Supplementary Methods

Quantifying a transcriptome using RSEM (LI and DEWEY 2011) can lead to spurious read assignment and differences in transcript abundance when the transcriptome is incomplete or consists of a surplus of transcript models (NEWMAN *et al.* 2018). To verify the biological results of our differential expression (DE) analysis, we quantified the full B73 v4 reference transcriptome (46,430 genes) and the subset of one-to-one B73-Mo17 syntenic genes in the B73 v4 reference and Mo17 Cau (25,342 genes) reference using genome alignments and coverage over exonic regions associated with each gene. Using expression values from genome alignments, we found patterns of DE and associated biological functions were the same when compared to the RSEM quantification approach.

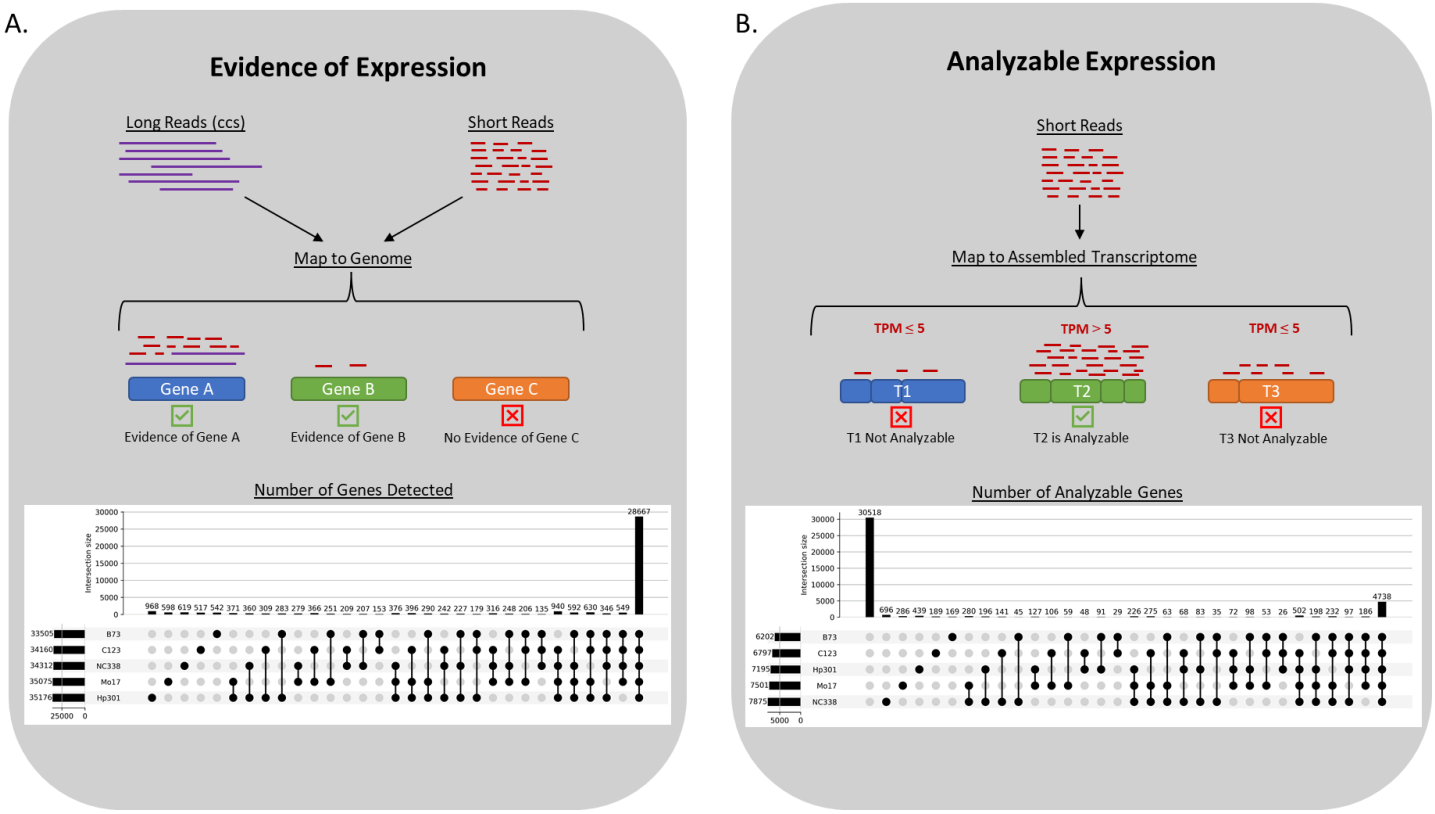
Quantification of the reference transcriptome: Paired-end Illumina short-read data was used for quantification of the B73 v4 reference transcriptome (Ensembl version 41, ftp://ftp.ensemblgenomes.org/pub/plants/release-41/fasta/zea_mays/) (JIAO *et al.* 2017) and Mo17 Cau reference transcriptome (SUN *et al.* 2018). FASTQ files were processed as described in the main text. Duplicate reads were removed and aligned to the each of the two genomes using BWA mem (version 0.7.17) (LI 2013). Uniquely mapped reads were used to quantify exonic regions of the reference transcriptome and the number of reads within each exonic region was estimated as $E_{rs} = (\sum D_{nrs})/L_r$ where D_{nrs} is the read depth for each nucleotide $n = 1, \dots, L_r$ of exonic region r in sample s, L_r is the length of exonic region r. TPM was calculated for each region $T_{rs} = \left((E_{rs})/K_s\right) * 10^6$, where K_s is $\sum((\sum D_{nis})/L_{is})$ for all exonic regions $i = 1, \dots m$ in sample s. Gene-level read counts were calculated by summing E_{rs} across all exonic regions r in gene g for sample s. Gene-level TPM was calculated using a weighted average of exonic region TPM values, $T_{gs} = \sum(T_{rs} * L_r)/\sum L_r$, for all exonic regions r in gene g for sample s. Genes were analyzable for DE if TPM values were greater than 5 in 50% of replicates in either ambient or elevated ozone condition.

