

RESEARCH ARTICLE

Open Access



Identification and functional characterization of intermediate-size non-coding RNAs in maize

Dandan Li^{1†}, Huili Qiao^{1†}, Wujie Qiu¹, Xin Xu¹, Tiemei Liu¹, Qianling Jiang¹, Renyi Liu², Zhujin Jiao¹, Kun Zhang¹, Lijun Bi³, Runsheng Chen³ and Yunchao Kan^{1*}

Abstract

Background: The majority of eukaryote genomes can be actively transcribed into non-coding RNAs (ncRNAs), which are functionally important in development and evolution. In the study of maize, an important crop for both humans and animals, aside from microRNAs and long non-coding RNAs, few studies have been conducted on intermediate-size ncRNAs.

Results: We constructed a homogenized cDNA library of 50–500 nt RNAs in the maize inbred line Chang 7–2. Sequencing revealed 169 ncRNAs, which contained 58 known and 111 novel ncRNAs (including 70 snoRNAs, 27 snRNAs, 13 unclassified ncRNAs and one tRNA). Forty of the novel ncRNAs were specific to the Panicoideae, and 24% of them are located on sense-strand of the 5' or 3' terminus of protein coding genes on chromosome. Target site analysis found that 22 snoRNAs can guide to 38 2'-O-methylation and pseudouridylation modification sites of ribosomal RNAs and small nuclear RNAs. Expression analysis showed that 43 ncRNAs exhibited significantly altered expression in different tissues or developmental stages of maize seedlings, eight ncRNAs had tissue-specific expression and five ncRNAs were strictly accumulated in the early stage of leaf development. Further analysis showed that 3 of the 5 stage-specific ncRNAs (Zm-3, Zm-18, and Zm-73) can be highly induced under drought and salt stress, while one snoRNA Zm-8 can be repressed under PEG-simulated drought condition.

Conclusions: We provided a genome-wide identification and functional analysis of ncRNAs with a size range of 50–500 nt in maize. 111 novel ncRNAs were cloned and 40 ncRNAs were determined to be specific to Panicoideae. 43 ncRNAs changed significantly during maize development, three ncRNAs can be strongly induced under drought and salt stress, suggesting their roles in maize stress response. This work set a foundation for further study of intermediate-size ncRNAs in maize.

Keywords: Maize, Intermediate-size ncRNAs, Stress

Background

With more and more genomes being sequenced, numerous non-coding RNAs (ncRNAs) have been identified, their functions are also being revealed [1–8]. In maize, a large number of miRNAs have been identified in different lines and developmental conditions [9–12], such as miR159, miR164, miR167, miR393, miR408 and miR528 are mainly

involved in root development and stress responses [13, 14]. miR160, miR167, miR164, miR169 and miR393 can respond to hormone signaling [14]. Genome-wide identification of long non-coding RNAs (lncRNAs) was also accomplished in maize, more than 40,000 lncRNAs have been identified, most of them were expressed in tissue-specific manner [15–20]. Some lncRNA can participate in maize pollen development, such as lncRNA *Zm401*, knocking down of which can significantly affect three key genes for pollen development (*ZmMADS2*, *MZm3-3*, and *AmC5*), then result in male sterility [21]. Other lncRNAs can response to stress conditions, Wang et al. found some

* Correspondence: yckan1974@nynu.edu.cn

†Dandan Li and Huili Qiao contributed equally to this work.

¹China-UK-NYNU-RRes Joint Laboratory of insect biology, Henan Key Laboratory of Insect Biology in Funiu Mountain, Nanyang Normal University, 1638 Wolong Road, Nanyang 473061, Henan, China

Full list of author information is available at the end of the article



transposable-element-derived long intergenic non-coding RNAs (TE-lincRNAs) can be induced or inhibited by cold, heat, drought, salt or high light stress in rice and maize [22]. However, studies focused on the identification and functional characterization of intermediate-size ncRNAs (50–500 nt) in maize is rare. Small nucleolar RNAs (snoRNAs) and small nuclear RNAs (snRNAs) are the major classes of intermediate-size ncRNAs, snoRNAs can guide site-specific RNA modifications of ribosomal RNAs (rRNAs), snRNAs, tRNAs as well as mRNAs [23], while snRNAs can participate in alternative splicing in mRNA processing. Nowadays, more and more evidence showed that snoRNA, the old dog, has new tricks [24]. Through sRNA deep sequencing studies, several reports have identified snoRNA-derived small RNAs (sdrRNAs) [25–29], which broaden the roles of snoRNAs. sdrRNAs in animals are preferentially derived from the 3' end of H/ACA snoRNAs and the 5' end of C/D snoRNAs [29–31]. While sdrRNAs of *Arabidopsis* are always associated with AGO7 [30]. Recently, a snoRNA-derived piRNA piR30840 was found to be accumulated in human CD4 primary T-lymphocytes, piR30840 can bind to the intron of interleukin-4 through AGO4/PIWI14/piRNA complex, then lead to the pre-messenger RNA (pre-mRNA) decay of interleukin-4 [32]. Meanwhile, snoRNAs can be the biomarkers for diseases and physiological changes [33, 34], such as SNORA21 and SNORD126, can be accumulated more in colorectal adenomas and hepatocellular carcinoma, respectively [35, 36]. Moreover, some snoRNAs were found to play roles in pre-mRNA alternative splicing [24, 37–42], suggesting that the potential roles of snoRNAs are just beginning to be appreciated.

Maize (*Zea mays* L.) is one of the most important food crops in the world. The maize genome has been sequenced, and nearly 85% of the 3.2 GB genome sequence is composed of hundreds of families of transposable elements [43]. Because the vast majority of the maize genome consists of non-protein-coding regions, there is great potential for discovering more ncRNAs in maize.

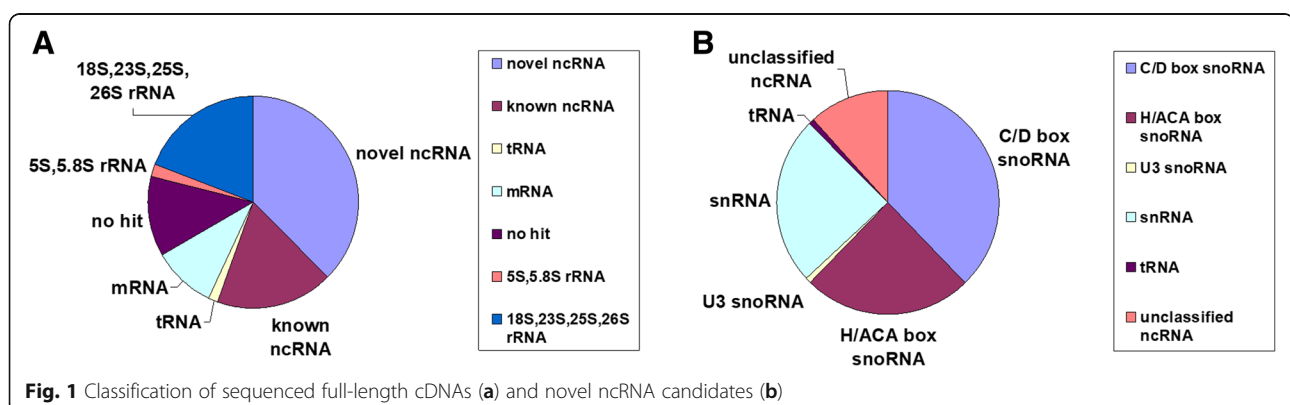
In this work, we constructed a cDNA library of 50–500 nt RNAs that were extracted and size-selected from

seedling and developing grain total RNAs of the maize inbred line Chang 7–2. From 1,273 full-length cDNA sequences, we identified 111 novel ncRNA candidates. The expression pattern of these ncRNAs from different tissues and different developmental stages were studied. Functional exploration of ncRNAs was also done under drought and salt stress. Our results provide the first genome-wide survey and functional characterization of intermediate-size ncRNAs in maize.

