Identification, Association of Natural Variation and Expression Analysis of *ZmNAC9* Gene Response to Low Phosphorus in Maize Seedling Stage

Javed Hussain Sahito [1], Xiao Zhang [1], Haixu Zhong [1], Xuan He [1], Chen Zhen [1], Peng Ma [1], Bowen Luo [1], Dan Liu [1], Ling Wu [1], Muhammad Hayder Bin Khalid [2], Hakim Ali Sahito [3], Zeeshan Ghulam Nabi Gishkori [4], Asif Ali [4] and Shibin Gao [1,*]

[1] Key Laboratory of Crop Gene and Exploration and Utilization in Southwest Region, Maize Research Institute, Sichuan Agriculture University, Chengdu 611130, China; javedhussainsahito@yahoo.com (J.H.S.); hunterzap@163.com (X.Z.); zhonghaixuccc@163.com (H.Z); hexuan1312@163.com (X.H.); cz765826100@126.com (C.Z.); mapeng561@163.com (P.M.); bowenluo_sicau@163.com (B.L.); liu2883508@126.com (D.L.); wtl.570@163.com (L.W.)

[2] Maize Research Institute, Sichuan Agriculture University, Chengdu 611130, China; haider2323@gmail.com

[3] Faculty of Natural Sciences, Shah Abdul Latif University, Khairpur 66020, Sindh, Pakistan; hakim.sahito@salu.edu.pk

[4] Rice Research Institute, Sichuan Agricultural University, Chengdu 611130, China; 2018611002@stu.sicau.edu.cn (Z.G.N.G.); asifali@sicau.edu.cn (A.A.)

* Correspondence: shibingao@163.com; Tel.: +86-186-1578-6279

Abstract: Phosphorus (P) is an essential macroelement supporting maize productivity and low-P stress is a limiting factor of maize growth and yield. Improving maize plant tolerance to low P through molecular breeding is an effective alternative to increase crop productivity. In this study, a total of 111 diverse maize inbred lines were used to identify the favorable alleles and nucleotide diversity of candidate *ZmNAC9*, which plays an important role in response to low P and regulation in root architecture. A significant difference was found under low- and sufficient-P conditions for each of the 22 seedling traits, and a total of 41 polymorphic sites including 32 single nucleotide polymorphisms (SNPs) and 9 insertion and deletions (InDels) were detected in *ZmNAC9* among 111 inbred lines. Among the 41 polymorphic studied sites, a total of 39 polymorphic sites were associated with 20 traits except for the dry weight of shoots and forks, of which six sites were highly significantly associated with a diverse number of low-P tolerant root trait index values by using a mixed linear model (MLM) at $-\log 10\ P = 3.61$. In addition, 29 polymorphic sites under P-sufficient and 32 polymorphic sites under P-deficient conditions were significantly associated with a diverse number of seedling traits, of which five polymorphic sites (position S327, S513, S514, S520, and S827) were strongly significantly associated with multiple seedling traits under low-P and normal-P conditions. Among highly significant sites, most of the sites were associated with root traits under low-P, normal-P, and low-P trait index values. Linkage disequilibrium (LD) was strong at ($r2 > 1.0$) in 111 inbred lines. Furthermore, the effect of five significant sites was verified for haplotypes in 111 lines and the favorable allele S520 showed a positive effect on the dry weight of roots under the low-P condition. Furthermore, the expression pattern confirmed that *ZmNAC9* was highly induced by low P in the roots of the P-tolerant 178 inbred line. Moreover, the subcellular localization of *ZmNAC9* encoded by protein was located in the cytoplasm and nucleus. Haplotypes carrying more favorable alleles exhibited superior effects on phenotypic variation and could be helpful in developing molecular markers in maize molecular breeding programs. Taken together, the finding of this study might lead to further functions of *ZmNAC9* and genes that might be involved in responses to low-P stress in maize.

1. Introduction

Maize (*Zea mays* L.) is one of the most important food and feed crops globally [1,2]. Improving maize yield is the most important goal of maize breeding programs [3]. It is also obvious that plants suffer from various abiotic stresses throughout their life cycles such as drought, heat, cold, salinity, and nutrient deficiency [4–7]. Phosphorus (P) is an essential macronutrient for plant growth and development and cell functions cannot be performed by any alternative nutrient [8,9]. Phosphorus has a certain association with water regulation; i.e., osmosis, photosynthesis regulations, plant metabolic activity, and drought resistance [10,11]. A phosphate (Pi) deficiency not only affects plant growth, development, and crop yield but also deteriorates the quality of the fruit and seed formation [12]. However, improving plant tolerance to P deficiency is an efficient alternative source to increase crop production. Therefore, clarifying P-deficient responses in plants at the molecular level is vital for developing improved genotypes that might perform well under low-P conditions.

However, plants take up P from soil through root hairs to cope with low-P stress and have evolved diverse adaptive mechanisms at physiological, morphological, and molecular levels to enhance P utilization [13,14]. In previous studies, *Arabidopsis thaliana* increased their root hairs by inhibiting the growth of primary roots and several lateral roots in order to adapt under low-P conditions [13,15–17]. Generally, plants increase the number of root hairs when exposed to low-P stress [8,17]. The adaptation mechanism of the acid phosphatase gene *AtPAP26* mutant was controlled due to low P [18]. In earlier studies, many acid phosphatase genes were cloned from maize, barley, rice, and wheat where these genes were induced to be expressed under low-P stress [19]. Recently, many researchers have identified the molecular mechanism of plants under low-P stress conditions which contained the suppressor of yeast gpa1, the yeast phosphatase 81, and the human Xenotropic and Polytropic Retrovirus receptor 1 (SPX) domain encoding proteins to phosphate and plants adapting low-P conditions [20,21]. However, many genes containing SPX domains to maintain P in homeostasis in plants play an important role in P signaling and dynamic balancing based on the structure and these SPX domain proteins are divided into four subfamilies including SPX, SPX-RING, SPX-MFS, and SPX-EXS [22]. The five major family members of phosphorus transporters are PHO1, PHO2, PHT1, PHT2, and PHT3, of which the PHT1 family is mainly responsible for P absorption and phosphorus transport [23]. However, the PHO1 protein family is only one protein family containing both SPX and EXS domains [22]. In the previous study, the AtPHO1 mutant gene was identified in wild and mutant varieties of *Arabidopsis* and phosphorus content increased in mutant roots as compared to wild and decreased in the leaves while shoot growth was affected and inhibited in mutants [24]. Later, the AtPHO1 was cloned in *Arabidopsis* containing both the SPX and EXS domain which was found to be involved in the loading process of inorganic P in the xylem of *Arabidopsis* roots [25]. However, two regulatory genes PHR1 and PHR2 are affected under P-deficient conditions. PHR1, micro399, and PHO2 respond to the low-P stress signal pathway.

Association analysis is a method to identify the genetic variation among the natural populations to establish genetic markers based on linkage disequilibrium (LD) [26]. LD or allelic association refers to a group of non-random associations between different loci alleles in the population [27]. Genetically, R2 and D′ are used to estimate the degree of LD between two alleles [28]. Candidate gene association studies the kinship and population structure mostly used to detect the false-positive association [29]. This model is useful for predicting the potential of different populations for an association analysis and response to marker-assisted selection. The association between non-random alleles of markers reflects the size of chromosomal segments remaining in a complete population [30,31]. Maize is a cross-pollinated crop, which is prone to genetic recombination, gene mutation, etc. LD declined rapidly during the process of evolution, so the use of candidate gene association analysis in maize can

identify effective alleles [32,33]. In recent years, genome wide association study (GWAS) analysis is the most widely used method to identify candidate genes and resolves some complex quantitative trait variations. In previous studies, 36 loci were significantly associated with 32 small-scale diseases of leaf in the maize population [34,35]. Five candidate genes were identified under low-P stress conditions, as were genes involved in the metabolic pathway and that are responsive to low-P tolerant traits at the seedling stage [36]. The single nucleotide polymorphism (SNP) sites, insertion and deletions (InDels), and SSR markers were developed from candidate genes and involved in phenotypic variation [37–39]. In previous studies, some functional molecular markers were associated with specific traits in maize [38]. Two functional molecular markers, HO06 and DGAT04, were associated with high oil content in maize [39].

The NAC (NAM, ATAF and CUC) transcription factors plays an important role in developmental and environmental stress-responsive mechanisms and transcriptional regulatory networks associated with plant stress [40–42]. The NAC acronym originates from three different protein (NAM, ATAF and CUC) domains [41,42]. Recently, many members of the NAC family have been identified by the availability of complete genome sequence in plants against biotic and abiotic stresses—for example, 117 NAC genes were identified in *Arabidopsis*, 151 genes in rice, 79 genes in grape, 26 genes in citrus, 163 genes in poplar, and 152 genes in soybean, tobacco, and also in other species [22,43–45]. In maize, 124 homologous members of the NAC family have been identified based on whole-genome duplication which may lead to an evaluation, molecular functions, and identification analysis in maize [44]. Three novels NAC candidate genes, *ZmNAC145*, *ZmNAC51*, *ZmNAC72*, were induced by drought resistance in maize [45]. Moreover, *ZmNAC84* responded against drought stress in maize [46]. In total, 148 NAC (*ZmNAC1*–*ZmNAC148*) non-redundant genes were identified in a genome-wide association study in maize and all genes belonging to the d group and play a role in plant growth [47]. NAC transcription factors (TFs) contained different domains with at least 150 amino acids in the N-terminal region and these are responsible for DNA binding in the nucleus while the C-terminal region is related to transcriptional regulation and diverse functions [48]. In rice, Os04g0477300 (encoded NAC protein) was identified which is responsible for multiple processes including pathogenic diseases, B toxicity tolerance, boron and P deficiency [49–52].

However, the findings of previous studies provided useful information about NAC transcription factors, such as the fact that NAC genes can regulate diverse functions and may be involved in several processes to improve nutrient efficiency and might be mediated with different signal pathways. These studies have shown the importance of natural variation in genes during the evolution of stress resistance and can be used directly for genetic improvement in maize. In the previous study, a total of 108 candidate genes were identified in a genome-wide association study (GWAS) in maize [36], and *ZmNAC9* was a reliable candidate gene that was not only associated to root dry weight with the low-phosphorus tolerant trait index, but also a differential expression gene in the transcriptome data under different phosphorus condition. Therefore, in this study we re-sequenced *ZmNAC9* in 111 maize inbred lines and aimed to investigate the association between the novel *ZmNAC9* candidate gene with low phosphorus stress-related seedling traits under normal and low-P stress conditions. Besides, molecular markers were identified in maize inbred lines, which can be used as a reference for exploring the molecular mechanism of the *ZmNAC9* gene. Furthermore, the expression pattern was investigated in low-P tolerant 178 and P-sensitive 9782 inbred lines under low-P stress conditions and provides new insights to improve the low-P tolerance in maize and the findings of this study will be helpful to understand the molecular mechanism of candidate *ZmNAC9* under low-P stress conditions in maize.

