Identification and Functional Verification of Cold Tolerance Genes in Spring Maize Seedlings Based on a Genome-Wide Association Study and Quantitative Trait Locus Mapping

Yukun Jin†, Zhongren Zhang‡, Yongjing Xi¹, Zhou Yang¹, Zhifeng Xiao¹, Shuyan Guan¹, Jing Qu¹, Piwu Wang* and Rengui Zhao*

¹Joint Laboratory of International Cooperation in Modern Agriculture Technology of Ministry of Education, Plant Biotechnology Center, Jilin Agricultural University, Changchun, China, ‡Novogene Bioinformatics Institute, Novogene Co., Ltd, Beijing, China

Maize (Zea mays L.) is a tropical crop, and low temperature has become one of the main abiotic stresses for maize growth and development, affecting many maize growth processes. The main area of maize production in China, Jilin province, often suffers from varying degrees of cold damage in spring, which seriously affects the quality and yield of maize. In the face of global climate change and food security concerns, discovering cold tolerance genes, developing cold tolerance molecular markers, and creating cold-tolerant germplasm have become urgent for improving maize resilience against these conditions and obtaining an increase in overall yield. In this study, whole-genome sequencing and genotyping by sequencing were used to perform genome-wide association analysis (GWAS) and quantitative trait locus (QTL) mapping of the two populations, respectively. Overall, four single-nucleotide polymorphisms (SNPs) and 12 QTLs were found to be significantly associated with cold tolerance. Through joint analysis, an intersection of GWAS and QTL mapping was found on chromosome 3, on which the Zm00001d002729 gene was identified as a potential factor in cold tolerance. We verified the function of this target gene through overexpression, suppression of expression, and genetic transformation into maize. We found that Zm00001d002729 overexpression resulted in better cold tolerance in this crop. The identification of genes associated with cold tolerance contributes to the clarification of the underlying mechanism of this trait in maize and provides a foundation for the adaptation of maize to colder environments in the future, to ensure food security.

Keywords: cold tolerance, GWAS, QTL mapping, gene cloning, functional annotation
INTRODUCTION

Various abiotic stresses, such as low temperature, drought, and high salinity, significantly affect the normal growth and yield of plants. Of these, cold stress has had an impact on the growth, yield, and spatial distribution of crops, including maize (Zea mays L.). Cold tolerance in maize is a complex quantitative hereditary trait, controlled by multiple genes; different periods of this crop are controlled by different genetic mechanisms, which are easily affected by environmental factors (Hodges et al., 1997; Gao et al., 2009). Jilin Province is one of the most important sites of maize cultivation in China; however, it is extremely prone to cold damage, resulting in reduced maize production (Jena et al., 2012; Ma et al., 2018; Yan et al., 2018).

Several reports identifying molecular markers and quantitative trait locus (QTLs) associated with cold tolerance have been published, but their practical application in maize breeding to improve resistance to cold conditions has so far been limited (Revilla et al., 2014). Many studies have shown that cold tolerance at the germination and seedling stages of maize is a quantitative trait, controlled by multiple genes (Trzczinska-Danielewicz et al., 2009); studies have also shown that epistasis and additive and dominant gene effects in maize germination under low temperature significantly affect tolerance (Ma et al., 2007).

Rodriguez et al. (2008) used a recombinant inbred line (RIL) constructed using a cross between the inbred lines B73 × Mo17 to locate QTLs affecting maize leaf color under low temperatures; they found that the major QTL that control maize leaf color under these conditions is located at the chromosome positions bin 3.01 and bin 6.03. Jompuk et al. (2005) used the F2 population ETH-DH7 × ETH-DL3 to identify the main QTL governing the operation of the photosynthetic system at low temperatures during the seedling stage of maize, which was found to be distributed along chromosome 6. Leipner et al. (2008) also used this F2 population ETH-DH7 × ETH-DL3 to assess cold tolerance in the maize seedling photosynthetic system under low temperature stress and detected that the major QTL for this trait was distributed along chromosome 3. Pimentel et al. (2005) used the constructed F2 population W6786 × IL731A to identify the QTL controlling the degree of photoinhibition in maize under low temperature stress; three QTLs on chromosomes 3, 4, and 8 were detected in plants cultivated under controlled indoor and outdoor conditions. The QTL in bin 3.09 on chromosome 3 was detected in both indoor and outdoor plants; the other two QTLs were only detected in indoor environments. Li et al. (2018) used two cold-tolerant inbred lines, 220 and P9-10, and two cold-sensitive inbred lines, Y1518 and PH4CV, to construct three F2:3 populations; and using the germination rate as the phenotype to QTL mapping, they identified 43 QTLs that each explained 0.62–39.44% of the phenotypic variation among plants in this population. Seven of these QTLs together explained more than 10% of the total phenotypic variation. Hu et al. (2016) identified a total of six QTLs governing seed germination under low temperature conditions, on chromosomes 4, 5, 6, 7, and 9, under different degrees of cold stress, which were found to explain 3.39–11.29% of phenotypic variation.