Gene expression with B73 vs. Mo17 genome alignment: One-to-one B73-Mo17 Cau syntenic gene pairs were identified using SynMap (HAUG-BALTZELL *et al.* 2017) and SynFind within CoGe (LYONS and FREELING 2008) and compared to the 12,604 genes in the assembled leaf transcriptome resulting in 10,614 gene pairs. Comparisons of gene-level read counts from mapping to the B73 v4 and Mo17 Cau reference genomes were visualized within each sample (Supplementary Figure 12). Pearson correlation coefficients of read count comparisons demonstrate high correlation (0.92 – 0.94) between alignments to the different reference genomes. The percent difference between gene-level read counts when mapped to B73 and Mo17 is calculated as the absolute value of the difference over the average $(2 * |G_{B73} - G_{Mo17}|/(G_{B73} + G_{B73}))$. The majority of genes had approximately the same read counts with at most a 5% difference (Supplementary Table 7), while the remaining genes had mostly higher read counts when mapped to B73.

Evidence for expression: CCS long reads were mapped as described in the main text. CCS reads and adapter trimmed short reads were evaluated for evidence of expression in each sample separately. Evidence of expression for each gene in each sample was defined as at least 1 read (short or long) mapping within at least one exonic region of the gene (see figure below). Evidence

of expression in a genotype was defined by evidence of expression in either the ambient or elevated ozone condition.

Differential expression: For each genotype, genes with analyzable expression in either ambient or elevated ozone conditions were evaluated for DE using tappAS (version 1.0.6) (DE LA FUENTE *et al.* 2019) at FDR level (BENJAMINI and HOCHBERG 1995) of 0.05. To focus on gene-level DE, each analyzable gene from Ensembl version 41 was assigned an associated transcript from Ensembl version 34 in order to use the tappAS provided *Z. mays* annotation. GO enrichment using Fisher's exact test and an FDR threshold of 0.05 was performed (Supplementary Table 8, https://github.com/McIntyre-Lab/papers/tree/master/nanni_maize_2022). Basic trends of DE were exactly the same as the RSEM-based analysis, with B73 having limited DE (220 genes) and NC338 having the largest DE (2,243 genes). Of the genes DE in either analysis, ~48% were DE in both for B73, ~54% were DE in both for C123, ~63% were DE in both for Hp301, ~64% were DE in both for Mo17, and ~60% were DE in both for NC338. It is clear that there is short-read alignment ambiguity in assignment when a larger set of transcripts is included (SONESON *et al.* 2016; NEWMAN *et al.* 2018; TARDAGUILA *et al.* 2018).



