Ribosome profiling elucidates the contribution of differential translation to differential gene expression in bundle sheath and mesophyll cells in maize

Prakitchai Chotewutmontri* and Alice Barkan

Institute of Molecular Biology, University of Oregon, Eugene, OR 97403 USA

*To whom correspondence should be addressed. pchotewu@uoregon.edu

Short title: Bundle sheath and mesophyll cell translatomes

Abstract

The efficiencies offered by C_4 photosynthesis have motivated efforts to understand its biochemical, genetic and developmental basis. Reactions underlying C_4 traits in most C_4 plants are partitioned between two cell types, bundle sheath (BS) and mesophyll (M) cells. RNA-seq has been used to catalog differential gene expression in BS and M cells in maize and several other C_4 species. However, the contribution of translational control to maintaining the distinct proteomes of BS and M cells has not been addressed. In this study, we used ribosome profiling (ribo-seq) and RNA-seq to describe translatomes, translational efficiencies, and microRNA abundance in BS and M-enriched fractions of maize seedling leaves. A conservative interpretation of our data revealed 39 genes exhibiting cell-type dependent differences in translational efficiency, fourteen of which encode proteins with core roles in C_4 photosynthesis. Use of more relaxed criteria that are widely used in ribo-seq studies increased the set of differentially translated genes to several hundred. In addition, our data expand the set of genes known to be differentially expressed in BS and M cells, including genes encoding transcription factors and microRNAs. These data add to the resources for understanding the evolutionary and developmental basis of C_4 photosynthesis and for its engineering into C_3 crops.

Introduction

Plant species are classified as C_3 or C_4 according to their mechanism of photosynthetic carbon fixation. C_4 photosynthesis offers advantages under hot, dry conditions due to its higher efficiency of carbon fixation and lower water and nitrogen consumption (Hatch, 1987), and it evolved many times from a C_3 progenitor (Sage et al., 2011; Schlüter and Weber, 2020). Many C_4 plants are characterized by a specialized leaf anatomy, denoted Kranz anatomy, that partitions enzymes of the C_4 pathway between two cell types: mesophyll (M) and bundle sheath (BS). M cells surround BS cells, which in turn surround vascular bundles. Atmospheric CO_2 is first fixed by phosphoenolpyruvate (PEP) carboxylase in M cells to produce four-carbon acids. These are transported to BS cells, where they are decarboxylated, providing CO_2 for fixation by ribulose bisphosphate carboxylase/oxygenase (Rubisco) according to the C_3 scheme. This organization reduces Rubisco's wasteful oxygenation reaction by increasing local CO_2 concentration and decreasing local O_2 concentration.

Research into the genetic and developmental mechanisms underlying C_4 traits has been propelled by interest in introducing C_4 traits into C_3 crops (Sedelnikova et al., 2018; Ermakova et al., 2020). A thorough understanding of patterns of gene expression during C_4 differentiation is essential to achieve this goal. Numerous studies have approached this problem through transcriptome analysis (Wang et al., 2016; Schlüter and Weber, 2020), including studies that profiled transcriptomes during development of C_4 leaves (Li et al., 2010; Pick et al., 2011; Liu et al., 2013; Wang et al., 2013b; Wang et al., 2014; Ding et al., 2015; Mattiello et al., 2015; Denton et al., 2017), or compared transcriptomes between BS and M cells (Li et al., 2010; Chang et al., 2012; Aubry et al., 2014; John et al., 2014; Tausta et al., 2014; Denton et al., 2017), among different C_4 lineages (Aubry et al., 2014; Chen et al., 2014; John et al., 2014; Ding et al., 2015; Li et al., 2015; Offermann et al., 2015; Covshoff et al., 2016; Rao et al., 2016) or among related C_3 and C_4 species (Bräutigam et al., 2011; Gowik et al., 2011; Chen et al., 2014; Külahoglu et al., 2014; Wang et al., 2014; Ding et al., 2015; Lauterbach et al., 2017; Schlüter et al., 2017; Dunning et al., 2019). These transcriptome data have been complemented by surveys of protein populations during C_4 differentiation and in isolated BS and M chloroplasts (Majeran et al., 2005; Bräutigam et al., 2008; Majeran et al., 2008; Friso et al., 2010; Majeran et al., 2010).

The results from these studies support the view that transcriptional control plays the major role in determining patterns of gene expression during C_4 differentiation and in mature BS and M cells. That said, evidence for post-transcriptional contributions emerged in several studies (John et al., 2014; Ponnala et al., 2014; Williams et al., 2016), yet translational regulation has barely been explored (Schlüter and Weber, 2020). Ribosome profiling (ribo-seq) provides the means to comprehensively address this issue. Ribo-seq uses deep sequencing to map and quantify ribosome-protected mRNA fragments (ribosome footprints). Because average translation elongation rates are generally similar among mRNAs under a given condition, the normalized abundance of ribosome footprints is a widely accepted proxy for relative rates of protein synthesis (Brar and Weissman, 2015). Comparison of ribosome footprint

abundance with the abundance of the corresponding mRNA (i.e. the ratio of ribo-seq to RNA-seq reads) allows inferences about translational efficiencies (TE) on a genome-wide scale.

Previously, we used ribo-seq in conjunction with RNA-seq to analyze chloroplast gene expression in BS- and M-enriched leaf fractions in maize, a C_4 species of the NADP-malic enzyme (NADP-ME) subtype (Chotewutmontri and Barkan, 2016). We found that differences in mRNA abundance largely account for differential expression of chloroplast genes in the two cell types, but differences in translational efficiency synergize with mRNA-level effects in some cases. We have now extended this analysis to cytosolic mRNAs. Similar to what we observed in chloroplasts, the differential expression of nuclear genes in BS and M cells results primarily from differences in mRNA abundance. However, our data identify a subset of mRNAs whose translational efficiencies are significantly different in the two cell types. Additionally, our results expand the set of genes known to be differentially expressed in maize BS and M cells, suggesting new genes of potential relevance to the establishment or maintenance of C_4 traits.

Results and Discussion

Overview of ribo-seq and RNA-seq data collected from BS- and M-enriched leaf fractions

Analyses reported here used BS- and M-enriched fractions generated from the apical region of seedling leaves with a rapid mechanical fractionation method similar to that used in our prior study of chloroplast translatomes (Chotewutmontri and Barkan, 2016). Marker proteins for each cell type were highly enriched in these fractions (Figure 1A). Ribosome footprints and RNA were purified from aliquots of the same BS and M preparations, in three biological replicates. The ribo-seq reads mapping to cytosolic mRNAs exhibit the expected characteristics of ribosome footprints: they map almost exclusively to protein coding regions, their size distribution is heavily weighted to 29-30 nucleotides, and their positions within open reading frames (ORFs) show strong 3-nucleotide periodicity (Supplemental Figure 1A-C). We used rRNA-depleted total RNA for RNA-seq to avoid the 3'-bias that can lead to false inferences about differential expression (Denton et al., 2017) (Supplemental Figure 2A-B). Correlation coefficients among replicates ranged from approximately 0.93 to 0.98, with replicate datasets clustering together as expected (Supplemental Figure 1D).

Four RNA-seq studies of maize BS and M fractions were reported previously (Li et al., 2010; Chang et al., 2012; Tausta et al., 2014; Denton et al., 2017), all of which used polyA-selected mRNA. We used the two most recent studies as points of comparison for our data. Tausta et al (2014) generated BS and M-enriched fractions by laser capture microdissection whereas Denton et al (2017) used a mechanical method that differed from ours. Both studies analyzed tissue slices at several positions along the leaf blade, which represent different stages along the pathway of photosynthetic differentiation. The apical sections were most analogous to our material, so we selected those data for our comparisons. To facilitate comparisons, we aligned both prior datasets to the current maize genome assembly (B73 RefGen_v4) and analyzed differential expression using the same analysis pipeline we used for our data.

Figure 1B compares the cell type-enrichment of sequence reads from four commonly used C_4 marker genes in our data to those in the prior datasets. In comparison with the Denton data, our RNA-seq data exhibited greater enrichment of both M and BS markers. In comparison with the Tausta data, our RNA-seq data exhibited similar enrichment of M markers and less enrichment of BS markers. We expected that our mechanical fractionation method would result in contamination of the M and BS fractions with epidermal and vascular tissue, respectively, and analysis of epidermal and phloem markers confirmed this to be the case (Figure 1C and Supplemental Figure 2C). Despite the considerable differences in sample preparation, a core gene set associated with C_4 photosynthesis (Schlüter and Weber, 2020) showed Pearson correlation coefficients of at least 0.85 among all three datasets (Supplemental Figure 3A).

Comparison of our ribo-seq and RNA-seq data (Supplemental Figure 3B) returned correlation coefficients of 0.72 and 0.89 for genome-wide and C_4 gene comparisons, respectively. These values are considerably lower than those for replicate samples, suggesting some differences in translational efficiency in the two cell types. In addition, we compared our data to that from a proteomic study of material obtained with a similar mechanical fractionation method (Friso et al., 2010) (Supplemental Figure 3C). This comparison was complicated by the many changes in maize gene models since the time of the proteome study. Nevertheless, proteins that were quantified with higher confidence levels showed considerable correlation with our ribo-seq data. Furthermore, our ribo-seq data correlated better with the proteomic data than did our RNA-seq data, consistent with a role for translational control in maintaining the distinct proteomes in the two cell types.

Differences in translational efficiency contribute to the differential expression of genes in BS and M cells.

To detect mRNAs that experience differential translation in M and BS cells, we used two different algorithms to calculate differences in translational efficiency (TE): XTAIL (Xiao et al., 2016) and anota2seq (Oertlin et al., 2019) (Supplemental Data Set 1). Both algorithms report normalized ratios of ribo-seq to RNA-seq reads, but they employ different statistical methods. To reduce false positives resulting from sampling variation, we filtered the outputs to remove genes expressed at low levels. Genes were retained if they met both of the following criteria: an average of at least 100 RNA-seq reads mapped to coding sequences in both the BS and M fractions, and an average of at least 100 ribo-seq reads in either the BS or M fraction. We defined a gene as being differentially translated if it passed these read count cutoffs and showed > 2-fold difference in TE between the two tissues ($|\log_2 (BS/M)| > 1$), with a false discovery rate (FDR) < 0.001 for XTAIL or < 0.05 for anota2seq. We chose different FDRs for the two algorithms due to striking differences in the number of differentially translated genes returned for any particular FDR (Figure 2A). Of the 9,476 genes meeting our read count cutoffs, XTAIL reported 928 differentially translated genes ($|\log_2 (BS/M)| >1$ and FDR < 0.001) while anota2seq reported 71 differentially translated genes ($|\log_2 (BS/M)| >1$ and FDR < 0.05) (Figure 2B, Supplemental Data Set 2).

Only 39 genes met our cutoffs in both algorithms (Table 1, Figure 2B). Fourteen genes in the 39 gene set encode proteins that are known to partition between BS and M cells (asterisks in Table 1): ten "C_4" genes discussed below, together with genes encoding fructose-1,6-bisphosphatase (FBP), ribose-5-phosphate isomerase (RPI), glycine decarboxylase system T-protein (GDCT) photorespiration, and alanine aminotransferase (AlaAT) (Schlüter et al., 2019; Schlüter and Weber, 2020). Interestingly, the gene encoding RPI had been suggested to be under translational control based on a comparison of Setaria and maize transcriptome data (John et al., 2014).

We then focused on two sets of genes highlighted in prior work as being relevant to C_4 physiology: (i) a list of "C_4" genes (Schlüter and Weber, 2020) that includes genes involved in photorespiration, the Calvin-Benson cycle, C_4 metabolism, redox regulation, and nitrogen and sulfur assimilation; and (ii) membrane transporters, subunits of PSII, and subunits of the NADH dehydrogenase-like complex (NDH) that are strongly enriched in one cell type or the other (Majeran and Wijk, 2009). Of the 204 genes on these lists, ten were reported by both XTAIL and anota2seq as having a greater than 2-fold difference in translational efficiency in the BS and M fractions (Figure 3, Supplemental Figure 4, Supplemental Data Set 3). These include four genes encoding proteins that function in C_4 metabolism, the Calvin-Benson cycle or photorespiration (PPDK-RP.2, CP12.1, PRK.2, GLYK.1), three genes encoding transporters (DIT1, MEP3, PIP2.3), and three genes encoding subunits of the NDH complex (NdhN, PnsL3/PQL2, and NdhU) (see Supplemental Data Set 3 for gene identifiers). The $\log_2$ (BS/M) ratios (Figure 3) as well as normalized read coverage plots (Figure 4) suggest that differential translational efficiency functions synergistically with differential transcription/mRNA stability to produce strong cell-type dependent differences in the outputs of these genes. Proteomic data are available for eight of these genes (Figure 3). The BS/M ratio for CP12.1, PRK.2 and GLYK.1 in the proteomic and ribo-seq data are similar, whereas the ribo-seq data show greater cell-type specific enrichment for the other genes.

We then expanded the highly stringent set of 39 differentially translated genes to include genes that passed the 2-fold change criterion in both programs, but passed the FDR criterion in just one of them. This produced 194 differentially translated genes (Figure 5A, "clustering" set), which were then used for *k*-means clustering based on differential RNA abundance and differential TE (Figure 5B, Supplemental Data Set 4). The majority of these genes showed synergistic changes in RNA abundance and TE that amplify their differential expression (labeled RNA&TE in Figure 5B). For 38 genes, the differential expression reported by ribo-seq arises predominantly from differing translational efficiency (labeled TE in Figure 5B). Interestingly, some genes exhibited differences in TE and RNA abundance of opposite direction, such that the two effects offset one another (labeled as "buffering" in Figure 5B). Buffering of this type has been detected in many ribo-seq studies (McManus et al., 2014; Schafer et al., 2015; Oertlin et al., 2019).

An expansion of the heatmap in Figure 5B (right) focuses on the 39 genes reported by both XTAIL and anota2seq to be differentially translated using our stringent FDR cutoffs, comparing our data with

prior RNA seq and proteome datasets (Supplemental Data Set 5). Some of these genes are regulated primarily at the level of translational efficiency (TE) (for example, SUS shown in Figure 5C) and others by a combination of RNA abundance and TE (for example CHUP1 shown in Figure 5C). None of these 39 genes fall into the "buffering" set.

