JIPB Journal of Integrative Plant Biology

Genetic mapping of folate QTLs using a segregated population in maize[FA]

Wenzhu Guo[1†], Tong Lian[2,3†], Baobao Wang[3†], Jiantao Guan[4†], Dong Yuan[5], Huan Wang[3], Fardous Mohammad Safiul Azam[3], Xing Wan[6], Weixuan Wang[3], Qiuju Liang[3], Haiyang Wang[3], Jinxing Tu[1], Chunyi Zhang[3] and Ling Jiang[3]*

1. College of Plant Science and Technology, Huazhong Agricultural University, Wuhan 430070, China
2. Plant Genetics, Gembloux Agro-Bio Tech, University of Liège, Gembloux 5030, Belgium
3. Biotechnology Research Institute, Chinese Academy of Agricultural Sciences, Beijing 100081, China
4. Institute of Crop Sciences, Chinese Academy of Agricultural Sciences, Beijing 100081, China
5. Department of Chemistry and Chemical Engineering, Qilu Normal University, Jinan 250200, China
6. School of Chemistry and Molecular Engineering, East China University of Science and Technology, Shanghai 200237, China
†These authors contributed equally to this work.
*Correspondence: Ling Jiang (jiangling@caas.cn)
doi: 10.1111/jipb.12811

Research Article

Abstract As essential B vitamin for humans, folates accumulation in edible parts of crops, such as maize kernels, is of great importance for human health. But its breeding is always limited by the prohibitive cost of folate profiling. The molecular breeding is a more executable and efficient way for folate fortification, but is limited by the molecular knowledge of folate regulation. Here we report the genetic mapping of folate quantitative trait loci (QTLs) using a segregated population crossed by two maize lines, one high in folate (GEMS31) and the other low in folate (DAN3130). Two folate QTLs on chromosome 5 were obtained by the combination of F_2 whole-exome sequencing and F_3 kernel-folate profiling. These two QTLs had been confirmed by bulk segregant analysis using F_6 pooled DNA and F_7 kernel-folate profiling, and were overlapped with QTLs identified by another segregated population. These two QTLs contributed 41.6% of phenotypic variation of 5-formyltetrahydrofolate, the most abundant storage form among folate derivatives in dry maize grains, in the GEMS31×DAN3130 population. Their fine mapping and functional analysis will reveal details of folate metabolism, and provide a basis for marker-assisted breeding aimed at the enrichment of folates in maize kernels.

Edited by: Jianbing Yan, Huazhong Agricultural University, China
Received Dec. 21, 2018; **Accepted** Mar. 25, 2019; **Online on** Apr. 2, 2019
FA: Free Access

Free Access

INTRODUCTION

Folates are water-soluble B vitamins that are essential for the growth and development of all species. Folates are coenzymes of many reactions in the synthesis of purine and pyrimidine, amino acid, pantothenic acid, and methionine. In microorganisms and plants, the basic pathways for folate synthesis and one-carbon metabolism have been well studied (Hanson and Gregory 2011; Strobbe and Van Der Straeten 2017). However, humans cannot biosynthesize folates *de novo*, and need a daily intake of more than 400 micrograms of folates per day to satisfy the needs of health. In pregnant women, folate deficiency can lead to growth retardation and neural tube defects in the fetus (Rader and Schneeman 2006; Silver et al. 2015). Moreover, folate deficiency directly leads to megaloblastic anemia and high homocysteine essential hypertension, and increases the risk of cancer, cardiovascular disease, Alzheimer's disease, coronary atherosclerosis and other diseases (Blancquaert et al. 2010; Guo et al. 2017). Folate deficiency has become a global public health problem, especially in China (Bhutta and Salam 2012). In 2016, the 5-year plan for national economic and social development proposed that the folate deficiency rate in pregnant women should be brought to below 5%

in 2020 in China, highlighting the urgency of fighting folate deficiency.

Although humans can acquire folates from fruits and vegetables, the intake of folates from the staple crops is more important to human health, especially in developing countries. In Chinese diets, staple crops (e.g., cereals and tubers) provide the majority of calories, and serve as important sources of necessary macronutrients and key vitamins and minerals (e.g., folates, and calcium) (Chang et al. 2018). However, the edible parts of crops tend to be low in folate, and this is especially true in cereals, such as maize and rice (https://ndb.nal.usda.gov/ndb/nutrients/index). Targeted breeding could be applied to increase folate contents in cereals. However, the prohibitive cost of folate measurement limits the phenotype-based selection, and it is difficult to screen for higher-folate-level varieties in breeding without any molecular genetic information (Dong et al. 2011).

Folates derivatives consist of a pteridine ring, a para-aminobenzoate ring, a tail of one or more L-glutamate residues, and different one-carbon units attached to the N5 and/or N10 positions of tetrahydrofolate (THF) molecules (Gorelova et al. 2017). The pathways for folates synthesis and one-carbon metabolism in different species are highly conserved. In plants, pteridine and para-aminobenzoate are biosynthesized in the cytoplasm and chloroplast, respectively, and then transported to the mitochondrion to form folpolyglutamates (Blancquaert et al. 2010). Based on the pathways of folate metabolism in plants, different strategies have been designed to increase the amount of folates in crops, including enhancing folate and one-carbon metabolism, and increasing the stability of folates by genetic modification (Jiang et al. 2017). For example, to enhance folate biosynthesis, increasing dihydrofolate synthetase or folylpolyglutamate synthetase alone, reducing the activity of gamma glutamyl hydrolase alone, or increasing the transcriptional levels of genes encoding GTP cyclohydrolase I and aminodeoxychorismate synthase individually or together are three successful strategies in *Arabidopsis* (*Arabidopsis thaliana*), maize (*Zea mays*), potato (*Solanum tuberosum*), rice (*Oryza sativa*), and tomato (*Solanum lycopersicum*) (Diaz de la Garza et al. 2004; Diaz de la Garza et al. 2007; Storozhenko et al. 2007; Naqvi et al. 2009; De Lepeleire et al. 2018; Liang et al. 2019). To increase the stability of folates, introducing mammals folate-binding proteins and improving folate polyglutamylation have been shown to enable long-term storage of biofortified high-folate rice grains (Blancquaert et al. 2015). But a better understanding of folate genetic networks that control the folate accumulation in the edible parts of crops is still required, so are new candidate genes for breeding aimed at the enrichment of folates in maize kernels.

5-Formyltetrahydrofolate (5-F-THF) is the main derivative among folates, and 5-methyltetrahydrofolate and THF are also present in maize kernels (Lian et al. 2015). 5-F-THF is a relatively stable natural folate and can be used in dietary supplementation and the treatment of cancer (Suh et al. 2001). In one-carbon metabolism, the conversion between 5-F-THF and 5, 10-methenyltetrahydrofolate needs two different enzymes, and 5, 10-methenyltetrahydrofolate can form 5-methyltetrahydrofolate and other derivatives through a series of enzymatic reactions (Blancquaert et al. 2010). The genes related to folate metabolism in maize share high sequence similarity with those in other organisms, such as *Arabidopsis*, rice and *Escherichia coli*. However, there is no significant correlation between the transcription levels of these genes and the accumulation of folates in grains, indicating that there might be other key genes regulating folate metabolism in maize (Lian et al. 2015).

Quantitative trait loci (QTL) mapping is an efficient way to identify underlying genes and elements for phenotype variation, but it is rarely conducted in folates in maize. Due to the complexity of plant genomes and low density of markers, exploring genetic base of quantitative traits faces huge challenges (Yu et al. 2011). QTL mapping depends on the marker density and population size, and higher genetic marker density and larger population size can improve the precision of QTL mapping (Chen et al. 2014a). Relatively small population size of advanced populations such as recombination inbred lines (RIL) can reduce the cost of genotyping and phenotyping but makes for time consuming and expensive advanced population construction. Conversely, early generation crosses such as F_2s and backcross containing abundant recombinant, combined with high through-put genotyping method are powerful to detect QTLs (Vales et al. 2005; Chen et al. 2014a). Gene pyramiding by marker-assisted selection is an effective means of fortifying vitamins, such as vitamin A and vitamin E (Jiang et al. 2017), and the lack of functional-loci identification and prohibitive cost for

folate profiling limits the molecular breeding of folates in maize to a large extent.

In this study, we used a segregated population that was constructed from two maize inbred lines, one high in folate (GEMS31) and the other low in folate (DAN3130). Two QTLs on chromosome 5 were obtained by the combination of F_2 whole-exome sequencing and F_3 kernel-folate profiling. The phenotypic variation contribution rates were 26.7% (*q5-F-THFa*) and 14.9% (*q5-F-THFb*), respectively. These two QTLs had been confirmed by bulk segregant analysis (BSA) using F_6 pooled DNA and F_7 kernel-folate profiling and QTL verification using another segregated population (K22×DAN340, Xu et al. 2012). BSA obtained several regions overlapped with *q5-F-THFa* and *q5-F-THFb*. And among QTLs obtained from the latter population, *q5-F-THFd* and *q5-F-THFe* could explain 10.6% and 7.1% of the phenotypic variation, and were also overlapped with *q5-F-THFa* and *q5-F-THFb*, respectively. Thus, we successfully obtained two QTLs (*q5-F-THFa* and *q5-F-THFb*) that control 5-F-THF variation in maize grains.