Results

Identification of intermediate-size ncRNAs in maize

To identify new ncRNA candidates in maize, we constructed a full-length intermediate-size ncRNA-enriched library (50–500 nt) from the maize (*Zea mays* L.) inbred line Chang 7–2, the wide-spread planted line in Henan province of China. Total RNAs were isolated from seven tissues harvested at eleven developmental stages of maize seedlings and five stages of developing grains. In order to obtain the 50–500 nt RNA fragments precisely and to remove the known RNAs with high abundance, cDNA library was constructed with the method of Deng et al. but not RNA-seq [44, 45]. Probes with poly (A) tails were used to remove the enriched RNAs such as mRNAs, rRNAs, and snRNAs to homogenize the library, then 50–500 nt RNAs were sliced from the gel and inserted into the pGEM-T vector for electro-transformation. Approximately 5,000 clones were picked out and tested by PCR, the products were identified by 6% Polyacrylamide gel electrophoresis (PAGE) and selected by their length. Totally, 1,273 clones were picked out and sequenced. After removing duplications with same sequence, 297 unique sequences were identified, among which 63 transcripts (21%) were rRNAs, 169 transcripts (57%) were annotated as other ncRNAs, the rest were error reads or mRNA degradation products (Fig. 1a), which indicated that the probes used to fish out rRNAs was not so efficient, but the yield of other ncRNAs in unique transcripts was high, accomplished 57%.



Of the 169 ncRNAs, 58 were known ncRNAs, which contained 42 snRNAs, 7 snoRNAs, 5 Signal recognition particle RNAs (SRPs) and 4 tRNAs. Meanwhile, 111 (66%) transcripts were identified as novel ncRNA candidates. To verify these ncRNAs were real transcripts but not degradation products, the expression of 84 randomly selected ncRNAs were identified by northern blots (data in the ncRNA expression results). In addition, full length of 14 ncRNAs was checked by 5' and 3' RACE. Results showed besides the RACE sequences differed from the corresponding cDNA sequences by at most one to two terminal nucleotides, the other results showed that these ncRNA sequences are indeed full-length transcripts.

Based on structural features and sequence similarity to known ncRNAs of other species (NONCODE version 5.0), 111 novel ncRNAs were classified into different categories. Seventy of them (63%) were annotated as snoRNA-like transcripts, including 42 C/D box, 27 H/ACA box snoRNAs, and one U3 snoRNA. Twenty-seven ncRNAs (24%) were classified as snRNAs and one as tRNA. The other 13 ncRNA candidates (12%) could not be assigned to any group due to lack of known motifs or secondary structure hallmarks and are referred to as “unclassified ncRNAs” (Fig. 1b). Comparing to the RNA-seq method used in rice, which got 125 novel snoRNAs (9%), 59 (4%) snRNAs and a large number of unclassified ncRNAs (781) [45], we got high percentage of novel snoRNAs but low ratio of unclassified ncRNAs, which indicated that our library covered the main part of ncRNAs like snoRNAs and snRNAs in this length range.

Conservation analysis shows a large number of Panicoideae-specific ncRNAs

To investigate whether the sequenced ncRNAs were conserved in flowering plants, we used the BLASTN program to search for their homologs in sorghum, rice, and *Arabidopsis* genomes. Results showed that 76 of the 169 unique ncRNAs had counterparts in the three plant genomes, including 3 snRNA families (57 unique snRNAs), 8 snoRNAs, 4 SRPs, 5 tRNAs, and 2 unclassified ncRNAs (Zm-79 and Zm-80). Almost all of the ncRNAs in maize (166 of 169 transcripts) had counterparts in the closely-related *Sorghum bicolor* genome (with identity more than 90%), among which 100 ncRNAs match the sorghum EST sequences, indicating that they are also expressed in sorghum. We also found 128 (76%) ncRNAs had counterparts in the rice genome (with identity more than 95%). The remaining 41 ncRNAs with little conservation were snoRNAs and unclassified ncRNAs (Additional file 1).

Targets predication showed that 16 C/D box and 6 H/ACA box snoRNAs could guide to 30 2'-O-methylation and 8 pseudouridylation sites of maize rRNAs and snRNAs, respectively. After removing the conserved

ncRNAs and transcripts with conserved functional elements, 40 ncRNAs (containing 9 C/D box, 18 H/ACA box snoRNAs and 13 unclassified ncRNAs) still could not identify any homolog or ortholog to *Arabidopsis* or rice, indicating that they were specific to Panicoideae.

Genomic organization of maize ncRNAs

We compared the genomic locations of the novel ncRNAs with annotated protein-coding genes in the maize genome. Results showed that more than half of novel ncRNAs were located in the intergenic regions and 15% in introns (Additional file 2: Table S1), 70% host genes of the intronic ncRNAs encode ribosomal proteins (Additional file 2: Figure S1), which is similar to other organisms [23]. Twenty-seven novel ncRNAs (24% of total novel ncRNAs) had overlaps with the 5' or 3' untranslated region (UTR) of protein coding genes on the sense strand (Additional file 2: Table S1). This is similar to rice, in which more than half of the intermediate-size ncRNAs were located around the plus-strand of the 5' and 3' terminus of the coding sequences [45]. However, different to rice (in which the number of ncRNAs on the 3' UTR is nearly two-fold as that on the 5' UTR), 67% of the UTR-origin ncRNAs in maize are located on the 5' UTR, and 78% of the UTR-origin ncRNAs are snoRNAs, but most of them have no predicated targets to rRNAs or snRNAs, which indicated that they might be orphan snoRNAs.

To verify that UTR-origin ncRNAs are not degradation products of protein-coding genes, 21 of the 27 UTR-origin ncRNAs were detected by the northern blot. Results showed that all of them can be hybridized with the correct size, but the corresponding host genes could not be detected (Additional file 2: Figure S2A and B), suggesting that the UTR-origin ncRNAs were full-length transcripts rather than degradation products of host genes. Moreover, 93% of the host genes of UTR-origin ncRNAs are annotated as hypothetical or pseudogenes (Additional file 2: Figure S3). To study whether these UTR-origin snoRNAs has the potential to be precursors of miRNAs, online software miRNAfold was used. Results showed that these snoRNAs had no potential to be miRNA precursors.

The ncRNAs that originated from Open Reading Frames (ORFs) had been found in other organisms, such as in *Drosophila* and rice [45, 46]. In maize, nine ncRNAs (8 U1 snRNAs variants and 1 C/D box snoRNA) were found to be located in ORFs. The host genes of the 8 snRNAs are annotated as pseudogenes with cDNA length between 141 nt and 400 nt. Northern blot results showed that only the U1 snRNA but not the host pseudogene had hybridization signals (Additional file 2: Figure S2D), indicating that the host pseudogenes might not be transcribed to coding RNAs.