Global analysis of BS and M translatomes expands the known set of differentially expressed genes

The abundance of ribosome footprints mapping to a gene reflects both mRNA abundance and translational efficiency. This is widely used as a proxy for relative rates of protein synthesis among genes in the sampled material, and there is considerable evidence that this is, with rare exceptions, a valid assumption (Brar and Weissman, 2015). Therefore, ribo-seq can provide a more accurate picture of differential expression than RNA-seq. With that in mind, we analyzed our ribo-seq data with DESeq2 to catalog genes that are differentially expressed in BS and M cells. We limited the analysis to the 13,182 genes having an average of at least 100 ribo-seq reads mapped to coding sequences in at least one of the two tissue types (Supplemental Data Set 6). Differentially expressed genes were defined as having $|\log_2 (BS/M)| > 1$ and FDR < 0.001. With these criteria, we detected 2,338 differentially expressed genes: 1,345 and 993 genes are BS and M-enriched, respectively (Figure 6A, Supplemental Datasets 7 and 8). We then compared these results to those from previous RNA-seq studies of maize BS and M fractions (Li et al., 2010; Chang et al., 2012; Tausta et al., 2014; Denton et al., 2017). Roughly 10% of the genes we identified as differentially expressed could not be matched in a 1-to-1 fashion to the outmoded gene models used in earlier studies ("conflict gene model" set in Figure 6A). Among the remaining genes, approximately 30% of the genes reported as differentially expressed in our ribo-seq analysis were not reported to be differentially expressed in prior RNA-seq studies (Supplemental Data Sets 7 and 8; "unique" sets in Figure 6A). These "unique" gene sets showed similar functional enrichment patterns as the complete sets of BS- and M-enriched genes (Figure 6B, Supplemental Data Set 9), instilling confidence that many of these genes are, in fact, differentially expressed in BS and M cells. One interesting example is a BS-enriched member of the Aux/IAA family of auxin-regulated transcription factors (Gene ID Zm00001d041416), which also emerged as relevant to C_4 biology based on evolutionary inferences (Huang et al., 2017). Aux/IAA proteins govern diverse developmental processes including vascular patterning (Zhang et al., 2014; Luo et al., 2018). These observations suggest the possibility that this Aux/IAA protein regulates C_4 vascular patterning in the maize leaf.

Three photosynthetic complexes harboring plastid-encoded subunits accumulate differentially in BS and M cells: Rubisco and the NADH dehydrogenase-like complex (NDH) accumulate preferentially in BS cells whereas PSII accumulates preferentially in M cells. The biogenesis of each complex involves a large number of nuclear genes, some of which are involved in complex assembly and others in the expression of plastid-encoded subunits. Differential ribo-seq and RNA-seq expression data for factors that function specifically in the biogenesis of these complexes (Supplemental Figure 5 and Supplemental

Dataset 10) show that many of the biogenesis factors are preferentially expressed in the same cell type as the cognate complex. Furthermore, the M-enriched expression of three *psbA* translational activators (HCF173, HCF244, OHP2) (Schult et al., 2007; Link et al., 2012; Chotewutmontri et al., 2020) can account for the preferential translation of *psbA* mRNA in M cells we had reported previously (Chotewutmontri and Barkan, 2016).

Differential expression of transcription factors and microRNAs in BS and M fractions.
The distinct transcriptomes of BS and M cells are presumably maintained, at least in part, by cell-type specific regulatory molecules, including transcription factors and microRNAs. We identified genes encoding 113 and 37 transcription factors that were enriched in BS and M translatomes, respectively (Supplemental Data Set 11). Approximately one third of these had not been reported as differentially expressed in prior RNA-seq studies, and only one fifth were reported as differentially expressed across all studies (Figure 6C). The low concordance among studies is likely due to a variety of factors, including varying sequencing depth, varying criteria used to identify differentially expressed genes, and imperfectly matched developmental stages. In addition, cell-type dependent translation accounted for the differential expression of one of the transcription factors detected uniquely in our set (Trihelix TF28 in Figure 5B). Transcription factors whose mRNAs were reported to accumulate in a similar cell-type specific fashion across multiple species (Huang and Brutnell, 2016) are marked in the final column in Figure 6C. Expression of GLK1, a transcription factor required specifically for M-cell differentiation in maize (Wang et al., 2013a) was M-enriched in our data as in all prior RNA-seq studies.

MicroRNAs regulate mRNA stability and translation. To address whether cell-type specific microRNA expression accounts for the differential expression of any genes in BS and M cells, we sequenced small RNAs in the same RNA preparations used for RNA-seq. DESeq2 returned six differentially expressed microRNAs [$|\log_2 (BS/M)| > 1$ and FDR < 0.05] (Supplemental Data Set 12), representing two microRNA families: MIR160 and MIR166 (Supplemental Figure 6). MicroRNAs are processed from precursors that form a stem-loop structure, yielding both a microRNA and its complement (microRNA*) (Wang et al., 2019). Five members of the MIR160 gene family encode identical mature microRNAs, and these are strongly enriched in our BS fraction (Figure 7A-B). One member of the MIR166 family (zma-MIR166a) (Supplemental Figure 6) is enriched in our M fraction (Figure 7A-B).

We then identified mRNA targets of these microRNAs in the DPMIND database (Fei et al., 2018). Zma-MIR160, which is enriched in our BS fraction, targets mRNAs encoding three auxin responsive factor (ARF) transcription factors. RNA-seq data show that expression of one of these, ARFTF15, is enriched in the M fraction (Figure 7C), supporting a role for zma-MIR160 in directing ARFTF15 mRNA decay in BS cells. The M-enriched microRNA zma-MIR166 targets mRNAs encoding three class III homeodomain leucine zipper (HD-ZIP III) transcription factors (Juarez et al., 2004; Fei et al., 2018). RNA-seq data reported by Chang et al (2012) indicate BS enrichment for all three of these mRNA targets, and our RNA-seq data support this for one of them (RLD1) (Figure 7C). Zma-MIR166 is known to control spatial

expression of RLD1 in maize leaf primordia (Juarez et al., 2004). Our data suggest further that zma-MIR166 regulates RLD1 in differentiated leaf tissue. Interestingly, a gain-of-function *Rld1* mutant that is no longer controlled by zma-MIR166 exhibited normal vasculature in leaf primodia (Juarez et al., 2004) but failed to increase vascular bundle density and lacked photosynthetic BS and M cells in mature leaf (Nelson et al., 2002). These results together with our findings suggest that zma-MIR166-mediated inhibition of RLD1 expression in mature M cells is important to maintain high vein density and differentiated BS/M cells, which are important C_4 traits (Ermakova et al., 2020).

Upstream open reading frames and non-AUG initiated translation associated with BS and M-specific expression

Upstream open reading frames (uORFs) often regulate the translation of "main" ORFs (mORFs) in eukaryotic mRNAs (Hinnebusch et al., 2016). To examine the contribution of uORFs to the translational regulation we observed, we inspected the 39 genes identified as differentially translated with our most stringent criteria for ribosome footprints upstream of their annotated start codon. We considered uORFs beginning with AUG as well as with the non-canonical start codons CUG and ACG (van der Horst et al., 2019). We identified seven genes that fit these criteria (Figure 8A). Of these, two had uORFs beginning with AUG, five had uORFs beginning with CUG or ACG, two had overlapping uORFs, and three had uORFs overlapping with mORFs.

The ribo-seq data suggest that the ratios of uORF to mORF reads differ in the two cell types (Figure 8A-B). All five genes with non-AUG uORFs (MEP3, CP12.1, Ubiquitin-like, AlaAT and PC-PLA1) had a higher ratio of uORF/mORF ribosome occupancy in BS cells than in M cells. By contrast, the two genes with AUG-initiated uORFs, showed a similar (CHUP1) or lower (BX4) ratio of uORF/mORF ribosome occupancy in BS compared to M cells. These results suggest a bias toward translation of non-AUG start codons in BS cells; however, a large-scale analysis of start codon usage will be needed to draw firm conclusions. In yeast, the translation initiation factors eIF1 and eIF5A promote translation from AUG codons while eIF5 antagonizes this function, thereby favoring non-AUG initiation (Hinnebusch et al., 2016; Eisenberg et al., 2020). We therefore considered the possibility that differing ratios of these factors in the two cell types underlie the differences in AUG versus non-AUG uORF initiation suggested by our data. However, this was not the case (Figure 8C). That said, expression of one eIF1 paralog (Zm00001d021668) and two eIF5A paralogs (Zm00001d006760, Zm00001d022042) was BS-enriched in our ribo-seq data (Figure 8C), and we speculate that these may play a role in favoring non-AUG initiations.

Conclusion

Results presented here provide strong evidence that translational regulation contributes to the differential expression of genes in BS and M cells in maize. For most genes impacted by translational control, differences in translational efficiency synergize with differences in mRNA abundance to produce strong cell-type dependent differences in gene output. This is similar to what we observed in an analogous study

of chloroplast gene expression in BS and M fractions (Chotewutmontri and Barkan, 2016). Our data add many new candidates to the set of genes reported to be differentially expressed in the two cell types, including numerous transcription factors and several microRNAs. For example, we detected BS-enriched expression of an AUX/IAA transcription regulator and of Trihelix Transcription Factor 28 (THX28), the latter of which is differentially expressed primarily due to differential translation. We present evidence that the differential expression of two microRNAs underlies the cell-type enriched expression of several transcription factors in mature BS and M cells: zma-MIR166a and zma-MIR160 are preferentially expressed in M and BS fractions, respectively, correlating with expression of their target mRNAs encoding transcription factors RLD1 and HD-ZIPIII proteins (respectively) in the opposite cell type.

The molecular basis for the differential translation events detected in this study provide interesting areas for future investigation. Factors such as alternative polyadenylation sites should be considered. Indeed, we found evidence that alternative polyadenylation sites correlate with differential translation of one mRNA (Supplemental Figure 7), indicating that a deeper analysis is warranted. In addition, our data hint that AUG *versus* non-AUG uORF initiation is relevant to the cell-type dependent translation of several genes. We are particularly intrigued with our data suggesting that initiation from non-AUG codons is favored in BS cells. A method called Translation Initiation Site Profiling revealed a shift to enhanced translation from non-AUG start codons during yeast meiosis, resulting in the synthesis of many novel protein isoforms (Eisenberg et al., 2020). Analysis of the BS and M translatomes with this assay could address whether translational reprogramming of this type affects BS and M proteomes in biologically meaningful ways.

Materials and Methods

Preparation of BS and M fractions

The fractions were prepared using a minor modification of the rapid mechanical procedure described previously (Chotewutmontri and Barkan, 2016). In brief, *Zea mays* (inbred line B73) was grown for 13 days under cycles of 12 h light (300 µmol $m^{-2}s^{-1}$) at 31°C and 12 h dark at 22°C. The apical one-third of the second and third leaves to emerge were harvested 2 h into the light cycle. The M fraction was taken as the supernatant following gentle grinding of fresh tissue in a mortar and pestle as described before (Chotewutmontri and Barkan, 2016). However, differing from our previous procedure which used the residual material from the M preparation to generate the BS fraction, the BS fraction was prepared from separate tissue aliquots as the tissue remaining after several rounds of gentle grinding. This fractionation procedure took roughly 3 min for M fractions and 10 min for BS fractions, at which point samples were flash frozen in liquid nitrogen.

Ribo-seq, RNA-seq and sRNA-seq libraries

Ribosome footprints and total RNA were prepared from the same lysates in three biological replicates as described previously (Chotewutmontri and Barkan, 2016). Ribosome protected fragments from ~20 to ~40 nucleotides were selected for ribo-seq library preparation, and libraries were prepared with the

NEXTflex Small RNA Sequencing Kit v2 (Bioo Scientific) in combination with rRNA depletion with customized biotinylated antisense oligonucleotides (Chotewutmontri et al., 2018). RNA-seq libraries were prepared from total RNA after rRNA depletion with the Ribo-Zero rRNA Removal Kit (Plant Leaf) (Illumina) using the NEXTflex Rapid Directional qRNA-Seq Kit (Bioo Scientific). To catalog microRNAs, the same RNA samples were analyzed by sRNA-seq with the same kit used to sequence ribosome footprints. The ribo-seq and RNA-seq libraries were sequenced using a HiSeq 2500 (Illumina) in single-read mode with read lengths of 75 nucleotides. The microRNA libraries were sequenced using a HiSeq 4000 (Illumina) in single-read mode with read lengths of 100 nucleotides.

Sequence read processing

Ribo-seq and RNA-seq data were processed as described previously (Chotewutmontri and Barkan, 2020) with minor modifications. In brief, adapter sequences were trimmed using cutadapt (Martin, 2011). For ribo-seq, only trimmed reads between 18 and 40 nucleotides were retained for alignments. Reads were aligned sequentially to the maize chloroplast genome (GenBank accession X86563), the maize mitochondrial genome (B73 RefGen_v4 assembly release 38), and the maize nuclear genome (B73 RefGen_v4 assembly release 38) using STAR version 2.5.3a (Dobin et al., 2013). Read splicing was restricted to annotated splice junctions. Read counting was performed using featureCounts from Subread package version 1.6.0 (Liao et al., 2014), counting only those reads that mapped to a unique site. Ribo-seq read counts for nuclear and organellar genes excluded the first 25 and 10 nucleotides of each ORF, respectively, to avoid counting ribosome pileups at start codons. RNA-seq read counts used full length ORFs. Read counts are provided in Supplemental Data Set 13.

The ribosome footprint length distribution, three-nucleotide periodicity, and metagene analysis were performed using in-house scripts. Comparison among replicates was evaluated with Pearson correlation and hierarchical clustering using $\log_{10}$ RPKM values. Transcript coverage and 3' bias were analyzed using Picard Tool version 2.6.0 (http://broadinstitute.github.io/picard/). Samtools (Li et al., 2009) was used to extract gene-specific read coverages. P-site assignment of ribo-seq reads was inferred from the ribosome footprint size-dependent placements of footprint 5' ends to the P-site location, which were observed from metagene analysis of reads mapping to start codons in the data discussed here.

MicroRNA data were trimmed with cutadapt (Martin, 2011). Trimmed reads between 15 and 45 nucleotides were retained and aligned to the B73 RefGen_v4 assembly using STAR version 2.5.3a (Dobin et al., 2013). For multi-mapped reads, STAR parameters were set to distribute each multi-mapped read randomly to one of the mapped locations (--outMultimapperOrder Random --outSAMmultNmax 1). Read counting was performed using featureCounts (Liao et al., 2014) with maize microRNA annotation release 22 (mirbase.org), where counting included multi-mapped reads. The read counts are provided in Supplemental Data Set 12.

