

Identification of novel loci and candidate genes for cucumber downy mildew resistance using GWAS

Xiaoping Liu

Chinese Academy of Agricultural Sciences Institute of Vegetables and Flowers

Hongwei Lu

Chinese Academy of Agricultural Sciences Institute of Vegetables and Flowers

Panna Liu

Chinese Academy of Agricultural Sciences Institute of Vegetables and Flowers

Han Miao

Chinese Academy of Agricultural Sciences Institute of Vegetables and Flowers

Zifu Yan

University of Agricultural and Horticultural Sciences

Yuling Bai (✉ baiyuling@wur.nl)

Wageningen University and Research Wageningen Plant Research

Xingfang Gu (✉ guxingfang@caas.cn)

Chinese Academy of Agricultural Sciences Institute of Vegetables and Flowers

Shengping Zhang (✉ zhangshengping@caas.cn)



Chinese Academy of Agricultural Sciences Institute of Vegetables and Flowers

Research article

Keywords: cucumber, downy mildew, loci, GWAS technology

Posted Date: July 16th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-41473/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Abstract

Background

Downy mildew (DM) is one of the most serious diseases in cucumber and brings the loss of yield and profit. Multiple QTLs for DM resistance have been detected, however, no loci related to resistance was reported using genome-wide association analysis (GWAS). In this study, the core germplasm (CG) of cucumber lines that had been constructed and resequenced were used to identify DM resistance Loci using GWAS technology.

Results

Thirteen loci (*dmG1.1*, *dmG1.2*, *dmG2.1*, *dmG2.2*, *dmG3.1*, *dmG4.1*, *dmG4.2*, *dmG5.1*, *dmG5.2*, *dmG6.1*, *dmG6.2*, *dmG7.1* and *dmG7.2*) associated with DM resistance were detected on all the seven chromosomes. Among these loci, *dmG2.1* and *dmG7.1* were novel loci compared with previous studies. Based on the annotation of homologous genes in *Arabidopsis* and pairwise LD correlations, *Csa1G575030* could be the most likely candidate gene of *dmG1.2*; *Csa2G059820* and *Csa2G060360* could be the candidate gene of *dmG2.1*. A WRKY transcription factor *Csa5G606470* could be the candidate gene of *dmG5.2*. *Csa7G004020* could be the candidate gene of *dmG7.1*.

Conclusions

These results identify five candidate genes for four loci related to DM resistance in cucumber which provide theoretical basis for gene cloning and genetic breeding of DM resistance in cucumber.

Introduction

Cucumber (*Cucumis sativus* L.) is an economically important vegetable crop grown all over the world. Downy mildew (DM) is one of the most serious diseases and causes significantly production lost, it has been found all over the world (Perez-Garcia et al. 2009; Thomas C E, 1986). The symptoms of DM mainly appeared on the upper of leaf surfaces. Although there are many studies on DM that have been reported, its mechanism was not clear and there were few genes been identified.

Downy mildew (DM) is caused by the obligate biotrophic oomycete *Pseudoperonospora cubensis* [(Berkeley & M.A. Curtis) Rostoyzev]. DM was occurred in many species, such as *pearl millet* [1], *Arabidopsis thaliana* [2–4], *Triticum aestivum* [5, 6], *Helianthus annuus* L.[7, 8] and *Cucumis melo*[9]. Downy mildew was also devastating in cucumber. Many researches on downy mildew in cucumber have been reported [10–12]. Zhang et al. (2013) constructed the F₂ and F_{2:3} population based on the cross of inbred line K8 and K18 and identified 5 QTLs (*dm1.1*, *dm5.1*, *dm5.2*, *dm5.3* and *dm6.1*) associated with DM on the Chr1, Chr5 and Chr6, respectively. And the QTLs in Chr5 could also be detected by other studies (Bai et al. 2008). By using F₂ population constructed by the cross between resistance line IL52 and susceptible line 'Changcunmici', Peng

et al. (2013) detected two QTLs (*DM_5.1* and *DM_5.2*). Yoshioka et al. (2014) found two QTLs on Chr5 by using RILs population derived from CS-PMR1 and the old Japanese cultivar Santou. Four QTLs (*dm2.1*, *dm4.1*, *dm5.1*, and *dm6.1*) were detected by the cross between DM-resistant inbred line WI7120 and susceptible '9930', and *dm4.1* and *dm5.1* were major effect QTLs [13]. Five QTLs (*dm2.2*, *dm4.1*, *dm5.1*, *dm5.2*, and *dm6.1*) were found using NGS-assisted BSA, and *dm2.2* and *dm5.2* were major effect loci for DM resistance [14]. Wang et al. (2018) reported that *dm5.1*, *dm5.2*, and *dm5.3* were major-effect QTLs contributing to DM resistance. Wang et al. (2019) identified a DM resistance gene STAYGREEN, which also located on Chr5 [12]. Therefore, we proposed that the sites located on Chr5 was the most likely loci on DM resistance. Thirteen MLO-like genes and three homolog DMR genes were identified in cucumber [15]

Genome-wide association analysis (GWAS) is an effective and feasible method of analyzing resistance traits [16]. With the development of sequencing technology and the reduction of sequencing costs, GWAS has been applied in many species, such as corn [17], rice [18–20], maize [21], sorghum [22] and foxtail millet [23]. In cucumber, multiple genes were identified by GWAS. The bitterness gene *Bi* was detected by the genome-wide association of 115 diverse cucumber lines [24]. Wang et al. (2018) found a candidate gene related to callus regeneration by GWAS combined with QTL mapping [25]. Bo et al. (2019) identified a novel locus *fsd6.1* combined with *fsd6.2* (*Csgl3*) which regulates fruit spine density by GWAS [26]. GWAS has been widely used for gene mapping of multiple species and complex traits.

In this study, 97 core germplasms (CG) lines [27] were used to perform genome-wide association analysis on downy mildew resistance and multiple loci were detected. These loci provided strong support for the comprehensive utilization of cucumber resistance to downy mildew genetic resources.

Materials And Methods

Plant materials

97 CG lines of cucumber were provided by the cucumber research group in the Institute of Vegetables and Flowers, Chinese Academy of Agricultural Sciences, China. This CG population was selected from more than 3,000 germplasms worldwide and resequenced [27]. Sequencing data was on the Cucumber Genome website (<http://www.icugi.org/cgi-bin/ICuGI/index.cgi>).

Investigation of disease index of the core germplasm resistance

All materials were grown in the greenhouse of Nankou farm (40°13'N, 116°09'E) and Shunyi farm (40°15'N, 116°83'E) in Beijing in 2014. Disease tests for DM resistance were performed twice, in Spring 2014 (*dm_2014S*) and Autumn 2014 (*dm_2014A*), respectively. In each test three replicates were applied and 6 plants were used per line. Naturally DM occurred at adult stage. Downy mildew symptom appeared at about five weeks after sowing. The disease was identified three times at about six weeks after sowing, and once a week.

