Large-scale translatome profiling annotates the functional genome and reveals the key role of genic 3’ untranslated regions in translatomic variation in plants

Wanchao Zhu1,4, Jing Xu1,4, Sijia Chen1, Jian Chen2, Yan Liang1, Cuijie Zhang3, Qing Li1, Jinshe Lai2 and Lin Li1,*

1National Key Laboratory of Crop Genetic Improvement, Huazhong Agricultural University, Wuhan 430070, China
2State Key Laboratory of Agrobiotechnology and National Maize Improvement Center, Department of Plant Genetics and Breeding, China Agricultural University, Beijing, 100193, China
3Novogene Co., Ltd., Beijing, 100015, China
4These authors contributed equally to this article.

*Correspondence: Lin Li (hzaulilin@mail.hzau.edu.cn)
https://doi.org/10.1016/j.xplc.2021.100181

ABSTRACT

The translatome, a profile of the translational status of genetic information within cells, provides a new perspective on gene expression. Although many plant genomes have been sequenced, comprehensive translatomic annotations are not available for plants due to a lack of efficient translatome profiling techniques. Here, we developed a new technique termed 3’ ribosome-profiling sequencing (3’Ribo-seq) for reliable, robust translatomic profiling. 3’Ribo-seq combines polysome profiling and 3’ selection with a barcoding and pooling strategy. Systematic translatome profiling of different tissues of Arabidopsis, rice, and maize using conventional ribosome profiling (Ribo-seq) and 3’Ribo-seq revealed many novel translational genomic loci, thereby complementing functional genome annotation in plants. Using the low-cost, efficient 3’Ribo-seq technique and genome-wide association mapping of translatome expression (eGWAS), we performed a population-level dissection of the translatomes of 159 diverse maize inbred lines and identified 1,777 translational expression quantitative trait loci (eQTLs). Notably, local eQTLs are significantly enriched in the 3’ untranslated regions of genes. Detailed eQTL analysis suggested that sequence variation around the polyadenylation (polyA) signal motif plays a key role in translatomic variation. Our study provides a comprehensive translatome annotation of plant functional genomes and introduces 3’Ribo-seq, which paves the way for deep translatomic analysis at the population level.

Key words: translatome, Ribo-seq, 3’Ribo-seq, translational eGWAS, functional genomics

Zhu W., Xu J., Chen S., Chen J., Liang Y., Zhang C., Li Q., Lai J., and Li L. (2021). Large-scale translatome profiling annotates the functional genome and reveals the key role of genic 3’ untranslated regions in translatomic variation in plants. Plant Comm. 2, 100181.

INTRODUCTION

Proteins are the final executors of many biological processes in organisms (Mergner et al., 2020). The genome—the initial vector of genetic information—is transcribed into the transcriptome and further translated into the proteome. These processes are regulated in a highly complex manner at the level of mRNA splicing (Bava et al., 2013), transport (Rousseau et al., 1996), localization (Holt and Schuman, 2013), translation (Cenik et al., 2015), and protein modification (Deschenes-Simard et al., 2014). However, despite the tremendous importance of the proteome, proteomic research is generally less advanced than nucleic acid research.

A major goal of proteomic research is to accurately profile proteins genome-wide. With the advent of high-throughput techniques, RNA sequencing (RNA-seq) is widely used as a global
transcript profiling technique to investigate the transcriptome. However, proteomic variation does not usually occur in concert with transcriptomic variation because the latter can be buffered by multifaceted regulatory mechanisms at the protein level (Laurent et al., 2010; Vogel and Marcotte, 2012; Khan et al., 2013; Jiang et al., 2019). Therefore, various proteomic techniques have been developed, including gel-based methods and mass spectrometry (MS)-based approaches. However, no techniques have been developed, including gel-based methods and mass spectrometry (MS)-based approaches. However, no more than 10,000 highly expressed proteins can be detected using current proteomic techniques (Khan et al., 2013; Jiang et al., 2019). Given that the number of expressed transcripts is much larger than 10,000, other alternative techniques are needed for the accurate detection or better estimation of the translatomic abundance of genes.

Translation, an important link between RNA and protein, governs protein production in response to many physiological conditions (Jackson et al., 2010; Hershey et al., 2012). The generalized translation complex contains translating mRNAs, tRNAs, ribosomes, translation factors, and some elements that directly participate in the translational process (Zhao et al., 2019a, 2019b). The term “translatome” usually refers to all translating mRNAs, which is a better predictor of protein expression than the transcriptome (Cenik et al., 2015; Chassé et al., 2016; Zhao et al., 2019a, 2019b). Various techniques have been developed to measure global translation levels; these include polysome profiling, ribosome profiling (Ribo-seq), and translating ribosome affinity purification (TRAP) (Ingolia et al., 2009; Heiman et al., 2014; Zhao et al., 2019a, 2019b). Polysome profiling based on sucrose gradient ultracentrifugation shows low sensitivity for small samples and is widely used to detect overall changes in translation. Ribo-seq provides accurate information about ribosome positions and density by sequencing ribosome-protected RNA fragments (RPFs) (Ingolia et al., 2009). However, the complicated, costly experimental procedures and the short lengths of RPFs have limited its usage to a small number of samples in a few sophisticated laboratories. TRAP, based on the affinity of tagged Rpl25p (the large ribosomal subunit protein), was devised to isolate translating mRNAs from specific tissues or cells (Heiman et al., 2014). Nevertheless, the transgenic process required for TRAP is time consuming and costly. Thus, an efficient, inexpensive, and time-saving technique with high throughput for profiling the translatome is currently lacking.

Three-prime untranslated regions (3′ UTRs) are vital parts of mRNA transcripts, as they determine the stability, localization, translation, and degradation of mRNA, thus influencing protein synthesis (Matoukova et al., 2012). A number of 3′-focused profiling techniques have been developed, such as polyadenylation (polyA) site sequencing (PAS-seq) (Shepard et al., 2011), polyA-position profiling by sequencing (3P-seq) (Jan et al., 2011), polyA sequencing (polyA-seq) (Derti et al., 2012), 3′T-fill (Wilkening et al., 2013), 3′ region extraction and deep sequencing (3′READS) (Hoque et al., 2013), and polyA-test RNA-seq (PAT-seq) (Harrison et al., 2015). The rapid development of these techniques (Jan et al., 2011; Shepard et al., 2011; Derti et al., 2012; Hoque et al., 2013; Wilkening et al., 2013; Harrison et al., 2015), along with powerful barcoding systems for single-cell sequencing (Hashimshony et al., 2012, 2016; Klein et al., 2015; Macosko et al., 2015; Arguel et al., 2017) and unique molecular identifiers (UMIs) (Kivioja et al., 2012; Islam et al., 2014; Smith et al., 2017), has paved the way for reducing the cost and increasing the efficiency of current ribosome profiling techniques.