Strigens et al. (2013) conducted a genome-wide association study (GWAS) in maize, locating 19 single-nucleotide polymorphisms (SNPs) related to cold tolerance, which explained 5.7–52.5% of the phenotypic genetic variation in maize seedlings. Meanwhile, a number of candidate cold tolerance genes were discovered in the Strigens’ study. Hund et al. (2004) used the F2:3 population, Lo964 × Lo1016, to identify QTL controlling cold-related traits of maize seedlings under cold stress. A major QTL was found on chromosome 5 that could explain 14% of the total phenotypic variation of root diameters. Fracheboud et al. (2002, 2004) chlorophyll fluorescence parameters, CO2 exchange rate, leaf greenness, shoot dry matter, and shoot nitrogen content as phenotype in an F2:3 population under cold stress and found that a major QTL located on chromosome 6 explained 37.4% of the phenotypic variance in the chronic photoinhibition at low temperature. Rodriguez et al. (2014) used the F2:3 population, EP42 and A661, to map QTLs using four cold tolerance-related traits, number of survival plants, dry weight, quantum yield of photosystem II, and total anthocyanin content. Four QTLs associated with cold stress were identified three genomic regions in chromosomes 2, 4, and 8.

In our study, we used GWAS and QTL mapping to identify major SNPs and QTL in two populations of maize. We used an integrative analysis approach, combining the results of the GWAS and QTL mapping, to further identify key candidate genes for cold tolerance. We then performed a functional annotation and cloning into Z. mays for functional verification of these candidate genes. Using these methods, we aimed to provide the basis for developing possible strategies for breeding new maize with improved cold tolerance, thus improving food security.

MATERIALS AND METHODS

Plant Materials and Experimental Design

There were two genetic populations in our study: the first (population 1) was a natural population consisting of 80 backbone inbred lines. These lines were selected from numerous maize materials based on their frequency and genetic diversity in maize breeding in Jilin Province, China (Supplementary File 1). The second population (population 2) was an F2 population of 210 offspring, with parents derived from population 1 (W72 × W10). Two lines W10 (chilling-tolerant) and W72 (chilling-sensitive) were used as parents to produce a segregating F2 population. To obtain F3 hybrids, we self-crossed F2 individuals, and then self-crossed F3 individuals to obtain F2:3 seeds. The above experimental materials were provided by the Biotechnology Center of Jilin Agricultural University, China.

Analysis of Phenotypic Data and Peroxidase Activity

We measured the peroxidase (POD) activity of each maize inbred line using the guaiacol method (Javadian et al., 2010) at seedling stage. Three consecutive batches of tests were
performed on each inbred line at a low temperature of 6°C for 24h. Ten strains of the same line were tested in each test, and average values of all 30 strains were obtained for subsequent analysis, which was conducted using SPSS 23.0 software (Pallant, 2013).

The formula for calculating POD activity was as:

\[
\text{POD} \left( \frac{U}{g \times \text{min}} \right) = \frac{A_{470} \cdot V_t}{0.01 \times W \cdot V_s \cdot t}
\]

where \(A_{470}\) was the change in absorbance within the reaction time; \(W\) was the fresh weight (g); \(V_t\) was the total volume of extracted enzyme solution (ml); \(V_s\) was the volume of enzyme solution during measurement (ml); and \(t\) was time (min).

**DNA Extraction for Sequencing in Populations 1 and 2**

The leaves of populations 1 and 2 were collected and flash frozen in liquid nitrogen. The DNA was extracted using the NuClean Plant Genomic DNA Kit (CWBio, Jiangsu, China); the concentration and purity of the extracted DNA were determined using a Nano Drop 2000 micro spectrophotometer (Thermo Fisher, Waltham, MA, United States). The integrity of the extracted DNA was confirmed using agarose gel electrophoresis (agarose concentration 1%, voltage 200V, time 25 min; Zhao et al., 2018).

**Whole-Genome Sequencing and SNP Identification in Population 1**

We used the NuClean Plant Genomic DNA Kit (CWBio, Jiangsu, China) to extract genomic DNA from the 80 inbred maize lines in population 1 at seedling stage. DNA quality was assessed using 1% agarose gel electrophoresis under 150 volts 70 milliamperes, 30 min and a Nano Drop 2000 spectrophotometer (NanoDrop, Wilmington, DE, United States). Qualified DNA samples with clear bands were used to construct a genomic library. Sequencing was performed using an Illumina HiSeq PE150 (Illumina Inc., San Diego, CA, United States) to obtain raw data. Then, read pairs containing the linker sequence were filtered out; paired reads with N content in a single-end sequencing read exceeding 10% of the length of the read, and those with a number of low-quality bases (defined as quality value \(Q \leq 5\)) in a single-end sequencing read exceeding 10% of the length of the read, were removed. Finally, the sequencing data were left high-quality, clean genomic data.

Maize B73 RefGen_v4 was the data mining resource of the Maize Genetics and Genome Database (MaizeGDB). We compared these clean data to the reference genome, Maize B73 RefGen_v4 (Shaminuzzaman et al., 2020), using the Burrows-Wheeler Aligner (BWA) software (Jia et al., 2013), and SAMtools software (Zhou et al., 2015) was used to remove duplicates (command: rmdup). Then, we used a Bayesian model to detect polymorphic sites in the population, and filtered and screened high-quality SNPs using a minor allele frequency (MAF) > 0.05 and call rate > 80%. Finally, we used ANNOVAR software (Yang and Wang, 2015) to functionally annotate the identified SNPs.