Disease rating scale for each seedling was as follows: 0: no symptom; 1 : $\leq 1/10$ of all leaves with downy mildew spots; 3: $1/10 \sim 1/4$ of all leaves with downy mildew spots; 5: $1/4 \sim 1/2$ of all leaves with downy mildew spots ; 7: $1/2 \sim 3/4$ of all leaves with downy mildew spots; 9: $\geq 3/4$ of all leaves with downy mildew spots, or the whole leaf dead [28]. $DI = 100 \times \frac{\sum (\text{Number of plants with disease rating} \times \text{Disease rating})}{(\text{Total number of plants} \times \text{Highest disease rating})}$ [28].

All statistics for significant differences among different materials were performed using two-way analysis of variance or two-tailed, two-sample Student's t-test.

Genetic diversity of DM resistance in germplasm

The phylogenetic tree was constructed based on the genotype and phenotype of germplasm using SAS 9.0 [29]. According to the mean of DI, three groups were clustered using the complete. The distance of the complete was 1.5.

Genomewide association analyses of DM resistance

FastLMM was used for genome wide association analyses, with an estimated relatedness matrix as covariate, and the genome-wide minimal P value was recorded. The 5% lowest minimal P values were as the threshold for genome-wide significance. The Manhattan map for GWAS was generated using R package CMplot [30]. SNP data used for the association analysis were from the cucumber genome website (<http://www.icugi.org/cgi-bin/ICuGI/index.cgi>) [27].

Linkage-disequilibrium analysis

The software Plink [31] was used to calculate the LD coefficient (r^2) between pairwise high-quality SNPs. the parameters were set as: '-r2 -ld-window 999999 -ld-window-kb 1000 -ld-window-r2 0', and the results were used to estimate LD decay.

Analysis of novel loci and stronger signal loci for the DM resistance

The novel loci were compared with previous studies. The 50 kb around peak SNPs were considered as candidate region. The candidate regions were predicted using the LDblock and the candidate genes were identified by the annotation of homologous gene in *Arabidodopsis* and the SNP variation. The candidate gene was analyzed by the gene annotation. The physical distance was based on the cucurbit genomics (<http://cucurbitgenomics.org/>). The effect of the QTLs was performed using the R/qtl package with the multiple-QTL model (MQM) [32, 33].

Results

Genetic diversity of DM resistance in CG

The DM resistance of the CG materials were phenotyped in the spring of 2014 (dm_2014S) and the autumn of 2014 (dm_2014A), respectively. The mean DI of dm_2014A was higher than that of dm_2014S, which indicated that DM was more severe in the autumn of 2014. The coefficient of variation (CV) was 27.4% and

59.3%, respectively. The DI in the two years were similarly (Fig. 1b), so the phenotype of DM resistance was effective and accurate for the GWAS analysis. The frequency distribution of DI from different environments were illustrated in Fig. 1a. Distribution of DI in two experiments showed a bimodal model, which suggested that DM resistance in cucumber was controlled by multiple QTLs.

According to the mean of DI in CG materials, there were four categories clustered based on the complete method (1.5) of SAS (Supplementary Fig. 1). Highly resistance (HR) ($DI < 10$) contained 6 lines (“CG64”, “CG70”, “CG72”, “CG75”, “CG81” and “CG86”). Resistance (R) ($10 < DI < 27$) contained 34 lines. Sensitive (S) ($27 < DI < 39$) contained 34 lines. And highly sensitive (HS) ($DI > 39$) contained 23 lines (Supplementary Table S1). South China Banna type was distributed in the resistance and highly resistance category. Japanese type was mainly sensitive to DM. Due to HR lines distributed in the India and Xishuangbanna, we proposed that the resistance gene to DM might be derived from India or Xishuangbanna (Fig. 2c).

Genome-associated analysis of DM resistance

Thirteen loci (*dmG1.1*, *dmG1.2*, *dmG2.1*, *dmG2.2*, *dmG3.1*, *dmG4.1*, *dmG4.2*, *dmG5.1*, *dmG5.2*, *dmG6.1*, *dmG6.2*, *dmG7.1* and *dmG7.2*) for DM resistance were detected, distributed across 7 chromosomes (Fig. 2). The detailed SNPs were listed in Table 1. Four loci (*dmG2.1*, *dmG6.1*, *dmG6.2* and *dmG7.2*) were detected in dm_2014A. Five loci (*dmG1.1*, *dmG2.2*, *dmG4.1*, *dmG4.2* and *dmG5.1*) were detected in dm_2014S. Three loci (*dmG1.2*, *dmG5.2* and *dmG7.1*) were repeatedly detected in two years, and could be considered as stable and major loci related to DM. *dmG2.1* and *dmG7.1* were novel loci based on previous study (Discussion). These novel and major loci (*dmG2.1*, *dmG1.2*, *dmG5.2* and *dmG7.1*) were used for further candidate genes analysis.

Table 1
Comparison of loci identified by GWAS with previously reported QTLs

Locus	SNP name	Chromosome	SNP Position	Flanking markers	Physical location (Mb)	Reference
dmG-1.1	SNP187059	1	10,520,443	UW084539	8.72	Zhang et al. 2018
				SSR16055	13.30	
dmG-1.2	SNP398090	1	21,761,223	Marker1_15701257	15.70	Wang et al.2018
				Marker1_26471365	26.47	
dmG-2.1	SNP639496	2	4,618,724	Novel		
dmG-2.2	SNP730945	2	8,449,502	SSR13532	5.64	Win et al. 2017
				SSR13105	10.28	
dmG-3.1	SNP1139014	3	8,524,603	SSR16264	9.15	Win et al. 2017
				SSR13312	10.25	
dmG-4.1	SNP1793602	4	5,302,408	SSR00012	4.11	Cavagnaro et al.2011
				SSR13159	20.10	
dmG-4.2	SNP1938892	4	13,242,382	SSR00012	4.11	Cavagnaro et al.2011
				SSR13159	20.10	
dmG-5.1	SNP2452535	5	18,119,570	Marker5_16446482	16.44	Wang et al.2018
				Marker5_18140438	18.14	
dmG-5.2	SNP2531703	5	22,899,752	SSR01498	22.66	Zhang et al. 2018
				InDel82	25.75	
dmG-6.1	SNP13977656	6	13,977,656		5.54	Win et al. 2016
					17.28	
dmG-6.2	SNP3104337	6	25,297,012	SSR01148	25.95	Wang et al. 2016
				SSR05946	28.09	
dmG-7.1	SNP3165616	7	84,900	Novel		
dmG-7.2	SNP3517753	7	18,251,423	SSR33278	17.52	Yoshioka et al. 2014
				SSR477	19.19	

Candidate gene analysis of novel locus dmG2.1

In order to identify potential candidate genes in novel and stable major loci, the 50 kb region around the peak SNPs were used for further analysis (Fig. 3a). Based on the cucurbit genome (<http://cucurbitgenomics.org/>), six genes were presented in the *dmG2.1* region. In this region (Fig. 3b), *Csa2G059820* encodes Pentatricopeptide repeat-containing protein; *Csa2G059830* encodes homogentisate phytyltransferase; *Csa2G059840* encodes Lysine-tRNA ligase; *Csa2G060350* is a transcription factor/ transcription regulator; *Csa2G060360* is an unknown protein; *Csa2G060370* encodes Fe/S biogenesis protein. Expression of six genes were analyzed based on the RPKM in the RNA-transcriptome of PRJNA285071 (<http://cucurbitgenomics.org/rnaseq/home>). The expression of *Csa2G059820*, *Csa2G060360* and *Csa2G060370* were up-regulated in the resistant line PI 197088 (Fig. 3c). The homolog gene of *Csa2G059820* (*AT2G18940*) and *Csa2G060360* (*AT5G57345*) in Arabidopsis were differentially expressed in infected rosette leaves at 12 days post inoculation of Cabbage leaf curl virus [34]. These two genes could be the candidate gene of *dmG2.1*.