Differential expression analysis

Differential expression analysis based on RNA-seq, ribo-seq or microRNA data was performed using DESeq2 (Love et al., 2014). Outputs are reported in Supplemental Data Sets 6, 12 and 14. Genes were

reported as differentially expressed if they showed an average of at least 100 reads in at least one of the tissue types, a $|\log_2$ fold change$|$ > 1 and FDR < 0.001 for RNA-seq and ribo-seq or FDR < 0.05 for microRNA data.

For analysis of differential TE, we used two different software packages: XTAIL (Xiao et al., 2016) and anota2seq (Oertlin et al., 2019). XTAIL is based on DESeq2 and uses a log ratio approach, whereas anota2seq uses an approach based on the analysis of partial variance. Only genes with an average of at least 100 RNA-seq reads in both tissue types and an average of at least 100 ribo-seq reads in at least one of the tissue types were reported. Genes were reported as differentially translated if they showed $|\log_2$ (BS/M)$|$ > 1 with FDR < 0.001 for XTAIL and FDR < 0.05 for anota2seq. The outputs of XTAIL and anota2seq are provided in Supplemental Data Set 1. *k*-means clustering analysis was done in R with the DESeq2-derived $\log_2$ (BS/M) values of RNA-seq data and the XTAIL-derived $\log_2$ (BS/M) values of TE.

Functional enrichment analysis

The enrichment analysis was performed with MapMan (Thimm et al., 2004) as described (Zones et al., 2015), using maize X4.2 functional assignments and Mapman terms Levels 1 to 3. In brief, a set of genes showing an average ribo-seq RPKM value > 1 in at least one tissue type (18,980 genes) was used to calculate background distribution of each MapMan term among the expressed gene set. To evaluate a pair of BS and M data sets of sizes *m* and *n* genes, a total of 10,000 permutations by random sampling without replacement for pair sets of *m* and *n* genes from the expressed gene pool were performed. The term occurrences in these 10,000 permutated data set pairs produced the mean and standard deviation for each term occurrence in the background. The observed term occurrence in each data set together with the mean and standard deviation of the term occurrence in the background from the corresponding permutation set were used to calculate z-score and p-value. The adjusted p-values or FDR values were calculated using the Benjamini-Hochberg method in R as described previously (Zones et al., 2015).

Reanalysis of external data

Previously published RNA-seq analyses of maize BS and M fractions used non-current maize genome assemblies and gene models. For comparison to our data, we re-analyzed the +9 cm data (also called section 14) from Tausta et al. (2014) and the slice 1 data from Denton et al. (2017) using our pipeline. To compare our ribo-seq data to proteome data from Friso et al. (2010), those data were matched to B73 v4 gene identifiers using the gene ID history (gramene.org). For comparison to our differential translatomes in Figures 6, 7 and Supplemental Figure 5, the originally reported $\log_2$ fold change (BS/M) values were compiled from the BS and M data from Li et. al. (2010), Chang et al. (2012), the section 14 data from Tausta et al. (2014), and the slice 1 data from Denton et al. (2017). The values reported with older gene IDs were matched using the gene ID history (gramene.org).

Antibodies

Antibodies to NdhH and PPDK were generously provided by Tsuyoshi Endo (Kyoto University) and Kazuko Aoyagi (UC Berkeley), respectively. Antibodies to PEPC, malic enzyme, and RbcL were gifts of William Taylor (University of California, Berkeley). The antibody to D2 was obtained from Agrisera.

Accession Numbers

The ribo-seq and RNA-seq data were deposited at the NCBI Sequence Read Archive (SRA) with accession number PRJNA667075. RNA-seq data from Tausta et al. (2014) and Denton et al. (2017) were downloaded from SRA (accession numbers SRP035577 and SRP052802, respectively). Alignments of reads to the maize chloroplast genome used Genbank accession X86563. B73 RefGen_v4 assembly with annotation release 38 (gramene.org) was used for alignments to the nuclear and mitochondrial genomes.

Supplemental Data

Supplemental Figure 1. Characteristics of ribo-seq and RNA-seq data from BS and M fractions.
Supplemental Figure 2. Comparison of RNA-seq data characteristics with those from prior studies.
Supplemental Figure 3. Comparison of RNA-seq, ribo-seq, and proteome data collected for maize BS and M leaf fractions.
Supplemental Figure 4. Summary of expression data for C_4-related genes involved in redox regulation, nitrogen assimilation, sulfur assimilation and Photosystem II function.
Supplemental Figure 5. Differential expression of biogenesis factors for PSII, Rubisco, and the NDH complex.
Supplemental Figure 6. Alignment of microRNA sequences of zma-MIR160 and zma-MIR166 families.
Supplemental Figure 7. Alternative polyadenylation sites of PRK.2 (Zm00001d000164) transcripts.
Supplemental Data Set 1. XTAIL and anota2seq outputs and related data.
Supplemental Data Set 2. Genes reported by XTAIL and/or anota2seq as being differentially translated in BS and M fractions.
Supplemental Data Set 3. Differential expression of 204 genes with known relevance to C4 physiology.
Supplemental Data Set 4. Clustering analysis of translational regulation modes.
Supplemental Data Set 5. Data for the 39 genes reported by both XTAIL and anota2seq as differentially translated in BS and M cells.
Supplemental Data Set 6. DESeq2 output of ribo-seq data.
Supplemental Data Set 7. BS-enriched genes in ribo-seq data.
Supplemental Data Set 8. M-enriched genes in ribo-seq data.
Supplemental Data Set 9. Enrichment analysis of MapMan terms in the enriched genes of ribo-seq data.
Supplemental Data Set 10. Differential expression of the nucleus-encoded proteins involved in expression and assembly of Rubisco, NDH-like and PSII complexes.
Supplemental Data Set 11. Data for differentially expressed transcription factors in ribo-seq data.
Supplemental Data Set 12. MicroRNA data and DESeq2 analysis.
Supplemental Data Set 13. Ribo-seq and RNA-seq read counts, including re-analyzed data from Denton S1 and Tausta Sec 14 samples.
Supplemental Data Set 14. DESeq2 outputs of RNA-seq data, including re-analyzed data from Denton S1 and Tausta Sec 14 samples.

Acknowledgments

We wish to thank Rosalind Williams-Carrier for assistance with sample preparation and for comments on the manuscript, Susan Belcher for comments on the manuscript, and the University of Oregon Genomics and Cell Characterization Core Facility for Illumina sequencing. This work was supported by grant IOS-1339130 to A.B. from the US National Science Foundation.

Author Contributions

Conceived and designed the experiments: P.C., A.B. Performed the experiments: P.C. Analyzed the data: P.C., A.B. Contributed reagents/materials/analysis tools: P.C. Wrote the paper: P.C., A.B.

REFERENCES

Aubry, S., Kelly, S., Kumpers, B.M., Smith-Unna, R.D., and Hibberd, J.M. (2014). Deep evolutionary comparison of gene expression identifies parallel recruitment of trans-factors in two independent origins of C4 photosynthesis. PLoS Genet **10**: e1004365.

Brar, G.A., and Weissman, J.S. (2015). Ribosome profiling reveals the what, when, where and how of protein synthesis. Nature Rev Molec Cell Biol **16**: 651-664.

Bräutigam, A., Hoffmann-Benning, S., and Weber, A.P. (2008). Comparative proteomics of chloroplast envelopes from C3 and C4 plants reveals specific adaptations of the plastid envelope to C4 photosynthesis and candidate proteins required for maintaining C4 metabolite fluxes. Plant Physiol **148**: 568-579.

Bräutigam, A., Kajala, K., Wullenweber, J., Sommer, M., Gagneul, D., Weber, K.L., Carr, K.M., Gowik, U., Mass, J., Lercher, M.J.*, et al.* (2011). An mRNA blueprint for C4 photosynthesis derived from comparative transcriptomics of closely related C3 and C4 species. Plant Physiol **155**: 142-156.

Chang, Y.M., Liu, W.Y., Shih, A.C., Shen, M.N., Lu, C.H., Lu, M.Y., Yang, H.W., Wang, T.Y., Chen, S.C., Chen, S.M.*, et al.* (2012). Characterizing regulatory and functional differentiation between maize mesophyll and bundle sheath cells by transcriptomic analysis. Plant Physiol **160**: 165-177.

Chen, T., Zhu, X.G., and Lin, Y. (2014). Major alterations in transcript profiles between C3-C4 and C4 photosynthesis of an amphibious species Eleocharis baldwinii. Plant Mol Biol **86**: 93-110.

Chotewutmontri, P., and Barkan, A. (2016). Dynamics of Chloroplast Translation during Chloroplast Differentiation in Maize. PLoS genetics **12**: e1006106.

Chotewutmontri, P., and Barkan, A. (2020). Light-induced *psbA* translation in plants is triggered by photosystem II damage via an assembly-linked autoregulatory circuit. Proc Natl Acad Sci USA **117**: 21775-21784.

Chotewutmontri, P., Stiffler, N., Watkins, K.P., and Barkan, A. (2018). Ribosome Profiling in Maize. Methods Mol Biol **1676**: 165-183.

Chotewutmontri, P., Williams-Carrier, R., and Barkan, A. (2020. Exploring the link between photosystem II assembly and translation of the chloroplast *psbA* mRNA. Plants **9:** pii: E152.

Covshoff, S., Szecowka, M., Hughes, T.E., Smith-Unna, R., Kelly, S., Bailey, K.J., Sage, T.L., Pachebat, J.A., Leegood, R., and Hibberd, J.M. (2016). C4 Photosynthesis in the Rice Paddy: Insights from the Noxious Weed Echinochloa glabrescens. Plant Physiol **170**: 57-73.

Denton, A.K., Mass, J., Kulahoglu, C., Lercher, M.J., Bräutigam, A., and Weber, A.P. (2017). Freeze-quenched maize mesophyll and bundle sheath separation uncovers bias in previous tissue-specific RNA-Seq data. J Exp Bot **68**: 147-160.

Ding, Z., Weissmann, S., Wang, M., Du, B., Huang, L., Wang, L., Tu, X., Zhong, S., Myers, C., Brutnell, T.P.*, et al.* (2015). Identification of Photosynthesis-Associated C4 Candidate Genes through Comparative Leaf Gradient Transcriptome in Multiple Lineages of C3 and C4 Species. PLoS One **10**: e0140629.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics **29**: 15-21.

Dunning, L.T., Moreno-Villena, J.J., Lundgren, M.R., Dionora, J., Salazar, P., Adams, C., Nyirenda, F., Olofsson, J.K., Mapaura, A., Grundy, I.M.*, et al.* (2019). Key changes in gene expression identified for different stages of C4 evolution in Alloteropsis semialata. J Exp Bot **70**: 3255-3268.

Eisenberg, A.R., Higdon, A.L., Hollerer, I., Fields, A.P., Jungreis, I., Diamond, P.D., Kellis, M., Jovanovic, M., and Brar, G.A. (2020). Translation Initiation Site Profiling Reveals Widespread Synthesis of Non-AUG-Initiated Protein Isoforms in Yeast. Cell Syst **11**: 145-160 e145.

Ermakova, M., Danila, F.R., Furbank, R.T., and von Caemmerer, S. (2020). On the road to C4 rice: advances and perspectives. Plant J **101**: 940-950.

Fei, Y., Wang, R., Li, H., Liu, S., Zhang, H., and Huang, J. (2018). DPMIND: degradome-based plant miRNA-target interaction and network database. Bioinformatics **34**: 1618-1620.

Friso, G., Majeran, W., Huang, M., Sun, Q., and van Wijk, K.J. (2010). Reconstruction of metabolic pathways, protein expression, and homeostasis machineries across maize bundle sheath and mesophyll chloroplasts: large-scale quantitative proteomics using the first maize genome assembly. Plant Physiol **152**: 1219-1250.

Gowik, U., Brautigam, A., Weber, K.L., Weber, A.P., and Westhoff, P. (2011). Evolution of C4 photosynthesis in the genus Flaveria: how many and which genes does it take to make C4? Plant Cell **23**: 2087-2105.
Hatch, M. (1987). C4 photosynthesis: a unique blend of modified biochemistry, anatomy and ultrastructure. BBA Rev Bioenerget **895**: 81-106.
Hinnebusch, A.G., Ivanov, I.P., and Sonenberg, N. (2016). Translational control by 5'-untranslated regions of eukaryotic mRNAs. Science **352**: 1413-1416.
Huang, P., and Brutnell, T.P. (2016). A synthesis of transcriptomic surveys to dissect the genetic basis of C4 photosynthesis. Curr Opin Plant Biol **31**: 91-99.
Huang, P., Studer, A.J., Schnable, J.C., Kellogg, E.A., and Brutnell, T.P. (2017). Cross species selection scans identify components of C4 photosynthesis in the grasses. J Exp Bot **68**: 127-135.
John, C.R., Smith-Unna, R.D., Woodfield, H., Covshoff, S., and Hibberd, J.M. (2014). Evolutionary convergence of cell-specific gene expression in independent lineages of C4 grasses. Plant Physiol **165**: 62-75.
Juarez, M.T., Kui, J.S., Thomas, J., Heller, B.A., and Timmermans, M.C. (2004). microRNA-mediated repression of rolled leaf1 specifies maize leaf polarity. Nature **428**: 84-88.
Külahoglu, C., Denton, A.K., Sommer, M., Mass, J., Schliesky, S., Wrobel, T.J., Berckmans, B., Gongora-Castillo, E., Buell, C.R., Simon, R.*, et al.* (2014). Comparative transcriptome atlases reveal altered gene expression modules between two Cleomaceae C3 and C4 plant species. Plant Cell **26**: 3243-3260.
Lauterbach, M., Schmidt, H., Billakurthi, K., Hankeln, T., Westhoff, P., Gowik, U., and Kadereit, G. (2017). De novo Transcriptome Assembly and Comparison of C3, C3-C4, and C4 Species of Tribe Salsoleae (Chenopodiaceae). Front Plant Sci **8**: 1939.
Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., and Genome Project Data Processing, S. (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics **25**: 2078-2079.
Li, P., Ponnala, L., Gandotra, N., Wang, L., Si, Y., Tausta, S.L., Kebrom, T.H., Provart, N., Patel, R., Myers, C.R.*, et al.* (2010). The developmental dynamics of the maize leaf transcriptome. Nat Genet **42**: 1060-1067.
Li, Y., Ma, X., Zhao, J., Xu, J., Shi, J., Zhu, X.G., Zhao, Y., and Zhang, H. (2015). Developmental genetic mechanisms of C4 syndrome based on transcriptome analysis of C3 cotyledons and C4 assimilating shoots in Haloxylon ammodendron. PLoS One **10**: e0117175.
Liao, Y., Smyth, G.K., and Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics **30**: 923-930.
Link, S., Engelmann, K., Meierhoff, K., and Westhoff, P. (2012) The atypical short-chain dehydrogenases HCF173 and HCF244 are jointly involved in translational initiation of the *psbA* mRNA of Arabidopsis. Plant Physiol **160:** 2202-2218.
Liu, W.Y., Chang, Y.M., Chen, S.C., Lu, C.H., Wu, Y.H., Lu, M.Y., Chen, D.R., Shih, A.C., Sheue, C.R., Huang, H.C.*, et al.* (2013). Anatomical and transcriptional dynamics of maize embryonic leaves during seed germination. Proc Natl Acad Sci U S A **110**: 3979-3984.
Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol **15**: 550.
Luo, J., Zhou, J.J., and Zhang, J.Z. (2018). Aux/IAA Gene Family in Plants: Molecular Structure, Regulation, and Function. Int J Mol Sci **19**.
Majeran, W., Cai, Y., Sun, Q., and van Wijk, K.J. (2005). Functional differentiation of bundle sheath and mesophyll maize chloroplasts determined by comparative proteomics. The Plant cell **17**: 3111-3140.
Majeran, W., Friso, G., Ponnala, L., Connolly, B., Huang, M., Reidel, E., Zhang, C., Asakura, Y., Bhuiyan, N.H., Sun, Q.*, et al.* (2010). Structural and metabolic transitions of C4 leaf development and differentiation defined by microscopy and quantitative proteomics in maize. The Plant cell **22**: 3509-3542.
Majeran, W., and Wijk, K.v. (2009). Cell-type-specific differentiation of chloroplasts in C4 plants. Trends Plant Sci **14**: 100-109. **i**
Majeran, W., Zybailov, B., Ytterberg, A.J., Dunsmore, J., Sun, Q., and van Wijk, K.J. (2008). Consequences of C4 differentiation for chloroplast membrane proteomes in maize mesophyll and bundle sheath cells. Molec Cell Proteomics **7**: 1609-1638.
Martin, M. (2011). Cutadapt Removes Adapter Sequences From High-Throughput Sequencing Reads. EMBnetjournal **17**: 10-12.