**Population Structure, Linkage Disequilibrium, GWAS, and Candidate Gene Annotation in Population 1**

The phylogenetic tree was inferred using TreeBест2 using the \(p\)-distances genetic distance estimation model and the identified SNPs. The Principal Component Analysis (PCA) was analyzed by Genome-wide Complex Trait Analysis (GCTA; Yang et al., 2011).3

We used Hapview (Barrett et al., 2005) to calculate linkage disequilibrium (LD) in population 1 and FarmCPU to perform a GWAS to identify associations between genetic loci and phenotypic traits (Kaler et al., 2020). Research in recent years had shown that the SNPs at which the maximum value of \(R^2\) (corresponding to the LD) decays to half are linked. When \(R^2\) decayed to half, the corresponding LD was about 5.0 kb. Therefore, we chose a LD of 5.0 kb to scan for functional genes. All identified functional genes were annotated using the SWISS-PROT,4 COG,5 GO,6 KEGG,7 and NCBI8 databases; possible candidate cold tolerance genes were screened based on previous studies.

**Genotyping by Sequencing in Population 2**

Given that the parental line of population 2 had already been sequenced among population 1, we used GBS to sequence population 2 only. We used qualified samples with clear bands in electrophoresis for GBS and quality analysis of SNPs. Sequencing was completed using Illumina NovaSeq 6,000 (Illumina Inc., San Diego, CA, United States).

The GBS data were filtered to obtain clean data, using the same criteria as for population 1, detailed in section “Whole-Genome Sequencing and SNP Identification in Population 1.”

We used BWA software (Jia et al., 2013) to compare the clean data with the reference genome (Maize B73 RefGen_v4) and used SAMtools to convert the compared result format into SAM/BAM files; in addition, the Perl script was used to calculate the mapping rate and coverage, and SAMtools “sort” command was used to detect mutations. Polymorphic tags of parental line were genotyped to facilitate subsequent genetic analysis. Because the parents were homozygous inbred lines with aa and bb genotypes, SNPs with the segregating pattern of aa×bb were screened (Zhang et al., 2016a). Three standards were used for the screening of SNP: first, in the offspring typing, there may be a few base types that did not appear in the parents, and we considered them as deletions; second, included genotypes covered at least 80% of SNPs in the offspring;

---

1https://www.maizegdb.org/gene/assembly/
Zm-B73-REFERENCE-GRAMENE-4.0

2http://treesofsourceforge.net/
3https://yanglab.westlake.edu.cn/software/gcta/#Overview
4http://www.uniprot.org/
5http://clov.org/docs/clusters-of-orthologous-groups-cogs/
6http://geneontology.org/
7https://www.genome.jp/kegg/
8https://www.ncbi.nlm.nih.gov/
Jin et al. Spring Maize Cold Tolerance Genes

by RT-PCR,® Stam, 1993). Jiang et al., 2018); and Meng Supplementary File 2 Zhao et al., 2016.

Frame et al., 2006) to design found the presence of Zm00001d002729 Zhang et al., 2016b. Wei et al., 2009 cloning vector ). Then, 9 ) software to construct a Lu

9, J8, J98, J183, and J191; and extremely cold-sensitive strains resistant strains J15, J75, and J144; general cold-resistant strains W48, and W72. In population 2, we selected extremely cold-sensitive strains W43, W10, W18, and W26; general cold-resistant strains W1, W15, W36, and W52; and extremely cold-sensitive strains W43, W48, and W72. In population 2, we selected extremely cold-resistant strains J15, J75, and J144; general cold-resistant strains J8, J98, J183, and J191; and extremely cold-sensitive strains J107, J123, and J200. We extracted RNA from leaves of each inbred line as a template at seedling stage and measured the relative expression of Zm00001d002729 in each line using quantitative real-time PCR (qRT-PCR) and the fluorescent dye SYBR Green (Takara Biomedical Technology Co., Ltd., China). The internal reference gene we chose was Actin-1 (Act; Accession: J01238). The primers used are given in Supplementary File 2.

Total RNA was extracted using the Plant RNA Kit R6827-02 (Omega Bio-tek, United States), and then reverse-transcribed into cDNA using the All-in-One™ First-Strand cDNA Synthesis Kit (GeneCopeia Inc., United States); qRT-PCR primers were designed using an online tool (https://sg.idtdna.com/ PrimerQuest/Home/Index). qRT-PCR was performed using the All-in-One™ qPCR Mix (GeneCopeia Inc., United States) and the Stratagene Mx3000P PCR system (Agilent Technologies, United States) with three repetitions, and the relative expression level of the genes was calculated using the 2-ΔΔCT method (Lu et al., 2018).

Cloning of Zm00001d002729 and Vector Construction

We used Primer Premier 5.0 (Wei et al., 2009) to design specific primers for cloning Zm00001d002729 by RT-PCR, using cDNA from leaf tissue of maize strain Dan 340 at seedling stage as the template, and annealing temperature was about 55.0°C (Supplementary File 2). Then, we constructed the pMD-18T-Zm00001d002729 cloning vector and sequenced it to confirm successful cloning. We used the Seamless Assembly Cloning Kit (Taike Biotechnology Co., Ltd., China) to construct the pCAMBIA3301-Zm00001d002729 expression vector. We used the CRISPR-P Web site® to design small guide RNA (sgRNA). We selected sgRNA with high-target scores, low miss-target rates, and used Seamless cloning to construct a pCXB053-Zm00001d002729-CRISPR/Cas9 vector with a single gene and dual targets. Meanwhile, we used U6 as a promoter to construct a CRISPR/Cas9 gene editing vector.