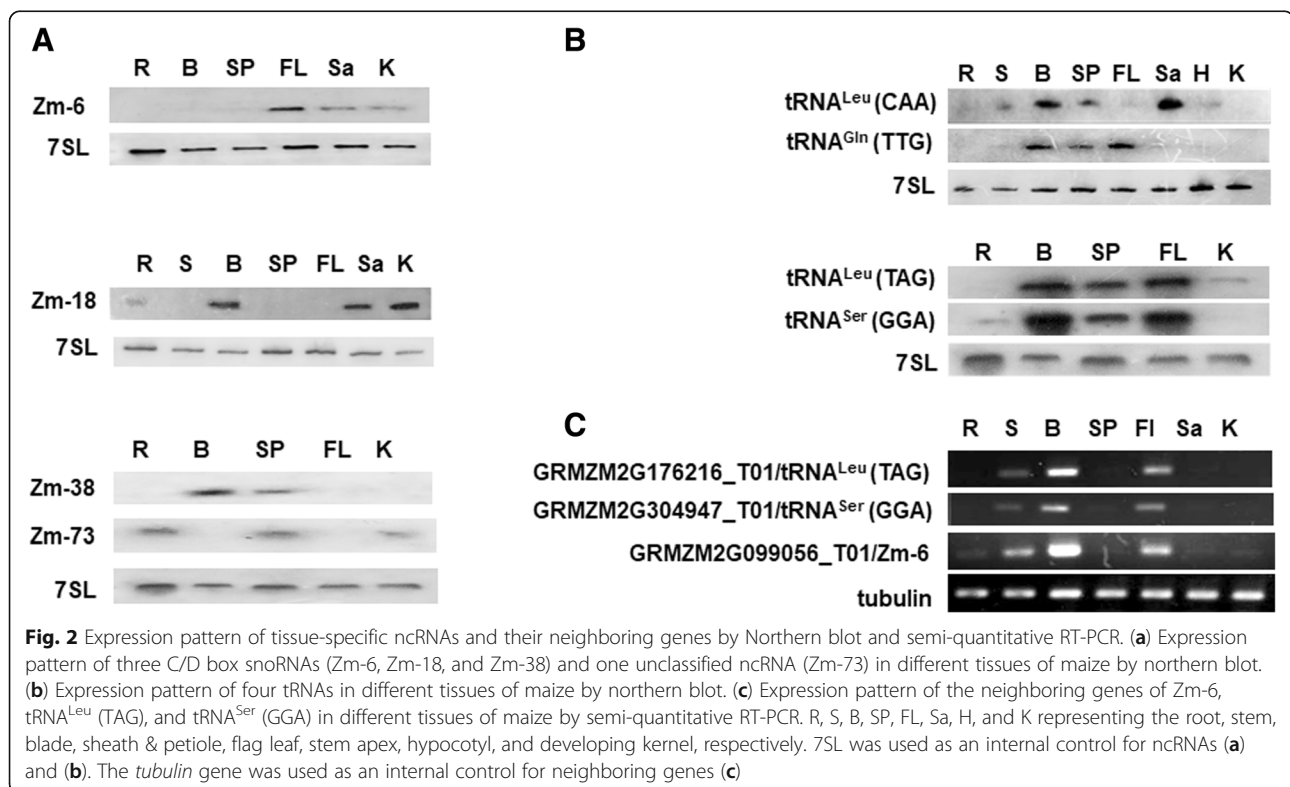
In plants, most snoRNAs are found in polycistronic clusters. We examined the relative locations of intermediate-size ncRNAs in maize and found that 82 ncRNA variants (58 unique ncRNAs) formed 28 clusters on chromosomes (Additional file 1), 71% of which were snoRNAs. In contrast to rice, in which all snoRNA clusters are in intron and intergenic regions, there are only five intronic and three intergenic snoRNA clusters in maize, and the other 11 snoRNA clusters are located in the UTR regions of protein-coding genes (Additional file 1).

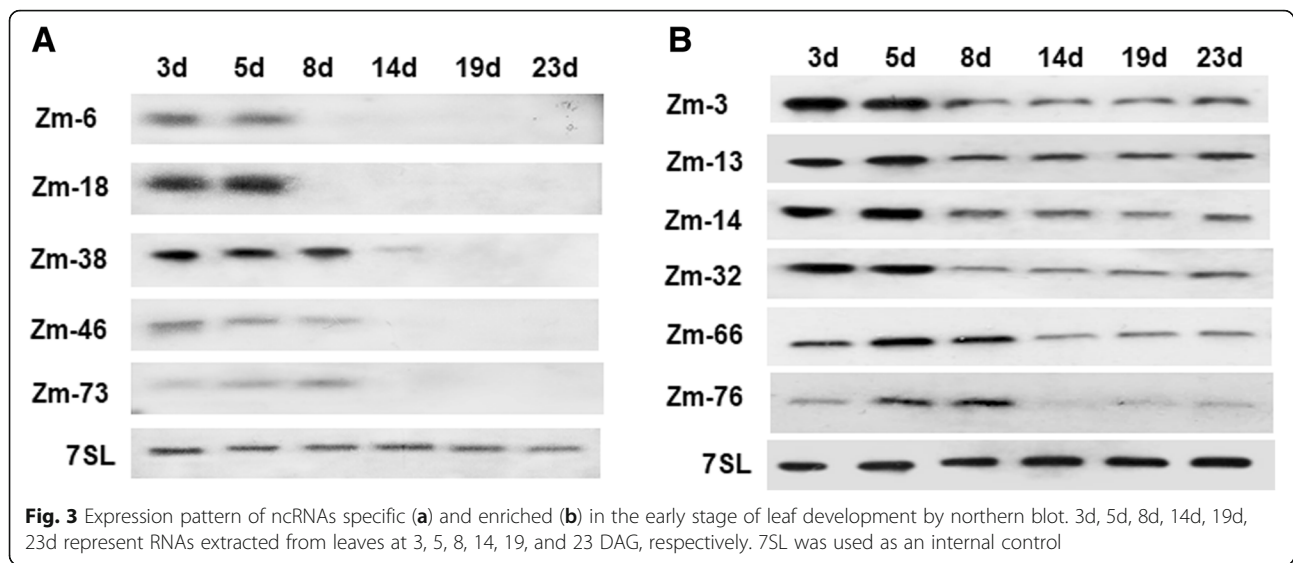
In rice, half of the snoRNA gene families expanded in the genome through tandem duplications [45]. In maize, many less tandem duplications were involved in the origination of ncRNAs. Only one C/D box snoRNA, Zm-19, which was transcribed from the second intron of gene GRMZM2G047727-T01, has five tandem repeats with an interval of 138 nt. The counterparts of these tandem repeats are also observed in the sorghum and rice genomes, but with different repeat numbers. Another two unclassified ncRNAs, Zm-79 and Zm-80, have nine tandem repeats on six chromosomes (chromosome 1, 2, 3, 5, 7, and 10) with intervals between 2791 bp and 2822 bp. They also have counterparts of tandem repeats in the sorghum, rice, and *Arabidopsis* genomes with less repeat numbers, indicating that these tandem repeats are highly conserved between monocots and dicots.

ncRNAs with tissue-specific expression

The ncRNAs are frequently expressed in specific tissues or developmental stages in eukaryotes, indicating their diverse roles in various cellular processes [42, 47–50]. We used northern blots to compare the expression of the intermediate-size ncRNAs in different tissues and developmental stages of maize. The expression of 64 unique novel ncRNAs and 20 known ncRNAs were detected, results showed that 43 ncRNAs had significantly altered expression (Figs. 2, 3, Additional file 2: Figures S2A, B, D and S5A and S6A). These include 31 snoRNAs (20 C/D box and 11 H/ACA box snoRNAs), 4 unclassified ncRNAs, 4 snRNAs (three U2 and one U5 snRNA), and four tRNAs. Nearly 65% of the differentially-expressed snoRNAs had no predicted targets and might be orphan snoRNAs.

Most of the intronic and UTR-origin ncRNAs exhibited a similar expression pattern with their host genes in different tissues (Additional file 2: Figure S2 and S5). However, the expression of intergenic ncRNAs had various correlations to their neighboring genes, consistent in some tissues (root, blade, stem, flag leaf and developing kernel), but different in other tissues (sheath & petiole and stem apex) (Additional file 2: Figure S6), indicating their different regulatory roles in various organs. Interestingly, Zm-46 had two transcript isoforms in the blade and developing kernel (Additional file 2: Figure S5),





indicating that alternative splicing products were generated in different tissues.

Eight ncRNAs exhibited tissue-differential expression (Fig. 2), including three C/D box snoRNAs (Zm-6, Zm-18, and Zm-38), one unclassified ncRNA, and four tRNAs. Such as Zm-6, had more expression in the flag leaf, stem apex, and developing kernel, but little or no detectable expression in the blade and root (Fig. 2a); however, its antisense upstream gene GRMZM2G099056-T01(*roothairless*) (genomic position was shown in Additional file 2: Figure S4) accumulated more in the blade and root (Fig. 2c), indicating that Zm-6 might play reverse role in the regulation of *roothairless* during maize root development.

All four tRNAs with altered expressions were accumulated more in leaves (Fig. 2b). To further study the functions of these tRNAs, we analyzed their genomic organization and the expression pattern of their neighboring genes. We found that each tRNA had multiple copies in the genome. For example, there are 25 copies of tRNA^{Ser} (GGA) in the maize genome, and all of them are located on the antisense strand of LTR/Copia retrotransposons. The expression of tRNAs were positively correlated with their neighboring genes (Fig. 2b and c), indicating that the tissue-differential expressed tRNAs might function together with their neighboring genes in the leaf development of maize.

ncRNAs accumulated more in juvenile stage of maize seedling

The change of gene expression in leaves during the seedling stage is important for grain yield. During this period, leaves exhibit photomorphogenic growth and accumulate much more nutriment in preparation for the reproductive stage. With Northern blots analysis, we found that 11 ncRNAs can accumulated more in the

early stages of leaf development (Fig. 3), including 2 unclassified ncRNAs and 9 snoRNAs. Furthermore, five ncRNAs (Zm-6, Zm-18, Zm-38, Zm-46, and Zm-73) were strictly expressed before 14 DAG (Fig. 3a). Further analysis showed that besides Zm-38, the other snoRNAs have no predicated targets to rRNAs and snRNAs, which appeared to be orphan snoRNAs. From 8 to 11 DAG, maize enters into the trefoil stage, the seedling begins transition from juvenile to adult, a lot of protein-coding genes had differential expressions before and after the conversion, such as *Teopod* genes, the accumulation of *Tp1*, *Tp2*, and *Tp3* before the transition can promote juvenility of maize seedling [51, 52]. In our results, the accumulation of snoRNAs in leaves before the 14th day of seedling indicated their roles in juvenile maintenance of maize.