Mattiello, L., Riano-Pachon, D.M., Martins, M.C., da Cruz, L.P., Bassi, D., Marchiori, P.E., Ribeiro, R.V., Labate, M.T., Labate, C.A., and Menossi, M. (2015). Physiological and transcriptional analyses of developmental stages along sugarcane leaf. BMC Plant Biol **15**: 300.
McManus, C.J., May, G.E., Spealman, P., and Shteyman, A. (2014). Ribosome profiling reveals post-transcriptional buffering of divergent gene expression in yeast. Genome Research **24**: 422-430.
Nelson, J.M., Lane, B., and Freeling, M. (2002). Expression of a mutant maize gene in the ventral leaf epidermis is sufficient to signal a switch of the leaf's dorsoventral axis. Development **129**: 4581-4589.
Oertlin, C., Lorent, J., Murie, C., Furic, L., Topisirovic, I., and Larsson, O. (2019). Generally applicable transcriptome-wide analysis of translation using anota2seq. Nucl Acids Res **47**: e70.
Offermann, S., Friso, G., Doroshenk, K.A., Sun, Q., Sharpe, R.M., Okita, T.W., Wimmer, D., Edwards, G.E., and van Wijk, K.J. (2015). Developmental and Subcellular Organization of Single-Cell C(4) Photosynthesis in Bienertia sinuspersici Determined by Large-Scale Proteomics and cDNA Assembly from 454 DNA Sequencing. J Proteome Res **14**: 2090-2108.
Pick, T.R., Brautigam, A., Schluter, U., Denton, A.K., Colmsee, C., Scholz, U., Fahnenstich, H., Pieruschka, R., Rascher, U., Sonnewald, U.*, et al.* (2011). Systems analysis of a maize leaf developmental gradient redefines the current C4 model and provides candidates for regulation. Plant Cell **23**: 4208-4220.
Ponnala, L., Wang, Y., Sun, Q., and van Wijk, K.J. (2014). Correlation of mRNA and protein abundance in the developing maize leaf. The Plant journal : for cell and molecular biology **78**: 424-440.
Rao, X., Lu, N., Li, G., Nakashima, J., Tang, Y., and Dixon, R.A. (2016). Comparative cell-specific transcriptomics reveals differentiation of C4 photosynthesis pathways in switchgrass and other C4 lineages. J Exp Bot **67**: 1649-1662.
Sage, R.F., Christin, P.A., and Edwards, E.J. (2011). The C(4) plant lineages of planet Earth. J Exp Bot **62**: 3155-3169.
Schafer, S., Adami, E., Heinig, M., Rodrigues, K.E.C., Kreuchwig, F., Silhavy, J., van Heesch, S., Simaite, D., Rajewsky, N., Cuppen, E.*, et al.* (2015). Translational regulation shapes the molecular landscape of complex disease phenotypes. Nature Comm **6**: 7200.
Schult, K., Meierhoff, K., Paradies, S., Toller, T., Wolff, P. and Westhoff, P. (2007) The nuclear-encoded factor HCF173 is involved in the initiation of translation of the *psbA* mRNA in Arabidopsis thaliana. Plant Cell **19:** 1329-1346.
Schlüter, U., Bräutigam, A., Droz, J.M., Schwender, J., and Weber, A.P.M. (2019). The role of alanine and aspartate aminotransferases in C4 photosynthesis. Plant Biol (Stuttg) **21 Suppl 1**: 64-76.
Schlüter, U., Bräutigam, A., Gowik, U., Melzer, M., Christin, P.A., Kurz, S., Mettler-Altmann, T., and Weber, A.P. (2017). Photosynthesis in C3-C4 intermediate Moricandia species. J Exp Bot **68**: 191-206.
Schlüter, U., and Weber, A.P.M. (2020). Regulation and Evolution of C4 Photosynthesis. Annu Rev Plant Biol **71**: 183-215.
Sedelnikova, O.V., Hughes, T.E., and Langdale, J.A. (2018). Understanding the Genetic Basis of C4 Kranz Anatomy with a View to Engineering C3 Crops. Annu Rev Genet **52**: 249-270.
Tausta, S.L., Li, P., Si, Y., Gandotra, N., Liu, P., Sun, Q., Brutnell, T.P., and Nelson, T. (2014). Developmental dynamics of Kranz cell transcriptional specificity in maize leaf reveals early onset of C4-related processes. J Exp Bot **65**: 3543-3555.
Thimm, O., Blasing, O., Gibon, Y., Nagel, A., Meyer, S., Kruger, P., Selbig, J., Muller, L.A., Rhee, S.Y., and Stitt, M. (2004). MAPMAN: a user-driven tool to display genomics data sets onto diagrams of metabolic pathways and other biological processes. Plant J **37**: 914-939.
van der Horst, S., Snel, B., Hanson, J., and Smeekens, S. (2019). Novel pipeline identifies new upstream ORFs and non-AUG initiating main ORFs with conserved amino acid sequences in the 5' leader of mRNAs in Arabidopsis thaliana. RNA **25**: 292-304.
Wang, J., Mei, J., and Ren, G. (2019). Plant microRNAs: Biogenesis, Homeostasis, and Degradation. Front Plant Sci **10**: 360.
Wang, L., Czedik-Eysenberg, A., Mertz, R.A., Si, Y., Tohge, T., Nunes-Nesi, A., Arrivault, S., Dedow, L.K., Bryant, D.W., Zhou, W.*, et al.* (2014). Comparative analyses of C(4) and C(3) photosynthesis in developing leaves of maize and rice. Nat Biotechnol **32**: 1158-1165.
Wang, P., Fouracre, J., Kelly, S., Karki, S., Gowik, U., Aubry, S., Shaw, M.K., Westhoff, P., Slamet-Loedin, I.H., Quick, W.P.*, et al.* (2013a). Evolution of GOLDEN2-LIKE gene function in C(3) and C (4) plants. Planta **237**: 481-495.

Wang, P., Kelly, S., Fouracre, J.P., and Langdale, J.A. (2013b). Genome-wide transcript analysis of early maize leaf development reveals gene cohorts associated with the differentiation of C4 Kranz anatomy. Plant J **75**: 656-670.
Wang, P., Vlad, D., and Langdale, J.A. (2016). Finding the genes to build C4 rice. Curr Opin Plant Biol **31**: 44-50.
Williams, B.P., Burgess, S.J., Reyna-Llorens, I., Knerova, J., Aubry, S., Stanley, S., and Hibberd, J.M. (2016). An Untranslated cis-Element Regulates the Accumulation of Multiple C4 Enzymes in Gynandropsis gynandra Mesophyll Cells. Plant Cell **28**: 454-465.
Wu, C.C., Zinshteyn, B., Wehner, K.A., and Green, R. (2019). High-Resolution Ribosome Profiling Defines Discrete Ribosome Elongation States and Translational Regulation during Cellular Stress. Molec Cell **73**: 959-970 e955.
Xiao, Z., Zou, Q., Liu, Y., and Yang, X. (2016). Genome-wide assessment of differential translations with ribosome profiling data. Nature Commun **7**: 11194.
Zhang, Y., Paschold, A., Marcon, C., Liu, S., Tai, H., Nestler, J., Yeh, C.T., Opitz, N., Lanz, C., Schnable, P.S.*, et al.* (2014). The Aux/IAA gene rum1 involved in seminal and lateral root formation controls vascular patterning in maize (Zea mays L.) primary roots. J Exp Bot **65**: 4919-4930.
Zones, J.M., Blaby, I.K., Merchant, S.S., and Umen, J.G. (2015). High-Resolution Profiling of a Synchronized Diurnal Transcriptome from Chlamydomonas reinhardtii Reveals Continuous Cell and Metabolic Differentiation. Plant Cell **27**: 2743-2769.

Table 1. Summary of data for the 39 genes reported by both XTAIL and anota2seq as differentially translated in BS and M cells. *Fourteen genes with core roles in C_4 photosynthesis.

Gene ID	Gene Product/Description	DESeq2				Translational Efficiency (log_2 BS/M)				Regulation mode
		Ribo-seq (log_2 BS/M)	FDR (ribo-seq)	RNA-seq (log_2 BS/M)	FDR (RNA-seq)	XTAIL log_2 BS/M	XTAIL FDR	Anota2seq log_2 BS/M	Anota2seq FDR	
Zm00001d022608*	PIP2-2, aquaporin family	4.71	0.0000	2.80	0.0000	1.75	0.0000	4.78	0.0311	RNA+TE
Zm00001d048703	BX4, benzoxazinone synthesis 4	3.93	0.0000	1.80	0.0000	1.98	0.0000	3.95	0.0315	RNA+TE
Zm00001d007319	subtilisin-like protease	4.97	0.0000	3.35	0.0000	1.46	0.0002	4.53	0.0364	RNA+TE
Zm00001d042050*	MEP3, chloroplast pyruvate transporter candidate	3.52	0.0000	2.04	0.0000	1.32	0.0000	3.07	0.0364	RNA+TE
Zm00001d000164*	PRK, phosphoribulokinase, Calvin-Benson cycle	3.46	0.0000	2.11	0.0000	1.20	0.0005	2.93	0.0372	RNA+TE
Zm00001d028562*	FBP, fructose-1,6-bisphosphatase, Calvin-Benson cycle	3.04	0.0000	0.92	0.0172	1.97	0.0000	3.02	0.0364	RNA+TE
Zm00001d018936*	RPI, ribose-5-phosphate isomerase, Calvin-Benson cycle/Pentose phosphate pathway	3.23	0.0000	1.33	0.0008	1.74	0.0000	3.21	0.0372	RNA+TE
Zm00001d002258*	GDCT, glycine decarboxylase system T-protein, Photorespiration	2.25	0.0000	1.09	0.0001	1.01	0.0000	2.37	0.0447	RNA+TE
Zm00001d018623*	PnsL3/PQL2, luminal subunit of NDH	3.32	0.0000	1.37	0.0039	1.79	0.0000	2.75	0.0463	RNA+TE
Zm00001d023757*	NdhU, subunit of NDH subcomplex EDB	2.97	0.0000	1.68	0.0000	1.14	0.0001	2.27	0.0472	RNA+TE
Zm00001d011006	ABC Transporter	3.86	0.0000	1.15	0.1557	2.54	0.0003	3.13	0.0372	RNA+TE
Zm00001d044925*	CP12, regulates GAPDH and PRK in Calvin-Benson cycle	3.64	0.0000	0.99	0.0024	2.49	0.0000	3.27	0.0404	RNA+TE
Zm00001d011063	MT, Metallothionein-like protein 2C	4.69	0.0000	1.08	0.0908	3.46	0.0000	4.98	0.0472	RNA+TE
Zm00001d004936	Ubiquitin-like protein	2.61	0.0000	0.55	0.2547	1.90	0.0000	2.46	0.0430	RNA+TE
Zm00001d042620*	NdhN, subunit of NDH subcomplex A	2.53	0.0000	0.36	0.4974	2.01	0.0000	2.50	0.0472	RNA+TE
Zm00001d018157	LHCB4, light harvesting complex a/b protein 4	2.30	0.0000	0.51	0.1394	1.63	0.0000	2.32	0.0472	RNA+TE
Zm00001d000045	Trihelix transcription factor 28	3.02	0.0000	0.39	0.6061	2.47	0.0000	3.50	0.0364	TE
Zm00001d024637	L-lectin kinase, receptor kinase	2.81	0.0000	-0.18	0.8327	2.84	0.0000	2.82	0.0372	TE
Zm00001d052875	GDPDL, glycerophosphodiester phosphodiesterase	2.92	0.0000	0.10	0.9091	2.67	0.0000	3.15	0.0409	TE
Zm00001d033194	subtilisin-like protease	2.21	0.0000	-0.18	0.8376	2.23	0.0000	2.24	0.0482	TE
Zm00001d029087	SUS, sucrose synthase	2.65	0.0000	0.24	0.7059	2.26	0.0000	2.88	0.0500	TE
Zm00001d046182	OFUT, O-fucosyltransferase family	2.95	0.0000	0.24	0.7947	2.58	0.0002	2.71	0.0500	TE
Zm00001d020008	IGPS, chloroplast indole-3-glycerol phosphate synthase	-3.65	0.0000	-1.56	0.0074	-2.25	0.0000	-3.13	0.0372	TE
Zm00001d004855	Actin	-3.31	0.0000	-1.12	0.0334	-2.34	0.0000	-3.59	0.0430	TE
Zm00001d006738	QSP, queuosine salvage protein candidate	-3.39	0.0000	-1.27	0.0150	-2.28	0.0000	-3.54	0.0447	TE
Zm00001d015138	Unknown protein	-2.54	0.0000	-0.83	0.0005	-1.86	0.0000	-2.44	0.0430	TE
Zm00001d036788	subtilisin-like protease	-3.74	0.0000	-1.00	0.0481	-2.90	0.0000	-4.74	0.0447	TE
Zm00001d014178	AARE, acylamino-acid-releasing enzyme	-5.36	0.0000	-1.96	0.0000	-3.56	0.0000	-5.00	0.0315	RNA+TE
Zm00001d000060	NAD(P)-linked oxidoreductase superfamily protein	-5.29	0.0000	-2.64	0.0000	-2.80	0.0000	-4.40	0.0364	RNA+TE
Zm00001d035737*	GLYK, glycerate kinase, Photorespiration	-4.91	0.0000	-2.53	0.0000	-2.52	0.0000	-4.17	0.0367	RNA+TE
Zm00001d039542	PC-PLA1, chloroplast phospholipase A1	-4.19	0.0000	-2.59	0.0000	-1.76	0.0000	-4.07	0.0467	RNA+TE
Zm00001d030557*	AlaAT, cytosolic alanine aminotransferase	-3.10	0.0000	-2.12	0.0000	-1.13	0.0000	-3.42	0.0372	RNA+TE
Zm00001d002490	CP33B, chloroplast RRM protein	-2.92	0.0000	-1.69	0.0000	-1.37	0.0000	-3.90	0.0372	RNA+TE
Zm00001d039104	CHUP1, chloroplast outer envelope protein	-2.92	0.0000	-1.86	0.0000	-1.21	0.0000	-3.32	0.0401	RNA+TE
Zm00001d005950*	DIT1, chloroplast dicarboxylic acid transporter 1	-3.66	0.0000	-2.46	0.0000	-1.36	0.0000	-4.50	0.0430	RNA+TE
Zm00001d006520*	PPDK-RP, PPDK regulatory protein, C_4 metabolism	-4.93	0.0000	-3.26	0.0000	-1.82	0.0000	-5.71	0.0311	RNA+TE
Zm00001d009022	PCBER, phenylcoumaran benzylic ether reductase	-4.44	0.0000	-3.37	0.0000	-1.23	0.0000	-3.97	0.0364	RNA+TE
Zm00001d020852	CSLE1, cellulose synthase-like protein E1	-5.63	0.0000	-3.61	0.0000	-2.17	0.0000	-5.01	0.0372	RNA+TE
Zm00001d032828	Citrate transporter family protein	-4.38	0.0000	-3.42	0.0000	-1.12	0.0001	-3.34	0.0474	RNA+TE