Candidate gene analysis of novel locus dmG7.1

Fifty kb around the peak SNPs were further used to analyze the candidate gene associated with the locus dmG7.1 (Fig. 3d). Based on the cucurbit genome (<http://cucurbitgenomics.org/>), 4 genes were predicted in the *dmG7.1* region. All the four genes were the isoflavone reductase related protein (Fig. 3e). Phylogenetic tree was constructed based on the amino acids in the reference genome. Phylogenetic tree showed that *Csa7G002050* was homologous with *Csa7G004020* (Fig. 3f). The results showed that *dmG7.1* could be related to the isoflavone reductase. The expression of *Csa7G004020* was up-regulated at the 3-day and 6-day after DM infection in the resistant line PI 197088 based on the RPKM in the RNA-transcriptome of PRJNA285071 in <http://cucurbitgenomics.org/rnaseq/home> (Fig. 3g). Therefore, we proposed that *Csa7G004020* could be the candidate gene of *dmG7.1*.

Candidate gene analysis for stronger loci dmG1.2

For the *dmG1.2*, the candidate region (Chr.1: 21,711 – 21,781 kb) were analyzed by pairwise LD correlations (Fig. 4a). We focused on the locus mapped from 21,756,442 bp to 21,774,330 bp by LD block. Based on the Cucumber Genome Browser (<http://www.icugi.org/cgi-bin/ICuGI/index.cgi>), two candidate genes were located in this region. The two genes *Csa1G575020* and *Csa1G575030* both encode kinase family protein. No SNPs located in the CDS of *Csa1G575020*. Seven SNPs variation which result in amino acid variation within the CDS region of *Csa1G575030* (Fig. 4c). We proposed that *Csa1G575030* was the candidate gene of *dmG1.2*.

Candidate gene analysis for stronger loci dmG5.2

For the *dmG5.2*, 50 kb around the peak SNPs region were further analyzed (Fig. 5a). The candidate region from 22,883 kb to 22,913 kb (~ 30 kb) was estimated using pairwise LD correlations ($r^2 \geq 0.6$) (Fig. 5b). According to the Cucumber Genome Browser (<http://cucurbitgenomics.org/organism/2>), five annotated genes were identified in this candidate region (Fig. 5b). *Csa5G606470* is a WRKY transcription factor 2–4; *Csa5G606480* is a eukaryotic translation initiation factor 5A; *Csa5G606490* is a cell division topological specificity factor; *Csa5G606500* encodes a putative uncharacterized protein RAF9-1; *Csa5G606510* is an unknown protein. Based on the homologous gene in *Arabidopsis*, the gene *AT1G13960* (Homologous of *Csa5G606470*) in response to salicylic acid in leaf proteome profile of *Arabidopsis thaliana* [35]. SNP_2531403 was on the CDS of *Csa5G606470*. On this site, 29 of 30 HR lines was G (Pro), the other one was A (Ser), whereas that 7 of 29 HS lines was G (Pro), the others were A (Pro) (Fig. 5c). *Csa5G606470* could be the candidate gene of *dmG5.2*.

Discussion

Inheritance of DM resistance in cucumber was complex. Angelov showed that the resistance of DM in cucumber was controlled by a recessive gene [36] and this gene was linked with dull green fruit skin color related gene *D* [37]. Most studies have reported that DM was controlled by multiple genes in cucumber, such as a dominant susceptible gene and a recessive gene [38], a pair of dominant and recessive interacting genes [39], three recessive genes [9]. In our study, the results showed that DM resistance was controlled by multiple genes. The different results could be due to different materials and different environments. The core germplasm we used possesses the DM resistance with most variation.

In recent years, GWAS is widely used to detect genes in many species. Our study firstly used the GWAS to detect loci and genes for DM resistance in cucumber. Several loci distributed on all 7 chromosomes have been identified by different populations. Based on previous studies, two loci were mapped on chr1, chr2, chr3, and chr6, respectively, one loci on chr4 and chr7, three loci (*dm5.1*, *dm5.2* and *dm5.3*) on chr5 [10, 13, 14, 28, 40–45]. In our study, a total of thirteen loci (*dmG1.1*, *dmG1.2*, *dmG2.1*, *dmG2.2*, *dmG3.1*, *dmG4.1*, *dmG4.2*, *dmG5.1*, *dmG5.2*, *dmG6.1*, *dmG6.2*, *dmG7.1* and *dmG7.2*) distributed on 7 chromosomes were detected by CG lines. Compared with the previous QTLs, the loci on the Chr1, Chr3, Chr4, Chr5 and Chr6 had been previously reported. Two loci on the Chr2, Chr7 was new reported, respectively. (Table 1). Based on the loci listed in Table 1, eleven of thirteen loci were also identified in QTL mapping. The candidate gene, *Csa5G606470*, mapped in *dmG5.2*, had a SNP in the CDS. And the SNP variation result in amino acid changes. The candidate gene was the same as *dm5.2* based on BSA-seq analysis in Zhang et al 's study [41]. The variation was effect on the 3' UTR, which could be due to different materials. Three loci could be identified in two years, which showed that the results in our study were consistent. Ten of thirteen loci could be identified in one year, which could be due to different environments. In our study, thirteen loci were consistent with previous studies, which showed that GWAS could detect more variations [46]. So, GWAS was efficient to detect DM resistance related QTLs.

DM resistance is a complex trait that involves in multiple genes and metabolic pathways. Some of the major loci were interacted with each other. Three major QTLs in Chr5 were negative additive effects [42]. Wang et al. (2016) detected the effect of four QTLs (*dm2.1*, *dm4.1*, *dm5.1* and *dm6.1*). *dm6.1* could be interacted

with *dm2.1*, *dm4.1* and *dm5.1*. And *dm2.1* could be affected by *dm5.1*. In our study, the peak SNPs in *dmG1.2*, *dmG2.1*, *dmG5.2* and *dmG7.1* were applied to detect the effect code of each other. *dmG1.2* interacted with *dmG7.1* in the two experiments. *dmG5.2* had effect on *dmG1.2* and *dmG7.1* in one season. *dmG1.2* was independent of *dmG2.1* and *dmG5.2*, and *dmG2.1* was independent with *dmG7.2*. (Supplementary Fig. 2) These results showed that the DM resistance is a complex trait and with multiple gene interactions, which was consistent with previous studies [42, 47]. Therefore, GWAS is an effective strategy to identify complex traits.

Conclusions

Thirteen loci associated with DM resistance were detected in cucumber. Two novel loci were identified compared with previous studies. Based on the annotation of homologous genes in *Arabidopsis* and pairwise LD correlations, five candidate genes for four loci related to DM resistance were identified in cucumber. This study provides theoretical basis for gene cloning and genetic breeding of DM resistance in cucumber.