Stress-regulated ncRNAs

Accumulated evidence supports important roles for ncRNAs in plant response and adaptation to abiotic stresses [53]. To investigate whether intermediate-size ncRNAs are regulated by abiotic stresses, we used polyethylene glycol 6000 (PEG6000), abscisic acid (ABA) and NaCl to simulate drought and salt stress to study the expression of ncRNAs in maize roots. Results showed that 3 ncRNAs (Zm-3, Zm-18, and Zm-73) were induced strongly at the early stage of PEG simulated drought stress (1 h), but their expression decreased sharply after 6 h treatment (Fig. 4a), a similar result was also found under ABA stress (Fig. 4b). Moreover, Zm-18 and Zm-73 can be induced strongly at 3 h under salt stress, and Zm-3 was induced quickly after 1 h NaCl treatment (Fig. 4c), which indicated their roles in stress response.

However, other ncRNAs can be inhibited by stress, like Zm-8, whose expression was decreased after 2 h of PEG-simulated drought stress (Fig. 4a). But which

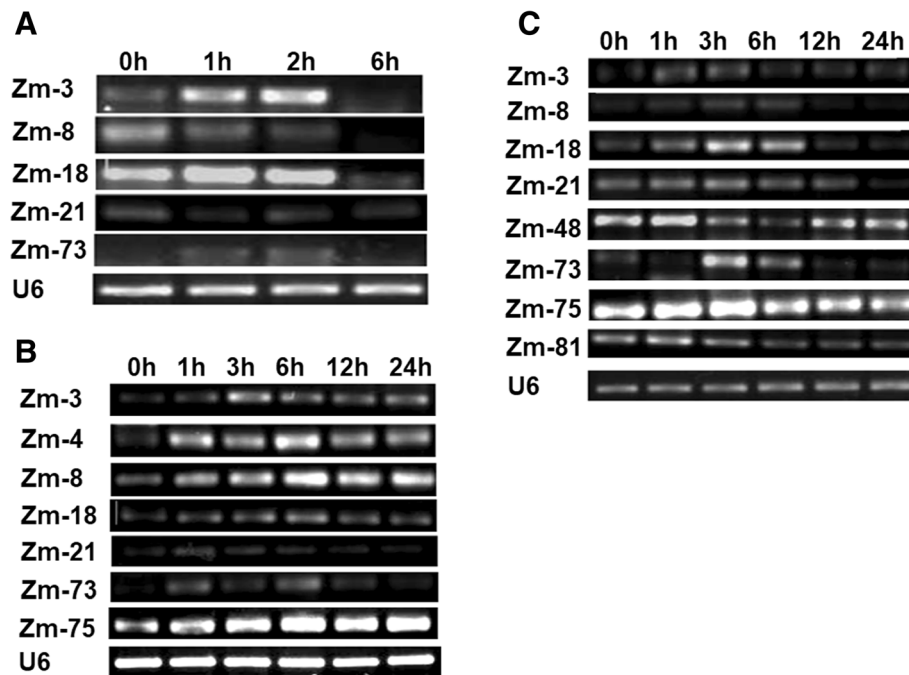


Fig. 4 Semi-quantitative RT-PCR results of nCrNA genes under PEG6000 simulated-drought stress (a), ABA simulated-drought stress (b), and NaCl simulated-salt stress (c). 0, 1, 3, 6, 12, and 24 h represented RNAs extracted from roots at the 0, 1, 3, 6, 12, and 24 h after different abiotic stress. U6 was used as an internal control

interesting is the expression of Zm-8 can be induced after 6 h ABA treatment (Fig. 4b), which indicated Zm-8 might participate in different signal pathways in different treatments although PEG and ABA both simulated the drought stress.

Discussion

Panicoideae-specific nCrNAs

Maize and sorghum are both Panicoideae plant, maize genome is three times larger than the sorghum genome, but 93% of their protein coding gene families were conserved [54, 55]. In our results, almost all the nCrNAs detected in maize can find counterparts on sorghum genome, and 100 nCrNAs can blast to the EST sequences of sorghum, which indicated that nCrNAs of same subfamily were also highly conserved.

But when comparing the nCrNAs between monocot and dicot, we found 40 nCrNAs were specific to Panicoideae, with more diversity to rice and wheat. Moreover, nearly half of them were located in intergenic regions. Divergency of Monocot and dicot was taken placed around 200 mya ago [56, 57], during the last ~3 million years, the size of the maize genome has expanded to 2.3 GB via a proliferation of long terminal repeat retrotransposons (LTR retrotransposons) in the intergenic region [43, 58–63]. The large number of Panicoideae-specific nCrNAs being located in the intergenic region of maize genome indicated that these

nCrNAs might also experience the selective pressure during maize evolution.

Pseudogene-derived nCrNAs

The distribution of maize-novel nCrNAs on chromosomes was varied, 51% loci were intergenic and 15% were in introns. Twenty-four percent of the novel nCrNAs originated from the 5' or 3' UTR of protein coding genes, moreover, a large number of their host genes are annotated as pseudogenes. Pseudogenes are complete or partial of protein-coding genes with no functional protein products [64, 65]. For the high conservation between pseudogenes and their parental protein-coding genes, some novel function of pseudogenes were verified [66]. Guo et al. found 145 pseudogene-derived siRNAs in rice genome, they can interact with their antisense RNAs (38 cases) or form double-strand RNAs with their adjacent pseudogenes (2 cases), some of them are abundant in specific rice developmental stages or physiological growth conditions, suggesting their potential roles in rice development [67]. Recent studies also established essential roles of pseudogene-derived lncRNAs in development and disease in mammalian [68–71]. Like the pseudogene-derived long non-coding RNA DUXAP10, can bind to histone demethylase lysine-specific demethylase 1 and silence the expression of *p21* and phosphatase and tensin homolog (*PTEN*), then promote colorectal cancer

cell growth [70]. In our results, 25 novel ncRNAs were originated from the 5' or 3' UTR of hypothetical or pseudogenes, many of them had tissue differential expression. Further analysis showed these pseudogene-derived ncRNAs had no potential to be precursor of miRNAs or siRNAs of maize, which indicated that they might play roles by themselves. Whether they can function as reverse regulator to the parental protein-coding genes of pseudogenes and how they play roles in maize development need much more evidence.

Expressions of tRNAs showed codon usage bias in leaves

In our results, four tRNAs had tissue-differential expression, which was similar to that have been found in other organisms. Like in human brain, nuclear encoded tRNAs had a high expression level than in other tissues [72]. In four sequenced cotton species, Wang et al. found that the pyrimidine-rich codons were used much more than purine-rich codons [73]. Tissue-specific expression of different tRNA species indicate that codon bias might play important roles in different tissues and developmental stages of organisms [72]. We found the expression of tissue-specific tRNAs accumulated more in the leaf organs of maize, with less or none in the root, meristem, and developing kernel. Leaf-related organs are the most productive parts for protein biogenesis, in which tRNAs preferred the codons of Leu, Gln, and Ser, the uncharged amino acids, which indicated the existence of codon bias in maize leaf development.

