1 **Development of a multiparent advanced generation intercross**
2 **(MAGIC) population for genetic exploitation of complex traits in**
3 ***Brassica juncea*: glucosinolate content as an example**

4

5 Tianya Wang[†], Wei Wan[†], Kunjiang Yu, Aimal Nawaz Khattak, Botao Ye, Renqin
6 Yang, Entang Tian^{1*}

7

8 ¹Agricultural College of Guizhou University, Guizhou University, Guiyang 550025,
9 China

10 * For correspondence: erictian121@163.com

11

12

13 [†]These authors contributed equally to the work.

14

15

16

17

18

19

20

21

22

23

24

Abstract

Multiparent advanced generation intercross (MAGIC) populations have recently been developed to allow the high-resolution mapping of complex quantitative traits. This article describes the development of one MAGIC population and verifies its potential application for mapping quantitative trait loci (QTLs) in *B. juncea*. The population was developed from eight founders with diverse traits and composed of 408 F₆ recombinant inbred lines (RILs). To develop one rapid and simplified way for using the MAGIC population, a subset of 133 RILs as the primary mapping population were genotyped using 346 intron-length polymorphism (ILP) polymorphic markers. The population lacks significant signatures of population structure that are suitable for the analysis of complex traits. Genome-wide association mapping (GWAS) identified three major glucosinolate (GSL) QTLs of *QGsl.ig01.1* on J01 for indole GSL (IG), *QGsl.atg09.1* on J09 and *QGsl.atg11.1* on J11 for aliphatic GSL (AG) and total GSL (TG). The candidate genes for *QGsl.ig01.1*, *QGsl.atg09.1* and *QGsl.atg11.1* are *GSH1*, *GSL-ALK* and *MYB28*, which are involved in converting glutamate and cysteine to γ -EC, the accumulation of glucoraphanin, and the whole process of AG metabolism, respectively. One effective method for association mapping of quantitative traits in the *B. juncea* MAGIC population is also suggested by utilization of the remaining 275 RILs and incorporation of the novel kompetitive allele specific PCR (KASP) technique. In addition to its QTL mapping purpose, the MAGIC population could also be potentially utilized in variety development by breeders.

Keywords: *Brassica juncea*; MAGIC population; QTL mapping; Population structure;
Plant breeding

Introduction

With the rapid development of next-generation sequencing technologies and the dramatic decrease in the cost of sequencing, the detection of quantitative trait loci (QTLs) in plant breeding is no longer limited by the availability of genetic marker information and genotyping throughput but rather by the genetic materials employed [1]. Traditional QTL mapping populations combine the genomes of two parents and can only capture a small snapshot of the genomic regions differing between the two parents. While the alternative of associating mapping has far greater diversity and sufficient power to detect genomic regions of interest, it needs a very large population and may have difficulty detecting rare alleles of importance [2].

In recent years, we have witnessed the rise of multiparent populations with a new type of experimental design, offering great advantages for genetic studies in plants. The multiparent advanced generation intercross (MAGIC) population is one of the most popular multiparent populations, in which the multiple inbred founders are intercrossed several times to combine all the genomes of the founders in a single line, resulting in a diverse population whose genomes are fine-scale mosaics of the contributions from all founders [3]. It generates a diverse population whose genomes recombine many times, which is suitable for high-resolution trait mapping [2]. In addition, MAGIC populations can also provide excellent materials for plant breeding

due to their features of high recombination and the resulting diverse phenotypic diversity. The term MAGIC was first coined in 2007 [4], and the first MAGIC population in plant species was developed in *Arabidopsis thaliana* (L.) [5]. Then, dozens of MAGIC populations were established in a wide range of various crops, including rice [6-10], tomato [11, 12], fava bean [13], cowpea [14], maize [15], barley [16, 17], strawberry [18], sorghum [19], and wheat [3, 20-23]. However, no MAGIC populations in Brassicaceae have been reported to our knowledge.

To verify the suitability of the newly developed *B. juncea* MAGIC population for gene mapping, several glucosinolate (GSL) traits were chosen for further association analysis using the population. GSLs were first discovered in mustard seeds during an exploration of the chemical origin of their sharp taste in the 17th century. In the Brassicaceae family, GSLs are the major secondary metabolites and can be synthesized in all species of this family via a three-part biosynthetic pathway from methionine, tryptophan and phenylalanine [24-26]. GSLs are mainly divided into aliphatic, indolic and benzyl GSLs in *Brassica* species. As important secondary metabolites, GSLs can be hydrolyzed by myrosinase to form different toxins, such as nitriles, isothiocyanates, and thiocyanate. These toxins have been widely studied in regard to their function in altering insect resistance [27, 28] and herbivory [29-31], nematode survival [32, 33], and microbial pathogens [34, 35]. More importantly, certain GSLs might have potential anticarcinogenic activity in mammals [36-38]. The importance of GSLs makes them a very special complex trait, and their metabolic mechanism is well studied in *A. thaliana*. More work is needed in *Brassica* crops in

regard to their polyploidy feature.

Brassica juncea exhibited better drought and heat tolerance, disease resistance, insect resistance and shattering resistance than *B. napus* [39-45]. In addition to its use as a condiment in Canada and China and as a vegetable in China, great efforts have been made to develop *B. juncea* as an alternative oilseed crop, especially in semiarid areas. In the present study, we try to develop and use one MAGIC population to exploit the complex traits in *B. juncea*. For genotype detection, we used ILP markers, which were designed in exons flanking the target intron, which is highly polymorphic, specific, codominant, neutral, convenient and reliable [46, 47]. ILP-type markers have been used for genotyping studies in many species, such as in rice, yellow mustard, foxtail millet, maize, tomato, *B. juncea*, *B. rapa* and *A. thaliana* [46-52]. The genotyping results are used to estimate the population structure, in linkage disequilibrium (LD) analysis, and in genome-wide association study (GWAS) analysis. The detected QTLs for GSLs demonstrate that the MAGIC population is an excellent platform for exploiting the complex traits in *B. juncea*.

Materials and Methods

Plant materials and field experiment

The eight founders of the MAGIC population were chosen from 34 *B. juncea* germplasm lines in our earlier field experiments [53]. Two rows were planted for each line of the 408 lines and the eight founders in the MAGIC population in 2018 and 2019, respectively. All the materials were planted in the field of Guizhou University,

Guiyang, China. Three self-pollinated plants were randomly chosen to harvest seeds for GSL content analysis in 2018 and 2019, respectively.

