



Integrated Analysis of Large-Scale Omics Data Revealed Relationship Between Tissue Specificity and Evolutionary Dynamics of Small RNAs in Maize (*Zea mays*)

OPEN ACCESS

Yu Xu^{1†}, Ting Zhang^{1,2†}, Yuchen Li¹ and Zhenyan Miao^{1,2*}

Edited by:

Yin Li,
Rutgers, The State
University of New Jersey,
United States

Reviewed by:

Chong Teng,
Donald Danforth Plant Science Center,
United States
Jianhong Xu,
Zhejiang University,
China
Paul Fourounjian,
The State University of New Jersey,
Newark, United States

*Correspondence:

Zhenyan Miao
miaozhenyan@163.com

[†]These authors have contributed
equally to this work and share first
authorship

Specialty section:

This article was submitted to
Evolutionary and Population Genetics,
a section of the journal
Frontiers in Genetics

Received: 16 September 2019

Accepted: 16 January 2020

Published: 11 February 2020

Citation:

Xu Y, Zhang T, Li Y and Miao Z (2020)
Integrated Analysis of Large-Scale
Omics Data Revealed Relationship
Between Tissue Specificity and
Evolutionary Dynamics of Small RNAs
in Maize (*Zea mays*).
Front. Genet. 11:51.
doi: 10.3389/fgene.2020.00051

¹ State Key Laboratory of Crop Stress Biology for Arid Areas, College of Life Sciences, Northwest A&F University, Yangling, China, ² Center of Bioinformatics, College of Life Sciences, Northwest A&F University, Yangling, China

The evolutionary dynamics and tissue specificity of protein-coding genes are well documented in plants. However, the evolutionary consequences of small RNAs (sRNAs) on tissue-specific functions remain poorly understood. Here, we performed integrated analysis of 195 deeply sequenced sRNA libraries of maize B73, representing more than 10 tissues, and identified a comprehensive list of 419 maize microRNA (miRNA) genes, 271 of which were newly discovered in this study. We further characterized the evolutionary dynamics and tissue specificity of miRNA genes and corresponding miRNA isoforms (isomiRs). Our analysis revealed that tissue specificity of isomiR events tends to be associated with miRNA gene abundance and suggested that the frequencies of isomiR types are affected by the local genomic regions. Moreover, genome duplication (GD) events have dramatic effect on evolutionary dynamics of maize miRNA genes, and the abundance divergence for tissue-specific miRNA genes is associated with GD events. Further study indicated that duplicate miRNA genes with tissue-specific expression patterns, such as *miR2275a*, a phased siRNA (phasiRNA) trigger, contribute to phenotypic traits in maize. Additionally, our study revealed the expression preference of 21- and 24-nt phasiRNAs in relation to tissue specificity. This large-scale sRNAomic study depicted evolutionary implications of tissue-specific maize sRNAs, which coordinate genome duplication, isomiR modification, phenotypic traits and phasiRNAs differentiation.

Keywords: maize, microRNA, phasiRNAs, tissue specificity, genome duplication

INTRODUCTION

Plants are multicellular eukaryotes with diverse tissues comprising various cell types, which carry out common processes essential for survival. However, within the physical context of the tissue environment, cells also exhibit unique functions that help define tissue-specific phenotypes (Edwards and Coruzzi, 1990). These common and tissue-specific processes are ultimately controlled by gene regulatory networks that alter the extent of gene expression. The tissue

specificity of these processes is often described based on the expression level of protein-coding genes (PCGs). These analyses only partially capture the variety of processes that distinguish different tissues, due to the ignorance of other regulatory elements, such as small RNAs (sRNAs). Plant sRNAs include microRNAs (miRNAs) and small interfering RNAs (siRNAs), which are typically 20–24 nucleotides (nt) in length. These sRNAs regulate the expression of PCGs involved in a variety of biological processes associated with growth, development, and stress responses in plants (Meyers et al., 2008). MiRNAs also play roles in triggering the biogenesis of secondary siRNAs, termed epigenetically activated siRNAs (easiRNAs), *trans*-acting siRNAs (ta-siRNAs) or phased siRNAs (phasiRNAs) (Allen et al., 2005; Creasey et al., 2014). As miRNAs, phasiRNAs function in a homology-dependent manner to suppress the expression of their targets (Fei et al., 2013). The number of phasiRNA-generating loci (*PHAS* loci) varies substantially among species, ranging from over 800 in wild rice (*Oryza rufipogon*) (Liu et al., 2013b) to less than 30 in *Arabidopsis thaliana* (Columbia-0 ecotype) (Fei et al., 2013). Therefore, the investigation of tissue-specific miRNAs and phasiRNAs will help to understand the regulatory mechanisms of tissue specific phenotypes, thereby facilitating crop improvement via genetic engineering and precision breeding.

Maize (*Zea mays*) is not only one of the most economically important crops but also a model monocot plant for genetic and genomic studies. A growing body of molecular genetics and transgenic studies support the role of sRNAs in regulating diverse agronomic traits in maize (Li et al., 2017; Tang and Chu, 2017). However, the large-scale functional characterization of sRNAs is still considerably lagging in maize compared with the progress made in the model species *Arabidopsis*. In the field of bioinformatics and computational biology, there are still plenty of obstacles to further progress. One of those is the deficiency of comprehensive and high-quality annotation of miRNA genes in maize. To date, only 202 distinct mature miRNAs encoded by 169 miRNA genes in maize have been reported in miRBase (release 22) (Kozomara and Griffiths-Jones, 2014). The number of annotated miRNAs in maize is unexpectedly lower than that in other model plants, such as *Arabidopsis* (349 miRNA genes), soybean (*Glycine max*; 482 miRNA genes), and rice (553 miRNA genes) reported in miRBase. Currently, more than 100 maize sRNA-seq datasets are available in the National Center for Biotechnology Information (NCBI) database. These sRNA-seq datasets are of great value in investigating the expression profiles

and potential function of miRNAs. However, these datasets were generated from different laboratories worldwide, which have different sequencing qualities and represent miRNAs in different tissues. Given that the incomplete maize reference genome and inappropriate approaches used to identify maize miRNAs ranging from prediction to validation of target cleavage (Jiao et al., 2017; Axtell and Meyers, 2018), it is inevitable that some miRNAs remain to be discovered. Thus, an integrated analysis of sRNA-seq datasets will be very helpful in enriching the set of annotated maize miRNAs, exploring tissue-specific miRNA genes, and uncovering the expression of different miRNA variants.

The complexity of maize genome is another major obstacle. Maize was domesticated from its wild progenitor species, teosinte (*Zea mexicana*), ~10,000 years ago (Doebley, 2004). It has been proposed that the *Zea* lineage has undergone several rounds of genomic duplication (GDs) events, such as whole genome duplications (WGDs), large-scale segmental duplications (SDs), and tandem duplications (TDs), after its divergence from the lineage that gave rise to sorghum (*Sorghum bicolor*) ~5–12 million years ago (Schnable et al., 2011). GDs are a common phenomenon in plant genome evolution, resulting in the expansion and diversification of many gene families. Post-duplication, many genes have been retained in the genome as paralogous pairs, and individual genes comprising the pair have been either sub-functionalized (partitioning and sharing the original gene function) and/or neo-functionalized (gaining novel functions) via sequence and/or expression divergence (Freeling, 2009). Studies in different species have revealed that ancient miRNA gene families have evolved via several of GD events (Baldrich et al., 2018). For example, the miR166 family contains seven members in *Arabidopsis*. These miRNA members resulted from WGD, SD, and TD, followed by tissue-specific sub-functionalization, as well as the deletion of one of the members (Maher et al., 2006). Similarly in soybean, post-WGD, the majority of miRNA gene singletons originated from rapid decay, and the retained miRNA gene duplicates evolved slower than singletons, regardless of whether they originated from WGD or TDs (Zhao et al., 2015). With the availability of the maize B73 reference genome sequence and genome resequencing data from dozens of maize accessions, this grass crop was made more accessible and has become an important model system suitable for investigation of the evolutionary dynamics and consequences of recurrent GD events.

In this study, we attempted to investigate evolutionary implications of tissue-specific sRNA genes in maize. To achieve this, we broadly collected 195 maize sRNA-seq datasets across 14 tissues and performed an integrated analysis to build a comprehensive list of maize miRNA and phasiRNA genes. A genome-wide study on maize miRNA genes was performed to examine isomiR modification, tissue specificity, and duplication status. We also investigated agronomic traits associated with tissue-specific miRNA genes and selected by artificial improvement. Furthermore, we dissected differentiation of phasiRNAs in relation to tissue specificity.

Abbreviations: DAP, days after pollination; DTA, days to anthesis; DTS, days to silk production; DCL1, DICER-LIKE1; DEMs, differentially expressed microRNAs; DS, drought-stressed; easiRNAs, epigenetically activated small interfering RNAs; GD, genome duplication; GWAS, genome-wide association study; miRNAs, microRNAs; isomiRs, microRNA isoforms; NCBI, National Center for Biotechnology Information; nt, nucleotide; phasiRNAs, phased small interfering RNAs; *PHAS* loci, phasiRNA-generating loci; PCGs, protein-coding genes; SD, segmental duplication; SRA, Sequence Read Archive; SNPs, single nucleotide polymorphisms; sRNAs, small RNAs; siRNAs, small interfering RNAs; TD, tandem duplication; ta-siRNAs, *trans*-acting small interfering RNAs; TPM, transcript per million; TE, transposable element; UIs, unclassified intergenic sequences; WGD, whole genome duplication; WW, well-watered.

MATERIALS AND METHODS

Libraries Construction for High-Throughput Sequencing

Maize B73 inbred lines were grown in pots under a controlled growth chamber (28 °C day / 26 °C night, 14 h light / 10 h dark). Well-watered (WW) seedlings were ensured a normal water supply with 80% of soil field moisture capacity. The drought-stressed (DS) seedlings were subjected to progressive stress by withholding water, and the soil relative water content were maintained at 40% of soil field moisture capacity. The two experimental samples (three biological replicates for each) were collected when three fully expanded leaves appeared, and immediately frozen in liquid nitrogen, and stored at -80 °C for RNA extraction.

Total RNA for small RNA sequencing was extracted using TRIzol reagent (Invitrogen, Carlsbad, USA). Total RNA was separated through 17% denaturing polyacrylamide gels and small RNAs between 10- and 60-nt were collected. Then, 5' and 3' RNA adaptors were ligated to small RNAs and followed by reverse transcription to produce cDNAs. These cDNAs were subsequently amplified by PCR and subjected to Illumina sequencing by Novogene Company (<http://www.novogene.com/>).

