nature genetics

Genetic variation in *ZmVPP1* contributes to drought tolerance in maize seedlings

Xianglan Wang[1,2,6], Hongwei Wang[1,2,6], Shengxue Liu[1], Ali Ferjani[3], Jiansheng Li[4], Jianbing Yan[5], Xiaohong Yang[4] & Feng Qin[1]

Maize production is threatened by drought stress worldwide. Identification of the genetic components underlying drought tolerance in maize is of great importance. Here we report a genome-wide association study (GWAS) of maize drought tolerance at the seedling stage that identified 83 genetic variants, which were resolved to 42 candidate genes. The peak GWAS signal showed that the natural variation in *ZmVPP1*, encoding a vacuolar-type H^+ pyrophosphatase, contributes most significantly to the trait. Further analysis showed that a 366-bp insertion in the promoter, containing three MYB *cis* elements, confers drought-inducible expression of *ZmVPP1* in drought-tolerant genotypes. Transgenic maize with enhanced *ZmVPP1* expression exhibits improved drought tolerance that is most likely due to enhanced photosynthetic efficiency and root development. Taken together, this information provides important genetic insights into the natural variation of maize drought tolerance. The identified loci or genes can serve as direct targets for both genetic engineering and selection for maize trait improvement.

Because of its high yield potential and nutritional value, maize (*Zea mays*) is a major crop grown for food, feed, and fuel. Increased yields of maize have been steadily achieved over the past decades, although its sensitivity to drought stress has also concomitantly increased[1]. Water scarcity, large fluctuations in weather patterns, and the unpredictable nature of drought result in a substantial threat to maize production worldwide[2,3]. Drought has even been thought of as a "cancer" to plants, owing to its complexity and destructiveness[4]. Thus, there is tremendous interest in and demand for improving maize drought tolerance through biotechnology.

Previous studies in the model plant *Arabidopsis thaliana* have elucidated complex drought-responsive pathways and metabolic networks[5–7]. However, the genetic basis of quantitative variation for drought tolerance in agriculturally important crops remains poorly understood. Recently, GWAS, based on linkage disequilibrium (LD), have gained more favorability in the genetic dissection of complex traits in crops because a greater number of alleles, broader genetic variations, and more historical recombinations can be investigated in an association study as compared to linkage analysis[8,9]. For example, the genetic basis of important agronomic traits was successfully characterized in rice by GWAS, with a large collection of rice germplasm[10,11]. In maize, as a kind of out-pollination species, the LD decay was estimated to be ≤2 kb in diverse inbred lines[12,13]. Thus, a maize GWAS is able to provide gene-level resolution[14], which facilitated the gene cloning of *ZmCCT*, controlling maize photoperiod sensitivity[15,16], and of a maize gene encoding glutathione *S*-transferase (GST), which impairs multiple-disease resistance[17].

Although several studies have reported the association of DNA polymorphisms with maize drought tolerance[18–22], genes underlying maize drought tolerance remain to be elucidated owing to a lack of functional validation. Here we report a GWAS for maize drought tolerance at the seedling stage in a natural-variation population[13]. On the whole-genome scale, 83 genetic variants, resolved to a total of 42 candidate genes, were significantly associated with drought tolerance in maize seedlings. The most significant variation occurs within a gene, *ZmVPP1*, encoding a vacuolar H^+ pyrophosphatase. A 366-bp insertion in the promoter of *ZmVPP1* confers drought-inducible expression of *ZmVPP1* in drought-tolerant maize inbred lines, and genetic selection of this *ZmVPP1* allele improved maize drought tolerance. Additionally, transgenic maize with increased *ZmVPP1* expression also exhibits enhanced tolerance to water deficit.

RESULTS

GWAS for maize drought tolerance at seedling stage

Drought tolerance is a complex quantitative trait, affected by both the timing and the severity of the stress exerted on plant growth and development. In this study, we focused on maize drought tolerance at the seedling stage. Because of the relatively high reproducibility of this phenotype, drought tolerance was indexed by plant survival rate (SR) under severe drought stress. Under this condition, water limitation was the predominant stress, as compared to other possible environmental fluctuations or developmental nuances. In order to investigate a wide range of maize genetic diversity, we phenotyped the drought tolerance of a maize natural-variation population (367 different

[1]Key Laboratory of Plant Molecular Physiology, Institute of Botany, Chinese Academy of Sciences, Beijing, China. [2]Graduate University of the Chinese Academy of Sciences, Beijing, China. [3]Department of Biology, Tokyo Gakugei University, Tokyo, Japan. [4]National Maize Improvement Center of China, Beijing Key Laboratory of Crop Genetic Improvement, China Agricultural University, Beijing, China. [5]National Key Laboratory of Crop Genetic Improvement, Huazhong Agricultural University, Wuhan, China. [6]These authors contributed equally to this work. Correspondence should be addressed to F.Q. (qinfeng@ibcas.ac.cn) or X.Y. (yxiaohong@cau.edu.cn).

Received 5 April; accepted 5 July; published online 15 August 2016; doi:10.1038/ng.3636

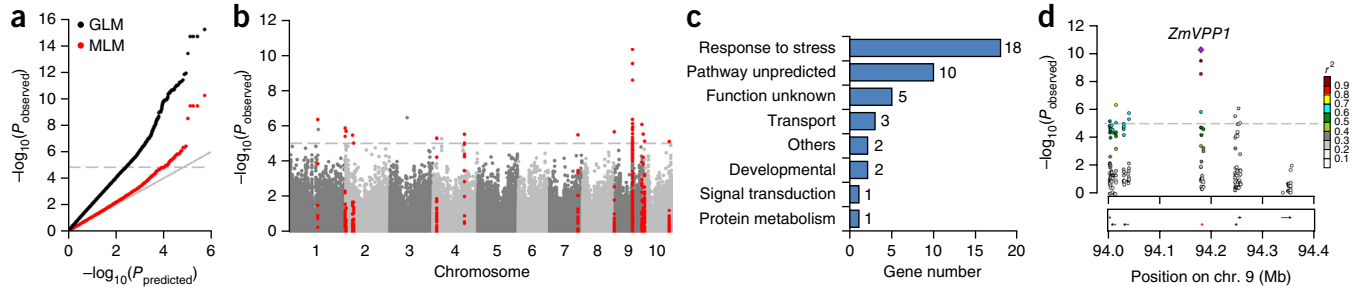

Figure 1 GWAS for drought tolerance in maize seedlings. (**a**) Quantile–quantile plot for the GWAS under a general linear model (GLM) and mixed linear model (MLM). (**b**) Manhattan plot for the GWAS. The dashed horizontal line depicts the Bonferroni-adjusted significance threshold ($P = 1.0 \times 10^{-5}$). The SNPs located within the candidate genes, as identified by the GWAS of the whole population, are labeled as red dots. (**c**) Histogram of biological pathway annotations of 42 candidate genes identified by GWAS. (**d**) Regional Manhattan plot of the *ZmVPP1* genomic region on chromosome 9. The 0.2-Mb genomic region on either side of the most significant SNP is shown. The lead SNP is shown with a purple diamond, and the other SNPs are colored according to their LD (r^2 value) with the lead SNP. The filtered gene models within this region are indicated as arrows at the bottom (MaizeGDB release 5b.60).

inbred lines). This population includes 152 tropical/subtropical (TST) inbred lines, 149 temperate lines, and 66 lines of mixed origin[14]. Large variations in SR were observed in repeated phenotypic assays, with SR ranging from 82.98% to 1.65% (**Supplementary Fig. 1** and **Supplementary Table 1**), and the repeatability of the phenotype was 0.85. Across different genetic resources, the TST subpopulation exhibited the greatest drought tolerance and the broadest phenotypic variation (**Supplementary Fig. 1**). Using 556,944 SNPs with a minor allele frequency (MAF) ≥ 0.05, covering the whole maize genome[13,14], we performed GWAS to identify the genetic loci underlying the drought tolerance. Under the standard mixed linear model (MLM)[23], which accounts for false positives arising from the population structure (Q) and familial relatedness (K) of the natural variation in the population on the basis of a general linear model (GLM), 37 SNPs were identified as significantly associated with drought tolerance (**Fig. 1a**,**b**). In addition, when these variations were regarded as additional covariates, conditional GWAS revealed five additional significant SNPs contributing to the phenotype (**Supplementary Fig. 2a**,**b**). Approximately 55.2% of the phenotypic variation could be explained by the 42 loci identified from the GWAS of the entire population. The most significant