Supplementary Materials

Figure S1: The phylogenetic tree of phenotype of CG; Figure S2: Genotypic effects of the *dmG1.2*, *dmG2.1*, *dmG5.2* and *dmG7.1* in the dm_2014S and dm_2014A; Table S1: Phenotype of core germplasm.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and materials

All data generated or analysed during this study are included in this published article and its supplementary information files.

Competing interests

The authors declare that they have no competing interests.

Funding

This work was supported by the National Natural Science Foundation of China (31672172); the earmarked fund for Modern Agro-industry Technology Research System (CARS-23), The Agricultural Science and Technology Innovation Program of the Chinese Academy of Agricultural Sciences (CAAS-ASTIP-IVFCAAS)

and the Key Laboratory of Biology and Genetic Improvement of Horticultural Crops, Ministry of Agriculture and Rural Affairs, P. R. China.

Authors' contributions

XL wrote and revised the manuscript. HL, PL and HM investigated disease index data. YB, XG and SZ conceived of the study and critically reviewed the manuscript. All authors read and approved the final manuscript.

Acknowledgements

The authors thank the Key Laboratory of Biology and Genetic Improvement of Horticultural Crops, Ministry of Agriculture, China.

References

1. Jones ES, Liu CJ, Gale MD, Hash CT, Witcombe JR. Mapping quantitative trait loci for downy mildew resistance in pearl millet. *Tagtheoretical Applied Geneticstheoretische Angewandte Genetik*. 1995;91(3):448–56.
2. McDowell JM, Cuzick A, Can C, Beynon J, Dangl JL, Holub EB. **Downy mildew (*Peronospora parasitica*) resistance genes in *Arabidopsis* vary in functional requirements for NDR1, EDS1, NPR1 and salicylic acid accumulation.** *Plant Journal* 2000, 22(6).
3. Bittner-Eddy PD, Crute IR, Holub EB, Beynon JL. **RPP13 is a simple locus in *Arabidopsis thaliana* for alleles that specify downy mildew resistance to different avirulence determinants in *Peronospora parasitica*.** 2010, 21(2):177–188.
4. Van Damme M, Zeilmaker T, Elberse J, Andel A, de Sain-van der Velden M, van den Ackerveken G. **Downy Mildew Resistance in *Arabidopsis* by Mutation of HOMOSERINE KINASE.** *Plant Cell*, 21(7):2179–2189.
5. Beniswal MS, Karwasra SS, Chhabra ML. **Simultaneous occurrence of flag smut and downy mildew in wheat plant.** 1994.
6. Li G, Xu X, Tan C, Carver BF, Cowger C. **Identification of powdery mildew resistance loci in wheat by integrating genome-wide association study (GWAS) and linkage mapping.** *The Crop Journal* 2019(3).
7. Liu Z, Zhang L, Ma GJ, Seiler GJ, Qi LL. Molecular mapping of the downy mildew and rust resistance genes in a sunflower germplasm line TX16R. *Mol Breeding*. 2019;39(2):19.
8. Radwan O, Bouzidi MF, Vear F, Philippon J, Mouzeyar S. Identification of non-TIR-NBS-LRR markers linked to the PI5/PI8 locus for resistance to downy mildew in sunflower. *Theoretical Applied Genetics*. 2003;106(8):1438–46.
9. Kenigsbuch D. Inheritance of Resistance to Downy Mildew in *Cucumis melo* PI 124112 and Commonality of Resistance Genes with PI 124111F. *Plant Dis*. 1992;76(6):615.
10. Vandenlangenberg KM. **Studies on Downy Mildew Resistance in Cucumber (*Cucumis sativus* L.).** 2015.

11. Wang Y, Bo K, Gu X, Pan J, Li Y, Chen J, Wen C, Ren Z, Ren H, Chen X. Molecularly tagged genes and quantitative trait loci in cucumber with recommendations for QTL nomenclature. *Horticulture Research*. 2020;7(1):1–20.
12. Wang Y, Tan J, Wu Z, VandenLangenberg K, Wehner TC, Wen C, Zheng X, Owens K, Thornton A, Bang HH. STAYGREEN, STAY HEALTHY: a loss-of-susceptibility mutation in the STAYGREEN gene provides durable, broad-spectrum disease resistances for over 50 years of US cucumber production. *New Phytol*. 2019;221(1):415–30.
13. Wang Y, VandenLangenberg K, Wehner TC, Kraan PA, Suelmann J, Zheng X, Owens K, Weng Y: **QTL mapping for downy mildew resistance in cucumber inbred line WI7120 (PI 330628)**. *Theoretical and applied genetics* 2016, **129**(8):1493–1505.
14. Win KT, Vegas J, Zhang C, Song K, Lee S. QTL mapping for downy mildew resistance in cucumber via bulked segregant analysis using next-generation sequencing and conventional methods. *Theoretical applied genetics*. 2017;130(1):199–211.
15. Schouten HJ, Krauskopf J, Visser RG, Bai Y. Identification of candidate genes required for susceptibility to powdery or downy mildew in cucumber. *Euphytica*. 2014;200(3):475–86.
16. Hindorff LA, Sethupathy P, Junkins HA, Ramos EM, Mehta JP, Collins FS, Manolio TA: **Potential etiologic and functional implications of genome-wide association loci for human diseases and traits**. *Proceedings of the National Academy of Sciences* 2009, **106**(23):9362–9367.
17. Ma Z, He S, Wang X, Sun J, Zhang Y, Zhang G, Wu L, Li Z, Liu Z, Sun G. Resequencing a core collection of upland cotton identifies genomic variation and loci influencing fiber quality and yield. *Nat Genet*. 2018;50(6):803.
18. Huang X, Sang T, Zhao Q, Feng Q, Zhao Y, Li C, Zhu C, Lu T, Zhang Z, Li M. Genome-wide association studies of 14 agronomic traits in rice landraces. *Nat Genet*. 2010;42(11):961.
19. Huang X, Kurata N, Wang Z-X, Wang A, Zhao Q, Zhao Y, Liu K, Lu H, Li W, Guo Y. A map of rice genome variation reveals the origin of cultivated rice. *Nature*. 2012;490(7421):497–501.
20. Yano K, Yamamoto E, Aya K, Takeuchi H, Lo P-c, Hu L, Yamasaki M, Yoshida S, Kitano H, Hirano K. Genome-wide association study using whole-genome sequencing rapidly identifies new genes influencing agronomic traits in rice. *Nat Genet*. 2016;48(8):927.
21. Li H, Peng Z, Yang X, Wang W, Fu J, Wang J, Han Y, Chai Y, Guo T, Yang N. Genome-wide association study dissects the genetic architecture of oil biosynthesis in maize kernels. *Nat Genet*. 2013;45(1):43.
22. Mace ES, Tai S, Gilding EK, Li Y, Prentis PJ, Bian L, Campbell BC, Hu W, Innes DJ, Han X. Whole-genome sequencing reveals untapped genetic potential in Africa's indigenous cereal crop sorghum. *Nature communications*. 2013;4(1):1–9.
23. Jia G, Huang X, Zhi H, Zhao Y, Zhao Q, Li W, Chai Y, Yang L, Liu K, Lu H. A haplotype map of genomic variations and genome-wide association studies of agronomic traits in foxtail millet (*Setaria italica*). *Nat Genet*. 2013;45(8):957–61.
24. Shang Y, Ma Y, Zhou Y, Zhang H, Duan L, Chen H, Zeng J, Zhou Q, Wang S, Gu W. Biosynthesis, regulation, and domestication of bitterness in cucumber. *Science*. 2014;346(6213):1084–8.