Computational Analysis of sRNA-Seq Data and miRNA Annotation

A total of 195 maize sRNA-seq libraries (189 downloaded from NCBI database and six generated in this study) were examined in this study (**Supplementary Table 1**). These libraries were constructed from sRNAs extracted from different tissues, including seedlings, leaves, ears, silks, tassels, anthers, pollen, seeds, embryos, endosperm, stalk, shoot apical meristem, early-prophase meiocytes, and roots at different developmental stages. To detect miRNAs, sRNA-seq datasets were analyzed using a bioinformatics pipeline referring to the recently updated pipeline (Axtell and Meyers, 2018), with slight modifications for the maize genome. Raw miRNA reads were normalized to transcripts per million (TPM) by multiplying a factor of 1,000,000 divided by the total number of mapped reads. For samples with two or more replications, final TPM values were means of all replications in corresponding samples. One mismatch between the genome and sRNA sequence reads was allowed, and structural RNAs (transfer RNAs, ribosomal RNAs, small nuclear RNAs, and small nucleolar RNAs), low abundance RNAs (total abundance ≤ 100 TPM and abundance in at least one library ≤ 10 TPM), RNA with irregular sizes (retaining 18–26 nt), and highly repetitive sRNAs (hit number on genome > 20) were removed. The modified script of miREAP software (Jeong et al., 2011) was used to predict miRNA precursors and potential pairing of miRNA and miRNA*, and miRNA precursors longer than 300 nt were discarded. For each miRNA precursor, the most abundant miRNA was regarded as canonical miRNA. Variants of canonical miRNAs (isomiRs) were identified using Jasmine pipeline (Zhong et al., 2019) and classified into 5' isomiR, 3' isomiR, and polymorphic isomiRs. To distinguish

between miRNAs and possible siRNAs, precursor sequences with strong bias (strand bias ≥ 0.9 and abundance bias ≥ 0.75) were selected, and stem-loop structures were predicted using CentroidFold (Sato et al., 2009). Strand bias was calculated by dividing the number of reads from the sense strand with the sum of the number of reads from both strands. Abundance bias was calculated as the ratio of miRNA, miRNA* and their isomiRs to the number of reads from the sense strand. Precursor sequences that do not have a typical stem-loop structure (mismatch between the miRNA and miRNA* no more than 5 and asymmetrical bulge number no more than 3) were removed. Identified miRNA precursors were divided into previously annotated precursors and novel precursors by BLAST searching against maize precursors in miRBase v22 (<http://www.mirbase.org/ftp.shtml>). Maize PCGs and transposable elements (TEs) were downloaded from Gramene database (http://ensembl.gramene.org/Zea_mays).

Identification of Differentially Expressed miRNAs

Raw miRNA reads were normalized to TPM by multiplying a factor of 1,000,000 divided by the total number of mapped reads. Final TPM values were means of three replications. The significance level of differential expression (*P*-value) was determined by the following equation (Audic and Claverie, 1997):

$$P(x|y) = \binom{N_2}{N_1} \frac{(x+y)!}{x!y!(1+\frac{N_2}{N_1})x+y+1}$$

In this equation, N_1 and N_2 were replaced with 1,000,000; x and y represent TPM values under well-watered and drought-stressed conditions, respectively.

Identification of Tissue-Specific miRNA Genes

112 of 195 sRNA-seq libraries representing 14 specific tissues and non-stressed conditions were used to define tissue-specific miRNA genes. Shannon entropy and *Z* score methods were combined to identify miRNA genes specifically expressed in a single tissue. Shannon entropy is used to measure the concentration ratio of gene expression levels in different samples and *Z* score is used to detect outliers. Considering an expression vector x for a miRNA gene, expression vectors x_1, x_2, \dots, x_n for n tissues, and an observation x_i for tissue I , the entropy (H) of miRNA genes was calculated as:

$$H = -\sum_{i=1}^n E_i \log_2 E_i$$

where, E_i is the relative expression of miRNA x for tissue i and is defined as:

$$E_i = x_i / \sum_{i=1}^n x_i$$

modH is calculated after using a one-step Tukey's biweight to improve robustness of the expression data as described previously (Kadota et al., 2006). The *Z* score is calculated as:

$$Z_i = (x_i - \mu) / \sigma$$

where, μ is the average expression of miRNA genes x_1, x_2, \dots, x_n in all tissues, and σ is the standard deviation. The quantity of the Z score represents the distance between the raw expression and average expression. For tissue-specific miRNAs, we used Shannon entropy $\text{modH} < 1.8$ and the maximum Z score ($Z_{\text{max}} > 3$).

Identification of PHAS Loci

The PHASIS software (minDepth was set as 3) (Kakrana et al., 2017) was used to identify PHAS loci from 195 sRNA-seq datasets. Candidate PHAS loci were selected using $P < 1 \times 10^{-5}$. The summarization of PHAS loci was generated using *phasmerge* component. The triggers were identified using *phastrigs* component. Recent studies have shown that preference for production of easiRNAs versus other secondary siRNAs (which include phasiRNAs) may in part be the result of mono-uridylation of 22-nt miRNAs such as miR170 and miR171a which triggers the production of easiRNAs (Zhai et al., 2013; Tu et al., 2015). The defining characteristic that sets easiRNAs apart from other secondary siRNAs is that they arise from transcriptionally active retrotransposons and function in the RNA-directed DNA methylation pathway (Creasey et al., 2014). Therefore, the PHAS loci originated from retrotransposons and triggered by 22-nt miR170 and miR171a were excluded from our analyses.

Chromosomal Location and Duplication Analysis of miRNA Genes

The chromosomal distribution of miRNA genes was visualized using MapChart software (Voorrips, 2002). Syntenic blocks were analyzed with the CoGe Synmap program using ZmB73 gramene v4.36 masked coding sequence with default parameters (<https://genomeevolution.org/coge/SynMap.pl>). The miRNA genes with no more than two mismatches with the genome were considered as conserved. PHAS loci with the same miRNA triggers and at least two pairs of similar phasiRNAs were considered as conserved. Conserved sequence pairs belonging to syntenic regions were defined as syntenic duplication. Conserved sequences with less than 200 kb between them were considered as tandem duplicates (Holub, 2001).

Identification of miRNA Genes Potentially Related to Agronomic Traits

Datasets from genome-wide association study (GWAS) of 35 agronomic traits in 10 different populations were downloaded (Supplementary Table 2) and analyzed using the compressed mixed linear model with the R package GAPIT (Lipka et al., 2012). For each population, single nucleotide polymorphisms (SNPs) were filtered using linkage disequilibrium-based variant pruner implemented in PLINK (Purcell et al., 2007), with the parameters as “-indep-pairwise 1000 100 0.2 -geno 0.1 -maf 0.05”. The parameters “-geno” and “-maf” were used for controlling missing rate and minor allele frequency (MAF), respectively. In addition, the first three principal components (“Q” matrix) were also used to control

population structure. All SNPs with $P < 1 \times 10^{-5}$ were considered as trait-related SNPs.

Expression Analysis of the miRNA Using Stem-Loop qRT-PCR

Maize B73 plants were cultivated in a phytotron at 25 °C with 65% relative humidity under a 14-h/10-h light/dark cycle. Tissues at the V5 stage (leaves and roots), R1 stage (tassels, ears, anther, pollen, and silks), and 20 days after pollination (DAP) stage (seeds, embryos, and endosperms) were collected and immediately frozen in liquid nitrogen. Samples from at least five plants were pooled for each biological replicate, and three biological replicates were performed. Total RNA was extracted following above described methods. To confirm the miRNA expression data, stem-loop qRT-PCR was performed referring to previously described methods (Chen et al., 2005). Typically, 200ng of RNA was digested by DNase I (Thermo scientific, USA) and reverse transcribed using a TaKaRa Mir-X miRNA First-Strand Synthesis Kit (Cat. No.638313). The stem-loop RT-PCR was performed on the ABI7500 Real-Time System (Applied Biosystems, USA) using EvaGreen 2 × qPCR MasterMix-ROX (Abm, Canada). The reactions were incubated in a 96-well plate at 55 °C for 2 min, 95 °C for 10 min, followed by 40 cycles of 95 °C for 30 s and 60 °C for 1 min. All reactions were assayed in three biological replicates, and 18S rRNA was used as the internal control for stem-loop RT-PCR. Normalized expression levels were calculated as previously described methods (Schmittgen and Livak, 2008). Primers used in the qRT-PCR are listed in Supplementary Table 3.

Identification of miRNA Target

Potential target transcripts of miRNAs were predicted using psRNATarget (2017 update; <http://plantgrn.noble.org/psRNATarget>) (Dai and Zhao, 2011), psRobot (v1.2; <http://omicslab.genetics.ac.cn/psRobot/downloads.php>) (Wu et al., 2012), and TargetFinder (<https://github.com/carringtonlab/TargetFinder>) (Fahlgren and Carrington, 2010) with default parameters, except for the maximum expectation of five in psRNATarget. Predicted sRNA targets were verified with degradome sequencing data of maize B73. The raw degradome sequencing datasets were downloaded from the NCBI's Sequence Read Archive (SRA) database (SRR768486, SRR768488, SRR768490, SRR768493, SRR895786, SRR895787, SRR895788, SRR895789, SRR1028862, SRR1028864, SRR1028865, SRR1028866, SRR2917866, SRR4183498, SRR4183499, SRR3347501, SRR3348073, SRR3470769, and SRR3470770). Raw reads were preprocessed using the FASTX toolkit (http://hannonlab.cshl.edu/fastx_toolkit) to remove adaptors, low-quality bases (score <20), and short reads (<18 nt). Clean reads and maize cDNA sequences were input into CleaveLand4 (setting: -p 0.05) (Addo-Quaye et al., 2009) to identify potential cleavage sites.

Statistical Analysis

The Student's t -test was performed using *t.test* function in R package. The χ^2 test was performed using *chisq.test* function in R package.

RESULTS

Genome-Wide Identification of Maize miRNA Genes and Their Mature Sequences

In this study, we developed a bioinformatics pipeline to identify maize miRNAs and their mature sequences referring to the recently updated pipeline (Axtell and Meyers, 2018) (Figure 1). This miRNA identification pipeline was applied to process ~3.83 billion sequencing reads from 195 maize sRNA-seq datasets covering diverse tissues and conditions (Supplementary Table 1). As a result, we identified a total of 419 maize miRNA genes (Supplementary Table 4), consisting of 148 annotated miRNA genes in miRBase v22, as well as 271 miRNA genes newly discovered in this study. Like the miRNA genes present in miRBase, these novel miRNA genes mapped throughout the maize B73 reference genome with an uneven distribution on 10 chromosomes, and most of these genes were in regions of chromosome arms (Figure 2A). The sequence structures of 419 maize miRNA genes were visualized in Supplementary Data 1. DICER-LIKE1 (DCL1) is one of central components of the plant miRNA biogenesis pathway. By analyzing previously published *dcl1* mutants sequencing data (Petsch et al., 2015), we found the expression levels of both known and novel miRNAs were reduced in mutation samples (Supplementary Figure 1). Of the 419 miRNA

genes, 125 (29.8%) were located within PCGs, 125 (29.8%) were surrounded by TEs, and 196 (46.8%) were located within unclassified intergenic sequences (UIs; Supplementary Table 5).