Genetic Transformation of Zm00001d002729 Into Z. mays H299 and Functional Verification

In order to verify the function of Zm00001d002729, we transformed Zm00001d002729 into strain H299 using an Agrobacterium-mediated method (Frame et al., 2006). In this study, Agrobacterium-mediated methods were used to transfer the constructed expression vector to the recipient. After embryoculture, subculture, pre-culture, co-culture, screening culture, elongation culture, and rooting culture, we obtained T0 maize and cultivated it to the T2 generation. We performed test strip detection for bar gene by GeneTure™ BAR Test Kit (Artron Laboratories Inc.) and qRT-PCR detection on the transformed maize. Finally, we cultivated the T2 generation containing pCAMBIA3301-Zm00001d002729 or CRISPR/

® http://crispr.hzau.edu.cn/CRISPR2/

and third, segregation distortion. We used the chi-square test to assess offspring with a significance threshold of p < 0.001.

Construction and Evaluation of the Genetic Linkage Map

We used Join Map 4.1(Stam, 1993) software to construct a genetic linkage map of high-quality SNPs obtained after screening. First, the linkage group was divided into chromosomes; second, a maximum likelihood method was used to sort the markers in each linkage group. That was, A and B were linked, B and C were linked, and then, A and C were linked; otherwise, B and C were not linked, then A and C were not linked. Third, the genetic distance between markers was calculated using Kosambi mapping function. Finally, we used the Perl SVG module to draw the genetic linkage map.

To ensure the quality of the genetic linkage map, we created a map of the haploid source of population 2 in 10 linkage groups, and possible double crossover sites were identified; we also evaluated linkage relationships between each of the markers in the same linkage group on the map (Zhao et al., 2016); and finally, we performed a collinearity analysis by the R programming language using marker positions on the genome and the genetic map to perform collinearity analysis (Zhang et al., 2016b).

Screening QTLs Associated With POD

QTLs were detected using composite interval mapping (CIM) in WinQTL Cart 2.5 software (Jiang et al., 2018) utilizing the phenotype and genotype data from population 2. The Kosambi function was used with the scan step set as 1 cM, and the maximum value of p for entering variables in a stepwise regression of residual phenotype on marker variables (PIN) was set as 0.001. LOD (=0.217 likelihood ratio) threshold for declaring a putative QTL for each trait, data set, and model was defined by 1,000 permutations. Finally, we selected a QTL with a threshold logarithm of the odds (LOD)>2.5 (Meng et al., 2015). Fracheboud et al. (2002) found the presence of a QTL was declared significant if the LOD value was >2.5 for a single trait analysis.

Integrated GWAS and QTL Mapping Analysis

We integrated the results of the GWAS and QTL mapping in populations 1 and 2 to identify SNPs that were significantly associated with the POD phenotype. We used an LD of 5.0 kb to scan for functional genes.

Expression Analysis of the Candidate Cold Tolerance Gene Zm00001d002729

We selected 10 strains in each of population 1 and 2 for expression analysis of the candidate gene Zm00001d002729; in population 1, we selected the extremely cold-resistant strains W10, W18, and W26; general cold-resistant strains W1, W15, W36, and W52; and extremely cold-sensitive strains W43, W48, and W72. In population 2, we selected extremely cold-resistant strains J15, J75, and J144; general cold-resistant strains J8, J98, J183, and J191; and extremely cold-sensitive strains J107, J123, and J200. We extracted RNA from leaves of each inbred line as a template at seedling stage and measured the relative expression of Zm00001d002729 in each line using quantitative real-time PCR (qRT-PCR) and the fluorescent dye SYBR Green (Takara Biomedical Technology Co., Ltd., China). The internal reference gene we chose was Actin-1 (Act; Accession: J01238). The primers used are given in Supplementary File 2.

Total RNA was extracted using the Plant RNA Kit R6827-02 (Omega Bio-tek, United States), and then reverse-transcribed into cDNA using the All-in-One™ First-Strand cDNA Synthesis Kit (GeneCopeia Inc., United States); qRT-PCR primers were designed using an online tool (https://sg.idtdna.com/ PrimerQuest/Home/Index). qRT-PCR was performed using the All-in-One™ qPCR Mix (GeneCopeia Inc., United States) and the Stratagene Mx3000P PCR system (Agilent Technologies, United States) with three repetitions, and the relative expression level of the genes was calculated using the 2-ΔΔCT method (Lu et al., 2018).

Cloning of Zm00001d002729 and Vector Construction

We used Primer Premier 5.0 (Wei et al., 2009) to design specific primers for cloning Zm00001d002729 by RT-PCR, using cDNA from leaf tissue of maize strain Dan 340 at seedling stage as the template, and annealing temperature was about 55.0°C (Supplementary File 2). Then, we constructed the pMD-18T-Zm00001d002729 cloning vector and sequenced it to confirm successful cloning. We used the Seamless Assembly Cloning Kit (Taike Biotechnology Co., Ltd., China) to construct the pCAMBIA3301-Zm00001d002729 expression vector. We used the CRISPR-P Web site® to design small guide RNA (sgRNA). We selected sgRNA with high-target scores, low miss-target rates, and used Seamless cloning to construct a pCXB053-Zm00001d002729-CRISPR/Cas9 vector with a single gene and dual targets. Meanwhile, we used U6 as a promoter to construct a CRISPR/Cas9 gene editing vector.