Plants are a large, diverse group of species worldwide. They cover the entire Earth ecosystem and are major sources of energy for all other life forms. The plant functional genome has been the focus of intensive research efforts, and draft sequences of thousands of plant genomes have been generated (Twyford, 2019). Near-complete genomes for the model plants Arabidopsis, rice, and maize have been obtained as well (Kawahara et al., 2013; Berardini et al., 2015; Jiao et al., 2017). However, since the release of the first draft of the maize reference genome 10 years ago, efforts at functional genome annotation remain incomplete and are ongoing, thanks to the development of new genome detection techniques (Li et al., 2014; Wang et al., 2016; Jiao et al., 2017; Han et al., 2019). A major topic of functional genomics research is the dissection of proteomic variation and its underlying mechanisms, given the direct link between the proteome and the phenotype. However, the landscape and molecular mechanisms of proteomic variation are largely unknown due to the limitations of current proteome profiling techniques. A complete annotated reference translatome, along with the genetic dissection of mechanisms associated with translatomic variation in a natural breeding population, could provide an unprecedented opportunity for the complete annotation of the functional genome in plants.

Here, we developed 3′Ribo-seq, a new translatome profiling technique that combines polysome profiling, 3′ selection, and a new strategy for barcoding and pooling. A series of comprehensive evaluations of UMI bias, read distribution, repeatability, number of translated genes, and bulked batch effects revealed the reliability and robustness of 3′Ribo-seq for translatome profiling at both the individual and population levels. We profiled the translatomes of multiple organs from three model plant species (Arabidopsis, rice, and maize) by combining conventional Ribo-seq and 3′Ribo-seq, providing a more complete translatome landscape and a new functional genome annotation in plants. A comparison between conventional Ribo-seq and 3′Ribo-seq demonstrated that 3′Ribo-seq is not only as efficient as Ribo-seq for translation detection but also less expensive. Finally, we conducted a large-scale analysis of the translatome using leaf samples from 159 diverse inbred maize lines at the population level by 3′Ribo-seq. Genome-wide mapping of translatome expression uncovered 1777 expression quantitative trait loci (eQTLs) associated with translatomic variation in plants. Therefore, 3′Ribo-seq is a reliable, robust, cost-effective translatome profiling technique that provides a panoramic landscape view of the plant translatome.

RESULTS AND DISCUSSION

3′Ribo-seq shows reliability and robustness for translatomic profiling

To accurately and effectively profile translatomic variation, we designed a technique termed 3′Ribo-seq that combines polysome profiling with 3′ selective library sequencing (supplementary materials). The reverse transcription reactions are performed
To determine whether 3’Ribo-seq can be used to accurately investigate the translatome, we conducted a comprehensive multi-omics comparison. The 3’Ribo-seq data were highly correlated with the FL_Ribo-seq data (Pearson correlation coefficient: 0.85–0.87) and also showed a significant correlation with the protein MS data (Pearson correlation coefficient: 0.57–0.60) (Figure 1F and 1H, Supplemental Figure 3B and 3C).

Importantly, far more genes were detected by 3’Ribo-seq than by protein MS (Figure 1G). In addition, most genes (93.3%) in the 3’Ribo-seq data could be detected by the two conventional methods: FL_Ribo-seq and Ribo-seq (Supplemental Figure 4). However, 3’Ribo-seq missed hundreds of genes detected by FL_Ribo-seq and Ribo-seq (Supplemental Figure 4), perhaps because of unsaturated sequencing or 3’ polyA selection by 3’Ribo-seq. The 3’Ribo-seq data were significantly more highly correlated with the MS data than were the FL_Ribo-seq and RNA-seq data (Figure 1H–1I), perhaps because of the non-uniform distribution and dynamic features of ribosomes along the coding sequence (CDS) (Figure 1J). 3’Ribo-seq captures only the 3’ mRNA fragments bound by ribosomes, which are more likely to accurately and instantaneously profile the completed translation of mRNA into proteins.

The 3’ selection strategy has been widely used for RNA-seq and many new transcriptome profiling techniques (Jan et al., 2011; Shepard et al., 2011; Derti et al., 2012; Hoque et al., 2013; Wilkening et al., 2013; Harrison et al., 2015; Chen et al., 2021). These techniques are reliable not only for analyzing alternative cleavage and polyA (APA) but also for accurately quantifying gene expression levels. With rapid advances in barcoding and UMI techniques for large-scale single-cell sequencing (Hashimshony et al., 2012, 2016; Klein et al., 2015; Macosko et al., 2015; Arguel et al., 2017), it has become feasible to combine 3’ selection and barcode systems for population-level transcriptome profiling at a relatively low cost. Compared with the current 3’ selection techniques, 3’ Ribo-seq is the first 3’ selection method for ribosome profiling. Unlike current translatome techniques such as polysome profiling, Ribo-seq, and TRAP, 3’Ribo-seq focuses on the 3’ ends of translated mRNAs. Although the ribosome positions and the translatomes of some specific tissues or cells cannot be investigated using this technique (Ingolia et al., 2009; Derti et al., 2012; Heiman et al., 2014), the 3’ end of a translated mRNA sequence can be used to quantify gene translation for more samples simultaneously. The reduced amount of total effective sequencing data and reduced demand for library construction of pooled samples dramatically decrease the cost of this technique to approximately $33.55 per sample (Supplemental Table 1), which is far less than the costs of Ribo-seq and protein MS (Figure 1K). Thus, 3’Ribo-seq is a better choice for profiling the translatomes of a large number of samples and exploring translational variation at the population level.

3’Ribo-seq shows high repeatability between technical replicates, has high accuracy, is more cost-effective for translatome profiling, and detects more translated genes than protein MS. Thus, 3’Ribo-seq is a new, reliable, robust tool for profiling the translatome and annotating the functional genome.