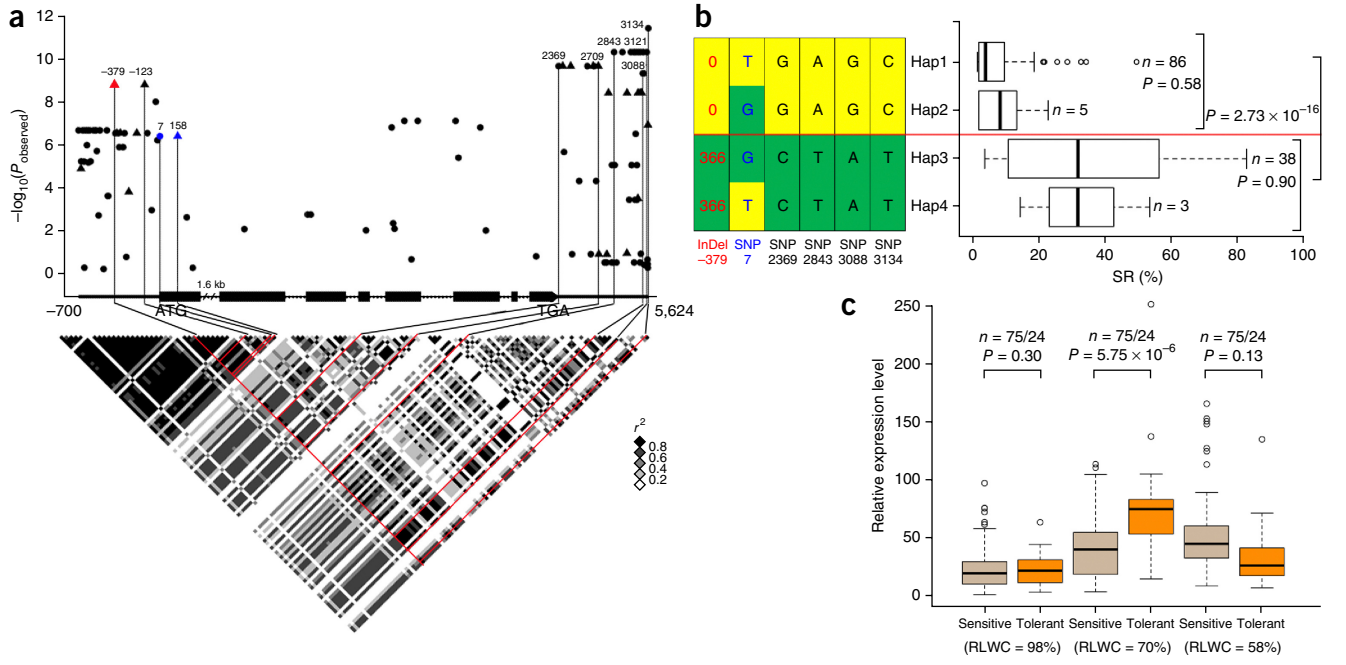

Figure 2 Natural variations in *ZmVPP1* were significantly associated with maize drought tolerance. (**a**) *ZmVPP1*-based association mapping and pairwise LD analysis. Triangles denote indels and dots represent SNPs. Indel −379 and two nonsynonymous variants are highlighted in red and blue, respectively. The significant variants ($P < 2 \times 10^{-9}$) and the two nonsynonymous variants are connected to the pairwise LD diagram with a solid line (the first and last polymorphisms in a string of variants in complete LD in the 3′ UTR are indicated). Red lines highlight the strong LD of indel −379 with the significant variants. (**b**) Haplotypes of *ZmVPP1* among maize natural variations. *n* denotes the number of genotypes belonging to each haplotype group. When a string of variations are in complete LD, only one is shown. Statistical significance was determined by a two-sided *t*-test. The SR distribution of each haplotype group is displayed by the box plot. (**c**) Comparison of *ZmVPP1* expression between the tolerant and sensitive alleles. The gene expression level was determined among 99 maize genotypes under well-watered, moderate drought and severe drought stress conditions, when RLWC was 98%, 70%, and 58%, respectively. "Tolerant" represents the Hap3 and Hap4 alleles, and "Sensitive" represents the Hap1 and Hap2 alleles. *n* is the genotype number of the two alleles. In box plots, center values are medians, dashed lines indicate variability outside the upper and lower quartiles, and dots denote outliers. Statistical significance was determined by analysis of variance (ANOVA).

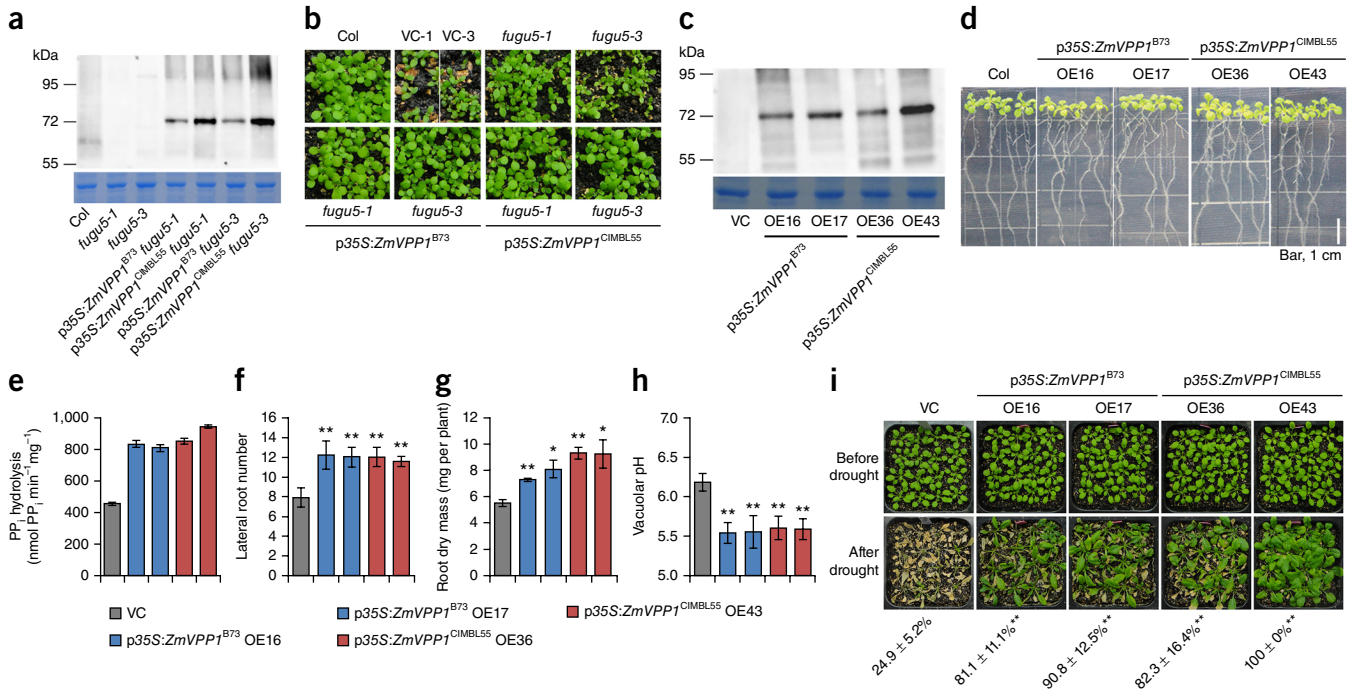

Figure 3 Functional characterization of ZmVPP1 as a H+-PPase and comparison of ZmVPP1^{B73} and ZmVPP1^{CIMBL55} protein function *in Arabidopsis*. (**a**) Top, protein-blot detection of ZmVPP1^{B73} and ZmVPP1^{CIMBL55} protein levels in the *fugu5-1* and *fugu5-3* mutants using a specific anti-ZmVPP1 antibody. Bottom, protein loading control of the total membrane fraction of *Arabidopsis* seedlings. (**b**) Both ZmVPP1^{B73} and ZmVPP1^{CIMBL55} complemented the defective phenotype of *fugu5-1* and *fugu5-3*. VC-1 and VC-3 represent empty-vector transformation of *fugu5-1* and *fugu5-3*, respectively. (**c**–**h**) Comparison of VC and transgenic *Arabidopsis* p*35S:ZmVPP1*^{B73} and p*35S:ZmVPP1*^{CIMBL55}. (**c**) Protein-blot analysis of ZmVPP1 protein levels in transgenic *Arabidopsis*, using a specific ZmVPP1 antiserum. The Coomassie Brilliant Blue staining indicates equivalent sample loading. (**d**) Phenotype of lateral-root development. (**e**) Activity assays of H+-PPase. (**f**) Lateral-root development statistical data for **d** were obtained from at least 20 seedlings in three repeated experiments. (**g**) Root dry mass and statistical data were obtained from 32 plants of each kind. (**h**) Root vacuolar pH and statistical data were obtained from at least 31 measurements from five seedlings of each kind. (**i**) Assessment of plant drought tolerance. Statistical data were obtained from a total of 100 seedlings for each plant line in three repeated experiments. The photographs were taken before and after drought stress. Error bars, s.d.; statistical significance was determined by a two-sided *t*-test: $*P < 0.05$, $**P < 0.01$.

SNP, chr9.S_94178074 (in the 3′ UTR of *ZmVPP1*), accounted for ~10% of the phenotypic variance, while the majority of the remaining loci individually contributed ~5% of the variation of the entire population. Within the TST and temperate subpopulations, the GWAS identified another 21 and 20 significant SNPs, respectively, contributing to drought tolerance (**Supplementary Fig. 2c–f**). Altogether, 83 genetic variants were discovered contributing to the phenotypic variation at $P < 1.0 \times 10^{-5}$. These were resolved to 42 candidate genes, located on all ten chromosomes with the exception of chromosome 5 (**Supplementary Table 2**). Among these 42 genes, 14 and 8 genes were up- or downregulated, respectively, by drought stress (**Supplementary Table 2**). A gene ontology analysis showed that 18 genes are linked to a biological pathway involved in plant stress response (**Fig. 1c** and **Supplementary Table 2**), including several genes related to drought stress, such as the maize homologs of *Arabidopsis ABO1*, *RCI2A*, *CYS6*, *AVP1*, *SRO1*, and *NAC047* (**Supplementary Table 2**). The other genes are predicted to be involved in transport, development, and metabolism, among other processes (**Fig. 1c** and **Supplementary Table 2**).

ZmVPP1 is significantly associated with drought tolerance

The five most significant SNPs ($P < 1.0 \times 10^{-8}$) are located within the 3′ UTR of a single gene (*GRMZM2G170927*) residing on chromosome 9 (**Fig. 1d**). *GRMZM2G170927* encodes a vacuolar-type H+-pyrophosphatase (H+-PPase), sharing high protein sequence homology with *Arabidopsis* AVP1 (ref. 24). Thus, we designated it "*ZmVPP1*". To verify this result, we constructed a BAC library for a drought-tolerant genotype, CIMBL55, and screened the library for the genomic sequence of *ZmVPP1*^{CIMBL55}. After the positive clone was obtained, a 106-kb genomic sequence harboring *ZmVPP1*^{CIMBL55} was sequenced. Two genes (*GRMZM2G105167* and *ZmVPP1*) were identified on the BAC clone, while ~85% of the sequences were intergenic and encoded different types of transposons (**Supplementary Fig. 3**). Genomic structure variations (inversions and deletions) were seen between B73 and CIMBL55 within this region as compared with the current B73 genome assembly (the reference genome, a drought-sensitive genotype). In comparison, the gene-encoding region displays better synteny than the intergenic region between the two genomes. We subsequently resequenced *ZmVPP1* from 140 maize inbred lines, which were representative of the entire association population. Two variants, indel −379 (366 bp) and indel −123 (5 bp), were identified upstream from the translation start codon (ATG) and found to be significantly associated with maize drought tolerance (SR) ($P = 1.68 \times 10^{-9}$; **Fig. 2a**). In addition, two nonsynonymous variations in the coding region (SNP 7 and indel 158 (6 bp)) contributed significantly to the phenotypic variation ($P = 4.16 \times 10^{-7}$). In comparison with the ZmVPP1^{B73} protein, SNP 7 and indel 158 resulted in a substitution of the third serine into alanine and a deletion of two amino acids (glycine and alanine) in ZmVPP1^{CIMBL55}, respectively (**Fig. 2a**). Subsequent analysis of the 3′-UTR sequences confirmed the GWAS result. Moreover, the SNPs in this region were in strong LD with indel −379 and indel −123 in the 5′ UTR among the sequenced genotypes ($r^2 > 0.8$; **Fig. 2a**).