Genetic Transformation of Zm00001d002729 Into Z. mays H299 and Functional Verification

In order to verify the function of Zm00001d002729, we transformed Zm00001d002729 into strain H299 using an Agrobacterium-mediated method (Frame et al., 2006). In this study, Agrobacterium-mediated methods were used to transfer the constructed expression vector to the recipient. After embryoculture, subculture, pre-culture, co-culture, screening culture, elongation culture, and rooting culture, we obtained T0 maize and cultivated it to the T2 generation. We performed test strip detection for bar gene by GeneTure™ BAR Test Kit (Artron Laboratories Inc.) and qRT-PCR detection on the transformed maize. Finally, we cultivated the T2 generation containing pCAMBIA3301-Zm00001d002729 or CRISPR/
Cas9-Zm00001d002729, and three wild-type maize plants in total, to under conditions of cold stress at 6°C for 24 h at seedling stage, monitoring their morphology and measuring POD activity, malondialdehyde (MDA) content, and relative conductivity of seedlings leaves.

RESULTS

Analysis of Phenotypic Data and POD Activity
The POD activity phenotype data of two population were shown in Supplementary File 3. The average POD activity in population 1 was 46.95 U/g × min, with a range of 22.97–100.75 U/g × min, and a coefficient of variation of 0.30. The average POD activity in population 2 was 55.19 U/g × min, with a range of 27.44–83.37 U/g × min, and a coefficient of variation of 0.28. These data were found to be normally distributed according to the Kolmogorov-Smirnov test (p = 0.20) of normality (Figure 1).

Whole-Genome Sequencing and SNP Identification in Population 1
The quality of DNA extracted from the sample was high and the fragments were clear (Figure 2), thus meeting the requirements of DNA library construction for sequencing.

We performed WGS on population 1 and obtained 1.58 TB of clean data after quality control. Comparison with the maize inbred line reference genome showed that the mapping rate of the population samples was 98.72%. The average sequencing depth was 17.10×, and the average genome coverage was 89.16%. The similarity between each sample and the reference genome met the requirements of WGS analysis and had a very good sequencing depth and coverage. We obtained a total of 24,860,241 high-quality SNPs, of which 85.4% were intergenic. The results of SNP genotyping are shown in Supplementary File 4.

Population Structure and Linkage Disequilibrium in Population 1
We constructed a phylogenetic tree for population 1, based on genotype data of 80 inbred lines (Figure 3A). The results
showed three subpopulations: subpopulation 1 (in red) contains lines mainly from Tang Si Ping Tou and Lv Da Red Bone; subpopulation 2 (in blue) contains lines mainly from Lancaster; and subpopulation 3 (in yellow) contains lines mainly from Reid. The PCA results were shown in Figure 3B and Supplementary File 5. The results showed the distributions of inbred lines, and they were separated accordingly.

The distribution of LD on chromosomes is described in the LD attenuation diagram (Figure 3C). The abscissa represents the distance at which LD occurs, and the ordinate is the correlation coefficient of LD, a way of displaying the LD value.

The attenuation distance was 3.7 kb in chromosome 1; 3.6 kb in chromosome 2; 3.5 kb in chromosome 3; 3.5 kb in chromosome 4; 3.1 kb in chromosome 5; 3.1 kb in chromosome 6; 2.6 kb in chromosome 7; 3.2 kb in chromosome 8; 3.8 kb in chromosome 9; and 3.2 kb in chromosome 10. The average attenuation distance of the whole genome was 3.4 kb.

**GWAS to Identify Genes Associated With POD**

The Manhattan chart and Q-Q chart for the GWAS for POD activity are shown in Figure 4. We used \( p < 0.000001 \) as the threshold; our analysis detected four SNPs that were significantly related to POD activity. These SNPs were all located on chromosome 3 (Table 1).

**GBS, Development, Encoding, and Screening of SNPs in Population 2**

The two parents from population 1 (whose genomes were sequenced using WGS) and 210 offspring from population 2 (GBS) were successfully sequenced with an average sequencing depth of 30× and 10×, respectively. Our results yielded a total of 261.005 GB raw data from the offspring in population 2, with an average of 1,243 GB of raw data for each sample, and the filtered clean data was 260.879 GB, with 1.242 GB on average in each sample. The sequencing quality for the offspring was very high (Q20 ≥ 94% and Q30 ≥ 85%), and the sample GC distribution was normal. The database was successfully constructed and sequenced.

On comparison with the maize reference genome, 9,141,219 and 6,127,709 SNPs were identified in the parent plants, W72 and W10, respectively. A total of 5,759,868 SNPs were identified in population 2 compared with the reference genome, and 1,847,696 polymorphic SNPs were assigned to the parental aa × bb segregation pattern (Supplementary File 6). After

![Figure 3](image1.png)

**Figure 3** | (A) The population structure, (B) Principal Component Analysis (PCA), and (C) linkage disequilibrium (LD) diagram for population 1.

![Figure 4](image2.png)

**Figure 4** | (A) Manhattan plot of significant SNPs correlating with POD activity using the FarmCPU model; (B) Quantile-Quantile (Q-Q) plot.
filtering, 26,693 SNPs remained. We used these SNPs for genetic linkage map construction. SNPs were screened and adjusted iteratively to optimize the map following its initial construction.