Growing evidence suggested that a single miRNA gene can generate multiple mature isomiRs that differ in their length and/or sequence composition (Zhai et al., 2013). From the 419 miRNA genes, we identified 20,315 isomiRs, which can be categorized into seven classes: nucleotide substitutions within the seed sequence (seed SNP; 41.1%), nucleotide substitutions within tail sequence (tail SNP; 32.8%), 3' trimming (10.7%), 5' trimming (8.1%), 5' addition (7.5%), 3' addition (6.8%), and non-templated 3' addition (3' nt addition; 3.1%) (Figure 2B). Biased nucleotide composition was observed at the ends of isomiRs. For instance, the 5' and 3' ends of truncated isomiRs were likely to be 'G' and 'C' nucleotides, respectively. By Contrast, nucleotides on the 3' end of non-template added isomiRs were generally 'A' and 'U' (Supplementary Figure 2).

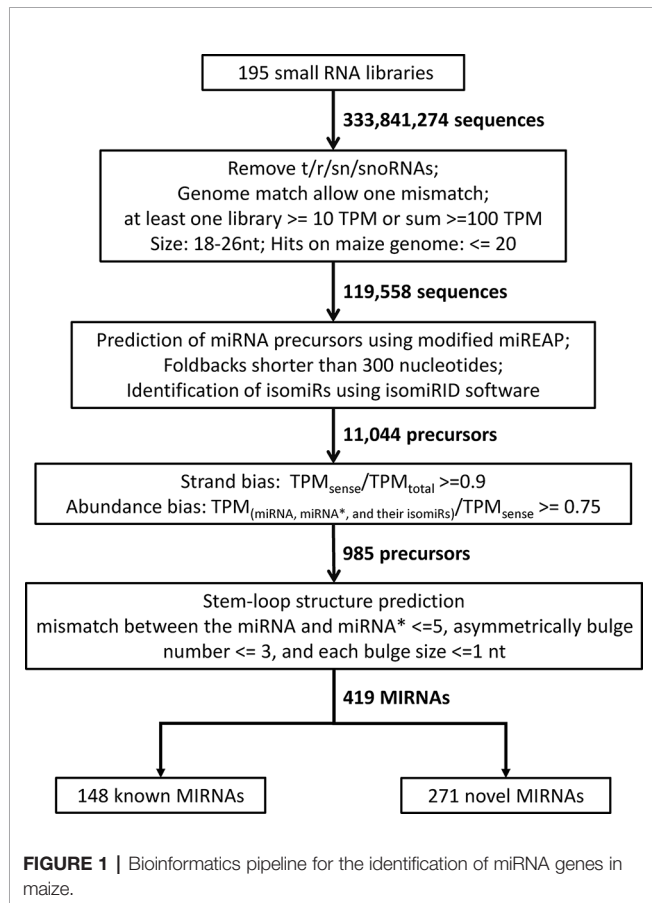
Together, the large-scale analysis of sRNA-seq datasets provides a more comprehensive profiling of maize miRNA genes and the corresponding isomiRs (Supplementary Table 6). An overview of the maize sRNA profiling supported by JBrowse (Buels et al., 2016) can be accessed in the Maize sRNA Data Browser, which is publicly available at <http://bioinfo.nwafu.edu.cn/MSDB/index.html>.

Survey of Differential Expressed miRNAs in Maize Response to Drought Stress

To study the expression dynamics of miRNAs and their potential roles in gene expression regulation in maize drought responses, we constructed and sequenced six sRNA libraries from WW and DS maize B73 inbred lines. Of the 419 miRNAs genes, 43 differentially expressed miRNAs (DEMs) induced by drought stress were detected (Supplementary Table 7). To understand the potential regulatory roles of drought stress-associated miRNAs, we identified the target genes of DEMs using three bioinformatics programs applying different alignment algorithms, then further validated by using sequencing datasets from 19 publicly available PARE libraries of maize B73. Of the 43 DEMs, 24 were found to target 30 transcripts (Supplementary Table 8). These include 13 transcription factors, such as MYB domain-containing, NAC domain-containing, Scarecrow like, and squamosa promoter binding, as well as 17 other functional genes, such as ubiquitin, zinc ion binding, and Syg1/Pho81/XPRI.

Ninety-Four miRNA Genes Show Highly Tissue-Specific Expression Patterns

A tissue-specific miRNA gene is defined as one that is highly expressed in a specific tissue. Based on the expression profile of maize sRNAs in 14 specific tissues under normal conditions, we identified a total of 94 miRNA genes (34 previously annotated and 60 newly discovered miRNA genes) using Shannon entropy and Z score. These genes were specifically expressed in nine tissues (Figure 3 and Supplementary Figure 3). We noticed that endosperm, pollen, and leaf tissues required relatively more specific miRNA genes than other tissues. Some miRNA gene families (e.g. *miR164* and *miR169*) were predominantly



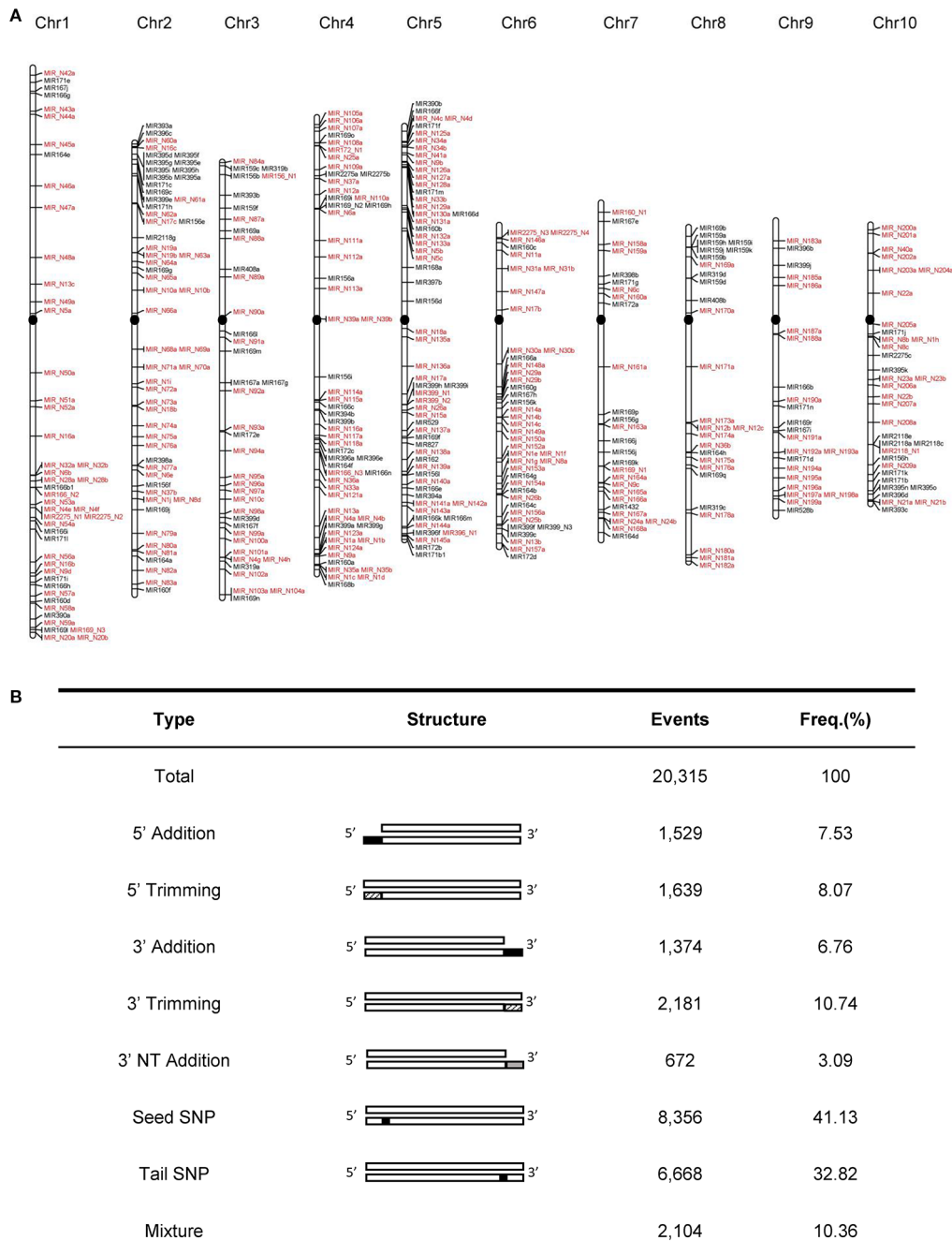


FIGURE 2 | Genome-wide characteristics of maize miRNA genes. **(A)** Chromosomal distribution of 419 maize miRNA genes. Newly identified miRNA genes were labeled as red. The black filled circles represent centromeres **(B)** Statistics of different isomiR types in the maize genome.