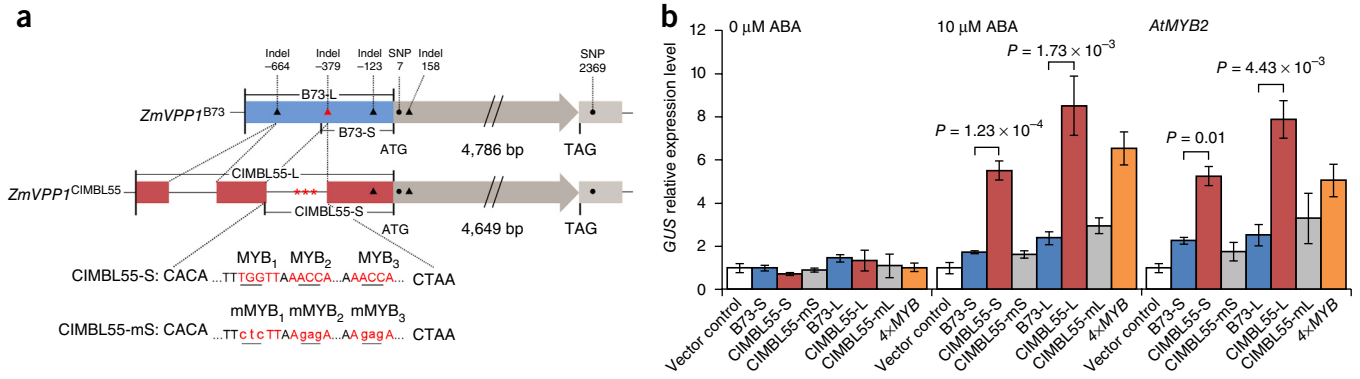

Figure 4 The 366-bp insertion (indel −379) in the *ZmVPP1*^{CIMBL55} allele confers stress-inducible expression of *ZmVPP1*. (**a**) Schematic diagram of the *ZmVPP1*^{B73} and *ZmVPP1*^{CIMBL55} promoters. Triangles and dots represent indels and SNPs, respectively. Three MYB-recognition sites within the 366-bp insertion are indicated by red asterisks. The mutated nucleotides in the MYB-binding sequence are underlined and presented in lowercase in the mutated MYB sequences "mMYB". The regions of short and long promoter fragments of *ZmVPP1*^{B73} (B73-S and B73-L) and *ZmVPP1*^{CIMBL55} (CIMBL55-S and CIMBL55-L) are denoted in the diagram. The *ZmVPP1*^{CIMBL55} promoter fragments with mutated MYB-binding sequences are designated as CIMBL55-mS and CIMBL55-mL. (**b**) *GUS* expression levels driven by different *ZmVPP1* promoter fragments. "*4×*MYB" denotes the artificial sequence of a four-time tandem repeat of the 63-bp sequence containing three MYB-recognition sites in indel −379. *Ubi:Luciferase* was co-transected and served as a reference gene as a control for the transfection efficiency. The data were based on three independent biological replicates. Error bars, s.d.; statistical significance was determined by a two-sided *t*-test.

The 140 maize genotypes were classified into four haplotype groups, based on the significant variants ($P < 2.0 \times 10^{-9}$) and the two nonsynonymous variants (frequency > 0.01; **Fig. 2b**). *ZmVPP1*^{B73} is a representative of haplotype 1 (Hap1, $n = 86$), which is the largest group, whereas *ZmVPP1*^{CIMBL55} belongs to Hap3 ($n = 38$), the second largest group. Hap2 and Hap4 are minor groups, comprising a few maize inbred lines. Statistically, lines with Hap1 and Hap2 have a significantly lower SR than those with Hap3 and Hap4 ($P = 2.73 \times 10^{-16}$), while there are no significant differences between lines with Hap1 and Hap2 or between those with Hap3 and Hap4. Therefore, we designated Hap1/Hap2 and Hap3/Hap4 as the sensitive and tolerant alleles, respectively, of *ZmVPP1* (**Fig. 2b**). As the most significant variants are located in the 5′ and 3′ UTRs, we quantified the abundance of *ZmVPP1* mRNA using quantitative reverse-transcription PCR (qRT–PCR) under well-watered (relative leaf water content (RLWC) = 98%), moderate drought (RLWC = 70%), and severe drought (RLWC = 58%) conditions, among 99 different maize inbred lines. We found that, under moderate drought stress, expression of the *ZmVPP1* sensitive allele was significantly lower than that of the tolerant allele ($P = 5.75 \times 10^{-6}$). There was no significant difference observed under either the well-watered or severe drought stress conditions, although expression of the sensitive allele trended higher under severe stress conditions (**Fig. 2c**). Collectively, these data suggest that drought-inducible expression of *ZmVPP1* contributes to drought tolerance in maize seedlings.

ZmVPP1 functions as a H$^+$-PPase

In order to clarify whether the two nonsynonymous variants in the *ZmVPP1* coding region alter gene function, we compared the protein function of ZmVPP1^{B73} and ZmVPP1^{CIMBL55} *in planta*. In *Arabidopsis*, mutations of the *AVP1* gene in *fugu5-1* and *fugu5-3* mutants result in heterotrophic growth defects in the plants after germination, with oblong cotyledons in the absence of exogenous sucrose supply[24]. We transformed the *ZmVPP1*^{B73} and *ZmVPP1*^{CIMBL55} alleles into each of these mutants. The morphological defects of both *fugu5-1* and *fugu5-3* were fully and comparably complemented, indicating that *ZmVPP1* has conserved biological function similar to that of its *Arabidopsis* ortholog (**Fig. 3a**,**b**). In addition, we also generated p*35S:ZmVPP1*^{B73} and p*35S:ZmVPP1*^{CIMBL55} transgenic plants where the *ZmVPP1* genes are under the control of the cauliflower mosaic virus (CaMV) *35S* promoter in wild-type *Arabidopsis* (Col). Two independent transgenic lines for each construct (p*35S:ZmVPP1*^{B73} (OE16 and OE17) and p*35S:ZmVPP1*^{CIMBL55} (OE36 and OE43)) exhibiting comparable levels of the ZmVPP1 protein variants were further analyzed (**Fig. 3c**). In comparison to vector-transformed (VC) plants, p*35S:ZmVPP1*^{B73} and p*35S:ZmVPP1*^{CIMBL55} transgenic plants had nearly doubled H$^+$-PPase activity (**Fig. 3c**,**e**). Notably, increased numbers of lateral roots and root mass were observed in the transgenic *Arabidopsis*, which may facilitate water uptake in soil (**Fig. 3d**,**f**,**g**). We further compared the pH values of transgenic vacuoles and VC plant vacuoles, as the H$^+$-PPase functions to pump protons from the cytoplasm into the vacuole via pyrophosphate (PP_i) hydrolysis. As expected, the vacuoles of *Arabidopsis* plants transgenic for either p*35S:ZmVPP1*^{B73} or p*35S:ZmVPP1*^{CIMBL55} were more acidic than those of their VC counterparts (**Fig. 3h**). Importantly, significantly improved drought tolerance was observed in both kinds of transgenic plants (**Fig. 3i**). Taken together, these data demonstrate that ZmVPP1 is indeed a functional H$^+$-PPase and that ZmVPP1^{B73} or ZmVPP1^{CIMBL55} overexpression can improve plant root development and drought tolerance.

366-bp insertion confers *ZmVPP1*^{CIMBL55} stress induction

As indel −379 and indel −123 in the *ZmVPP1* promoter are highly associated with drought tolerance, we cloned the promoter fragments from the *ZmVPP1*^{B73} and *ZmVPP1*^{CIMBL55} alleles upstream of a *GUS* (β-glucuronidase) reporter gene to compare their activities in driving gene expression (**Fig. 4a**). Sequence analysis found that the 366-bp insertion (indel −379) within the *ZmVPP1*^{CIMBL55} promoter contains three MYB transcription factor (TF) binding sequences, AACCA, while the other DNA variants (indels and SNPs) lacked clear features. When the short promoter fragments of the two alleles, which differed only at indel −379 and indel −123, were compared, *GUS* expression was at similarly low levels (**Fig. 4b**). However, application of 10 μM ABA (abscisic acid) greatly enhanced the *GUS* expression driven by the *ZmVPP1*^{CIMBL55} promoter fragment. To find out whether the MYB-recognition sequence within this insertion is essential for the ABA-inducible *GUS* expression, targeted mutations were introduced

