



# Characterization of Genome-Wide Variation in Four-Row Wax, a Waxy Maize Landrace with a Reduced Kernel Row Phenotype

Hanmei Liu<sup>1†</sup>, Xuewen Wang<sup>2†</sup>, Bin Wei<sup>3†</sup>, Yongbin Wang<sup>3</sup>, Yinghong Liu<sup>3</sup>, Junjie Zhang<sup>1</sup>, Yufeng Hu<sup>4</sup>, Guowu Yu<sup>4</sup>, Jian Li<sup>5</sup>, Zhanbin Xu<sup>5</sup> and Yubi Huang<sup>3,4\*</sup>

<sup>1</sup> College of Life Science, Sichuan Agricultural University, Ya'an, China, <sup>2</sup> Germplasm Bank of Wild Species, Kunming Institute of Botany, Chinese Academy of Sciences, Kunming, China, <sup>3</sup> Maize Research Institute, Sichuan Agricultural University, Chengdu, China, <sup>4</sup> College of Agronomy, Sichuan Agricultural University, Chengdu, China, <sup>5</sup> Seed Station of Xishuangbanna, Jinghong, China

## OPEN ACCESS

### Edited by:

Jean-Philippe Vielle-Calzada,  
Centro de Investigación y de Estudios  
Avanzados del Instituto Politécnico  
Nacional, Mexico

### Reviewed by:

Serena Varotto,  
University of Padova, Italy  
Luis Enrique Eguarte,  
Universidad Nacional Autónoma  
de México, Mexico

### \*Correspondence:

Yubi Huang  
yubihuang@sohu.com

<sup>†</sup> These authors should be regarded  
as joint first authors.

### Specialty section:

This article was submitted to  
Plant Genetics and Genomics,  
a section of the journal  
Frontiers in Plant Science

**Received:** 12 February 2016

**Accepted:** 01 May 2016

**Published:** 18 May 2016

### Citation:

Liu H, Wang X, Wei B, Wang Y, Liu Y,  
Zhang J, Hu Y, Yu G, Li J, Xu Z  
and Huang Y (2016) Characterization  
of Genome-Wide Variation  
in Four-Row Wax, a Waxy Maize  
Landrace with a Reduced Kernel Row  
Phenotype. *Front. Plant Sci.* 7:667.  
doi: 10.3389/fpls.2016.00667

In southwest China, some maize landraces have long been isolated geographically, and have phenotypes that differ from those of widely grown cultivars. These landraces may harbor rich genetic variation responsible for those phenotypes. Four-row Wax is one such landrace, with four rows of kernels on the cob. We resequenced the genome of Four-row Wax, obtaining 50.46 Gb sequence at 21.87× coverage, then identified and characterized 3,252,194 SNPs, 213,181 short InDels (1–5 bp) and 39,631 structural variations (greater than 5 bp). Of those, 312,511 (9.6%) SNPs were novel compared to the most detailed haplotype map (HapMap) SNP database of maize. Characterization of variations in reported kernel row number (KRN) related genes and KRN QTL regions revealed potential causal mutations in *fea2*, *td1*, *kn1*, and *te1*. Genome-wide comparisons revealed abundant genetic variations in Four-row Wax, which may be associated with environmental adaptation. The sequence and SNP variations described here enrich genetic resources of maize, and provide guidance into study of seed numbers for crop yield improvement.

**Keywords:** Four-row Wax, genome-wide variations, SNPs, kernel row number, landraces of Chinese waxy maize

## INTRODUCTION

As the most widely grown C4 crop in the world, the first maize genome sequence was published in 2009 for a commercial variety called B73 (Schnable et al., 2009). Subsequently, much progress has been made to reveal genetic variation among maize varieties. A popcorn variety called *Palomero* was the second maize genome sequence published, and its comparison to B73 revealed variations, as well as monomorphic gene such as heavy metal detoxification genes that may have been selected during domestication (Vielle-Calzada et al., 2009). Additional genetic variations were detected from 27 inbred lines and a NAM population using NGS technology, and the resulting data was used in compiling the first generation maize HapMap (Gore et al., 2009; McMullen et al., 2009). The first HapMap revealed the genome-wide variation pattern of maize for the first time, although

**Abbreviations:** GO, gene ontology; HapMap, haplotype map; InDel, insertion and deletion; KRN, kernel row number; NAM, nested association mapping; NGS, next generation sequencing; PAV, presence/absence variation; SNP, single nucleotide polymorphism; SV, structure variation; UTR, untranslated region.

only ~38% of the maize genome was covered. Whole genome sequencing of 103 maize landraces and elite lines identified 55 million genome-wide SNPs (Chia et al., 2012; Hufford et al., 2012). Three types of variations including SNPs, SVs and genome size variations were included in the second generation maize HapMap (Chia et al., 2012). Characterization of the variations may help to suggest their roles and functions during maize domestication and improvement (Hufford et al., 2012). A growing collection of individual and population studies (Lai et al., 2010; Jiao et al., 2012; Xu et al., 2014) have also contributed to knowledge of maize variation and deducing functions of these variations. Despite these extensive studies of maize variation, sequencing of additional novel maize varieties can further enrich the variation dataset and may also benefit research into important traits found in the novel varieties.

Waxy maize has a dramatic reduction in synthesis of amylose, caused by a recessive mutation of the waxy gene (*wx*; Wessler and Varagona, 1985; Zheng et al., 2013). Waxy maize may have originated in southwest China several 100 years ago (Tian et al., 2009). Four-row Wax is one of the waxy landraces, however, it has a special phenotype of only four rows of kernels on the cob (Figures 1A,B). It has a long history of production in Menghai County, Yunnan Province in southwest China as food by local Dai minority people (Tian et al., 2009). Four-row Wax is mainly grown in red latosol soils of mountainous areas, longitudes 99°56'E to 100°41'E and latitudes 21°28'N to 22°28'N (Figure 1C). This region has a tropical and subtropical monsoon climate, an annual average temperature of 18.7°C and total rainfall of 1341 mm. Four-row Wax is not known to have been subjected to scientific breeding (Tian et al., 2009). Researchers have been focused on this variety for some time, especially for its systematics and waxy trait (Fan et al., 2008, 2009; Tian et al., 2009). A 15-bp deletion at the 10th exon in *wx* is the mutation thought to be responsible for low amylose content in Four-row Wax (Fan et al., 2008, 2009). Phylogenetic analysis indicates that Four-row Wax is related to other Chinese waxy maize genotypes and landraces, in particular from neighboring Yunnan and Guangxi provinces, and suggests that Chinese waxy maize might have been derived from varieties of cultivated flint maize (Fan et al., 2008).

Both waxy endosperm and seed number are of importance for food quality and for crop improvement. Thus it is of interest and significance to analyze the genome of Four-row Wax for its genetic basis, domestication and evolution, to obtain diagnostic DNA markers for breeding, and also to provide clues to features of its features and those of other local maize varieties in China. Advances in NGS technologies make it possible to 'resequence' entire genomes efficiently and economically, providing the opportunity to identify and investigate genome-wide DNA polymorphisms, including SNPs, short InDels, SVs (Huang et al., 2010; Chia et al., 2012; Hufford et al., 2012; Xu et al., 2012; Romay et al., 2013). Further association of DNA-level variations with phenotypic differences will provide clues to interpret the genetic basis of important traits.

To gain insight into its genomic diversity, whole-genome resequencing of Four-row Wax was carried out using NGS technology. Genome-wide variations were identified through

comprehensive bioinformatics comparisons. Analysis of novel variations and candidate genes were conducted in this paper with the aim of helping to explain the KRN trait of Four-row Wax. The data resources and information described here will increase our knowledge of maize diversity, help to speed up genetic dissection of KRN and guide future breeding applications to improve maize yield.

## MATERIALS AND METHODS

### DNA Preparation and Sequencing

Four-row Wax, a Chinese waxy maize landrace, is not known to have been subjected to scientific breeding (Tian et al., 2009). We randomly selected a single individual from a traditional plot of this landrace with typical phenotypic characteristics, such as four rows of kernels on the cob, which was growing in the field with natural pollination at the flowering stage. We chose 10 open pollinated waxy seeds for genome sequencing by detecting KI/I2-staining of seeds and the 15-bp deletion mutation of *wx*. Thereafter, we ground 7-day-old etiolated seedlings in liquid nitrogen and extracted DNA using a CTAB (Hexadecyltrimethylammonium bromide)-based method, then randomly sheared the DNA. Fragments of ~500 bp were selected from a gel and used for a library preparation according to Illumina's instructions. The libraries were sequenced at paired ends using an Illumina HiSeq 2000 in four sequencing lanes. The average length of sequencing reads was 90 bp.