GSL component analysis

The 2-propenyl and 3-butenyl GSL contents of the mature seeds from each plot were analyzed following published methods [54, 55] with minor modifications. Each seed sample was crushed, and 200 mg of each sample was extracted twice with 2 ml boiling 70% methanol. The concentration of GSLs in the seeds was determined by high-performance liquid chromatography (Waters 2487/600/717) using the ISO9167-1 (1992) standard method. The AG is the sum of 2-propenyl GSL, 3-butenyl GSL and 4-pentenyl GSL. The TG is the sum of AG, IG and 4-hydroxy-3-indolylmethy GSL (HOI GSL).

ILP analysis

Genomic DNA was extracted from young leaves of eight founders and the 408 MAGIC lines using the modified sodium dodecyl sulfate method [56]. PCR of the ILP markers was carried out according to our previous studies [57, 58]. Each PCR (20 μ l) contained 1 $\times$ standard PCR buffer (NEB), 1 U of Taq polymerase (NEB), 0.25 μ M forward primer, 0.25 μ M reverse primer, 100 μ M each dNTP and 50 ng of genomic DNA in a total volume of 20 μ L. The PCR amplification consisted of an initial denaturation step at 94°C for 5 min; 35 cycles consisting of 94°C (45 sec), 55°C (45 sec) and 72°C (1 min); followed by termination at 72°C for 7 min. All PCR products were analyzed by electrophoresis in 2% agarose gels in 1 $\times$ tri-acetate-ethylene diamine tetra acetic acid buffer. Gels were visualized by staining in ethidium bromide

and photographed on a digital gel documentation system.

Population structure and LD analysis

PCA was used to assess the MAGIC population using the genotyping data, which was performed in SPSS 20.0. The first two principal components (PCs) were used to visualize the dispersion of the lines within a population in a graph. Pairwise similarity coefficients were calculated for all pairwise combinations within a population according to the method developed by (Nei and Li) implemented in the software TASSEL v5.0 [59]. The neighbor-joining (NJ) tree based on this genetic similarity matrix was constructed accordingly. To display the clustering feature, clustering analysis was conducted in R. The linkage disequilibrium (LD) analysis of the MAGIC population was performed by computing the r^2 values between pairs of ILP markers using Tassel v3.0 [55]. A significance threshold of $p < 0.001$ was used to screen the LD locus pairs, and the remaining pairs were not considered informative. The parameter r^2 was used to graphically represent the LD curves with the R software.

Marker-trait association analyses

Marker-trait association analyses between the three GSL quality traits of two consecutive years and the ILP genotyping data were conducted using a general linear model (GLM) with the TASSEL 3.0 software package [59]. The significance threshold for the associations between ILP markers and traits was set as $p < 0.001$ ($-\log_{10}(p) = 3$). The ILP markers prefixed with “At” were mapped on the chromosomes of the *B. juncea* genome according to Panjabi’s genetic map [52]. In addition, the ILP markers prefixed with “Bnap” and “Brap” were mapped on the chromosomes of the *B.*

juncea genome by blasting their originated DNA sequence with the *B. juncea* genome (Brassica_juncea_v1.5, http://brassicadb.org/brad/datasets/pub/Genomes/Brassica_juncea/V1.5/Bju.genome.fasta.gz) and assigned a genome location using a BLAST-like alignment tool [60].

Data availability

The MAGIC lines and founders are maintained by Institute of Oil Crops of Guizhou University (IOC-GZU) and are available upon request. S1 Figure contains supplemental figures and tables cited in this article. S1 Table contains the information of the distribution of ILP type markers in different chromosomes. S2 Table contains the data of GSL contents (umol/g) of the PAM-MAGIC population in 2018 and 2019. Supplemental File 1, File 2 and File 3 contain genotype, phenotype and population structure for each MAGIC line used for GWAS analysis, respectively.

Results

Development of the MAGIC population

The initial *B. juncea* MAGIC population was developed by intercrossing 8 founder lines that are diverse in agronomic and quality traits (Table 1). The first stage followed a half-diallel mating system by intermating the 34 Chinese native *B. juncea* germplasm lines, and 17 biparental crosses were performed. The agronomic and quality traits of the 34 germplasm lines were estimated in our earlier study [53], and the eight diverse lines and their derived 4 F₁s were selected for the next step. The resulting 4 F₁s were intercrossed to derive 4-way crosses for which 2 such 4-way

crosses integrated the four founders' genomes of ABCD and EFGH. The last stage involved intercrossing the 2 4-way crosses to derive one 8-way cross integrating the eight founders' genomes of ABCDEFGH. Finally, 600 seeds of the 8-way crosses were self-pollinated to produce the final MAGIC population including 408 F₆ RILs with single-seed descent methods. The crossing scheme for the *B. juncea* MAGIC populations is shown in Fig 1.

Amplification fragment polymorphisms of the ILP markers

To increase the working efficiency, the MAGIC population is divided into two subpopulations for primary association mapping (PAM) and subsequent fine association mapping (FAM). The PAM-MAGIC and the FAM-MAGIC populations include 133 and 275 randomly selected lines, respectively. In this study, the genetic structure of the PAM-MAGIC population was analyzed to estimate the genetic structure of the whole MAGIC population. A total of 1,272 ILP primers, 284 from *Arabidopsis thaliana* [52], 745 from *B. napus* and 243 from *B. rapa* available in the Potential Intron Polymorphism (PIP) database [46], were used to screen the eight founders for polymorphic primers.

Of the 1,272 ILP primers, 296 (23.3%) generated clear and scorable polymorphic bands, varying in size from 150 to 1250 bp, among the eight founders. The 296 primers were used for genotyping the PAM-MAGIC population and produced a total of 346 polymorphic markers for estimation of the subpopulation's genetic structure. The polymorphic markers are distributed evenly across the whole *B. juncea* genome

(S1 Table), as reported in our previous publications [57, 58, 61].

Population structure and LD analysis

Neighbor-joining tree and PCA methods are introduced to estimate the genetic structure of the PAM-MAGIC population. Fig 2A and 2B show a dendrogram from this cluster analysis. The eight founder cultivars are distributed almost evenly in the dendrogram, although 7H881 and SL105 were closer to each other than to the other cultivars. Fig 2C shows the proportion and the cumulative proportion of all the PCs derived from the PCA. The proportions of variance of PC1 and PC2 are 6.26% and 5.33%, respectively. The scatter plot made with PC1 and PC2 indicates that the distribution of the PAM-MAGIC lines is almost even and not clustered (Fig 2D).