Stress-induced intermediate-size ncRNAs of maize

Plants are exposed to ever-changing environmental conditions including drought, freezing, and salinity. Modulation of gene expression is the key method for plants to respond, adjust and adapt to stress conditions. Besides protein-coding genes, non-coding RNAs also play essential roles in this process, such as miRNAs, trans-acting siRNAs, heterochromatic siRNAs as well as long non-coding RNAs. Molecular mechanisms of miRNAs and siRNAs in stress conditions have been well-studied [16, 74–77]. Functional roles of lncRNA in plant are gradually revealed, such as acting as ceRNAs to block the interaction between miRNAs and their target genes [78], or through chromatin modification or DNA methylation (RdMD) with other proteins [16, 79–82], like the cold-induced lncRNAs, *COOLAIR* and *COLDAIR*, they can reduce the H3K36me3 or H3K4me2 level at *FLOWERING LOCUS C (FLC)* and recruit polycomb repressive complex 2 (PRC2) to promote H3K27me3 accumulation at *FLC* to inhibit the expression of *FLC*, and lead to the flowering repression [80–83]. When we used PEG6000, ABA, and NaCl to simulate the drought and salt stress, we found many ncRNAs can be induced. For example, Zm-3, Zm-18, and Zm-73 were strongly induced at the

early stage of drought stress, and Zm-18 and Zm-73 were strongly induced at 3 h after NaCl treatment. The same phenomenon has been found in *Arabidopsis*, the lncRNA *npc60* showed a 100-fold increase under salt treatment [84], which indicated their roles in stress response. Further analysis showed that Zm-3, Zm-18 and Zm-73 had no potential to be precursor or sponge to maize known miRNAs, so whether these snoRNAs function as scaffold of protein complex or affect histone modification need much more evidence.

Conclusions

In this study, we provided a genome-wide identification and functional analysis of ncRNAs with a size range of 50–500 nt in maize. One-hundred and eleven novel ncRNAs were cloned and 40 ncRNAs were determined to be specific to Panicoideae. Target site analysis predicted a total of 38 2'-O-methylation and pseudouridylation modification sites of rRNAs and snRNAs. Furthermore, analysis of the expression profiles of the novel ncRNAs showed that 43 ncRNAs changed significantly during maize development, three ncRNAs can be strongly induced under drought and salt stress, suggesting their roles in maize stress response.

Methods

Construction of the cDNA library of 50–500 nt ncRNAs in maize

To construct a full-length intermediate-size (50–500 nt) ncRNA library of maize (*Zea mays* L.) inbred line Chang 7–2 (gift from Dr. Jihua Tang of Henan Agricultural University), total RNAs were isolated from seven tissues harvested at eleven developmental stages of maize seedlings and five stages of developing grains. First, the seeds of maize were surface-sterilized in 0.1% HgCl₂ for 10 min and rinsed in distilled water 10 times, then soaked in the dark overnight and placed on filter papers soaked with distilled water for another 24 h. The germinated seeds were transferred into 25 cm seedling pot with vermiculite at a photoperiod of 16/8 h (light/darkness) and 25/15 °C with a relative humidity of 65% in phytotron, the light intensity was 200 μmol photons m⁻² s⁻¹. Tissues of root, stem, blade, sheath & petiole, flag leaf, stem apex and hypocotyl were collected from eleven developmental stages of maize seedlings, the 3, 5, 8, 11, 14, 17, 19, 23, 26, 34, and 42 days after germination (DAG). Meanwhile, maize inbred line Chang 7–2 was also planted at the farmland of Henan Agricultural University (Zhengzhou, China). Immature seeds were collected from 5, 7, 9, 14 and 19 days after artificial pollination (DAP). Samples were immediately frozen in liquid nitrogen. TRIzol (Thermo Fischer Scientific) was used for RNA extraction. Equal amount of RNA

from 82 samples were mixed together to generate a homogenized library.

Small RNA library (50–500 nt) was constructed with the method of Deng et al. [44]. The small RNA fraction was isolated with a QIAGEN tip (QIAGEN) from 200 µg total RNAs. Messenger RNAs (mRNAs) and rRNAs were removed with the Ambion® Poly(A) Purist™ MAG (Thermo Fisher Scientific) and MICROExpress™ kits (Thermo Fisher Scientific) with the magnetic oligo(dT) cellulose, the probes used to fish out rRNAs were listed in Additional file 2: Table S2. RNAs were dephosphorylated with calf intestine alkaline phosphatase (Thermo Fisher Scientific) and then ligated to the 3-adaptor oligonucleotide (3 AD, with a restriction endonuclease site of SacI) by T4 RNA ligase (Thermo Fisher Scientific). The ligation product was split into two aliquots: one was treated with polynucleotide kinase (Thermo Fisher Scientific) for the uncapped RNA, and the other was treated with tobacco acid pyrophosphatase (Thermo Fisher Scientific) to remove 5'-end methyl-guanosine caps from capped RNA. The two parts of RNAs were ligated to the 5'-adaptor oligonucleotide (5 AD, with a restriction endonuclease site of KpnI). cDNA was prepared with Thermoscript™ reverse transcriptase (Thermo Fisher Scientific) at 50 °C, using oligo 3RT (in Additional file 2: Table S3) as the reverse transcription primer. The cDNA was PCR-amplified and digested with SphI and SacI (NEB) and cloned in pGEM-4Z vector (Promega). Transformation was performed with *E. coli* DH5α electrocompetent cells. Mono-clones were picked out and identified by PCR with 5CD (same DNA sequence with 5 AD) and 3RT (reverse complementary DNA sequence to 3 AD) primers (in Additional file 2: Table S3). PCR was performed at an annealing temperature of 55 °C with 20 cycles, PCR products were identified on 6% native PAGE, clones with same length of PCR product were reserved not more than 3, clones with different length of PCR product were picked out directly. Sanger sequencing method was used to identify different transcripts (Thermo Fisher Scientific).

Bioinformatics analysis

mfold (version 4.6) was used to predict the secondary structures of RNAs [85]. MEME (version 4.9.1) was used to predicate the conserved box of ncRNAs [86]. Potential targets of snoRNAs were predicated by snoScan and snoGPS with the same criteria as Li et al. [87]. The *Zea mays* L. genome sequence and annotation data were downloaded from maizesequence.org (version 5b+), genome and EST sequences of *Oryza sativa* (version 7) and *Sorghum bicolor* (version JGI Sbi1) were download from PlantGDB, genome of *Arabidopsis thaliana* was download from TAIR (version 10). The homologs or orthologs of ncRNAs were identified with BLASTN (version 2.2.22

+), the cutoff of sequence identity is 90% or 95%, the e-value is 1×10^{-10} . The online tool of miRNAfold was used to predicate whether ncRNAs could be precursors of miRNAs (<http://evryrna.ibisc.univ-evry.fr/miRNA-Fold>) [88], the parameters were default, species parameters were *Zea mays*.

Northern blot and 5'- and 3' - rapid amplification of cDNA end (RACE)

Northern blot was performed according the method of Li et al. [87]. RACE was performed as following, total RNAs were isolated from maize seedling with TRIzol method, RNAs were ligated to the 5'-adaptor (5 AD) and 3'-adaptor (3 AD) oligonucleotides with the method described above, cDNAs were reverse transcribed with Thermoscript™ reverse transcriptase (Thermo Fisher Scientific) at 50 °C, using oligo 3RT as the reverse transcription primer. Rapid amplification of cDNA ends (RACE) was performed by PCR amplification, with one primer specific to the ncRNA sequence and the other primer specific to the 5'-adaptor (5CD) or reverse complement to the 3'-adaptor (3RT) for 5'- or 3'-RACE, respectively. All primers were listed in Additional file 2: Table S3.

Stress treatments of maize seedling

Maize seeds were surface-sterilized in 0.1% HgCl₂ for 10 min, then rinsed in distilled water 10 times and soaked in the dark overnight at room temperature. Then the seeds were placed on filter papers soaked with distilled water until germination. The germinated seeds were transferred into trays (one seedling per well) with their roots soaked into the Hoagland solution (solution was changed every 2–3 days). When the seedling went to the three-leaf stage, 20% polyethylene glycol 6000 (PEG6000, Sigma Aldrich), 100 µM abscisic acid (ABA, Sigma Aldrich) and 200 mM NaCl (Sigma Aldrich) were added to simulate drought and salt stress. Roots were collected at 0, 1, 2 and 6 h after PEG6000 treatment, and at 0, 1, 3, 6, 12 and 24 h after ABA and NaCl treatment. Every treatment had three groups, each group had 3 duplicates. Root was immediately frozen in liquid nitrogen, total RNAs were extracted with TRIzol method.