25. Wang Y, Zhou Q, Zhu GT, Wang SH, Ma YS, Miao H, Zhang SP, Huang SW, Zhang ZH, Gu XF. Genetic analysis and identification of a candidate gene associated with in vitro regeneration ability of cucumber. *Theor Appl Genet.* 2018;131(12):2663–75.
26. Bo K, Miao H, Wang M, Xie X, Song Z, Xie Q, Shi L, Wang W, Wei S, Zhang S: **Novel loci *fsd6. 1* and *Csagl3* regulate ultra-high fruit spine density in cucumber.** *Theoretical and Applied Genetics* 2019, **132**(1):27–40.
27. Qi J, Liu X, Shen D, Miao H, Xie B, Li X, Zeng P, Wang S, Shang Y, Gu X. A genomic variation map provides insights into the genetic basis of cucumber domestication and diversity. *Nat Genet.* 2013;45(12):1510.
28. Zhang S, Liu M, Miao H, Zhang S, Yang Y, Xie B, Wehner T, Gu X. Chromosomal mapping and QTL analysis of resistance to downy mildew in *Cucumis sativus*. *Plant Dis.* 2013;97(2):245–51.
29. Institute S: **SAS/OR 9.3 User's Guide: Mathematical Programming Examples.** SAS institute; 2012.
30. Yin L. **CMplot: circle manhattan plot.** 2018.
31. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, De Bakker PI, Daly MJ. PLINK: a tool set for whole-genome association and population-based linkage analyses. *The American journal of human genetics.* 2007;81(3):559–75.
32. Broman KW, Wu H, Sen S, Churchill GA. R/qtl: QTL mapping in experimental crosses. *Bioinformatics.* 2003;19(7):889–90.
33. Arends D, Prins P, Jansen RC, Broman KW. R/qtl: high-throughput multiple QTL mapping. *Bioinformatics.* 2010;26(23):2990–2.
34. Ascencio-Ibáñez JT, Sozzani R, Lee T-J, Chu T-M, Wolfinger RD, Cella R, Hanley-Bowdoin L. Global analysis of *Arabidopsis* gene expression uncovers a complex array of changes impacting pathogen response and cell cycle during geminivirus infection. *Plant physiology.* 2008;148(1):436–54.
35. Datta R, Sinha R, Chattopadhyay S. Changes in leaf proteome profile of *Arabidopsis thaliana* in response to salicylic acid. *Journal of biosciences.* 2013;38(2):317–28.
36. Angelov D. **Inheritance of resistance to downy mildew, *Pseudoperonospora cubensis*.** In: *Berk & Curt) Rostow Rep 2nd Natl Symp Plant Immunity (Plovdiv): 1994;* 1994: 99–105.
37. Van Vliet G, Meysing W. Inheritance of resistance to *Pseudoperonospora cubensis* Rost. in cucumber (*Cucumis sativus* L.). *Euphytica.* 1974;23(2):251–5.
38. El-Hafez A, Shehata S, El-Din S, El-Doweny H, Awad M. Inheritance of downy mildew disease and its nature of resistance in cucumber. *Annals of Agricultural Science Moshtohor.* 1990;28(3):1681–97.
39. Badr L, Mohamed F. Inheritance and nature of resistance to downy mildew disease in cucumber (*Cucumis sativus* L.). *Ann Agric Sci Moshtohor.* 1998;36(4):2517–44.
40. Yoshioka Y, Sakata Y, Sugiyama M, Fukino N. Identification of quantitative trait loci for downy mildew resistance in cucumber (*Cucumis sativus* L.). *Euphytica.* 2014;198(2):265–76.
41. Zhang K, Wang X, Zhu W, Qin X, Xu J, Cheng C, Lou Q, Li J, Chen J. Complete resistance to powdery mildew and partial resistance to downy mildew in a *Cucumis hystrix* introgression line of cucumber were controlled by a co-localized locus. *Theoretical applied genetics.* 2018;131(10):2229–43.

42. Wang Y, VandenLangenberg K, Wen C, Wehner TC, Weng Y. QTL mapping of downy and powdery mildew resistances in PI 197088 cucumber with genotyping-by-sequencing in RIL population. *Theoretical applied genetics*. 2018;131(3):597–611.
43. Bai Z, Yuan X, Cai R, Liu L, He H, Zhou H, Pan J. QTL analysis of downy mildew resistance in cucumber. *Prog Nat Sci*. 2008;18:706–10.
44. Pang X, Zhou X, Wan H, Chen J. QTL mapping of downy mildew resistance in an introgression line derived from interspecific hybridization between cucumber and *Cucumis hystrix*. *J Phytopathol*. 2013;161(7–8):536–43.
45. Szczechura W, Staniaszek M, Klosinska U, Kozik E. Molecular analysis of new sources of resistance to *Pseudoperonospora cubensis* (Berk. et Curt.) Rostovzev in cucumber. *Russian Journal of Genetics*. 2015;51(10):1134.
46. Tang W, Ye J, Yao X, Zhao P, Xuan W, Tian Y, Zhang Y, Xu S, An H, Chen G. Genome-wide associated study identifies NAC42-activated nitrate transporter conferring high nitrogen use efficiency in rice. *Nature communications*. 2019;10(1):1–11.
47. Zhang YM, Zhao FJ, Kong MM, Wang SW, Nan L, Hu BJ, Olszewski MA, Miao YX, Ji DN, Jiang WR, et al: **Validation of a High-Throughput Multiplex Genetic Detection System for *Helicobacter pylori* Identification, Quantification, Virulence, and Resistance Analysis**. *Front Microbiol* 2016, 7.