LD analysis was conducted to explore the genome landscape of the PAM-MAGIC population for association studies. The squared Pearson correlation coefficients (r^2) and D' (measures of LD) were measured according to these 346 ILP polymorphic markers to describe the LD relationship between markers. The LD indicates a very low level in PAM-MAGIC lines (S1 Figure). A total of 59,686 possible pair combinations were obtained with an average of $r^2 = 0.034$. The distribution of r^2 values was concentrated in the < 0.1 interval that accounted for 95.3% of all pairwise coefficients, and 58.2% of the coefficients were < 0.01 . The distribution of D' values differed, with approximately 5% and 42.8% of the values being between 0.1 and 0.3. In addition, only 4.46, 5.16, and 8.11% of the coefficients (r^2) were revealed to be significant at $p < 0.0001$, $p < 0.001$ and $p < 0.01$, respectively.

Performance of phenotype and QTL mapping

To verify the suitability of our new *B. juncea* MAGIC population for gene mapping, as predicted from our genetic analysis, we conducted GWAS for the GSL traits of IG, AG and TG. Phenotypic dissection and depiction of genetic variability from a 2-year GSL dataset is summarized in S2 Table. The GSL contents were stable and significantly correlated for IG (0.987), AG (0.917) and TG (0.920) between 2018 and 2019. The broad-sense heritability of IG, AG and TG are 82.6%, 96.9% and 97.1%, respectively.

The associations between ILP markers and the GSL traits were analyzed with a naive general linear model (GLM) in Tassel 3.0 [59]. The markers from the same chromosomes with lower significance thresholds of $p < 0.001$ ($-\log_{10}(p)=3$) are used to analyze the real QTL region. To determine the confidence interval for the real QTL identified in this study, markers significantly associated with the GSL content that were colocated together and/or adjacent within intervals inferior to the LD decay distance of up and down region from the association peak were considered as being in the same QTL region. Finally, three QTL regions for IG, AG and TG in 2018 and 2019 were identified (Fig 3). Each QTL in one year is mapped onto Panjabi's genetic map [52] and the physical map of *B. juncea* (Fig 3). The *B. juncea* physical map was constructed by blasting the DNA sequence of ILP markers prefixed with BnapPIP and BrapPIP to the *B. juncea* genome. The overlapped QTL regions of At4g37510 ($R^2=20.2\%$ and 19.6%) and BnapPIP1698 ($R^2=21.1\%$ and 20.5%), for IG, are mapped

onto J01 and renamed *QGsl.ig01.1* (Table 2). The overlapped QTL regions of At5g66290 ($R^2=23.3$ and 22.0 for AG and $R^2=23.4$ and 21.1 for TG) and BnapPIP1437 ($R^2=10.1$ and 7.9 for AG and $R^2=10.3$ and 9.5 for TG), for AG and TG, are mapped onto J11 and renamed *QGsl.atg11.1* (Table 2). In addition, another overlapped QTL regions of At5g63905 ($R^2=22.9$ and 18.9 for AG and $R^2=23.0$ and 19.9 for TG) and BnapPIP1490 ($R^2=26.2$ and 25.0 for AG and $R^2=26.4$ and 24.1 for TG), for AG and TG, are mapped onto J09 and renamed *QGsl.atg09.1* (Table 2).

Candidate genes that may be associated with GSL synthesis

The physical regions of consensus *QGsl.ig01.1* on J01, *QGsl.atg11.1* on J11 and *QGsl.atg09.1* on J09 are located at 3,924,530-5,949,433 bp, 5,668,412-14,027,300 bp and 9,541,599-1,164,980 bp, respectively (Fig 3). The sequences of these two QTL regions were downloaded and utilized to predicate candidate genes using the annotated *B. juncea* reference sequence (*Brassica_juncea_v1.5*) [62]. The candidate gene of *QGsl.ig01.1* is *GSH1*, which is responsible for converting glutamate and cysteine to γ -EC [63-66]. For the consensus QTL of *QGsl.atg11.1*, the candidate gene is *Myb28*, which is involved in the whole process of aliphatic GLS (AG) metabolism *Myb28* [67-69]. Moreover, the candidate gene of *QGsl.atg09.1* is *GSL-ALK*, which is responsible for the accumulation of glucoraphanin [70, 71].

Discussion

Dozens of studies have reported the utilization of MAGIC populations in many crops,

such as rice and wheat. However, no such population has been reported for Brassicaceae. This paper provides another alternative for QTL mapping in addition to traditional biparental mapping populations and diverse germplasm populations in Brassicaceae. Within comparison to the MAGIC population, which has a higher number of parents and recombination events, the biparental population has limited genetic diversity and complexity. Even though the genetic complexity and diversity of the natural population is higher than that of the MAGIC population, its confounding population structure [72-74] enhances the risk of detecting false positives [75-77]. While the population structure of MAGIC is ignorable, the population structure obtained in this study is suitable for QTL mapping. Thus, to date, the MAGIC mapping population is the most ideal mapping population.

To verify the potential application of the MAGIC mapping population in mapping QTLs in *B. juncea*. We selected the three GSL traits of IG, AG and TG as candidates for association mapping. In this study, we mapped the three QTLs of *QGsl.ig01.1* for IG, *QGsl.atg11.1* and *QGsl.atg09.1* for AG and TG through association mapping. The three candidate genes are well studied in *Arabidopsis thaliana* and less studied in other plants. *GSH1* is reported to be responsible for converting glutamate and cysteine to γ -EC in *A. thaliana* [63-66]. *GSH1* is not specific to IG biosynthesis, which is involved in the biosynthesis pathway of all GSLs. Therefore, some other types of regulation, such as methylation, RNA and protein levels, might function and result in the phenotypic variation of IG content in the MAGIC population. As an important transcription factor, *Myb28* is specific for

regulating the whole process of aliphatic GLS (AG) metabolism [67-69]. Moreover, *GSL-ALK* corresponding to the QTL of *QGsl.atg09.1* is responsible for the accumulation of glucoraphanin [70, 71]. Specifically, *GSL-ALK* can convert 4-methyl sulfinyl butyl and 4-methyl sulfinyl pentyl to 3-butenyl GSLs and 4-pentenyl GSLs, respectively [65, 78, 79]. The 3-butenyl GSLs and 4-pentenyl GSLs are the main individual components of AG, and AG is the main component of TG. Therefore, *GSL-ALK* and *MYB28*, corresponding to QTLs of *QGsl.atg09.1* and *QGsl.atg11.1*, respectively, must be the key factors involved in the accumulation and phenotypic variation of AG and TG in *B. juncea* MAGIC population in the present study. Moreover, the IG content is also one of the individual components of AG and TG, while the QTL of *QGsl.ig01.1* for IG was not detected when performing association mapping analysis for AG and TG. The main reason should be the small ratio of the IG content in AG and TG. This work can provide useful information for the mechanism of GSL biosynthesis and molecular marker-assisted breeding (MAS); however, more work is needed to clone and verify the function of these two genes in the future.