A comprehensive annotation of the functional genomes of plants

To profile the translatome landscape of plants, we performed 3’Ribo-seq and Ribo-seq (or ribosome footprinting; Ingolia et al., 2009) on different tissues of Arabidopsis thaliana (Columbia), rice (Nipponbare), and maize (B73) to annotate the functional genomes of plants. We collected 62 samples, including seven tissues from Arabidopsis (five mature and two...
seedling tissues), nine tissues from rice (five from V2-stage seedlings and four from booting-stage plants), and seven tissues from maize (four tissues from 14-day-old seedlings, ears and tassels from V12 plants, and kernels at 9 days after pollination), most with two biological replicates, for both 30Ribo-seq and Ribo-seq (Supplemental Figure 5 and Supplemental Table 2). We also performed FL_Ribo-seq using one root and one stem tissue from Arabidopsis, one stem and one leaf tissue from rice, and two ear tissues from maize. As expected, the 3' Ribo-seq data were highly correlated with the FL_Ribo-seq data for all species (Pearson = 0.91 in Arabidopsis; Pearson = 0.89 in rice; Pearson = 0.89 in maize; Supplemental Figure 6). The Ribo-seq data showed clear three-nucleotide periodicity, and RPFs were enriched surrounding the start and stop codons (Supplemental Figure 7). These results point to the high quality of data produced from 3'Ribo-seq and Ribo-seq.

To further evaluate the robustness of the 3'Ribo-seq and Ribo-seq data, we performed PCA, which showed reliable clustering of distinct tissues (after removing the batch effect) in all three plant species (Figure 2A–2C and Supplemental Figure 8). We also performed RNA-seq of the same samples from multiple
tissues and protein MS of leaf samples from all three species (Supplemental Table 2). Across different tissues, the Ribo-seq data were more highly correlated with the RNA-seq data than were the 3′Ribo-seq data, suggesting that the 3′Ribo-seq data are more divergent from the RNA-seq data (Figure 2D and Supplemental Figure 9A–9D). The 3′Ribo-seq data showed slightly higher correlations with protein MS data in leaf samples than did the Ribo-seq and RNA-seq data (Figure 2E and 2F and Supplemental Figure 9E and 9F). Importantly, when we compared the 3′Ribo-seq and Ribo-seq data in the three species, the correlations were mainly distributed in the range from 0.6 to 0.7 (Figure 2G), indicating that the two techniques are not only related but also divergent. Notably, slightly fewer genes were detected by 3′Ribo-seq than by Ribo-seq, perhaps because translating loci without standard 3′ UTRs were excluded (Figure 2H).

The availability of two datasets across different tissues/stages in all three plant species provided us with an unprecedented chance to update functional genome annotations in plants. More comprehensive translational annotations were obtained for all three plant species (supplemental data 1, 2, 3, 4, 5, and 6). In total, 26,404,
40,558, and 41,657 translated isoforms were detected in Arabidopsis, rice, and maize, respectively (fragments per kilobase per million [FPKM] > 0.5, Figure 3A and supplemental data 7). Although most translated isoforms were consistent with the annotated isoforms of the transcriptome (using the known isoforms as a reference), 896, 4,602, and 2,572 new isoforms were found to be translated in Arabidopsis, rice, and maize, respectively (Figure 3A). As expected, most isoforms were from protein-coding genes, although 645, 777, and 338 genes that were annotated as long non-coding RNAs (lncRNAs) or...
In *Arabidopsis*, we detected more translated genes compared with a recent study that used protein MS on 30 tissues (18 210 genes) (Supplemental Figure 10A) (Mergner et al., 2020). In maize, 17 862 proteins (corresponding to 15 583 genes in the V4 reference genome) were previously detected using protein MS across 33 tissues (Walley et al., 2016), which is also far less than the number of translated genes detected by our translatome annotation (Supplemental Figure 10C). We also identified many intergenic isoforms derived from unannotated gene loci in plants (Figure 3C). These intergenic loci were usually detected in different tissues or samples of the three species, pointing to possible tissue-specific translation of these newly identified translated loci (Figure 3D–3F). As expected, most tissues of the three species could be clustered reasonably based on these loci (Figure 3D–3F), highlighting the reliability of the identified loci. We also compared the lengths of intergenic loci annotated by our study with the lengths of reference annotated loci and found that the newly annotated intergenic loci were significantly shorter than those of the reference annotation (Figure 3G), indicating that shorter amino acid chains or small peptides might be translated. Both 3’Ribo-seq and Ribo-seq identified novel translated loci. We randomly selected 20 translated loci that were detected by the two methods and amplified their cDNAs from RNA obtained by polyscale profiling. Of these 20 cDNAs, 13 were successfully amplified and produced bands of the expected size (Figure 3H).

Interestingly, some annotated noncoding RNAs were also found to be translated, perhaps because they were wrongly annotated, produced small peptides, or were contaminated by RNAs as false-positives. These alternatives could be clarified by searching against these RNA loci using a comprehensive MS dataset (Guttman et al., 2013). Furthermore, we used the mass spectra of maize seedlings to search against the newly annotated intergenic loci and identified 57 uncharacterized proteins, 52 of which could be quantified (Figure 3I and Supplemental Data 8). Based on the quantified proteins, four samples of seedlings (BRA1, BRA2, BRAB, and BRB) were clustered in a reasonable pattern, showing the reliability of the identified proteins (Figure 3I). These newly identified translated loci can complement the current reference genome annotation and can serve as additional functional research targets in the future. Finally, we performed Gene Ontology (GO) annotation based on sequences homologous to these newly identified loci and found that they were significantly associated with the GO terms “cell part,” “cellular process,” “metabolic process,” and “response to stimulus” (Supplemental Figure 11).

**3’ UTRs play key roles in regulating translatomic variation in plants**

Compared with conventional Ribo-seq, 3’Ribo-seq is a cost-saving method for profiling the translatomic landscape, providing an opportunity to profile natural translatomic variations and uncover the underlying genetic mechanisms. We collected 159 diverse maize inbred lines (Supplemental Data 9) and subjected them to 3’Ribo-seq profiling using seedling leaves at the V7 stage. More than 10 000 genes were found to be translated in most samples (Figure 4A). To eliminate any potential batch effect, some samples were analyzed with two or three replicates in different libraries. Although these natural populations exhibited extensive translatomic variation, the population structure could not be clearly classified based on translatomic variation (Figure 4B), suggesting that highly complex mechanisms are associated with population-level translatomic variation in maize.