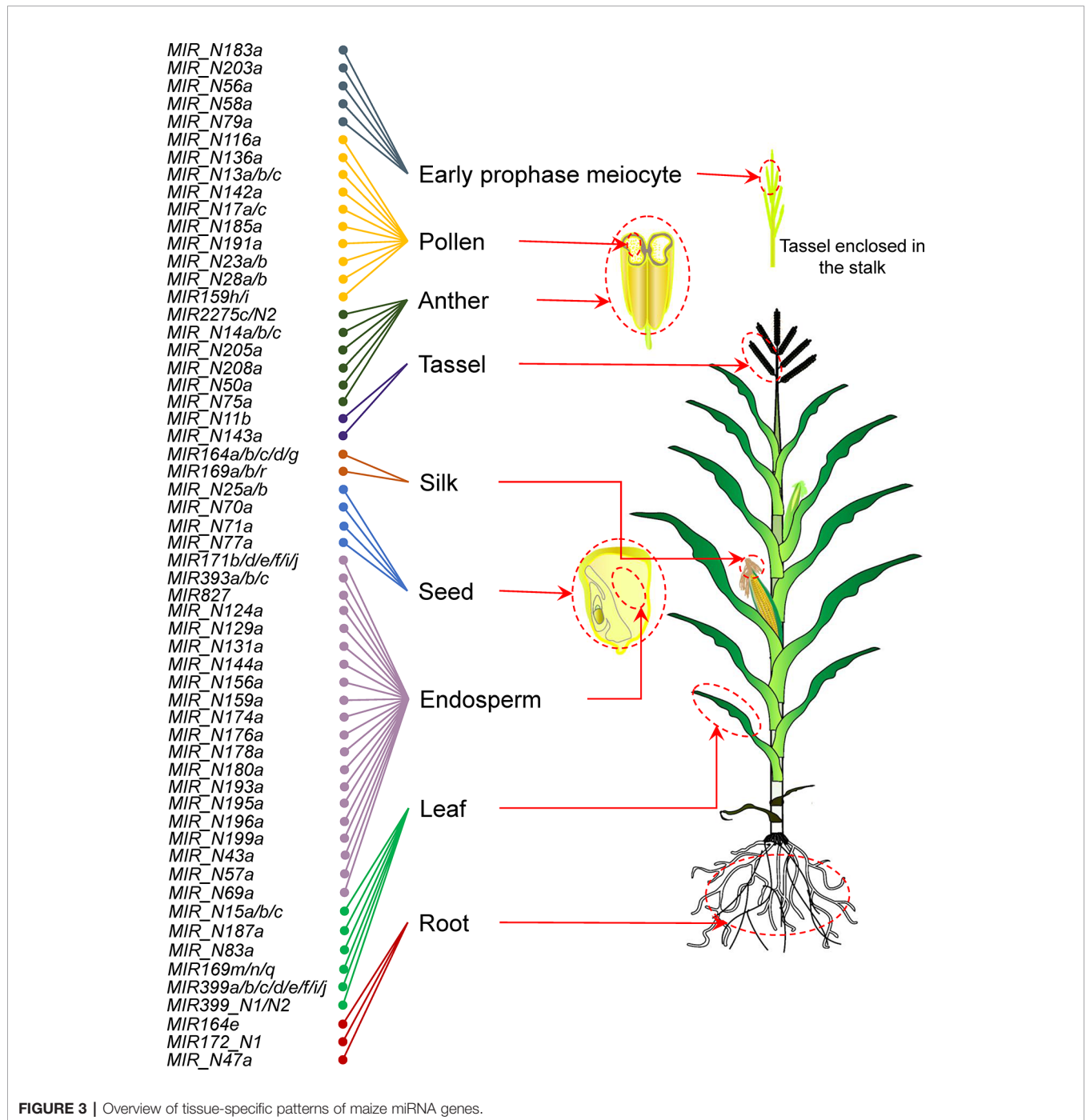
expressed in multiple tissues. For example, *miR164e* was specifically expressed in root, whereas *miR164a*, *miR164b*, *miR164c*, *miR164d*, and *miR164g* showed silk-specific expression. Similarly, *miR169m*, *miR169n*, and *miR169q* were preferentially expressed in the leaf, whereas *miR169a*, *miR169b*, and *miR169r* were preferentially expressed in the silk. Then, we

performed qRT-PCR assay to further confirm the expression levels of nine miRNAs. As shown in **Supplementary Figure 4**, the results of qRT-PCR assay are consistent with those of sRNA-seq. These results indicated that tissue-specific miRNA gene family members, which differ by only a few nucleotides may have redundant or diverse functions.

Dynamic Patterns of Isomir Events in Terms of Tissue Types and Genomic Environments

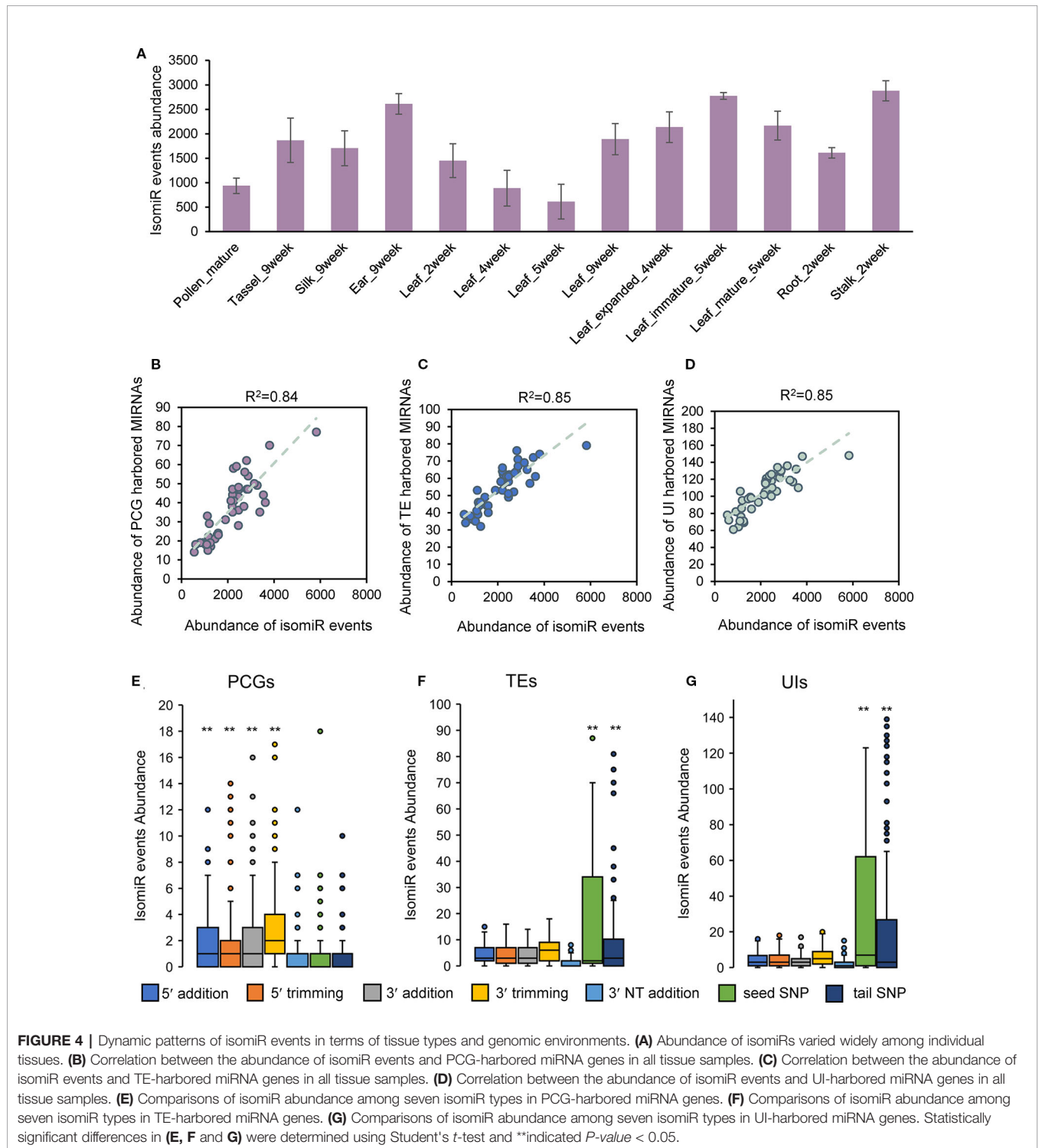
Considering the tissue specificity of miRNA expression and regulation, we are interested that whether isomiRs exhibit distinctive properties in different tissues. Statistical analysis showed that the abundance of isomiRs varied widely among tissues (Figure 4A). There were more isomiRs events in the ear and stalk tissues, and the abundance of isomiRs varied during

development of leaf tissue (Figure 4A). The composition of seven isomiR types changed remarkably among the individual tissues (Supplementary Figure 5). Nucleotide bias was also observed at the 5' and 3' ends of isomiRs among tissues (Supplementary Figure 6). Correlation analysis using data from all individual tissue samples showed that the abundance of isomiRs were highly correlated with the abundance of expressed miRNA genes located in PCGs, TEs, and UIs (Figures 4B–D). Comparisons of isomiR abundance among



seven isomiR types were conducted in three categories of miRNA genes (PCGs, TEs, and UIs) separately. For miRNA genes located in PCGs, 5' and 3' isomiR types (5' addition, 5' trimming, 3' addition and 3' trimming) exhibited significantly higher abundance than polymorphic isomiR types (seed SNP and tail SNP; **Figure 4E**). In contrast, the abundance of 5' and

3' isomiR types were significantly lower for miRNA genes harbored by TEs and UIs (**Figures 4F, G**). These observations indicated that tissue specificity of isomiR events tends to be associated with miRNA gene abundance and suggested that the frequencies of isomiR types are affected by the local genomic regions.



Evolutionary Implications of Duplicated miRNA Genes on Tissue Development

It has been proposed that maize underwent a series of GD events after its divergence from sorghum ~5–12 million years ago (Swigonova et al., 2004; Schnable et al., 2011). GDs are recognized as essential sources of gene family expansion and functional specialization (Panchy et al., 2016). Due to evolutionary dynamics of duplicated genes, one copy of many duplicated gene pairs in maize was eliminated, leaving the other one as a singleton after the genomic duplication events. Therefore, the singleton-to-duplicate ratio represents the degree of fragmentation of duplicated genomes. To understand the evolutionary process and consequences of tissue-specific miRNA genes after GD, based on the newly annotated set of PCGs in B73 reference genome (Jiao et al., 2017), we reanalyzed the large GDs encompassing the 419 miRNA genes and found that these genes can be classified into 305 singletons and 114 duplicates, with a ratio of 2.68:1 (Table 1). For miRNAs located in PCGs, TEs, and UIs, the corresponding singleton-to-duplicate ratio were 8.62:1, 2.79:1, and 1.76:1, respectively (Table 1). The chi-square "goodness-of-fit" test showed that the singleton-to-duplicate ratios for miRNA genes in PCGs and UIs were significantly different from that for miRNA genes in the whole maize genome (Table 1). In addition, the average abundance of miRNA genes in PCGs was the lowest compared with those in TEs and UIs, and the genome-wide miRNA duplicates showed relatively higher expression than singletons (Supplementary Table 9). These analyses suggested that the duplication status of miRNA genes was affected by their locations in the genome.

For the 94 tissue-specific miRNA genes, 72 were singletons and 22 were duplicates, corresponding to a singleton-to-duplicate ratio of 3.27:1, which was significantly higher than that in the whole maize genome (2.68:1; χ^2 test, $P < 0.05$). This indicated an overrepresentation of singletons compared with duplicates in tissue-specific miRNA genes. To further investigate if and how GD contributed to the tissue specificity in maize, we compared the relative proportions of miRNA singletons and duplicates expressed in individual tissue samples. The ratios of expressed miRNA singletons to duplicates (ranging from 0.82:1 to 2.28:1) in all tissues, except endosperm, were significantly lower than that in the whole maize genome (χ^2 test, $P < 0.05$) (Table 2). With leaf development, the ratio of expressed miRNA singletons to duplicates decreased after 2 weeks after germination and then increased after 6 weeks, however, the opposite trend was observed during silk development. Moreover, the singleton-

to-duplicate ratio decreased during ear development, and the similar trend was also observed in the tassel (Table 2). We also compared the ratios of expressed miRNA singletons to duplicates located in the three categories of genomic compartments in each individual tissue sample. In general, among the three categories of miRNA genes, the singleton-to-duplicate ratio was the lowest in UIs and the highest in PCGs in most tissues (Supplementary Table 10). Collectively, these observations revealed extensive divergence in the duplication status of miRNA genes among tissues.