RESULTS

Development of a folate segregated population from phenotype-extreme-differing parents

We investigated folate levels in maize kernels using a global germplasm collection (Yang et al. 2011). 5-F-THF accounted for over 70.0% of the total folates in 70.5% of the inbred lines, which confirms that 5-F-THF is the main derivative among kernel folates (Table 1; Lian et al. 2015). Further phenotyping indicated that, irrespective of significant variations across the environments, GEMS31 contained more 5-F-THF than DAN3130, and the 5-F-THF level in GEMS31 was always around 20 times that in DAN3130 (Table 2). Therefore, we used these two inbred lines to construct a segregated population by single seed descent beginning in 2014, with the seeds of F_7 harvested in 2017.

During the development of this segregated population, genetic mapping using F_2 whole-exome sequencing and F_3 kernel folate profiling was carried out to obtain QTLs. Later, BSA using F_6 bulked segregant sequencing and F_7 kernel-folate profiling was performed to confirm the location of QTLs.

Table 1. The distribution of the percentage of 5-F-THF in total folates in a maize global germplasm collection and percentages of inbred lines corresponding to the former in 2013, $n = 359$

The percentage of 5-F-THF in total folates (%)	The percentage of inbred lines (%)
30.0–39.9	0.8
40.0–49.9	2.5
50.0–59.9	10.0
60.0–69.9	16.2
70.0–79.9	23.1
80.0–89.9	34.3
90.0–99.9	13.1

Bin map construction using whole-exome sequencing data

In all, 216 F_2 plants were sequenced using the exon-capture sequencing strategy. For single nucleotide polymorphism (SNP) detection, the sequences of each plant were aligned to the B73 reference genome Version 2 (only to the SNPs positions retained from the two parents). The called SNP were filtered by the criteria of MAF > 0.1 and missing rate < 0.9. After the filtering, a total of 114,183 SNPs was selected for the following analysis. A sliding window method that calculates the ratio between the numbers of SNPs from the three genotypes (GEMS31, DAN3130 and Heterozygous) was used for genotype calling (Huang et al. 2009). The uninterrupted segments across the population were regarded as bins used for recombination map construction (see Methods). The resulting 3,780 recombination bins captured the majority of the recombination events detected in the population (Figure 1A). The average physical length of the recombination bins was 548 kb, ranging from 479 Kb to 726 Kb across chromosomes (Table 3). A linkage map was constructed with these bins, which served as markers. The map had a total genetic distance of 1482

Table 2. 5-F-THF profiles in GEMS31 and DAN3130 inbred lines across the environments

		5-F-THF (nmol/g DW)	
Year	Location	GEMS31	DAN3130
2009	Hainan, China	17.83 ± 0.56	1.05 ± 0.08
2010	Yunnan, China	8.01 ± 0.43	0.37 ± 0.02
2013	Beijing, China	5.27 ± 0.32	0.26 ± 0.01

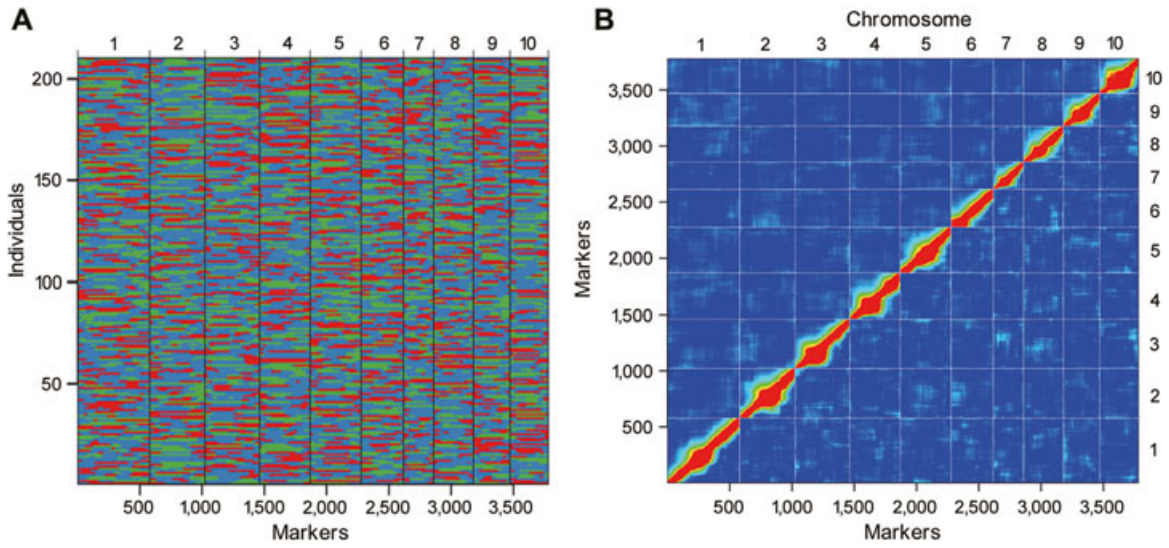


Figure 1. Recombination bin-map and heatmap for all markers of F_2 whole-exome sequencing from the GEMS31×DAN3130 population

(A) Bin-map consists of 3,780 bin markers inferring 114,183 high quality SNPs in F_2 population. Physical position is based on B73 RefGen V2 sequence. Red: GEMS31 genotype; Green: DAN3130 genotype; Blue: heterozygote. **(B)** Pairwise recombination fractions are shown (upper left) as well as logarithm (base 10) of odds (LOD) which score for tests of linkage (bottom right) for all markers. Red corresponds to low recombination or a high LOD score, while blue represents unlinked markers.

cM with an average interval of 0.729 cM between the bins (Table 3). There were 12 bins that were >10 Mb, and 36 bins that were >5 Mb but <10 Mb. All of the >10 Mb bins and 22 of the 5-10 Mb bins were within 20 Mb of centromeres. Pairwise recombinational fraction analysis revealed that the linkage signals were diagonally displayed, validating the quality of the bin map (Figure 1B).

To evaluate the power and accuracy of this genetic map for QTL mapping, QTL analysis of cob color was first performed in this population. There were 53 white cobs and 163 red cobs in the F_2 population. The ratio of

Table 3. The information of the linkage map obtained from F_2 whole-exome sequencing from the GEMS31×DAN3130 population

	Bins			Linkage map (cM)				
Chr.	Number	Mean (Mb)	Max (Mb)	Total	Mean	Max	Recombination (cM/Mb)	NO. of crossover
1	579	0.519	19.872	201.57	0.349	1.44	0.671	4.1
2	443	0.534	14.771	160.49	0.363	2.909	0.68	3.3
3	439	0.523	14.98	196.95	0.449	8.623	0.858	4
4	406	0.593	16.547	151.88	0.375	1.924	0.632	3.1
5	409	0.531	27.083	141.54	0.347	2.41	0.653	2.9
6	340	0.494	8.611	126.91	0.374	2.412	0.757	2.6
7	242	0.726	22.692	119.04	0.489	1.925	0.674	2.4
8	320	0.546	20.041	139.12	0.436	2.659	0.797	2.8
9	292	0.53	14.223	115.84	0.396	3.068	0.747	2.4
10	310	0.479	5.85	128.78	0.415	2.414	0.865	2.7
Total	3,780	0.548	27.083	1482.12	0.399	8.623	0.729	3

white: red fitted the ratio of 1:3 ($P = 0.875$, χ^2 test), indicating that there might be only one Mendelian locus responsible for the trait variation. QTL mapping yielded only one locus, which was at Bin1_171-Bin1_175, with a physical position of 47.325–51.286 Mb on chromosome 1. The peak bin, Bin1_172 (Chr1: 47.81–49.22 Mb) just encompassed the cloned gene *pericarp color* 1 (*P1*) at position 48.1Mb (Figure 2). *P1* has been reported to regulate red pigmentation in cob, pericarp, tassel glumes, and husks (Robbins et al. 2009). Exact mapping of this locus for cob color demonstrated the quality of our bin map.

QTL analysis for 5-F-THF variation

GEMS31, DAN3130 and 216 F_2 plants of the GEMS31×DAN3130 population were grown in Langfang, Hebei, China, in 2015. 5-F-THF contents of GEMS31 and DAN3130 were 4.61 ± 0.21, and 0.21 ± 0.02 nmol/g DW, respectively. 5-F-THF contents of the F_2 population ranged from 0.08 ± 0.01, to 4.88 ± 0.32 nmol/g DW, and the average level was 1.51 nmol/g DW. The phenotypes in the F_2 population showed a continuous distribution, indicating the quantitative characteristics of this trait (Figure 3).

Next, we used our bin map to identify QTLs controlling kernel 5-F-THF variation. Our analysis of the 216 F_2 plants detected two QTLs (*q5-F-THFa* and *q5-F-THFb*), with likelihood of odds (LOD) peaks overlapping with Bin5_17 and Bin5_128 on chromosome 5, which explained 26.7% and 14.9% of the phenotypic variation, respectively (Figure 4A). *q5-F-THFa* occupied a physical position of 1.532–2.672 Mb, and *q5-F-THFb* was at 16.638–22.402 Mb on chromosome 5, respectively (Figure 4B). The positive contributors to these two QTLs were both derived from the inbred line GEMS31.