To dissect the genetic mechanisms that underlie translatomic variation in plants, we performed expression quantitative trait locus (eQTL) mapping by genome-wide association analysis of diverse maize inbred lines. We identified 1777 eQTLs (Supplemental Data 10) associated with translatomic variation (< 8.03e-08, Figure 4C and 4D). Of these eQTLs, genes, most (95%) are under only distal regulation, 1% are under both distal and local regulation, and only 4% are under local regulation (Figure 4E). However, no distal eQTL hotspots for the regulation of population-level translatomic variation were identified. eQTL mapping is an efficient method for dissecting molecular mechanisms that underlie variation in gene abundance (Kliebenstein, 2009). Previous studies involving eQTL dissection of transcriptional variation identified numerous cis-eQTLs and a cis and trans eQTL ratio of approximately 3:7 in maize (Li et al., 2013). However, our first eQTL mapping of translatome variation at the population level revealed a different regulatory scenario in which only 4%–5% of the eQTLs were found to be cis-eQTLs; these results are similar to previous results at the proteome level (Blein-Nicolas et al., 2020). Translational regulation is the most important and complicated regulatory step and accounts for more than half of all regulatory amplitudes (RNA synthesis, RNA degradation, protein synthesis, and protein degradation) estimated by omics measurements and mathematical models (Schwanhäusser et al., 2011; Zhao et al., 2019a). Thus, ultra-complex regulatory mechanisms involving thousands of regulators may confer translatomic variation in plants. Alternatively, post-transcriptional regulatory mechanisms may be dominant compared with the local effects of genomic structural variations on protein biogenesis, or other noisy factors may interfere with the translation process. Overall, our study represents the first translatome eQTL mapping in plants and unravels ultra-complex regulatory mechanisms that underlie population-level translational variation, which differs greatly from that at the transcriptome level.

Of the locally regulated eQTLs, the most significant SNPs are located in the 3’ UTRs of their target genes, followed by introns, whereas the fewest significant SNPs are located in the exon regions of their target genes (Figure 4F and Supplemental Figure 12A). Mutations in the 3’ UTRs of genes can significantly affect translational expression (Supplemental Figure 12B–12D). To investigate whether any motif located in the 3’ UTR contributes to the local regulation of translatomic variation, we aligned the last 2-kb sequences of genes with detectable locally regulated eQTLs and subjected them to motif enrichment analysis. We identified a specific motif containing polyA in the 3’ UTRs of genes that was significantly associated with locally regulated eQTLs (Supplemental Figure 12E).

The sequence elements involved in polyA in the 3’ regions of genes contain a polyA signal, a polyA site, and a downstream element (DSE) (Figure 5A) (Proudfoot, 2011). These three...
sequence elements in the 3' UTRs of genes are vital for the polyA of mature RNAs. The polyA signal is usually a conserved AATAAA motif, although there are many other variants. The polyA signal is retained during polyA. The enrichment of cis-translational eQTL regions suggests that mutations near the polyA signal may alter the translational levels of genes. The polyA site is usually located approximately 20 nt downstream of the polyA signal and is the cleavage site where the polyA tail is attached to the mRNA (Manley and Takagaki, 1996). Of the 74 genes with local regulation, 31 genes had a detectable polyA signal in their genic 3' UTRs. We identified genomic variation around the polyA signal in 31 genes (Supplemental Table 3), and it showed dramatic association signals with translation-level variation.

We identified one significant translational association signal that peaked around 20 nt downstream of the polyA signal and may be the polyA site (Figure 5B). Notably, one other significant translational association peak was identified around 20 nt upstream of the polyA signal, suggesting that a novel motif probably associated with polyA may be located there and that mutation in this region affects translation (Figure 5B). For example, an SNP (A to C) 20 nt upstream of the polyA signal in the gene Zm00001d012635 is significantly associated with translation-level variation (Figure 5C). An SNP (C to A) 19 nt upstream of the polyA signal in the gene Zm00001d053834 is also associated with translation-level variation (Figure 5D). Interestingly, such genomic SNP associations are more likely to occur at the translatome level than at the transcriptome level (association P values: 0.76 at the transcriptome level versus 4.2E-04 at the translatome level for Zm00001d053834; 9.6E-04 versus 1.0E-04 for Zm00001d012635; 0.52 versus 3.1E-02 for Zm0001d005814). By removing redundant SNPs with significant linkage disequilibrium, we identified different haplotypes of genomic variation around the polyA elements of Zm00001d012635 and Zm00001d053834 that showed dramatic translational variations (Figure 5E and 5F, Supplemental Figure 13A and 13B). These results indicate that 3' UTRs play important roles in regulating gene translation in plants.

Notably, the haplotype frequency of the nonredundant translation-associated SNPs around the polyA signal of Zm00001d005814 varied dramatically among teosinte, landrace, and modern maize (Figure 5G) and may be associated with the changes in translation levels that were verified by luciferase assay (LUC) (Figure 5H and Supplemental Figure 13C). The frequency of haplotype Hap1 increased significantly during maize domestication and improvement, suggesting that translation-associated haplotypes may be selection targets. LHCA6, the Arabidopsis homolog of Zm0001d005814, encodes a subunit of the light-harvesting complex (LHC) of photosystem I that affects light-harvesting efficiency (Otani et al., 2018). During domestication from teosinte to modern maize, the maize growing area expanded from tropical to temperate regions, which exhibit dramatic variations in daylength. Hap1, the haplotype with the highest translational level, may have been selected to improve light harvesting in modern maize. This
finding suggests that selection can occur at the translatome level during domestication and improvement.

Translation is the final step in the flow of genetic information for protein biosynthesis. Monitoring translation shows a better correlation with proteomic profiling than monitoring transcription. In the current study, we determined that genic 3’ UTRs play a vital role in translatomic variation, consistent with previous findings (Zhao et al., 2019b). For the first time, we uncovered two translation-associated regions around genic polyA signals that

---

**Figure 5. Mutations around the polyA site domain affect the translational levels of genes.**

(A) Sequence elements for polyA include the polyA signal (PolyA_signal), polyA sites (PolyA_site), and downstream elements.

(B) Association analysis between SNPs 50 nt upstream or downstream of PolyA_signal and the translational levels of genes. The green triangle points to a new site significantly associated with translation.

(C and D) Two genes with SNPs 20 nt (C) and 19 nt (D) upstream of PolyA_signal show divergent translational expression patterns in different genotypes (nonparametric test, ***p < 0.001).

(E and F) Different haplotypes show divergent translational levels in Zm00001d012635 (E) and Zm00001d005814(F) (ANOVA, ***p < 0.001). The haplotype information is provided in Supplemental Figure 15A and 15B.

(G) HapMap divergence around PolyA_signal in teosinte, landrace, and modern maize. Haplotype1 in Zm00001d005814 became dominant during maize domestication and improvement.

(H) Similar translation variation of four haplotypes was detected using LUC (Student’s t-test, *p < 0.05).
are vital for translation, and we propose that the 20-nt region upstream of the polyA signal may be a new motif for polyA. It is worth noting that the regions 20 nt upstream and downstream of the polyA signal are crucial for promoting or reducing translatome levels. These sites represent possible targets for accurate crop improvement by fine-tuning the translatome and proteome levels of important genes. Therefore, the current study provides not only comprehensive translatome landscapes of plant genomes and a new cost-saving ribosome-profiling technique but also valuable targets for crop improvement.