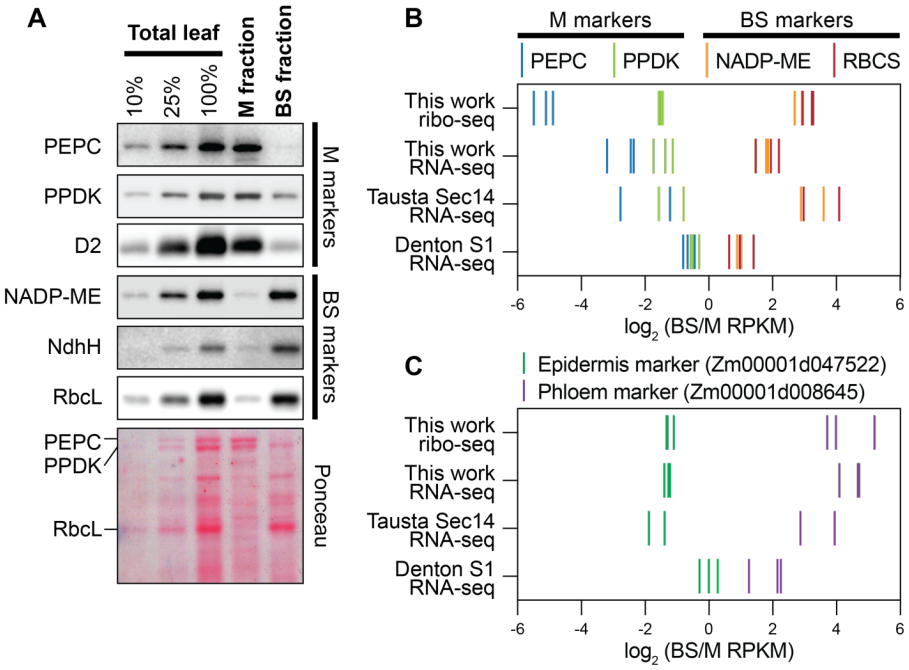


Figure 1. Enrichment of cell-type specific markers in BS and M fractions.
(A) Immunoblots showing abundance of marker proteins in BS and M fractions. Samples were loaded on the basis of equal protein. Replicate blots were probed with the indicated antibodies. BS markers: NADP-dependent malic enzyme (NADP-ME), H subunit of the NDH-like complex (NdhH), Rubisco large subunit (RbcL). M markers: phosphoenolpyruvate carboxylase (PEPC), pyruvate orthophosphate dikinase (PPDK), D2 subunit of Photosystem II (D2). An image of one blot stained with Ponceau S is shown below.
(B) Relative abundance of RNA-seq and ribo-seq reads from BS and M marker genes. BS markers: NADP-ME (Zm00001d000316) and Rubisco small subunit (RBCS; Zm00001d052595). M markers: PEPC (Zm00001d046170) and PPDK (Zm00001d038163). RNA-seq data from Tausta et al (2014) (leaf section 14 with 2 replicates) and Denton et al (2017) (leaf section S1 with 3 replicates) are compared with our data. Each line symbol represents normalized reads for one gene (see color key at top) in one replicate. RPKM, reads per kilobase per million reads mapped to nuclear coding sequences.
(C) Relative abundance of RNA-seq and ribo-seq reads from epidermis and phloem marker genes. Data for additional marker genes are shown in Supplemental Figure 2C.

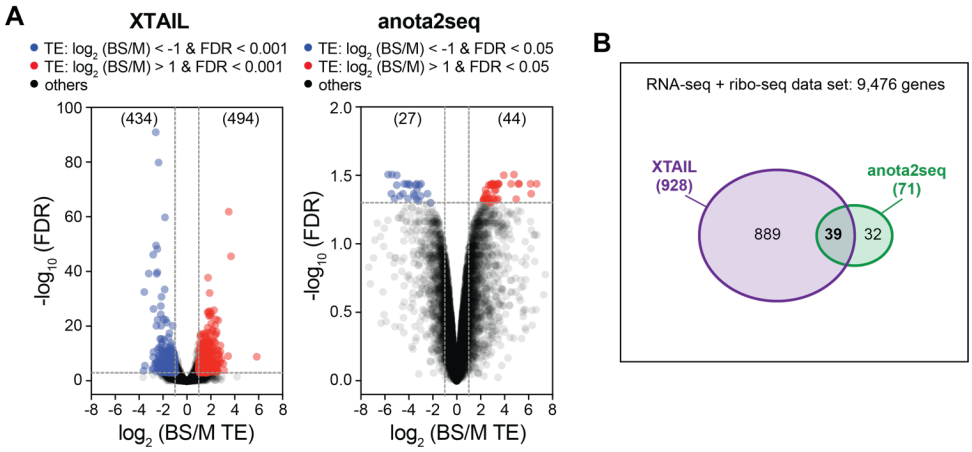


Figure 2. Overview of analysis of differential translational efficiency (TE).

(A) Volcano plots summarizing the output of XTAIL and anota2seq. Each symbol represents one gene. Genes exhibiting preferential translation in mesophyll (M) or bundle sheath (BS) fractions are shown in blue or red symbols, respectively. Numbers of genes are shown in parentheses. Dashed lines indicate the FDR and fold-change cutoffs we used to define differentially translated genes.

(B) Comparison of gene sets reported by XTAIL [$|\log_2 (BS/M)| > 1$ and FDR < 0.001] and anota2seq [$|\log_2 (BS/M)| > 1$ and FDR < 0.05] to exhibit differences in translational efficiency in BS and M fractions.

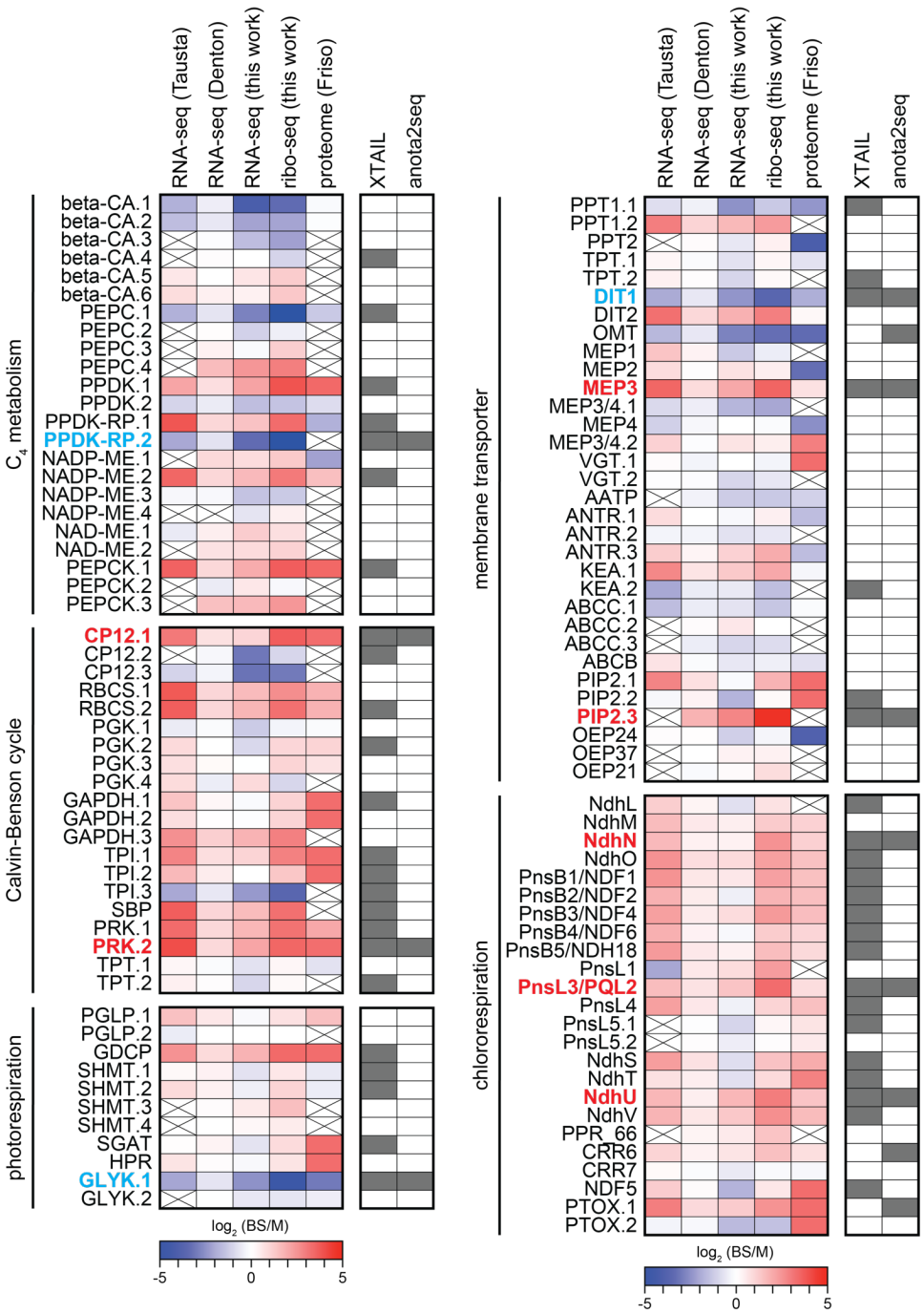


Figure 3. Differential expression of genes of particular relevance to C_4 metabolism.

RNA-seq and ribo-seq data reported here are compared with data from Tausta et al (2014) (leaf section 14), Denton et al (2017) (leaf section S1), and Friso et al (2010). The previous RNA-seq data were re-analyzed with our pipeline. Heatmaps show $\log_2$ (BS/M) values. Crossed boxes indicate read counts below our cutoff or lack of data. Genes reported as differentially translated by XTAIL or anota2seq are marked with gray boxes. The ten genes reported by both algorithms are highlighted in red (BS-enriched) or blue (M-enriched) font. The genes involved in C_4 metabolism, photorespiration, and the Calvin-Benson cycle were taken from (Schlüter and Weber, 2020). Genes encoding membrane transporters came from (Majeran and Wijk, 2009). Genes associated with chlororespiration and the NDH complex come from MapMan X4.2 BIN 1.1.8. The gene IDs and data values are provided in Supplemental Data Set 3.