Validation of folate QTLs

We had used two methods to validate the above two folate QTLs: BSA using the GEMS31×DAN3130 RIL population and QTL verification using another segregated population K22×DAN340. First, GEMS31, Dan3130 and 216 F_6 plants of the GEMS31×DAN3130 population were additionally grown in Langfang, Hebei, China, in 2017. The 5-F-THF contents of GEMS31 and DAN3130 were 5.29 ± 0.45, and 0.70 ± 0.01 nmol/g DW, respectively. The 5-F-THF contents of the GEMS31×DAN3130 F_7 seeds ranged from 0.57 ± 0.05 to 8.01 ± 0.51 nmol/g DW, and the average level was 1.57 nmol/g DW. For bulk sergeant analysis, 30 DNA samples from 15 lines (two siblings for each line) of F_6 plants that exhibited either extreme-high or extreme-low 5-F-THF in the F_7 seeds were pooled to form the extreme-high and extreme-low groups, respectively (Table 4). Both pools were sequenced and analyzed via BSA (Takagi et al. 2013), identifying several regions at the physical position from 1.40 to 21.20 Mb on chromosome 5 (Figure 5). *q5-F-THFa* (1.532–2.672 Mb) was found to be overlapped with one region (1.40-2.65 Mb); *q5-F-THFb* (16.638–22.402 Mb) was found to be overlapped with five regions (16.55-17.10 Mb, 17.75-17.85 Mb, 18.55–19.50 Mb, 19.80–20.55 Mb, 20.60–21.20 Mb, Figure 5B). Second, we used another segregated population K22×DAN340

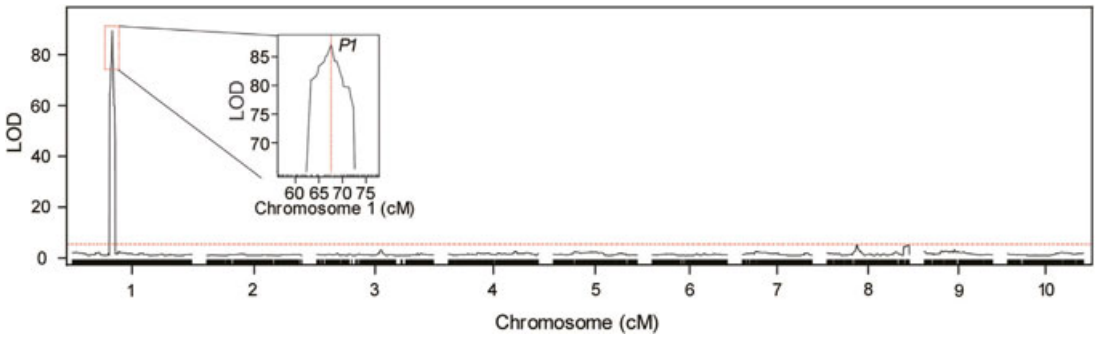


Figure 2. Mapping of a QTL controlling cob color in F_2 population and the location of *P1*
Curve in plot indicate the genetic coordinate (X-axis) and LOD score (Y-score) of detected QTL, and the horizontal dashed line presents the LOD threshold. The peak encompasses the cloned gene *pericarp color* 1 (*P1*) at position 48.1 Mb on chromosome 1. The box inside is the zoom-in image of the peak on chromosome 1. The vertical red dashed line presents the peak position of *P1* gene.

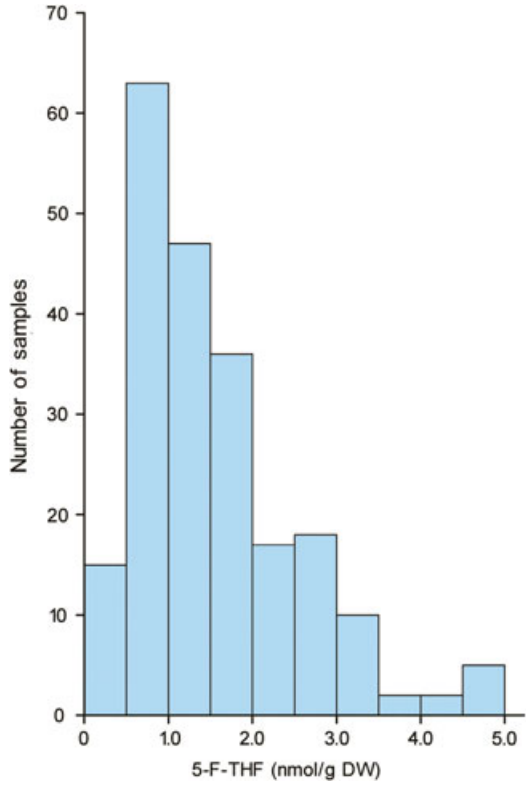
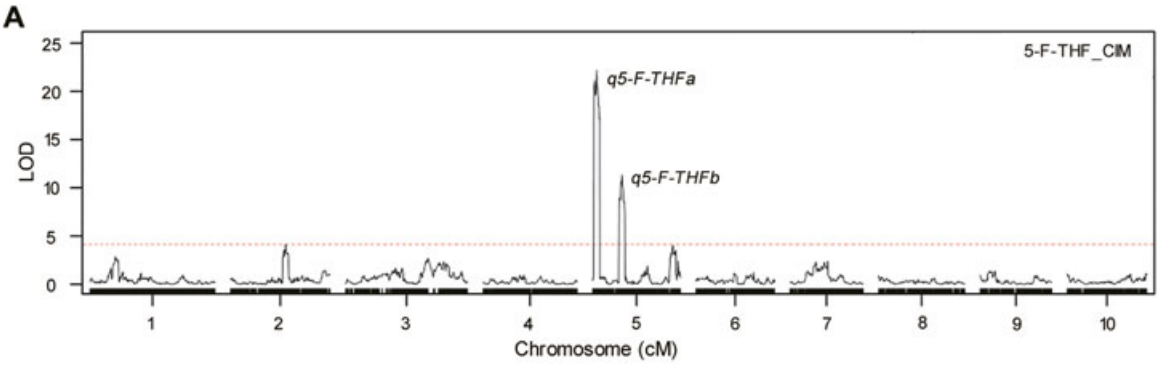


Figure 3. Variation of 5-F-THF levels in F_3 kernels of the F_2 population
The number of samples is shown as gray columns in the variation of 5-F-THF levels. 5-F-THF levels ranged from 0.08 ± 0.01, to 4.88 ± 0.32 nmol/g DW, and the average level was 1.51 nmol/g DW.

originally constructed for tocopherols QTL analysis to verify the folate QTLs (Xu et al. 2012). The 5-F-THF levels have been measured in K22, DAN340 and 188 F_3 kernels of the K22×DAN340 population. 5-F-THF level in DAN340 kernels was about 2.1 times of that in K22 kernels (2.85 ± 0.26 nmol/g DW vs 1.35 ± 0.11 nmol/g DW). 5-F-THF content displays abundant phenotypic variation ranging from 0.62 ± 0.05 to 2.89 ± 0.19 nmol/g DW in the population. QTL mapping results showed that in total four QTLs detected, three are on chromosome 5 in this population (*q5-F-THFd* to *q5-F-THFf*) (Figure 6A). Among them, there are two main genetic loci (*q5-F-THFd*, 0.000-6.860 Mb; *q5-F-THFe*, 18.349–58.604 Mb) that have overlapped with the localization of *q5-F-THFa* and *q5-F-THFb* obtained from the GEMS31×DAN3130 population (Figures 4B, 6B). *q5-F-THFd* and *q5-F-THFe* could explain 10.6% and 7.1% of the phenotypic variation in the K22×DAN340 population, respectively. These results thus confirmed that we obtained two folate QTLs (*q5-F-THFa* and *q5-F-THFb*) related to 5-F-THF variations in maize kernels.

Candidate gene prediction

According to maize gene annotation database accessible at MaizeGDB (www.maizegdb.org), there are 88 protein-coding genes and 142 protein-coding genes in *q5-F-THFa* and *q5-F-THFb*, respectively. We have compared these protein-coding genes with genes expressed in seeds, and 68 and 102 genes have been obtained in *q5-F-THFa* and *q5-F-THFb*, respectively (Chen et al. 2014b). Then we filtered these genes with tissue-specific expressions by using high density NimbleGen Microarrays (Winter et al. 2007; Sekhon et al. 2011) and RNA Sequencing (Stelpflug et al. 2015). Twenty-two and 32 genes have relatively high transcript levels during milking stage (DAP 18 to DAP 24) in embryo, endosperm or whole seeds in *q5-F-THFa* and *q5-F-THFb*, respectively (Table 5). Among these genes after filtration, there are 22 genes coding functional unclassified proteins, 13 genes involved in protein synthesis, modification and degradation, seven genes involved in RNA regulation. Based on the annotated function involved in metabolic pathway, two genes might be involved in folate-related metabolism. GRMZM2G121840, which located in the region of *q5-F-THFa*, encodes a S-adenosyl-L-methionine-dependent methyltransferase. GRMZM2G124863, which located in the region of *q5-F-THFb*, encodes a transferase which contains a folic acid binding domain. The different SNPs in both GEMS31 X DAN3130 population and K22 X DAN340 population were screened by checking with the RNA-seq data set from kernels in the public maize genome sequence (Li et al. 2013), also confirmed in F_2 whole-exome sequencing data in GEMS31 X DAN3130 population. Six SNPs in GRMZM2G121840 and four SNPs in GRMZM2G124863 were obtained respectively. In these SNPs, some could lead to the amino acid substitution. For example, G to C transition on Chr5.s_19676846 displayed a substitution from alanine to proline (Figure 7). Haplotype analysis in GEMS31 X DAN3130 population has been conducted using these SNPs. In GRMZM2G121840, the results showed that 5-F-THF levels with six female genotypes (GEMS31) were significantly higher than those with six male genotypes (DAN3130, 2.34 ± 1.14 nmol/g DW vs 0.96 ± 0.14 nmol/g DW; Figure 7A). In GRMZM2G124863, 5-F-THF levels with four female genotypes