2. Results

2.1. Basic Molecular Information of ZmNAC's Gene Family

A total of 208 ZmNAC family genes were identified in maize. Homologous genes of *ZmNAC9* were identified using a protein sequence by a protein blast tool search similarity in NCBI, maize GDB

(https://www.maizegdb.org/) and Gramene online databases at 1×10^{10} value. The results showed that the *ZmNAC9* consisted of 780 bp and encoded putative protein 259 amino acids (aas), with an isoelectric point (pI) of 9.70, and six homologous genes were similar to *ZmNAC9* based on high similarity scores. The CDS of *ZmNAC9* homologous genes were between 594 and 1053 bp and their proteins were between 197 and 350 amino acids with one, two or three transcripts. The molecular weight was different among genes analyzed by ExPASy (http://www.expasy.org/) online software and the isoelectric point varied between pI 5.52 and pI 10.64 (Table 1). The subcellular localization of genes was predicted by the cello online database (http://cello.life,netu.edu.tw) and most of the genes were found to be located in the nucleus [51].

Table 1. Molecular characteristics of *ZmNAC19* homologous genes in maize.

Gene Names	Accession Number	CDSbp	Amino Acids	Ch	T	kDa	pI	GRAVY	Subcellular Location
ZmNAC9	Zm00001d011969	780	259	8	2	25,390.81	9.70	−0.327	N
ZmNAC	Zm00001d032208	1053	350	3	1	39,270.58	5.70	−0.680	N
ZmNAC90	Zm00001d042288	741	246	2	3	26,239.40	8.21	−0.510	N
ZmNAC35	Zm00001d006106	783	260	7	1	28,013.65	10.07	−0.582	N
ZmNAC63	Zm00001d021069	750	249	7	2	27,645.22	10.07	−0.752	N
ZmNAC75	Zm00001d049860	687	228	4	1	24,313.91	10.46	−0.288	N
ZmNAC48	Zm00001d031655	594	197	1	3	21,592.24	10.64	−0.178	M/C/P

Note. Ch, T, kDa, pI, GRAVY, N, C, M, C and P stand for chromosome, transcripts, molecular weight, isoelectric point, grand average of hydropathicity, nucleus, cytoplasm and mitochondrial plasma membrane, respectively.

2.2. Phylogenetic Tree from Different Species

To identify the evolutionary relationship of *ZmNAC9* and their homologous genes within maize, rice and *Arabidopsis thaliana*, a tree was constructed by using MEGA 7.0 with a neighbor-joining method by the amino acid sequence of *ZmNAC9* homologous genes from rice and *Arabidopsis thaliana* and the uniform rate was tested using 1000 bootstrap replicates (Figure 1). Results showed that *ZmNAC9* and their homologous genes were in closed relationships with rice members, respectively.
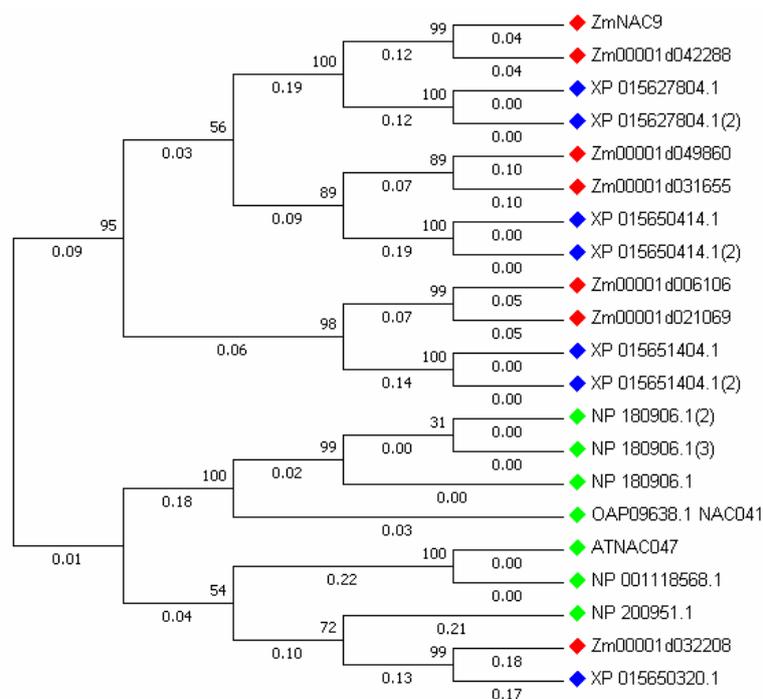
Figure 1. The phylogenetic tree of *ZmNAC9* and their homologous genes. Red mark indicates the homologous *ZmNAC9*; blue—rice genes and green—*Arabidopsis thaliana*.

2.3. Gene Structure, Domain and Chromosomal Location of ZmNAC9 and Their Homologous Genes in Maize

To detect the evolutionary relationship of the ZmNAC gene family, six ZmNAC homologous genes of *ZmNAC9* were selected from 10 subclasses of an evolutionary tree. Among the 10 subclasses, six genes belonged to the same subclade. Results showed that these genes contained 3 exons and 2 introns and conserved the domain which might provide support to the relationship of genes homologous with *ZmNAC9* (Figure 2). Besides, we performed chromosomal locations with a physical relative distance model. Results showed that *ZmNAC9s* and their homologous genes were distributed on different chromosome numbers and two homologous genes were on seven chromosomes in maize (Figure 3).
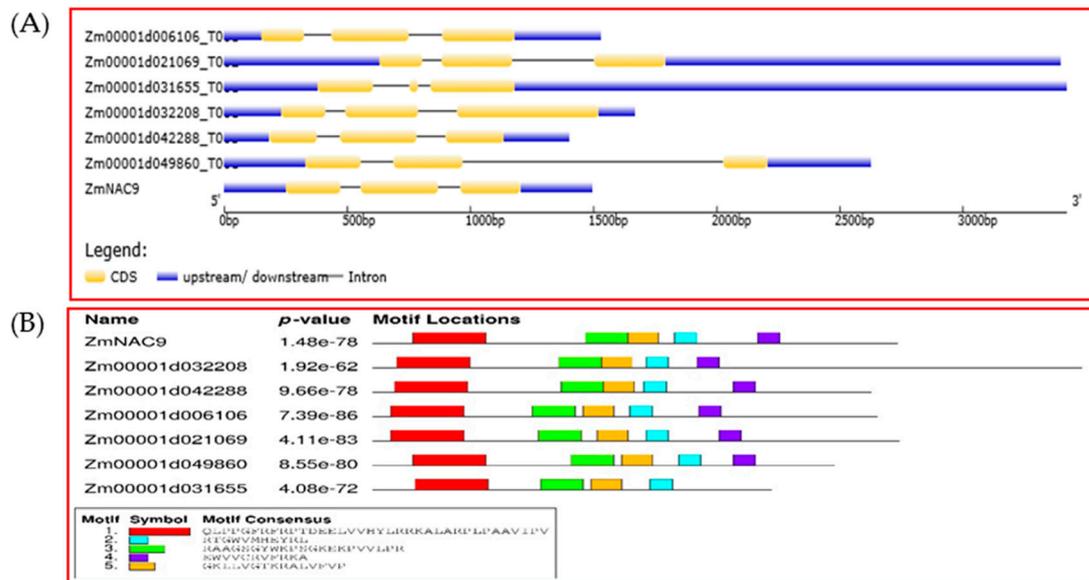
Figure 2. The gene structure and conserved domain of maize 7 genes. (**A**) Untranslated and exons are represented in blue and yellow boxes while introns are presented by blank line. (**B**) Conserved domains are presented in different color boxes.
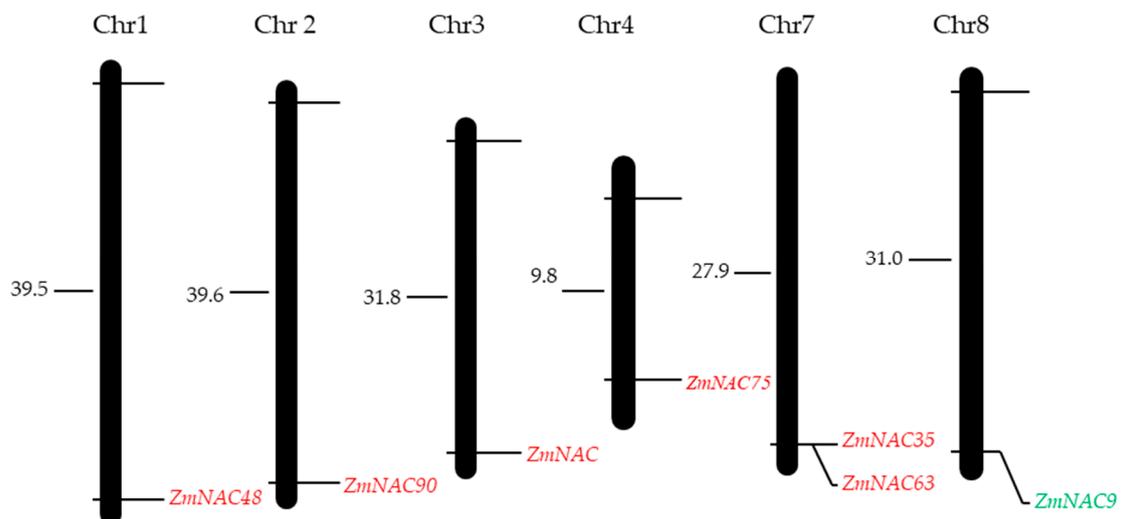
Figure 3. Chromosomal distribution of *ZmNAC9* and their homologous genes. *ZmNAC9* indicated by green color while homologous by red color. Numbers on chromosomes represented by relative physical distance. Chr—chromosomes.