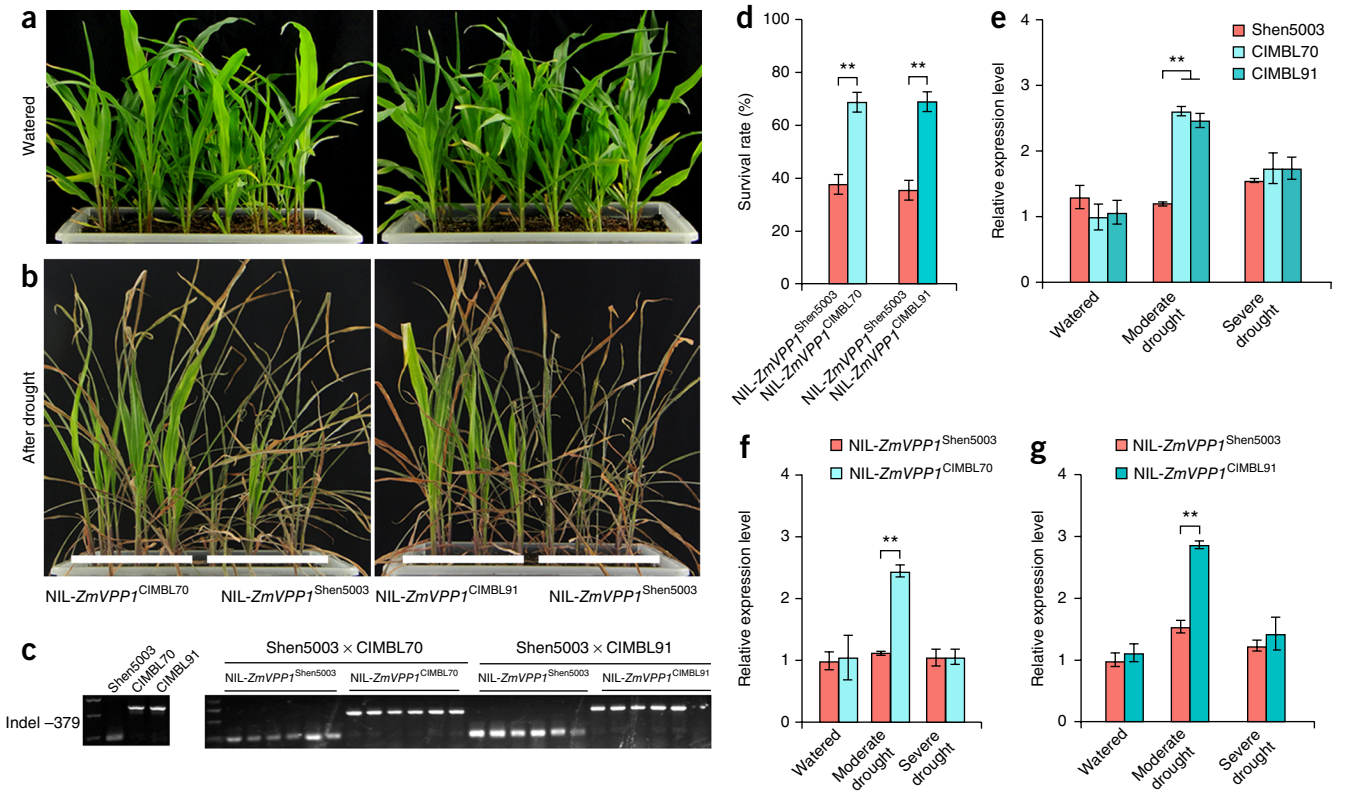

Figure 5 The *ZmVPP1* tolerant allele improves drought tolerance in maize seedlings. (**a**) Morphological phenotype of NILs carrying tolerant allele *ZmVPP1*^{CIMBL70} or *ZmVPP1*^{CIMBL91} or sensitive allele *ZmVPP1*^{Shen5003}. (**b**) Comparison of drought tolerance of the NILs. (**c**) The molecular marker (indel –379) selection of segregating NILs carrying a homozygous allele of either *ZmVPP1*$^{CIMBL70/CIMBL91}$ or *ZmVPP1*^{Shen5003}. (**d**) Statistical data for SR of the NILs when exposed to drought stress. Data were obtained from 45 seedlings of each kind of plant in three independent experiments. (**e**) Relative expression level of *ZmVPP1* in the parental maize inbred lines of CIMBL70, CIMBL91, and Shen5003 under well-watered and moderate and severe drought stress conditions. (**f**,**g**) Relative expression levels of *ZmVPP1* in NIL-*ZmVPP1*^{Shen5003} and NIL-*ZmVPP1*^{CIMBL70} (**f**) and in NIL-*ZmVPP1*^{CIMBL91} (**g**), under well-watered and moderate and severe drought stress conditions. RNA was obtained from three seedlings in two independent experiments. Error bars, s.d.; statistical significance was determined by a two-sided *t*-test: $*P < 0.05$, $**P < 0.01$.

into these MYB-binding sequences (**Fig. 4a**). The result was that the *GUS* expression was clearly reduced to a similar level, as that driven by the *ZmVPP1*^{B73} promoter fragment (**Fig. 4b**). Besides ABA application, when a MYB TF gene (*AtMYB2*) was coexpressed, the *GUS* expression driven by the *ZmVPP1*^{CIMBL55} promoter was also significantly increased (**Fig. 4b**). Furthermore, when three MYB-binding sequences (63 bp) were tandemly inserted four times upstream of *GUS*, ABA-responsive and AtMYB2-effective *GUS* expression was induced to similar levels (**Fig. 4b**). In addition, comparable expression was observed with longer promoter fragments of *ZmVPP1*^{B73} and *ZmVPP1*^{CIMBL55}, which included an additional variant (indel −664, $P = 1.35 \times 10^{-5}$), despite increased basal expression (**Fig. 4b**). Moreover, in order to understand whether the genetic variation in the 3′ UTR also contributes to stress-inducible *ZmVPP1* expression, constructs with ~1-kb 3′-UTR fragments from *ZmVPP1*^{B73} and *ZmVPP1*^{CIMBL55} inserted downstream of green fluorescent protein (GFP) were generated. No significant induction was observed for mRNA levels of GFP after ABA or osmotic-stress treatment (**Supplementary Fig. 4**). Taken together, these data clearly indicate that the MYB-binding sequence residing in the 366-bp insertion is the causal region for the stress-inducible expression of the tolerant allele *ZmVPP1*^{CIMBL55} (**Fig. 4b**).

ZmVPP1 tolerant allele improves drought tolerance

To verify that the tolerant allele of *ZmVPP1* truly contributes to drought tolerance in maize, we introgressed the tolerant allele of *ZmVPP1* from two additional drought-tolerant inbred lines (CIMBL70 and CIMBL91) into a drought-sensitive inbred line (Shen5003) by generating near-isogenic lines (NILs) through four generations of successive backcrossing of the F_1 plants (Shen5003 × CIMBL70 and Shen5003 × CIMBL91) with Shen5003. For each generation, the *ZmVPP1*-heterozygous plants were selected and backcrossed with Shen5003. After one generation of self-pollination, we obtained segregating BC_4F_2 plants carrying either the homozygous-tolerant *ZmVPP1*$^{CIMBL70/CIMBL91}$ or sensitive *ZmVPP1*^{Shen5003} allele (**Fig. 5a**,**c**). Evaluations of drought tolerance confirmed that the NIL-*ZmVPP1*^{CIMBL70} and NIL-*ZmVPP1*^{CIMBL91} plants were more tolerant than the NIL-*ZmVPP1*^{Shen5003} siblings. Approximately 70% of the NIL-*ZmVPP1*^{CIMBL70} and NIL-*ZmVPP1*^{CIMBL91} plants survived the drought stress, as compared to ~40% of the NIL-*ZmVPP1*^{Shen5003} plants (**Fig. 5b**,**d**). Further gene expression analysis showed that, under moderate drought stress, higher *ZmVPP1* expression was detected in NIL-*ZmVPP1*^{CIMBL70} and NIL-*ZmVPP1*^{CIMBL91} plants, as compared with NIL-*ZmVPP1*^{Shen5003} plants (**Fig. 5e–g**), indicating that selection of the tolerant allele of *ZmVPP1* is effective in improving drought tolerance.

Overexpression of *ZmVPP1* improves drought tolerance

In order to understand how increased *ZmVPP1* expression enhances maize drought tolerance, we generated transgenic maize expressing the cDNA of *ZmVPP1*^{B73} under the constitutive *ZmUbiquitin1* (*Ubi*)

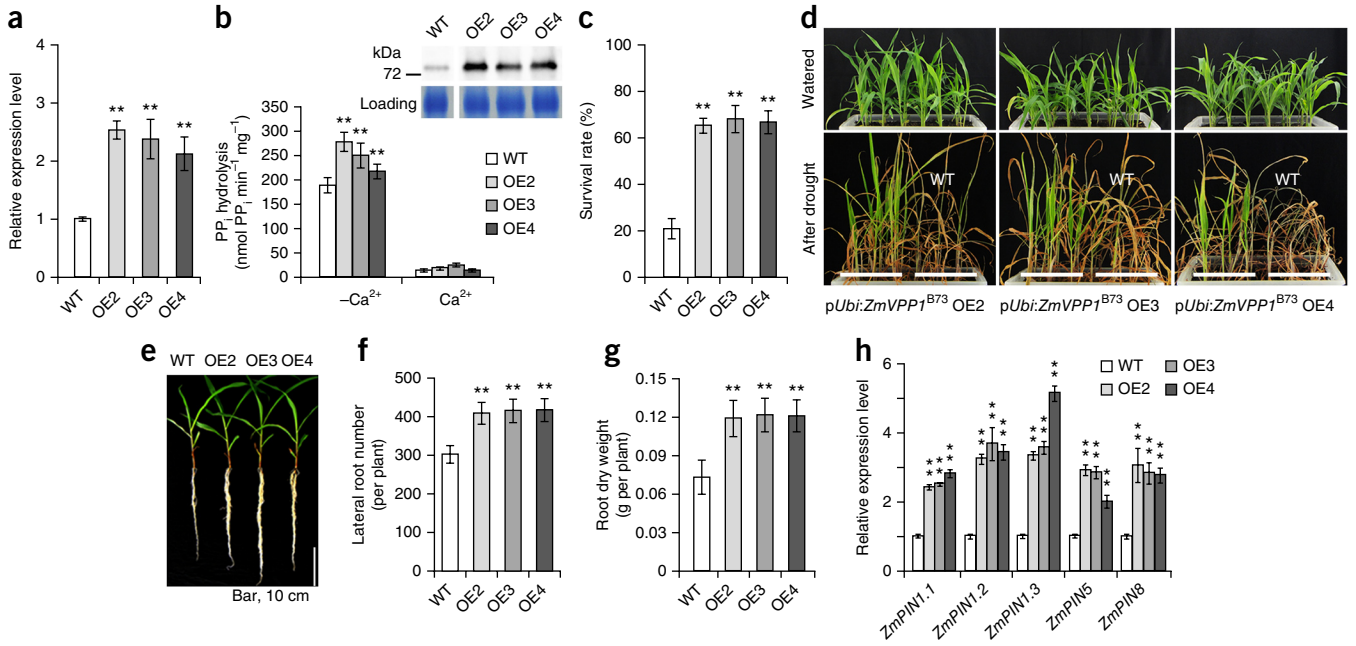

Figure 6 Drought tolerance of *ZmVPP1* transgenic maize. (**a**) Relative expression levels of *ZmVPP1* in WT and three independent *Ubi*:*ZmVPP1*^{B73} transgenic lines (OE2, OE3, OE4) grown under normal conditions. (**b**) Top, protein-blot analysis of ZmVPP1^{B73} protein in WT and transgenic maize. Bottom, H^+-PPase activity in WT and transgenic maize that was specifically inhibited by Ca^{2+} (an inhibitor of H^+-PPase). Data were obtained from 15 seedlings for each WT and transgenic plant line in **a** and **b**. (**c**) Statistical data for the SR of WT and transgenic maize exposed to drought stress. Data were obtained from 75 seedlings for each kind of plant in five independent tests. (**d**) Drought tolerance of transgenic maize as compared to WT. Photographs were taken before and after drought treatment. (**e**–**g**) Representative pictures (**e**) and statistical comparisons of lateral root numbers (**f**) and root dry mass (**g**) for WT and transgenic maize under normal conditions. Data were obtained from 15 individuals for each plant line. (**h**) Comparison of *ZmPIN* gene expression in WT and transgenic maize root under normal conditions. Error bars, s.d.; statistical significance was determined by a two-sided *t*-test: $*P < 0.05$, $**P < 0.01$.