B

QTL	Chr	Peak_Bin	Peak_Lod	Peak_Pos_Phy (Mb)	QTL_int		Var%	Add	Dominance	Positive_Allele
					Left_Phy	Right_Phy				
q5-F-THF5a	5	Bin5_17	22.187	2.297	1.532	2.672	26.723	0.616	-0.527	GEMS31
q5-F-THF5b	5	Bin5_128	11.340	20.139	16.638	22.402	14.942	0.544	0.154	GEMS31

Figure 4. Mapping of QTLs for 5-F-THF variation using the GEMS31×DAN3130 population
(A) Location of 5-F-THF QTLs. Peaks of QTLs are shown as *q5-F-THFa* and *q5-F-THFb* on chromosome 5, respectively. Curve in plot indicate the genetic coordinate (X-axis) and LOD score (Y-score) of detected QTL, and the horizontal dashed line presents the LOD threshold. **(B)** Details of 5-F-THF QTLs. Chr, chromosome; Var, phenotypic variation contribution rate; Add, additive effect.

(GEMS31) were significantly higher than those with four male genotypes (DAN3130, 2.05 ± 1.28 nmol/g DW vs 0.75 ± 0.14 nmol/g DW; Figure 7B). These results indicated that these two candidate genes could regulate folate levels in maize seeds, and also confirmed that the positive contributors to these two genes were both derived from the inbred line GEMS31. More biochemical and genetic work is needed to dissect the relationship between these genes and folate metabolism.

DISCUSSION

The insistent demands for genetic mapping of new folate-candidate regions in maize

Increasing folate intake is an efficient way to battle folate deficiency, and the current international practice is to take folic acid tablets and add folic acid to flour in developed countries (Bhutta and Salam 2012). However, the latest research shows that excessive intake of synthetic folic acid can lead to changes in the levels and patterns of DNA methylation in human bodies, accelerate the occurrence and development of some kinds of cancers, and also mask the deficiency of vitamin B12 (Cole et al. 2007; Bae et al. 2014; Moore et al. 2014). More importantly, some populations carry a homozygous mutation in the gene for methylenehydrofolate reductase (the 677TT genotype); in those cases, even supplying folic acid tablets does not effectively control the homocysteine level (Waskiewicz et al. 2011). Thus, intake of bioactive folates from edible parts of crops is healthier and more important.

In maize, there is rarely a correlation between folate content and expression of conserved genes of folate biosynthesis and metabolism in kernels (Lian et al. 2015). Even though the increase in folate levels was observed in transgenic maize (Naqvi et al. 2009; Liang et al. 2019), the understanding of folate genetic networks in grains remains unclear. Furthermore, the measurement of folates is unacceptably expensive for breeders, and different from most other agronomic traits, the breeding-based phenotype selection is untenable for folate breeding. Thus, genetic mapping of new folate-candidate regions is imperative. These regions could be new genetic regulators, and serve a key role in regulating the accumulation of folates in grains, and the development of corresponding molecular makers will be an efficient way to promote efforts for breeding higher-folate-level varieties of maize.

Table 4. 5-F-THF profiles of the extreme-high and extreme-low groups from F_7 seeds of the GEMS31×DAN3130 population

High 5-F-THF		Low 5-F-THF	
Sample name	5-F-THF (nmol/g DW)	Sample name	5-F-THF (nmol/g DW)
A180-1	4.40 ± 0.35	E116-2	0.57 ± 0.05
A180-2	4.50 ± 0.40	E116-1	0.58 ± 0.04
A111-1	4.66 ± 0.38	E139-2	0.59 ± 0.05
A111-2	4.67 ± 0.23	E139-20	0.60 ± 0.02
B101-1	4.67 ± 0.34	C138-2	0.69 ± 0.03
A129-1	4.68 ± 0.28	B136-2	0.72 ± 0.06
D171-3	4.72 ± 0.44	B134-3	0.79 ± 0.04
B101-3	4.81 ± 0.42	B136-3	0.83 ± 0.02
A136-2	4.82 ± 0.37	B140-2	0.83 ± 0.07
B130-4	4.90 ± 0.32	C136-3	0.88 ± 0.05
C163-1	5.03 ± 0.47	B127-5	0.88 ± 0.08
B130-3	5.14 ± 0.23	B141-4	0.90 ± 0.05
A179-1	5.15 ± 0.35	C133-2	0.91 ± 0.06
D180-2	5.16 ± 0.46	C133-3	1.01 ± 0.05
B103-2	5.16 ± 0.16	C139-1	1.02 ± 0.07
B103-4	5.23 ± 0.34	C138-1	1.06 ± 0.03
C163-2	5.24 ± 0.27	B141-3	1.10 ± 0.06
A136-3	5.45 ± 0.51	E170-3	1.10 ± 0.08
A171-4	5.53 ± 0.26	C137-3	1.11 ± 0.03
A171-2	5.62 ± 0.29	B127-1	1.13 ± 0.06
F121-1	5.74 ± 0.41	B140-3	1.16 ± 0.08
D171-4	5.84 ± 0.36	E170-1	1.19 ± 0.07
F121-2	5.87 ± 0.31	C135-2	1.22 ± 0.08
E103-1	6.04 ± 0.57	C135-1	1.26 ± 0.12
A179-2	6.12 ± 0.15	B134-5	1.26 ± 0.05
D180-3	6.22 ± 0.56	B179-3	1.30 ± 0.07
E103-3	6.25 ± 0.21	C136-1	1.58 ± 0.06
A125-3	6.49 ± 0.33	C139-3	1.61 ± 0.04
A129-2	7.13 ± 0.52	B179-5	1.61 ± 0.09
A125-1	8.01 ± 0.51	C137-1	1.68 ± 0.11

Rapidly mapping QTLs using a segregated population derived from phenotype-extreme-differing parents

5-F-THF is the most dominant derivative in mature maize grains (Table 1). In this study, a folate segregated population developed from a cross of maize inbred lines with extreme-high and extreme-low levels of 5-F-THF was used to dissect the genetic architecture of folate accumulation in maize grains. This population consisted of 216 F_2 plants and was genotyped by whole-exome sequencing (Lu et al. 2018). Two folate QTLs were obtained and validated by BSA using the GEMS31×DAN3130 RIL population and QTL verification using another segregated population K22×DAN340 (Figures 4–6).

The mapping resolution using high density markers produced by whole-exome sequencing for folate QTLs is higher than in previous findings (Figure 4; Price 2006; Chen et al. 2014a). Previously, the mapping resolution was 2 cM or less when using QTL mapping to position genes with a population size of 100–200 lines and hundreds of markers (Price 2006); 0.8–56.6 Mb for tassel and ear architecture QTLs were also obtained from a large early generation population (708 F_2 progeny) combined with sequencing-based genotyping (Chen et al. 2014a). In our case, a similar strategy, but with a much smaller population size (216 F_2 progeny) was used here and this analysis achieved higher resolution than previous methods, with 1.1 and 5.8 Mb for each folate QTLs interval, respectively (Figure 4).

Normally the aim of QTL analysis with replication in different environments was to check the repeatability of the phenotype and to acquire more reliable results. Validations are usually needed when QTL analysis is using an F_2 population without replication (Vales et al. 2005; Chen et al. 2014a). In our case, the reliability of mapping loci had been validated by BSA using the GEMS31×DAN3130 RIL population and QTL verification using another segregated population K22×DAN340. Both regions from BSA and *q5-F-THFd* and *q5-F-THFe* from another QTL analysis were overlapped with *q5-F-THFa* and *q5-F-THFb*, respectively (Figure 5B and Figure 6B). Thus, we inferred that combining a small population derived from parents having extremely different phenotypes with genotyping by sequencing strategy would be another efficient method for rapid QTL identification, while reducing costs through the number of individuals subjected to genotyping and phenotyping.