### Construction and Evaluation of the Genetic Linkage Map

Each chromosome constituted a linkage group. A total of 8,213 SNPs were shown on the map, with a total genetic distance of 6516.35 cM. The average genetic distance between two adjacent markers was 0.79 cM. There were 1,163 markers in the first linkage group covering a genetic distance of 1114.99 cM: the longest genetic distance among the 10 linkage groups. On the other hand, there were 619 markers in the fifth linkage group over a genetic distance of 414.49 cM, covering the shortest genetic distance among the 10 linkage groups (Supplementary File 7).

We also analyzed the distribution of gaps in each linkage group (Supplementary File 8); most gaps were less than 5 cM. We used the Perl SVG module to draw a genetic linkage map (Supplementary File 9) and evaluated the quality of the map by constructing a map of the haploid source with all markers in each linkage group from population 2. The haploid source of the large intervals observed in every offspring was one of the parents, and the proportion of double crossing-over was less than 3%, indicating high quality of the linkage map (Supplementary File 9). We constructed a heat map of linkage relationships between each marker in the same linkage group from the genetic linkage map. The linkage relationship between adjacent markers in each group was very strong. With increasing distance, the linkage relationship between markers weakened gradually, indicating that the order of markers in the map was correct (Supplementary File 9). Finally, we analyzed the collinearity of the markers on the genome and genetic map. The collinearity map can not only reflect the collinearity of the position of the markers on the genome and genetic map, but also reflect the distribution of the markers on the genetic map mapped to the chromosome in general. It can also reflect the relationship of each marker on the genome and genetic map. It can be seen from this figure that the distribution of markers on the chromosomes is relatively uniform, most of the marker sequences on each linkage group are basically consistent with the genome, the collinearity is good, and the calculation accuracy of the genetic recombination rate is high (Supplementary File 9).

### Screening of QTLs Associated With POD

We detected a total of 12 QTLs related to POD. Among them, 11 QTLs were located on chromosome 2, and 1 QTL was located on chromosome 3. The QTL qPOD2b was located on chromosome 2, which had the largest contribution to POD activity of all QTLs, at 6.27%, with an additive effect of 52.80 U/g × min and a dominant effect of 6.51 U/g × min. qPOD3 was located on chromosome 3, with a contribution to the POD phenotype of 5.60%, an additive effect of 1.17, and a dominant effect of 4.52 (Figure 5; Table 2).

### Annotation and Screening of Cold Tolerance Candidate Genes in Populations 1 and 2

We integrated the results of the GWAS and QTL mapping in two populations to identify SNPs and QTLs that were significantly associated with the POD activity phenotype. Then, we used an LD of 5.0 kb to scan for functional genes. Our integrated GWAS and QTL mapping analysis revealed that the SNP (sPOD-20,592,562) was located in the QTL qPOD3 and that this SNP was associated with the gene Zm00001d002729. It meant that Zm00001d002729 has been identified in both populations, which greatly increased the credibility of this gene. Our functional annotation indicated that this gene was an acyl-acyl carrier protein thioesterase.

We used Cell-PLoc 2.010 to predict the subcellular localization of the Zm00001d002729 protein (Chou and Shen, 2008). The result showed that the predicted location was chloroplast and nuclear.

### Table 1

| Name        | Start (bp) | End (bp) | peakPOS (bp) | chr | ref | alt | peak_value | Candidate gene | LD |
|-------------|------------|----------|--------------|-----|-----|-----|------------|----------------|----|
| sPOD-20,592,562 | 20,587,562 | 20,597,562 | 20,592,562 | 3   | A   | G   | 6.10       | Zm00001d002729 | 0   |
| sPOD-96,084,381 | 96,079,381 | 96,089,381 | 96,084,381 | 3   | C   | T   | 6.17       | Zm00001d0041082 | 3,183 |
| sPOD-133,504,984 | 133,499,956 | 133,509,984 | 133,504,984 | 3   | T   | A   | 6.60       | Zm00001d0041691 | 3,551 |
| sPOD-140,877,969 | 140,872,969 | 140,882,969 | 14,087,7,969 | 3   | T   | C   | 6.33       | Zm00001d0041853 | 1,384 |

10http://www.csbio.sjtu.edu.cn/bioinf/Cell-PLoc-2/
Expression Analysis of Cold Tolerance Candidate Gene Zm00001d002729

We analyzed the relative expression levels of the candidate gene Zm00001d002729 seedling stage in the two populations. The results showed that the expression level of this cold tolerance candidate gene in maize strains with better cold tolerance (W10, W18, W26, J15, J75, and J144) was significantly (*p* < 0.05) higher than that in cold-sensitive maize strains (W43, W48, W72, J107, J123, and J200). From this we concluded that higher expression levels of Zm00001d002729 in maize result in greater cold tolerance (Figure 6).

Cloning and Vector Construction of Zm00001d002729

We successfully obtained Zm00001d002729 cDNA by RT-PCR method by using RNA extracted from the leaves seedling stage Dan 340 as the template. The full length of Zm00001d002729 coding sequence was 690 bp.

After construction of the cloning vector, we used Seamless cloning technology to construct the plant over expression vector and CRISPR/Cas9 vector. Firstly, we used pCAMBIA3301 as the vector backbone and constructed pCAMBIA3301-Zm00001d002729. Meanwhile, we used U6 as a promoter to construct a CRISPR/Cas9 gene editing vector. Two carrier structures are shown in Supplementary File 10.