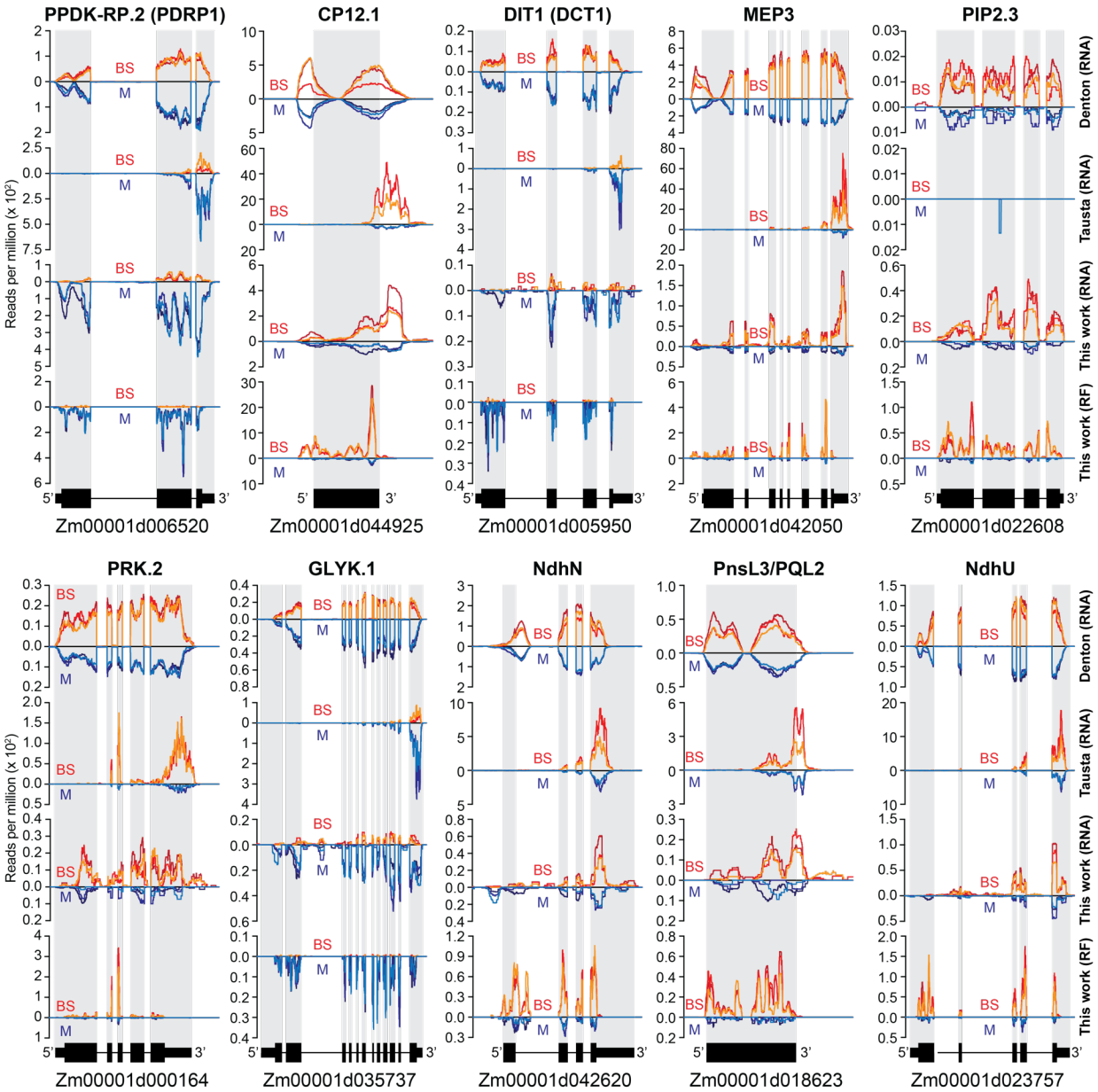


Figure 4. Normalized read coverage of C_4-related genes reported by both XTAIL and anota2seq as being differentially translated in BS and M fractions.

Normalized read coverage from RNA-seq (RNA) and ribo-seq (RF) data are shown for the ten genes highlighted as exhibiting differential translation in Figure 3. Read coverage in BS or M fractions is shown above or below the line, respectively. Coverage was normalized by million reads mapped to nuclear coding sequences. Each line represents one replicate. As expected based on a prior analysis (Denton et al., 2017), the Tausta data show strong 3' bias and the Denton data show less robust cell-type specific enrichment. PDRP1, pyruvate orthophosphate dikinase regulatory protein 1; DCT1, dicarboxylic acid transporter 1.

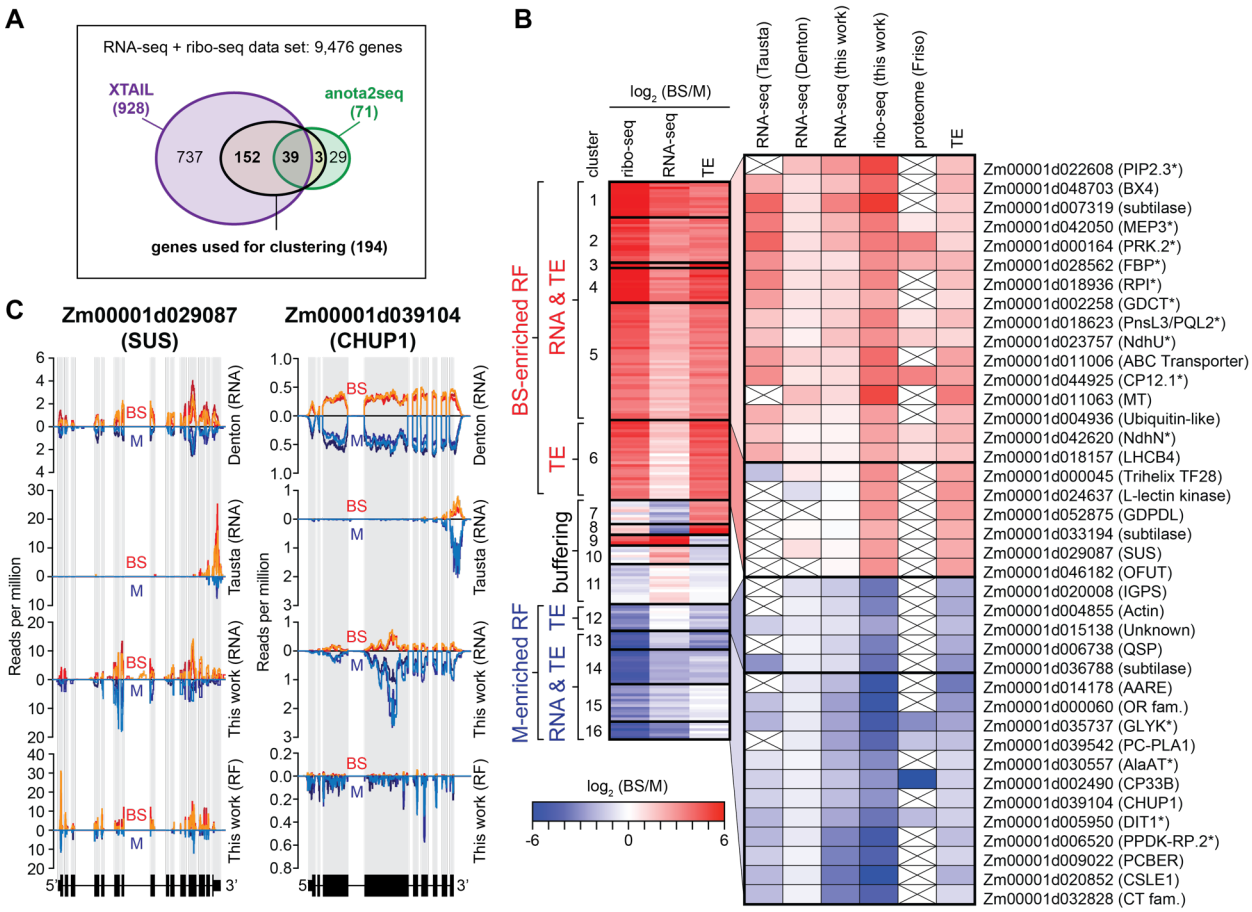


Figure 5. Relative contributions of translational control and differential mRNA abundance to differential expression of translationally regulated genes.

(A) Expanded set of differentially translated genes used for *k*-means clustering. The 194 genes in the clustering set includes genes reported by both XTAIL and anota2seq to exhibit $|\log_2 (BS/M)| > 1$ and that passed the FDR criterion (XTAIL FDR < 0.001, anota2seq FDR < 0.05) in at least one of the programs.
(B) The Clustering gene set (see panel A) was used for *k*-means clustering based on $\log_2$ (BS/M) values of RNA-seq reported by DESeq2 and translational efficiency (TE) reported by XTAIL. Ribo-seq BS/M ratios (left column) were added after clustering. Clusters are labeled according to the dominant feature underlying differential expression (TE, RNA, or both). Buffering refers to clusters in which the difference in TE is opposite to that of the difference in RNA abundance. The heatmap to the right compares our data to prior RNA-seq and proteomic data for the 39 genes reported by both XTAIL and anota2seq to be differentially translated. The values shown for prior RNA-seq datasets are based on re-analysis with our pipeline. Crossed boxes indicate genes that did not meet our read count cutoffs or that have no data. Asterisks mark core C_4 genes discussed elsewhere. Values underlying these plots are provided in Supplemental Data Sets 4 and 5.
(C) Normalized read coverage of genes encoding CHUP1 and SUS in BS and M fractions. Coverage was normalized by million reads mapped to nuclear coding sequence. Each line represents one replicate. The re-analyzed values of published RNA-seq data are shown for comparison. RNA, RNA-seq data; RF, ribo-seq data.

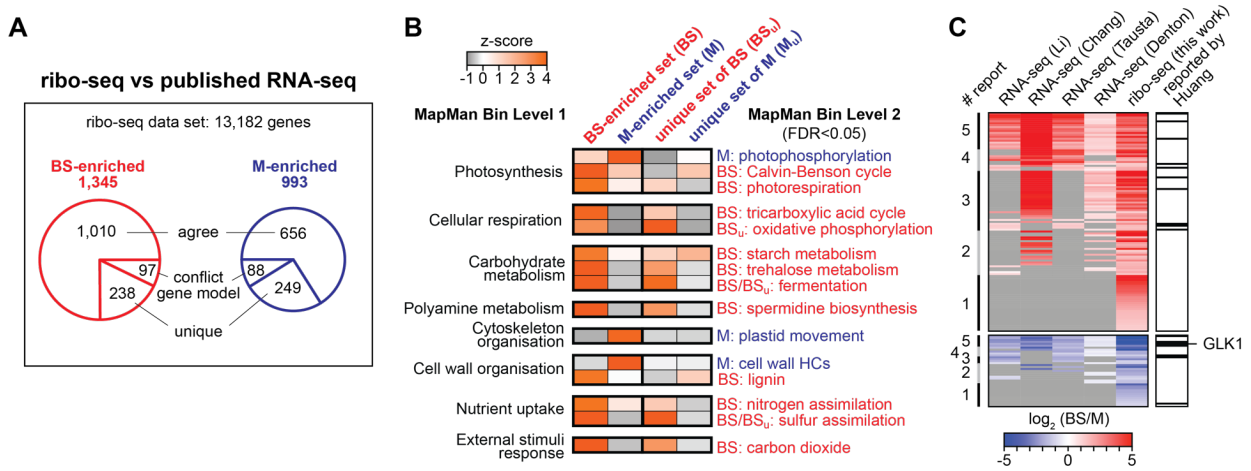


Figure 6. Comparison of translatomes in BS and M fractions.
(A) Data overview. 13,182 genes met the read-count cutoff to be used in this analysis (average of ≥100 reads in at least one of the two fractions). The BS- and M-enriched gene sets [$|\log_2 (BS/M)| > 1$ and FDR < 0.001] were compared to genes that had been reported to be differentially expressed in previous RNA-seq studies: Chang et al (2015), Li et al (2010), Denton et al (2017) section S1, Tausta et al (2014) section 14. The "conflict gene model" sets consist of genes that cannot be compared among studies because of conflicting gene models. The "agree" sets include genes that were reported by at least one prior RNA-seq study to be preferentially expressed in the same tissue type as in our ribo-seq data. The "unique" sets are genes that our ribo-seq data report as differentially expressed but that were not reported as differentially expressed in prior studies.
(B) Functional enrichment analysis of genes identified as differentially expressed in our ribo-seq data that had not been reported as differentially expressed in previous maize studies. Heatmaps display the term enrichment z-scores when comparing BS- and M-enriched gene sets to the set of all expressed genes. Data for the complete BS and M-enriched sets (BS and M) are compared with data for the unique sets described in (A). HC, hydroxycinnamic acid.
(C) Heatmap of $\log_2$ (BS/M) values of ribo-seq data for transcription factor (TF) genes that exhibited BS or M-enriched ribosome footprint abundance. Values from prior RNA-seq studies are shown for comparison. Genes are ranked based on the number of reports concluding that the gene was differentially expressed (# report; left). Genes not reported as differentially expressed are in gray. The TFs that were identified in a cross-species transcriptome analysis by Huang et al (2016) are marked in the right column. Transcription factors reported to be differentially expressed among all studies are listed in Supplemental Data Set 10. GLK1, Golden2-like 1.

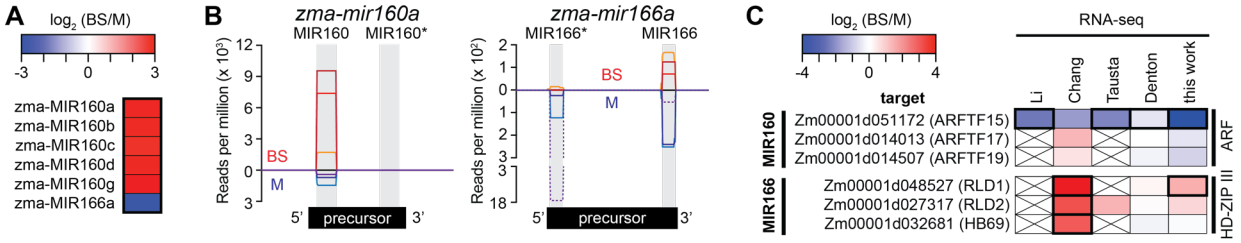
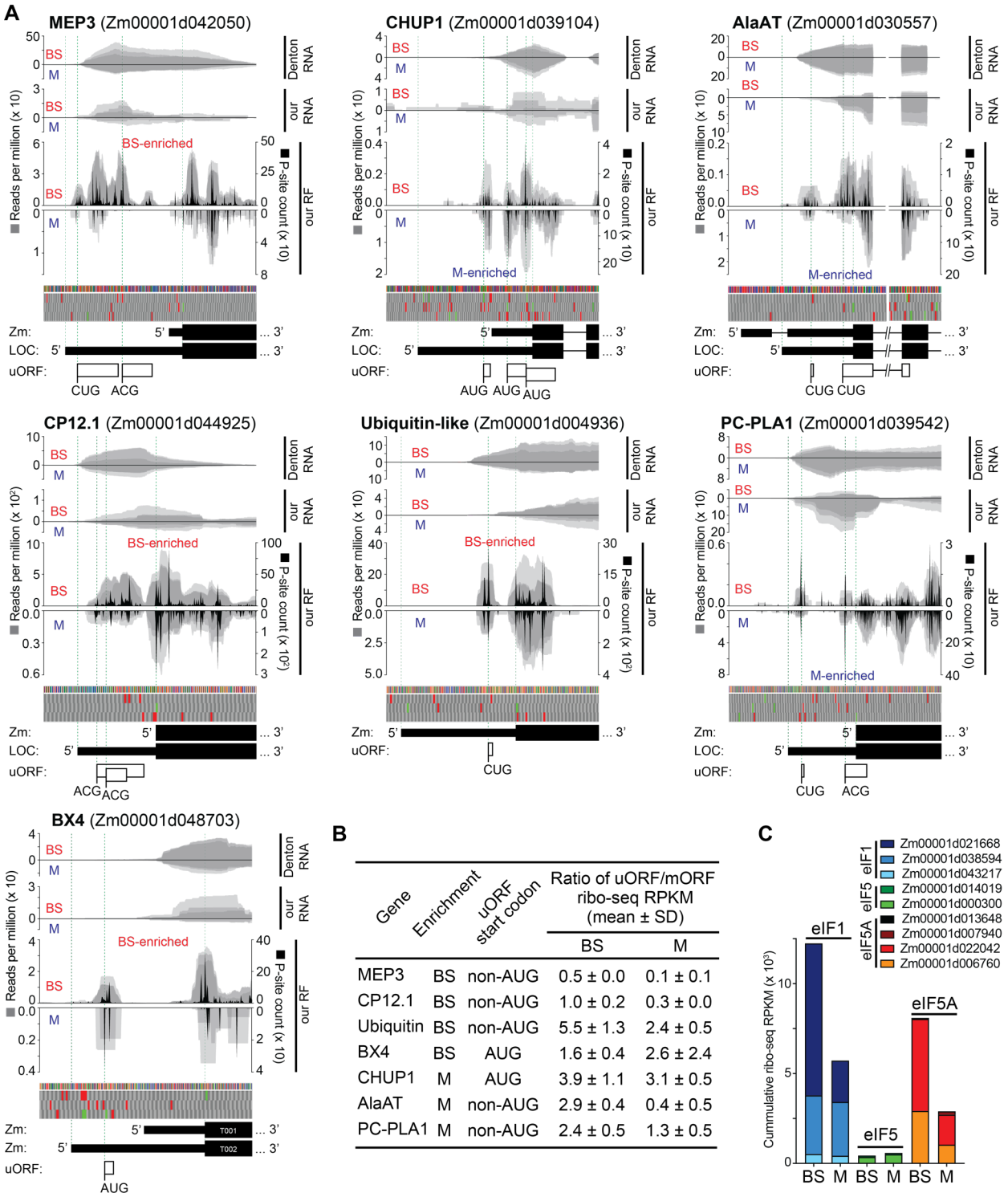