promoter. We focused on this experimental approach because variations in the protein-coding sequence did not result in a functional variation of *ZmVPP1* (**Fig. 3**). Three independent p*Ubi*:*ZmVPP1*^{B73} transgenic lines were analyzed in the T_2 generation. Protein blot analyses and H^+-PPase enzyme-activity assays indicated that the protein was properly produced and functional in transgenic maize (**Fig. 6a**,**b**). We subsequently compared drought tolerance with side-by-side planting of transgenic and transgene-negative siblings (WT) in soil-containing pots. Significantly enhanced drought tolerance was repeatedly observed. The transgenic maize with enhanced *ZmVPP1*^{B73} gene expression exhibited two- to three-fold increased SR relative to WT under water-deficit conditions (**Fig. 6c**,**d**). Consistent with observations from the transgenic *Arabidopsis* lines, the number of lateral roots and root dry mass of the transgenic maize were remarkably increased in comparison with WT (**Fig. 6e**–**g**). The expression levels of five auxin transporter genes in maize (*ZmPIN1.1*, *ZmPIN1.2*, *ZmPIN1.3*, *ZmPIN5*, *ZmPIN8*) were examined. Expression of each of these genes was elevated in the transgenic plants relative to WT plants, which may contribute to the enhanced development of their root systems (**Fig. 6h**). These data suggest that ectopic expression of *ZmVPP1* in plants may somehow affect auxin transport and auxin-related plant development in plants[25]. Furthermore, we compared water use efficiency (WUE) of the transgenic and WT maize seedlings. Plant leaf photosynthesis rates (PS), stomatal conductance (SC), and transpiration rates (TR) were recorded every other day during a progressive water-stress regime that resulted in a reduction of soil water content (SWC) from 40% to ~2%. Significantly higher PS, SC, and TR were consistently observed in the transgenic plants, until SWC reduced to ~10% (**Fig. 7a**). When the SWC was within the range of ~15% to ~25%, WUE (calculated as PS in relation to TR) was elevated in the transgenic maize seedlings in comparison to WT (**Fig. 7a**). Additionally, in order to determine whether the improved drought tolerance would confer yield advantages to the transgenic plants under drought stress, we performed evaluations under field conditions. p*Ubi*:*ZmVPP1*^{B73} OE3 and OE4 lines were planted and compared with WT in three replicated plots that were designed for irrigated and controlled-drought experiments (drought-1 and drought-2). In comparison to the irrigated plot, the total irrigation in drought-1 and drought-2 was about 60% and 40%, respectively. Under irrigated conditions, the transgenic plants were almost identical to WT regarding 14 important agronomic traits, with an exception being that the leaves of OE4, but not OE3, were slightly wider than WT leaves (**Fig. 7b** and **Supplementary Fig. 5a**). Importantly, under water-deficit stress, both transgenic lines exhibited greater grain yield than WT, especially in the drought-2 plot. It is plausible that the yield-gain advantage was attributed to a reduced anthesis-silking interval (ASI) in transgenic plants as compared to WT; this is essential for pollination and yield under drought stress (**Fig. 7b** and **Supplementary Fig. 5b**). We also monitored photosynthesis efficiency and compared the WUE of the transgenic plants with WT in the drought-2 plot. Similar to controlled experimental conditions, both transgenic lines exhibited enhanced photosynthesis efficiency before and after drought stress in comparison to WT under field conditions. Higher WUE was seen in the transgenic plants with the progression of drought stress (**Supplementary Fig. 5c**). Collectively, these data indicated that increased expression of *ZmVPP1* may promote plant root growth and enhance photosynthetic efficiency as a mechanism to avoid or relieve the stress, but not by restricting plant growth to save water.

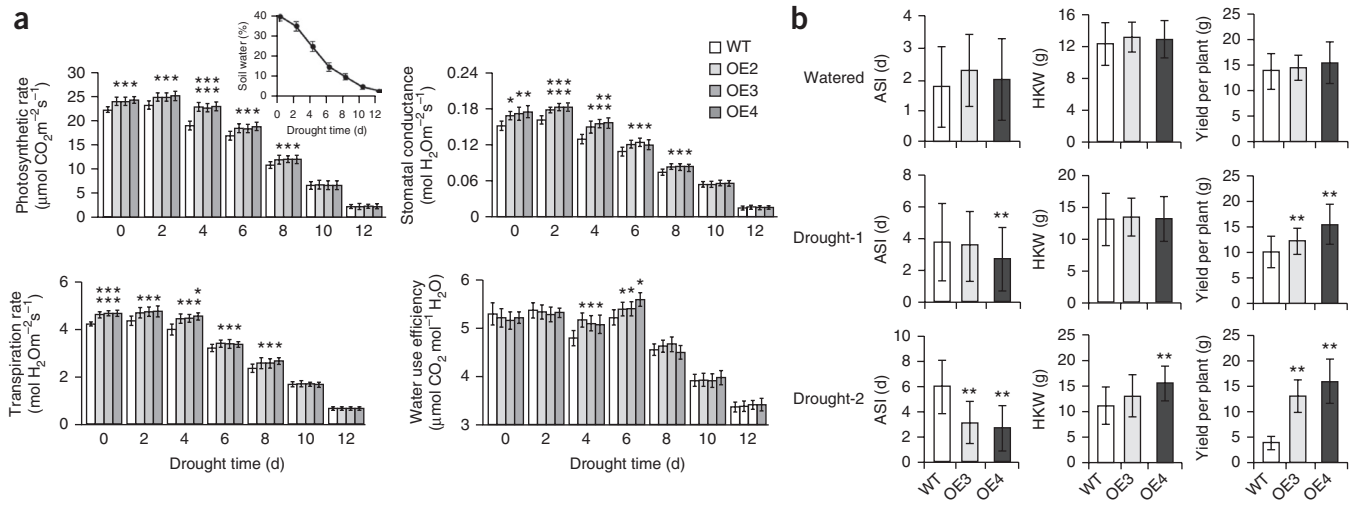

Figure 7 Photosynthetic capacity and yield performance of transgenic maize. (**a**) Photosynthetic capacity of transgenic and WT maize was compared under progressive drought stress under laboratory conditions. Data were obtained from seven seedlings, and similar trends were observed in repeated experiments. (**b**) Statistical data for the anthesis-silking interval (ASI), hundred kernel weight (HKW), and yield per plant of WT and transgenic maize under well-watered and drought stress conditions in the field. Data were obtained from at least 25 plants for each kind. Error bars, s.d.; statistical significance was determined by a two-sided *t*-test: $*P < 0.05$, $**P < 0.01$.

DISCUSSION

Drought tolerance is a complex and inherent trait of plants. In addition to a transgenic approach, the selection and accumulation of tolerant or superior alleles of key genes functioning in stress tolerance may be an effective strategy for genetic improvement of crops. Thus, identifying the genetic components underlying drought tolerance is of great agronomic importance. In this study, we analyzed maize drought tolerance under severe drought stress at the seedling stage (indexed by SR). The greatest phenotypic variation was observed within the TST subpopulation, and maize genotypes belonging to this subpopulation exhibited a relatively strong tolerance to severe drought stress on average (**Supplementary Fig. 1**). Maize has been domesticated over the past ~10,000 years from the wild grass progenitor teosinte (*Zea may* spp.), which originated in southwestern Mexico[26]. Maize has been cultivated and subjected to extensive selection for traits suited to temperate regions, and consequently developed into a major food and staple resource. As a result of the domestication and artificial selection of maize, genetic diversity has been reduced in many cases, and favorable alleles or genes may have been lost from the wild ancestor that were once linked to environmental stress tolerance. For example, the *ZmWAK* locus, which confers resistance to head smut, has been lost from the teosinte ancestry[27]. In this research, we found that, on average, maize TST germplasm is more drought tolerant than maize temperate germplasm (**Supplementary Fig. 1**). Thus, the identification and utilization of valuable genetic resources harbored within TST germplasm, such as the drought-tolerant allele of *ZmVPP1* in CIMBL55, CIMBL70, and CIMBL91, can be highly valuable for the improvement of maize drought tolerance.

In total, GWAS identified 83 genetic variants, resolved to 42 candidate genes that contribute to 55.2% natural variation of drought tolerance in maize seedlings (**Fig. 1** and **Supplementary Table 2**). These data are consistent with the fact that drought resistance is a complicated trait that may be controlled by multiple quantitative trait loci (QTLs) (**Supplementary Table 2**). Further experiments are necessary to validate individual contributions to the phenotype. We compared our association studies with a meta-QTL analysis using 18 biparental populations that identified 68 QTLs for maize drought tolerance[28]. The peak SNP (chr9.S_94178074) overlapped with QTL9.3, which was identified in two populations[28].

Among the 42 candidate genes, 18 genes (42.86%) are predicted to be involved in stress-related biological pathways, and the expression of 22 genes is drought responsive (**Fig. 1c** and **Supplementary Table 2**). Many *Arabidopsis* homologs of the candidate genes have been previously reported to be involved in plant stress response and tolerance. For example, CCS (copper chaperone for superoxide dismutase) delivers copper into copper–zinc superoxide dismutase (CuZnSOD) and activates the enzyme to remove superoxide radicals upon exposure to oxidative stress[29,30]. ABO1 (ABA overly sensitive 1) is a subunit of the histone acetyltransferase complex, and its mutant is hypersensitive to ABA during seed germination, stomatal closure, and drought resistance, suggesting important roles in ABA signaling[31]. *NAC047* belongs to a NAC-type TF gene family, and its homologous gene in maize was designated as *ZmNAC111* (ref. 32). Recently, an 82-bp miniature inverted-repeat transposable element insertion in the *ZmNAC111* promoter was associated with maize drought tolerance[33]. Several members of this gene family have been characterized to function in rice drought tolerance, such as *SNAC1* (ref. 34) and *OsNAC6* (ref. 35). *SRO1* is similar to *RCD1* (ref. 36), which was reported to be responsive to osmotic and oxidative stresses. Its rice homolog, *OsSRO1*, was characterized as a target gene of *SNAC1* and functions in stomatal closure and oxidative stress[37]. Ectopic expression of *RCI2A* (rare cold-inducible 2A), a hydrophobic protein with transmembrane domains, was reported to confer cold tolerance in transgenic plants[38], whereas loss of its function compromised plant salt tolerance[39]. Overexpression of *CYS6* (phytocystatin 6) in transgenic *Arabidopsis* increased plant resistance to multiple stresses including high salt, drought, oxidative, and cold stresses[40].

We report the first findings, to our knowledge, that the natural variation of a H^+-PPase significantly contributes to drought tolerance in maize, although previous studies have suggested that enhancement of this activity confers drought and salt tolerance to transgenic plants[41,42]. The vacuolar H^+-ATPase and H^+-PPase have synergistic roles in maintaining a proton gradient across tonoplasts, and the H^+-PPase was considered as a backup system for the H^+-ATPase

under ATP-limiting conditions, such as abiotic stresses[43]. Recently, *Arabidopsis* vacuolar H^+-PPase was shown to have an important role in acidifying the vacuole in response to ABA, which is required for ABA-induced stomatal closure[44,45]. Also, it has been shown *in vitro* that, while inorganic phosphate (P_i) inhibits glycolysis, PP_i is inhibitory to the reactions of gluconeogenesis, which lead to sucrose biosynthesis in plants[46]. Indeed, *fugu5* mutants accumulated higher levels of PP_i and required exogenous sucrose supply for proper seedling establishment[24]. Therefore, it is plausible that, besides its role in vacuolar acidification, *ZmVPP1* overexpression can simultaneously improve plant metabolic performances by shifting the photoassimilate flux toward sucrose formation[24]. Together, these data are in good accordance with our finding that *ZmVPP1* contributes to maize drought tolerance and leaf photosynthesis efficiency.