2.4. Sequence Polymorphism and Nucleotide Diversity of ZmNAC9 Gene

The total genomic DNA reference sequence of the candidate ZmNAC9 (1459 bp) gene was downloaded from the maize GDB (https://www.maizegdb.org/) database. The candidate *ZmNAC9* gene was successfully amplified in 111 maize inbred lines. The coding sequence region of the *ZmNAC9* gene was 780 bp and a total of 32 single nucleotide polymorphisms (SNPs) and 9 insertion and deletions (InDels) were identified with a minor allele frequency (MAF) >0.05 in the coding region of *ZmNAC9* among 111 inbred lines—the results are shown in Table S2. The results revealed that the average nucleotide diversity of *ZmNAC9* was Pi (π) 0.01055 for all inbred lines, for polymorphic sites this was 35, for haplotypes 16 and haplotypes diversity was 0.861 in the coding region of *ZmNAC9* among 111 maize inbred lines. Besides, nucleotide diversity was analyzed using a sliding window of 100 bp with a 25 bp step size, and two parameters—π and θ values—were used to measure nucleotides diversity. The results indicated that the π value was rapidly decreased at 0.005 around 200 bp and slightly increased at 0.03 and maintained around 800 bp. While the θ value decreased at 0.005 and slightly increased at 0.01 and minted around 400 bp, it then increased at 0.02 and maintained around 800 bp (Figure 4). Additionally, the Tajima D test was mostly applied for neutral theory in the genetic population to measure the allelic frequency in the nucleotide sequence. In the present study, the Tajima D test exhibited non-significance at $p < 0.05$ whereas Fu and Li's tests showed a significant neutral variation in the coding region of *ZmNAC9* at $p < 0.05$ in the tested maize population (Table 2). Even though these results cannot deny the hypothetical mutation drift equilibrium, a lack of a footprint of positive selection in *ZmNAC9* was suggested. It also implies that there are functional components of the gene-encoding proteins in the process of evolution.
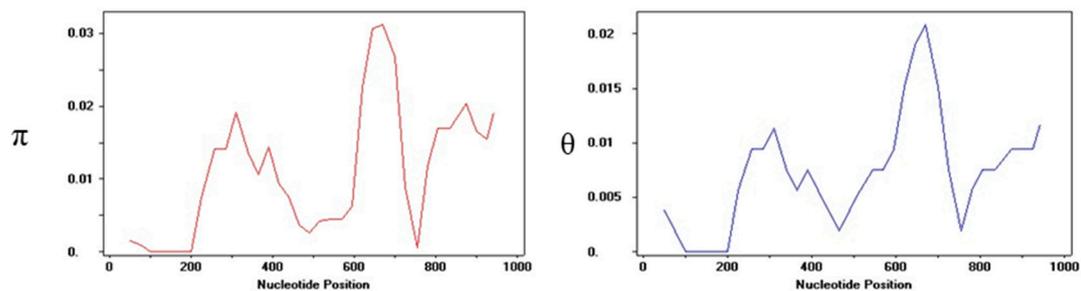
Figure 4. The nucleotide diversity π and θ value in the coding region of *ZmNAC9* using the sliding windows of 100 bp with the step of 25 bp in 111 inbred lines.

Table 2. Summary of single nucleotide polymorphism (SNP) sites, nucleotide diversity (π), nucleotide polymorphism (θ) and Tajima's D, Fu and Li's D test of *ZmNAC9* gene in a different maize population, and ** representing the significance value, $p < 0.05$.

TF	Coding Region	SNP /InDels	Nucleotide Diversity	Nucleotide Polymorphic Sites	Haplotypes	Haplotypes Diversity (Hd)	Tajima's D	Fu and Li's D *
ZmNAC9	780 bp	32/10	Pi(π)0.01055	35	16	0.861	1.47198	1.73019 **

2.5. Linkage Disequilibrium (LD) Analysis in the Coding Sequence of ZmNAC9 Gene

LD is an effective method to investigate the functional allelic variation in association mapping analysis. The LD was analyzed using Tassel software v5.0. In this experiment, the LD was analyzed between all pairs of polymorphic sites (SNPs) and insertion deletion polymorphisms (InDels) in the coding region of the ZmNAC9 gene in 111 inbred lines. The significant LD was found at ($r2 > 0.1$) and also various polymorphic sites at ($p < 0.00001$) in 111 maize inbred lines and different polymorphic sites showed a different level of LD (Figure 5A), respectively. Besides, the LD decay of ZmNAC9 was rapidly decreased in 500 bp physical distance base pairs at ($r2 > 0.1$) (Figure 5B), respectively.
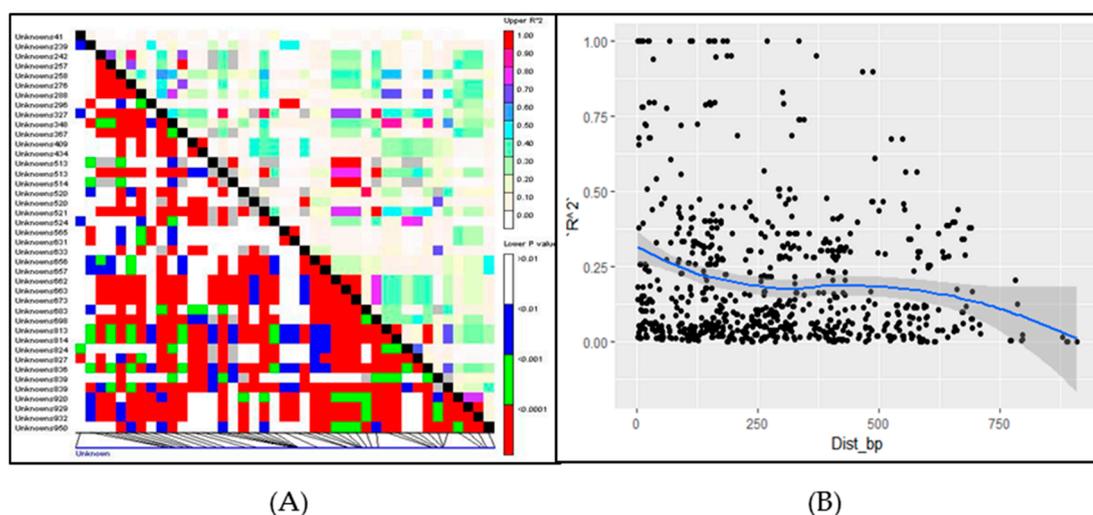
Figure 5. (**A**) Linkage disequilibrium pattern; (**B**) linkage disequilibrium (LD) decay of *ZmNAC9* in 111 maize lines.

2.6. Association Analysis of Low-P Tolerant Traits with ZmNAC9 Gene

The phenotypic difference of 22 seedlings traits was calculated in three independent treatments under low-P, normal-P, and low-P tolerant trait index values. Therefore, 22 significant traits were evaluated for association with the ZmNAC9 gene. A total of 41 polymorphic loci were identified in the coding region of the ZmNAC9 gene including 32 SNPs and 9 InDels. The association between 41 polymorphic sites with 22 low-P tolerance traits was performed using a statistically standard mixed linear model (MLM+Q+K) including a population structure and kinship matrix and a threshold was set at (–log P value 3.61) –log10(0.01/41 polymorphic loci). Among 41 polymorphic sites, 39 sites were associated with 20 traits and a total of 191 associations were significant with the P-tolerant traits index except for the dry weight of shoot and forks. Of which six sites (S327, S513, S514, S520, S827, and S839) were highly significantly associated with a diverse number of P-tolerant trait index values (Figure 6a, Table 3 and Table S3). At the same time, among 41 sites 29 sites were associated with 22 traits, of which six polymorphic sites (position S327, S513, S514, S520, S633, and S839) were highly significantly associated with a variable number of low-P seedling traits under normal-P conditions (Figure 6b, Table 3 and Table S3). However, among 41 polymorphic loci, 32 polymorphic loci included SNPs and InDels were significantly associated with the different number of seedling traits and the remaining nine polymorphic loci were non-significantly associated with each trait under low-P conditions (Figure 6c, Table 3 and Table S3). Of them five sites (position S242, S513, S521, S824, and S827) were strong significantly associated with multiple seedling traits using the same threshold under low-P stress conditions. Among all significant sites, most of the sites were associated with root traits with low-P, normal-P and low-P trait index values. Linkage disequilibrium was strong at (r2 > 1.0) in 111 inbred lines.
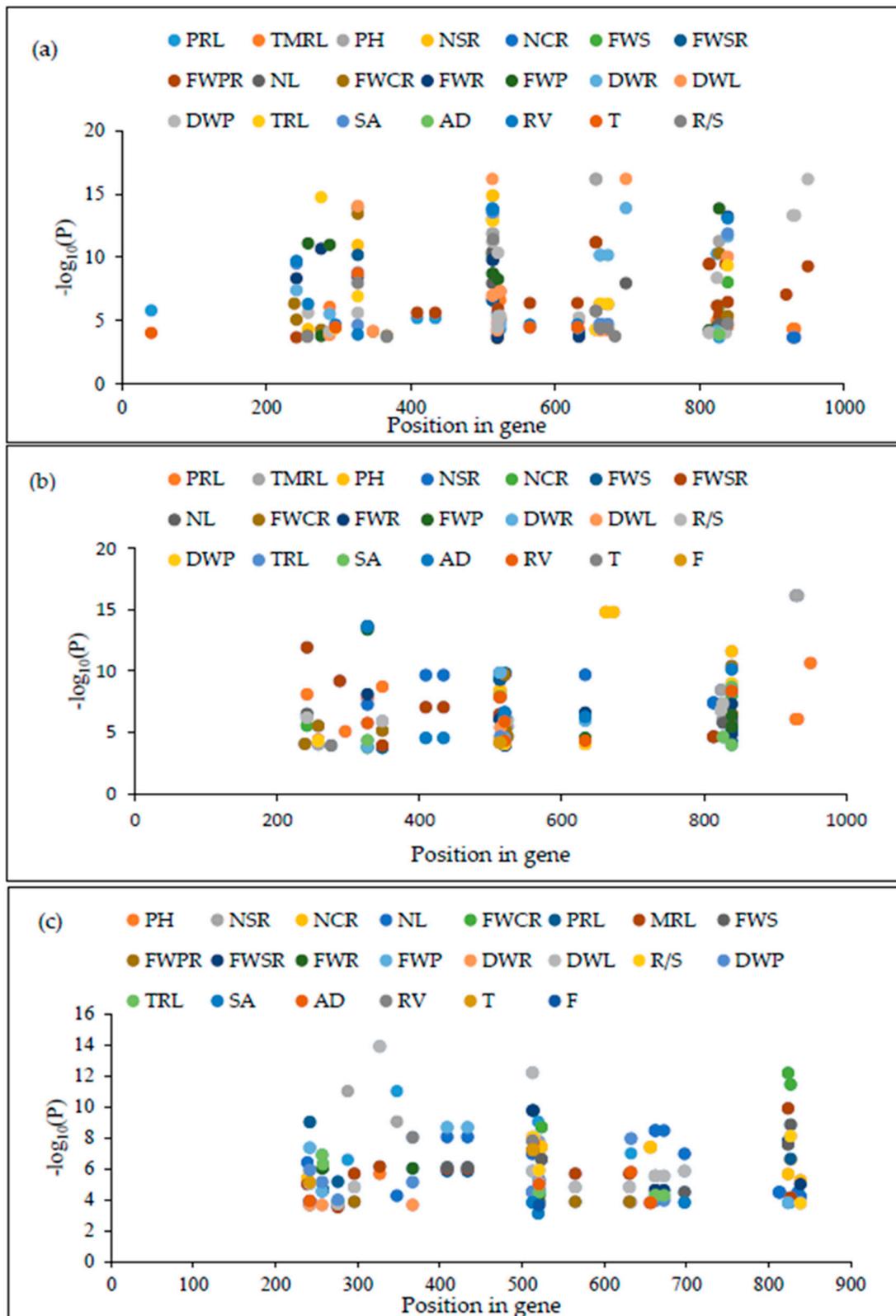
Figure 6. Association analysis of polymorphic sites of *ZmNAC19* and low-phosphorus tolerance seedling traits at threshold (–log P value 3.43). (**a**) Polymorphic sites with low-phosphorus (P) tolerant trait index; (**b**) polymorphic sites with low-P tolerant trait under normal-P conditions; (**c**) polymorphic sites with low-P tolerant trait under low-P conditions.