In summary, we devised a new, cost-effective translatome profiling technique, 3’ Ribo-seq, and annotated the genomes of three model plants at the translatome level by combining 3’ Ribo-seq and conventional Ribo-seq. Translatome profiling in all three model plants uncovered many novel translational genic loci, permitted the systematic annotation of plant genomes, and complemented functional genome annotation in plants. Although conventional Ribo-seq successfully decoded translating codons and detected more loci, 3’ Ribo-seq is much less expensive and is suitable for population-level translational analysis. We used 3’ Ribo-seq to perform the first translational eGWAS on a diverse natural population and uncovered complex regulatory mechanisms that underlie translational variation and a novel, critical role for the region 20 nt upstream of the polyA signal in the translational variation of genes.

The 3’ Ribo-seq technique is an efficient tool for translatome profiling and has multiple advantages, including low cost and the ability to quantify expression levels. Nevertheless, two major shortcomings must be noted. First, its ability to identify alternative splicing is limited because it relies on 3’-end selection. Second, rRNA contamination can be a problem. Because 3’ Ribo-seq combines polysome profiling and 3’ RNA-seq, RNAs are readily degraded, and this can easily increase the proportion of rRNA sequences generated during the process. The use of probe hybridization or mRNA enrichment may be helpful for reducing rRNA contamination. Finally, some optimization steps are required for the successful use of 3’ Ribo-seq to profile the translatomes of many samples or populations: (1) appropriately increasing the number of mixed samples during library construction can further improve the efficiency of this technique and reduce the cost of obtaining a reasonable amount of data; and (2) distributing the same samples in different libraries is essential for removing potential batch effects.

**MATERIALS AND METHODS**

**Plant materials**

We devised a comprehensive technique for evaluating the robustness and reliability of 3’ Ribo-seq using two independent biological replicate samples. Seeds of the maize reference inbred line B73 were sown every 2 days in growth chambers under similar conditions to serve as two biological replicates. The aboveground tissues of 2-week-old seedlings (at least six plants per replicate) were collected and ground into a powder for RNA-seq, full-length ribosome profiling (FL_Ribo-seq), 3’ Ribo-seq, and protein MS.

We also collected a series of samples (most with two biological replicates) from different Arabidopsis, rice, and maize tissues at different developmental stages. Multiple A. thaliana (Col-0) tissues, including roots, stems, leaves, flowers, and fruits, were collected from mature plants, and whole seedlings were collected at 4 weeks after sowing and during the bolting stage. Stem, sheath, and leaf samples were collected from rice (Oryza sativa L.) plants at the V2 stage and booting stage. Root and seedling samples were collected from rice at the V2 stage, and spikelets were collected from rice at the booting stage. The following tissues were collected from maize reference inbred line B73: root, stem, leaf, and whole-seedling tissues from 14-day-old seedlings; kernel tissue at 9 days after pollination; and ear and tassel tissue at the V12 stage.

We also collected a large set of samples from a natural diverse population of maize to dissect population-level translational variation. One-hundred and fifty-nine genetically diverse maize inbred lines were sown in a field in Hainan in the winter of 2018. Leaf tissues were collected from at least three plants per inbred line at the V7 stage and equally mixed for 3’ Ribo-seq at the population level.

**Extraction of ribosome complexes**

Ribosome complexes were extracted from approximately 600 mg of tissue using PE buffer (20 mM HEPES-KOH, pH 7.5; 100 mM KCl; 10 mM MgCl2; 1 mM DTT; 100 μg/mL cycloheximide). The samples were loaded onto a sucrose density gradient (10%-50%) and subjected to ultracentrifugation at 38 000 rpm (Beckman, SW40 rotor) for 3 h at 4°C. A gradient profiler (BioComp, http://www.biocompinstruments.com/index.php) with an EM-1 UV data acquisition system (Bio-Rad, http://www.bio-rad.com/) was employed for component analysis according to the manufacturer’s instructions. All tubes contained monosone and polysome fractions, which were mixed for ribosome-embedded RNA isolation.

**RNA-seq, FL_Ribo-seq, and Ribo-seq**

RNA was extracted from frozen tissues and from fractions embedded with monosomes and polysomes using TRIzol reagent (Invitrogen) according to the manufacturer’s instructions. The sequencing libraries were prepared using a standard mRNA-seq library preparation kit according to the manufacturer’s protocol. The libraries were sequenced by paired-end 100-bp/150-bp sequencing on the BGISEQ-500 platform.

Ribosome complexes were extracted using extraction buffer (44 mM Tris-HCl, pH 7.5, 175 mM KCl, 13 mM MgCl2, 100 μg/mL cycloheximide, 15 mM 2-mercaptoethanol, 1% Triton X-100, 10 units/mL DNase I) and treated with RNase I (10 units/μg RNA) for 1 h at room temperature. The reaction was terminated using RNase inhibitor (20 units/μL), and the sample was immediately transferred into a MicroSpin S-400 column to enrich the RNA-ribosome complex (monosomes). The RPFs were extracted from the samples using an miRNeasy RNA isolation kit following the manufacturer’s instructions (Qiagen). After removing the rRNA and purifying the remaining RNA, the microRNA libraries were sequenced on the Illumina HiSeq X Ten platform according to the manufacturer’s instructions.

**Protein preparation and label-free MS**

Fresh tissues from maize seedlings (2 weeks old) and the leaves of Arabidopsis (mature), rice (V2 and booting stages), and maize (14 days old) were ground into a powder in liquid nitrogen. The samples were fully dissolved in four volumes of phenol extraction buffer (10 mM dithiothreitol, 1% protease inhibitor, 3 mM TSA, 50 mM NAM) and lysed by ultrasonication. An equal volume of Tris-phenol was added to each mixture, and the supernatant was collected by centrifugation at 38 000 rpm (Beckman, SW40 rotor) for 3 h at 4°C. The supernatant was mixed for 10 min at 4°C, and the supernatant was immediately transferred into a MicroSpin S-400 column to enrich the RNA-ribosome complex (monosomes). The RPFs were extracted from the samples using an miRNeasy RNA isolation kit following the manufacturer’s instructions (Qiagen). After removing the rRNA and purifying the remaining RNA, the microRNA libraries were sequenced on the Illumina HiSeq X Ten platform according to the manufacturer’s instructions.