Figure 7. Differential expression of microRNAs and their transcription factor targets in BS and M fractions.

(A) Differential expression of microRNAs from sRNA-seq analysis of BS and M fractions. Six genes identified by DESeq2 are shown [$|\log_2 (BS/M)| > 1$ and FDR < 0.05]. The mature zma-MIR160 microRNAs have identical sequences, so sRNA-seq was unable to distinguish among them.

(B) Normalized read coverage of genes encoding zma-MIR160a and zma-MIR166a. Coverage was normalized by million reads mapped to all microRNA precursors. Single sequence reads spanned the full length of each microRNA, resulting in uniform coverage across their length. Each line represents one replicate, with reads from the BS and M fractions shown above and below the horizontal line, respectively. Each microRNA precursor (diagrammed at bottom) folds into a hairpin via the pairing of complementary regions (gray). One strand in the duplex becomes mature microRNA and the other is typically degraded. An asterisk marks the degraded strand identified by miRBase (http://www.mirbase.org/).

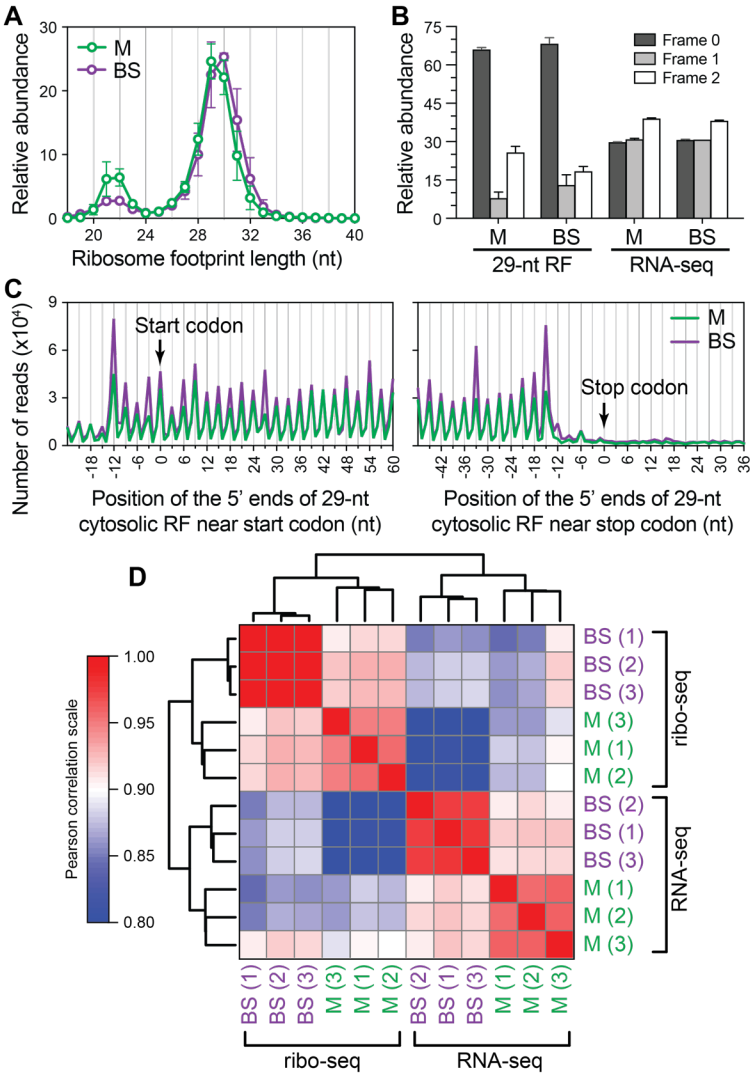
(C) Cell-type specific expression of mRNA targets of differentially expressed microRNAs. The mRNA targets of MIR166 and MIR160 were recovered from the DPMIND database (Fei et al., 2018). The original $\log_2$ (BS/M) values reported in prior RNA-seq datasets are summarized. Genes reported as differentially expressed are bordered in black. Crossed boxes indicate absence of data. ARFTF, auxin response transcription factor; HD-ZIP III, Class III homeodomain leucine zipper; RLD1/2, Rolled leaf1/2; HB69, Homeobox-transcription factor 69.



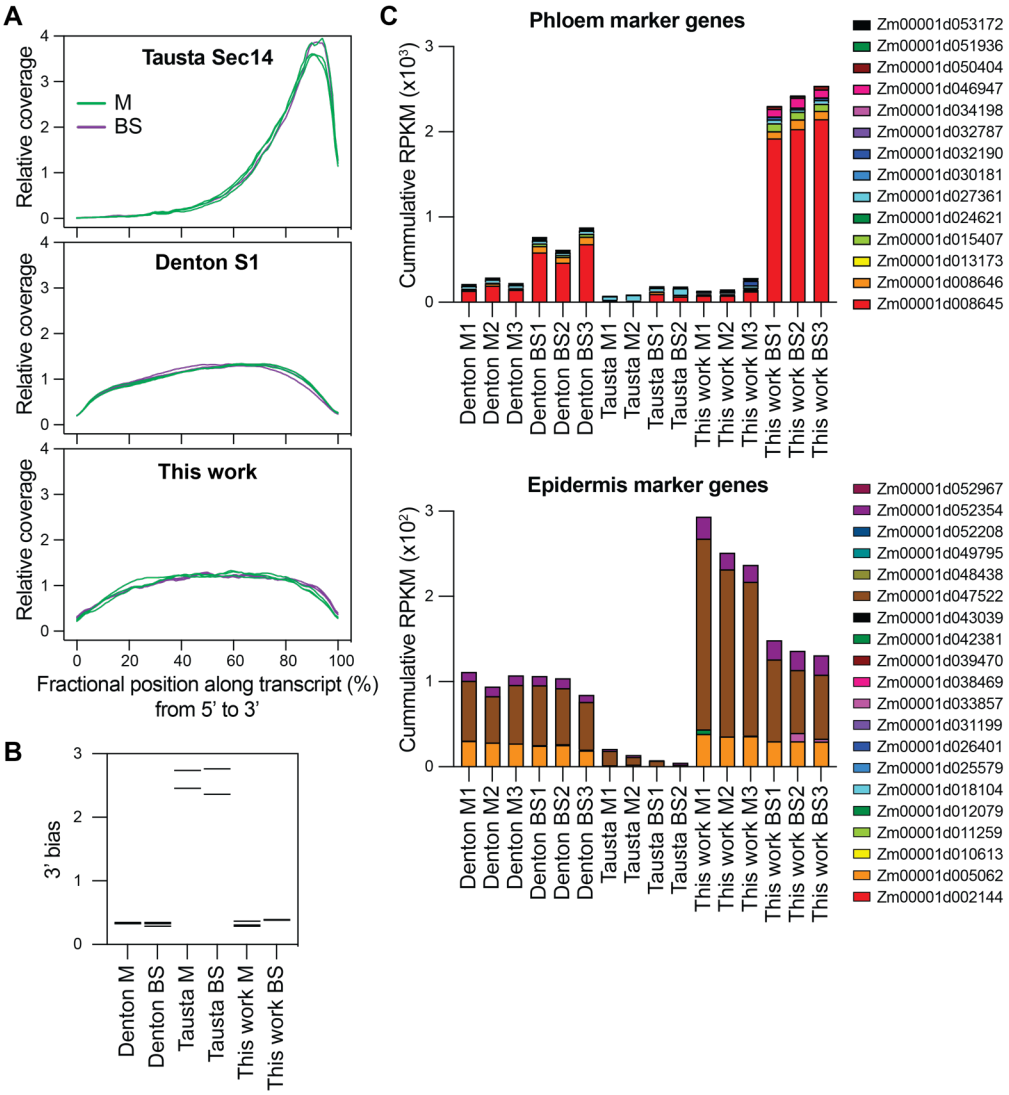
Gene	Enrichment	uORF start codon	Ratio of uORF/mORF ribo-seq RPKM (mean ± SD)	
			BS	M
MEP3	BS	non-AUG	0.5 ± 0.0	0.1 ± 0.1
CP12.1	BS	non-AUG	1.0 ± 0.2	0.3 ± 0.0
Ubiquitin	BS	non-AUG	5.5 ± 1.3	2.4 ± 0.5
BX4	BS	AUG	1.6 ± 0.4	2.6 ± 2.4
CHUP1	M	AUG	3.9 ± 1.1	3.1 ± 0.5
AlaAT	M	non-AUG	2.9 ± 0.4	0.4 ± 0.5
PC-PLA1	M	non-AUG	2.4 ± 0.5	1.3 ± 0.5

Figure 8. Ribosome occupancy in uORFs of differentially translated genes.
(A) Normalized read coverage from the 5'-regions of differentially translated genes with uORFs. Gray shading indicates coverage from each replicate such that overlapping coverage from different replicates appears as a darker shade. Values were normalized by million reads mapped to nuclear coding sequence. The re-analyzed values of the slice 1 samples from Denton et al. (2017) are shown for comparison. P-site occupancies (sum of replicates) are shown in black (scale to the right). Nucleotide

track and 3-frame translation tracks are shown above gene models. The green and red boxes in the translation tracks indicate AUG start and stop codons, respectively. Zm and LOC indicate gene models from Gramene and GenBank, respectively. RNA, RNA-seq; RF, ribo-seq.
(B) Ratio of ribosome occupancies in uORF *versus* mORF. uORF regions overlapping with mORFs were excluded from calculations. mORF values excluded ribosomes bound to the start codon.
(C) Expression of eIF1, eIF5 and eIF5A in BS and M fractions based on ribo-seq data (sum of three replicates). RPKM, reads per kilobase per million reads mapped to nuclear coding sequences.



Supplemental Figure 1. Characteristics of ribo-seq and RNA-seq data from BS and M fractions.
(A) Size distribution of ribosome footprints. Values are the mean ± SD from three biological replicates. Our protocol purifies RNA fragments between 20 and 40 nucleotides for sequencing, so this size distribution is not skewed by the selection of an expected size. The peak at ~21 nucleotides presumably represents footprints of ribosomes with an empty A site (Wu et al., 2019).
(B) Three-nucleotide periodicity of ribo-seq data. The frame placements were inferred from the locations of the 5' ends of footprints at start and stop codons. Values are the mean ± SD from three replicates.
(C) Metagene analysis of reads mapping near all cytosolic start and stop codons. The sum from three replicates is shown.
(D) Pearson correlation coefficients were calculated using $\log_{10}$ RPKM (reads per million reads mapped to nuclear coding sequences) values of protein coding genes in the nuclear genome. Hierarchical clustering based on Pearson correlation is shown. The number in parentheses indicates replicate number.

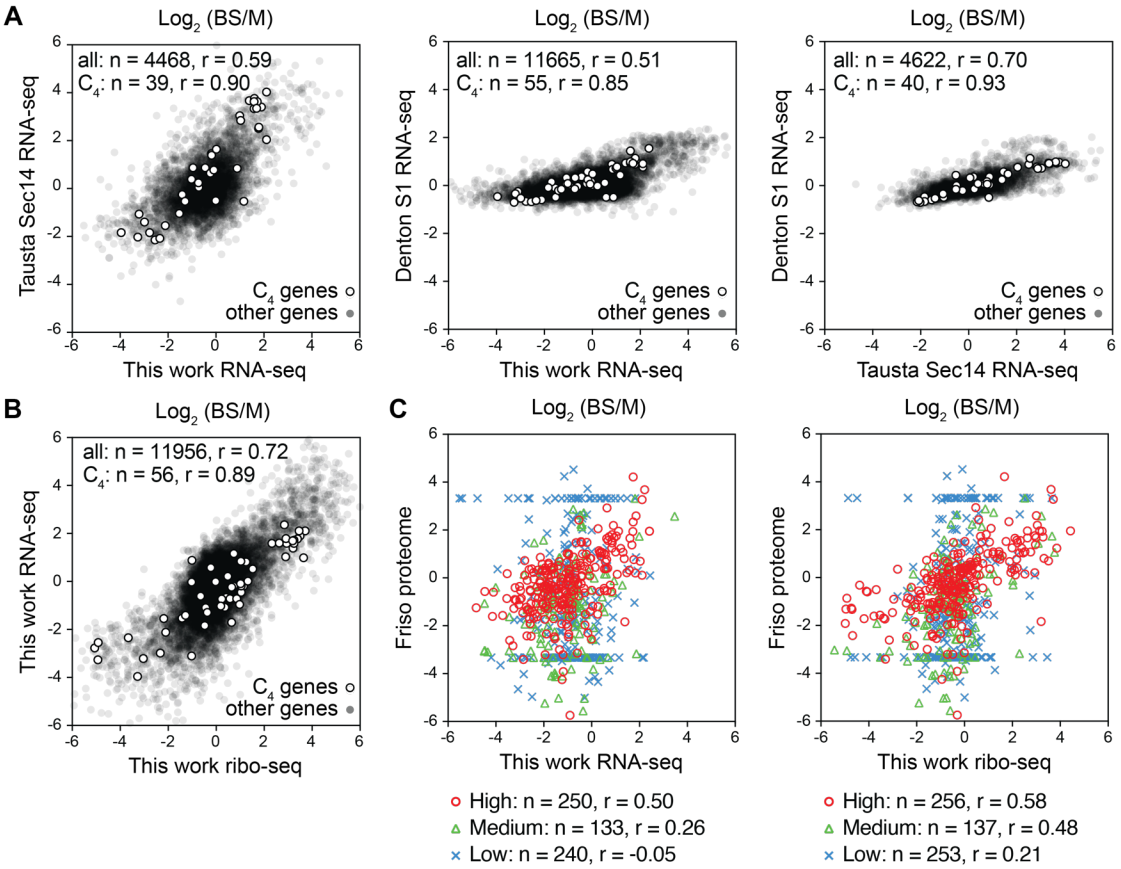


Supplemental Figure 2. Comparison of RNA-seq data characteristics with those from prior studies. Values come from Tausta et al (2014) Leaf Section 14 (Sec 14) and Denton et al (2017) Slice 1 (S1) after reanalysis of those data with our pipeline.