The maize B73 reference genome sequence and gene set were downloaded from MaizeGDB (release version 4a.53; Schnable et al., 2009). We then filtered genes using the following criteria: (1) genes with CDS shorter than 150 bp were removed; (2) genes was discarded if a stop codon occurred within the coding region; (3) the longest transcript of a gene locus was retained to represent all transcripts at that locus; (4) overlapping genes were removed, retaining the longest one. After filtration, 75,747 genes were retained and were used for analysis.

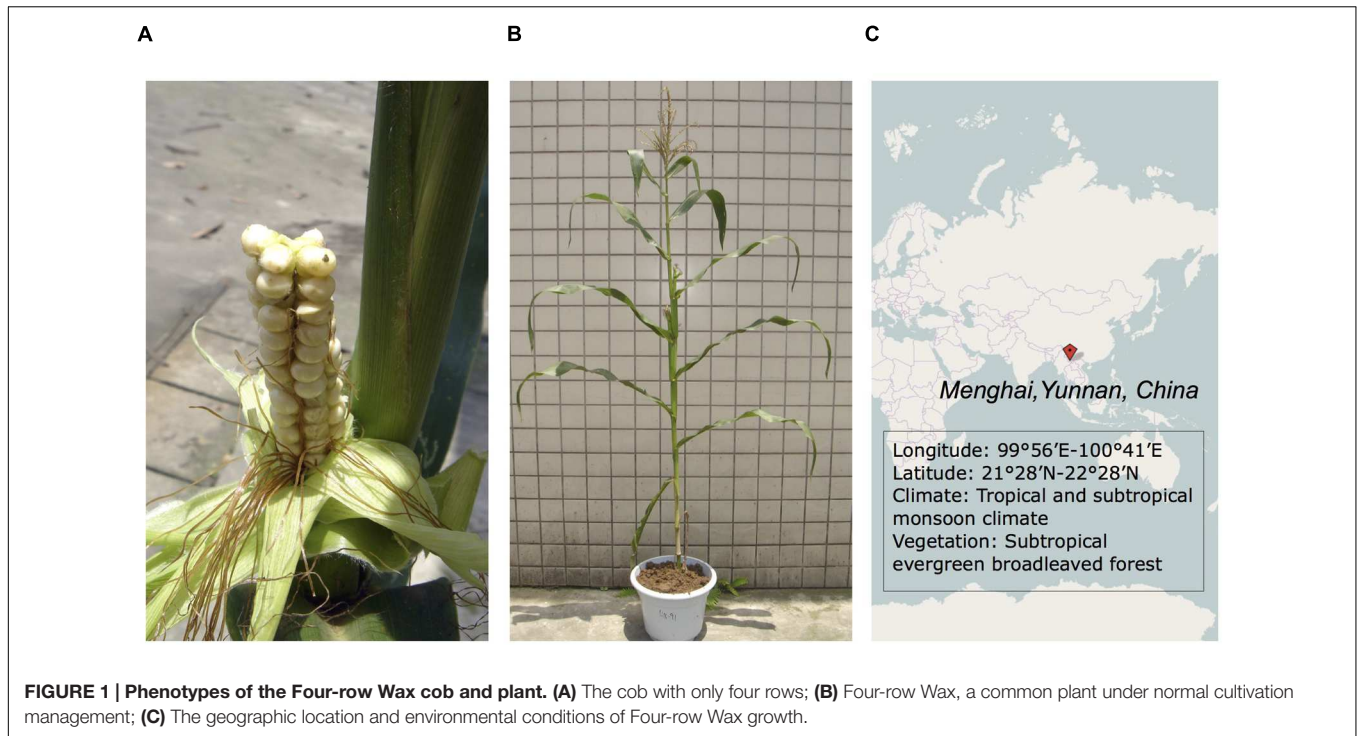
### Assembly and Novel Gene Detection

*De novo* assembly was performed for all reads sequenced from Four-row Wax using SOAPdenovo software<sup>1</sup> (version 1.05; Li et al., 2010). The k-mer parameter was set to 17. We then aligned the assembly against the B73 reference sequence. Regions which could not be mapped to the B73 genome were extracted and BLASTed against proteins of *Arabidopsis thaliana*, *Oryza sativa*, and *Sorghum bicolor* using default settings. Sequences with similarity higher than 70% to known proteins were considered as candidate novel genes.

### Short Read Mapping

We filtered low quality reads before mapping, including reads: (1) with more than 5% Ns or As; (2) with 20 bp or more having quality lower than 7; (3) contaminated by the adaptor (i.e., more than 10 bp aligned to the adaptor when allowing fewer than 3 bp mismatches); (4) with the first and second ends overlapping

<sup>1</sup><http://soap.genomics.org.cn>



(10 bp aligned when allowing 10% mismatches); (5) with both ends identical to other reads, suggesting PCR duplication. The filtered reads were then aligned to the B73 reference genome sequence using SOAPaligner (version 2.18; Li et al., 2009b). We set the mapping parameter to allow at most two mismatches within one read. The sequencing depth and coverage were then calculated based on the alignment result. To detect small InDels, we conducted further mapping, setting the parameter ‘-g 5’, which allowed less than or equal to 5 bp gaps.

### Identification of SNPs, InDels, and SVs

To detect SNPs, we used SOAPsn software (version 1.03; Li et al., 2009a) using the following criteria: (1) The quality score given by SOAPsn (which takes sequencing quality and ambiguity into consideration) should be greater than 20; (2) the sequencing depth should be higher than 2; (3) SNPs within 5 bp of nearby SNPs were considered as other variations; (4) SNPs with copy number greater than 1.5. SNPs meeting these filtering criteria were considered to be supported by high quality data, and to minimize possible errors caused by repeated sequences.

For short InDels the ‘gaps allowed’ mapping result above was subjected to InDel detection using SOAPindel (Li et al., 2013), and gaps supported by more than one third of the mapped reads were retained as qualified InDels. In this study, InDels are defined as insertions or deletions of 5 bp or less in length, while insertions or deletions of more than five nucleotides are classified as SVs.

The identified SNPs and InDels were classified based on occurrence in coding regions, non-coding regions, start codons, stop codons or splice sites. SNPs in coding sequences were annotated as synonymous or non-synonymous. GO or PFAM annotation of genes was also conducted.

Structural variations were detected using SOAPsv (version 1.02), for InDel (more than 5 bp in length), replication, reversion, transposition, and other variations. The minimum read depth was set to 3.

### Identification of Regions without SNPs

We first calculated the distances between adjacent SNPs, and the distribution of interval distance was depicted. Genes located between two adjacent SNPs were extracted as candidate genes without SNPs. Then, to further verify that there were no SNPs and other variations within those genes, we filtered the genes located in, or overlapping with, repeated elements (according to the annotation information), and genes that were not sequenced (no reads mapped to the gene region).

### Novel SNP Identification

We obtained the maize HapMap SNPs (in B73 RefGen\_v1 coordinates) from the Panzea website (Chia et al., 2012). SNPs present in Four-row Wax (identified by mapping and aligning our reads to the B73 genome) but absent from the HapMap were considered novel.

### Variations Detection in Genomic Regions Associated with Kernel Row Number

We focused on four reported genes (Supplementary Table S10; Vollbrecht et al., 1991; Veit et al., 1998; Taguchi-Shiobara et al., 2001; Bommert et al., 2005, 2013; Lunde and Hake, 2009) and five large-effect QTLs (Supplementary Table S12; Yan et al., 2006; Li et al., 2007; Ma et al., 2007; Lu et al., 2010; Brown et al., 2011; Cook et al., 2012; Liu L. et al., 2015) related to KRN. We also compared variations in these regions between six elite inbred

lines (Zheng58, 5003, 478, 178, Chang7-2 and Mo17) reported previously (Lai et al., 2010) and Four-row Wax.

## RESULTS

### Resequencing, Mapping, and Assembly of the Four-Row Wax Genome

Using Four-row Wax genomic DNA, a total of 560.65 million sequencing reads of length 90 bp, totaling 50.46 Gb, were generated from four Illumina sequencing lanes (Supplementary Table S1). All sequence information was deposited at the Sequence Read Archive (SRA) of National Center for Biotechnology Information (NCBI), under accession number SRA090053. A total of 499.62 M reads (89.11%) were successfully mapped to the B73 maize reference genome (version: 4a.53) using software SOAP2. These reads covered 89.86% of the reference genome at average  $21\times$  depth.