Although variations in the coding region and 3′ UTR were also significantly associated with the phenotype, they did not result in clear functional differences (**Fig. 3** and **Supplementary Fig. 4**). It is likely that their strong LD with indel −379, the functional variant, leads to the statistical significance (**Fig. 2a**). It is also possible that the gene model structure is variable between genotypes and can result in compensation or de-compensation of functional effects of variants. The observation that SNP 7 and indel −158 do not alter ZmVPP1 H^+-PPase activity in transgenic *Arabidopsis* is also supported by the crystal structure analysis of a homologous protein in *Vigna radiata* (VrH^+-PPase)[47]. This study confirmed that the two mutations do not affect the key residues for either formation of the substrate-binding pocket or PP_i hydrolysis to facilitate H^+ transport across vacuolar membranes[47]. In addition, a rice vacuolar-type H^+-PPase gene (*Chalk5*) was cloned as a QTL underlying chalky endosperm of the grain[48]. Its homologous gene in maize, *GRMZM2G041065*, is specifically expressed within endosperm and anthers, whereas *ZmVPP1* is expressed in various tissues and has a relatively low expression level in maize endosperm (**Supplementary Fig. 6**). The diversified gene expression pattern may underlie their distinct biological functions.

Taken together, *ZmVPP1* and its favorable allele can serve as a direct target for both genetic engineering and selection for improvement of maize drought tolerance. *ZmVPP1* confers a yield advantage under water deficit, without obvious negative effects on plant growth and development. Specifically, photosynthetic efficiency and root development are enhanced under both stress and non-stress conditions. Other genes or loci underlying drought tolerance identified in this research may also be of great value for trait improvement in maize as well as in other crops.

URLs. MaizeGDB, http://www.maizegdb.org/; Maizego, http://www.maizego.org/Resources.html; R Project for Statistical Computing, http://www.r-project.org; TAIR, http://www.arabidopsis.org/.

METHODS

Methods and any associated references are available in the online version of the paper.

Accession codes. The 106-kb genomic sequence of CIMBL55 is available under GenBank accession code KT923087.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

ACKNOWLEDGMENTS

The authors would like to thank T. Komori (Japan Tobacco, Inc.) for kindly providing us with the maize transformation plasmids pSBI and pSBII. Y. Ma helped in maize transformation vector construction. We would also like to thank S. Li and Z. Li for their excellent technical support in maize transformation and seed propagation. This research was supported by grants from the National High-Tech Research and Development Program of China (2012AA10A306-4), the National Basic Research Program of China (2012CB114302-4), Chinese Academy of Sciences grant (XDA08010206), and the National Natural Science Foundation of China (31471505) to F.Q.

AUTHOR CONTRIBUTIONS

X.W. resequenced *ZmVPP1*, analyzed gene and protein expression levels, and performed *Arabidopsis* transformation and transgenic analysis and the yield test of transgenic maize in fields. H.W. carried out the GWAS of maize drought tolerance and identified the *ZmVPP1* gene, and analyzed the phenotype of transgenic maize in the lab. S.L. helped with the phenotypic analysis of maize drought tolerance. X.Y., J.Y., and J.L. provided the maize materials and the SNP information, and X.Y. advised on the experiments. A.F. provided the *fugu5* mutant seeds and advised on the experiments. F.Q. designed and advised on the experiments and wrote the manuscript.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

1. Lobell, D.B. *et al.* Greater sensitivity to drought accompanies maize yield increase in the U.S. Midwest. *Science* **344**, 516–519 (2014).
2. Boyer, J.S. *et al.* The U.S. drought of 2012 in perspective: A call to action. *Global Food Security.* **2**, 139–143 (2013).
3. Yu, C. China's water crisis needs more than words. *Nature* **470**, 307 (2011).
4. Pennisi, E. Plant genetics. The blue revolution, drop by drop, gene by gene. *Science* **320**, 171–173 (2008).
5. Maruyama, K. *et al.* Metabolic pathways involved in cold acclimation identified by integrated analysis of metabolites and transcripts regulated by DREB1A and DREB2A. *Plant Physiol.* **150**, 1972–1980 (2009).
6. Qin, F., Shinozaki, K. & Yamaguchi-Shinozaki, K. Achievements and challenges in understanding plant abiotic stress responses and tolerance. *Plant Cell Physiol.* **52**, 1569–1582 (2011).
7. Yamaguchi-Shinozaki, K. & Shinozaki, K. Transcriptional regulatory networks in cellular responses and tolerance to dehydration and cold stresses. *Annu. Rev. Plant Biol.* **57**, 781–803 (2006).
8. Mackay, T.F.C. Quantitative trait loci in *Drosophila*. *Nat. Rev. Genet.* **2**, 11–20 (2001).
9. Yu, J. & Buckler, E.S. Genetic association mapping and genome organization of maize. *Curr. Opin. Biotechnol.* **17**, 155–160 (2006).
10. Huang, X. *et al.* Genome-wide association studies of 14 agronomic traits in rice landraces. *Nat. Genet.* **42**, 961–967 (2010).
11. Huang, X. *et al.* Genome-wide association study of flowering time and grain yield traits in a worldwide collection of rice germplasm. *Nat. Genet.* **44**, 32–39 (2011).
12. Gore, M.A. *et al.* A first-generation haplotype map of maize. *Science* **326**, 1115–1117 (2009).
13. Yang, X. *et al.* Characterization of a global germplasm collection and its potential utilization for analysis of complex quantitative traits in maize. *Mol. Breed.* **28**, 511–526 (2011).
14. Li, H. *et al.* Genome-wide association study dissects the genetic architecture of oil biosynthesis in maize kernels. *Nat. Genet.* **45**, 43–50 (2013).
15. Yang, Q. *et al.* CACTA-like transposable element in *ZmCCT* attenuated photoperiod sensitivity and accelerated the postdomestication spread of maize. *Proc. Natl. Acad. Sci. USA* **110**, 16969–16974 (2013).
16. Hung, H.Y. *et al. ZmCCT* and the genetic basis of day-length adaptation underlying the postdomestication spread of maize. *Proc. Natl. Acad. Sci. USA* **109**, E1913–E1921 (2012).
17. Wisser, R.J. *et al.* Multivariate analysis of maize disease resistances suggests a pleiotropic genetic basis and implicates a *GST* gene. *Proc. Natl. Acad. Sci. USA* **108**, 7339–7344 (2011).
18. Xue, Y. *et al.* Genome-wide association analysis for nine agronomic traits in maize under well-watered and water-stressed conditions. *Theor. Appl. Genet.* **126**, 2587–2596 (2013).
19. Thirunavukkarasu, N. *et al.* Functional mechanisms of drought tolerance in subtropical maize (*Zea mays* L.) identified using genome-wide association mapping. *BMC Genomics* **15**, 1182 (2014).
20. Setter, T.L. *et al.* Genetic association mapping identifies single nucleotide polymorphisms in genes that affect abscisic acid levels in maize floral tissues during drought. *J. Exp. Bot.* **62**, 701–716 (2011).
21. Lu, Y. *et al.* Joint linkage-linkage disequilibrium mapping is a powerful approach to detecting quantitative trait loci underlying drought tolerance in maize. *Proc. Natl. Acad. Sci. USA* **107**, 19585–19590 (2010).
22. Liu, S. *et al.* Genome-wide analysis of *ZmDREB* genes and their association with natural variation in drought tolerance at seedling stage of *Zea mays* L. *PLoS Genet.* **9**, e1003790 (2013).

23. Yu, J. *et al.* A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nat. Genet.* **38**, 203–208 (2006).
24. Ferjani, A. *et al.* Keep an eye on PPi: the vacuolar-type H^+-pyrophosphatase regulates postgerminative development in *Arabidopsis*. *Plant Cell* **23**, 2895–2908 (2011).
25. Li, J. *et al. Arabidopsis* H^+-PPase AVP1 regulates auxin-mediated organ development. *Science* **310**, 121–125 (2005).
26. Doebley, J.F., Gaut, B.S. & Smith, B.D. The molecular genetics of crop domestication. *Cell* **127**, 1309–1321 (2006).
27. Zuo, W. *et al.* A maize wall-associated kinase confers quantitative resistance to head smut. *Nat. Genet.* **47**, 151–157 (2015).
28. Semagn, K. *et al.* Meta-analyses of QTL for grain yield and anthesis silking interval in 18 maize populations evaluated under water-stressed and well-watered environments. *BMC Genomics* **14**, 313 (2013).
29. Huang, C.H., Kuo, W.Y. & Jinn, T.L. Models for the mechanism for activating copper-zinc superoxide dismutase in the absence of the CCS Cu chaperone in *Arabidopsis*. *Plant Signal. Behav.* **7**, 428–430 (2012).
30. Negi, N.P., Shrivastava, D.C., Sharma, V. & Sarin, N.B. Overexpression of CuZnSOD from *Arachis hypogaea* alleviates salinity and drought stress in tobacco. *Plant Cell Rep.* **34**, 1109–1126 (2015).
31. Chen, Z. *et al.* Mutations in ABO1/ELO2, a subunit of holo-Elongator, increase abscisic acid sensitivity and drought tolerance in *Arabidopsis thaliana*. *Mol. Cell. Biol.* **26**, 6902–6912 (2006).
32. Voitsik, A.M., Muench, S., Deising, H.B. & Voll, L.M. Two recently duplicated maize NAC transcription factor paralogs are induced in response to Colletotrichum graminicola infection. *BMC Plant Biol.* **13**, 85–100 (2013).
33. Mao, H. *et al.* A transposable element in a *NAC* gene is associated with drought tolerance in maize seedlings. *Nat. Commun.* **6**, 8326 (2015).
34. Hu, H. *et al.* Overexpressing a NAM, ATAF, and CUC (NAC) transcription factor enhances drought resistance and salt tolerance in rice. *Proc. Natl. Acad. Sci. USA* **103**, 12987–12992 (2006).
35. Nakashima, K. *et al.* Functional analysis of a NAC-type transcription factor OsNAC6 involved in abiotic and biotic stress-responsive gene expression in rice. *Plant J.* **51**, 617–630 (2007).
36. Vainonen, J.P. *et al.* RCD1-DREB2A interaction in leaf senescence and stress responses in *Arabidopsis thaliana*. *Biochem. J.* **442**, 573–581 (2012).
37. You, J. *et al.* The SNAC1-targeted gene *OsSRO1c* modulates stomatal closure and oxidative stress tolerance by regulating hydrogen peroxide in rice. *J. Exp. Bot.* **64**, 569–583 (2013).
38. Sivankalyani, V., Geetha, M., Subramanyam, K. & Girija, S. Ectopic expression of *Arabidopsis RCI2A* gene contributes to cold tolerance in tomato. *Transgenic Res.* **24**, 237–251 (2015).
39. Mitsuya, S., Taniguchi, M., Miyake, H. & Takabe, T. Disruption of *RCI2A* leads to over-accumulation of Na^+ and increased salt sensitivity in *Arabidopsis thaliana* plants. *Planta* **222**, 1001–1009 (2005).
40. Zhang, X., Liu, S. & Takano, T. Two cysteine proteinase inhibitors from *Arabidopsis thaliana*, AtCYSa and AtCYSb, increasing the salt, drought, oxidation and cold tolerance. *Plant Mol. Biol.* **68**, 131–143 (2008).
41. Gaxiola, R.A. *et al.* Drought- and salt-tolerant plants result from overexpression of the AVP1 H^+-pump. *Proc. Natl. Acad. Sci. USA* **98**, 11444–11449 (2001).
42. Park, S. *et al.* Up-regulation of a H^+-pyrophosphatase (H^+-PPase) as a strategy to engineer drought-resistant crop plants. *Proc. Natl. Acad. Sci. USA* **102**, 18830–18835 (2005).
43. Maeshima, M. Vacuolar $H(^+)$-pyrophosphatase. *Biochim. Biophys. Acta* **1465**, 37–51 (2000).
44. Bak, G. *et al.* Rapid structural changes and acidification of guard cell vacuoles during stomatal closure require phosphatidylinositol 3,5-bisphosphate. *Plant Cell* **25**, 2202–2216 (2013).
45. Kriegel, A. *et al.* Job sharing in the endomembrane system: vacuolar acidification requires the combined activity of V-ATPase and V-PPase. *Plant Cell* **27**, 3383–3396 (2015).
46. Stitt, M. Product inhibition of potato tuber pyrophosphate:fructose-6-phosphate phosphotransferase by phosphate and pyrophosphate. *Plant Physiol.* **89**, 628–633 (1989).
47. Lin, S.M. *et al.* Crystal structure of a membrane-embedded H^+-translocating pyrophosphatase. *Nature* **484**, 399–403 (2012).
48. Li, Y. *et al. Chalk5* encodes a vacuolar $H(^+)$-translocating pyrophosphatase influencing grain chalkiness in rice. *Nat. Genet.* **46**, 398–404 (2014).