Table 3. Statistics of association analysis between the polymorphisms in *ZmNAC9* and low-P tolerance traits in maize seedlings.

Site No.	Loci Name	Position in Gene (bp)	SNPs	Minor Allele Frequency	MLM(Q+K) Number of Significant Traits			
					T/CK Index	T	CK	−log10 *P* Value
1	S41	41	C/G	0.05405	3	0	0	ns
2	S239	239	C	0.17117	1	3	1	*
3	S242	242	C	0.10909	6	10	5	***
4	S257	257	C/T	0.10811	2	5	1	**
5	S258	258	G/A	0.3964	4	3	5	**
6	S276	276	T	0.14414	4	5	1	***
7	S288	288	A/T	0.47748	5	1	1	**
8	S296	296	T/A	0.09009	4	3	1	**
9	S327	327	C	0.40541	13	3	10	***
10	S348	348	G/A	0.45946	2	2	5	**
11	S367	367	G/C	0.10811	2	4	0	*
12	S409	409	G/A	0.24324	2	5	3	**
13	S434	434	T/A	0.24324	2	5	3	**
14	S513	513	G/A	0.26415	16	11	10	***
15	S513	513	G	0.26415	0	0	0	ns
16	S514	514	C/A	0.26415	14	5	11	***
17	S520	520	C/T	0.25773	11	5	18	***
18	S520	520	0/3	0.10834	0	0	0	ns
19	S521	521	G/C	0.28866	6	10	5	***
20	S524	524	G/C	0.11712	7	3	3	**
21	S565	565	C/T	0.09009	4	3	0	**
22	S631	631	T/C	0.09009	4	3	0	**
23	S633	633	A/T	0.12613	5	3	8	***
24	S656	656	T/G	0.37838	5	2	0	**
25	S657	657	G/T	0.37838	5	2	0	**
26	S662	662	A/G	0.2973	6	5	1	***
27	S663	663	T/C	0.2973	6	5	1	***
28	S673	673	G	0.2973	6	5	1	***
29	S683	683	C/A	0.23423	1	0	0	ns
30	S698	698	T/G	0.25225	3	4	0	**
31	S813	813	A/G	0.48649	3	1	2	**
32	S814	814	G/C	0.48649	3	1	2	**
33	S824	824	A/G	0.0991	6	6	2	***
34	S827	827	A	0.12037	9	7	5	***
35	S836	836	G/A	0.48649	3	1	2	**
36	S839	839	G/A	0.28261	11	4	20	***
37	S839	839	A	0.28261	1	0	0	ns
38	S920	920	G/A	0.21622	3	0	0	ns
39	S929	929	T/C	0.26126	3	0	2	**
40	S932	932	C/T	0.26126	3	0	2	***
41	S950	950	T/C	0.27027	2	0	1	**

Note: The letters represent nucleotides; letters represent the major alleles and underlined letters minor alleles. T/CK indicate low-P tolerant trait index; T represents P-deficient conditions; CK, normal-P conditions; MLM (Q+K), mixed linear model, population structure and kinship matrix; numerical numbers indicate the significant traits and – indicates data unavailable, ** indicates the −log10(0.01/41), *** indicate the −log10(0.001/41) according to the previously described method for correction test for p-values [52].

2.7. Haplotypes Diversity in the Coding Sequence of ZmNAC9 Gene

Haplotypes were identified in the coding sequence of the *ZmNAC9* gene. A total of 15 haplotypes were identified in tested 111 maize inbred lines in the *ZmNAC9* gene. Among 15 haplotypes, haplotype-1 contained 25 inbred lines of which the most frequent was CDS haplotype 1 and other haplotypes contained more than 10 inbred lines included hap-2, hap-3, hap-4 comprising 23, 13, and 12 inbred lines, respectively (Table 4 and Table S4). Besides, the effects of polymorphic loci on the dry weight of low-P tolerant trait index values were evaluated by haplotype analysis in ZmNAC9. The results showed that, among the identified haplotypes, five major haplotypes contained more than 10 inbred lines which emerged from five significant sites and these haplotypes were used for a pairwise haplotype

comparison analysis. hap4 with a more favorable allele showed the highest significant difference with hap2, hap3, and hap5 while was non-significant with hap1 at $p < 0.05$. A pairwise comparison analysis and one-way ANOVA test were applied to determine significant differences between haplotypes (Figure 7).

Table 4. Division of five haplotypes and corresponding base changes in ZmNAC9 gene.

Haplotypes	Lines	S327	S513	S514	S520	S524	S827
Hap1	25	C	T	G	C	G	0 bp
Hap2	23	C	T	G	T	G	0 bp
Hap3	13	C	A	G	C	C	0 bp
Hap4	12	C	T	G	C	G	0 bp
Hap5	10	C	T	G	C	C	16
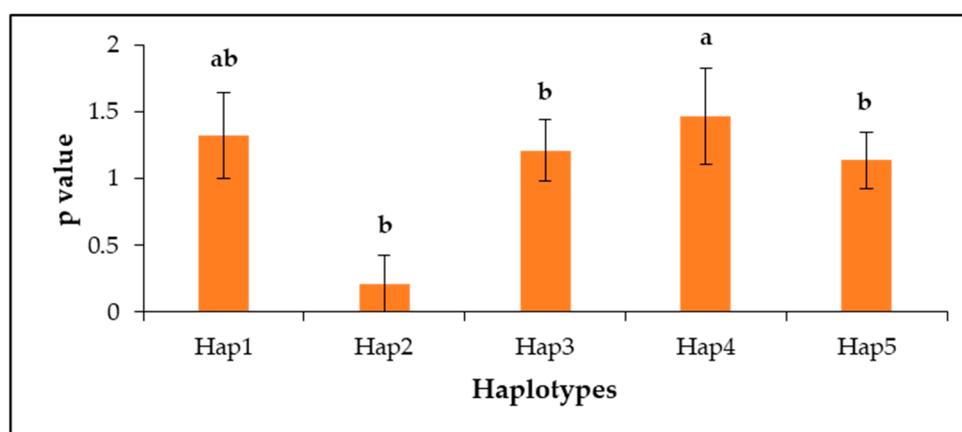
Figure 7. Effect of haplotypes consisting of five significant loci on low-phosphorus tolerance index of dry weight of roots in 111 inbred lines and a, b and ab represented the significance difference.

2.8. Expression Pattern of ZmNAC9 Gene in Low-P Tolerant and P-Sensitive Inbred Lines by qRT-PCR

A gene expression analysis is important in many fields of biological research. Understanding the expression pattern of genes provides useful information about complex regulatory networks occurring in living organisms and gene functions. In this study, the expression pattern of *ZmNAC9* was examined in two maize inbred lines, P-tolerant 178 and P-sensitive 9782, under low-P stress treatments during the seedling growth stage by using quantitative real-time polymerase chain reaction (qRT-PCR). Results showed that the expression pattern was increased in the early stage but decreased in the later stage of P treatment, and the *ZmNAC9* gene was up-regulated on 1, 3, 7 and 9 days in roots and 1, 3, 7, 9, and 12 days in the leaves of the P-tolerant 178 inbred line under low-P treatments, while down-regulated at others time points. In the P-sensitive 9782 inbred line, the *ZmNAC9* gene was induced on 1, 3, 12 and 16 days in roots and 1, 3, 9 days in leaves (Figure 8). The expression pattern was different between 178 and 9782 under low-P stress treatments and results suggested that the *ZmNAC9* gene was more induced in P-tolerant 178 under P-deficient conditions.

2.9. Subcellular Localization of ZmNAC9 Gene

Transcription factors (TFs) are typically localized in the nucleus of the cell and perform a transcriptional activation role and DNA binding. Therefore, in the present study the subcellular localization of *ZmNAC9* was detected in tobacco (Nicotiana benthamiana) leaves by using the transient transformation method and tested method by green fluorescence protein (GFP) and the fusion protein of *ZmNAC9*-eGFP driven by the 35S promoter. The GFP fluorescence signals of *ZmNAC9*-eGFP were localized in the nucleus and cytoplasm by using laser confocal microscopy (LSCM). The empty vector pCAMBIA2300-35S-eGFP without any foreign gene fragment was observed in the nucleus