Two QTLs for folates accumulation in maize kernels

Even though folates have the characteristics of quantitative traits (Figure 3), the genetic regulation of folate accumulation in kernels is much different from that of traditional flowering time, morphological and stress-related traits. Those traits have a highly polygenic genetic architecture that is mainly controlled by many small effect QTLs (Mickelson et al. 2002;

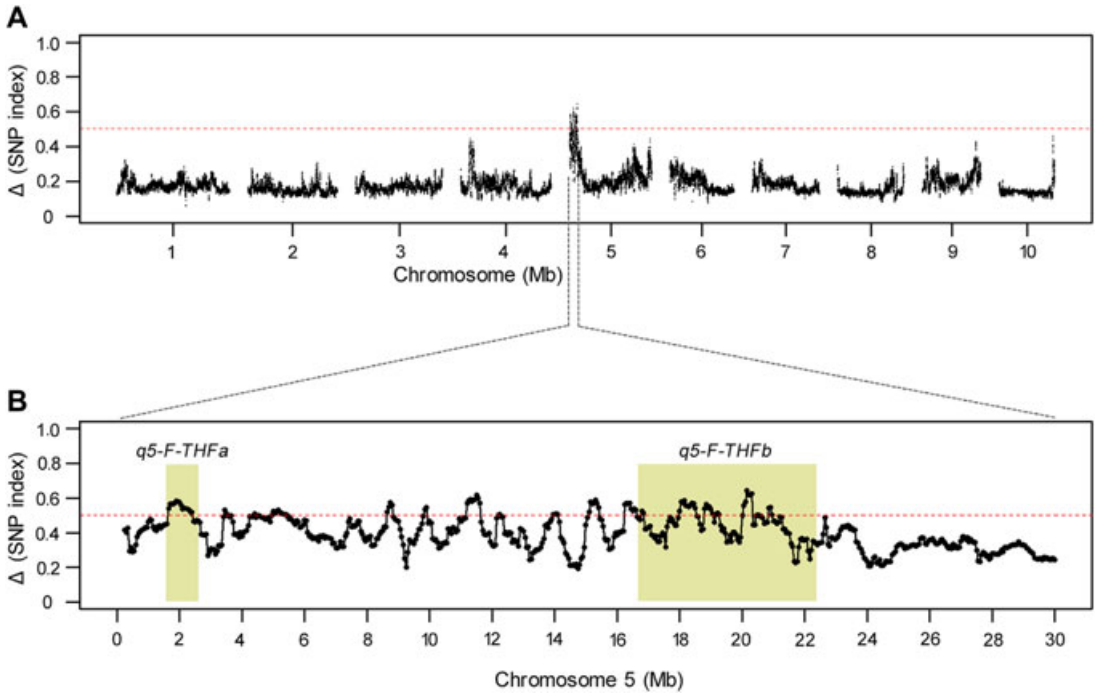
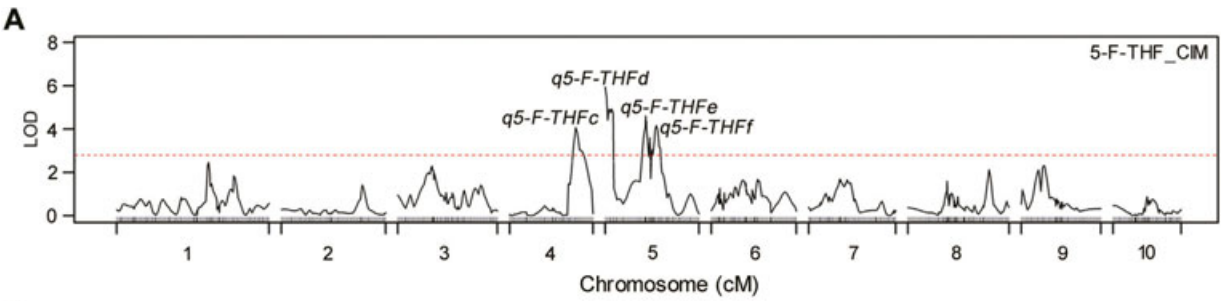


Figure 5. BSA results for the GEMS31×DAN3130 RIL population

(A) BSA regions at the physical position from 1.40 to 21.20 Mb on chromosome 5. Curve in plot indicate the physical coordinate (X-axis) and ΔSNP index (Y-score) of detected regions, and the horizontal dashed line presents the ΔSNP index threshold. **(B)** The zoom-in image of the above regions on chromosome 5. Yellow columns indicate *q5-F-THFa* and *q5-F-THFb* from QTL analysis.



B

QTL	Chr	Peak_LOD	Peak_Pos_Phy (Mb)	QTL_int (Mb)		Var%	Add	Dominance
				Left_Phy	Right_Phy			
q5-F-THFc	4	4.07	235.44	233.628	239.510	7.027	-0.155	-0.138
q5-F-THFd	5	6.01	3.539	0.000	6.860	10.624	-0.228	-0.089
q5-F-THFe	5	4.61	35.277	18.349	58.604	7.099	-0.062	-0.270
q5-F-THFf	5	4.16	91.814	66.432	188.609	6.335	-0.042	-0.279

Figure 6. Mapping of QTLs for 5-F-THF variation using the K22×DAN340 population

(A) Location of 5-F-THF QTLs. Peaks of QTLs are shown as *q5-F-THFc* to *q5-F-THFf* on chromosome 5, respectively. Curve in plot indicate the genetic coordinate (X-axis) and LOD score (Y-score) of detected QTL, and the horizontal dashed line presents the LOD threshold. **(B)** Details of 5-F-THF QTLs. Add, additive effect; Chr, chromosome; Var, phenotypic variation contribution rate.

Table 5. Filtered candidate genes in *q5-F-THFa* and *q5-F-THFb*

Position (Chr 5)	Gene ID	Annotation
1565340..1568691	GRMZM2G125300	Ribosomal protein S21e
1603428..1605321	GRMZM2G010861	Acyl-CoA binding protein 4
1608896..1612025	GRMZM2G326195	Expressed protein
1621338..1624924	GRMZM2G326272	Translocon at the outer envelope membrane of chloroplasts 159
1682058..1686853	GRMZM2G004140	Expressed protein
1822186..1829125	GRMZM2G138523	Expressed protein
1829948..1838934	GRMZM2G138511	Chaperone DnaJ-domain superfamily protein
1846096..1853085	GRMZM2G138421	Squamosa promoter-binding protein-like 12
1898216..1902113	GRMZM2G100419	Expressed protein
2116106..2121486	GRMZM2G121826	Protein kinase protein with tetratricopeptide repeat domain
2126865..2129228	GRMZM2G121840	S-adenosyl-L-methionine-dependent methyltransferase
2181763..2184186	GRMZM2G085301	Major facilitator superfamily protein
2259767..2263658	GRMZM2G144008	GTPase activity
2265501..2268450	GRMZM2G144020	Expressed protein
2275248..2277812	GRMZM2G445169	Barwin-like endoglucanases superfamily protein
2284054..2290332	GRMZM2G144042	Protein kinase superfamily protein
2350546..2357084	GRMZM2G003108	SEC14 cytosolic factor / phosphoglyceride transfer family protein
2358646..2361757	GRMZM2G003068	Zinc knuckle (CCHC-type) family protein
2362606..2365060	GRMZM2G002825	Actin depolymerizing factor 4
2387257..2400960	GRMZM2G002765	NAD(P)-binding Rossmann-fold superfamily protein
2524531..2535342	GRMZM2G458266	Transducin/WD40 repeat-like superfamily protein
2536464..2538475	GRMZM2G160046	Expressed protein
16701221..16703382	GRMZM5G895064	DNA-directed RNA polymerase family protein
17077279..17083849	GRMZM2G018943	NagB/RpiA/CoA transferase-like superfamily protein
17123432..17138841	GRMZM2G095239	Telomere repeat binding factor 1
17290718..17296942	GRMZM2G417223	RNA-binding KH domain-containing protein
17340297..17343212	GRMZM2G023347	ABI3-interacting protein 3
17881123..17884825	GRMZM2G046676	SNF7 family protein
18192142..18193721	GRMZM2G060940	2-Oxoglutarate and Fe(II)-dependent oxygenase superfamily protein
18358282..18370223	GRMZM2G067306	Ribosomal protein L25/Gln-tRNA synthetase anti-codon-binding domain
18469442..18472522	GRMZM2G066867	Sucrose nonfermenting 1(SNF1)-related protein kinase
18473592..18476671	GRMZM2G066996	Serine protease inhibitor (SERPIN) family protein
18650274..18656914	GRMZM2G044900	Seed maturation protein PM23
18789191..18791839	GRMZM2G087326	Reversibly glycosylated polypeptide 2
19014015..19016625	GRMZM2G045192	No annotation
19069664..19077156	GRMZM2G375984	DEK domain-containing chromatin associated protein
19145071..19148243	GRMZM2G074898	Ribosomal protein L24e family protein
19207810..19212507	GRMZM2G081221	SMAD/FHA domain-containing protein
19263506..19269985	GRMZM2G403149	GTP-binding family protein
19269959..19277452	GRMZM2G403218	FRIGIDA-like protein
19460235..19471332	GRMZM2G386714	Lysyl-tRNA synthetase 1
19478193..19481880	GRMZM2G086904	RNA polymerase Rpb6

(Continued)

Table 5. Continued

Position (Chr 5)	Gene ID	Annotation
19696664..19698973	GRMZM2G124863	Transferases, folic acid binding
19939913..19945652	GRMZM2G001631	Dicer-like 2
20784279..20795347	GRMZM2G002874	RNA-binding protein 45A
21005810..21011591	GRMZM2G024054	Ubiquitin-protein ligase 2
21337753..21339972	GRMZM5G871425	Embryo defective 2016
21366640..21370672	GRMZM5G823318	Aspartate kinase 1
21459531..21465540	GRMZM2G413544	Protein kinase superfamily protein
21592571..21599690	GRMZM2G073934	Tetratricopeptide repeat (TPR)-like superfamily protein
21601089..21603875	GRMZM2G074169	Cytochrome B5 isoform E
21719063..21724696	GRMZM2G153815	Mitochondrial HSO70 2, DnaK family protein
21903403..21908619	AC209835.4_FG004	ChaC-like family protein
21954101..21955171	GRMZM2G096372	Expressed protein