The *B. juncea* MAGIC population provided a template for the development and utilization of this type of population in other species of Brassicaceae. MAGIC populations have been constructed because of their attractive combination of gene discovery and breeding utility for geneticists and breeders. For breeders, the MAGIC population showed wider phenotypic variation in almost all traits than the eight founders [8], as observed in this study, which suggests the formation of transgressive

segregants in the breeding direction. Interestingly, one breeding program in IRRI was started, which includes 700 lines from a total of 21 high-yielding lines [8]. Moreover, high-throughput SNP genotyping platforms are becoming cheap, and statistical approaches that can be applied to such populations are now available [8]. In this study, we used one subset of 133 randomly chosen RILs to estimate the population structure and conduct primary association mapping. The population was used to successfully map two QTLs for IND, AG and TG contents. The FAM-MAGIC population with the remaining 275 F_6 lines is not useless and could be used for fine mapping through the utilization of the kompetitive allele specific PCR (KASP) technique for more complex quantitative traits. The KASP technique has been developed and successfully used for the fine mapping of the *Cf-10* gene in tomato through the utilization of bulk sequencing and the subsequent use of the KASP technique [80]. The KASP technique is useful in QTL mapping, especially for joint analysis by fully utilizing the genetic recombinant information of multiple populations. For the association mapping of more complex agronomic traits, the marker density should be increased by the utilization of the resequencing technique in case some minor QTL information is missed in the future. Then, the MAGIC population could play a greater role in gene discovery and breeding in *B. juncea*.

Conclusion

In this study, we reported the first MAGIC population developed in Brassicaceae. The population structure and LD analysis indicate that the population is suitable for association mapping. To verify the usefulness of the MAGIC population, the

association mapping of IG, AG and TG were analyzed. Finally, one QTL of *QGsl.ig01.1* for IND and two overlapping QTLs of *QGsl.atg09.1* and *QGsl.atg11.1* for AG and TG were identified, all of which play an important role in GSL metabolism. Thus, the MAGIC population provides an excellent platform for gene discovery and breeding in *B. juncea*.

Acknowledgments

This work was funded by National Natural Science Foundation of China (Grant No. 31560422), Agricultural Science and Technology Support Program of Guizhou Province (Qiankehe zhicheng No. [2019]2396), Science and Technology Foundation of Guizhou Province of China (Grant No. Qiankehe J zi [2015]2052), Scientific Research Foundation for Returned Scholars, Ministry of Education of China (Grant No. Jiaowaisiliu [2015]1098), Construction Program of Biology First-class Discipline in Guizhou (Grant No. GNYL[2017]009).

References

1. Stadlmeier M, Hartl L, Mohler V: **Usefulness of a Multiparent Advanced Generation Intercross Population With a Greatly Reduced Mating Design for Genetic Studies in Winter Wheat.** *Frontiers in plant science* 2018, **9**(1825).
2. Huang BE, Verbyla KL, Verbyla AP, Raghavan C, Singh VK, Gaur P, Leung H, Varshney RK, Cavanagh CR: **MAGIC populations in crops: current status and future prospects.** *Theoretical and Applied Genetics* 2015, **128**(6):999-1017.
3. Cavanagh CR, Chao S, Wang S, Huang BE, Stephen S, Kiani S, Forrest K, Saintenac C, Brown-Guedira GL, Akhunova A *et al*: **Genome-wide comparative diversity uncovers multiple targets of selection for improvement in hexaploid wheat landraces and cultivars.** *Proceedings of the National Academy of Sciences of the United States of America* 2013, **110**(20):8057-8062.
4. Mackay I, Powell W: **Methods for linkage disequilibrium mapping in crops.** *Trends in plant science* 2007, **12**(2):57-63.

- 361 5. Kover PX, Valdar W, Trakalo J, Scarcelli N, Ehrenreich IM, Purugganan MD, Durrant C, Mott R:
362 **A Multiparent Advanced Generation Inter-Cross to Fine-Map Quantitative Traits in**
363 **Arabidopsis thaliana.** *PLoS genetics* 2009, 5(7):e1000551.
- 364 6. Bandillo N, Raghavan C, Muyco PA, Sevilla MAL, Lobina IT, Dilla-Ermita CJ, Tung C-W, McCouch
365 S, Thomson M, Mauleon R *et al*: **Multi-parent advanced generation inter-cross (MAGIC)**
366 **populations in rice: progress and potential for genetics research and breeding.** *Rice* 2013,
367 6(1):11.
- 368 7. Ogawa D, Nonoue Y, Tsunematsu H, Kanno N, Yamamoto T, Yonemaru J-i: **Discovery of QTL**
369 **Alleles for Grain Shape in the Japan-MAGIC Rice Population Using Haplotype Information.**
370 *G3: Genes/Genomes/Genetics* 2018, 8(11):3559-3565.
- 371 8. Ponce KS, Ye G, Zhao X: **QTL Identification for Cooking and Eating Quality in indica Rice**
372 **Using Multi-Parent Advanced Generation Intercross (MAGIC) Population.** *Frontiers in plant*
373 *science* 2018, 9(868).
- 374 9. Meng L, Zhao X, Ponce K, Ye G, Leung H: **QTL mapping for agronomic traits using**
375 **multi-parent advanced generation inter-cross (MAGIC) populations derived from diverse**
376 **elite indica rice lines.** *Field Crops Research* 2016, 189:19-42.
- 377 10. Meng L, Guo L, Ponce K, Zhao X, Ye G: **Characterization of Three Indica Rice Multiparent**
378 **Advanced Generation Intercross (MAGIC) Populations for Quantitative Trait Loci**
379 **Identification.** *The Plant Genome* 2016, 9(2).
- 380 11. Pascual L, Desplat N, Huang BE, Desgroux A, Bruguier L, Bouchet J-P, Le QH, Chauchard B,
381 Verschave P, Causse M: **Potential of a tomato MAGIC population to decipher the genetic**
382 **control of quantitative traits and detect causal variants in the resequencing era.** *Plant*
383 *biotechnology journal* 2015, 13(4):565-577.
- 384 12. Campanelli G, Sestili S, Acciarri N, Montemurro F, Palma D, Leteo F, Beretta M: **Multi-Parental**
385 **Advances Generation Inter-Cross Population, to Develop Organic Tomato Genotypes by**
386 **Participatory Plant Breeding.** *Agronomy* 2019, 9(3).
- 387 13. Sallam A, Martsch R: **Association mapping for frost tolerance using multi-parent advanced**
388 **generation inter-cross (MAGIC) population in faba bean (Vicia faba L.).** *Genetica* 2015,
389 143(4):501-514.
- 390 14. Huynh B-L, Ehlers JD, Huang BE, Muñoz-Amatriaín M, Lonardi S, Santos JRP, Ndeve A, Batieno
391 BJ, Boukar O, Cisse N *et al*: **A multi-parent advanced generation inter-cross (MAGIC)**
392 **population for genetic analysis and improvement of cowpea (Vigna unguiculata L. Walp.).**
393 *The Plant Journal* 2018, 93(6):1129-1142.
- 394 15. Dell'Acqua M, Gatti DM, Pea G, Cattonaro F, Coppens F, Magris G, Hlaing AL, Aung HH,
395 Nelissen H, Baute J *et al*: **Genetic properties of the MAGIC maize population: a new**
396 **platform for high definition QTL mapping in Zea mays.** *Genome Biology* 2015, 16(1):167.
- 397 16. Sannemann W, Huang BE, Mathew B, Léon J: **Multi-parent advanced generation inter-cross**
398 **in barley: high-resolution quantitative trait locus mapping for flowering time as a proof of**
399 **concept.** *Molecular Breeding* 2015, 35(3):86.
- 400 17. Jamali S, Mohammadi S, Sadeghzadeh B: **Association mapping for morphological traits**
401 **relevant to registration of barley varieties.** *Spanish Journal of Agricultural Research* 2017,
402 15:e0704.
- 403 18. Wada T, Oku K, Nagano S, Isobe S, Suzuki H, Mori M, Takata K, Hirata C, Shimomura K,
404 Tsubone M *et al*: **Development and characterization of a strawberry MAGIC population**