Figures

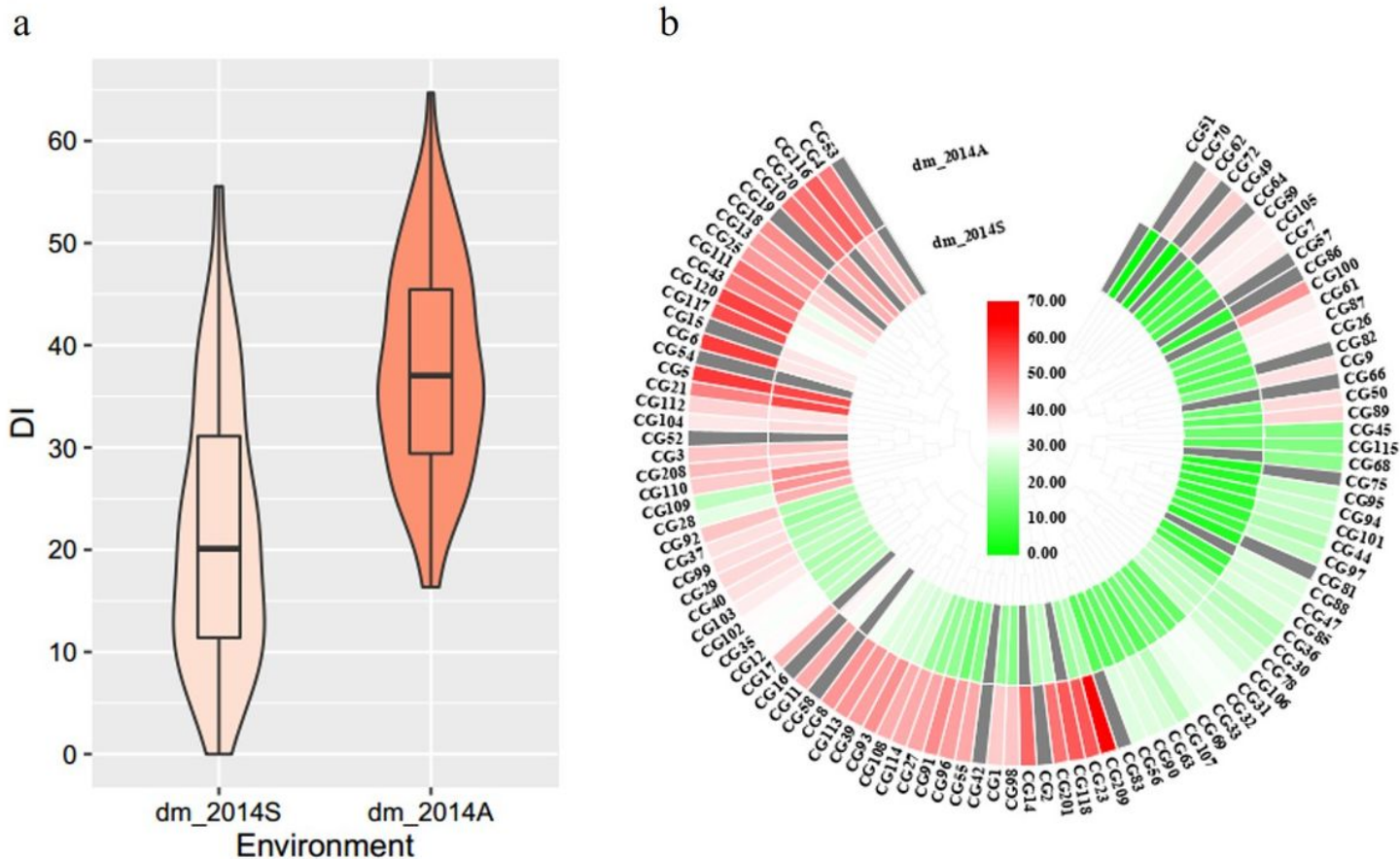


Figure 1

Box plots and heatmap depicting the phenotypic distribution of DI of DM resistance in two environments.

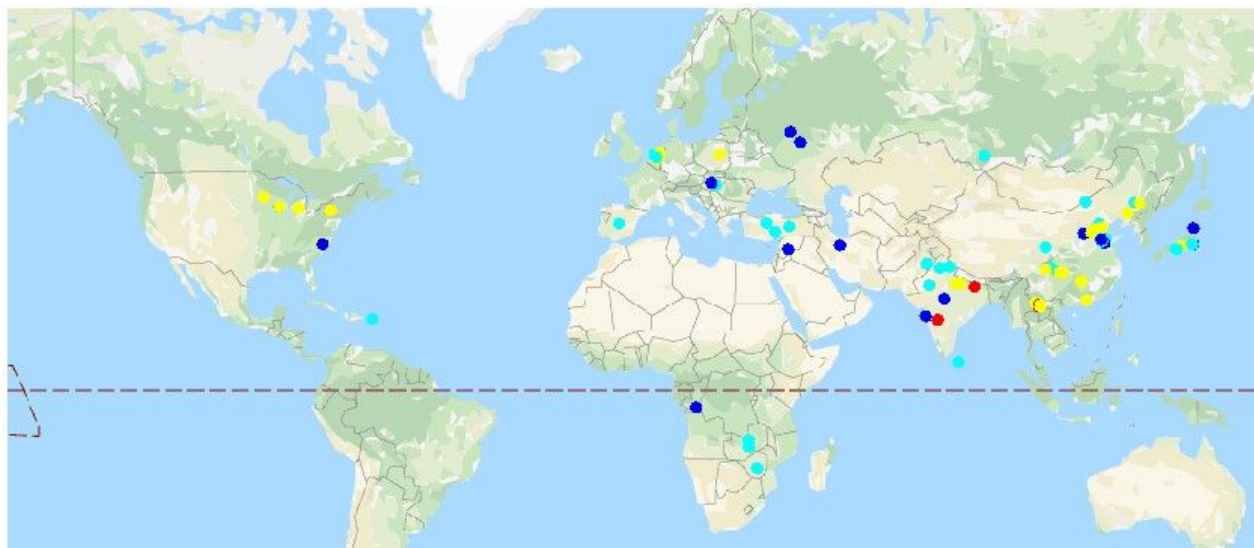
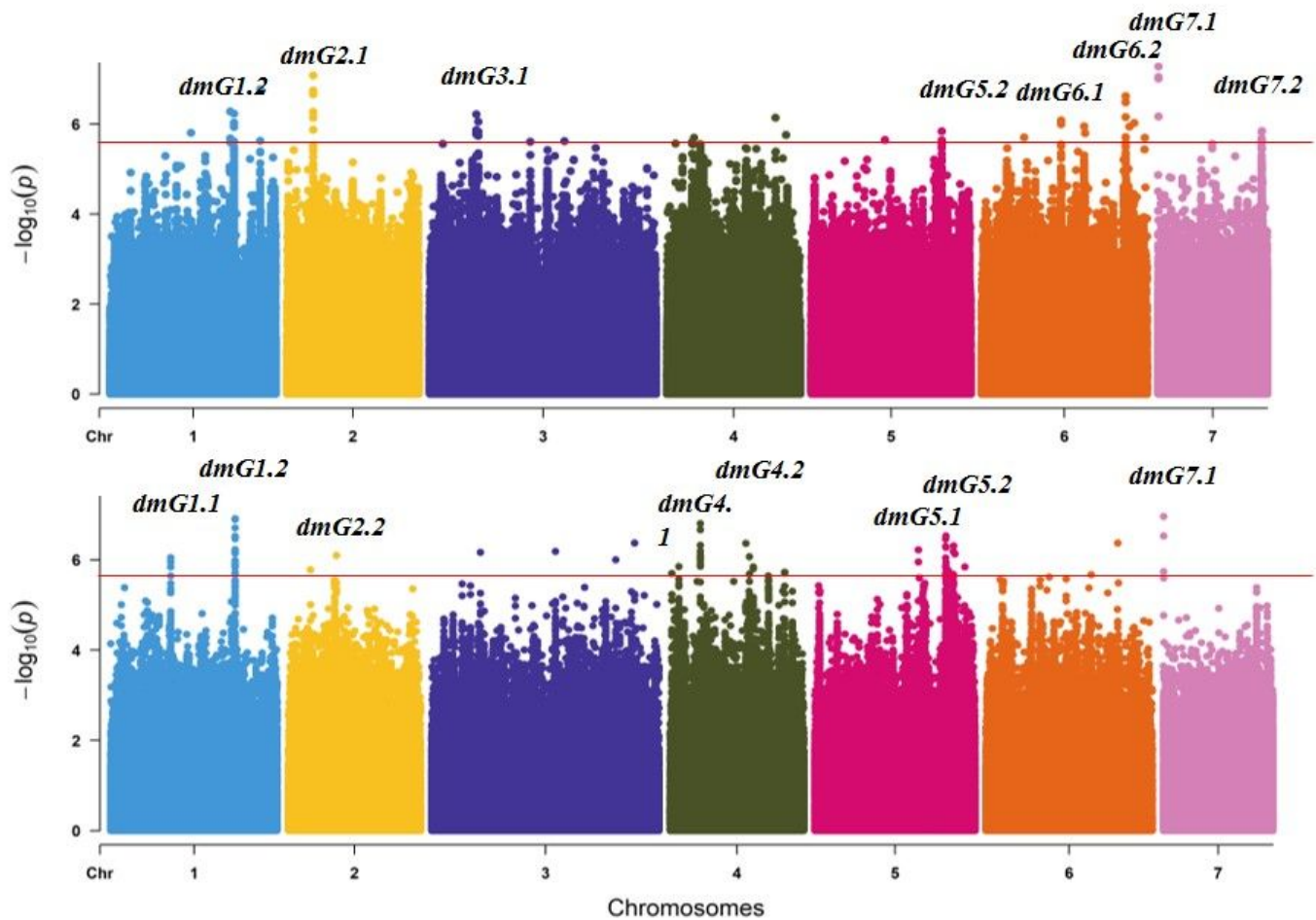