In order to detect possible novel genes in Four-row Wax, we conducted a *de novo* genome assembly with software SOAPdenovo, and obtained a total assembled length of 996,022,706 bp (48.63% of the reference genome), believed to be mostly the low copy part of the maize genome. The contig N50 was 1022 bp, with the longest contig being 24,347 bp. The scaffold N50 was 2,410 bp, with the longest scaffold being 49,840 bp.

### Detection and Characterization of SNPs, InDels, and SVs

Sequencing depth affects accuracy of DNA variation detection (Arai-Kichise et al., 2011). Therefore, we compared the numbers of SNPs, InDels, and SVs detected from a set of combinations of resequencing data using different numbers of sequencing lanes, i.e., data with different coverages (Supplementary Table S1). The number of detected variations became stable above  $10\times$  coverage. Therefore, our  $21.87\times$  average depth of resequencing data is sufficient to identify genome-wide variations. A total of 3,252,194 SNPs, 213,181 InDels (1~5 bp), and 39,631 SVs were identified in the Four-row Wax genome as compared to the maize B73 reference sequence (Table 1).

We characterized the distributions of SNPs, InDels, and SVs in the Four-row Wax genome (Table 1). The average densities of detected SNPs, InDels, and SVs were 1578, 103, and 19 per Mb, respectively. The density of SNPs or InDels was uneven among chromosomes (Figure 2). For SNPs, 10 and 9 regions with high ( $>3500/\text{Mb}$ ) and low density ( $<500/\text{Mb}$ ) were found, respectively. High SNP regions were on chromosomes 1, 2, 4, 7, and 8, while low SNP regions were on chromosomes 1, 2, 3, 4, 6, and 9. Interestingly, four of ten high SNP regions were on chromosome 2, the highest being  $3962/\text{Mb}$ . Three of nine low SNP regions were on chromosome 4, the lowest being  $181/\text{Mb}$ . There were eleven regions with high InDel density ( $>250/\text{Mb}$ ) on chromosomes 1, 3, 4, 7, and 9, with the highest ( $285/\text{Mb}$ ) on chromosome 9. Five regions with low InDel density ( $<10/\text{Mb}$ ) were identified on chromosome 3, 4, and 6, the lowest ( $2/\text{Mb}$ ) located on chromosome 4. The majority of InDels (63.87%) were mononucleotide variations, including 61,433 insertions and

76,778 deletions (Figure 3). The abundance of InDels decreased with increasing length, while different InDels of similar length had similar abundance.

Four regions had high SV density ( $>80/\text{Mb}$ ) on chromosomes 1, 4, and 8. The highest SV density ( $145/\text{Mb}$ ) region was on chromosome 1 and the lowest ( $1/\text{Mb}$ ) was on chromosome 4. The average SV density on each chromosome was similar, although SV variations were not uniformly distributed within a chromosome (Table 1). Of the 39,631 SVs, 30,046, 6,879, and 1,887 were classified as deletion type, insertion type and tandem duplication, respectively. The types, numbers and length distributions of SVs are shown in Figure 4.

### Annotation of SNPs and InDels in Four-Row Wax

We annotated the sequences containing detected SNPs and InDels according to the gene annotation of the B73 reference genome (Schnable et al., 2009). Only 587,303 SNPs (18.06%) and 63,927 InDels (29.99%) were located in genic regions, the majority being in intergenic regions. The number of SNPs and InDels in gene coding sequences (CDSs), introns and UTRs are shown in Table 2, and their distributions on each chromosome are in Supplementary Tables S2 and S3. Of those variations in genic regions, 26.15% SNPs and 5.95% InDels were in CDSs while the remainder were in introns. Of those in CDSs, 46.36% of SNPs were non-synonymous, and 62.36% of InDels induced translation frame-shifts, while the remaining InDels (1,431) caused deletion/insertion of amino acids.

We further classified the non-synonymous SNPs for inferred effects on gene function (Table 3). Of all non-synonymous SNPs, 2,617 located in 2,259 transcripts induced premature stop codons; 681 abolished stop codons, resulting in longer open reading frames; 298 altered initiation methionine residues. A total of 2,769 SNPs were located within splice sites (1–3 bp into exons or 2 bp into introns), 718 of which affected essential splice donor or acceptor sites (first 2 bp or last 2 bp of an intron), resulting in 704 transcripts with open reading frame changes. In addition, 48,523 SNPs, located in 21,663 transcripts, encoded amino acids with significant property difference from those of B73, based on four types of hydrophobic, negative, positive, and polar without charge properties.

A total of 22,264 transcripts were predicted to contain 'large-effect' SNPs, which changed start codons, stop codons, splice sites or amino acids encoded (Table 3). GO analysis (Supplementary Table S4 and Supplementary Figure S1) showed that large-effect SNPs were significantly enriched in genes encoding serine/threonine kinase, tyrosine kinase, nucleoside-triphosphatase, serine hydrolase, and binding proteins, such as ATP binding, sugar binding, and heat shock protein.

### Regions without SNPs

We calculated the distance between each adjacent pair of SNPs, and found many with short intervening distances (Supplementary Figure S2). The distance between 1,793,165 adjacent SNPs (55.14%) was shorter than 100 bp, while 1,102,673 adjacent SNPs (33.9%) were more than 100 bp but less

**TABLE 1 | Distribution and density of SNPs, InDels, and SVs identified between Four-row Wax and B73 maize.**

Chromosome	SNPs	SNPs/Mb	InDels	InDels/Mb	Ins	Ins/Mb	Del	Del/Mb	SV	SV/Mb
Chr1	473,967	1,579	32,998	110	15,324	51	17,674	59	5,960	20
Chr2	386,104	1,645	23,981	102	11,111	47	12,870	55	4,580	20
Chr3	364,401	1,581	24,643	107	11,335	49	13,308	58	4,310	19
Chr4	424,860	1,719	24,735	100	11,214	45	13,521	55	4,890	20
Chr5	313,126	1,444	22,491	104	10,269	47	12,222	56	4,019	19
Chr6	249,643	1,475	17,063	101	7,681	45	9,382	55	3,025	18
Chr7	287,446	1,681	17,952	105	8,185	48	9,767	57	3,360	20
Chr8	273,747	1,569	17,799	102	8,301	48	9,498	54	3,325	19
Chr9	237,802	1,561	16,279	107	7,508	49	8,771	58	2,984	20
Chr10	227,839	1,522	14,643	98	6,689	45	7,954	53	2,655	18
Chr unknown <sup>1</sup>	13,259	903	597	41	211	14	386	26	523	36
Total	3,252,194	1,578	213,181	103	97,828	47	115,353	56	39,631	19

<sup>1</sup>Chr unknown indicates unplaced or unlocalized contigs.

than 1 kb apart. Only 10.96% of adjacent SNPs are 1 kb or more apart. Interestingly, we identified 39,865 regions of more than 10 kb without SNPs, i.e., which have identical sequences between B73 and Four-row Wax. A total of 28,952 genes were located in these invariant regions, including many encoding proteins with important biological functions essential for life, and mainly participating in primary metabolic processes (Supplementary Table S5 and Supplementary Figure S3), such as ligand-dependent nuclear receptor, electron carrier and enzymes with catalytic activity (NADH dehydrogenase, protein disulfide oxidoreductase, cysteine-type peptidase, and aspartic-type endopeptidase). Additionally, some proteins with DNA binding and metal ion (copper or zinc ion) binding activity were also enriched.