**3’ Ribo-seq library construction and sequencing**

The 3’ selective library was constructed as follows: (1) RNA isolated from a mixture of monosomes and polysomes was quantified with a NanoDrop
Large-scale translome profiling in plants

2000 spectrophotometer. Approximately 500 ng of RNA from each sample was subjected to mRNAs enrichment, and half of the resulting sample was subjected to reverse transcription. (2) Each 2.75-μl enriched mRNAs sample and 0.5 μl of 10 mM dNTPs were placed into the well of a 96-well plate, followed by the addition of 0.5 μl of 50 μM reverse transcription (RT) primer (Ad2-1 to Ad2-12, Supplemental Table 4) containing anchored polyT, UMIs, barcodes, and one unique adapter. The plate was incubated at 65°C for 5 min and placed back on ice for more than 1 min. For first-strand cDNA synthesis, a mixture of RNase inhibitor (Invitrogen), SuperScript IV first-strand buffer, DTT, betaine, MgCl2, Ad1 primer (linking with three guanine ribonucleotides at the 3’ end, Supplemental Table 4), and SuperScript IV reverse transcriptase (Invitrogen) at final concentrations of 1 U/μl, 1 x, 5 mM, 1 M, 6 mM, 5 μl, and 10 U/μl, respectively, was added to each well. The RT reaction was performed by incubating the samples for 15 min at 50°C, followed by 10 cycles at 55°C for 2 min, 50°C for 2 min, and a final denaturation step at 80°C for 10 min. (3) For second-strand cDNA synthesis, 12 samples per row were pooled together and treated with RNase H. Using one quarter of the mixture as the template, PCR amplification was performed using one unique adapter (Ad1k, Supplemental Table 4) overlapping with Ad1 and another adapter (Ad2-bio, Supplemental Table 4) with biotin at its 5’ end overlapping with Ad2-n (Supplemental Table 4). The resulting cDNA was cleaned with KAPA Pure Beads (KAPA Biosystems) and quantified using a Qubit fluorometer. (4) Fifty nanograms of cDNA were subjected to tagmentation using a Regene DNA Sample prep kit (Illumina-4)-Bio with Tn5 transposase containing Illumina sequencing adapter. The 3’ ends of each cDNA fragment was captured using Dynabeads MyOne Streptavidin C1 beads (Life Technologies). (5) The 3’ ends of the cDNA fragment was enriched by PCR amplification using the following cycling conditions: 98°C for 3 min; 10-12 cycles of 98°C for 10 sec, 64°C for 30 sec, 72°C for 1 min; final extension at 72°C for 3 min. We added two indexes at both ends during this step to guarantee the specificity of each library. Indexed primer 1 (gAd3-P5, Supplemental Table 4) consisted of P5, index1, and Illumina sequencing adapter, and indexed nested primers (Ad2-ad and ad2-P7, Supplemental Table 4) consisted of P7, index2, Illumina sequencing adapter, and one unique adapter. (6) Size selection of the ampiclon was performed using KAPA Pure Beads (KAPA Biosystems), and the libraries were sequenced on the Illumina XTen platform.

Bioinformatic analysis of 3’Ribo-seq, RNA-seq, FL_Ribo-seq, and Ribo-seq data

Different libraries were separated based on the two indexes. The 36-bp fragments containing unique adapter, barcode, and UMI sequences at the 5’ end of R2.q were excised and linked to the 3’ end of the corresponding R1.q. The reads in each sample were extracted based on barcodes. The rRNA and tRNA reference sequences were downloaded from NCBI by searching with the keywords “ribosomal DNA” and “transfer RNA.” The 3’Ribo-seq reads were aligned against these sequences with Bowtie 2 software to remove the rRNA and tRNA reads (Langmead and Salzberg, 2012). Using RSEM-1.3.1 software (Li and Dewey, 2011) combined with the Bowtie 2 aligner, the remaining reads from maize, Arabidopsis, and rice were aligned to the maize reference genome (Zea_mays.AGPv4), the Arabidopsis reference genome (TAIR10), and the rice reference genome (Oryza_sativa.IRGPSP-1.0), respectively. Based on the UMIs, the duplicated reads generated by PCR amplification were removed. FPKM was used to measure gene expression levels. Genes with FPKM values ≥0.5 were considered to be expressed.

The reads from maize, Arabidopsis, and rice produced by RNA-seq or FL_Ribo-seq were also aligned to the maize reference genome (Zea_mays.AGPv4), the Arabidopsis reference genome (TAIR10), and the rice reference genome (Oryza_sativa.IRGPSP-1.0), respectively, using RSEM-1.3.1 (Li and Dewey, 2011) combined with the Bowtie 2 aligner (Langmead and Salzberg, 2012). FPKM was used to measure gene expression levels. Genes with FPKM values ≥0.5 were considered to be expressed.

Translomic annotation of the genome

Based on translomic data from multiple tissues, the Arabidopsis, rice, and maize genomes were re-annotated using the following steps: first, the single-end reads were mapped to the reference genome using HISAT2 (v2.0.4) software (Kim et al., 2015) with the parameter “dta.” Second, the alignments were assembled into potential transcripts using StringTie (v1.3.0) software (Pertea et al., 2015) with the parameter “G,” which produced known and new transcripts. Third, the results were compared with the reference annotation using Cuffcompare software (Trapnell et al., 2012), and new isoforms and translated loci were identified. The extracted sequences of the new genes were subjected to BLAST analysis against the UniProt protein database, and GO annotation was performed using TBoots (Chen et al., 2020) and WEGO (v 2.0, http://wego.genomics.org.cn/) software (Ye et al., 2018).

Genome-wide association study of translomic variation in maize

After removing low-quality samples and samples with high batch effects, 159 inbred samples were retained for further analysis. A genome-wide association study (GWAS) was performed to reveal the genetic architecture of translomic data in this diverse natural population of maize (Yang et al., 2010). The translation level of each gene was initially normalized using rank-based inverse normal transformation as the molecular phenotype. Genetic variants from the resequencing data of the association population were used as markers (Li et al., 2012). After filtering based on minor allele frequency (MAF) >0.05, 12 452 367 SNPs were retained for further analysis. GWAS was performed using EMMAX (Kang et al., 2010), controlling for population structure and family relatedness. Population structure and family relatedness were calculated using EMMAX and ADMIXTURE (Alexander et al., 2009), respectively. Significant associations were determined based on Bonferroni-corrected p < 8.03 x 10^-8 (1/N, N = 12 452 367). Haploview (Barrett et al., 2005) was used to estimate the correlation between SNPs to filter out cases with high linkage disequilibrium (LD) between markers. Only independent SNPs (r^2 < 0.1) were retained for further analysis. If the distance between a gene and the corresponding significant SNP was less than 10^5 bp, the SNP was defined as a local SNP. If the distance was greater than 10^5 bp, the SNP was defined as a distant SNP.