(A) Metagene analysis of read coverage of all transcripts analyzed by Picard software (http://broadinstitute.github.io/picard/). Each line represents one replicate. As expected based on a prior analysis (Denton et al., 2017), the Tausta data show strong 3' bias.

(B) 3' bias reported by Picard software. Each line symbol represents one replicate.

(C) Contamination of phloem and epidermis marker genes in RNA-seq data. Genes with MapMan descriptions including the words 'phloem' or 'epidermal' were assigned as phloem and epidermis marker genes, respectively.

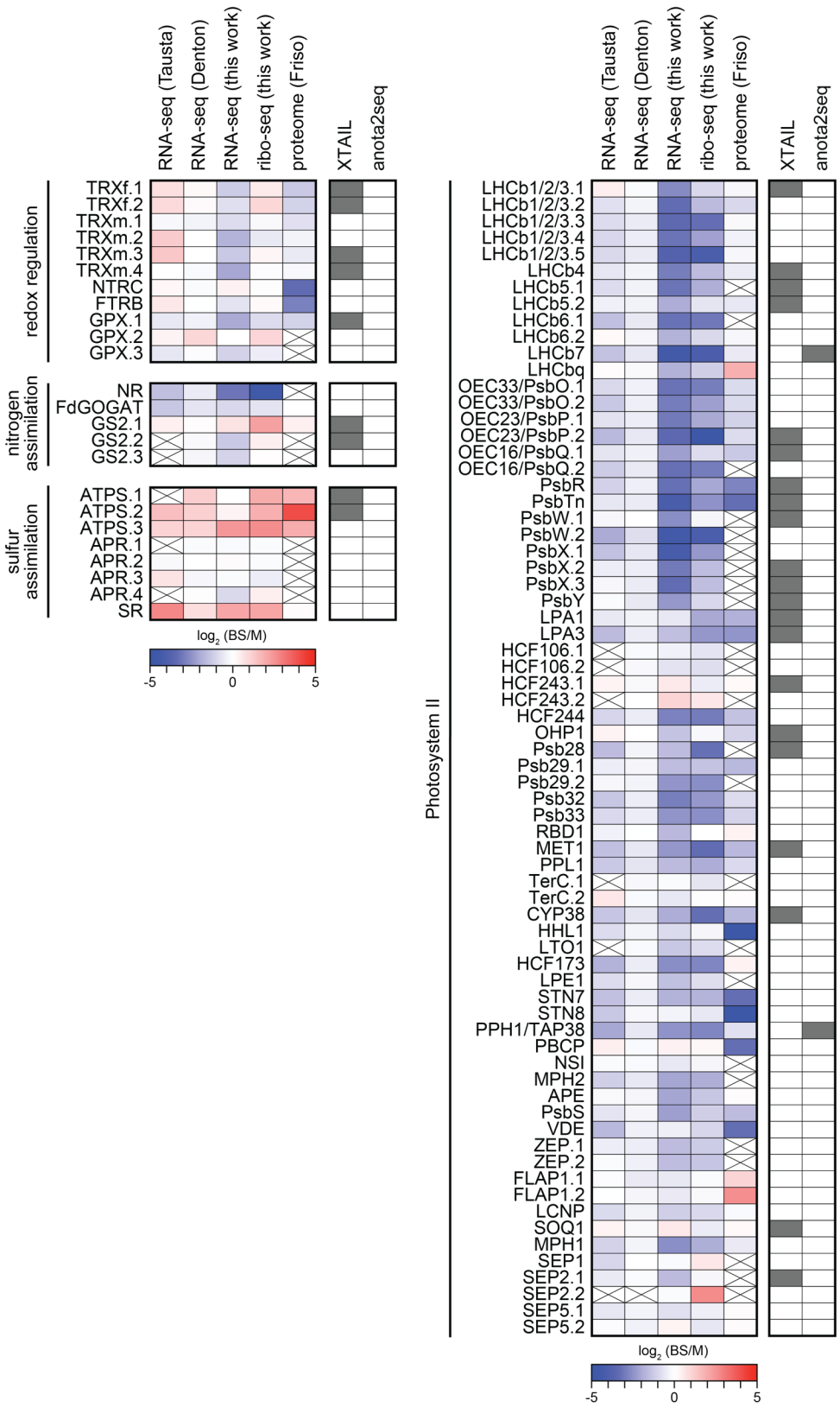


Supplemental Figure 3. Comparison of RNA-seq, ribo-seq, and proteome data collected for maize BS and M leaf fractions. Each symbol represents one gene. n, number of genes in the set; r, Pearson correlation.

(A) Comparison of our RNA-seq data to RNA-seq data from leaf section 14 in Tausta et al (2014) and slice 1 in Denton et al (2017). Data from the prior studies were re-analyzed with our pipeline prior to making these comparisons. C_4 genes are a set of genes involved in C_4 metabolism, photophosphorylation and Calvin-Benson cycle (Schlüter and Weber, 2020), and are the same genes shown in Figure 3.

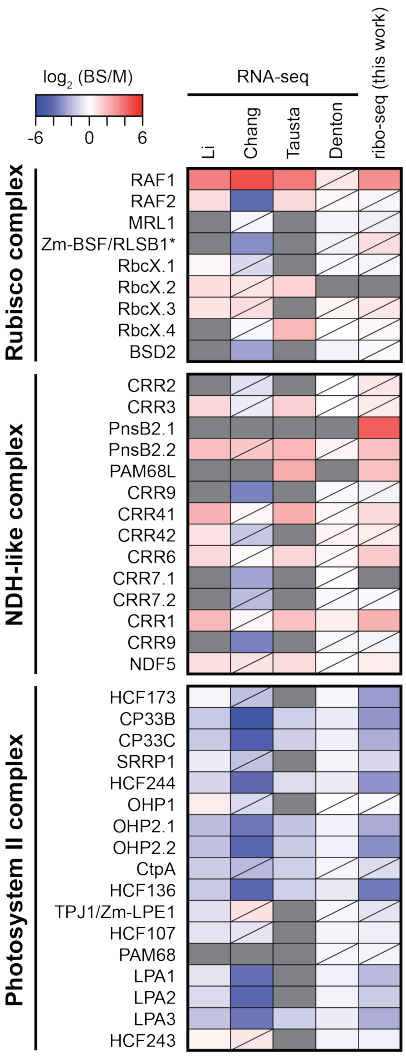
(B) Comparison of our ribo-seq and RNA-seq data.

(C) Comparison of our RNA-seq and ribo-seq data with proteome data from Friso et al (2010). The proteins identified in the proteome analysis were matched to the corresponding gene ID in maize genome annotation version APGv4. Only genes with unambiguous 1-to-1 matches between the two genome versions are shown. Proteins identified with high, medium, or low confidence in the proteome study are labeled.



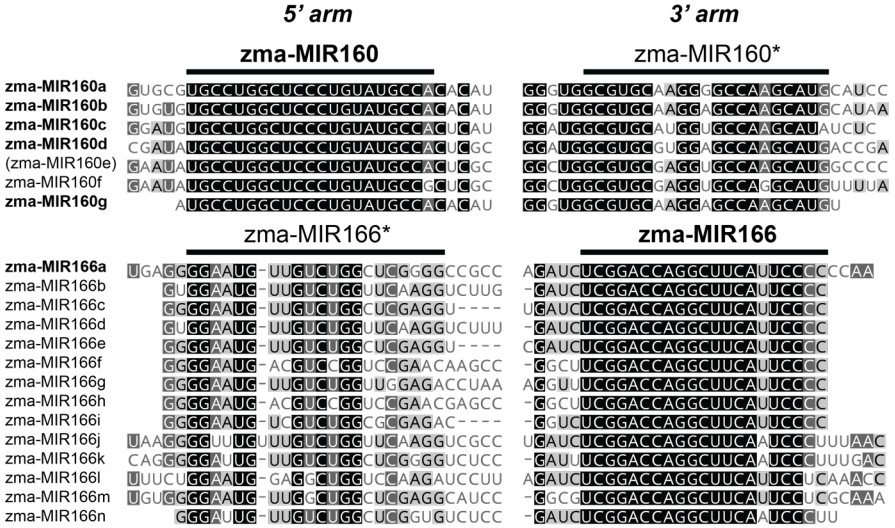
Supplemental Figure 4. Summary of expression data for C_4-related genes involved in redox regulation, nitrogen assimilation, sulfur assimilation and Photosystem II function.

RNA-seq and ribo-seq data reported here are compared with RNA-seq and proteome data from prior studies, as explained in Supplementary Figure 3. The previous RNA-seq data were re-analyzed with our pipeline and the values of $\log_2$ (BS/M) were determined by DESeq2. Only genes meeting our read count cutoffs are shown. Heatmaps to the left show $\log_2$ (BS/M) of RNA-seq, ribo-seq and proteome data. Crossed boxes indicate read counts below our cutoff or absence of data. Genes reported as exhibiting differential translation by XTAIL or anota2seq are marked with gray boxes to the right. The genes involved in redox regulation, nitrogen assimilation and sulfur assimilation were selected from Schlueter & Weber (Schlüter and Weber, 2020). The Photosystem II genes are taken from MapMan X4.2 BIN 1.1.1 (photosystem II). The gene IDs, data and additional information are provided in Supplemental Data Set 3.



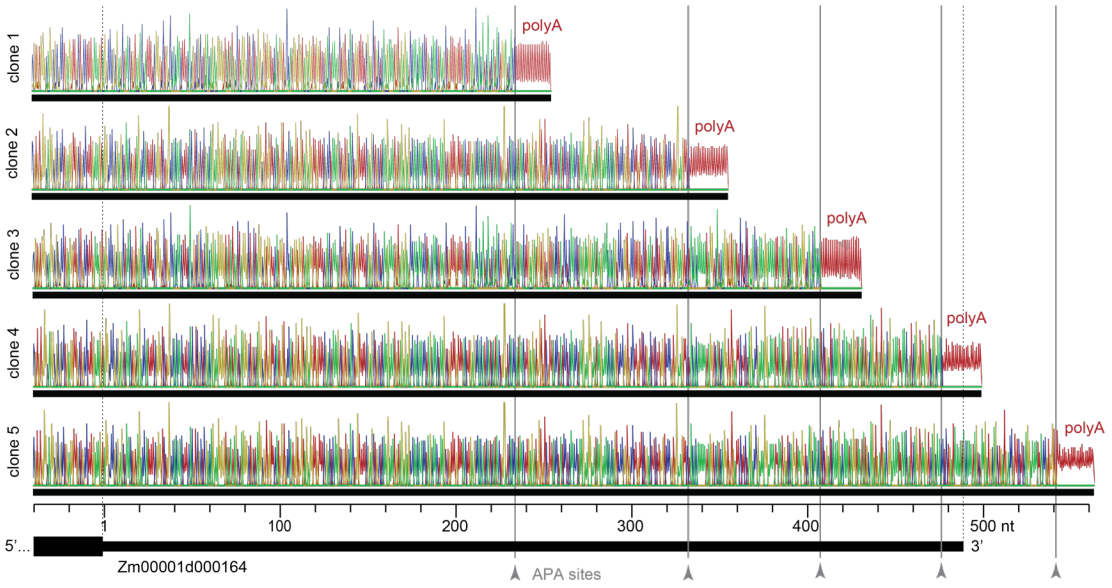
Supplemental Figure 5. Differential expression of biogenesis factors for PSII, Rubisco, and the NDH complex.

Grey boxes mark genes with missing values or for which the read counts are below cut-offs. Slashes mark genes that were not differentially expressed in Chang et al. (2012) or did not meet FDR cutoffs for differential expression (Denton FDR <0.05, ribo-seq FDR < 0.001). The gene IDs and data are provided in Supplemental Data Set 10. *RLSB1 had been reported to be involved specifically in *rbcL* translation (Bowman et al., 2013) but a subsequent report called that conclusion into question (Jiang et al., 2019). The gene was subsequently renamed Zm-BSF.



Supplemental Figure 6. Alignment of microRNA and microRNA* sequences of zma-MIR160 and zma-MIR166 families.

Each microRNA precursor folds into a hairpin structure by complementary base pairing of the 5' arm and 3' arm, followed by processing into an RNA duplex. One RNA strand in the duplex becomes mature microRNA and the other is typically degraded (marked with an asterisk). The information displayed was obtained from miRBase (http://www.mirbase.org/). The differentially expressed microRNAs in our BS and M data are shown with bold font. The differentially expressed zma-MIR160 could come from any of the five indicated genes because they all encode a microRNA of identical sequence. zma-MIR160e lacks a genome coordinate assignment and was not part of the DESeq2 analysis.



Supplemental Figure 7. Alternative polyadenylation sites of PRK.2 (Zm00001d000164) transcripts. Sequence of five clones from a 3' rapid amplification of cDNA ends (3'RACE) assay showed alternative polyadenylation (APA) sites of the RPK.2 gene. Sequencing chromatograms are aligned to the genomic region near the 3' end of the PPK.2 gene.