## Detection of Novel Variations and Novel Gene Candidates

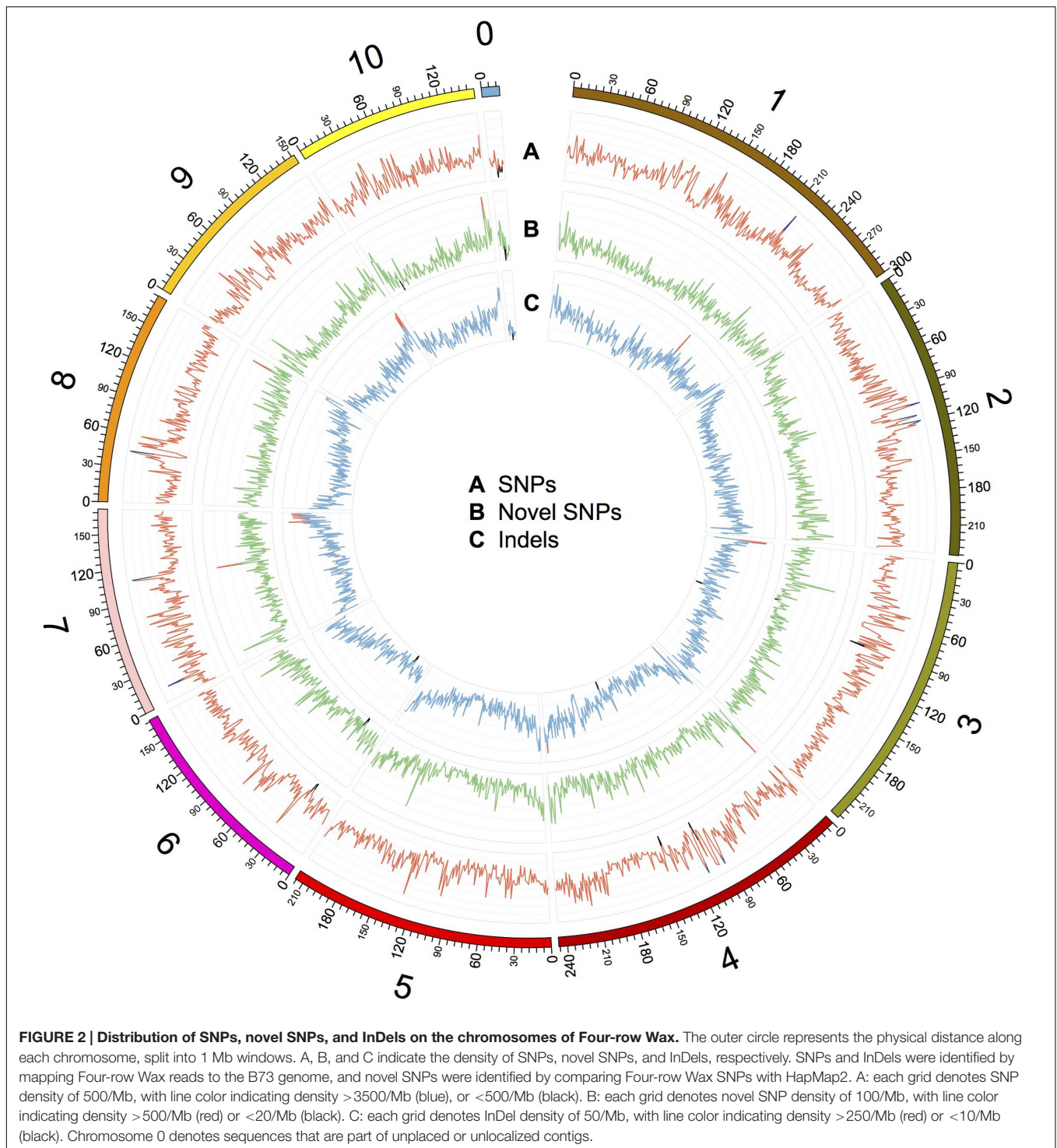
Specific variations in an individual may explain its unique phenotypes. We further analyzed novel variations in Four-row Wax, finding that 2,043,063 (62.8%) of the 3,252,194 SNPs differentiating it from B73 can also be found in the most comprehensive genome-wide maize SNP dataset, maize HapMap (Chia et al., 2012), while the remaining SNPs are unique to Four-row Wax. After filtering out SNPs in repetitive regions of the maize genome where SNP detection may be unreliable, 312,511 (9.6%) novel SNPs were identified in non-repetitive regions. The chromosome distributions and annotations of those novel SNPs were analyzed and listed in Supplementary Table S6. The average density of novel SNPs on each chromosome was similar, but there were high-density or low-density regions (Figure 2). Four regions with high novel SNP density (>500/Mb) were located on chromosomes 4, 7, 8, and 10, respectively. Seven regions with low novel SNP density (<20/Mb) were located on chromosomes 3, 4, 6, 9, and 10, respectively. Comparing the distributions of SNPs and InDels, we found some overlapping regions, such as a region (located about 216M of chromosome 1) with high SNPs and InDels; two regions (~90M of chr 3 and ~138M of chr 4) with low SNPs and InDels; and two regions (~27M of chr 6 and ~7M of chr 0) with low densities of all three kinds of variations

(SNPs, novel SNPs, and InDels; Figure 2). Most novel SNPs, 248,518 (79.52%), were located in intergenic regions and introns; with 37,198 (11.9% of) novel SNPs in coding sequences. The distribution of novel SNPs was similar to that of total detected SNPs described above. Novel SNPs were distributed in 24,285 genes, mainly in introns and UTRs, with only a few in CDSs. Among the ten genes containing the highest number of novel SNPs (Supplementary Table S7), six were predicted to encode binding proteins, including proteins binding zinc ions, DNA or ATP. Three genes have known functions encoding GTPase, tyrosine protein kinase and pentatricopeptide.

Similar to total SNPs, 63.42% of novel SNPs in Four-row Wax were non-synonymous. Of those, 16,809 (71.25%) belonged to the large effect SNP class. The types and numbers of large effect SNPs are provided (Table 3). Among all SNPs, the number that were non-synonymous was far more than the number of novel SNPs, so the large effect SNP number for each type was more than that of novel SNPs. However, for total SNPs and novel SNPs, the proportions that were non-synonymous were similar. Two types of mutations (premature stop codons and essential splice sites) occurred at higher proportions in novel SNPs than total SNPs. GO term analysis revealed that of genes enriched for these large effect novel SNPs were in four categories based on molecular function (Supplementary Table S8 and Supplementary Figure S4), including genes encoding ATP binding proteins, nucleoside-triphosphatase, serine/threonine kinase, and tyrosine kinase.

In the current maize database, InDels were not well documented and SVs were not accurate due to short reads and low depth of sequencing, so novel InDels and SVs in Four-row Wax were not predicted in this study.

The phenotypes that are characteristic of Four-row Wax may be caused by large sequence variations such as novel genes. We annotated the assembled Four-row Wax genome to seek novel genes, selecting scaffolds longer than 1 Kb and without corresponding annotation information in the reference genome. By comparison to gene sets of *Arabidopsis*, rice and sorghum, 583 novel genes were predicted, of which 123 have homologs in the current plant protein database

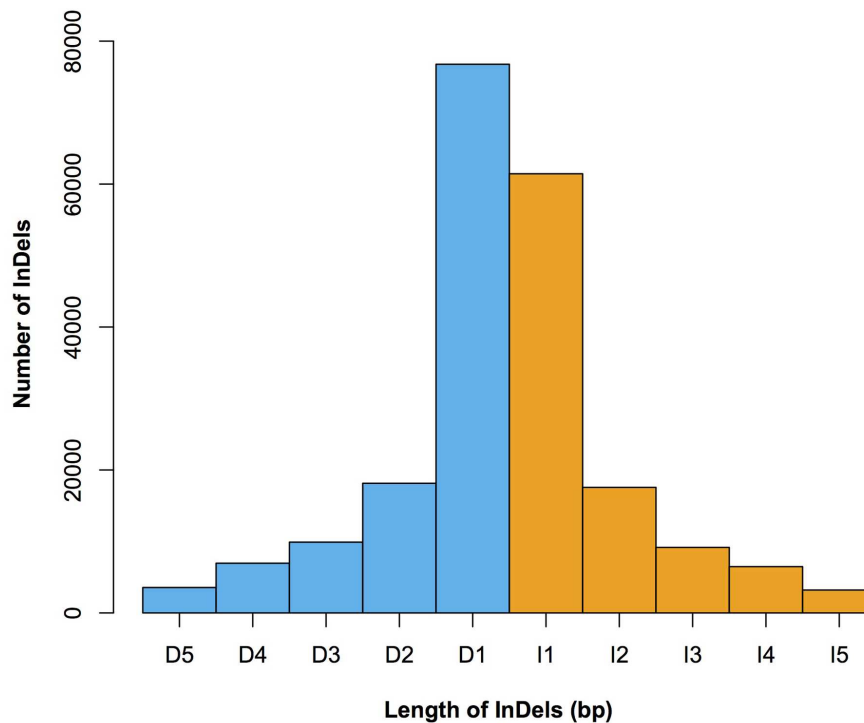


(Uniprot and KEGG). Four novel maize genes were identified based on less than 20% similarity to other genes in the maize protein database (Supplementary Table S9), encoding a zinc transporter 2, pentatricopeptide repeat-containing protein, phloem-specific lectin and saposin-like type B, respectively. Such striking divergence from other maize genes makes these of interest to study further as potential contributors to

phenotypic differences between Four-row Wax and other maize genotypes.