B73 reference genes have evidence of expression. (A) Evidence of expression, or detection, is defined as the presence of a single long read (CCS) or short read mapping to the exonic regions of the gene. A gene is considered to have evidence for expression in a genotype if it is detected in either ambient or ozone conditions. The upset plot rows are the total number of genes with evidence of expression in each genotype. The columns indicate the unique combinations. The sum

of the columns is the number of reference genes with at least 1 transcript with evidence of expression (40,731). There are 12,604 genes without evidence of expression in all 5 genotypes. (B) Analyzable expression is defined as at least one transcript with TPM greater than 5 in at least 50% of RNA-seq replicates. The upset plot has the same format as in panel A. The majority of genes (75%) are not analyzable in any of the five genotypes.

GO enrichments are consistent between DE approaches.

Differentially expressed genes were evaluated for enrichment of Gene Ontology (GO) terms (Supplementary Table 8). Genes differentially expressed in all 5 genotypes were enriched for fewer GO categories than with the selected method of quantification: biological process (heat acclimation) and cellular component category (photosystem II). However, when focusing on genes differentially expressed in the more treatment-sensitive genotypes (C123, Hp301, Mo17, and NC338), GO categories enriched include biological processes related to metabolic processes (aromatic amino acid family biosynthetic process, coenzyme biosynthetic process, lipoate metabolic process, regulation of lipid metabolic process, unsaturated fatty acid biosynthetic process, vitamin metabolic process), photosynthesis (light harvesting in photosystem I, light reaction, photosynthetic electron transport chain, photosynthetic electron transport in photosystem I, photosystem I assembly, reductive pentose-phosphate cycle, response to low light intensity stimulus, response to red or far red light), protein-chromophore linkage, iron-sulfur cluster assembly, detection of biotic stimulus, and regulation of response to biotic stimulus. While the genes determined to be differentially expressed within each genotype were different between the methods of quantification, the GO results were nearly identical, resulting in the same biological interpretations regardless of detection method used.

References

Benjamini, Y., and Y. Hochberg, 1995 Controlling the false discovery rate - a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society Series B-Statistical Methodology 57**:** 289-300.

de la Fuente, L., Á. Arzalluz-Luque, M. Tardáguila, H. del Risco, C. Martí *et al.*, 2019 tappAS: a comprehensive computational framework for the analysis of the functional impact of differential splicing. bioRxiv**:** 690743.

Haug-Baltzell, A., S. Stephens, S. Davey, C. Scheidegger and E. Lyons, 2017 SynMap2 and SynMap3D: web-based whole-genome synteny browsers. Bioinformatics 33**:** 2197-2198.

Jiao, Y., P. Peluso, J. Shi, T. Liang, M. Stitzer *et al.*, 2017 Improved maize reference genome with single-molecule technologies. Nature 546**:** 524-+.

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

Li, H., 2013 Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM, pp. arXiv:1303.3997.

Lyons, E., and M. Freeling, 2008 How to usefully compare homologous plant genes and chromosomes as DNA sequences. Plant Journal 53**:** 661-673.

Newman, J., P. Concannon, M. Tardaguila, A. Conesa and L. McIntyre, 2018 Event Analysis: Using Transcript Events To Improve Estimates of Abundance in RNA-seq Data. G3-Genes Genomes Genetics 8**:** 2923-2940.

Soneson, C., K. L. Matthes, M. Nowicka, C. W. Law and M. D. Robinson, 2016 Isoform prefiltering improves performance of count-based methods for analysis of differential transcript usage. Genome Biology 17.

Sun, S., Y. Zhou, J. Chen, J. Shi, H. Zhao *et al.*, 2018 Extensive intraspecific gene order and gene structural variations between Mo17 and other maize genomes. Nature Genetics 50**:** 1289-+.

Tardaguila, M., L. de la Fuente, C. Marti, C. Pereira, F. J. Pardo-Palacios *et al.*, 2018 SQANTI: extensive characterization of long-read transcript sequences for quality control in full-length transcriptome identification and quantification. Genome Res.