- 405 **derived from crosses with six strawberry cultivars. *Breeding science* 2017, 67.**
- 406 19. Ongom PO, Ejeta G: **Mating Design and Genetic Structure of a Multi-Parent Advanced**
- 407 **Generation Intercross (MAGIC) Population of Sorghum (*Sorghum bicolor* (L.) Moench). *G3: Genes/Genomes/Genetics* 2018, 8(1):331-341.**
- 408
- 409 20. Yao F, Zhang X, Ye X, Li J, Long L, Yu C, Li J, Wang Y, Wu Y, Wang J *et al*: **Characterization of**
- 410 **molecular diversity and genome-wide association study of stripe rust resistance at the**
- 411 **adult plant stage in Northern Chinese wheat landraces. In: *BMC genetics*. vol. 20; 2019: 38.**
- 412 21. Emma Huang B, Clifford D, Cavanagh C: **Selecting subsets of genotyped experimental**
- 413 **populations for phenotyping to maximize genetic diversity. *Theoretical and Applied Genetics***
- 414 **2013, 126(2):379-388.**
- 415 22. Rebetzke GJ, Verbyla AP, Verbyla KL, Morell MK, Cavanagh CR: **Use of a large multiparent**
- 416 **wheat mapping population in genomic dissection of coleoptile and seedling growth. *Plant***
- 417 ***biotechnology journal* 2014, 12(2):219-230.**
- 418 23. Scutari M, Howell P, Balding DJ, Mackay I: **Multiple Quantitative Trait Analysis Using**
- 419 **Bayesian Networks. *Genetics* 2014, 198(1):129-137.**
- 420 24. Kjær A, Olesen Larsen P: **Non-protein amino acids, cyano-genic glycosides and**
- 421 **glucosinolates. In: Geissman, T.A. (Ed.), *Specialist Periodical Reports*. The Chemical Society,**
- 422 **London. 1973:71-105.**
- 423 25. Hogge LR, Reed DW, Underhill EW, Haughn GW: **HPLC Separation of Glucosinolates from**
- 424 **Leaves and Seeds of *Arabidopsis thaliana* and Their Identification Using Thermospray Liquid**
- 425 **Chromatography/Mass Spectrometry*. *Journal of Chromatographic Science* 1988,**
- 426 **26(11):551-556.**
- 427 26. Halkier BA, Du L: **The biosynthesis of glucosinolates. *Trends in plant science* 1997,**
- 428 **2(11):425-431.**
- 429 27. Zhang Z, Ober JA, Kliebenstein DJ: **The gene controlling the quantitative trait locus**
- 430 **EPITHIOSPECIFIER MODIFIER1 alters glucosinolate hydrolysis and insect resistance in**
- 431 ***Arabidopsis*. *The Plant cell* 2006, 18(6):1524-1536.**
- 432 28. Kliebenstein D, Pedersen D, Barker B, Mitchell-Olds T: **Comparative analysis of quantitative**
- 433 **trait loci controlling glucosinolates, myrosinase and insect resistance in *Arabidopsis***
- 434 ***thaliana*. *Genetics* 2002, 161(1):325-332.**
- 435 29. Bidart-Bouzat MG, Kliebenstein DJ: **Differential levels of insect herbivory in the field**
- 436 **associated with genotypic variation in glucosinolates in *Arabidopsis thaliana*. *Journal of***
- 437 ***chemical ecology* 2008, 34(8):1026-1037.**
- 438 30. Mewis I, Tokuhisa JG, Schultz JC, Appel HM, Ulrichs C, Gershenzon J: **Gene expression and**
- 439 **glucosinolate accumulation in *Arabidopsis thaliana* in response to generalist and specialist**
- 440 **herbivores of different feeding guilds and the role of defense signaling pathways.**
- 441 ***Phytochemistry* 2006, 67(22):2450-2462.**
- 442 31. Li Q, Eigenbrode SD, Stringam GR, Thiagarajah MR: **Feeding and Growth of *Plutella xylostella***
- 443 **and *Spodoptera eridania* on *Brassica juncea* with Varying Glucosinolate Concentrations and**
- 444 **Myrosinase Activities. *Journal of chemical ecology* 2000, 26(10):2401-2419.**
- 445 32. Zasada IA, Ferris H: **Nematode suppression with brassicaceous amendments: application**
- 446 **based upon glucosinolate profiles. *Soil Biology and Biochemistry* 2004, 36(7):1017-1024.**
- 447 33. Lazzeri L, Tacconi R, Palmieri S: **In vitro activity of some glucosinolates and their reaction**
- 448 **products toward a population of the nematode *Heterodera schachtii*. *Journal of agricultural***