Semi-quantitative RT-PCR

Total RNAs were extracted from different tissues, different developmental stages or roots under different stress conditions by TRIzol method. Synthesis of cDNAs was performed with 2 µg of total RNAs, 50 ng of random hexamer primers (for ncRNAs) or oligo d(T)15 (for protein coding genes) were used in reverse transcription with the first strand cDNA synthesis kit (Promega). PCR primers were designed using VectorNTI software (Thermo Fisher Scientific). The

primers used for ncRNAs and protein coding genes are listed in Additional file 2: Table S4 (for Quantitative real-time PCR) and Table S5 (for Semi-quantitative PCR). Semi-quantitative PCR was performed at an annealing temperature of 55 °C or 58 °C. U6 and tubulin were used as internal controls for ncRNAs and protein coding genes, respectively.

Additional files

Additional file 1: Information of all novel ncRNAs that being identified in maize inbred line Chang 7-2. (XLSX 35 kb)

Additional file 2: Figure S1. Functional class distribution of host genes of intron-origin ncRNAs. **Figure S2.** Expression of UTR-origin ncRNAs (A and B) and their host genes (C) as well as ORF-origin U1 snRNAs (D) in different tissues of maize. **Figure S3.** Functional class distribution of protein-coding genes of UTR and ORF-origin ncRNAs. **Figure S4.** Genome location of ncRNAs and their neighboring genes. **Figure S5.** Expression of intron-origin ncRNAs and their host genes in different tissues of maize. **Figure S6.** Expression of intergenic-origin ncRNAs (A) and their neighboring genes (B) in different tissues of maize. **Table S1.** Genome location of novel ncRNAs of maize. **Table S2.** Probe sets used for removal of known ncRNAs like rRNAs and U snRNAs. **Table S3.** Primer sets for 5' and 3' RACE. **Table S4.** Primer sets for real-time PCR. **Table S5.** Primer sets for Semi-quantitative RT-PCR. (PDF 409 kb)

Abbreviations

5-HT₂CR: Serotonin (5-HT) 2C Receptor; ABA: Abscisic acid; DAG: Day after germination; DAP: Days after pollination; FLC: FLOWERING LOCUS C; miRNA: microRNA; ncRNAs: Non-coding RNAs; ORFs: Open Reading Frames; PEG: Polyethylene glycol; piRNA: Piwi-interacting RNA; PRC2: Polycomb repressive complex 2; siRNA: Small interfering RNA; snoRNA: Small nucleolar RNA; snRNA: Small nuclear RNA; SRP: Signal recognition particle; TE-lincRNAs: Transposable-element-origin long intergenic non-coding RNAs; UTR: Untranslated region

Acknowledgments

Maize were planted in the farmland of Henan Agricultural University by Dr. Jihua Tang, we thank Dr. Jihua Tang for providing the seeds and immature grains of *Zea mays* inbred line Chang 7-2. Seedlings of maize were cultured in the phytotron of Nanyang Normal University. Samples collection was permitted by government-owned land. We also thank Dr. Xiangqian Guo for assistance with the bioinformatics analysis.

Funding

This work was supported by grants from the National Science Foundation of China (31172158 and 31100938), Young Elite Scientist Sponsorship Program by China Association for Science and Technology (YESS 20150026) and Scientific and Technological Projects of Henan Province (No. 112102110158).

Availability of data and materials

The 111 novel non-coding RNAs generated during the current study are available in NCBI nucleotides with ID number of KJ526628-KJ526738.

Authors' contributions

All authors contributed to various aspects of this work: YK supervised the project, YK, DL and HQ aided in study design, YK and DL obtained the research funds, maize planting and library construction were performed by DL, HQ and WQ, expression analysis were performed by WQ, XX, QJ and ZJ, TL and KZ performed the stress experiments, DL, RL and RC performed the bioinformatics analysis, RC and LB provided technical support. All authors read and approved the final manuscript.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details

¹China-UK-NYNU-RRes Joint Laboratory of insect biology, Henan Key Laboratory of Insect Biology in Funiu Mountain, Nanyang Normal University, 1638 Wolong Road, Nanyang 473061, Henan, China. ²Center for Agroforestry Mega Data Science and FAFU-UCR Joint Center for Horticultural Biology and Metabolomics, Haixia Institute of Science and Technology, Fujian Agriculture and Forestry University, Fuzhou 350002, China. ³Bioinformatics Laboratory and National Laboratory of Biomacromolecules, Institute of Biophysics, Chinese Academy of Sciences, Beijing 100101, China.

Received: 23 April 2018 Accepted: 21 September 2018

Published online: 04 October 2018

References

- Mercer TR, Dinger ME, Mattick JS. Long non-coding RNAs: insights into functions. *Nat Rev Genet.* 2009;10(3):155–9.
- Ponting CP, Oliver PL, Reik W. Evolution and functions of long noncoding RNAs. *Cell.* 2009;136(4):629–41.
- Budak H, Akpınar BA. Plant miRNAs: biogenesis, organization and origins. *Funct Integr Genomics.* 2015;15(5):523–31.
- Unver T, Budak H. Conserved microRNAs and their targets in model grass species *Brachypodium distachyon*. *Planta.* 2009;230(4):659–69.
- Lucas SJ, Budak H. Sorting the wheat from the chaff: identifying miRNAs in genomic survey sequences of *Triticum aestivum* chromosome 1AL. *PLoS One.* 2012;7(7):e40859.
- Budak H, Kantar M. Harnessing NGS and big data optimally: comparison of miRNA prediction from assembled versus non-assembled sequencing data—the case of the grass *Aegilops tauschii* Complex genome. *OMICS.* 2015;19(7):407–15.
- Huang L, Dong H, Zhou D, Li M, Liu Y, Zhang F, Feng Y, Yu D, Lin S, Cao J. Systematic identification of long non-coding RNAs during pollen development and fertilization in *Brassica rapa*. *Plant J.* 2018;96:203.
- Cho J. Transposon-derived non-coding RNAs and their function in plants. *Front Plant Sci.* 2018;9:600.
- Zhang B, Pan X, Anderson TA. Identification of 188 conserved maize microRNAs and their targets. *FEBS Lett.* 2006;580(15):3753–62.
- Ding D, Zhang L, Wang H, Liu Z, Zhang Z, Zheng Y. Differential expression of miRNAs in response to salt stress in maize roots. *Ann Bot.* 2009;103(1):29–38.
- Budak H, Kantar M, Bulut R, Akpınar BA. Stress responsive miRNAs and isomiRs in cereals. *Plant Sci.* 2015;235:1–13.
- Li D, Liu Z, Gao L, Wang L, Gao M, Jiao Z, Qiao H, Yang J, Chen M, Yao L, et al. Genome-wide identification and characterization of microRNAs in developing grains of *Zea mays* L. *PLoS One.* 2016;11(4):e0153168.
- Liu Z, Kumari S, Zhang L, Zheng Y, Ware D. Characterization of miRNAs in response to short-term waterlogging in three inbred lines of *Zea mays*. *PLoS One.* 2012;7(6):e39786.
- Zhao Z, Xue Y, Yang H, Li H, Sun G, Zhao X, Ding D, Tang J. Genome-wide identification of miRNAs and their targets involved in the developing internodes under maize ears by responding to hormone signaling. *PLoS One.* 2016;11(10):e0164026.
- Li L, Eichten SR, Shimizu R, Petsch K, Yeh CT, Wu W, Chetoor AM, Givan SA, Cole RA, Fowler JE, et al. Genome-wide discovery and characterization of maize long non-coding RNAs. *Genome Biol.* 2014;15(2):R40.
- Zhang W, Han Z, Guo Q, Liu Y, Zheng Y, Wu F, Jin W. Identification of maize long non-coding RNAs responsive to drought stress. *PLoS One.* 2014;9(6):e98958.
- Fan C, Hao Z, Yan J, Li G. Genome-wide identification and functional analysis of lincRNAs acting as miRNA targets or decoys in maize. *BMC Genomics.* 2015;16:793.
- Boerner S, McGinnis KM. Computational analysis of lincRNA from cDNA sequences. *Methods Mol Biol.* 2016;1402:255–69.