Functional Verification of Zm00001d002729

We transformed Zm00001d002729 into Z. mays H299 via Agrobacterium-mediated method; the growth process is shown in Supplementary File 11. After testing to confirm successful transformation using the GeneTure™ BAR Test Kit (Artron Laboratories Inc.; Supplementary File 12), we cultivated the genetically modified maize to the T2 generation.

We also measured the POD activity, MDA content, and relative conductivity of these nine inbred maize lines after 24 h at 6°C (Figure 7). We found that in maize with overexpression of the target gene, the POD activity was significantly increased, the MDA content significantly decreased, and the relative conductivity was also significantly decreased (all *p* < 0.05); while in maize where the target gene was silenced, POD activity was decreased, and the MDA content and relative conductivity significantly increased. This indicates that Zm00001d002729 is a positive regulator of cold resistance in maize. The higher the expression of Zm00001d002729, which could increase cold tolerance of maize.

| Name | chr | Left marker | Right marker | Left position (bp) | Right position (bp) | LOD | Additive | Dominance | PVE (%) |
|------|-----|-------------|--------------|-------------------|---------------------|-----|-----------|-----------|--------|
| qPOD2a | 2 | mk-2-411 | mk-2-1,108 | 73,584,577 | 143,233,161 | 3.14 | − | 57.66 | 6.30 |
| qPOD2b | 2 | mk-2-1,107 | mk-2-561 | 143,233,142 | 92,502,399 | 2.75 | − | 52.90 | 6.11 |
| qPOD2c | 2 | mk-2-561 | mk-2-801 | 92,502,399 | 123,536,831 | 3.11 | − | 55.91 | 16.18 |
| qPOD2d | 2 | mk-2-801 | mk-2-950 | 123,536,831 | 133,026,652 | 2.75 | − | 58.22 | 6.45 |
| qPOD2e | 2 | mk-2-972 | mk-2-1,062 | 133,419,121 | 143,771,596 | 2.93 | − | 119.36 | 5.84 |
| qPOD2f | 2 | mk-2-947 | mk-2-1,133 | 132,958,156 | 115,063,134 | 3.15 | − | 119.36 | 6.19 |
| qPOD2g | 2 | mk-2-1,133 | mk-2-731 | 143,771,596 | 115,063,134 | 2.53 | − | 119.36 | 5.89 |
| qPOD2h | 2 | mk-2-731 | mk-2-1,152 | 115,063,134 | 144,527,652 | 2.53 | − | 119.36 | 5.89 |
| qPOD2i | 2 | mk-2-1,157 | mk-2-1,113 | 144,789,539 | 143,233,161 | 2.97 | − | 30.49 | 5.11 |
| qPOD2j | 2 | mk-2-1,113 | mk-2-998 | 143,233,161 | 133,848,401 | 2.81 | − | 120.31 | 5.75 |
| qPOD3 | 3 | mk-3-76 | mk-3-78 | 20,191,362 | 23,118,526 | 2.61 | − | 119.36 | 5.99 |

TABLE 2 | Quantitative trait locus (QTLs) associated with POD activity.
DISCUSSION

POD Activity and Maize Cold Resistance
In our study, we demonstrated that cold-tolerant maize had higher POD activity than cold-sensitive maize. The maize antioxidant stress pathway is a known important mechanism underlying cold tolerance (Özlem et al., 2013; Zhang et al., 2020; Yu et al., 2021). Most research on antioxidant enzymes has so far focused on maize catalase, which removes excess H₂O₂ from mitochondria and protects plants from cold stress (Prasad, 1997). Studies had shown that in the early stages of cold stress, POD activity in maize seedlings declines and then rises. With increased time and enhanced intensity of cold stress, POD activity has been shown to increased significantly (Guy et al., 1997; Roncatto and Pascholati, 1998; Šukalović et al., 2010; Ao et al., 2013). POD activity has been shown in previous studies to be higher in cold-tolerant than cold-sensitive varieties (Hodges et al., 1996; Wang et al., 2015), which is consistent with our results.

QTLs Associated With Maize Cold Tolerance
The QTL we found to be associated with cold tolerance was on chromosome 3 in maize, located in bin 3.04. Fracheboud et al. (2002, 2004) studied the effect of low temperature on the photosynthetic mechanism of maize and analyzed five QTLs related to this function; one of these QTLs was located near the mmc0022 marker on chromosome 3, and explained 28% of the phenotypic variation, believed to be related to photosynthetic function at low temperatures in maize. The relevant gene is likely to be located on chromosome 3, close to the centromere; this is very close to bin 3.04. Interestingly, there are previous reports of other biotic or abiotic stress-related QTLs located near the QTL we identified as: Messmer et al. (2009) detected three drought-related stable QTLs located on chromosome 3 bin 3.04, and their contributions of phenotypic variation in this trait were 8.9, 10.8, and 6.3%. Balint-Kurti identified QTLs related to southern leaf blight resistance in two RIL populations (H99 × B73 and B73 × B52), also located in bin 3.04 (Balint-Kurti et al., 2008). Based on a genetic map of 246 polymorphic simple sequence repeat markers with an average genetic distance of 9.1 cM, Ding et al. (2008) used the compound interval mapping method (CIM) for QTL mapping to analyze associations with maize fusarium ear rot under four environments in a RIL population. Two QTLs, both located at bin 3.04 of chromosome 3, were detected in all environments, and these QTLs were located on either side of the markers umc1025 and umc1742, respectively; the QTL we identified in this study was very close to umc1742. Guan et al. (2012) found that a Ragged leaves 1
(Rg1) maize mutant is susceptible to disease caused by abiotic stress, resulting in the extensive accumulation of H₂O₂ in diseased tissues and lesions on the leaves, sheaths, and bracts of maize. Interestingly, this is similar to the mechanism by which plants are subjected to cold stress (Weig et al., 1997). However, Rg1 is located on chromosome 3 bin 3.04. In summary, for the QTL we discovered, there were almost no previous reports of association with maize cold tolerance traits, though there were some with tolerance to other stresses. This indicates that we have discovered a new QTL related to maize cold tolerance.