Evolution of Tissue-Specific miRNA Genes Contributed to Phenotypic Traits in Maize

Maize was domesticated from its wild relative teosinte approximately 10,000 years ago and was subsequently subjected to intensive breeding efforts to improve its adaptation to modern agricultural practices, resulting in dramatic morphological and physiological modifications (Lyu, 2017). Recent studies demonstrated that SNPs that occurred at miRNA-related regions (pre-miRNAs and mature miRNAs) may affect miRNA biogenesis and function and cause serious phenotype changes (Sun et al., 2009; Hung et al., 2012; Liu et al., 2013a). Therefore, we investigated whether miRNA loci and quantitative trait loci were simultaneously selected by linkage drag, or whether miRNA loci have effects on tissue-specific traits. To address this question, we collected 10 representative maize B73 populations comprising and integrated the results of GWAS on the genetic diversity in various traits, such as ear height, leaf width, days to anthesis (DTA), days to silk production (DTS), and plant height (Supplementary Table 2).

We detected 13 trait-associated mutation sites co-localizing with seven miRNA loci (Supplementary Figure 7 and Figure 5A). Of these seven miRNA loci, four loci were associated with at least two traits. An interesting example is *miR2275a*, which belongs to the anther-specific *miR2275* family and is associated with DTA and DTS (Figure 5A). According to our previous analyses of duplicated miRNA loci in maize, we observed that *miR2275_N1/miR2275_N2*, *miR2275_N3/miR2275_N4*, and *miR2275a/miR2275b* were three homoeologous pairs formed by TD (Figure 5B). These results promoted us to try to understand the process of selection of *miR2275* family members during evolution. Divergencies of the dominant arms were observed among the three homoeologous pairs (Figure 5C). Further analysis revealed that *miR2275c* co-localized with a helitron transposon, implying that TE-mediated mechanisms were responsible for the amplification of this miRNA gene family. To shed light on how genomic duplication affects the functional diversity of these homoeologous pairs, we analyzed the expression levels of seven *miR2275* loci in individual tissues. All *miR2275* loci were highly abundant in early-prophase meiocytes and anthers (Figure 5D). These observations, together with the association analysis, indicated that although these miRNA loci dominated in anthers, the evolutionary consequences of the *miR2275* family may have been impacted by duplication events leading to functional divergence of individual members. Subsequently,

TABLE 1 | Statistical analysis of singleton-to-duplicate ratios of miRNA genes located in three categories of genomic components.

	Singleton	Duplication	Ratio (S/D) ^a	P^b
Total	305	114	2.68:1	
PCG	112	13	8.62:1	9.68×10^{-5}
TE	92	33	2.79:1	0.858
UI	125	71	1.76:1	0.023

^aThe ratio of singletons to duplicates.

^bThe statistical analysis was conducted on the ratios of singletons to duplicates between total miRNAs and miRNAs in each category of genomic components by χ^2 goodness of fit test.

TABLE 2 | Statistical analysis of singleton-to-duplicate ratios of miRNA genes in individual tissue samples.

	Singleton	Duplicate	Ratio (S/D) ^a	P ^b
Total	305	114	2.68:1	
Anther	142	78	1.82:1	0.031
Ear_6week	72	64	1.12:1	1.64 × 10 ⁻⁵
Ear_8week	74	68	1.09:1	5.39 × 10 ⁻⁶
Ear_9week	72	74	0.97:1	2.17 × 10 ⁻⁷
Ear_10week	55	54	1.02:1	8.21 × 10 ⁻⁶
Early prophase meiocytes	131	85	1.54:1	3.91 × 10 ⁻³
Embryo_1d	87	63	1.38:1	7.84 × 10 ⁻⁴
Embryo_3week	57	46	1.24:1	5.78 × 10 ⁻⁴
Embryo_9DAP	85	73	1.16:1	1.38 × 10 ⁻⁵
Embryo_15DAP	89	66	1.35:1	4.24 × 10 ⁻⁴
Embryo_20DAP	67	61	1.10:1	1.42 × 10 ⁻⁵
Endosperm_9DAP	146	68	2.15:1	0.230
Endosperm_15DAP	155	70	2.21:1	0.296
Endosperm_20DAP	157	69	2.28:1	0.372
Leaf1_2week	126	74	1.70:1	1.32 × 10 ⁻²
Leaf2_2week	102	83	1.23:1	1.98 × 10 ⁻⁵
Leaf1_3week	122	71	1.72:1	1.65 × 10 ⁻²
Leaf2_3week	101	74	1.36:1	3.16 × 10 ⁻⁴
Leaf3_3week	92	76	1.21:1	2.44 × 10 ⁻⁵
Leaf_4week	88	85	1.04:1	2.81 × 10 ⁻⁷
Leaf_5week	77	82	0.94:1	3.29 × 10 ⁻⁸
Leaf_6week	112	67	1.67:1	1.27 × 10 ⁻²
Leaf_9week	116	85	1.36:1	1.67 × 10 ⁻⁴
Leaf_expanded_4week	135	89	1.52:1	1.13 × 10 ⁻³
Leaf_immature_5week	120	84	1.43:1	4.41 × 10 ⁻⁴
Leaf_mature_5week	138	85	1.62:1	4.43 × 10 ⁻³
Leaf_wrapped_4week	116	87	1.33:1	9.11 × 10 ⁻⁵
Pollen_early	89	66	1.35:1	4.24 × 10 ⁻⁴
Pollen_germination	54	45	1.20:1	4.00 × 10 ⁻⁴
Pollen1_mature	70	49	1.43:1	3.43 × 10 ⁻³
Pollen2_mature	84	46	1.83:1	4.31 × 10 ⁻²
Root1_2week	96	79	1.22:1	2.09 × 10 ⁻⁵
Root2_2week	112	66	1.70:1	1.62 × 10 ⁻²
Seed	123	68	1.81:1	3.56 × 10 ⁻²
Shoot apical meristem_4week	81	67	1.21:1	5.08 × 10 ⁻⁵
Stalk_2week	103	81	1.27:1	4.81 × 10 ⁻⁵
Silk_9week	64	78	0.82:1	1.77 × 10 ⁻⁹
Silk_11week	115	84	1.37:1	1.88 × 10 ⁻⁴
Silk_12week	101	89	1.13:1	1.92 × 10 ⁻⁶
Tassel_4week	101	76	1.33:1	1.66 × 10 ⁻⁴
Tassel_8week	65	58	1.12:1	2.93 × 10 ⁻⁵
Tassel_9week	100	91	1.10:1	7.21 × 10 ⁻⁷
Vegetative apex_2week	72	64	1.12:1	1.64 × 10 ⁻⁵

^aSingleton-to-duplicate ratios.

^bP values were calculated using the χ^2 goodness of fit test to compare the ratios of singletons to duplicates between miRNA genes in the whole genome and those expressed in different tissues.

specific *miR2275* family member was evolved as causative locus for the phenotypic transition in anthers by artificial selection.

Differentiation of 21- and 24-Nt PHAS Loci in Relation to Tissue Specificity

The 22-nt mature sequences of *miR2275* family members were recognized as critical triggers for 24-nt PHAS loci to generate phasiRNAs (Fei et al., 2013). In addition to the *miR2275* family, there is growing evidence supporting the tissue specificity of numerous PHAS triggers, such as *miR156*, *miR529*, and

miR2118. Therefore, we further examined whether the expression patterns of PHAS loci are tissue-specific in maize. To address this question, a genome-wide identification of PHAS loci was performed using sRNA-seq data collected from 14 tissues. A total of 469 21-nt and 190 24-nt PHAS loci were identified with high confidence ($P \leq 10^{-5}$), respectively (Supplementary Table 11, Supplementary Data 2, and Supplementary Data 3). Genome-wide identification of PHAS loci in maize W23 inbred line was previously performed by Zhai et al (Zhai et al., 2015). In our study, 375 21-nt PHAS loci and 113 24-nt PHAS loci were overlapped with those identified by Zhai et al, respectively (Supplementary Table 11). Meanwhile, 94 21-nt PHAS loci and 77 24-nt PHAS loci were novel. Integrating the results from two studies together, we found that the two classes of PHAS loci were highly expressed in reproductive tissues (TPM ≥ 5 ; Figure 6A). The number of PHAS loci expressed in individual tissues varied from 0 to 580, indicating tissue-specific expression (Supplementary Table 12). To gain insight into the contribution of tissue type on the production of phasiRNAs, the accumulations of phasiRNAs and corresponding trigger miRNAs were compared in individual tissues. The accumulations of phasiRNAs were positively correlated with the expression levels of corresponding trigger miRNAs (Figure 6A). The 21- and 24-nt phasiRNAs exhibited strikingly different spatiotemporal regulation. Compared with 21-nt phasiRNAs, the 24-nt phasiRNAs were preferentially expressed in pollen tissues and suppressed in early anther development (Figure 6A).

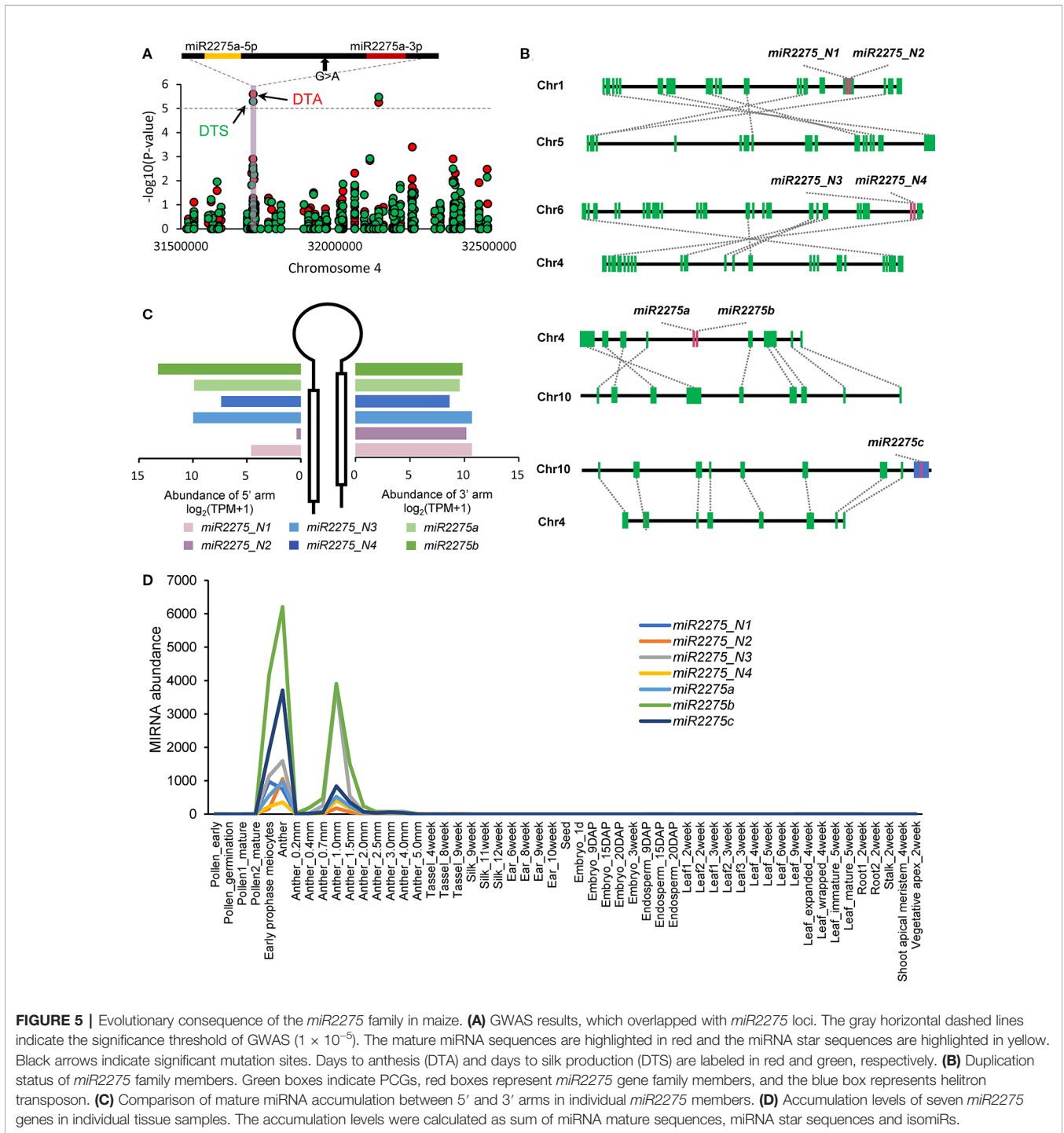
Then we analyzed the genomic distribution of PHAS loci and found that the ratio of 21-nt PHAS loci in TEs to those in PCGs (2.94:1) was significantly higher than that of 24-nt PHAS loci (1.85:1; χ^2 test, $P < 0.05$). This observation indicated evolutionary divergence of 21- and 24-nt PHAS loci, which was also reflected by the distinct expression patterns of the two categories PHAS loci in Figure 6B. Although the singleton-to-duplicate ratio of 21-nt PHAS loci (2.34:1) was approximately consistent with that of 24-nt PHAS loci (2.21:1), there were obvious differences between the expression patterns of 21- and 24-nt PHAS loci (Figure 6B).