Identification of the polyA signal motif

Sequences downstream of the stop codons of genes were extracted, and the polyA signal motif was identified with Dragon PolyA Spotter (DPS, v 1.200, https://www.cbrc.kuast.edu.sa/dps/Capture.html) (Kalkatawi et al., 2013) to identify 6-bp sequence variants with the sequences AATAAA, AAAAAA, AAGAAA, AATACA, AATAA, ACTAAA, AGTAAA, ATTAAA, CATAAA, GATAAA, or ATAAA.

ACCESSION NUMBERS

The datasets analyzed during the current study (including biological and technical replicates, transcriptome and translome profiles of three model plant species, and population-level translome profiles of maize) are available at the Sequence Read Archive (SRA) repository
SUPPLEMENTAL INFORMATION
Supplemental information is available at Plant Communications Online.

AUTHOR CONTRIBUTIONS
L.L. designed the research. W.Z. and J.X. performed the experiments. W.Z., J.X., S.C., J.C., and Y.L. also performed experiments or analyzed data. W.Z. and L.L. analyzed the data and wrote the manuscript. J.C., L.L., and Q.L. provided constructive suggestions and revised the manuscript.

ACKNOWLEDGMENTS
We thank Prof. Zhipeng Zhou (Huazhong Agricultural University) for help with ribosome complex extraction. We thank Dr. Deningxiang Du and Dr. Leiming Wu (Huazhong Agricultural University) for providing a subset of the samples. We are grateful for help from the high-throughput sequencing platform of Huazhong Agricultural University and Dr. Qinhua Zhang. We are also grateful for help from Novogene in sequencing and from PTM Biolabs in protein MS. There are no potential conflicts of interest.

Received: November 13, 2020
Revised: March 15, 2021
Accepted: March 22, 2021
Published: March 23, 2021

REFERENCES
Alexander, D.H., Novembre, J., and Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. Genome Res. 19:1655–1664.
Arguel, M., LeBrigand, K., Paquet, A., Ruiz Garcia, S., Zaragosi, L., Barbry, P., and Waldmann, R. (2017). A cost effective 5’ selective single cell transcriptome profiling approach with improved UMI design. Nucleic Acids Res. 45:e48.
Barrett, J.C., Fry, B., Maller, J., and Daly, M.J. (2005). Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics 21:263–265.
Bava, F.A., Elistrovich, C., Ferreira, P.G., Minana, B., Ben-Dov, C., Guigo, R., Valcarcel, J., and Mendez, R. (2013). CPEB1 coordinates alternative 3’-UTR formation with translational regulation. Nature 495:121–125, B2.
Berdarini, T.Z., Reiser, L., Li, D., Mezheritsky, Y., Muller, R., Strait, E., and Huala, E. (2016). The Arabidopsis Information Resource: making and mining the “gold standard” annotated reference plant genome. Genesis 53:474–485.
Blein-Nicolas, M., Negro, S.S., Balliau, T., Welcker, C., Cabrera-Bosquet, L., Nicolas, S.D., Charcosset, A., and Zivy, M. (2020). A systems genetics approach reveals environment-dependent associations between SNPs, protein coexpression, and drought-related traits in maize. Genome Res. 30:1593–1604.

For full references, please see the original publication.
Large-scale translatome profiling in plants

Jan, C.H., Friedman, R.C., Ruby, J.G., and Bartel, D.P. (2011). Formation, regulation and evolution of Caenorhabditis elegans 3'UTRs. Nature 469:97–101.

Jiang, L., Li, B., Liu, S., Wang, H., Li, C., Song, S., Beatty, M., Zastrow-Hayes, G., Yang, X., Qin, F., et al. (2019). Characterization of proteome variation during modern maize breeding. Mol. Cell Proteomics 18:253–276.

Jiao, Y., Peluso, P., Shi, J., Liang, T., Stitzer, M.C., Wang, B., Campbell, M.S., Stein, J.C., Wei, X., Chin, C., et al. (2017). Improved maize reference genome with single-molecule technologies. Nature 546:524–527.

Kalkatawi, M., Rangkuti, F., Schramm, M., Jankovic, B.R., Kamau, A., Chowdhary, R., Archer, J.A., and Bajic, V.B. (2013). Dragon PolyA Spotted: predictor of poly(A) motifs within human genomic DNA sequences. Bioinformatics 29:1484.

Kang, H.M., Sul, J.H., Service, S.K., Zaitlen, N.A., Kong, S., Freimer, N.B., Slatkin, M., and Eskin, E. (2010). Variance component model to account for sample structure in genome-wide association studies. Nat. Genet. 42:348–354.

Kawahara, Y., de la Bastide, M., Hamilton, J.P., Kanamori, H., Kwon, T., Craig, S.A., Boutz, D.R., Huse, H.K., Laurent, J.M., Vogel, C., Kalkatawi, M., Rangkuti, F., Schramm, M., Jankovic, B.R., Kamau, A., et al. (2015). Droplet of molecules using unique molecular identifiers. Nat. Methods 12:348–354.

Kim, D., Langmead, B., and Salzberg, S.L. (2015). HISAT: a fast splice aligner with low memory requirements. Nat. Methods 12:357–360.

Kivioja, T., V H Rautio, A., Karlsson, K., Bonke, M., Enge, M., Kim, D., Langmead, B., and Salzberg, S.L. (2015). HISAT: a fast gapped-read alignment tool with excellent sensitivity. Nat. Biotechnol. 33:20–25.

Klein, A.M., Mazutis, L., Akartuna, I., Tallapragada, N., Veres, A., Li, C., Petsch, K., Sliz, P., and Gregory, R.I. (2020). RiboToolkit: an environment to reconstruct and analyze the host transcriptome in single cells. Cell 161:1187–1201.

Kleibenstein, D. (2009). Quantitative genomics: analyzing intraspecific diversity using comparative genomics. Annu. Rev. Plant Biol. 60:93–114.

Laurent, J.M., Vogel, C., Kwon, T., Craig, S.A., Boutz, D.R., Huse, H.K., Nozue, K., Walia, H., Whiteley, M., Ronald, P.C., et al. (2010). Protein abundances are more conserved than mRNA abundances across diverse taxa. Proteomics 10:4209–4212.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat. Methods 9:357–359.