- 449 *and food chemistry* 1993, **1993 v.41 no.5**(no. 5):pp. 825-829.
- 450 34. Tierens KF, Thomma BP, Brouwer M, Schmidt J, Kistner K, Porzel A, Mauch-Mani B, Cammue
451 BP, Broekaert WF: **Study of the role of antimicrobial glucosinolate-derived isothiocyanates**
452 **in resistance of Arabidopsis to microbial pathogens.** *Plant physiology* 2001,
453 **125**(4):1688-1699.
- 454 35. Clay NK, Adio AM, Denoux C, Jander G, Ausubel FM: **Glucosinolate metabolites required for**
455 **an Arabidopsis innate immune response.** *Science* 2009, **323**(5910):95-101.
- 456 36. Verhoeven DTH, Verhagen H, Goldbohm RA, van den Brandt PA, van Poppel G: **A review of**
457 **mechanisms underlying anticarcinogenicity by brassica vegetables.** *Chemico-biological*
458 *interactions* 1997, **103**(2):79-129.
- 459 37. Faulkner K, Mithen R, Williamson G: **Selective increase of the potential anticarcinogen**
460 **4-methylsulphinybutyl glucosinolate in broccoli.** *Carcinogenesis* 1998, **19**(4):605-609.
- 461 38. Zhang Y, Talalay P: **Anticarcinogenic activities of organic isothiocyanates: chemistry and**
462 **mechanisms.** *Cancer Res* 1994, **54**(7 Suppl):1976s-1981s.
- 463 39. Wilson RA, Sangha MK, Banga SS, Atwal AK, Gupta S: **Heat stress tolerance in relation to**
464 **oxidative stress and antioxidants in Brassica juncea.** *J Environ Biol* 2014, **35**(2):383-387.
- 465 40. Kumar A, Singh DP: **Use of Physiological Indices as a Screening Technique for Drought**
466 **Tolerance in Oilseed Brassica Species.** *Annals of Botany* 1998, **81**(3):413-420.
- 467 41. Phutela A, Jain V, Dhawan K, Nainawatee HS: **Proline Metabolism Under Water Stress in the**
468 **Leaves and Roots of Brassica juncea Cultivars Differing in Drought Tolerance.** *Journal of*
469 *Plant Biochemistry and Biotechnology* 2000, **9**(1):35-39.
- 470 42. Srivastava SK: **Peroxidase and Poly-Phenol Oxidase in Brassica juncea Plants Infected with**
471 **Macrophomina phaseolina (Tassai) Goid. and their Implication in Disease Resistance.**
472 *Journal of Phytopathology* 1987, **120**(3):249-254.
- 473 43. Roy NN: **Interspecific transfer of Brassica juncea-type high blackleg resistance to Brassica**
474 **napus.** *Euphytica* 1984, **33**(2):295-303.
- 475 44. Østergaard L, Kempin SA, Bies D, Klee HJ, Yanofsky MF: **Pod shatter-resistant Brassica fruit**
476 **produced by ectopic expression of the FRUITFULL gene.** *Plant biotechnology journal* 2006,
477 **4**(1):45-51.
- 478 45. Liu Y, Wei W, Ma K, Darmency H: **Spread of introgressed insect-resistance genes in wild**
479 **populations of Brassica juncea: a simulated in-vivo approach.** *Transgenic Research* 2013,
480 **22**(4):747-756.
- 481 46. Yang L, Jin G, Zhao X, Zheng Y, Xu Z, Wu W: **PIP: a database of potential intron polymorphism**
482 **markers.** *Bioinformatics* 2007, **23**(16):2174-2177.
- 483 47. Wang X, Zhao X, Zhu J, Wu W: **Genome-wide Investigation of Intron Length Polymorphisms**
484 **and Their Potential as Molecular Markers in Rice (Oryza sativa L.).** *DNA Research* 2005,
485 **12**(6):417-427.
- 486 48. Javidfar F, Cheng B: **Single locus, multiallelic inheritance of erucic acid content and linkage**
487 **mapping of FAE1 gene in yellow mustard.** *Crop Sci* 2013, **53**(3):825-832.
- 488 49. Muthamilarasan M, Venkata Suresh B, Pandey G, Kumari K, Parida SK, Prasad M: **Development of 5123**
489 **Intron-Length Polymorphic Markers for Large-Scale Genotyping**
490 **Applications in Foxtail Millet.** *DNA Research* 2014, **21**(1):41-52.
- 491 50. Liu H, Lin Y, Chen G, Shen Y, Liu J, Zhang S: **Genome-scale identification of resistance gene**
492 **analogs and the development of their intron length polymorphism markers in maize.**