Taken together, these results suggested that divergence on local genomic environments and expression patterns occurred between 21- and 24-nt PHAS loci, appearing to promote the formation of intraspecific homogenization and interspecific heterogenization in terms of tissue-specific function.

DISCUSSION

Comprehensive and High-Quality Annotation of miRNA Genes in Maize

It is increasingly recognized that miRNAs are key components of gene regulation, and a diversity of mechanisms of action have been proposed (Yu et al., 2017). However, the evolutionary implications of maize miRNAs on tissue specificity are still uncharacterized. Studies of miRNA genes have been frequently confounded by questionable miRNA annotations and a lack of



standardized criteria for the identification of plant miRNAs from the sRNA-seq datasets. In addition, miRNA genes are highly tissue specific, making it difficult to detect accurate abundance levels, unless massive amounts of tissues are available. In our study, we addressed these shortcomings by integrating 195 deeply sequenced sRNA libraries and by surveying 14 tissues in maize. This large-scale analysis was

performed based on the recently updated criteria which can simultaneously emphasize accuracy and minimize false positives in the genome-wide annotation of plant miRNAs (Axtell and Meyers, 2018; Kuang et al., 2018). To our knowledge, this collection of sRNA-seq datasets is the largest effort to date for the genome-wide identification of miRNA genes in maize.

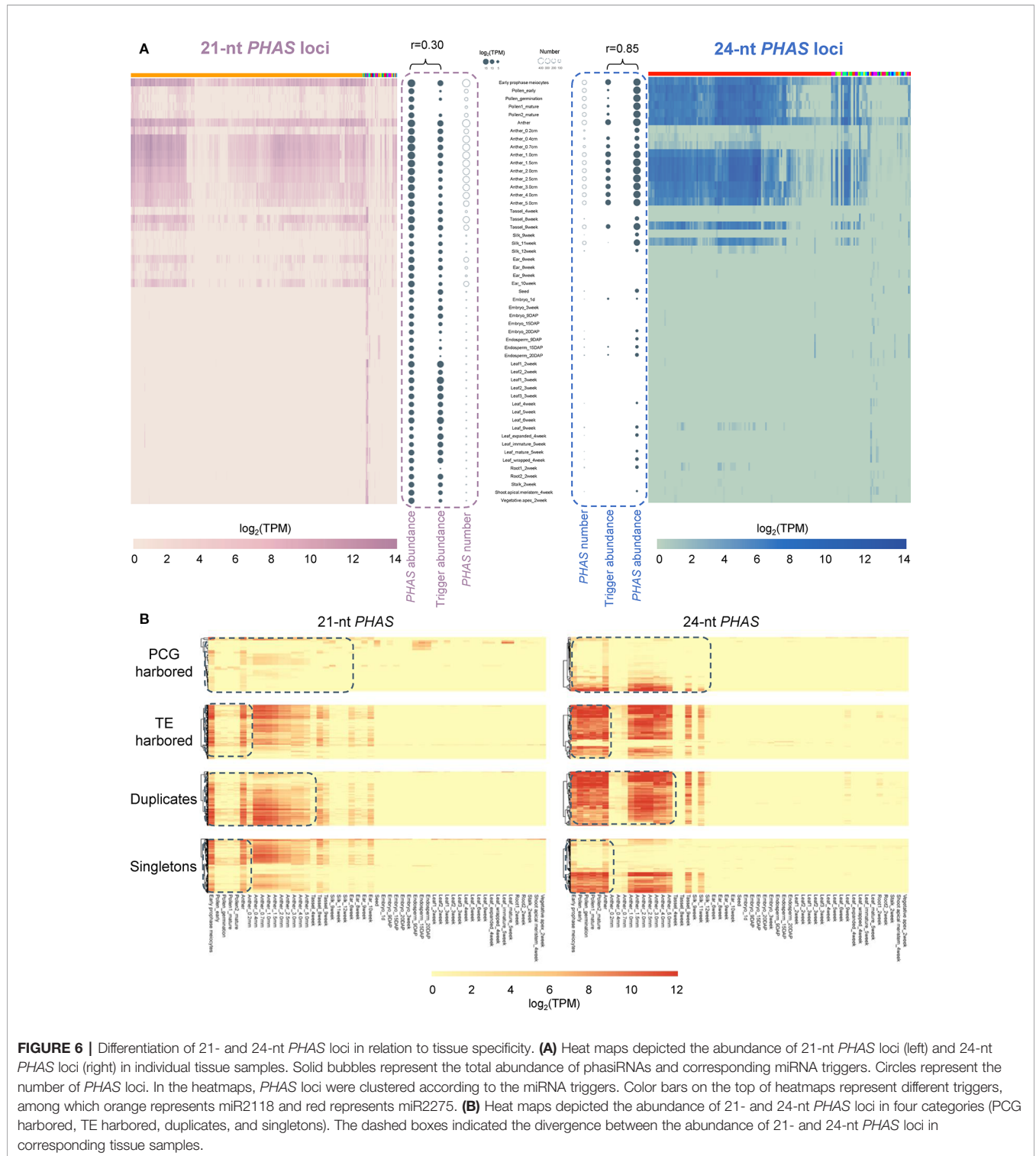


FIGURE 6 | Differentiation of 21- and 24-nt *PHAS* loci in relation to tissue specificity. **(A)** Heat maps depicted the abundance of 21-nt *PHAS* loci (left) and 24-nt *PHAS* loci (right) in individual tissue samples. Solid bubbles represent the total abundance of phasiRNAs and corresponding miRNA triggers. Circles represent the number of *PHAS* loci. In the heatmaps, *PHAS* loci were clustered according to the miRNA triggers. Color bars on the top of heatmaps represent different triggers, among which orange represents miR2118 and red represents miR2275. **(B)** Heat maps depicted the abundance of 21- and 24-nt *PHAS* loci in four categories (PCG harbored, TE harbored, duplicates, and singletons). The dashed boxes indicated the divergence between the abundance of 21- and 24-nt *PHAS* loci in corresponding tissue samples.

miRNAs as gene regulators are reported to participate in the regulation of drought responsive PCGs in many plant species (Ferdous et al., 2015). For instance, *miR162* (Tian et al., 2015) and *miR164* (Fang et al., 2014) are induced by drought stress in rice. Our study confirmed this regulation that, to our knowledge, had not been reported in maize. This implied functional

conservation of these miRNAs between rice and maize. However, some of the drought-responsive miRNAs in plants, such as *miR319* (Zhou et al., 2013), *miR393* (Bian et al., 2012), and *miR396* (Liu et al., 2009), had very low abundance in our samples and relatively stable expressions under drought stress. This is probably because the promoters of these miRNA genes

are not active in the tissues analyzed in our experiment. Nevertheless, we cannot rule out the possibility that these miRNAs do not function in response to drought stress in maize.

Local Genomic Stability Contributes to the Tissue Specificity of miRNAs and Variants

Local genomic features, such as the density of PCGs, proportion of TEs and rates of recombination, affect the removal of both genic and TE sequences (Wicker et al., 2007; Tian et al., 2009). We demonstrated that most of maize miRNA genes were distributed in chromosome arms, thus reflecting a preferential evolution driven by genetic recombination. The disruptive effects of TEs have been documented extensively, as they integrate into the regulatory or coding region of host genes or induce ectopic/nonallelic recombination (Jangam et al., 2017). We observed that there were significant differences among expression levels of miRNAs located in PCGs, TEs, and UIs. Although TEs constitute approximately 90% of the maize genome (SanMiguel et al., 1996), only 31.1% of miRNA genes were located in TEs. Such a low retention rate in TEs is more likely due to a faster purging rate of miRNA genes from TEs, which generally has limited deleterious effects on gene function and is, thus, more easily purified. We observed an extremely high abundance of miRNA genes in UIs compared with PCGs and TEs. If UIs are nearly neutral, the detected higher abundance for miRNA genes in UIs indicated that these intergenic sequences are under some level of specific functional constraints.

The modification of miRNAs, which gives rise to many isomiRs, occurs extensively and affects miRNA diversity or functional specificity (Sablok et al., 2015). Our analysis demonstrated that the number of isomiRs and frequency of isomiR types varied dramatically in different tissues. The first nucleotides on both 5' and 3' ends of miRNAs were generally 'U'. It is known that 3' adenylation increases miRNA stability, and 3' uridylation enhances miRNA degradation in plants (Lu et al., 2009; Wang et al., 2019). Thus, the presence of 'U' on the terminus of 5' addition isomiRs in all maize tissues suggested that the modification of isomiRs probably occurred to avoid degradation. Additionally, we observed diverse terminal nucleotides on the 3' end of isomiRs among the individual maize tissues, implying a tissue-specific development landmark, where the activities of enzymes involved in miRNA degradation or modification are dramatically adjusted. The abundance of isomiRs was highly associated with the genomic location of miRNA genes, implying that local genomic stability might be a principal factor contributing to the evolution of tissue specificity among maize isomiRs.