### Variations in the Regions Associated with KRN

Kernel row number is an important trait for maize yield. Previous research showed that it was a quantitative trait controlled by



**FIGURE 3 | The length distribution of 1–5 bp InDels in Four-row Wax.** Lengths (x-axis) and numbers (y-axis) of insertions (orange) and deletions (blue) are shown. I1–I5 denotes insertions of 1–5 bp in length, respectively; D1–D5 denotes deletions of 1–5 bp in length, respectively.

multiple genes, each with small phenotypic effects (Lu et al., 2010). Several QTLs and known genes have been associated with KRN (Lu et al., 2010; Brown et al., 2011; Cook et al., 2012; Bommert et al., 2013). We examined variations in Four-row Wax for four known KRN related genes (Supplementary Table S10), finding no SVs or novel SNPs. Some InDels and SNPs (all synonymous) were detected but were distributed in CDS, intron or 3'UTR regions (Supplementary Table S11). An InDel caused by a 3-bp deletion of AGG in the CDS of terminal ear1 (*te1*), resulted in the fourth position deletion of five consecutive glutamic acids, which may alter gene function and contribute to the low KRN phenotype in Four-row Wax.

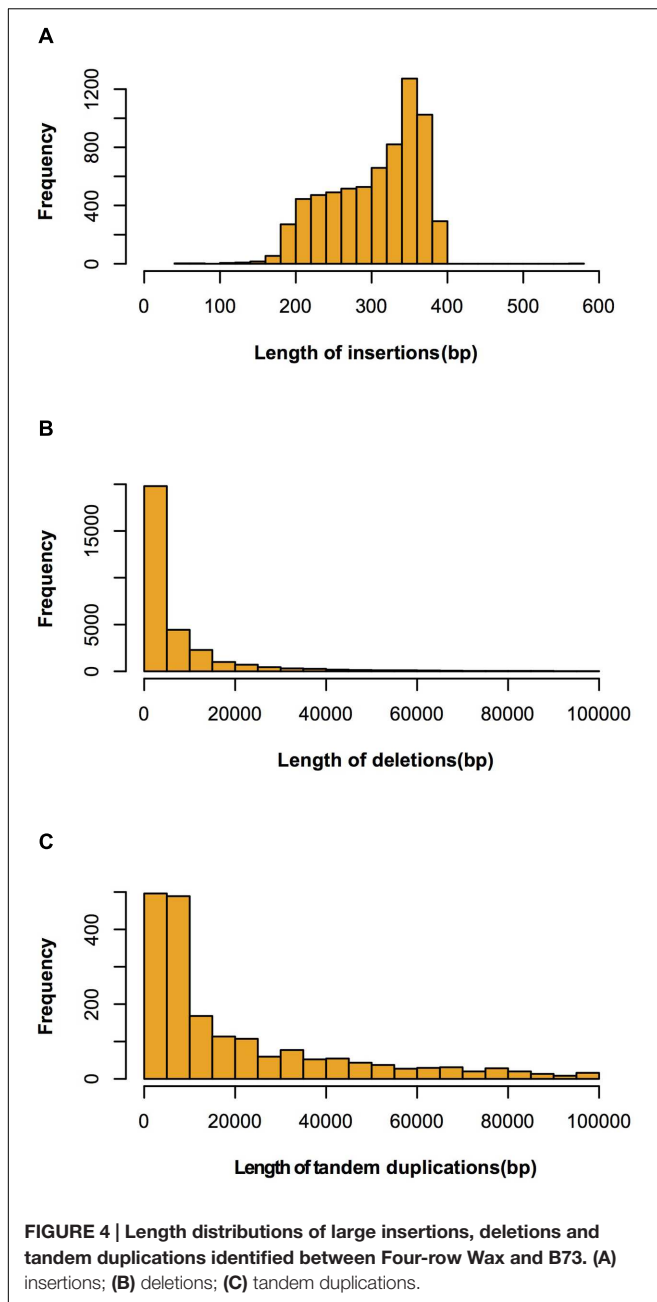
To further investigate possible variations responsible for KRN in maize, we analyzed five large-effect QTLs (Supplementary Table S12; Yan et al., 2006; Li et al., 2007; Ma et al., 2007; Lu et al., 2010; Brown et al., 2011; Cook et al., 2012), accounting for high proportions of the KRN phenotype. DNA variations in Four-row Wax for the five QTL regions are listed in Supplementary Table S13. The average density of each variation in the QTL regions was similar to that in the whole genome. The reported QTLs, obtained from linkage mapping, cover large genomic regions while only one or a few sites may cause the KRN phenotype. Thus, we investigated patterns of variation in smaller windows, of 10 Kb. We also compared the SNP densities with those of six previously published elite inbred lines (Lai et al., 2010), which had significant KRN differences and are widely used in Chinese maize breeding (Figure 5 and Supplementary Figure S6). SNP densities in a few 10 kb regions were substantially high in Four-row Wax than

other genotypes. For each QTL, the highest SNP-density region in Four-row Wax was scrutinized as a potential causal region of low KRN. There are a total of 88 genes (Supplementary Table S14) in the five QTL regions, of which 52 were functionally known. Among the known genes, we found one auxin responsive gene called small auxin up-regulated RNA (SAUR), for which family members were reported to reduce growth and seed yield, such as *SAUR39* in rice (Kant and Rothstein, 2009; Kant et al., 2009). Whether the SAUR gene may affect KRN in Four-row Wax is worthy of further investigation.

## DISCUSSION

### Rich Variation in the Four-Row Wax Genome

Modern maize has experienced domestication as a result of selection from *Zea mays* ssp. *Parviglumis*, and subsequent genetic improvement. A single domestication event is estimated to have taken place between 6000 and 9000 years ago in southern Mexico (Matsuoka et al., 2002), which resulted in the original maize landrace varieties. In the 20th century, breeders selected maize varieties and inbred lines from landraces for hybrid production, improving yield, resistance to biotic and abiotic stresses and seed nutritional quality (Jiao et al., 2012; Liu H. et al., 2015). Glutinous maize was first recorded in China in 1909, and was believed to have arisen from improvement of domesticated non-glutinous maize (Fan et al., 2008). Most Chinese waxy maize



landraces maintained rich genetic diversity, with different genetic improvement processes from modern improved maize lines (Fan et al., 2009). Some landraces had phenotypic variations in important agronomic traits, and were helpful to identify and characterize genes responsible for those traits. Using this knowledge, plant breeders can further improve crop varieties.

With rapid development of sequencing technologies, maize genome sequencing studies have been widely undertaken. Building on the completion of the B73 reference genome sequence (Schnable et al., 2009), resequencing of a number of maize lines (Chia et al., 2012; Hufford et al., 2012; Romay et al., 2013) provided opportunities to unravel the tremendous genetic

variation and diversity of maize at the whole genome scale. DNA polymorphisms among maize lines could be utilized for genetic analysis as well as crop improvement.

In our study, the maize landrace called Four-row Wax was resequenced and mapped to the B73 reference genome to discover genome-wide DNA variations. We uncovered more than three million SNPs and InDels, along with large numbers of SVs. Six resequenced elite maize inbred lines (Lai et al., 2010) include the parents of the most productive commercial hybrids in China and the USA, and are representatives of the main heterotic groups (B73, 478, 5003 and Zheng58 are in the Reid heterotic group; Mo17 and 178 are in the Lancaster heterotic group; Chang7-2 is in the Tangsipingtuo group). Compared to the six inbred lines, Four-row Wax was much richer in SNPs and InDels. The abundance of variation in Four-row Wax may be partly related to higher sequencing depth (21.87 $\times$ ) in our data than in those inbred lines (5.4 $\times$ ). However, the greater number of SNPs and InDels in Four-row Wax may also reflect greater genetic diversity. A phylogenetic tree constructed from identified SNPs (Supplementary Figure S5), showed Four-row Wax to be more distant from B73 than the elite maize inbreds. A relatively close genetic relationship was found between Four-row Wax and Chang7-2, with a more distant relationship to the other 5 lines and B73. Huangzaosi is an important inbred line widely used in China, derived from landraces in Huanghuaihai area. Chang7-2 is a line derived from Huangzaosi, and it has become a member of the important Tangsipingtuo heterotic group used in China. This finding shows that Four-row Wax is more closely related to the Tangsipingtuo group originated from Huanghuaihai area of China than to members of the important Reid or Lancaster heterotic groups. The results of phylogenetic analysis are consistent with a previous report of the relationships and characteristics of inbred lines (Lai et al., 2010), supporting the validity of genome resequencing and variation identification.