(Figure 9). These results were not consistent with those predicted by the CELLO v.2.5 program [53], which expected transcription factors to be only located in the nucleus. Generally, model plant tobacco or onion epidermal cells were mostly used for subcellular localization due to successful results. However, the subcellular location of genes is the main issue in cell biology and important study in molecular biology [54].
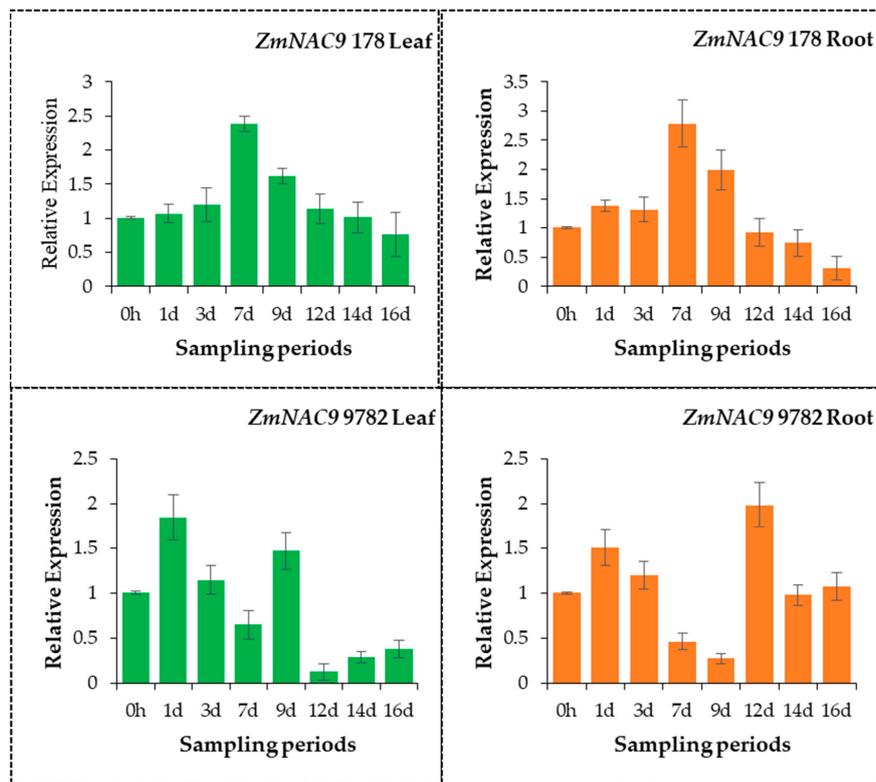
Figure 8. Expression pattern of ZmNAC9 in roots and leaves of P-tolerant 178 and P-sensitive 9782 inbred lines under low-P conditions in maize seedling.
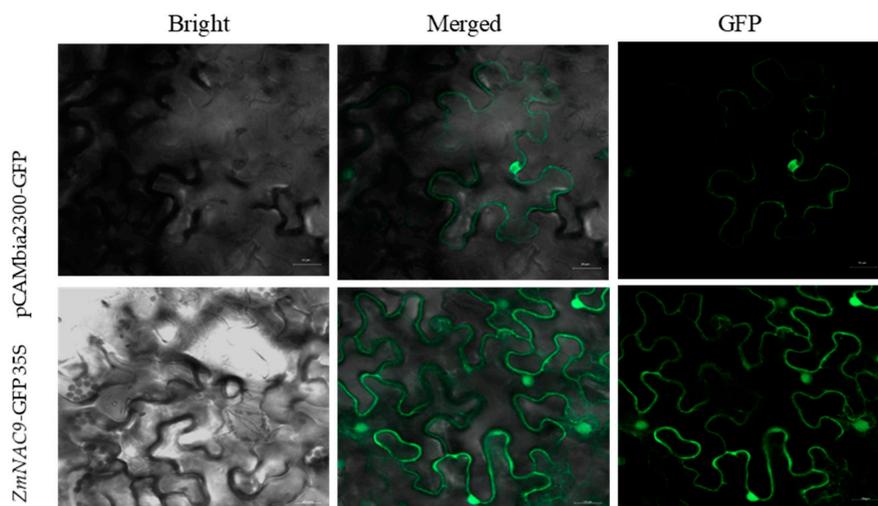
Figure 9. Subcellular localization of *ZmNAC9*–green fluorescence protein (GFP) green fusion protein in the tobacco (Nicotiana benthamiana) epidermal cells of leaf.

3. Discussion

Maize is one of the model plants used for basic and advanced studies of plant domestication and genetic improvement due to its genetic diversity and large phenotypic variation [55,56]. Phosphorus is an essential major element for plant growth and yield and induces biological processes (e.g., when there is a P deficiency) that affect the development, growth, yield, and also the biological functions of plants [10,33,57]. NAC proteins play vital roles in plant development, including signaling pathways, and act as regulators in various abiotic and biotic stresses [44,47]. In this study, we identified the 208 ZmNACs family members in maize and these family members were classified into 10 subclades on a phylogenetic tree. Among 10 subgroups the *ZmNAC9* and their homologous genes were on five subclades according to their position. Besides, the homologous genes were BLASTP with Arabidopsis thaliana and rice, and the phylogenetic tree was constructed in maize, Arabidopsis thaliana, and rice, indicating the evolutionary process of homologous genes in diverse species. Although, the homologous genes of *ZmNAC9* were distributed in different subgroups with members of Arabidopsis and rice, this indicated that the evolutionary process of these genes had not let to functional difference but it might provide a reference to find the functions of maize the same as previously reported for the NAC family in maize [47]. Furthermore, homologous genes were evaluated for the distribution of chromosomal locations, gene structure, and conserved motif. The homologous genes were distributed at different chromosomal locations. The chromosome location indicated that two homologous genes of *ZmNAC9* were distributed on seven chromosomes at a short distance. The gene structure of *ZmNAC9* homologous genes showed that all genes had a similar number of exons and conserved motif and results indicated that all genes have the same functions, such as stress responses. Our findings are in agreement with previous studies that the NAC family has a highly conserved domain [48,58]. Therefore, *ZmNAC9* might have the same functions as their homologous genes in maize as well as in other species.

The natural variation mostly caused by genetic variation is in the ancestor of maize teosinte [59]. Previous studies showed that the average nucleotide frequency of one SNP included 104 bp on 1 chromosome [60]. One single nucleotide polymorphism per 31 bp was found in non-coding regions and one polymorphism per 124 bp was found in the coding regions of 18 maize genes in 36 elite maize inbred lines [61]. In this study, candidate gene association was performed on 111 maize inbred lines at the seedling stage under low-P and normal-P conditions. A total of 41 polymorphic loci including 32 single nucleotide polymorphism (SNP) and 9 insertion and deletion (InDels) markers were identified in the coding sequence region of *ZmNAC9* in 111 inbred lines with a minor allele frequency of $p < 0.05$ (Table 3 and Table S2). Less nucleotide diversity was found in the *ZmNAC9* gene with one polymorphic site. Similar results were previously observed, finding that the level of nucleotide diversity might be contributed to by gene properties in different population sizes and diverse plant species and the nucleotide might have occurred on an average of 47.7, 61 and 104 bp in the gene region [62–64]. The findings of this study suggested that the SNPs and InDels could be essential to generate new changes during the evolution of maize species. Besides, Tajima's D and Fu and Li's D* tests showed that *ZmNAC9* had non-significant neutral variation in the coding region based on Tajima's D and significant variation on Fu and Li's D* tests at $p < 0.10$. This was a not-obvious artificial selection and the negative value of genes indicated that the neutral selection hypothesis could not be excluded and the gene was naturally selected in the screening of germplasm. The effect of selection is small but it is easier to produce gene drift and so on. The results were changed only in the sequence of genes in maize inbred lines. The findings of this study are similar with previous studies that found that the artificial selection of genes might affect the coding region of genes and depends upon the number of germplasm [57,62,63]. Furthermore, LD analysis was analyzed because it is an essential method to examine the genomic variation for the association mapping of complex traits [64]. In our study, the LD level was high in some sites and low in other sites which might be due to genetic mutations in the gene segment. Besides, the LD decay of *ZmNAC9* declined lower than the expected r2 value within 500 bp (Figure 5A). LD decay indicated that *ZmNAC9* was less affected by natural variation and it was

not under the selection pressure during germplasm and these loci might be functional sites for the response to low phosphorus (Figure 5B). Our results are in agreement with the previous studies that LD decay could be different between 200 and 2000 bp and selective sweep in maize lines [52].

Moreover, the candidate gene association can identify the natural variation and is best on haplotypes for targeted trait responses to low-P stress conditions in maize at the seedling stage. In this study, we used 22 traits for the *ZmNAC9* association analysis. The results revealed that *ZmNAC9* has good potential in low-P stress; among 41 polymorphic sites, 39 sites were associated with 20 traits and a total of 191 associations were significant with the P-tolerant traits index except for the dry weight of shoots and forks using the mixed linear model (MLM) including Q+K. Six sites were highly significantly associated with a diverse number of P tolerant trait index values (Figure 6a, Table 3 and Table S3). At the same time, among 41 sites 29 sites were associated with 22 traits, of which six polymorphic sites (position S327, S513, S514, S520, S633, and S839) were highly significantly associated with the variable number of low-P seedling traits under normal-P conditions (Figure 6b, Table 3 and Table S3), whereas, 32 polymorphic loci included SNPs and InDels were significantly associated with the different number of seedling traits, and the remaining nine polymorphic loci were non-significantly associated with each trait under low-P conditions (Figure 6c, Table 3 and Table S3). Five sites (position S242, S513, S521, S824, and S827) were strong significantly associated with multiple seedling traits using the same threshold under low-P stress conditions. Furthermore, haplotypes were analyzed in the tested population. Results showed that among the identified haplotypes five major haplotypes contained more than 10 inbred lines, of which hap4 with a more favorable allele showed the highest significant difference at $p < 0.05$. Previous studies have identified favorable natural variation in candidate genes—for example, *ZmDREB2.7*, *ZmNAC111*, *ZmPP2C-A*, *ZmVPP1*, and *ZmHKT1* (5, by using the MLM statistical model) [13,57,65,66].

Plants have evolved many physiological, biochemical, morphological, and molecular mechanism adaptation strategies to cope with P-deficiency stress [13]. Recently, many researchers have been identified the phosphorus transporter members in different crops such as, PHO1, PHO2, PHT1, PHT2, and PHT3 five major family members, of which most member of PHT1 family mainly responsible for phosphorus absorption from the soil, distribution within plants and up-regulated under low-P stress condition [23]. In our study, the expression pattern was increased in the early stage but decreased in the later stage of low-P treatment and *ZmNAC9* was dominant in roots but also increased in leaves and was significantly induced by the low-P stress condition in the P-tolerant 178 and P-sensitive 9782 inbred lines. There is no report about the P transporter *ZmNAC9* in maize. Results suggested that ZmNAC9 was probably involved in response to the P-starvation condition and might play an important role in maize and improve the low-P tolerance in maize.

The functions and characterization of any gene product in a post-genome sequence was the main challenges for plant biologists because it can be limiting to understand the biological roles of many genes [65]. The localization of gene expression products in cells is the main issue in cell biology and an important study in molecular biology [54]. Therefore, we identified the subcellular localization of *ZmNAC9* through the transient transformation method and a targeted fragment was fused with the green fluorescence protein (GFP) introduced in tobacco leaves driven by the CamV35S promoter and the green fluorescence protein (GFP) of *ZmNAC9* was detected in nucleus and cytoplasm (Figure 9). This is different from the natural transcription factors only localized in the nucleus by the predication cello program [51]. The subcellular localization of *ZmNAC9* might be disturbed due to various cellular processes because NAC transcription factors have a different conserved (NAM, ATAF and CUC) domain and each domain has a diverse function and environmental stress and the location might be changed [66].

4. Materials and Methods

4.1. Plant Materials, Growth Conditions and Phenotyping

A total of 111 diverse genotypically maize inbred lines (46 tropical 65 temperate), were screened under low-P stress conditions to identify the low-P tolerant inbred lines from the Southwest China breeding program [67]. These diverse heterotic groups of 111 maize inbred lines were used for a candidate gene association study (Table S1). The phenotypic traits difference was carefully identified after 15 days of low-P-stress and control P-sufficient condition. The manually evaluated traits were: primary root length (PRL); root maximum length (MRL); plant height (PH); number of seminal roots (NSRs); number of crown roots (NCRs); number of leaves (NLs); shoot fresh weight (FSW); seminal root fresh weight (FWSR); primary root fresh weight (FWPR); crown root fresh weight (FWCR); total root fresh weight (FWR); total plant fresh weight (FWP); root biomass (BR); shoot biomass (SB); root/shoot ratio (R/S); total plant biomass (DWP). The maize root architecture system was classified according to the protocol described in [68]. In addition, the total root length (TRL), root surface area (RS), root average diameter (AD), root volume (RV), number of root tips (T), and number of root forks (F) under low- and normal-P conditions have been previously analyzed [36]. These twenty-two significant traits were used for natural variation in the *ZmNAC9* gene.