Analysis of Zm00001d002729 Associated With Maize Cold Tolerance

Functional annotation suggested that Zm00001d002729 is an acyl-acyl carrier protein thioesterase. We found that overexpression of Zm00001d002729 increases the cold tolerance of maize, while inhibiting the expression of this gene increases sensitivity to low temperatures. In previous studies, acyl-acyl carrier protein thioesterase is the key enzyme in the regulation of fatty acid carbon chain elongation during fatty acid synthesis. It can catalyze the hydrolysis of thioester bonds and terminate the synthesis of fatty acids. Its activity thus affects the content of various fatty acids in plants (Nieto et al., 2007). However, there are relatively few functional verifications of this activity in plant resistance to stress (Dormann et al., 2000; Wu et al., 2009). In a study by Raffaele et al. (2008), it was shown that mutation in acyl-acyl carrier protein thioesterase in Arabidopsis thaliana leads to serious defects in the fatty acid supply for long-chain fatty acid biosynthesis. Arabidopsis thaliana also lost its hypersensitive response to pathogen attack. Zhang et al. (2012) found that the overexpression of acyl-acyl carrier protein thioesterase in tobacco improves drought tolerance. The phase transition hypothesis of cold stress damage in plants states that when temperate plants suffer low temperature damage below a certain threshold, the biological membrane will first undergo a phase transformation of membrane lipids, from a liquid crystal phase to a gel phase. The fatty acid chains in the membrane lipids change from a disorderly to an orderly arrangement; the fluidity of the membrane is greatly reduced, and pores or cracks form in the membrane, increasing the permeability of the membrane and changing the structure of membrane-bound proteins. The efflux of a large amount of materials in the membrane destroys the ion balance between the inside and outside of the cell, leading to changes in physiological metabolism and functional disorders in plant cells (Lyons and Raison, 1970). We quoted the expression levels of Zm00001d002729 from MaizeGDB (Supplementary File 13). We found that RNA-seq showed that the expression of target genes in V1 and V3 phases was extremely low in B73 (Winter et al., 2007; Makarevitch et al., 2015; Hoopes et al., 2019). In previous studies, it was found that B73 belong to cold-sensitive inbred line (Joets et al., 2018; Grzybowski et al., 2019; Jonczyk et al., 2020). Therefore, we speculated that the more cold-sensitive maize, the lower the expression levels of Zm00001d002729. It is consistent with the conclusions we got before. We do not yet understand the mechanism by which Zm00001d002729 confers greater cold tolerance, and how this might fit with the phase transition hypothesis; however, our identification of the general function of this gene has laid the foundation for elucidating and exploiting this mechanism through genetic engineering in crops.

In conclusion, we successfully used GWAS and QTL mapping to successfully identify Zm00001d002729, a gene on maize chromosome 3 that is significantly related to POD activity. We used genetic transformation to overexpress or inhibit the expression of Zm00001d002729 in maize, verifying its function as a cold tolerance gene. We have shown for the first time that Zm00001d002729 positively regulates cold resistance in maize and may be a potentially useful tool to generate cold-resistant crops to ensure food security in the face of climate change.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found at: https://www.ncbi.nlm.nih.gov/, PRJNA495031 and PRJNA763107.

AUTHOR CONTRIBUTIONS

YJ drafted the original manuscript. ZZ reviewed and edited the manuscript. YX and ZY conducted the experiments and performed the analysis. ZX, SG, and JQ performed the phenotyping. PW conducted formal analysis. RZ supervised the experiments. All authors contributed to the article and approved the submitted version.

FUNDING

This research was supported by the Department of Science and Technology of Jilin Province (20200402026NC) and the National Key Research and Development Program of China (2019YFD1002603-1).

ACKNOWLEDGMENTS

We would like to express our gratitude to the Novogene Co., Ltd. (Beijing, China) for their kind support in sequencing. We also thank Jilin Agricultural University, Biotechnology Center for facilitating the working environment to conduct the experiment.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at https://www.frontiersin.org/articles/10.3389/fpls.2021.776972/full#supplementary-material
Published in: Frontiers in Plant Science
Vol 12 | Article 776972

The researchers identified 23 maize genes involved in cold stress resistance. These genes are responsible for different aspects of maize's cold tolerance, including changes in peroxidase activity and electrophoretic pattern. The study also examined the role of catalase in inducing chilling tolerance in pre-emergent maize seedlings. The results suggest potential targets for improving maize's cold tolerance.