- 493 *Molecular Breeding* 2012, **29**(2):437-447.
- 494 51. Wang Y, Chen J, Francis DM, Shen H, Wu T, Yang W: **Discovery of intron polymorphisms in**
495 **cultivated tomato using both tomato and Arabidopsis genomic information.** *Theoretical and*
496 *Applied Genetics* 2010, **121**(7):1199-1207.
- 497 52. Panjabi P, Jagannath A, Bisht N, Padmaja K, Sharma S, Gupta V, Pradhan A, Pental D:
498 **Comparative mapping of Brassica juncea and Arabidopsis thaliana using intron**
499 **polymorphism (IP) markers: homoeologous relationships, diversification and evolution of**
500 **the A, B and C brassica genomes.** *BMC genomics* 2008, **9**(1):1-19.
- 501 53. Tian E, Li L, Jia S, Lin S: **The variation of fatty acids of Brassica juncea germplasm, correlation**
502 **analysis and design of specific primer of FAE1 gene for the control of erucic acid.** *Guihaia*
503 2016, **36**(12).
- 504 54. Agerbirk N, Olsen CE, Nielsen JK: **Seasonal variation in leaf glucosinolates and insect**
505 **resistance in two types of Barbarea vulgaris ssp. arcuata.** *Phytochemistry* 2001,
506 **58**(1):91-100.
- 507 55. Wathelet J-P, Iori R, Leoni O, Rollin P, Quinsac A, Palmieri S: **Guidelines for glucosinolate**
508 **analysis in green tissues used for biofumigation.** *Agroindustria* 2004, **3**.
- 509 56. Somers DJ, Friesen KRD, Rakow G: **Identification of molecular markers associated with**
510 **linoleic acid desaturation in Brassica napus.** *TAG Theoretical and applied genetics*
511 *Theoretische und angewandte Genetik* 1998, **96**(6):897-903.
- 512 57. Tian E, Zeng F, MacKay K, Roslinsky V, Cheng B: **Detection and Molecular Characterization of**
513 **Two FAD3 Genes Controlling Linolenic Acid Content and Development of Allele-Specific**
514 **Markers in Yellow Mustard (Sinapis alba).** *PloS one* 2014, **9**(5):e97430.
- 515 58. Tian E, Roslinsky V, Cheng B: **Molecular marker-assisted breeding for improved Ogura cms**
516 **restorer line (RfoRfo) and mapping of the restorer gene (Rfo) in Brassica juncea.** *Molecular*
517 *Breeding* 2014, **34**(3):1361-1371.
- 518 59. Bradbury PJ, Zhang Z, Kroon DE, Casstevens TM, Ramdoss Y, Buckler ES: **TASSEL: software for**
519 **association mapping of complex traits in diverse samples.** *Bioinformatics* 2007,
520 **23**(19):2633-2635.
- 521 60. Kent WJ: **BLAT--the BLAST-like alignment tool.** *Genome research* 2002, **12**(4):656-664.
- 522 61. Javidfar F, Cheng B: **Construction of a genetic linkage map and QTL analysis of erucic acid**
523 **content and glucosinolate components in yellow mustard (Sinapis alba L.).** *BMC plant*
524 *biology* 2013, **13**:142.
- 525 62. Yang J, Liu D, Wang X, Ji C, Cheng F, Liu B, Hu Z, Chen S, Pental D, Ju Y et al: **The genome**
526 **sequence of allopolyploid Brassica juncea and analysis of differential homoeolog gene**
527 **expression influencing selection.** *Nature Genetics* 2016, **48**:1225.
- 528 63. Hofberger JA, Lyons E, Edger PP, Chris Pires J, Eric Schranz M: **Whole genome and tandem**
529 **duplicate retention facilitated glucosinolate pathway diversification in the mustard family.**
530 *Genome biology and evolution* 2013, **5**(11):2155-2173.
- 531 64. Urbancsok J, Bones AM, Kissen R: **Arabidopsis mutants impaired in glutathione biosynthesis**
532 **exhibit higher sensitivity towards the glucosinolate hydrolysis product allyl-isothiocyanate.**
533 *Scientific reports* 2018, **8**(1):9809.
- 534 65. Sønderby IE, Geu-Flores F, Halkier BA: **Biosynthesis of glucosinolates – gene discovery and**
535 **beyond.** *Trends in plant science* 2010, **15**(5):283-290.
- 536 66. Bednarek P, Pislewska-Bednarek M, Svatos A, Schneider B, Doubek J, Mansurova M,

- 537 Humphry M, Consonni C, Panstruga R, Sanchez-Vallet A *et al*: **A glucosinolate metabolism**
538 **pathway in living plant cells mediates broad-spectrum antifungal defense.** *Science* 2009,
539 **323**(5910):101-106.
- 540 67. Gigolashvili T, Berger B, Flügge U-I: **Specific and coordinated control of indolic and aliphatic**
541 **glucosinolate biosynthesis by R2R3-MYB transcription factors in Arabidopsis thaliana.**
542 *Phytochemistry reviews* 2009, **2009 v.8 no.1**(no. 1):pp. 3-13.
- 543 68. Schweizer F, Fernandez-Calvo P, Zander M, Diez-Diaz M, Fonseca S, Glauser G, Lewsey MG,
544 Ecker JR, Solano R, Reymond P: **Arabidopsis basic helix-loop-helix transcription factors MYC2,**
545 **MYC3, and MYC4 regulate glucosinolate biosynthesis, insect performance, and feeding**
546 **behavior.** *The Plant cell* 2013, **25**(8):3117-3132.
- 547 69. Seo MS, Kim JS: **Understanding of MYB Transcription Factors Involved in Glucosinolate**
548 **Biosynthesis in Brassicaceae.** *Molecules* 2017, **22**(9).
- 549 70. Augustine R, Bisht NC: **Biofortification of oilseed Brassica juncea with the anti-cancer**
550 **compound glucoraphanin by suppressing GSL-ALK gene family.** *Scientific reports* 2015,
551 **5**:18005.
- 552 71. Bisht NC, Gupta V, Ramchiary N, Sodhi YS, Mukhopadhyay A, Arumugam N, Pental D, Pradhan
553 AK: **Fine mapping of loci involved with glucosinolate biosynthesis in oilseed mustard**
554 **(Brassica juncea) using genomic information from allied species.** *TAG Theoretical and*
555 *applied genetics Theoretische und angewandte Genetik* 2009, **118**(3):413-421.
- 556 72. Flint-Garcia SA, Thornsberry JM, IV ESB: **Structure of Linkage Disequilibrium in Plants.** *Annu*
557 *Rev Plant Biol* 2003, **54**(1):357-374.
- 558 73. Caldwell KS, Russell J, Langridge P, Powell W: **Extreme Population-Dependent Linkage**
559 **Disequilibrium Detected in an Inbreeding Plant Species, Hordeum vulgare.** *Genetics* 2006,
560 **172**(1):557-567.
- 561 74. Vilhjálmsson BJ, Nordborg M: **The nature of confounding in genome-wide association**
562 **studies.** *Nature Reviews Genetics* 2012, **14**:1.
- 563 75. Ewens WJ, Spielman RS: **Locating Genes by Linkage and Association.** *Theoretical Population*
564 *Biology* 2001, **60**(3):135-139.
- 565 76. Dickson SP, Wang K, Krantz I, Hakonarson H, Goldstein DB: **Rare variants create synthetic**
566 **genome-wide associations.** *Plos Biology* 2010, **8**(1):e1000294.
- 567 77. Korte A, Farlow A: **The advantages and limitations of trait analysis with GWAS: a review.**
568 *Plant methods* 2013, **9**(1):29.
- 569 78. Zhang X, Liu T, Duan M, Song J, Li X: **De novo Transcriptome Analysis of Sinapis alba in**
570 **Revealing the Glucosinolate and Phytochelatin Pathways.** *Frontiers in plant science* 2016,
571 **7**:259.
- 572 79. Petersen A, Wang C, Crocoll C, Halkier BA: **Biotechnological approaches in glucosinolate**
573 **production.** *Journal of integrative plant biology* 2018.
- 574 80. Liu G, Zhao T, You X, Jiang J, Li J, Xu X: **Molecular mapping of the Cf-10 gene by combining**
575 **SNP/InDel-index and linkage analysis in tomato (Solanum lycopersicum).** *BMC plant biology*
576 **2019, 19**(1):15.