Complex Interplays May Exist Among Duplication Status, Local Genome Stability, Tissue Specificity, and miRNA Abundance

Gene duplication is a major factor responsible for the evolution of genes with novel functions. When both members of a duplicated gene pair are retained, it is believed that neither member was involved in local genomic recombination. By

contrast, a singleton originates solely from a local deletion, insertion, or translocation (Panchy et al., 2016). In general, PCGs often accumulate in genomic regions with active recombination. In this study, we found that the singleton-to-duplicate ratio of miRNA genes in PCGs was significantly higher than that in TEs and UIs, suggesting that miRNA singletons located in PCGs were likely formed through the accumulation of small deletions generated via genetic recombination, which removed single gene. These observations suggested that most miRNA singletons in maize, particularly those located in PCGs, were formed by the rapid removal of their duplicate copies after neofunctionalization. Like miRNA genes in PCGs, relatively high rates of preservation of singletons were also observed in tissue-specific miRNAs. This indicated that relatively stronger selection acted on tissue-specific miRNA singletons, which contributed to functional constraints associated with tissue specificity.

In addition, we observed that the singleton-to-duplicate ratios of expressed miRNAs in individual tissues were significantly lower than the whole genome-wide ratio. This implied a relatively higher number of non-expressed miRNA singletons than duplicates. Together with the higher genome-wide abundance miRNA duplicates than singletons, these phenomena support the gene balance hypothesis, according to which a successful genome has evolved an optimum balance of gene products binding with one another to produce multi-subunit complexes. If a gene pair is fractionated, dosage imbalance may reduce fitness or have lethal consequences. Thus, gene duplicates are generally retained (Freeling, 2009). We found that the singleton-to-duplicate ratio of miRNA genes in PCGs was significantly higher than those in TEs and UIs, and in PCGs, the accumulations of miRNA singletons were significantly lower than those of miRNA duplicates. These observations implied that the ratio of miRNA singletons in PCGs may be affected largely by gene dosage of the resident PCGs, and highly expressed miRNA genes were preferentially retained as duplicates.

GD Events Have Effects on Evolutionary Dynamics of Maize miRNA Loci in Relation to Tissue Specificity

Gene families develop via duplication of an original ancestral gene, and gene duplicates subsequently diverge from the original gene under selection (Chauve et al., 2008). This divergence occurs at varied rates, depending on the selective forces. In the case of *miR164* and *miR399* families, gene duplication and divergence appeared to result in the sub-functionalization of members of these two families. Although both families were reported to function in the development of maize roots (Tang and Chu, 2017; Hashimoto et al., 2018), we showed that *miR164a* and *miR164d* were specifically expressed in silk, and *miR399b* and *miR399e* exhibited leaf-specific expression. These two cases are typical examples of sub-functionalization, where gene copies originated from an ancestral gene eventually acquire specialized function in different tissues. Often, the functional diversification of miRNA gene duplicates occurs over time through a series of modifications to archetype miRNAs, such as nucleotide polymorphisms, resulting in the specialization of miRNA gene

function in a particular tissue (Ameres and Zamore, 2013). This typically involves changes in complementary pairing with target genes that drive the expression dynamics in relevant pathways (Baldrich et al., 2018). In animals, the seed region is considered as the basic determinant of miRNA specificity (Chandradoss et al., 2015). Here, we showed that despite the divergence in the expression profiles of *miR164* and *miR399* family members, the seed region was considerably conserved (**Supplementary Figure 8**). Based on these observations, we proposed that SNPs preferentially occurred in the 3' tail region of miRNA genes after GD, thus leading to the functional divergence of each duplicate copy. In the case of *miR2275* family, of which one member showed anther-specific expression, three homoeologous pairs formed by TD were detected. However, their syntenic copies were absent from the maize genome. Together with their consistent expression profiles, we hypothesized that paralogs produced by the WGD event lost function and were eventually deleted prior to the TD event.

Further Perspectives: Exploring the Dynamics of Newly Identified miRNAs in Different Maize Tissues With Spatial, Temporal, and Environmental Dimensions

In this study, we newly identified 271 maize miRNA genes and uncovered the tissues-specific expression patterns. As important regulatory elements of PCGs, miRNAs have gained increasing interests in the last few years to understand their dynamics roles of regulation mechanism involved in a variety of biological processes associated with growth, development, and stress responses in plants. We expect that, in the future, these new miRNA genes will be explored using large-scale sRNA-seq data from different developmental stages and tissues of maize under diverse environmental conditions. Moreover, further in-depth structural and functional analysis using genetic and molecular biology technologies will be performed to validate these new miRNA genes in maize.

REFERENCES

- Addo-Quaye, C., Miller, W., and Axtell, M. J. (2009). CleaveLand: a pipeline for using degradome data to find cleaved small RNA targets. *Bioinformatics* 25 (1), 130–131. doi: 10.1093/bioinformatics/btn604
- Allen, E., Xie, Z., Gustafson, A. M., and Carrington, J. C. (2005). microRNA-directed phasing during trans-acting siRNA biogenesis in plants. *Cell* 121 (2), 207–221. doi: 10.1016/j.cell.2005.04.004
- Ameres, S. L., and Zamore, P. D. (2013). Diversifying microRNA sequence and function. *Nat. Rev. Mol. Cell Biol.* 14 (8), 475–488. doi: 10.1038/nrm3611
- Audic, S., and Claverie, J. M. (1997). The significance of digital gene expression profiles. *Genome Res.* 7 (10), 986–995. doi: 10.1101/gr.7.10.986
- Axtell, M. J., and Meyers, B. C. (2018). Revisiting criteria for plant MicroRNA annotation in the era of big data. *Plant Cell* 30 (2), 272–284. doi: 10.1105/tpc.17.00851
- Baldrich, P., Beric, A., and Meyers, B. C. (2018). Despacito: the slow evolutionary changes in plant microRNAs. *Curr. Opin. Plant Biol.* 42, 16–22. doi: 10.1016/j.pbi.2018.01.007
- Bian, H., Xie, Y., Guo, F., Han, N., Ma, S., Zeng, Z., et al. (2012). Distinctive expression patterns and roles of the miRNA393/TIR1 homolog module in regulating flag leaf inclination and primary and crown root growth in rice

DATA AVAILABILITY STATEMENT

The six sRNA-seq datasets generated in our study have been deposited into the NCBI's SRA database under the accession numbers SRR6322649, SRR6322650, SRR6322651, SRR6322652, SRR6322653, and SRP125654.

AUTHOR CONTRIBUTIONS

YX and TZ performed the research and analyzed the data. YL contributed analytic tools and methods. ZM conceived the project and wrote the article.

FUNDING

This work was supported by the Natural Science Foundation Research Project of Shaanxi Province of China (2019JQ-096), the College Students' Platform for Innovation Training Program from Northwest A&F University (201803094), and the Fund of Northwest A&F University (Z111021603).

ACKNOWLEDGMENTS

We thank Chuang Ma at Northwest A&F University for his discussion on the project.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2020.00051/full#supplementary-material>

(*Oryza sativa*). *New Phytol.* 196 (1), 149–161. doi: 10.1111/j.1469-8137.2012.04248.x

- Buels, R., Yao, E., Diesh, C. M., Hayes, R. D., Munoz-Torres, M., Helt, G., et al. (2016). JBrowse: a dynamic web platform for genome visualization and analysis. *Genome Biol.* 17, 66. doi: 10.1186/s13059-016-0924-1
- Chandradoss, S. D., Schirle, N. T., Szczepaniak, M., MacRae, I. J., and Joo, C. (2015). A dynamic search process underlies microRNA targeting. *Cell* 162 (1), 96–107. doi: 10.1016/j.cell.2015.06.032
- Chauve, C., Doyon, J. P., and El-Mabrouk, N. (2008). Gene family evolution by duplication, speciation, and loss. *J. Comput. Biol.* 15 (8), 1043–1062. doi: 10.1089/cmb.2008.0054
- Chen, C., Ridzon, D. A., Broome, A. J., Zhou, Z., Lee, D. H., Nguyen, J. T., et al. (2005). Real-time quantification of microRNAs by stem-loop RT-PCR. *Nucleic Acids Res.* 33 (20), e179. doi: 10.1093/nar/gni178
- Creasey, K. M., Zhai, J., Borges, F., Van Ex, F., Regulski, M., Meyers, B. C., et al. (2014). miRNAs trigger widespread epigenetically activated siRNAs from transposons in Arabidopsis. *Nature* 508 (7496), 411–415. doi: 10.1038/nature13069
- Dai, X., and Zhao, P. X. (2011). psRNA target: a plant small RNA target analysis server. *Nucleic Acids Res.* 39 (Web Server issue), W155–W159. doi: 10.1093/nar/gkr319