4.2. Bioinformatics Analysis of NAC TFs Family Members in Maize

The amino acid sequence of *ZmNAC9* gene was used as a query to search for the exact number of ZmNAC candidate genes in the maize family by using the Pfam http://pfam.sanger.ac.uk and http://http://smart.embl-heidelberg.de/ online databases. To evaluate the evolutionary relationship of ZmNAC family members in maize, the protein sequences of ZmNAC members were aligned in ClustalW to generate the phylogenetic tree of the ZmNAC family members of maize by using the neighbor-joining method in MEGA v7.0 [69]. The uniform rate was tested using 1000 bootstrap replicates and named according to their position correction in the phylogenetic tree. Besides, the relationship of *ZmNAC9* and its homologous genes was compared with other species, and the protein sequence of *ZmNAC9* and its homologous genes were used in the BLASTP program against Arabidopsis and rice with an E-value of 1×10^{-10}. The phylogenetic tree was constructed using the same methods as ZmNACs family members. Furthermore, the basic molecular characteristics of *ZmNAC9* and its homologous genes were identified through the ExPASy online database (https://www.expasy.org/) and genes were summarized according to their physical and chemical properties such as the accession number of genes, genes on a chromosome, number of transcriptions, molecular weight, isoelectric points, and grand average of hydropathicity (GRAVY) for each putative protein. Furthermore, the chromosome location of *ZmNAC9* homologous genes was downloaded from the Gramene database and chromosomal location distance was calculated by using Map-Draw software 2.1 as previously described [70]. In addition, the gene structure of *ZmNAC9* homologous genes including the exons, introns, the 5' and 3' UTR regions were evaluated through the gene structure display server (GSDS) online program [71]. Moreover, the prediction of the conserved domain of ZmNAC9 homologous genes was evaluated through the MEME online prediction database using the full length of an amino acid's sequence of genes. The parameter settings are as follows: the maximum number of motifs was 5 and length of the amino acid was 6–600 (aa).

4.3. Sequencing Polymorphism, Nucleotide Diversity and LD of ZmNAC9 Candidate Gene

The total genomic DNA of 111 inbred lines was extracted from fresh roots and leaves by using the cetyltrimethyl ammonium bromide (CTAB) method described in [72]. The genomic DNA sequence of *ZmNAC9* of the B73 inbred line reference sequence was downloaded from Maize GDB (http://www.maizegdb.org). The forward (AGTGTTTGGACGGGGTTGAGTCCATTTCA) and reverse (AGCGTTGCTGGAAGGGAACTGATTGATTT) primers were designed according to the target sequence of a gene by using primer premier 5v software. The high fidelity phanta max enzyme

polymerase was used for the amplification of a targeted fragment of genes and DNA was used as a template; the total reaction volume was 25μL, containing 12.5 μL phanta max buffer, 0.5 μL dNTP, 0.5 μL phanta max super fidelity, 8.5 μL ddH2O, 1 μL genomic DNA and 1 μL of each sense and anti-sense. The PCR cycle procedure was set as: initial denaturation at 95 °C for 3 min followed by 30 cycles of denaturation at 95°C for 15 sec and an annealing step for 15 s at 63°C, extension at 72°C for 1.15 min, an extension for 5 min at 72 and 12 °C for preservation. Amplified fragments for all inbred lines were determined by gel electrophoresis (Biorad, California) using 1% agarose gels in Tris-EDTA (TE) buffer. Agarose gel was stained with 2 uL of ethidium bromide and the running time was 25 min at 220 mV and then target fragments were direct and subjected to TsingKe Biological Technology Co., Ltd for sequencing. The targeted sequence of the ZmNAC9 gene was compared with the reference genomic sequence of the B73 inbred line by using DNAMAN v6.0 software. The sequence was adjusted manually by using Bio-Edit 7.1 software [56], and a multiple sequence alignment of ZmNAC9 gene was performed for 111 inbred lines by using clustalX 2.0 [73]. We analyzed single nucleotide polymorphisms (SNPs) and insertion and deletions (InDels) in all tested lines with a <0.05 minor allele frequency by using Tassel software v5.0 [74]. Furthermore, the linkage disequilibrium (LD) between two polymorphic sites in the coding region of the ZmNAC9 gene was generated in an LD plot by using TASSELv5.0 [74]. In addition, the average of the r2 values was used to calculate the LD decay with physical distance and a loess regression test was applied by using the R language in the gg plot package. In addition, nucleotide diversity, nucleotide polymorphism, haplotype diversity, haplotypes, haplotype distribution, and Tajima's D and Fu and Li's D tests were analyzed among 111 sequenced maize inbred lines for a neutrality test by using DnaSp v5.0 [75]. Furthermore, a sliding window of 100 bp with a 25 bp step size and π and θ values were used to measure the nucleotides' diversity.

4.4. Candidate ZmNAC9 Gene Association Analysis

An association analysis between single nucleotide polymorphism (SNPs) and insertion and deletion (InDel) sites in the *ZmNAC9* gene with 22 phenotypic traits were calculated using Tassel software v5.0 [74]. The standard mixed linear model (MLM) included a population structure and kinship matrix which were considered for false-positive control in association analysis under low-P stress conditions as previously described in [55]. In which both the population (Q) and kinship (k) matrix were calculated as earlier described [67]. The identified 32 SNPs and 9 InDels polymorphic loci with a MAF at $p \geq 0.05$ were used for candidate gene association analysis. The association analysis was considered significant at $p < 0.01$ and the calculated p-values were converted into $-\log10(P)$ to improve the accuracy of the association between polymorphic loci and low-P tolerant traits under low-P, normal-P conditions and the corresponding (T/CK) low-P tolerant trait index using a uniform Bonferroni-corrected test as described [36,76].

4.5. Low-P Stress Treatments, RNA Extraction and Expression Pattern of Candidate ZmNAC9 Gene

The seeds of 178 P-tolerant and 9782 P-sensitive inbred lines were provided by Maize Research Institute, Sichuan Agricultural University, Chengdu, China. Seeds were soaked at 37 °C for 3 days. After soaking, seeds were surface-sterilized using a 10% sodium hypochlorite solution for 20–30 min. Seeds were rinsed with distilled water two to four times and kept for 12 h at 28 °C and then transferred into a plastic vessel (5 L capacity of distill water) fitted with brown thermocol sheets. Standardized Hoagland solution was applied in plastic vessels for the initial three days of seedling to adapt to the hydroponic environment. The composition of Hoagland nutrients contained low P (−P, 1 μmol l−1 KH$_2$PO$_4$) and sufficient P (+P, 1 mmol l−1 KH$_2$PO$_4$). The pH was maintained for a nutrient solution around 5.5 and the aquarium pump was used during the experiment. Root and leaves were harvested separately at 0 h, 1, 3, 7, 9, 12, 14, and 16 days under sufficient-P and deficient-P conditions. Total RNA was extracted from each sample according to the manufacturer's protocol using TRIzol Reagent® (Life Technologies now ThermoFisher Scientific, https://www.thermofisher.com). Reverse transcription of cDNA using a Prime Script TM II 1st Strand cDNA Synthesis Kit and cDNA was used as a template

using Fast Start Essential DNA Green Master (Roche, https://www.roche.com) on a Roche Cobas Z480 system. The specific primers were designed using BEACON DESIGNER 7 software; qRT-PCR was programmed in initial step at 95 °C for 10 m, followed by 40 cycles including denaturation at 95 °C for 10 s, annealing at 58.4 °C for 30 s and extension at 60 °C for 15 s. The myosin housekeeping gene was used as an internal control and data were calculated in the $2^{-\Delta\Delta CT}$ method as described [77] with at least three biological repeats for each sample and three technical repeats were performed for each biological replicate. Three technical replications were used to calculate the average and standard deviation (SD) of expression levels for each sample.

4.6. Subcellular Localization of ZmNAC9 Gene

The full-length CDS region of the *ZmNAC9* gene was cloned using the specific primers; forward ATTTGGAGAGGACAGGGTACCATGGAGGCCCGGCCCGG (SAM1 underlined enzyme site) and reverse GGTACTAGTGTCGACTCTAGAATTAGAGGAGGCGACGCTGCAGCT (Xba1 underlined enzyme site) were added in both primers. The subcellular localization was assayed in tobacco leaves according to the reported transit transformation method [78,79]. The fragment of the gene was fused with pCAMBIA2300-35S-GFP; the position of *ZmNAC9* gene was between the cauliflower mosaic virus (CaMV) 35S promoter and GFP gene. The empty pCAMBIA2300 vector without GFP was used as a control and the loaded vector pCAMBIA2300-35S; *ZmNAC9*-eGFP (recombinant plasmid) was transformed into the *Agrobacterium tumefaciens* EHA105 strain and then introduced into tobacco (*Nicotiana benthamiana*) leaves as previously described [80]. The transformed tobacco epidermal cells were observed by the A1R-si laser confocal microscope (LSCM, Nikon, Japan) to imagine the GFP fluorescent signals. At least three times the experiment was repeated for consistent results.

5. Conclusions

In this study, a total of 208 ZmNAC family members were identified and *ZmNAC9* homologous genes, their molecular characteristics, gene structure, and conserved domains were studied. A bioinformatics analysis did not lead the functional difference but provided the reference for the functional analysis. Furthermore, the natural variation of *ZmNAC9* revealed an association of candidate gene with low-P tolerant traits and might allow the investigation of molecular markers under low-P stress conditions and genetic background. In addition, we validated the expression pattern of *ZmNAC9* under low-P stress by using the low-P tolerant 178 and P-sensitive 9782 inbred lines. Our results provide useful information for the improvement of low-P tolerance in maize and to understand the molecular characteristics and functions of *ZmNAC9* under low-P conditions in future molecular breeding programs in maize.