Chia et al. (2012) identified 55 million SNPs in 103 inbred lines representing a wide breadth of the *Z. mays* lineage, including 60 improved lines, 23 landraces, and 19 wild relatives, which are resources in the most comprehensive maize SNP datasets currently available, HapMap2. Most SNPs detected in Four-row Wax were also found in the HapMap2 dataset, but 312,511 (9.6% of) SNPs were unique to Four-row Wax, excluding 774,231 (23.81%) SNPs in repeated regions. The new SNPs detected in Four-row wax enrich the current HapMap database. Novel SNPs in Four-row Wax may have resulted from the following scenarios: (1) Four-row Wax was a landrace, derived from southwestern China, with a different genetic improvement history from modern maize lines; (2) maize is a cross-pollinated plant harboring rich genetic diversity, especially among different landraces.

## Genes with or without Abundant Variation

We found that protein kinases (such as serine/threonine kinase and tyrosine kinase) were enriched for large-effect SNPs. This tendency was also found in previous resequencing in maize (Lai et al., 2010) and other species, such as rice (McNally et al.,



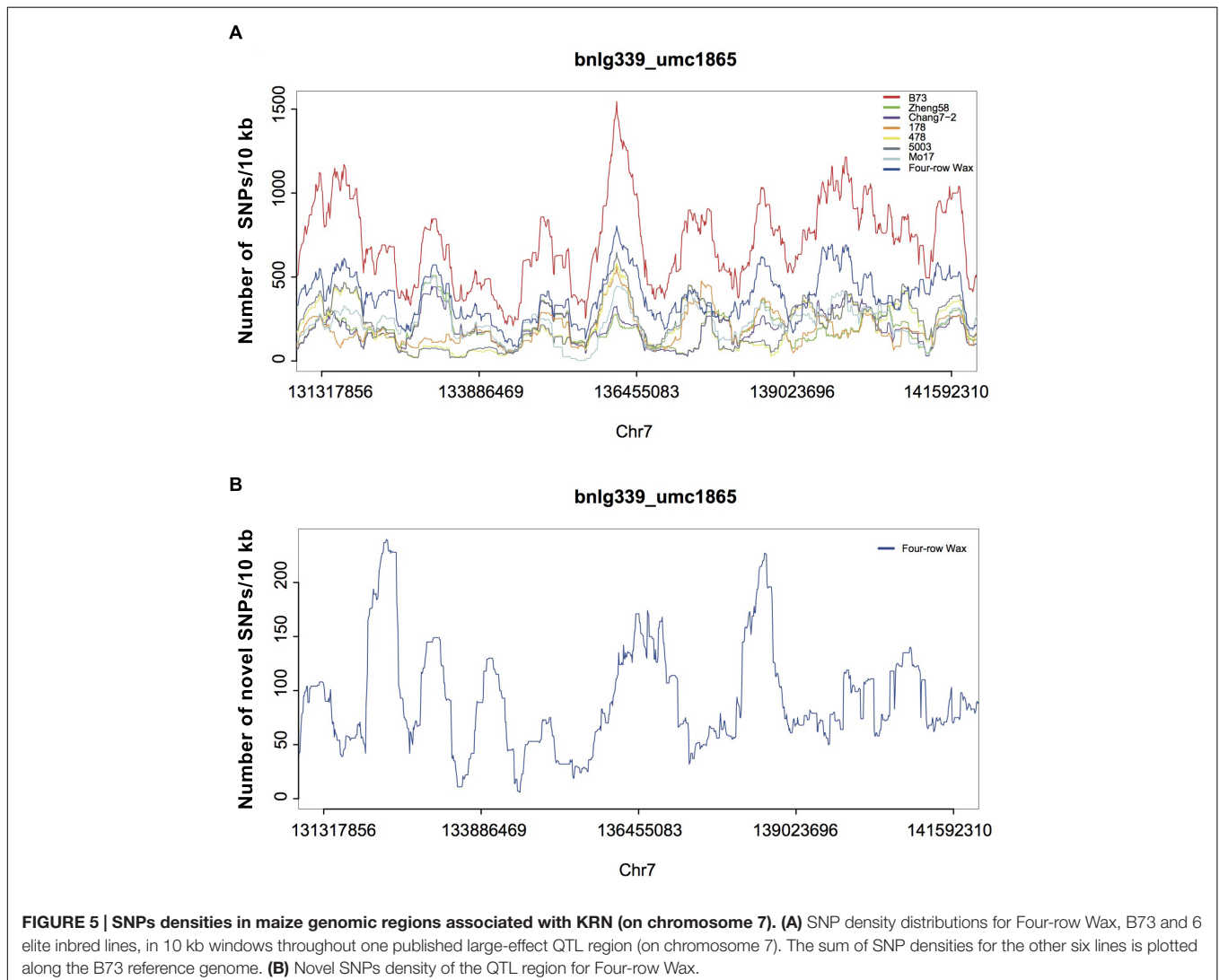
**TABLE 2 | Annotation of SNPs and InDels identified between Four-row Wax and B73 maize.**

Variation types	Total	Intergenic	Genic	Intron	UTR <sup>1</sup>	CDS <sup>2</sup>	Non-Synonymous/Frame_shift	Synonymous/X3_shift
SNPs	3,252,194	2,664,891	587,303	314,146	119,588	153,569	71,198	82,371
InDels	213,181	149,254	63,927	42,556	17,569	3,802	2,371	1,431

<sup>1</sup>UTR is the untranslated region of exons. <sup>2</sup>CDS is coding DNA sequence, composed of exons coding for protein.

**TABLE 3 | Types and numbers of large effect SNPs identified between Four-row Wax and B73 maize.**

Types	Total number	Inducing premature stop codons	Abolishing stop codons	Altering initiation Met	Altering splice sites	Essential splice sites	Changing properties of amino acids
Non-synonymous of total SNPs	71,198	2,617 (3.68%)	681 (0.96%)	298 (0.42%)	2,769 (3.89%)	718 (1.01%)	48,523 (68.15%)
Non-synonymous of novel SNPs	23,590	1,129 (4.79%)	187 (0.79%)	76 (0.32%)	952 (4.04%)	358 (1.52%)	14,465 (61.32%)



2009), sorghum (Zheng et al., 2011), and *Arabidopsis* (Clark et al., 2007). It was reported that these kinases participated in responses to stresses such as drought, high or low temperature; and plant diseases and pests. To adapt to new environments and

improve resistance to pests, plant stress-resistance genes may commonly experience diversifying selection. Among novel large-effect SNPs in Four-row Wax, we identified enrichment in genes related to biotic and abiotic stress, such as pathogenesis-related

transcription factor, tyrosine protein kinase and armadillo-type fold. Novel SNPs in these Four-row Wax genes may result from the selection and adaptation to local environments.

In the Palomero landrace (Vielle-Calzada et al., 2009), genes involved in heavy-metal detoxification have low nucleotide variability, such as the cadmium transporter, a P-type copper translocator and a phosphatidylinositol transporter. A similar result was found in Four-row Wax maize, as proteins with metal ion binding activity were enriched in conservative genomic regions, without SNPs. On one hand, proteins binding metal ions generally possess transcriptional regulator activity or catalytic activity, and participate in basic metabolic processes (Berg, 1986). Meanwhile, these proteins may be involved in metal transport for heavy-metal detoxification. In partial summary, our results are similar to those based on resequencing of Mexican landrace Palomero.

Novel traits in Four-row Wax may also be caused by novel genes. We *de novo* assembled the genome sequence using  $21.87 \times$  coverage reads, though sequence from only one insert size library was not enough to fully assemble the genome. Only four novel maize genes were predicted in the limited assembled sequence, which did not exist in the reference genome and were not homologous to annotated genes. One gene is homologous to a rice gene, and the remaining three are homologous to sorghum genes. The four genes showed the same overall depth of sequence coverage as other annotated genes, making it unlikely that these reads were due to contaminant DNA. The four novel genes may be really present only in the genome of Four-row Wax but not in other maize lines, because the genome of maize is complex and diversifying, with PAVs well known among genomes (Lai et al., 2010). We cannot rule out the possibility that the four genes exist in B73 and other maize genomes but were not included in the current B73 assembly, or not identified in other maize lines by molecular biology experiments. However, the chance is very low.