Li, B., and Dewey, C.N. (2011). RSEM: accurate transcript quantification from RNA-seq data with or without a reference genome. BMC Bioinformatics 12:323.

Li, H., Peng, Z., Yang, X., Wang, W., Fu, J., Wang, J., Han, Y., Chai, Y., Guo, T., Yang, N., et al. (2012). Genome-wide association study dissectes the genetic architecture of oil biosynthesis in maize kernels. Nat. Genet. 44:43–50.

Li, L., Eichten, S.R., Shimizu, R., Petsch, K., Yeh, C., Wu, W., Chettoor, A.M., Givan, S.A., Cole, R.A., Fowler, J.E., et al. (2014). Genome-wide discovery and characterization of maize long non-coding RNAs. Genome Biol. 15:1–15.

Li, L., Petsch, K., Shimizu, R., et al. (2013). Mendelian and non-mendelian regulation of gene expression in maize. PLoS Genet. 9:e1003202.

Lindstein, T., June, C.H., Ledbetter, J.A., Stella, G., and Thompson, C.B. (1989). Regulation of lymphokine messenger RNA stability by a surface-mediated T cell activation pathway. Science 244:339–343.

Liu, Q., Shvarts, T., Sliz, P., and Gregory, R.J. (2020). RiboToolkit: an integrated platform for analysis and annotation of ribosome profiling data to decode mRNA translation at codon resolution. Nucleic Acids Res. 48:W218–W229.

Macosko, E.Z., Basu, A., Satij, R., Nemesh, J., Shekhar, K., Goldman, M., Tirosh, I., Bialas, A.R., Kamitaki, N., Martersteck, E.M., et al. (2015). Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. Cell 161:1202–1214.

Manley, J.L., and Takagaki, Y. (1996). Molecular biology - the end of the message - another link between yeast and mammals. Science 247:1481–1482.

Matoukova, E., Michalova, E., Vojtesek, B., and Hrstka, R. (2012). The role of the 3’ untranslated region in post-transcriptional regulation of protein expression in mammalian cells. RNA Biol. 9:563–576.

Mergner, J., Frejno, M., List, M., Papacek, M., Chen, X., Chaudhary, A., Samaras, P., Richter, S., Shikata, H., Messerer, M., et al. (2020). Mass-spectrometry-based draft of the Arabidopsis proteome. Nature 579:1–6.

Otani, T., Kato, Y., and Shikinai, T. (2018). Specific substituitions of light-harvesting complex I proteins associated with photosystem I are required for supercomplex formation with chloroplast NADH dehydrogenase-like complex. Plant J. 94:122–130.

Pertea, M., Pertea, G.M., Antonescu, C.M., Chang, T., Mendell, J.T., and Salzberg, S.L. (2015). StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. Nat. Biotechnol. 33:290–295.

Proudfoot, N.J. (2011). Ending the message: poly(A) signals then and now. Gene Dev. 25:1770–1782.

Rousseau, D., Kaspar, R., Rosenwald, I., Gehrke, L., and Sonenberg, N. (1996). Translation initiation of mitochondrial decarboxylase and nucleoeytoplasmic transport of cyclin D1 mRNA are increased in cells overexpressing eukaryotic initiation factor 4E. Proc. Natl. Acad. Sci. U S A. 93:1065–1070.

Schwanhausser, B., Busse, D., Li, N., Dittmar, G., Schuchhardt, J., Wolf, J., Chen, W., and Selbach, M. (2011). Global quantification of mammalian gene expression control. Nature 473:337–342.

Shepard, P.J., Choi, E., Lu, J., Flanagan, L.A., Hertel, K.J., and Shi, Y. (2011). Complex and dynamic landscape of RNA polyadenylation revealed by PAS-Seq. RNA 17:761–772.

Smith, T., Andreas, H., and Sudbery, I. (2017). UMI-tools: modeling sequencing errors in Unique Molecular Identifiers to improve quantification accuracy. Genome Res. 27:491–499.

Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D.R., Pimentel, H., Salzberg, S.L., Rinn, J.L., and Pachter, L. (2012). Differental gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nat. Protocol. 7:562–578.

Twyford, A.D. (2018). The road to 10,000 plant genomes. Nat. Plants 4:312–313.

Vogel, C., and Marcotte, E.M. (2012). Insights into the regulation of protein abundance from proteomic and transcriptomic analyses. Nat. Rev. Genet. 13:227–232.

Walley, J.W., Sartor, R.C., Shen, Z., Schmitz, R.J., Wu, K.J., Urich, M.A., Nery, J.R., Smith, L.G., Schnable, J.C., Ecker, J.R., et al. (2016). Integration of omic networks in a developmental atlas of maize. Science 353:814–818.

Wang, B., Tseng, E., Regulski, M., et al. (2016). Unveiling the complexity of the maize transcriptome by single-molecule long-read sequencing. Nat. Commun. 7:11708.
Plant Communications

Wilkening, S., Pelechano, V., Jaervelin, A.I., Tekkedil, M.M., Anders, S., Benes, V., and Steinmetz, L.M. (2013). An efficient method for genome-wide polyadenylation site mapping and RNA quantification (vol 41, pg e65, 2013). Nucleic Acids Res. 41:6370.

Yang, X., Yan, J., Shah, T., Warburton, M.L., Li, Q., Li, L., Gao, Y., Chai, Y., Fu, Z., Zhou, Y., et al. (2010). Genetic analysis and characterization of a new maize association mapping panel for quantitative trait loci dissection. Theor. Appl. Genet. 121:417–431.

Ye, J., Zhang, Y., Cui, H., Liu, J., Wu, Y., Cheng, Y., Xu, H., Huang, X., Li, S., Zhou, A., et al. (2018). Wego 2.0: a web tool for analyzing and plotting GO annotations, 2018 update. Nucleic Acids Res. 46:W71–W75.

Zhao, J., Qin, B., Nikolay, R., Spahn, C.M.T., and Zhang, G. (2019a). Translatomics: the global view of translation. Int. J. Mol. Sci. 20:212.

Zhao, T., Huan, Q., Sun, J., Liu, C., Hou, X., Yu, X., Silverman, I.M., Zhang, Y., Gregory, B.D., Liu, C., et al. (2019b). Impact of poly(A)-tail G-content on Arabidopsis PAB binding and their role in enhancing translational efficiency. Genome Biol. 20:189.

Zhu, Y.Y., Machleder, E.M., Chenchik, A., Li, R., and Siebert, P.D. (2001). Reverse transcriptase template switching: a SMART (TM) approach for full-length cDNA library construction. Biotechniques 30:892–897.