19. Lv Y, Liang Z, Ge M, Qi W, Zhang T, Lin F, Peng Z, Zhao H. Genome-wide identification and functional prediction of nitrogen-responsive intergenic and intronic long non-coding RNAs in maize (*Zea mays* L.). *BMC Genomics*. 2016;17:350.
20. Huanca-Mamani W, Arias-Carrasco R, Cardenas-Ninasivinchu S, Rojas-Herrera M, Sepulveda-Hermosilla G, Caris-Maldonado JC, Bastias E, Maracaja-Coutinho V. Long non-coding RNAs responsive to salt and boron stress in the hyper-arid Lluteno maize from Atacama Desert. *Genes*. 2018;9(3):170.
21. Ma J, Yan B, Qu Y, Qin F, Yang Y, Hao X, Yu J, Zhao Q, Zhu D, Ao G. Zm401, a short-open reading-frame mRNA or noncoding RNA, is essential for tapetum and microspore development and can regulate the floret formation in maize. *J Cell Biochem*. 2008;105(1):136–46.
22. Wang D, Qu Z, Yang L, Zhang Q, Liu ZH, Do T, Adelson DL, Wang ZY, Searle I, Zhu JK. Transposable elements (TEs) contribute to stress-related long intergenic noncoding RNAs in plants. *Plant J*. 2017;90(1):133–46.
23. Brown JW, Echeverria M, Qu LH. Plant snoRNAs: functional evolution and new modes of gene expression. *Trends Plant Sci*. 2003;8(1):42–9.
24. Falaleeva M, Welden JR, Duncan MJ, Stamm S. C/D-box snoRNAs form methylating and non-methylating ribonucleoprotein complexes: old dogs show new tricks. *Bioessays*. 2017;39(6):1600264.
25. Patterson CG, Roberts JT, King VM, Houserova D, Barnhill EC, Crucello A, Polska CJ, Brantley LW, Kaufman GC, Nguyen M, et al. Human snoRNA-93 is processed into a microRNA-like RNA that promotes breast cancer cell invasion. *NPJ Breast Cancer*. 2017;3:25.
26. Falaleeva M, Stamm S. Processing of snoRNAs as a new source of regulatory non-coding RNAs: snoRNA fragments form a new class of functional RNAs. *Bioessays*. 2013;35(1):46–54.
27. Ender C, Krek A, Friedlander MR, Beitzinger M, Weinmann L, Chen W, Pfeffer S, Rajewsky N, Meister G. A human snoRNA with microRNA-like functions. *Mol Cell*. 2008;32(4):519–28.
28. Brameier M, Herwig A, Reinhardt R, Walter L, Gruber J. Human box C/D snoRNAs with miRNA like functions: expanding the range of regulatory RNAs. *Nucleic Acids Res*. 2011;39(2):675–86.
29. Wang Y, Li H, Sun Q, Yao Y. Characterization of small RNAs derived from tRNAs, rRNAs and snoRNAs and their response to heat stress in wheat seedlings. *PLoS One*. 2016;11(3):e0150933.
30. Taft RJ, Glazov EA, Lassmann T, Hayashizaki Y, Carninci P, Mattick JS. Small RNAs derived from snoRNAs. *RNA*. 2009;15(7):1233–40.
31. Saraiya AA, Wang CC. snoRNA, a novel precursor of microRNA in *Giardia lamblia*. *PLoS Pathog*. 2008;4(11):e1000224.
32. Chang L, Li C, Lan T, Wu L, Yuan Y, Liu Q, Liu Z. Decreased expression of long non-coding RNA GAS5 indicates a poor prognosis and promotes cell proliferation and invasion in hepatocellular carcinoma by regulating vimentin. *Mol Med Rep*. 2016;13(2):1541–50.
33. Williams GT, Farzaneh F. Are snoRNAs and snoRNA host genes new players in cancer? *Nat Rev Cancer*. 2012;12(2):84–8.
34. Su H, Xu T, Ganapathy S, Shadfan M, Long M, Huang TH, Thompson I, Yuan ZM. Elevated snoRNA biogenesis is essential in breast cancer. *Oncogene*. 2014;33(11):1348–58.
35. Fang X, Yang D, Luo H, Wu S, Dong W, Xiao J, Yuan S, Ni A, Zhang KJ, Liu XY, et al. SNORD126 promotes HCC and CRC cell growth by activating the PI3K-AKT pathway through FGFR2. *J Mol Cell Biol*. 2016;9:243.
36. Yoshida K, Toden S, Weng W, Shigeyasu K, Miyoshi J, Turner J, Nagasaka T, Ma Y, Takayama T, Fujiwara T, et al. SNORA21 - an oncogenic small nucleolar RNA, with a prognostic biomarker potential in human colorectal cancer. *EBioMedicine*. 2017;22:68–77.
37. Espinoza CA, Allen TA, Hieb AR, Kugel JF, Goodrich JA. B2 RNA binds directly to RNA polymerase II to repress transcript synthesis. *Nat Struct Mol Biol*. 2004;11(9):822–9.
38. Nguyen VT, Kiss T, Michels AA, Bensaude O. 7SK small nuclear RNA binds to and inhibits the activity of CDK9/cyclin T complexes. *Nature*. 2001;414(6861):322–5.
39. Bratkovic T, Modic M, Camargo Ortega G, Drukker M, Rogelj B. Neuronal differentiation induces SNORD115 expression and is accompanied by post-transcriptional changes of serotonin receptor 2c mRNA. *Sci Rep*. 2018;8(1):5101.
40. Falaleeva M, Pages A, Matuszek Z, Hidmi S, Agranat-Tamir L, Korotkov K, Nevo Y, Eyraas E, Sperling R, Stamm S. Dual function of C/D box small nucleolar RNAs in rRNA modification and alternative pre-mRNA splicing. *Proc Natl Acad Sci U S A*. 2016;113(12):E1625–34.
41. Kishore S, Khanna A, Zhang Z, Hui J, Balwierz PJ, Stefan M, Beach C, Nicholls RD, Zavolan M, Stamm S. The snoRNA MBII-52 (SNORD 115) is processed into smaller RNAs and regulates alternative splicing. *Hum Mol Genet*. 2010;19(7):1153–64.
42. Kishore S, Stamm S. The snoRNA HBII-52 regulates alternative splicing of the serotonin receptor 2C. *Science*. 2006;311(5758):230–2.
43. Schnable PS, Ware D, Fulton RS, Stein JC, Wei F, Pasternak S, Liang C, Zhang J, Fulton L, Graves TA, et al. The B73 maize genome: complexity, diversity, and dynamics. *Science*. 2009;326(5956):1112–5.
44. Deng W, Zhu X, Skogerbo G, Zhao Y, Fu Z, Wang Y, He H, Cai L, Sun H, Liu C, et al. Organization of the *Caenorhabditis elegans* small non-coding transcriptome: genomic features, biogenesis, and expression. *Genome Res*. 2006;16(1):20–9.
45. Liu TT, Zhu D, Chen W, Deng W, He H, He G, Bai B, Qi Y, Chen R, Deng XW. A global identification and analysis of small nucleolar RNAs and possible intermediate-sized non-coding RNAs in *Oryza sativa*. *Mol Plant*. 2013;6(3):830–46.
46. Yuan G, Klambt C, Bachellerie JP, Brosius J, Huttenhofer A. RNomics in *Drosophila melanogaster*: identification of 66 candidates for novel non-messenger RNAs. *Nucleic Acids Res*. 2003;31(10):2495–507.
47. Kiss T. Small nucleolar RNA-guided post-transcriptional modification of cellular RNAs. *EMBO J*. 2001;20(14):3617–22.
48. Szell M, Bata-Csorgo Z, Kemeny L. The enigmatic world of mRNA-like ncRNAs: their role in human evolution and in human diseases. *Semin Cancer Biol*. 2008;18(2):141–8.
49. Storz G, Altuvia S, Wassarman KM. An abundance of RNA regulators. *Annu Rev Biochem*. 2005;74:199–217.
50. Prasanth KV, Spector DL. Eukaryotic regulatory RNAs: an answer to the 'genome complexity' conundrum. *Genes Dev*. 2007;21(1):11–42.
51. Dudley M, Poethig RS. The effect of a heterochronic mutation, Teopod2, on the cell lineage of the maize shoot. *Development*. 1991;111(3):733–9.
52. Poethig RS. Heterochronic mutations affecting shoot development in maize. *Genetics*. 1988;119(4):959–73.
53. Leung AK, Sharp PA. MicroRNA functions in stress responses. *Mol Cell*. 2010;40(2):205–15.
54. Paterson AH, Bowers JE, Bruggmann R, Dubchak I, Grimwood J, Gundlach H, Haberer G, Hellsten U, Mitros T, Poliakov A, et al. The Sorghum bicolor genome and the diversification of grasses. *Nature*. 2009;457(7229):551–6.
55. Wang B, Regulski M, Tseng E, Olson A, Goodwin S, McCombie WR, Ware D. A comparative transcriptional landscape of maize and sorghum obtained by single-molecule sequencing. *Genome Res*. 2018;28(6):921–32.
56. Wolfe KH, Gouy M, Yang YW, Sharp PM, Li WH. Date of the monocot-dicot divergence estimated from chloroplast DNA sequence data. *Proc Natl Acad Sci U S A*. 1989;86(16):6201–5.
57. Gaut BS. Evolutionary dynamics of grass genomes. *New Phytol*. 2002;154(1):15–28.
58. Lai J, Ma J, Swigonova Z, Ramakrishna W, Linton E, Llac V, Tanyolac B, Park YJ, Jeong OY, Bennetzen JL, et al. Gene loss and movement in the maize genome. *Genome Res*. 2004;14(10A):1924–31.
59. SanMiguel P, Gaut BS, Tikhonov A, Nakajima Y, Bennetzen JL. The paleontology of intergene retrotransposons of maize. *Nat Genet*. 1998;20(1):43–5.
60. Bruggmann R, Bharti AK, Gundlach H, Lai J, Young S, Pontaroli AC, Wei F, Haberer G, Fuks G, Du C, et al. Uneven chromosome contraction and expansion in the maize genome. *Genome Res*. 2006;16(10):1241–51.
61. Biemont C, Vieira C. Genetics: junk DNA as an evolutionary force. *Nature*. 2006;443(7111):521–4.
62. Kazazian HH Jr. Mobile elements: drivers of genome evolution. *Science*. 2004;303(5664):1626–32.
63. Kidwell MG, Lisch D. Transposable elements as sources of variation in animals and plants. *Proc Natl Acad Sci U S A*. 1997;94(15):7704–11.
64. Prince VE, Pickett FB. Splitting pairs: the diverging fates of duplicated genes. *Nat Rev Genet*. 2002;3(11):827–37.
65. Mighell AJ, Smith NR, Robinson PA, Markham AF. Vertebrate pseudogenes. *FEBS Lett*. 2000;468(2–3):109–14.
66. Milligan MJ, Lipovich L. Pseudogene-derived lncRNAs: emerging regulators of gene expression. *Front Genet*. 2014;5:476.
67. Guo X, Zhang Z, Gerstein MB, Zheng D. Small RNAs originated from pseudogenes: cis- or trans-acting? *PLoS Comput Biol*. 2009;5(7):e1000449.
68. Ma H, Ma T, Chen M, Zou Z, Zhang Z. The pseudogene-derived long non-coding RNA SFTA1P suppresses cell proliferation, migration, and invasion in gastric cancer. *Biosci Rep*. 2018;38(2):BSR20171193.
69. Zhang H, Xiong Y, Xia R, Wei C, Shi X, Nie F. The pseudogene-derived long noncoding RNA SFTA1P is down-regulated and suppresses cell migration and invasion in lung adenocarcinoma. *Tumour Biol*. 2017;39(2):1010428317691418.