## Potential Variations for Low KRN

Low KRN is a characteristic phenotypic trait of Four-row Wax, making this genotype a valuable resource to investigate KRN. The ancestor of maize, teosinte, makes two rows of kernels while modern varieties and inbred lines have 10–20 rows (Bommert et al., 2013). Fewer kernel rows leads to lower yield, thus KRN is an important trait that human selection has altered during domestication and genetic improvement. Many studies have identified QTLs for maize KRN (Lu et al., 2010; Brown et al., 2011; Cook et al., 2012) but only a few genes have been isolated by map-based cloning or mutant analysis (Vollbrecht et al., 1991; Bommert et al., 2005, 2013), and the molecular mechanism underlying this trait is largely unknown. QTL for inflorescence traits have been discovered by both joint linkage mapping and GWAS in a maize NAM population (Brown et al., 2011). QTL for grain yield components involved in early maize domestication have been found in a population of maize-teosinte BC2S3 RILs (Shannon, 2013), segregating for allelic variation not present in the NAM population. Based on the two maize populations, some QTL associated with KRN were identified. Additionally, a few genes affecting KRN and required

for shoot meristem establishment, maintenance and size, have been isolated (Supplementary Table S10). Among them, *fea1* and *fea2* cause massive over-proliferation of the ear inflorescence meristem and lead to increased KRN, but a mutation of *kn1* displayed small ears, with length, diameter and KRN smaller than normal plants (Bommert et al., 2005).

In the CDS of the four genes previously implicated in genetic control of KRN, *fea2*, *td1*, *kn1*, and *te1*, all SNPs in Four-row Wax were synonymous. However, one InDel caused loss of a glutamic acid residue in the C terminus of *te1* although the InDel is not in a conserved domain (Supplementary Table S11). Clarifying the function of the InDels in *te1* needs further molecular and physiological investigation *in vivo*. Bommert et al. (2013) sequenced the *fea2* locus and compared Mo17 and B73, finding no non-synonymous changes, although B73 had significantly higher KRN and a larger inflorescence meristem than Mo17. Hence, changes in the *fea2* protein sequence are not responsible for variation in KRN between B73 and Mo17. However, *fea2* expression from the B73 allele was threefold lower than that of the Mo17 allele as revealed by RT-PCR, likely due to regulatory changes in the *fea2* alleles. Likewise, we inferred that SNPs and InDels in the genic region of the gene may interact with additional genes together to alter the KRN. Li et al. (2012) reported that trait-associated SNPs were enriched in the non-genic regions, particularly within a 5-kb window upstream of genes. To get further information, we investigated variation within 5 kb upstream of the four genes in Four-row Wax (Supplementary Table S11), finding that *fea2* and *td1* had more variation than *kn1* and *te1*. Some of these variations may contribute to the low KRN of Four-row Wax.

Besides the above four genes, *ids1* (Chuck et al., 1998), indeterminate spikelet 1, specifies a determinate spikelet meristem fate and thereby limits the number of floral meristems produced. In the absence of *ids1* gene function, the spikelet meristem becomes indeterminate and produces additional florets. So KRN increased in the *ids1* mutation. However, in the annotation of the reference genome version 1 (4a.53), we could not find the gene, thus could not analyze variation in it.

Moreover, we compared DNA polymorphisms between Four-row Wax and six inbred lines in five QTLs related to KRN. SNPs, InDels, and SVs in these regions may be responsible for low KRN. Because the reported QTL regions delimited by linked molecular markers were large, with many differences among compared varieties, the causal DNA polymorphisms contributing to low KRN have not been identified. Some genes located in these regions are enriched in novel SNPs and identified as potential candidate genes for the low KRN of Four-row Wax, such as the auxin responsive gene of SAUR. It is of high priority to investigate the relationship between these genes and the KRN phenotype.

## CONCLUSION

In summary, we identified 3.3 M SNPs, 213.2 K InDels, and 39.6 K SVs distinguishing Four-row Wax, a waxy maize landrace with low KRN from southwestern China, from the B73 reference genome. This variety harbored novel SNPs (absent from the

present maize Hapmap), and novel genes were also identified by comparing to known maize variation or gene datasets. By analyzing the Four-row Wax variations in reported genes and QTLs related to KRN, candidate genes and mutations related to its low KRN phenotype were inferred. These data and results, combined with other variation and population studies of maize, may facilitate the genetic analysis and improvement of important agronomic traits of maize.

## AUTHOR CONTRIBUTIONS

YH and HL conceived and designed the experiments. YL, JL, and ZX contributed materials. BW, HL, JZ, and GY contributed to data analysis. YW and YH performed the experiments. HL and XW wrote and revised the manuscript. All authors discussed the results and commented on the manuscript.

## REFERENCES

- Arai-Kichise, Y., Shiwa, Y., Nagasaki, H., Ebana, K., Yoshikawa, H., Yano, M., et al. (2011). Discovery of genome-wide DNA polymorphisms in a landrace cultivar of Japonica rice by whole-genome sequencing. *Plant Cell Physiol.* 52, 274–282. doi: 10.1093/pcp/pcr003
- Berg, J. M. (1986). Potential metal-binding domains in nucleic acid binding proteins. *Science* 232, 485–487. doi: 10.1126/science.2421409
- Bommert, P., Lunde, C., Nardmann, J., Vollbrecht, E., Running, M., Jackson, D., et al. (2005). Thick tassel dwarf1 encodes a putative maize ortholog of the *Arabidopsis* CLAVATA1 leucine-rich repeat receptor-like kinase. *Development* 132, 1235–1245. doi: 10.1242/dev.01671
- Bommert, P., Nagasawa, N. S., and Jackson, D. (2013). Quantitative variation in maize kernel row number is controlled by the FASCIATED EAR2 locus. *Nat. Genet.* 45, 334–337. doi: 10.1038/ng.2534
- Brown, P. J., Upadaya, N., Mahone, G. S., Tian, F., Bradbury, P. J., Myles, S., et al. (2011). Distinct genetic architectures for male and female inflorescence traits of maize. *PLoS Genet.* 7:e1002383. doi: 10.1371/journal.pgen.1002383
- Chia, J. M., Song, C., Bradbury, P. J., Costich, D., De Leon, N., Doebley, J., et al. (2012). Maize HapMap2 identifies extant variation from a genome in flux. *Nat. Genet.* 44, 803–807. doi: 10.1038/ng.2313
- Chuck, G., Meeley, R. B., and Hake, S. (1998). The control of maize spikelet meristem fate by the APETALA2-like gene indeterminate spikelet1. *Genes Dev.* 12, 1145–1154. doi: 10.1101/gad.12.8.1145
- Clark, R. M., Schweikert, G., Toomajian, C., Ossowski, S., Zeller, G., Shinn, P., et al. (2007). Common sequence polymorphisms shaping genetic diversity in *Arabidopsis thaliana*. *Science* 317, 338–342. doi: 10.1126/science.1138632
- Cook, J. P., McMullen, M. D., Holland, J. B., Tian, F., Bradbury, P., Ross-Ibarra, J., et al. (2012). Genetic architecture of maize kernel composition in the nested association mapping and inbred association panels. *Plant Physiol.* 158, 824–834. doi: 10.1104/pp.111.185033
- Fan, L., Bao, J., Wang, Y., Yao, J., Gui, Y., Hu, W., et al. (2009). Post-domestication selection in the maize starch pathway. *PLoS ONE* 4:e7612. doi: 10.1371/journal.pone.0007612
- Fan, L., Quan, L., Leng, X., Guo, X., Hu, W., Ruan, S., et al. (2008). Molecular evidence for post-domestication selection in the Waxy gene of Chinese waxy maize. *Mol. Breed.* 22, 329–338. doi: 10.1007/s11032-008-9178-2
- Gore, M. A., Chia, J. M., Elshire, R. J., Sun, Q., Ersoz, E. S., Hurwitz, B. L., et al. (2009). A first-generation haplotype map of maize. *Science* 326, 1115–1117. doi: 10.1126/science.1177837
- Huang, X., Wei, X., Sang, T., Zhao, Q., Feng, Q., Zhao, Y., et al. (2010). Genome-wide association studies of 14 agronomic traits in rice landraces. *Nat. Genet.* 42, 961–967. doi: 10.1038/ng.695
- Hufford, M. B., Xu, X., Van Heerwaarden, J., Pyhajarvi, T., Chia, J. M., Cartwright, R. A., et al. (2012). Comparative population genomics of maize domestication and improvement. *Nat. Genet.* 44, 808–811. doi: 10.1038/ng.2309

## FUNDING

This work was supported by the National Program on Key Basic Research Project (973 Program; 2014CB138202) and the National Natural Science Foundation of China (91435114).