Author Contributions: J.H.S., S.G. conceived and designed the experiment; J.H.S., B.L. and X.H. conducted the experiments; J.H.S., X.Z., D.L., L.W., H.Z., C.Z., P.M., M.H.B.K. and A.A. data analyzed; J.H.S. writing—original draft preparation; J.H.S., H.Z., B.L., H.A.S., and Z.G.N.G. writing—review and editing; S.G. supervision. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Blümmel, M.; Grings, E.; Erenstein, O. Potential for dual-purpose maize varieties to meet changing maize demands: Synthesis. *Field Crop. Res.* **2013**, *153*, 107–112. [CrossRef]
2. Ramakrishnan, M.; Ceasar, S.A.; Duraipandiyan, V.; Ignacimuthu, S. Efficient plant regeneration from shoot apex explants of maize (Zea mays) and analysis of genetic fidelity of regenerated plants by ISSR markers. *Plant Cell Tissue Organ Cult.* **2014**, *119*, 183–196. [CrossRef]
3. Prado, S.A.; López, C.G.; Senior, M.L.; Borrás, L. The genetic architecture of maize (Zea mays L.) kernel weight determination. *G3 Genes Genomes Genet.* **2014**, *4*, 1611–1621.
4. Lobell, D.B.; Wolfram, S.; Justin, C.R. Climate trends and global crop production since 1980. *Science* **2011**, *333*, 616–620. [CrossRef]
5. Shao, H.B.; Chu, L.Y.; Jaleel, C.A.; Manivannan, P.; Panneerselvam, R.; Shao, M.A. Understanding water deficit stress-induced changes in the basic metabolism of higher plants—Biotechnologically and sustainably improving agriculture and the ecoenvironment in arid regions of the globe. *Crit. Rev. Biotechnol.* **2009**, *29*, 131. [CrossRef]
6. Zhu, J.K. Plant salt tolerance. *Trends Plant Sci.* **2001**, *6*, 66–71. [CrossRef]
7. Khalid, M.H.B.; Raza, M.A.; Yu, H.Q.; Khan, I.; Sun, F.A.; Feng, L.Y.; Qu, J.T.; Fu, F.L.; Li, W.C. Expression, Subcellular Localization, and Interactions of CPK Family Genes in Maize. *Int. J. Mol. Sci.* **2019**, *20*, 6173. [CrossRef]
8. Schachtman, D.P.; Reid, R.J.; Ayling, S.M. Phosphorus Uptake by Plants: From Soil to Cell. *Plant Physiol.* **1998**, *116*, 447–453. [CrossRef]
9. Zhang, Y.; Peng, Y.U.; Yun-Feng, P.E.N.G.; Xue-Xian, L.I.; Fan-Jun, C.H.E.N.; Chun-Jian, L.I. Fine Root Patterning and Balanced Inorganic Phosphorus Distribution in the Soil Indicate Distinctive Adaptation of Maize Plants to Phosphorus Deficiency. *Pedosphere* **2012**, *22*, 870–877. [CrossRef]
10. Cordell, D.; Drangert, J.O.; White, S. The story of phosphorus: Global food security and food for thought. *Glob. Environ. Chang.* **2009**, *19*, 292–305. [CrossRef]
11. Jianbo, S.; Lixing, Y.; Junling, Z.; Haigang, L.; Zhaohai, B.; Xinping, C.; Weifeng, Z.; Fusuo, Z. Phosphorus dynamics: From soil to plant. *Plant Physiol.* **2011**, *156*, 997–1005.
12. Schröder, J.J.; Smit, A.L.; Cordell, D.; Rosemarin, A. Improved phosphorus use efficiency in agriculture: A key requirement for its sustainable use. *Chemosphere* **2011**, *84*, 822–831. [CrossRef]
13. Raghothama, K.G.; Karthikeyan, A.S. Phosphate acquisition. *Plant Soil* **2005**, *274*, 37. [CrossRef]
14. Rausch, C.; Bucher, M. Molecular mechanisms of phosphate transport in plants. *Planta* **2002**, *216*, 23–37. [CrossRef]
15. Péret, B.; Clément, M.; Nussaume, L.; Desnos, T. Root developmental adaptation to phosphate starvation: Better safe than sorry. *Trends Plant Sci.* **2011**, *16*, 442–450. [CrossRef]
16. Williamson, L.; Ribrioux, S.P.C.P.; Fitter, A.H.; Leyser, H.M.O. Phosphate Availability Regulates Root System Architecture in Arabidopsis. *Plant Physiol.* **2001**, *126*, 875–882. [CrossRef]
17. Lambers, H.; Colmer, T.D. Root Physiology—From Gene to Function. *Plant Soil* **2005**, *274*, 7–15. [CrossRef]
18. Hurley, B.A.; Tran, H.T.; Marty, N.J.; Park, J.; Snedden, W.A.; Mullen, R.T.; Plaxton, W.C. The Dual-Targeted Purple Acid Phosphatase Isozyme AtPAP26 Is Essential for Efficient Acclimation of Arabidopsis to Nutritional Phosphate Deprivation. *Plant Physiol.* **2010**, *153*, 1112–1122. [CrossRef]
19. Dionisio, G.; Madsen, C.K.; Holm, P.B.; Welinder, K.G.; Jorgensen, M.M.; Stoger, E.; Arcalis, E.; Brinchpedersen, H. Cloning and Characterization of Purple Acid Phosphatase Phytases from Wheat, Barley, Maize, and Rice. *Plant Physiol.* **2011**, *156*, 1087–1100. [CrossRef]
20. Chiou, T.J.; Shui, L. Signaling Network in Sensing Phosphate Availability in Plants. *Annu. Rev. Plant Biol.* **2011**, *62*, 185–206. [CrossRef]
21. Yang, X.J.; Finnegan, P.M. Regulation of phosphate starvation responses in higher plants. *Ann. Bot.* **2010**, *105*, 513–526. [CrossRef] [PubMed]
22. Secco, D.; Wang, C.; Arpat, B.A.; Wang, Z.; Poirier, Y.; Tyerman, S.D.; Wu, P.; Shou, H.; Whelan, J. The emerging importance of the SPX domain-containing proteins in phosphate homeostasis. *New Phytol.* **2012**, *193*, 842–851. [CrossRef]

23. Li, Z.; Qiang, G.; Liu, Y.; He, C.; Zhang, X.; Zhang, J. Erratum to: Overexpression of transcription factor ZmPTF1 improves low phosphate tolerance of maize by regulating carbon metabolism and root growth. *Planta* **2011**, *233*, 1145. [CrossRef]

24. Poirier, Y.; Thoma, S.; Somerville, C.; Schiefelbein, J. Mutant of Arabidopsis deficient in xylem loading of phosphate. *Plant Physiol.* **1991**, *97*, 1087–1093. [CrossRef] [PubMed]

25. Dirk, H.; Enea, R.; Jean, M.D.C.P.; Chris, S.; Yves, P. Identification and characterization of the Arabidopsis PHO1 gene involved in phosphate loading to the xylem. *Plant Cell* **2002**, *14*, 889–902.

26. Yang, X.; Yan, J.; Zheng, Y.; Li, J. Reviews of Association Analysis for Quantitative Traits in Plants. *Acta Agron. Sin.* **2007**, *33*, 523.

27. Gupta, P.K.; Rustgi, S.; Kulwal, P.L. Linkage disequilibrium and association studies in higher plants: Present status and future prospects. *Plant Mol. Biol.* **2005**, *57*, 461–485. [CrossRef]

28. Flint-Garcia, S.A.; Thornsberry, J.M.; Th, B.E. Structure of linkage disequilibrium in plants. *Annu. Rev. Plant Biol.* **2003**, *54*, 357–374. [CrossRef]

29. Flintgarcia, S.; Thuillet, A.; Yu, J.; Pressoir, G.; Romero, S.M.; Mitchell, S.E.; Doebley, J.; Kresovich, S.; Goodman, M.M.; Buckler, E.S. Maize association population: A high-resolution platform for quantitative trait locus dissection. *Plant J.* **2005**, *44*, 1054–1064. [CrossRef]

30. Caldwell, K.S.; Joanne, R.; Peter, L.; Wayne, P. Extreme population-dependent linkage disequilibrium detected in an inbreeding plant species, Hordeum vulgare. *Genetics* **2006**, *172*, 557–567. [CrossRef]

31. Mohlke, K.L.; Lange, E.M.; Valle, T.T.; Ghosh, S.; Magnuson, V.L.; Silander, K.; Watanabe, R.M.; Chines, P.S.; Bergman, R.N.; Tuomilehto, J. Linkage Disequilibrium Between Microsatellite Markers Extends Beyond 1 cM on Chromosome 20 in Finns. *Genome Res.* **2001**, *11*, 1221–1226. [CrossRef] [PubMed]

32. Buckler, E.S.; Gaut, B.S.; Mcmullen, M.D. Molecular and functional diversity of maize. *Curr. Opin. Plant Biol.* **2006**, *9*, 172–176. [CrossRef] [PubMed]

33. Buckler, E.S.; Thornsberry, J.M. Plant molecular diversity and applications to genomics. *Curr. Opin. Plant Biol.* **2002**, *5*, 107–111. [CrossRef]

34. Mcmullen, M.D.; Stephen, K.; Hector Sanchez, V.; Peter, B.; Huihui, L.; Qi, S.; Sherry, F.G.; Jeffry, T.; Charlotte, A.; Christopher, B. Genetic properties of the maize nested association mapping population. *Science* **2009**, *325*, 737–740. [CrossRef] [PubMed]

35. Tian, F.; Bradbury, P.J.; Brown, P.J.; Hung, H.; Sun, Q.; Flintgarcia, S.; Rocheford, T.; Mcmullen, M.D.; Holland, J.B.; Buckler, E.S. Genome-wide association study of leaf architecture in the maize nested association mapping population. *Nat. Genet.* **2011**, *43*, 159–162. [CrossRef]

36. Luo, B.; Ma, P.; Nie, Z.; Zhang, X.; He, X.; Ding, X.; Feng, X.; Lu, Q.; Ren, Z.; Lin, H. Metabolite profiling and genome-wide association studies reveal response mechanisms of phosphorus deficiency in maize seedling. *Plant J.* **2019**, *97*, 947–969. [CrossRef]

37. Kage, U.; Kumar, A.; Dhokane, D.; Karre, S.; Kushalappa, A.C. Functional molecular markers for crop improvement. *Crit. Rev. Biotechnol.* **2016**, *36*, 917–930. [CrossRef]

38. Assenov, B.; Andjelkovic, V.; Ignjatovic-Micic, D.; Vancetovic, J.; Todorovska, E. Identification of SNP Mutations in MYBE-1 Gene Involved in Drought Stress Tolerance in Maize. *Bulg. J. Agric. Sci.* **2013**, *19*, 181–185.