70. Lian Y, Xu Y, Xiao C, Xia R, Gong H, Yang P, Chen T, Wu D, Cai Z, Zhang J, et al. The pseudogene derived from long non-coding RNA DUXAP10 promotes colorectal cancer cell growth through epigenetically silencing of p21 and PTEN. *Sci Rep*. 2017;7(1):7312.
71. Ma HW, Xie M, Sun M, Chen TY, Jin RR, Ma TS, Chen QN, Zhang EB, He XZ, De W, et al. The pseudogene derived long noncoding RNA DUXAP8 promotes gastric cancer cell proliferation and migration via epigenetically silencing PLEKHO1 expression. *Oncotarget*. 2017;8(32):52211–24.
72. Dittmar KA, Goodenbour JM, Pan T. Tissue-specific differences in human transfer RNA expression. *PLoS Genet*. 2006;2(12):e221.
73. Wang L, Xing H, Yuan Y, Wang X, Saeed M, Tao J, Feng W, Zhang G, Song X, Sun X. Genome-wide analysis of codon usage bias in four sequenced cotton species. *PLoS One*. 2018;13(3):e0194372.
74. Sunkar R, Chinnusamy V, Zhu J, Zhu JK. Small RNAs as big players in plant abiotic stress responses and nutrient deprivation. *Trends Plant Sci*. 2007;12(7):301–9.
75. Covarrubias AA, Reyes JL. Post-transcriptional gene regulation of salinity and drought responses by plant microRNAs. *Plant Cell Environ*. 2010;33(4):481–9.
76. Zhai L, Liu Z, Zou X, Jiang Y, Qiu F, Zheng Y, Zhang Z. Genome-wide identification and analysis of microRNA responding to long-term waterlogging in crown roots of maize seedlings. *Physiol Plant*. 2013;147(2):181–93.
77. Zhang Z, Lin H, Shen Y, Gao J, Xiang K, Liu L, Ding H, Yuan G, Lan H, Zhou S, et al. Cloning and characterization of miRNAs from maize seedling roots under low phosphorus stress. *Mol Biol Rep*. 2012;39(8):8137–46.
78. Wang J, Yu W, Yang Y, Li X, Chen T, Liu T, Ma N, Yang X, Liu R, Zhang B. Genome-wide analysis of tomato long non-coding RNAs and identification as endogenous target mimic for microRNA in response to TYLCV infection. *Sci Rep*. 2015;5:16946.
79. Zhang H, Chen X, Wang C, Xu Z, Wang Y, Liu X, Kang Z, Ji W. Long non-coding genes implicated in response to stripe rust pathogen stress in wheat (*Triticum aestivum* L.). *Mol Biol Rep*. 2013;40(11):6245–53.
80. Michaels SD, Amasino RM. FLOWERING LOCUS C encodes a novel MADS domain protein that acts as a repressor of flowering. *Plant Cell*. 1999;11(5):949–56.
81. Heo JB, Sung S. Vernalization-mediated epigenetic silencing by a long intronic noncoding RNA. *Science*. 2011;331(6013):76–9.
82. Csorba T, Questa JI, Sun Q, Dean C. Antisense COOLAIR mediates the coordinated switching of chromatin states at FLC during vernalization. *Proc Natl Acad Sci U S A*. 2014;111(45):16160–5.
83. Marquardt S, Raitskin O, Wu Z, Liu F, Sun Q, Dean C. Functional consequences of splicing of the antisense transcript COOLAIR on FLC transcription. *Mol Cell*. 2014;54(1):156–65.
84. Ben Amor B, Wirth S, Merchan F, Laporte P, d'Aubenton-Carafa Y, Hirsch J, Maizel A, Mallory A, Lucas A, Deragon JM, et al. Novel long non-protein coding RNAs involved in Arabidopsis differentiation and stress responses. *Genome Res*. 2009;19(1):57–69.
85. Zuker M. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res*. 2003;31(13):3406–15.
86. Bailey TL, Elkan C. Fitting a mixture model by expectation maximization to discover motifs in biopolymers. *Proc Int Conf Intell Syst Mol Biol*. 1994;2:28–36.
87. Li D, Wang Y, Zhang K, Jiao Z, Zhu X, Skogerboe G, Guo X, Chinnusamy V, Bi L, Huang Y, et al. Experimental RNomics and genomic comparative analysis reveal a large group of species-specific small non-message RNAs in the silkworm *Bombyx mori*. *Nucleic Acids Res*. 2011;39(9):3792–805.
88. Tav C, Tempel S, Poligny L, Tahi F. miRNAfold: a web server for fast miRNA precursor prediction in genomes. *Nucleic Acids Res*. 2016;44(W1):W181–4.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