Peiffer et al. 2014; Raihan et al. 2016; Xu et al. 2017). Parental phenotypic and genetic distances are positively associated with the effects and numbers of QTLs yielded (Pan et al. 2017), and major-effect QTLs can be more easily mapped to a higher resolution (Wang et al. 2018). In our case, the phenotypic variation contribution rates were 26.7% (*q5-F-THFa*) and 14.9% (*q5-F-THFb*), respectively, reflecting that both are major-effect QTLs. This feature makes them well suited for marker-assisted selection, which is mainly used in maize breeding for

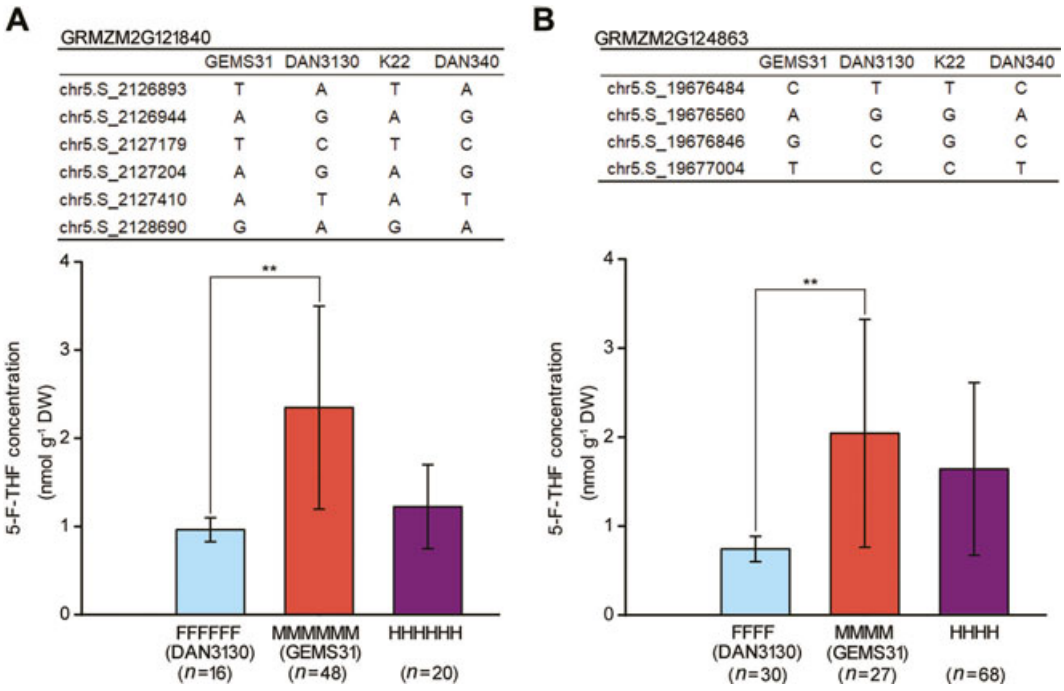


Figure 7. The segregation status of selected SNPs in GEMS31 X DAN3130 population

(A) The segregation status of 6 SNPs in GRMZM2G121840 in GEMS31 X DAN3130 population and 5-F-THF levels comparison between two types of haplotypes from 6 SNPs. **(B)** The segregation status of four SNPs in GRMZM2G124863 in GEMS31 X DAN3130 population and 5-F-THF levels comparison between two types of haplotypes from four SNPs. F, female; M, male; H, heterozygosis. N, number of haplotypes. **, $P < 0.01$.

qualitative-trait genes or major-effect QTLs (Pixley et al. 2010).

Furthermore, we have filtered protein-coding genes in QTLs using tissue-specific expressions data (Winter et al. 2007; Sekhon et al. 2011; Chen et al. 2014b; Stelpflug et al. 2015). Twenty-two and 32 genes have relatively high transcript levels during milking stage (DAP 18 to DAP 24) in embryo, endosperm or whole seeds in *q5-F-THFa* and *q5-F-THFb*, respectively (Table 5). Based on the annotated function involved in metabolic pathway, two genes might be involved in folate-related metabolism. In *q5-F-THFa*, GRMZM2G121840 encodes a methyltransferase using S-adenosyl-L-methionine as coenzyme. Folate is critical for cell division, metabolism of proteins, synthesis of purine and pyrimidine, and increasing the *de novo* delivery of methyl groups and S-adenosyl-L-methionine. S-adenosyl-L-methionine obtains a methyl group from 5-methyltetrahydrofolate during methyl cycle (Abbasi et al. 2018). In *q5-F-THFb*, GRMZM2G124863 encodes a transferase protein containing folic acid binding domain. Some proteins involved in folate metabolism can bind folates, such as serine hydroxymethyltransferase (Besson et al. 1993), and methionine synthetase (Ferrer et al. 2004). Additionally, the segregation status analysis of selected SNP variations in GEMS31 X DAN3130 population also indicated that 5-F-THF levels with female genotypes (GEMS31) were significantly higher than those with male genotypes (DAN3130, Figure 7). Thus, these two genes might be candidate genes involved in folate-related metabolism. Fine mapping of these two QTLs and functional analysis of candidate genes will reveal how 5-F-THF accumulates in maize grains, and provide a basis for marker-assisted breeding aimed at the enrichment of folates in maize kernels.

MATERIALS AND METHODS

Plant material and the development of a folate segregated population

The set of dry seeds powder ($n = 359$) for the folate profiling from the global germplasm collection were kindly provided by Dr. Jianbing Yan in 2009, 2010, and 2013, respectively. The set of dry seeds powder for the folate profiling from K22, DAN340 and 188 F_3 kernels of segregated population K22×DAN340 were kindly provided by Dr. Jianbing Yan in 2009. DAN3130 is originated from China, belonging to the NSS subpopulation with pedigree being American hybrid P78599; GEMS31 is from the United States, belonging to the TST subpopulation with pedigree being 2282-01_XL380_S11_F2S4_9226-Blk26/00 (Yang et al. 2011). Both inbred lines were grown at Langfang, Hebei, China in the summer of 2014. The experimental field was loamy soil with pH 6.8, organic matter 0.7%, phosphorus 13.8 mg/L, and potassium 48 mg/kg. During field preparation, 440 kg/acre of urea (46-0-0) were applied. The herbicides were applied 5 d after planting. Plants were hand planted in 5-m-long rows with row and plant spacing of 60 and 25 cm, respectively. The segregated population was developed by single seed descent from the cross (Murigneux et al. 1993). F_1 seeds were obtained, and later grew at Ledong, Hainan, China in the winter of 2014. F_2 plants were grown at Langfang, Hebei, China in the summer of 2015 under the same conditions. All the DNA of 21-d-old leaves have been collected and frozen in −80°C for the whole-exome sequencing. And F_3 kernels have been collected and used for folate analysis. F_3 plants have been grown at Ledong, Hainan, China in the winter of 2015; F_4 plants have been grown at Langfang, Hebei, China in the summer of 2016. F_5 plants have been grown at Ledong, Hainan, China in the winter of 2016. F_6 plants have been grown at Langfang, Hebei, China in the summer of 2017. All the DNA of 21-d-old leaves have been collected and frozen in −80°C for the BSA analysis. And F_7 kernels have been collected and used for folate analysis.

Whole-exome sequencing

Genomic DNA was isolated from young leaves of 216 F_2 plants using the Plant Genomic DNA Extraction Kit (Karroten). DNA was quantified using Nanodrop 2000 (Thermo Scientific). Then 4.5 μg of genomic DNA (35 ng/μL) was fragmented into 200–300 bp using Bioruptor UCD-200 (Diagenode). Libraries were constructed using NEBNext Ultra DNA Library Prep Kit for Illumina (New England Biolabs), the exomes were captured using the Capture Beads from the SeqCap EZ Pure Capture Bead Kit, washed, and amplified by Ligation-Mediated PCR, and were sequenced by Hiseq Xten (Illumina) to generate 150-bp paired-end reads and 8-bp indexed reads (Lu et al. 2018).

Folate analysis

The samples of dry seeds collected from field were used for identification of folate profiles. The methods for

sample preparation and metabolite measurement were described as previous study (Lian et al. 2015). Folates in seeds were measured in four biological replicates, and each sample consisted of 50 mg of plant material. The average content of folate was used as the phenotypic data.

SNP calling

After being trimmed, the sequencing data was mapped to the reference genome using BWA software (Li and Durbin 2009), we used the default parameter and pair-end mem model. The software Samtools was used to genotype every single nucleotide polymorphism of each samples within the background of polymorphism (191,223) between the parental samples (Li et al. 2013). Then we obtain the genotype of each F2 sample in the reference of parental polymorphism. Totally, 172,271 SNPs were obtained in the F2 population. The raw SNPs were further filtered by the criterions of MAF $>$ 0.1 and missing rate $<$0.9, thus 114,183 high-quality SNPs were used for subsequently analysis finally.