39. Chai, Y.; Hao, X.; Yang, X.; Allen, W.B.; Li, J.; Yan, J.; Shen, B.; Li, J. Validation of DGAT1-2 polymorphisms associated with oil content and development of functional markers for molecular breeding of high-oil maize. *Mol. Breed.* **2012**, *29*, 939–949. [CrossRef]

40. Nuruzzaman, M.; Sharoni, A.M.; Kikuchi, S. Roles of NAC transcription factors in the regulation of biotic and abiotic stress responses in plants. *Front. Microbiol.* **2013**, *4*, 248. [CrossRef] [PubMed]

41. Aida, M.; Ishida, T.; Fukaki, H.; Fujisawa, H.; Tasaka, M. Genes involved in organ separation in Arabidopsis: An analysis of the cup-shaped cotyledon mutant. *Plant Cell* **1997**, *9*, 841–857. [CrossRef]

42. Souer, E.; Houwelingen, A.V.; Kloos, D.; Mol, J.; Koes, R. The No Apical Meristem Gene of Petunia Is Required for Pattern Formation in Embryos and Flowers and Is Expressed at Meristem and Primordia Boundaries. *Cell* **1996**, *85*, 159–170. [CrossRef]

43. Rushton, P.J.; Bokowiec, M.S.; Zhang, H.; Brannock, J.F.; Chen, X.; Laudeman, T.W.; Timko, M.P. Tobacco transcription factors: Novel insights into transcriptional regulation in the solanaceae. *Plant Physiol.* **2008**, *147*, 280–295. [CrossRef]

44. Fan, K.; Wang, M.; Miao, Y.; Ni, M.; Bibi, N.; Yuan, S.; Li, F.; Wang, X. Molecular Evolution and Expansion Analysis of the NAC Transcription Factor in Zea mays. *PLoS ONE* **2014**, *9*, e111837. [CrossRef]

45. Shiriga, K.; Sharma, R.; Kumar, K.; Yadav, S.K.; Hossain, F.; Thirunavukkarasu, N. Genome-wide identification and expression pattern of drought-responsive members of the NAC family in maize. *Meta Gene* **2014**, *2*, 407–417. [CrossRef]

46. Zhu, Y.; Yan, J.; Liu, W.; Liu, L.; Sheng, Y.; Sun, Y.; Li, Y.; Scheller, H.V.; Jiang, M.; Hou, X. Phosphorylation of a NAC transcription factor by ZmCCaMK regulates abscisic acid-induced antioxidant defense in maize. *Plant Physiol.* **2016**, *171*, 1651. [CrossRef]

47. Peng, X.; Zhao, Y.; Xiaoming, L.I.; Min, W.U.; Chai, W.; Sheng, L.; Wang, Y.U.; Dong, Q.; Jiang, H.; Cheng, B. Genomewide identification, classification and analysis of NAC type gene family in maize. *J. Genet.* **2015**, *94*, 377–390. [CrossRef]

48. Olsen, A.N.; Ernst HALeggio, L.L.; Skriver, K. NAC transcription factors: Structurally distinct, functionally diverse. *Trends Plant Sci.* **2005**, *10*, 79–87. [CrossRef]

49. Nilsson, L.; Müller, R.; Nielsen, T.H. Dissecting the plant transcriptome and the regulatory responses to phosphate deprivation. *Physiol. Plant.* **2010**, *139*, 129–143. [CrossRef]

50. Waters, B.M.; Uauy, C.; Dubcovsky, J.; Grusak, M.A. Wheat (Triticum aestivum) NAM proteins regulate the translocation of iron, zinc, and nitrogen compounds from vegetative tissues to grain. *J. Exp. Bot.* **2009**, *60*, 4263–4274. [CrossRef]

51. Chin-Sheng, Y.; Chih-Wen, C.; Wen-Chi, S.; Kuei-Chung, C.; Shao-Wei, H.; Jenn-Kang, H.; Chih-Hao, L. CELLO2GO: A web server for protein subCELlular LOcalization prediction with functional gene ontology annotation. *PLoS ONE* **2014**, *9*, e99368.

52. Sahito, J.H.; Zheng, F.; Tang, H.; He, X.; Luo, B.; Zhang, X.; Ma, P.; Ding, X.; Liu, D.; Wu, L.; et al. Identification, association, and expression analysis of ZmNAC134 gene response to phosphorus deficiency tolerance traits in maize at seedling stage. *Euphytica* **2020**, *216*, 100. [CrossRef]

53. Yu, J.; Buckler, E.S. Genetic association mapping and genome organization of maize. *Curr. Opin. Biotechnol.* **2006**, *17*, 155–160. [CrossRef]

54. Sara, F.; Thorpe, F.L.; Julia, E.; Jessika, A.; Xue, D.; Liron, S.K.; Justin, G.; Franziska, T. Tissue-specific expression of FLOWERING LOCUS T in Arabidopsis is maintained independently of polycomb group protein repression. *Plant Cell* **2011**, *23*, 3204–3214.

55. Shengxue, L.; Xianglan, W.; Hongwei, W.; Haibo, X.; Xiaohong, Y.; Jianbing, Y.; Jiansheng, L.; Lam-Son Phan, T.; Kazuo, S.; Kazuko, Y.S. Genome-wide analysis of ZmDREB genes and their association with natural variation in drought tolerance at seedling stage of Zea mays L. *Plos Genet.* **2013**, *9*, e1003790.

56. Strable, J.; Scanlon, M.J. Maize (Zea mays): A Model Organism for Basic and Applied Research in Plant Biology. *Cold Spring Harb. Protoc.* **2009**, *2009*, pdb-emo132. [CrossRef]

57. Wu, F.; Liu, Z.; Xu, J.; Gao, S.; Lin, H.; Liu, L.; Liu, Y.; Lu, Y. Molecular Evolution and Association of Natural Variation inZmARF31with Low Phosphorus Tolerance in Maize. *Front. Plant Sci.* **2016**, *7*, 1076. [CrossRef]

58. Ooka, H.; Satoh, K.; Doi, K.; Nagata, T.; Otomo, Y.; Murakami, K.; Matsubara, K.; Osato, N.; Kawai, J.; Carninci, P. Comprehensive analysis of NAC family genes in Oryza sativa and Arabidopsis thaliana. *Dna Res. Int. J. Rapid Publ. Rep. Genes Genomes* **2003**, *10*, 239. [CrossRef]

59. Tian, F.; Stevens, N.M.; Buckler, E.S. Tracking footprints of maize domestication and evidence for a massive selective sweep on chromosome 10. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 9979–9986. [CrossRef]

60. Tenaillon, M.I.; Sawkins, M.C.; Long, A.D.; Gaut, R.L.; Doebley, J.F.; Gaut, B.S. Patterns of DNA sequence polymorphism along chromosome 1 of maize (Zea mays ssp. mays L.). *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 9161–9166. [CrossRef]

61. Ching, A.; Caldwell, K.S.; Jung, M.; Dolan, M.; Smith, O.S.; Tingey, S.; Morgante, M.; Rafalski, A.J. SNP frequency, haplotype structure and linkage disequilibrium in elite maize inbred lines. *BMC Genet.* **2002**, *3*, 19. [CrossRef]

62. Li, Q.; Yang, X.; Bai, G.; Warburton, M.L.; Mahuku, G.; Gore, M.; Dai, J.; Li, J.; Yan, Y. Cloning and characterization of a putative GS3 ortholog involved in maize kernel development. *Tag.Theor. Appl. Genet. Theor. Angew. Genet.* **2010**, *120*, 753–763.

63. Su, Z.; Li, X.; Hao, Z.; Xie, C.; Li, M.; Weng, J.; Zhang, D.; Liang, X.; Wang, Z.; Gao, J. Association Analysis of the nced and rab28 Genes with Phenotypic Traits Under Water Stress in Maize. *Plant Mol. Biol. Report.* **2011**, *29*, 714–722. [CrossRef]

64. Clark, R.M.; Eric, L.; Joachim, M.; Doebley, J.F. Pattern of diversity in the genomic region near the maize domestication gene tb1. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 700–707. [CrossRef]

65. Venugopal Rao, K.; Peng, Y.; Sadanandam, A.; Zhen, Z.; Pavan, U. Subcellular localization of proteins of Oryza sativa L. in the model tobacco and tomato plants. *Plant Signal. Behav.* **2010**, *5*, 1336–1341.

66. Wang, D.; Yu, Y.; Liu, Z.; Li, S.; Wang, Z.; Xiang, F. Membrane-bound NAC transcription factors in maize and their contribution to the oxidative stress response. *Plant Sci.* **2016**, *250*, 30–39. [CrossRef]

67. Zhang, X.; Zhang, H.; Li, L.; Lan, H.; Ren, Z.; Liu, D.; Wu, L.; Liu, H.; Jaqueth, J.; Li, B. Characterizing the population structure and genetic diversity of maize breeding germplasm in Southwest China using genome-wide SNP markers. *BMC Genom.* **2016**, *17*, 697. [CrossRef]

68. Hochholdinger, F.; Tuberosa, R. Genetic and genomic dissection of maize root development and architecture. *Curr. Opin. Plant Biol.* **2009**, *12*, 172–177. [CrossRef]

69. Kumar, S.; Stecher, G.; Tamura, K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol. Biol. Evol.* **2016**, *33*, 1870. [CrossRef]

70. Wu, X.; Yang, H.; Qu, C.; Xu, Z.; Li, W.; Hao, B.; Yang, C.; Sun, G.; Liu, G. Sequence and expression analysis of the AMT gene family in poplar. *Front. Plant Sci.* **2015**, *6*, 337. [CrossRef]

71. Guo, A.Y.; Zhu, Q.H.; Xin, C. GSDS:a gene structure display server. *Hereditas* **2007**, *29*, 1023–1026.

72. Porebski, S.; Bailey, L.G.; Baum, B.R. Modification of a CTAB DNA extraction protocol for plants containing high polysaccharide and polyphenol components. *Plant Mol. Biol. Report.* **1997**, *15*, 8–15. [CrossRef]

73. Thompson, J.D.; Gibson, T.J.; Higgins, D.G. Multiple Sequence Alignment Using ClustalW and ClustalX. *Curr Protoc Bioinform.* **2002**. [CrossRef]

74. Bradbury, P.; Zhang, Z.; Kroon, D.; Casstevens, T.Y.; Buckler, E. TASSEL: Software for association mapping of complex traits in diverse samples. *Bioinformatics* **2007**, *23*, 2633–2635. [CrossRef]

75. Librado, P.; Rozas, J. DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* **2009**, *25*, 1451–1452. [CrossRef]

76. Yang, N.; Lu, Y.; Yang, X.; Huang, J.; Zhou, Y.; Ali, F.; Wen, W.; Liu, J.; Li, J.; Yan, J. Genome Wide Association Studies Using a New Nonparametric Model Reveal the Genetic Architecture of 17 Agronomic Traits in an Enlarged Maize Association Panel. *PLoS Genet.* **2014**, *10*, e1004573. [CrossRef]

77. Livak, K.J.; Schmittgen, T.D. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* **2001**, *25*, 402–408. [CrossRef]

78. Andersen, J.R.; Lubberstedt, T. Functional markers in plants. *Trends Plant Sci.* **2003**, *8*, 554–560. [CrossRef]

79. Domon, E.; Yanagisawa, T.; Saito, A.; Takeda, K. Single nucleotide polymorphism genotyping of the barley waxy gene by polymerase chain reaction with confronting two-pair primers. *Plant Breed.* **2004**, *123*, 225–228. [CrossRef]

80. Fang, Y.; Liao, K.; Du, H.; Xu, Y.; Song, H.; Li, X.; Xiong, L. A stress-responsive NAC transcription factor SNAC3 confers heat and drought tolerance through modulation of reactive oxygen species in rice. *J. Exp. Bot.* **2015**, *66*, 6803–6817. [CrossRef]