ONLINE METHODS

Phenotyping of maize drought tolerance at seedling stage. Survival rate tests were done in an association-mapping panel composed of 367 diverse inbred lines[13,14], which are publicly available at Maizego. The drought tolerance of the panel was phenotyped in three independent repeats in 2011 (summer) and 2012 (spring and summer) in a cultivation pool (6 × 1.4 × 0.22 m, length × width × depth) in Beijing, China. Each independent repeat contained two replicated assays to estimate random errors. The cultivation pool was filled with uniformly mixed loam (5 tons) and chicken manure (0.25 tons). The saturated RSWC was approximately 40%, as measured by SU-LA (Mengchuangweiyie Technology Co., Ltd). Each pool was divided into 460 plots, and the marginal 82 outlier plots were designed as guard rows. All of the genotypes were randomly planted, and 12 plants of each genotype were grown per plot in each assay. Soil saturation was maintained until all genotypes developed three true leaves. Subsequently, water was withheld and RSWC was measured every other day. When the first symptom of leaf-wilting occurred, RSWC began to be recorded daily. Approximately 6 d after the SWC reached ~0% (at this time point, all seedlings were severely wilted), watering of the whole pool was resumed to recover the surviving plants at the same time. One week after rehydration, the survival rate of each genotype was recorded. Seedlings with green and viable stems were regarded as survivors.

Statistical analysis of the phenotype. To obtain the best linear unbiased prediction (BLUP) of survival rate, the linear mixed effect function lmer in the lme4 package of R Version 3.0.1 (R Project for Statistical Computing) was fitted to each genotype: $Y_{ijk} = u + \text{Environment}_i + \text{Replicate (Environment)}_{ij} + \text{Genotype}_k + (\text{Environment} \times \text{Genotype})_{ik} + e_{ijk}$. Repeatability of the phenotype was based on broad-sense heritability (h^2), which was estimated as $h^2 = \sigma_g^2 \Big/ \left(\sigma_g^2 + \sigma_{gy}^2 / r + \sigma_e^2 / yr\right)$, where σ_g^2 is the genetic variance, σ_{gy}^2 is the interaction of genotype with environment, σ_e^2 is the residual error, r is the number of replications, and y was the number of environments[49]. All the variances were calculated by the same function in the lme4 package of R.

Genome-wide association study. Principal components (PC) of the association panel were calculated by EIGENSTRAT[50] using the high-quality data for the 525,104 SNPs[13,14] with MAF ≥ 0.05. The first two PCs were used to estimate Q, which could explain 11.01% of the phenotypic variation. These data were comparable to those that were calculated by Structure[51]. In order to avoid an over-adjustment, a subset of markers with no statistical association with the phenotype, as determined by a one-way ANOVA analysis ($P \geq 0.995$), was used to construct the non-candidate-marker kinship matrix by SPAGeDi[52]. GWAS was performed using Tassel 3.0 (refs. 23,53,54) under both GLM and MLM. The EMMA algorithm was used for exact variance component estimation under the MLM[55]. For conditional GWAS, the SNPs ($P \leq 1.0 \times 10^{-6}$) identified in the GWAS were treated as additional fixed covariates to identify additional independent SNPs. For association analysis within the TST and NSS (Non-Stiff Stalk) subpopulations, only SNPs were treated as fixed effects. Because of the non-independence of SNPs caused by strong LD, it is usually too strict for significant association detection when the threshold is derived from the total number of markers[56,57]. The suggestive threshold to control the type I error rate was 1.0×10^{-5} (1/85,806), when independent marker numbers[55,56] were determined by PLINK[58] (window size 50, step size 50, $r^2 \geq 0.2$). Candidate genes were identified by the presence of significant SNPs directly located within the gene. Alternatively, genes residing within a corresponding LD region ($r^2 \geq 0.2$) were considered. To estimate the phenotypic variance explained by each significant SNP, ANOVA was used to construct linear models of $Y = \alpha X + \beta P + \varepsilon$. Thus, the variance explained by each SNP was reported after adjusting for population structure effects, as previously described[59]. In this model, Y is the phenotype, X is the SNP genotype, P is the first two PCs, α is the SNP effect, β is the PC effects, and ε is error, which follows a $N(0,\sigma_e^2)$ distribution. Phenotypic variance of the entire population, explained by the total significant lead SNPs identified by GWAS and conditional GWAS, was estimated by a stepwise regression method as previously described[14]. A custom 60k Agilent Maize Oligo Microarray (Agilent Technologies) was applied to maize transcriptome analysis (representing 49,041 transcripts) in response to dehydration treatment. All microarray experiments, including data analysis, were performed by Agilent Technologies under a standard pipeline. Briefly, total RNA from at least five seedlings with or without dehydration treatment was prepared and labeled with the Cy5 and Cy3 fluorescent dyes. Dual-labeling was performed in the array hybridization experiment. For technical repeats, dye-swap assays were performed through reverse labeling of the treated and untreated samples with Cy5 and Cy3. Two independent biological repeats were performed. Genes with a fold change > 1.5 and $P < 0.05$ were considered to be significantly up- or downregulated.

***ZmVPP1*-based association analysis.** According to the sequences of B73 and CIMBL55, eight pairs of primers were designed (Prime 5.0) to amplify *ZmVPP1* from 140 maize genotypes (**Supplementary Table 3**). The sequences were assembled by DNAMAN and aligned by MEGA 5.0. DNA variations, including SNPs and indels, were identified among these genotypes and their association with the phenotype and pairwise LD were calculated by Tassel 3.0 (refs. 23,53,54).

RNA isolation and quantitative RT–PCR analysis. Total RNA was extracted using TRIzol reagent (Biotopped) from at least three seedlings of each genotype. Subsequently, all of the samples were treated with DNase I (Takara) and the concentration of RNA was determined by IMPLANE. cDNA was prepared using M-MLV Reverse Transcriptase and qRT–PCR analyses were conducted using the SYBR Premix Ex Taq kit (Takara) on a Step One System (Applied Biosystems). The quantification method ($2^{-\Delta Ct}$) was used and the variation in expression was estimated using three biological replicates. The maize *Ubi2* (UniProtKB/TrEMBL, Q42415) gene was used as an internal control to normalize the data. PCR conditions consisted of an initial denaturation step at 95 °C for 10 min, followed by 40 cycles at 95 °C for 15 s and 60 °C for 30 s.

Transgenic maize construction and phenotypic analyses. The coding sequence of *ZmVPP1* was amplified from the cDNA library of B73 and inserted into the pSB II vector under the control of the maize *Ubi1* promoter. The LBA4404 *Agrobacterium tumefaciens* strain, carrying the helper plasmid pSB I, was then transformed by the pSB II plasmid containing the *ZmVPP1* insert, which resulted in the generation of the integrated pSB III plasmid. The p*Ubi1*:*ZmVPP1* expression cassette was introduced into the A188 maize inbred line as described[60]. Transgenic T_0, T_1, and T_2 plants were grown in a greenhouse under a 16-h-light/8-h-dark cycle. Transgene-positive and sibling transgene-negative (WT) plants were identified in each generation with *ZmVPP1*-specific PCR analysis. The T_0 transgene-positive plants were backcrossed to A188, and T_1 transgene-positive plants were self-crossed to obtain the T_2 lines for further analyses. Three independent *ZmVPP1* transgenic T_2 lines were germinated in soil for 3 d under dark conditions. After germination, the transgene-positive and WT plants were determined by PCR and transplanted side by side in a cultivation box (35 × 20 × 10 cm, length × width × depth) that was filled with 3.5 kg of enriched soil (turf to vermiculite in a ratio of 1:1). A drought treatment was applied to the soil-grown plants at the three-leaf seedling stage by withholding water. After approximately 20 d, watering was resumed to allow plants to recover and the number of surviving plants was recorded 7 d later. At least 15 plants of each line were compared in each test and statistical analyses were based on data obtained from five independent experiments. Nineteen-day-old soil-grown maize seedlings were removed from soil. Roots were subsequently washed and root volume was measured as previously described[61]. Root material from both transgenic and WT plants were then dried under 65 °C for 72 h to obtain the root dry mass. A LiCor-6400 portable photosynthesis system (LiCor) was used according to the manufacturer's instructions to obtain photosynthesis measurements on the third fully expanded leaves from *ZmVPP1*-transgenic and WT plants. Soil water content was recorded every other day after the initiation of water withholding. Statistical data were based on seven seedlings for each line and the experiment was repeated twice. *ZmPINs* expression levels were determined by qRT–PCR from root samples obtained from hydroponically cultivated 16-d-old maize seedlings (16-d-old) and normalized to the maize *Ubi2* internal control gene.

Introgression of the *ZmVPP1* tolerant allele into Shen5003. F_1 seeds were obtained by crossing Shen5003 with CIMBL70 and CIMBL91, respectively. At each generation, *ZmVPP1* was genotyped and the heterozygous plants were

doi:10.1038/ng.3636

backcrossed with Shen5003 to the BC_4F_1 generation. The BC_4F_1 plants were then self-pollinated to obtain BC_4F_2 plants for further analysis. The morphological and drought phenotypes were compared between NILs homozygous for either the tolerant or sensitive allele of *ZmVPP1*.

