14, 1195–1206. doi: 10.1111/pbi.12485

Zhu, C., Gore, M., Buckler, E. S., and Yu, J. (2008). Status and prospects of association mapping in plants. *Plant Genome J.* 1:5. doi: 10.3835/plantgenome2008.02.0089

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

*Copyright © 2018 Stadlmeier, Hartl and Mohler. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.*