Bin map construction and QTL analysis

The bin map was constructed using a sliding window approach as described by Huang et al. (2009) with a little modification. Raw SNPs were scanned in 15-SNP-windows with a sliding step of one SNP. In each window, the genotype was defined by the ratio of three kinds of SNPs (GEMS31, DAN3130 and Heterozygous): the widows was called homozygous GEMS31, or DAN3130 genotype when more than 12/15 of SNPs were from either GEMS31, or DAN3130, otherwise it would be defined as heterozygous genotype Regions with no more than 15 continuous heterozygous windows were set as breakpoint, since there will be several heterozygous windows for transition at the breakpoints, but these should not span more than 15 uninterrupted windows. Adjacent windows with same genotypes were merged as a block. Blocks with length less than 500 kb, or windows less than 15 were set as missing to avoid false double crossover. Adjacent windows and successive small blocks (less than 100 kb or five mapped SNPs) with frequently transient genotypes were merged and defined as a larger heterozygous block. 5 kb incremental segments with the same genotype across the 216 F_2 were merged together as a bin marker. Bins with extreme distortion (χ^2, $P < 1.0$ E-10) were discarded. The bins with missing genotype were imputed by R/qtl software using the "argmax" method. The linkage map was constructed with the est.map function of R/qtl (Broman et al. 2003).

Composite interval mapping (CIM) was performed for QTL analysis using the R/qtl package (Broman et al. 2003) with a scanning window size of 10-cM. The LOD threshold for QTL defining was calculated by 1,000 times permutations ($P < 0.05$) for each trait. The confidence intervals were estimated using the 1.5 LOD-drop method (Wang et al. 2018).

In the F_2 K22×DAN340 population, composite interval mapping (CIM) implemented in Windows QTL Cartographer V2.5 was used for QTL mapping. Zmap (model 6) with a 10 cM window and an interval-mapping increment of 2 cM were selected. The confidence interval was determined 1.5 LOD value down flanking the peak (Zeng 1994; Silva et al. 2012).

BSA analysis

Two libraries (extreme-high and extreme-low 5-F-THF) were sequenced by Hiseq Xten, and we obtained 110.27 and 128.18 million cleaned paired-end reads in each library (GEMS31 and DAN3130), respectively. The Q20 percentages were from 92.19 to 98.38, respectively. Next, the reads ere aligned to the maize genome (ftp://ftp.ensemblgenomes.org/pub/plants/release-37/fasta/zea_mays/dna/), and the SNPs and Indels were called using HaplotypeCaller in Genome Analysis Toolkit (GATK) after bam sorting, index using SAMtools v0.1.18 (Li et al. 2009; McKenna et al. 2010). The raw variants were then filtered using GATK VariantFiltration (for SNPs: QD < 2, MQ < 40, FS > 60, MQRankSum < -12.5, ReadPosRankSum < -8.0 and SOR > 3; for Indels: QD < 2, FS > 200, SOR > 10, MQRankSum < -12.5 and ReadPosRankSum < -8). Besides that, the SNPs and Indels with only two alleles, a coverage of more than ten reads and detected in both pools were retained for further analysis. Finally, we obtained 3.88 M SNPs and 0.473 M Indels for GEMS31 and DAN3130 group, respectively. A SNP-index was calculated for each detected SNPs. An average SNP-index was computed for at least five consecutive SNPs in a 500 Kb sliding window and 50 Kb step. The Δ (SNP-index) was calculated combining the information of SNP-index in both pools and plotted against the chromosome position (Takagi et al. 2013).

ACKNOWLEDGEMENTS

This work was financially supported by the Ministry of Science and Technology of China (2016YFD0100503 to L.J.), the National Natural Science Foundation of China (31870283 to L.J.), Shanghai Agriculture Applied Technology Development Program (Z20180103 to L.J.), Beijing Natural Science Foundation (6172032 to B.W.).

AUTHOR CONTRIBUTIONS

W.G., L.J., T.L., and B.W. analyzed the data and wrote the manuscript. L.J., B.W., and C.Z. designed the study. L.J., T.L., and D.Y. developed the folate RIL population. B.W. and W.G. performed QTL analysis using different segregated populations, respectively. T.L. and W.W. constructed the libraries for sequencing. J.G. and H.W. assembled the genome and performed BSA analysis. L.J., T.L., F.M.S.A., X.W. and Q.L. performed the folate analysis and cob color analysis. H.W., J.T. and C.Z. supervised the project and revised the manuscript. All authors read and approved the final manuscript.

REFERENCES

Abbasi IHR, Abbasi F, Wang L, Abd El Hack ME, Swelum AA, Hao R, Yao J, Cao Y (2018) Folate promotes S-adenosyl methionine reactions and the microbial methylation cycle and boosts ruminants production and reproduction. **AMB Exp** 8: 65

Bae S, Ulrich CM, Bailey LB, Malysheva O, Brown EC, Maneval DR, Neuhouser ML, Cheng TYD, Miller JW, Zheng YY, Xiao LR, Hou LF, Song XL, Buck K, Beresford SAA, Caudill MA (2014) Impact of folic acid fortification on global DNA methylation and one-carbon biomarkers in the Women's Health Initiative Observational Study cohort. **Epigenetics** 9: 396–403

Besson V, Rebeille F, Neuburger M, Douce R, Cossins EA (1993) Effects of tetrahydrofolate polyglutamates on the kinetic parameters of serine hydroxymethyltransferase and glycine decarboxylase from pea leaf mitochondria. **Biochem J** 292: 425–430

Bhutta ZA, Salam RA (2012) Global nutrition epidemiology and trends. **Ann Nutr Metab** 61: 19–27

Blancquaert D, Storozhenko S, Loizeau K, De Steur H, De Brouwer V, Viaene J, Ravanel S, Rebeille F, Lambert W, Van Der Straeten D (2010) Folates and folic acid: From fundamental research toward sustainable health. **Crit Rev Plant Sci** 29: 14–35

Blancquaert D, Van Daele J, Strobbe S, Kiekens F, Storozhenko S, De Steur H, Gellynck X, Lambert W, Stove C, Van Der Straeten D (2015) Improving folate (vitamin B-9) stability in biofortified rice through metabolic engineering. **Nat Biotechnol** 33: 1076–1078

Broman KW, Wu H, Sen S, Churchill GA (2003) R/qtl: QTL mapping in experimental crosses. **Bioinformatics** 19: 889–890

Chang X, DeFries RS, Liu L, Davis K (2018) Understanding dietary and staple food transitions in China from multiple scales. **PLoS ONE** 13: e0195775

Chen ZL, Wang BB, Dong XM, Liu H, Ren LH, Chen J, Hauck A, Song WB, Lai JS (2014a) An ultra-high density bin-map for rapid QTL mapping for tassel and ear architecture in a large F-2 maize population. **BMC Genomics** 15: 433

Chen J, Zeng B, Zhang M, Xie SJ, Wang GK, Hauck A, Lai JS (2014b) Dynamic transcriptome landscape of maize embryo and endosperm development. **Plant Physiol** 166: 252–264

Cole BF, Baron JA, Sandler RS, Haile RW, Ahnen DJ, Bresalier RS, McKeown-Eyssen G, Summers RW, Rothstein RI, Burke CA, Snover DC, Church TR, Allen JI, Robertson DJ, Beck GJ, Bond JH, Byers T, Mandel JS, Mott LA, Pearson LH, Barry EL, Rees JR, Marcon N, Saibil F, Ueland PM, Greenberg ER (2007) Folic acid for the prevention of colorectal adenomas: A randomized clinical trial. **Jama-J Am Med Assoc** 297: 2351–2359

De Lepeleire J, Strobbe S, Verstraete J, Blancquaert D, Ambach L, Visser RGF, Stove C, Van Der Straeten D (2018) Folate biofortification of potato by tuber-specific expression of four folate biosynthesis genes. **Mol Plant** 11: 175–188

Diaz de la Garza R, Quinlivan EP, Klaus SM, Basset GJ, Gregory JF 3rd, Hanson AD (2004) Folate biofortification in tomatoes by engineering the pteridine branch of folate synthesis. **Proc Natl Acad Sci USA** 101: 13720–13725

Diaz de la Garza R, Gregory JF 3rd, Hanson AD (2007) Folate biofortification of tomato fruit. **Proc Natl Acad Sci USA** 104: 4218–4222

Dong W, Cheng Z, Wang X, Wang B, Zhang H, Su N, Yamamaro C, Lei C, Wang J, Wang J, Zhang X, Guo X, Wu F, Zhai H, Wan J (2011) Determination of folate content in rice germplasm (*Oryza sativa* L.) using tri-enzyme extraction and microbiological assays. **Int J Food Sci Nutr** 62: 537–543

Ferrer JL, Ravanel S, Robert M, Dumas R (2004) Crystal structures of cobalamin-independent methionine synthase complexed with zinc, homocysteine, and methyltetrahydrofolate. **J Biol Chem** 279: 44235–44238

Gorelova V, De Lepeleire J, Van Daele J, Pluim D, Mei C, Cuypers A, Leroux O, Rebeille F, Schellens JHM, Blancquaert D, Stove CP, Van Der Straeten D (2017) Dihydrofolate reductase/thymidylate synthase fine-tunes the folate status and controls redox homeostasis in plants. **Plant Cell** 29: 2831–2853

Guo QN, Wang HD, Tie LZ, Li T, Xiao H, Long JG, Liao SX (2017) Parental genetic variants, MTHFR 677C > T and MTRR 66A>G, associated differently with fetal congenital heart defect. **BioMed Res Int** 2017: doi: 10.1155/2017/3043476