Figure 2

GWAS Mahattan plots of DM resistance in two experiments, and the distribution of 109 cucumber CG lines in different continents based on DI of DM resistance. a: dm_2014A; b:dm_2014S; c: distribution of 109 cucumber CG lines. The red spot represents HR; Blue HS; Yellow represents R; Green represents S.

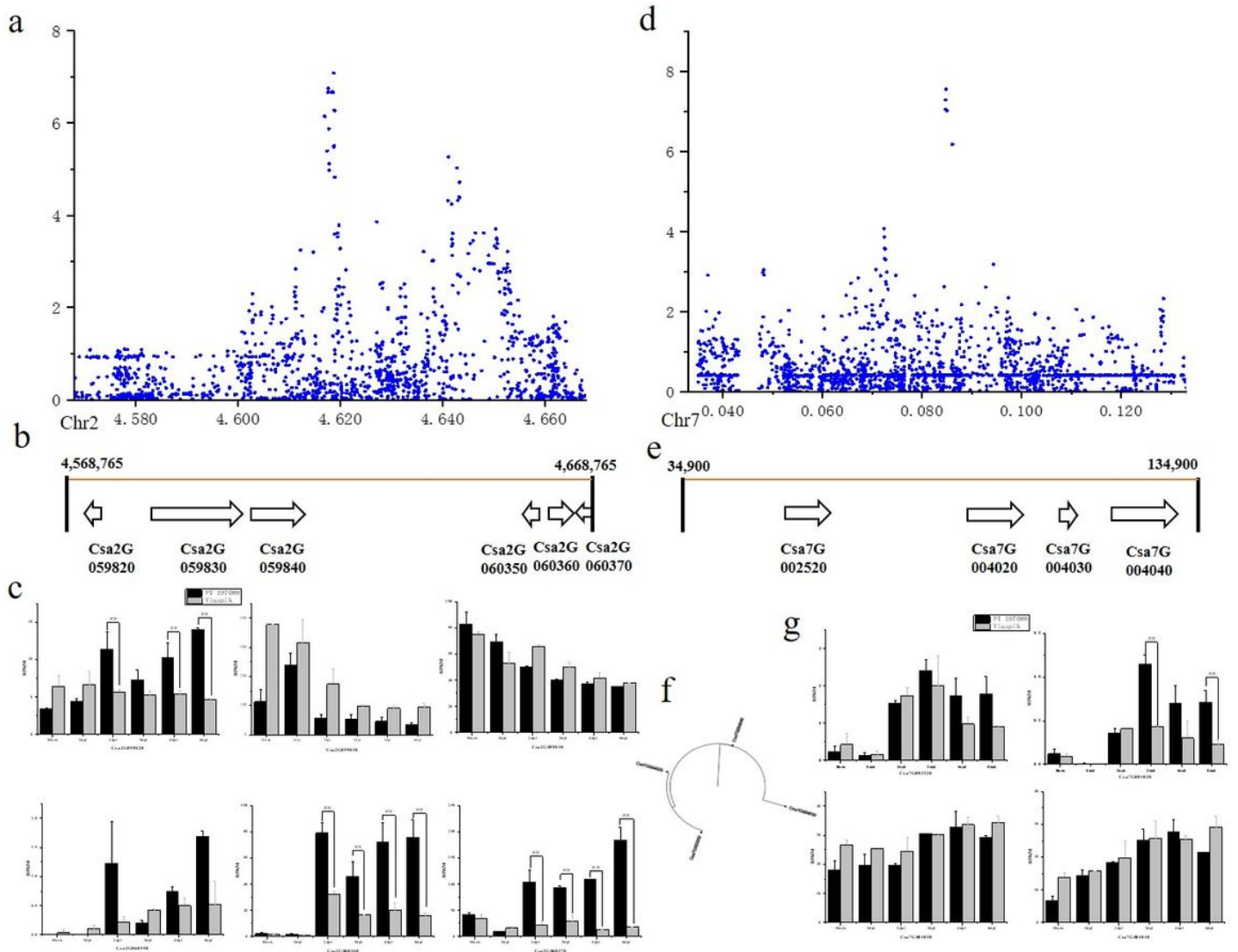


Figure 3

Identification of the causal gene for the peak dmG2.1 and dmG7.1. a: Local Manhattan plot surrounding the peak dmG2.1; b: Six genes were predicted in 50 kb around the peak dmG2.1. c: RPKM of these six genes based on the transcriptome of PRJNA285071; d: Local Manhattan plot surrounding the peak dmG7.1. e: Four genes were predicted in 50 kb around the peak dmG7.1. f: Phylogenetic analysis of four genes in dmG7.1 candidate region. g: RPKM of four genes in dmG7.1 candidate region based on the transcriptome of PRJNA285071.