***Arabidopsis* transformation and phenotypic analyses.** The coding region sequences of *ZmVPP1* amplified from maize inbred lines B73 (2,316 bp) and CIMBL55 (2,310 bp) were cloned into the BamHI and XhoI (Takara) sites of a pGreen0029-35SΩ vector. The constructed plasmids carrying the desired genes were transformed into *Agrobacterium tumefaciens* (GV3101+pSoup). *Arabidopsis* ecotypes Col-0 and the *fugu5-1* and *fugu5-3* mutants were transformed. Several T_1 transgenic lines were identified with kanamycin selection and grown for production of T_2 and T_3 seeds. Seedlings (T_2 or T_3) with single *ZmVPP1* insertions were subsequently identified and used for further analysis. For assessment of drought tolerance in T_2 seedlings, plants were grown on MS medium with 3% sucrose containing 30 μg/ml kanamycin for 10 d. Plants were subsequently transplanted into soil and maintained in greenhouse conditions at 22 °C in a 16-h-light/8-h-dark cycle. After growth for 30 d under well-watered conditions, water was withheld for 12 d to impose drought stress, followed by re-watering. Surviving plants were counted after a 3-d recovery period. The *fugu5-1* and *fugu5-3* mutants and T_2 plants of p*35S*:*ZmVPP1*^{B73} *fugu5-1*, p*35S*:*ZmVPP1*^{B73} *fugu5-3*, p*35S*:*ZmVPP1*^{CIMBL55} *fugu5-1*, and p*35S*:*ZmVPP1*^{CIMBL55} *fugu5-3* were sown in soil immediately after stratification of seeds under 4 °C. Cotyledon morphology was compared after a 12-d cultivation period. For analysis of lateral-root development, seedlings were cultured on MS medium for 7 d and then transferred to a new MS plate. Images were taken 16 d later and root dry mass was obtained from 44-d-old soil-grown plants. After cleaning, total roots were dried in a 65 °C incubator for 24 h to enable determination of total root mass.

Crude membrane preparation. Crude membranes of maize seedlings, cultured in hydroponic solution for 15 d, were extracted as previously described[62], with minor modifications. Seedlings (~20 g) were washed once in distilled water and ground. Buffer was added to each sample (80 ml) containing 50 mM HEPES/Tris (pH 7.8), 0.25 M sucrose, 10% glycerol, 2 mM EGTA, and 0.5% BSA. 5 mM dithiothreitol (DTT), 2 mM $MgSO_4$, 1% polyvinylpolypyrrolidone (PVPP) and 1 mM phenylmethylsulfonyl fluoride (PMSF) were added to the buffer before use. Homogenates were vortexed and centrifuged at 4,200*g* for 10 min at 4 °C. After filtration, supernatant was centrifuged at 100,000*g* for 1 h. The resulting pellet was resuspended in buffer containing 5 mM HEPES/Tris (pH 7.4), 2 mM $MgSO_4$, and 2 mM DTT. *Arabidopsis* seedlings were grown on MS medium containing 3% sucrose for 3 weeks and ground (~30 plants) after rapid freezing in liquid nitrogen. 15 ml of extraction buffer (50 mM Tris-acetate, pH 7.5, 0.25 M sorbitol, 1% PVPP, 1 mM EGTA, 2 mM DTT, and 1 mM PMSF) was added to samples and samples were vortexed. Homogenates were filtered through four layers of gauze and subsequently centrifuged for 10 min at 10,000*g* at 4 °C. Supernatants were removed to new tubes and centrifuged at 100,000*g* for 1 h at 4 °C. The pellet was then dissolved in storage buffer (20 mM Tris-acetate, pH 7.5, 0.25 M sorbitol, 1 mM EGTA, 1 mM $MgCl_2$, and 2 mM DTT). Crude membranes were stored at −80 °C before use and protein concentrations were determined through the G250 method.

Protein blot and PPase enzyme activity assays. Extracted proteins were detected with a ZmVPP1 antibody raised against an antigenic peptide sequence (DLVGKVERNIPEDDPRN). 20 μg of protein was loaded onto an 8% SDS–PAGE gel and immunodetection was performed with a cooled CCD camera. The original photo for protein-blot analysis of ZmVPP1^{B73} protein in WT and transgenic maize is shown in **Supplementary Figure 7** and PPase activity was measured as previously described[63]. The reaction medium contained 30 mM Tris/MES (pH 8.0), 1 mM $MgSO_4$, 0.5 mM sodium fluoride, 50 mM KCl, 1 mM PP_i, 1.5 μg/ml gramicidin D, and 30 μg/ml of microsome protein. The tubes were incubated at 37 °C for 20 min and PPase activity was calculated through the concentration of P_i.

Vacuolar pH measurement assay. Vacuolar pH was measured using the pH-sensitive dye 2′,7′-difluorofluorescein (Oregon Green 488, Molecular Probes) as previously described[64]. A calibration curve was generated based on the 7-d-old *Arabidopsis* seedlings (Col) that were incubated in the pH equilibration buffer from 5.0 to 7.0 (half-strength Hoagland nutrient solution, containing 50 mM HEPES or MES, 50 mM ammonium acetate, and 20 μM Oregon Green 488) for 15 min before detection. To determine the vacuolar pH of the samples, 7-d-old seedlings were incubated in half-strength Murashige & Skoog medium, containing 50 mM MES (pH 6.0) and 20 μM Oregon Green 488 for 15 min in the dark. The root elongation zone was observed and photographs of fluorescence excitation at 488 nm and 458 nm were recorded. Emission was collected from 525 to 550 nm using a laser confocal microscope (Olympus FluoView FV1000). The fluorescence ratio of 488 nm/458 nm was measured by ImageJ software and the value of vacuolar pH was calculated based on the calibration curve.

Transient gene expression assay. Fragments of the *ZmVPP1* promoter from the B73 and CIMBL55 inbred lines were cloned into a plant transient expression vector (pIG46) by a HindIII site in front of the minimum *35S* CaMV promoter. After identification of transformants, colonies were cultured and plasmid DNA was purified with Midi kits (Qiagen) and subsequently sequenced. Protoplast isolation from etiolated seedlings of a maize inbred line (A188) and transient gene expression analyses were performed as previously described[65]. *Luciferase* was co-transfected and used as a reference gene for the analysis. After transfection, protoplasts were cultured for 6 h before the detection of *GUS* expression. For ABA treatment, 10 μM ABA was applied to each sample after 2 h of cultivation. Total RNA was extracted from each sample 4 h later *GUS* relative expression levels were quantified with qRT–PCR. The data were obtained from three biological replicates.

Evaluation of drought tolerance under field conditions. Drought tolerance of WT and p*Ubi*:*ZmVPP1* OE3 and OE4 plants was compared under field conditions during the winter of 2015 at a breeding field in Sanya, China (18° 15′ N, 109° 30′ E), with an average temperature of 23.25 °C during the growing season. All materials were planted on 22 November 2015 at a density of 58,000 plants/hectare and harvested on 21 February 2016. A rain-off shelter was built to shield plants from receiving irrigation from rainfall. Three replicated plots were designed for one watered and two drought experiments (drought-1 and drought-2). Each plot was separated by a 4-m-wide interval with a 0.5-m-deep waterproof layer. The watered plot was irrigated with an ample water supply during the whole growing season. Conversely, the water supply was stopped 29 d after sowing (DAS) for the drought-1 plot, with the exception of a single irrigation at silking (R1) stage. For the drought-2 plot, watering was stopped 15 DAS, with the exception of two irrigations at the V7 and R1 stages. The total irrigation of the drought-1 and drought-2 plots was approximately 60% and 40% of that for the watered plot, respectively. Photosynthesis parameters were measured in parallel on the sixth to eighth leaves in the drought-2 plot between 9:30 and 11:00 a.m. on the specified dates. After harvest, all ears were naturally dried until the SMC was approximately 11–14%, as measured by PM-8188-A (Kett). Yields per plant were weighed and calculated by normalizing SMC to 14%.

49. Knapp, S.J., Stroup, W.W. & Ross, W.M. Exact confidence intervals for heritability on a progeny mean basis. *Crop Sci.* **25**, 192–194 (1985).
50. Price, A.L. *et al.* Principal components analysis corrects for stratification in genome-wide association studies. *Nat. Genet.* **38**, 904–909 (2006).
51. Falush, D., Stephens, M. & Pritchard, J.K. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics* **164**, 1567–1587 (2003).
52. Loiselle, B.A., Sork, V.L., Nason, J. & Graham, C. Spatial genetic structure of a tropical understory shrub, *Psychotria officinalis* (Rubiaceae). *Am. J. Bot.* **82**, 1420–1425 (1995).
53. Bradbury, P.J. *et al.* TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics* **23**, 2633–2635 (2007).
54. Zhang, Z. *et al.* Mixed linear model approach adapted for genome-wide association studies. *Nat. Genet.* **42**, 355–360 (2010).
55. Kang, H.M. *et al.* Efficient control of population structure in model organism association mapping. *Genetics* **178**, 1709–1723 (2008).
56. Li, M.X., Yeung, J.M., Cherny, S.S. & Sham, P.C. Evaluating the effective numbers of independent tests and significant p-value thresholds in commercial genotyping arrays and public imputation reference datasets. *Hum. Genet.* **131**, 747–756 (2012).

57. Yang, W. *et al.* Combining high-throughput phenotyping and genome-wide association studies to reveal natural genetic variation in rice. *Nat. Commun.* **5**, 5087 (2014).
58. Purcell, S. *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81**, 559–575 (2007).
59. Zhao, K. *et al.* Genome-wide association mapping reveals a rich genetic architecture of complex traits in *Oryza sativa*. *Nat. Commun.* **2**, 467 (2011).
60. Ishida, Y., Hiei, Y. & Komari, T. *Agrobacterium*-mediated transformation of maize. *Nat. Protoc.* **2**, 1614–1621 (2007).
61. Musick, G.J., Fairchild, M.L., Fergason, V.L. & Zuber, M.S. A method of measuring root volume in corn (*Zea mays* L.). *Crop Sci.* **5**, 601–602 (1965).
62. Li, B., Wei, A., Song, C., Li, N. & Zhang, J. Heterologous expression of the *TsVP* gene improves the drought resistance of maize. *Plant Biotechnol. J.* **6**, 146–159 (2008).
63. Hsiao, Y.Y., Van, R.C., Hung, S.H., Lin, H.H. & Pan, R.L. Roles of histidine residues in plant vacuolar $H^{(+)}$-pyrophosphatase. *Biochim. Biophys. Acta* **1608**, 190–199 (2004).
64. Wilson, G.H., Grolig, F. & Kosegarten, H. Differential pH restoration after ammonia-elicited vacuolar alkalisation in rice and maize root hairs as measured by fluorescence ratio. *Planta* **206**, 154–161 (1998).
65. Yoo, S.D., Cho, Y.H. & Sheen, J. *Arabidopsis* mesophyll protoplasts: a versatile cell system for transient gene expression analysis. *Nat. Protoc.* **2**, 1565–1572 (2007).

doi:10.1038/ng.3636