- Doebley, J. (2004). The genetics of maize evolution. *Annu. Rev. Genet.* 38, 37–59. doi: 10.1146/annurev.genet.38.072902.092425
- Edwards, J. W., and Coruzzi, G. M. (1990). Cell-specific gene expression in plants. *Annu. Rev. Genet.* 24, 275–303. doi: 10.1146/annurev.ge.24.120190.001423
- Fahlgren, N., and Carrington, J. C. (2010). miRNA target prediction in plants. *Methods Mol. Biol.* 592, 51–57. doi: 10.1007/978-1-60327-005-2_4
- Fang, Y., Xie, K., and Xiong, L. (2014). Conserved miR164-targeted NAC genes negatively regulate drought resistance in rice. *J. Exp. Bot.* 65 (8), 2119–2135. doi: 10.1093/jxb/eru072
- Fei, Q., Xia, R., and Meyers, B. C. (2013). Phased, secondary, small interfering RNAs in posttranscriptional regulatory networks. *Plant Cell* 25 (7), 2400–2415. doi: 10.1105/tpc.113.114652
- Ferdous, J., Hussain, S. S., and Shi, B. J. (2015). Role of microRNAs in plant drought tolerance. *Plant Biotechnol. J.* 13 (3), 293–305. doi: 10.1111/pbi.12318
- Freeling, M. (2009). Bias in plant gene content following different sorts of duplication: tandem, whole-genome, segmental, or by transposition. *Annu. Rev. Plant Biol.* 60, 433–453. doi: 10.1146/annurev.arplant.043008.092122
- Hashimoto, K., Miyashima, S., Sato-Nara, K., Yamada, T., and Nakajima, K. (2018). Functionally diversified members of the MIR165/6 gene family regulate ovule morphogenesis in *Arabidopsis thaliana*. *Plant Cell Physiol.* 59 (5), 1017–1026. doi: 10.1093/pcp/pcy042
- Holub, E. B. (2001). The arms race is ancient history in *Arabidopsis*, the wildflower. *Nat. Rev. Genet.* 2 (7), 516. doi: 10.1038/35080508
- Hung, P. S., Chang, K. W., Kao, S. Y., Chu, T. H., Liu, C. J., and Lin, S. C. (2012). Association between the rs2910164 polymorphism in pre-mir-146a and oral carcinoma progression. *Oncol.* 48 (5), 404–408. doi: 10.1016/j.oraloncology.2011.11.019
- Jangam, D., Feschotte, C., and Betran, E. (2017). Transposable element domestication as an adaptation to evolutionary conflicts. *Trends In Genet.* 33 (11), 817–831. doi: 10.1016/j.tig.2017.07.011
- Jeong, D. H., Park, S., Zhai, J., Gurazada, S. G., De Paoli, E., Meyers, B. C., et al. (2011). Massive analysis of rice small RNAs: mechanistic implications of regulated microRNAs and variants for differential target RNA cleavage. *Plant Cell* 23 (12), 4185–4207. doi: 10.1105/tpc.111.089045
- Jiao, Y., Peluso, P., Shi, J., Liang, T., Stitzer, M. C., Wang, B., et al. (2017). Improved maize reference genome with single-molecule technologies. *Nature* 546 (7659), 524–527. doi: 10.1038/nature22971
- Kadota, K., Ye, J., Nakai, Y., Terada, T., and Shimizu, K. (2006). ROKU: a novel method for identification of tissue-specific genes. *BMC Bioinf.* 7, 294. doi: 10.1186/1471-2105-7-294
- Kakrana, A., Li, P., Patel, P., Hammond, R., Anand, D., Mathioni, S., et al. (2017). PHASIS: a computational suite for de novo discovery and characterization of phased, siRNA-generating loci and their miRNA triggers. *bioRxiv.* 158832. doi: 10.1101/158832
- Kozomara, A., and Griffiths-Jones, S. (2014). miRBase: annotating high confidence microRNAs using deep sequencing data. *Nucleic Acids Res.* 42 (Database issue), D68–D73. doi: 10.1093/nar/gkt1181
- Kuang, Z., Wang, Y., Li, L., and Yang, X. (2018). miRDeep-P2: accurate and fast analysis of the microRNA transcriptome in plants. *Bioinformatics* 35 (14), 2521–2522. doi: 10.1093/bioinformatics/bty972
- Li, S., Castillo-Gonzalez, C., Yu, B., and Zhang, X. (2017). The functions of plant small RNAs in development and in stress responses. *Plant J.* 90 (4), 654–670. doi: 10.1111/tpj.13444
- Lipka, A. E., Tian, F., Wang, Q., Peiffer, J., Li, M., Bradbury, P. J., et al. (2012). GAPIT: genome association and prediction integrated tool. *Bioinformatics* 28 (18), 2397–2399. doi: 10.1093/bioinformatics/bts444
- Liu, D., Song, Y., Chen, Z., and Yu, D. (2009). Ectopic expression of miR396 suppresses GRF target gene expression and alters leaf growth in *Arabidopsis*. *Physiol. Plant.* 136 (2), 223–236. doi: 10.1111/j.1399-3054.2009.01229.x
- Liu, Q., Wang, H., Zhu, L., Hu, H., and Sun, Y. (2013a). Genome-wide identification and analysis of miRNA-related single nucleotide polymorphisms (SNPs) in rice. *Rice* 6 (1), 10. doi: 10.1186/1939-8433-6-10
- Liu, Y., Wang, Y., Zhu, Q. H., and Fan, L. (2013b). Identification of phasiRNAs in wild rice (*Oryza rufipogon*). *Plant Signaling Behav.* 8 (8), e25079. doi: 10.4161/psb.25079
- Lu, S., Sun, Y. H., and Chiang, V. L. (2009). Adenylation of plant miRNAs. *Nucleic Acids Res.* 37 (6), 1878–1885. doi: 10.1093/nar/gkp031
- Lyu, J. (2017). Maize domestication: an ancient genome speaks. *Nat. Plants* 3, 16215. doi: 10.1038/nplants.2016.215
- Maher, C., Stein, L., and Ware, D. (2006). Evolution of *Arabidopsis* microRNA families through duplication events. *Genome Res.* 16 (4), 510–519. doi: 10.1101/gr.4680506
- Meyers, B. C., Axtell, M. J., Bartel, B., Bartel, D. P., Baulcombe, D., Bowman, J. L., et al. (2008). Criteria for annotation of plant MicroRNAs. *Plant Cell* 20 (12), 3186–3190. doi: 10.1105/tpc.108.064311
- Panchy, N., Lehti-Shiu, M., and Shiu, S. H. (2016). Evolution of gene duplication in plants. *Plant Physiol.* 171 (4), 2294–2316. doi: 10.1104/pp.16.00523
- Petsch, K., Manzotti, P. S., Tam, O. H., Meeley, R., Hammell, M., Consonni, G., et al. (2015). Novel DICER-LIKE1 siRNAs bypass the requirement for DICER-LIKE4 in maize development. *Plant Cell* 27 (8), 2163–2177. doi: 10.1105/tpc.15.00194
- Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A., Bender, D., et al. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* 81 (3), 559–575. doi: 10.1086/519795
- Sablok, G., Srivastva, A. K., Suprasanna, P., Baev, V., and Ralph, P. J. (2015). isomiRs: Increasing evidences of isomiRs complexity in plant stress functional biology. *Front. In Plant Sci.* 6, 949. doi: 10.3389/fpls.2015.00949
- SanMiguel, P., Tikhonov, A., Jin, Y. K., Motchoulskaia, N., Zakharov, D., Melake-Berhan, A., et al. (1996). Nested retrotransposons in the intergenic regions of the maize genome. *Science* 274 (5288), 765–768. doi: 10.1126/science.274.5288.765
- Sato, K., Hamada, M., Asai, K., and Mituyama, T. (2009). CENTROIDFOLD: a web server for RNA secondary structure prediction. *Nucleic Acids Res.* 37 (Web Server issue), W277–W280. doi: 10.1093/nar/gkp367
- Schmittgen, T. D., and Livak, K. J. (2008). Analyzing real-time PCR data by the comparative C(T) method. *Nat. Protocols* 3 (6), 1101–1108. doi: 10.1038/nprot.2008.73
- Schnable, J. C., Springer, N. M., and Freeling, M. (2011). Differentiation of the maize subgenomes by genome dominance and both ancient and ongoing gene loss. *Proc. Natl. Acad. Sci. U. S. A.* 108 (10), 4069–4074. doi: 10.1073/pnas.1101368108
- Sun, G., Yan, J., Noltner, K., Feng, J., Li, H., Sarkis, D. A., et al. (2009). SNPs in human miRNA genes affect biogenesis and function. *RNA* 15 (9), 1640–1651. doi: 10.1261/rna.1560209
- Swigonova, Z., Lai, J., Ma, J., Ramakrishna, W., Llaca, V., Bennetzen, J. L., et al. (2004). Close split of sorghum and maize genome progenitors. *Genome Res.* 14 (10A), 1916–1923. doi: 10.1101/gr.2332504
- Tang, J., and Chu, C. (2017). MicroRNAs in crop improvement: fine-tuners for complex traits. *Nat. Plants* 3, 17077. doi: 10.1038/nplants.2017.77
- Tian, Z., Rizzon, C., Du, J., Zhu, L., Bennetzen, J. L., Jackson, S. A., et al. (2009). Do genetic recombination and gene density shape the pattern of DNA elimination in rice long terminal repeat retrotransposons? *Genome Res.* 19 (12), 2221–2230. doi: 10.1101/gr.083899.108
- Tian, C., Zuo, Z., and Qiu, J. L. (2015). Identification and characterization of ABA-responsive microRNAs in rice. *J. Genet. Genomics* 42 (7), 393–402. doi: 10.1016/j.jgg.2015.04.008
- Tu, B., Liu, L., Xu, C., Zhai, J., Li, S., Lopez, M. A., et al. (2015). Distinct and cooperative activities of HESO1 and URT1 nucleotidyl transferases in microRNA turnover in *Arabidopsis*. *PLoS Genet.* 11 (4), e1005119. doi: 10.1371/journal.pgen.1005119
- Voorrips, R. (2002). MapChart: software for the graphical presentation of linkage maps and QTLs. *J. Hered.* 93 (1), 77–78. doi: 10.1093/jhered/93.1.77
- Wang, J., Mei, J., and Ren, G. (2019). Plant microRNAs: biogenesis, homeostasis, and degradation. *Front. Plant Sci.* 10, 360. doi: 10.3389/fpls.2019.00360
- Wicker, T., Yahiaoui, N., and Keller, B. (2007). Illegitimate recombination is a major evolutionary mechanism for initiating size variation in plant resistance genes. *Plant J.* 51 (4), 631–641. doi: 10.1111/j.1365-3113.2007.03164.x
- Wu, H. J., Ma, Y. K., Chen, T., Wang, M., and Wang, X. J. (2012). PsRobot: a web-based plant small RNA meta-analysis toolbox. *Nucleic Acids Res.* 40 (Web Server issue), W22–W28. doi: 10.1093/nar/gks554
- Yu, Y., Jia, T., and Chen, X. (2017). The 'how' and 'where' of plant microRNAs. *New Phytol.* 216 (4), 1002–1017. doi: 10.1111/nph.14834

- Zhai, J., Zhao, Y., Simon, S. A., Huang, S., Petsch, K., Arikat, S., et al. (2013). Plant microRNAs display differential 3' truncation and tailing modifications that are ARGONAUTE1 dependent and conserved across species. *Plant Cell* 25 (7), 2417–2428. doi: 10.1105/tpc.113.114603
- Zhai, J., Zhang, H., Arikat, S., Huang, K., Nan, G. L., Walbot, V., et al. (2015). Spatiotemporally dynamic, cell-type-dependent premeiotic and meiotic phasiRNAs in maize anthers. *Proc. Natl. Acad. Sci. U. S. A.* 112 (10), 3146–3151. doi: 10.1073/pnas.1418918112
- Zhao, M., Meyers, B. C., Cai, C., Xu, W., and Ma, J. (2015). Evolutionary patterns and coevolutionary consequences of MIRNA genes and microRNA targets triggered by multiple mechanisms of genomic duplications in soybean. *Plant Cell* 27 (3), 546–562. doi: 10.1105/tpc.15.00048
- Zhong, X., Pla, A., and Rayner, S. (2019). Jasmine: a Java pipeline for isomiR characterization in miRNA-seq Data. *Bioinformatics* btz806 doi: 10.1093/bioinformatics/btz806
- Zhou, M., Li, D., Li, Z., Hu, Q., Yang, C., Zhu, L., et al. (2013). Constitutive expression of a miR319 gene alters plant development and enhances salt and drought tolerance in transgenic creeping bentgrass. *Plant Physiol.* 161 (3), 1375–1391. doi: 10.1104/pp.112.208702

Conflict of Interest: 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 © 2020 Xu, Zhang, Li and Miao. 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) and the copyright owner(s) 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.