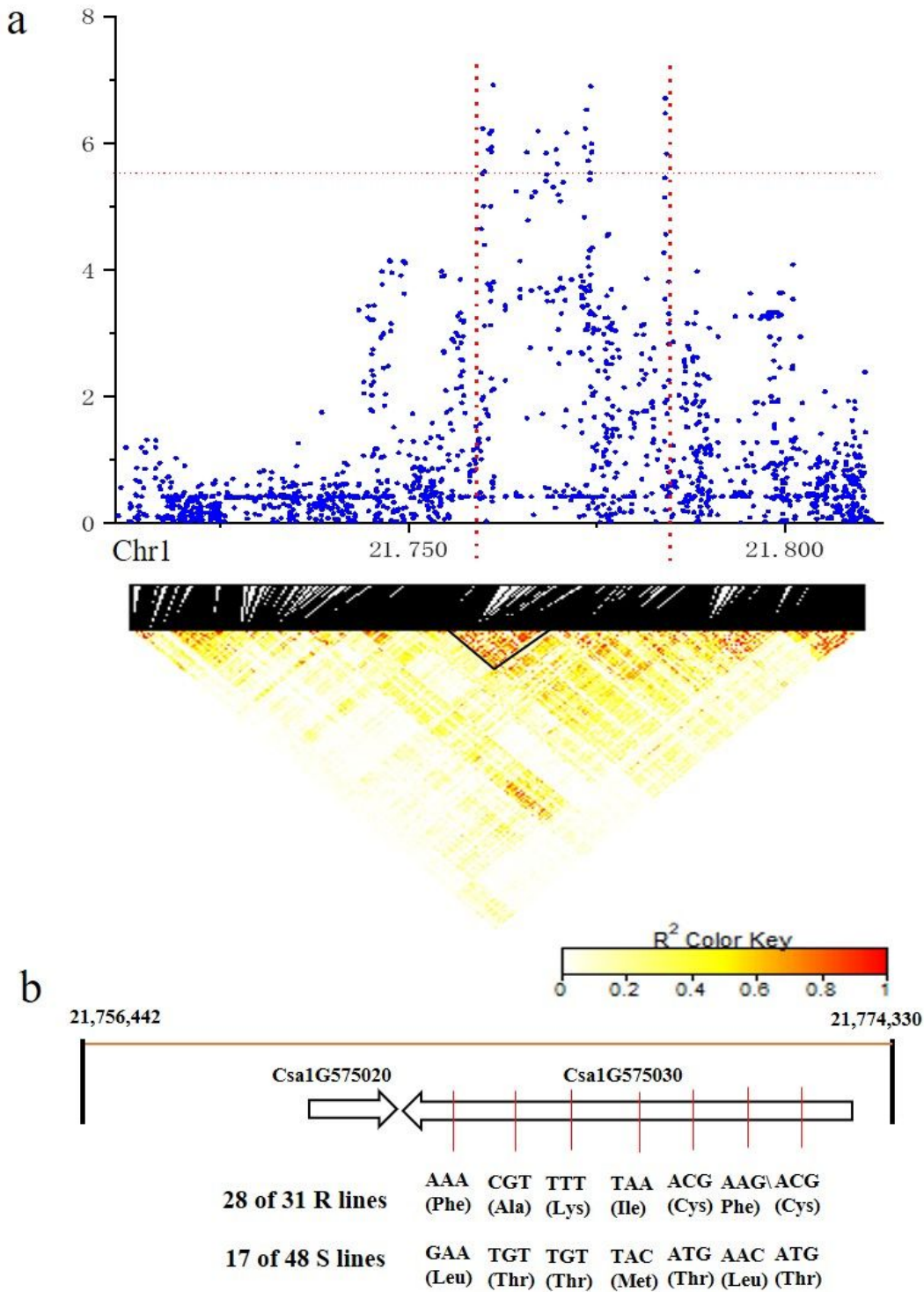


Figure 4

Identification of the causal gene for the peak dmG1.2. a: Local Manhattan plot (top) and LD heatmap (bottom) surrounding the peak dmG5.2; b: SNP variation of candidate gene in the core germplasm.

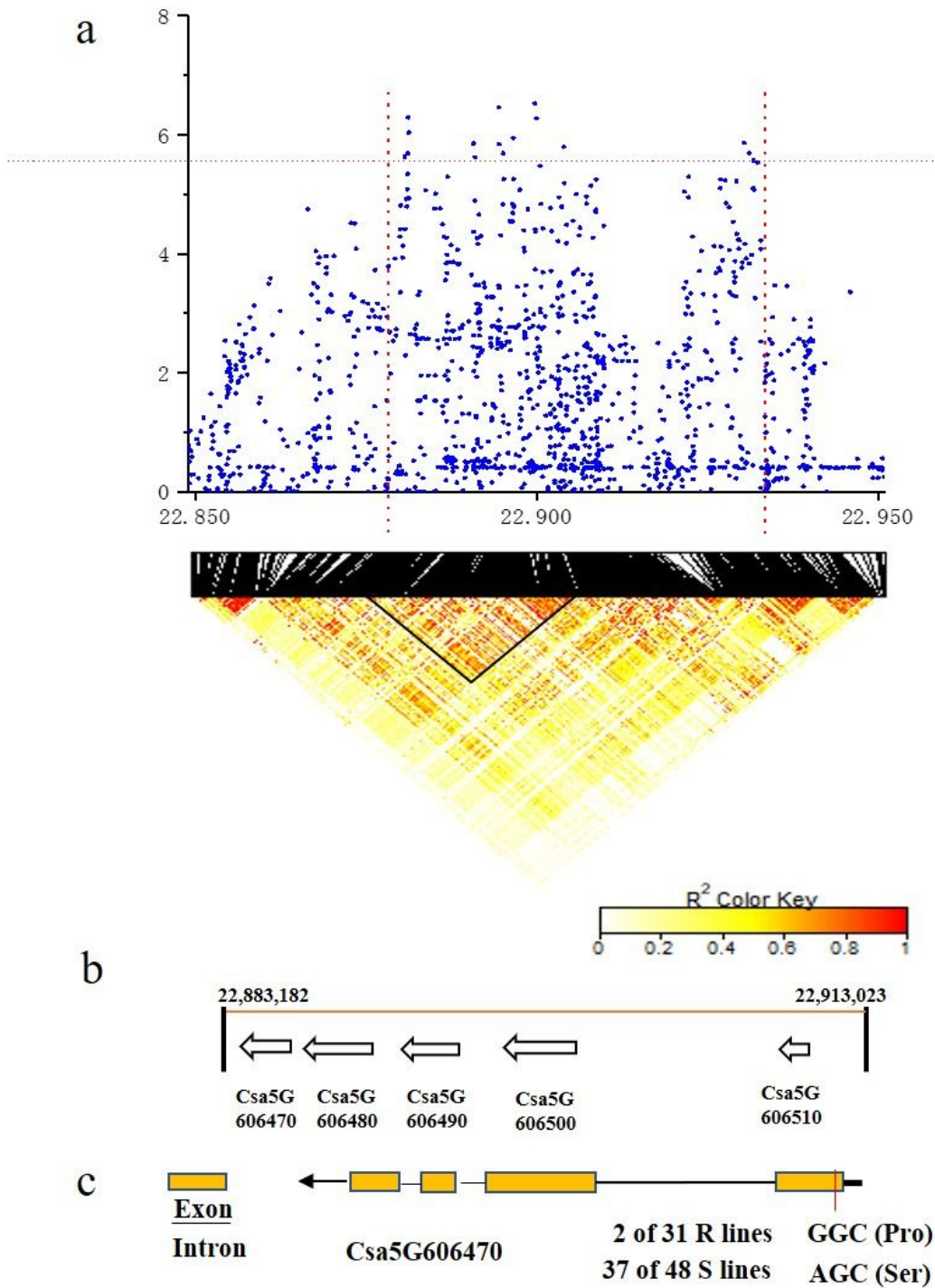


Figure 5

Identification of the causal gene for the peak dmG5.2. a: Local Manhattan plot (top) and LD heatmap (bottom) surrounding the peak dmG5.2; b: Dashed lines indicate the candidate region (~ 30 kb) for the peak. Five genes were predicted in the dmG2.1 candidate region. c: SNPs variation of candidate gene in the core germplasm.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [TableS1.docx](#)
- [FigureS1.jpg](#)
- [FigureS2.jpg](#)