Tables and Figures

Table 1 Features of the eight founder lines in developing the *B. juncea* MAGIC population.

Code	Name	Features
P _A	7H881	larger seed size, low number of grains per pod, low total GSL content
P _B	YufengZC	larger seed size, shorter, high total GSL content
P _C	K100	taller, early flowering, small seed size,
P _D	SL105	taller, early flowering, small seed size, long length of main inflorescence, thinner stem perimeter
P _E	SL14	thinner stem perimeter, taller
P _F	ShilunYC	taller, more number of primary branches, bigger number of grains per pod
P _G	LengjiaoYC	shorter, small seed size
P _H	T6342	shorter, smaller number of grains per pod

Table 2 The QTLs obtained from association mapping of IG, AG and TG.

GSLs	QTL Name	Markers	R ² (%) ¹	-LOG ₁₀ P	Chr ²	Position ³	Consensus QTL name	Candidate Genes
IG	IG01-18	At4g37510	20.2	7.5	J01	6.4cM	<i>QGsl.ig01.1</i>	<i>GSH1</i> [63-66]
	IG01-19	At4g37510	19.6	7.3	J01	6.4cM		
	IG01-18'	BnapPIP1698	21.1	7.9	J01	4,290,074bp		
	IG01-19'	BnapPIP1698	20.5	7.7	J01	4,290,074bp		
AG	AG11-18	At5g66290	23.3	8.5	J11	51.2cM	<i>QGsl.atg11.1</i>	<i>Myb28</i> [67-69]
	AG11-19	At5g66290	22.0	8.1	J11	51.2cM		
	AG11-18'	BnapPIP1437	10.1	3.7	J11	1,500,610bp		
	AG11-19'	BnapPIP1437	7.9	3.0	J11	1,500,610bp		
	AG09-18	At5g63905	22.9	8.1	J09	30.1cM	<i>QGsl.atg09.1</i>	<i>GSL-ALK</i> [70, 71]
	AG09-19	At5g63905	18.9	6.8	J09	30.1cM		
	AG09-18'	BnapPIP1490	26.2	9.8	J09	10,280,466bp		
	AG09-19'	BnapPIP1490	25.0	9.4	J09	10,280,466bp		
TG	TG11-18	At5g66290	23.4	8.5	J11	51.2cM	<i>QGsl.atg11.1</i>	<i>Myb28</i> [67-69]
	TG11-19	At5g66290	21.1	7.8	J11	51.2cM		
	TG11-18'	BnapPIP1437	10.3	3.7	J11	1,500,610bp		
	TG11-19'	BnapPIP1437	9.5	3.5	J11	1,500,610bp		
	TG09-18	At5g63905	23.0	8.2	J09	51.2cM	<i>QGsl.atg09.1</i>	<i>GSL-ALK</i> [70, 71]
	TG09-19	At5g63905	19.9	7.2	J09	51.2cM		
	TG09-18'	BnapPIP1490	26.4	9.9	J09	10,280,466bp		
	TG09-19'	BnapPIP1490	24.1	9.1	J09	10,280,466bp		

¹: Percentage of the total phenotypic variation explained by the QTL.

²: Abbreviation of chromosomes.

³: The position from the genetic map with a postfix of cM [52] and the physical position with a postfix of bp obtained by blasting the DNA sequence of ILP markers against the *B. juncea* genome [60].

594

595 **Fig 1 The scheme and flow of the MAGIC population development.** Eight

596 founders: A: 7H881, B: YufengZC, C: K100, D: SL105, E: SL14, F: ShilunYC, G:

597 LengjiaoYC, H: T6342.

598

599

600

601

602

603

604

605

606

607

608

609

610

611

612

613

614

Fig 2 Population structure analysis through clustering (A and B), PCA (C) and a scatter plot (D).

Fig 3 QTLs detected on *B. juncea* chromosomes J01, J09 and J11 for IG, AG and TG. The physical maps of J01', J09' and j11' were constructed by blasting the sequence of PIP markers against the *B. juncea* genome. The chromosomes of J01, J09 and J11 are based on the published genetic map [1], and the physical position of each marker was deduced according to the published *B. juncea* genome [2]. The physical positions of the consensus QTLs were marked on different chromosomes using different colors. The unit of the ruler on the left is Mb.

Supplementary materials

S1 Table The distribution of ILP type markers in different chromosomes.

Chromosomes	ILP markers from <i>B. napus</i> and <i>B. rapa</i> ¹	ILP markers from <i>A. thaliana</i> ²	Total
J01	39	8	47
J02	16	10	26
J03	44	14	58
J04	28	3	31
J05	36	6	42
J06	38	6	44
J07	22	4	26
J08	42	10	52
J09	46	14	60
J10	31	4	35
J11	29	5	34
J12	33	16	49
J13	36	5	41
J14	30	6	36
J15	41	2	43
J16	27	4	31
J17	35	5	40
J18	38	5	43
Total	611	127	738 ³

¹: The marker position of these markers prefixed "Bnap and Brap" is mapped onto *B. juncea* genome by the BLAST-like alignment tool [60].

²: The marker position of these markers prefixed "At" is from the published *B. juncea* genetic map [52].

³: Some primers are mapped onto more than one chromosome.

S2 Table Glucosinolate contents (umol/g) of the PAM-MAGIC population in 2018 and 2019.

Traits	Year	Mean	SD ¹	Range	CV (%) ²
IG	2018	0.62	0.51	0.10~2.62	82.26
IG	2019	0.56	0.51	0.00~2.62	91.07
AG	2018	101.33	33.54	28.98~197.91	33.10
AG	2019	103.32	38.67	28.48~218.76	37.43
TG	2018	106.75	33.71	33.91~204.44	31.58
TG	2019	106.79	38.22	31.07~211.82	35.79

¹: SD: standard deviation.

²: CV: coefficient of variation, CV=SD/Mean

S1 Figure Linkage disequilibrium (LD) analysis for exploring the genome landscape of the PAM-MAGIC population for association studies. The squared Pearson correlation coefficients (r^2 , A) and D' (measures of LD, B) are measured and analyzed (C).