## ACKNOWLEDGMENT

We thank members of our research groups and collaborators for technical assistance and discussions.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://journal.frontiersin.org/article/10.3389/fpls.2016.00667>

- Jiao, Y., Zhao, H., Ren, L., Song, W., Zeng, B., Guo, J., et al. (2012). Genome-wide genetic changes during modern breeding of maize. *Nat. Genet.* 44, 812–815. doi: 10.1038/ng.2312
- Kant, S., Bi, Y. M., Zhu, T., and Rothstein, S. J. (2009). SAUR39, a small auxin-up RNA gene, acts as a negative regulator of auxin synthesis and transport in rice. *Plant Physiol.* 151, 691–701. doi: 10.1104/pp.109.143875
- Kant, S., and Rothstein, S. (2009). Auxin-responsive SAUR39 gene modulates auxin level in rice. *Plant Signal. Behav.* 4, 1174–1175. doi: 10.4161/psb.4.12.10043
- Lai, J., Li, R., Xu, X., Jin, W., Xu, M., Zhao, H., et al. (2010). Genome-wide patterns of genetic variation among elite maize inbred lines. *Nat. Genet.* 42, 1027–1030. doi: 10.1038/ng.684
- Li, R., Li, Y., Fang, X., Yang, H., Wang, J., Kristiansen, K., et al. (2009a). SNP detection for massively parallel whole-genome resequencing. *Genome Res.* 19, 1124–1132. doi: 10.1101/gr.088013.108
- Li, R., Yu, C., Li, Y., Lam, T. W., Yiu, S. M., Kristiansen, K., et al. (2009b). SOAP2: an improved ultrafast tool for short read alignment. *Bioinformatics* 25, 1966–1967. doi: 10.1093/bioinformatics/btp336
- Li, R., Zhu, H., Ruan, J., Qian, W., Fang, X., Shi, Z., et al. (2010). De novo assembly of human genomes with massively parallel short read sequencing. *Genome Res.* 20, 265–272. doi: 10.1101/gr.097261.109
- Li, S., Li, R., Li, H., Lu, J., Li, Y., Bolund, L., et al. (2013). SOAPindel: efficient identification of indels from short paired reads. *Genome Res.* 23, 195–200. doi: 10.1101/gr.132480.111
- Li, X., Zhu, C., Yeh, C. T., Wu, W., Takacs, E. M., Petsch, K. A., et al. (2012). Genic and nongenic contributions to natural variation of quantitative traits in maize. *Genome Res.* 22, 2436–2444. doi: 10.1101/gr.140277.112
- Li, Y. L., Niu, S. Z., Dong, Y. B., Cui, D. Q., Wang, Y. Z., Liu, Y. Y., et al. (2007). Identification of trait-improving quantitative trait loci for grain yield components from a dent corn inbred line in an advanced backcross BC2F2 population and comparison with its F2:3 population in popcorn. *Theor. Appl. Genet.* 115, 129–140. doi: 10.1007/s00122-007-0549-6
- Liu, H., Wang, X., Warburton, M. L., Wen, W., Jin, M., Deng, M., et al. (2015). Genomic, transcriptomic, and phenomic variation reveals the complex adaptation of modern maize breeding. *Mol. Plant.* 8, 871–884. doi: 10.1016/j.molp.2015.01.016
- Liu, L., Du, Y., Huo, D., Wang, M., Shen, X., Yue, B., et al. (2015). Genetic architecture of maize kernel row number and whole genome prediction. *Theor. Appl. Genet.* 128, 2243–2254.
- Lu, M., Xie, C.-X., Li, X.-H., Hao, Z.-F., Li, M.-S., Weng, J.-F., et al. (2010). Mapping of quantitative trait loci for kernel row number in maize across seven environments. *Mol. Breed.* 28, 143–152.
- Lunde, C., and Hake, S. (2009). The interaction of knotted1 and thick tassel dwarf1 in vegetative and reproductive meristems of maize. *Genetics* 181, 1693–1697. doi: 10.1534/genetics.108.098350

- Ma, X. Q., Tang, J. H., Teng, W. T., Yan, J. B., Meng, Y. J., and Li, J. S. (2007). Epistatic interaction is an important genetic basis of grain yield and its components in maize. *Mol. Breed.* 20, 41–51. doi: 10.1007/s11032-006-9071-9
- Matsuoka, Y., Mitchell, S. E., Kresovich, S., Goodman, M., and Doebley, J. (2002). Microsatellites in *Zea* - variability, patterns of mutations, and use for evolutionary studies. *Theor. Appl. Genet.* 104, 436–450. doi: 10.1007/s001220100694
- McMullen, M. D., Kresovich, S., Villeda, H. S., Bradbury, P., Li, H., Sun, Q., et al. (2009). Genetic properties of the maize nested association mapping population. *Science* 325, 737–740. doi: 10.1126/science.1174320
- McNally, K. L., Childs, K. L., Bohnert, R., Davidson, R. M., Zhao, K., Ulat, V. J., et al. (2009). Genomewide SNP variation reveals relationships among landraces and modern varieties of rice. *Proc. Natl. Acad. Sci. U.S.A.* 106, 12273–12278. doi: 10.1073/pnas.0900992106
- Romay, M. C., Millard, M. J., Glaubitz, J. C., Peiffer, J. A., Swarts, K. L., Casstevens, T. M., et al. (2013). Comprehensive genotyping of the USA national maize inbred seed bank. *Genome Biol.* 14:R55. doi: 10.1186/gb-2013-14-6-r55
- Schnable, P. S., Ware, D., Fulton, R. S., Stein, J. C., Wei, F., Pasternak, S., et al. (2009). The B73 maize genome: complexity, diversity, and dynamics. *Science* 326, 1112–1115. doi: 10.1126/science.1178534
- Shannon, L. M. (ed.) (2013). *The Genetic Architecture of Maize Domestication and Range Expansion*. Ph.D. thesis, University of Wisconsin-Madison, Madison, WI.
- Taguchi-Shiobara, F., Yuan, Z., Hake, S., and Jackson, D. (2001). The fasciated ear2 gene encodes a leucine-rich repeat receptor-like protein that regulates shoot meristem proliferation in maize. *Genes Dev.* 15, 2755–2766. doi: 10.1101/gad.208501
- Tian, M., Tan, G., Liu, Y., Rong, T., and Huang, Y. (2009). Origin and evolution of Chinese waxy maize: evidence from the Globulin-1 gene. *Genet. Resour. Crop Evol.* 56, 247–255. doi: 10.1007/s10722-008-9360-8
- Veit, B., Briggs, S. P., Schmidt, R. J., Yanofsky, M. F., and Hake, S. (1998). Regulation of leaf initiation by the terminal ear 1 gene of maize. *Nature* 393, 166–168. doi: 10.1038/30239
- Vielle-Calzada, J. P., Martinez De La Vega, O., Hernandez-Guzman, G., Ibarra-Laclette, E., Alvarez-Mejia, C., Vega-Arrequin, J. C., et al. (2009). The Palomero genome suggests metal effects on domestication. *Science* 326:1078. doi: 10.1126/science.1178437
- Vollbrecht, E., Veit, B., Sinha, N., and Hake, S. (1991). The developmental gene Knotted-1 is a member of a maize homeobox gene family. *Nature* 350, 241–243. doi: 10.1038/350241a0
- Wessler, S. R., and Varagona, M. J. (1985). Molecular basis of mutations at the waxy locus of maize: correlation with the fine structure genetic map. *Proc. Natl. Acad. Sci. U.S.A.* 82, 4177–4181. doi: 10.1073/pnas.82.12.4177
- Xu, J., Yuan, Y., Xu, Y., Zhang, G., Guo, X., Wu, F., et al. (2014). Identification of candidate genes for drought tolerance by whole-genome resequencing in maize. *BMC Plant Biol.* 14:83. doi: 10.1186/1471-2229-14-83
- Xu, X., Liu, X., Ge, S., Jensen, J. D., Hu, F., Li, X., et al. (2012). Resequencing 50 accessions of cultivated and wild rice yields markers for identifying agronomically important genes. *Nat. Biotechnol.* 30, 105–111. doi: 10.1038/nbt.2050
- Yan, J.-B., Tang, H., Huang, Y.-Q., Zheng, Y.-L., and Li, J.-S. (2006). Quantitative trait loci mapping and epistatic analysis for grain yield and yield components using molecular markers with an elite maize hybrid. *Euphytica* 149, 121–131.
- Zheng, H., Wang, H., Yang, H., Wu, J., Shi, B., Cai, R., et al. (2013). Genetic diversity and molecular evolution of Chinese waxy maize germplasm. *PLoS ONE* 8:e66606. doi: 10.1371/journal.pone.0066606
- Zheng, L. Y., Guo, X. S., He, B., Sun, L. J., Peng, Y., Dong, S. S., et al. (2011). Genome-wide patterns of genetic variation in sweet and grain sorghum (*Sorghum bicolor*). *Genome Biol.* 12:R114. doi: 10.1186/gb-2011-12-11-r114

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Liu, Wang, Wei, Wang, Liu, Zhang, Hu, Yu, Li, Xu and Huang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.