Hanson AD, Gregory JF 3rd (2011) Folate biosynthesis, turnover, and transport in plants. **Annu RevPlant Biol** 62: 105–125

Huang X, Feng Q, Qian Q, Zhao Q, Wang L, Wang A, Guan J, Fan D, Weng Q, Huang T, Dong G, Sang T, Han B (2009) High-throughput genotyping by whole-genome resequencing. **Genome Res** 19: 1068–1076

Jiang L, Wang W, Lian T, Zhang C (2017) Manipulation of metabolic pathways to develop vitamin-enriched crops for human health. **Front Plant Sci** 8: 937

Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. **Bioinformatics** 25: 1754–1760

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, Proc GPD (2009) The sequence alignment/map format and SAMtools. **Bioinformatics** 25: 2078–2079

Li H, Peng Z, Yang X, Wang W, Fu J, Wang J, Han Y, Chai Y, Guo T, Yang N, Liu J, Warburton ML, Cheng Y, Hao X, Zhang P, Zhao J, Liu Y, Wang G, Li J, Yan J (2013) Genome-wide association study dissects the genetic architecture of oil biosynthesis in maize kernels. **Nat Genet** 45: 43–50

Lian T, Guo WZ, Chen MR, Li JL, Liang QJ, Liu F, Meng HY, Xu BS, Chen JF, Zhang CY, Jiang L (2015) Genome-wide identification and transcriptional analysis of folate metabolism-related genes in maize kernels. **BMC Plant Biol** 15: 204

Liang QJ, Wang K, Liu XN, Riaz B, Jiang L, Wan X, Ye XG, Zhang CY (2019) Improved folate accumulation in genetically modified maize and wheat. **J Exp Bot** 70: 1539–1551

Lu XD, Liu JS, Ren W, Yang Q, Chai ZG, Chen RM, Wang L, Zhao J, Lang ZH, Wang HY, Fan YL, Zhao J, Zhang CY (2018) Gene-indexed mutations in maize. **Mol Plant** 11: 496–504

McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Garimella K, Altshuler D, Gabriel S, Daly M, DePristo MA (2010) The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. **Genome Res** 20: 1297–1303

Mickelson SM, Stuber CS, Senior L, Kaeppler SM (2002) Quantitative trait loci controlling leaf and tassel traits in a B73 x MO17 population of maize. **Crop Sci** 42: 1902–1909

Moore EM, Ames D, Mander AG, Carne RP, Brodaty H, Woodward MC, Boundy K, Ellis KA, Bush AI, Faux NG, Martins RN, Masters CL, Rowe CC, Szoeke C, Watters DA (2014) Among vitamin B12 deficient older people, high folate levels are associated with worse cognitive function: Combined data from three cohorts. **J Alzheimers Dis** 39: 661–668

Murigneux A, Barloy D, Leroy P, Beckert M (1993) Molecular and morphological evaluation of doubled haploid lines in maize. 1. Homogeneity within DH lines. **Theor Appl Genet** 86: 837–842

Naqvi S, Zhu C, Farre G, Ramessar K, Bassie L, Breitenbach J, Perez Conesa D, Ros G, Sandmann G, Capell T, Christou P (2009) Transgenic multivitamin corn through biofortification of endosperm with three vitamins representing three distinct metabolic pathways. **Proc Natl Acad Sci USA** 106: 7762–7767

Pan QC, Xu YC, Li K, Peng Y, Zhan W, Li WQ, Li L, Yan JB (2017) The genetic basis of plant architecture in 10 maize recombinant inbred line populations. **Plant Physiol** 175: 858–873

Peiffer JA, Romay MC, Gore MA, Flint-Garcia SA, Zhang ZW, Millard MJ, Gardner CAC, McMullen MD, Holland JB, Bradbury PJ, Buckler ES (2014) The genetic architecture of maize height. **Genetics** 196: 1337–1356

Pixley K, Warburton M, Xie C (2010) Molecular marker-assisted breeding options for maize improvement in Aisa. **Mol Breed** 26: 339–356

Price AH (2006) Believe it or not, QTLs are accurate! **Trends Plant Sci** 11: 213–216

Rader JI, Schneeman BO (2006) Prevalence of neural tube defects, folate status, and folate fortification of enriched cereal-grain products in the United States. **Pediatrics** 117: 1394–1399

Raihan MS, Liu J, Huang J, Guo H, Pan QC, Yan JB (2016) Multi-environment QTL analysis of grain morphology traits and fine mapping of a kernel-width QTL in Zheng58 x SK maize population. **Theor Appl Genet** 129: 1465–1477

Robbins ML, Wang PH, Sekhon RS, Chopra S (2009) Gene structure induced epigenetic modifications of pericarp color1 alleles of maize result in tissue-specific mosaicism. **PLoS ONE** 4: e8231

Sekhon RS, Lin H, Childs KL, Hansey CN, Buell CR, de Leon N, Kaeppler SM (2011) Genome-wide atlas of transcription during maize development. **Plant Physiol** 66: 553–563

Silva Lda C, Wang S, Zeng ZB (2012) Composite interval mapping and multiple interval mapping: Procedures and guidelines for using Windows QTL Cartographer. **Methods Mol Biol** 871: 75–119

Silver MJ, Kessler NJ, Hennig BJ, Dominguez-Salas P, Laritsky E, Baker MS, Coarfa C, Hernandez-Vargas H, Castelino JM, Routledge MN, Gong YY, Herceg Z, Lee YS, Lee K, Moore SE, Fulford AJ, Prentice AM, Waterland RA (2015) Independent genomewide screens identify the tumor suppressor VTRNA2-1 as a human epiallele responsive to periconceptional environment. **Genome Biol** 16: 118

Stelpflug S, Sekhon R, Vaillancourt B, Hirsch C, Buell C, de Leon N, Kaeppler S (2015) An expanded maize gene expression atlas based on RNA sequencing and its use to explore root development. **Plant Genome** 9: 1

Storozhenko S, De Brouwer V, Volckaert M, Navarrete O, Blancquaert D, Zhang GF, Lambert W, Van Der Straeten D (2007) Folate fortification of rice by metabolic engineering. **Nat Biotechnol** 25: 1277–1279

Strobbe S, Van Der Straeten D (2017) Folate biofortification in food crops. **Curr Opin Biotech** 44: 202–211

Suh JR, Herbig AK, Stover PJ (2001) New perspectives on folate catabolism. **Annu Rev Nutr** 21: 255–282

Takagi H, Abe A, Yoshida K, Kosugi S, Natsume S, Mitsuoka C, Uemura A, Utsushi H, Tamiru M, Takuno S, Innan H, Cano LM, Kamoun S, Terauchi R (2013) QTL-seq: Rapid

mapping of quantitative trait loci in rice by whole genome resequencing of DNA from two bulked populations. **Plant J** 74: 174–183

Vales M, Schön C, Capettini F, Chen X, Corey A, Mather D, Mundt C, Richardson K, Sandoval-Islas J, Utz H, Hayes P (2005) Effect of population size on the estimation of QTL: A test using resistance to barley stripe rust. **Theor Appl Genet** 111: 1260–1270

Wang BB, Zhu YB, Zhu JJ, Liu ZP, Liu H, Dong XM, Guo JJ, Li W, Chen J, Gao C, Zheng XM, E L, Lai JS, Zhao HM, Song WB (2018) Identification and fine-mapping of a major maize leaf width QTL in a re-sequenced large recombinant inbred lines population. **Front Plant Sci** 9: 101

Waskiewicz A, Piotrowski W, Broda G, Sobczyk-Kopciol A, Ploski R (2011) Impact of *MTHFR C677T* gene polymorphism, folate, B6 and B12 vitamin intake on homocysteine concentration in polish adult population. **Eur Heart J** 32: 720–720

Winter D, Vinegar B, Nahal H, Ammar R, Wilson G, Provart N (2007) An "Electronic Fluorescent Pictograph" browser for exploring and analyzing large-scale biological data sets. **PLoS ONE** 2: e718

Xu GH, Wang XF, Huang C, Xu DY, Li D, Tian JG, Chen QY, Wang CL, Liang YM, Wu YY, Yang XH, Tian F (2017) Complex genetic architecture underlies maize tassel domestication. **New Phytol** 214: 852–864

Xu ST, Zhang DL, Cai Y, Zhou Y, Shah T, Ali F, Li Q, Li ZG, Wang WD, Li JS, Yang XH, Yan JB (2012) Dissecting tocopherols content in maize (*Zea mays* L.), using two segregating populations and high-density single nucleotide polymorphism markers. **BMC Plant Biol** 12: 201

Yang XH, Gao SB, Xu ST, Zhang ZX, Prasanna BM, Li L, Li JS, Yan JB (2011) Characterization of a global germplasm collection and its potential utilization for analysis of complex quantitative traits in maize. **Mol Breed** 28: 511–526

Yu HH, Xie WB, Wang J, Xing YZ, Xu CG, Li XH, Xiao JH, Zhang QF (2011) Gains in QTL detection using an ultra-high density SNP map based on population sequencing relative to traditional RFLP/SSR markers. **PLoS ONE** 6: e17595

Zeng Z (1994) Precision mapping of quantitative trait loci. **Genetics** 136: 1457–1468

Scan using WeChat with your smartphone to view JIPB online

Scan with iPhone or iPad to view JIPB online