Title: Variation in leaf transcriptome responses to elevated ozone corresponds with physiological sensitivity to ozone across maize inbred lines

Adalena V. Nanni[1,4], Alison M. Morse[1,4], Jeremy R. B. Newman[3,4], Nicole E. Choquette[5], Jessica M. Wedow[5], Zihao Liu[1,4] Andrew D. B. Leakey[5], Ana Conesa[2,7], Elizabeth A. Ainsworth[5,6], Lauren M McIntyre*[1,4]

[1]Department of Molecular Genetics and Microbiology, University of Florida, Gainesville Florida 32611

[2]Department of Cell and Microbial Sciences, University of Florida, Gainesville Florida 32611

[3]Department of Pathology, University of Florida, Gainesville Florida 32611

[4]Genetics Institute, University of Florida, Gainesville Florida 32611

[5]Carl R. Woese Institute for Genomic Biology, Department of Plant Biology, and Department of Crop Sciences, University of Illinois at Urbana-Champaign

[6]USDA ARS Global Change and Photosynthesis Research Unit, Urbana, IL

[7]Institute for Integrative Systems Biology, Spanish National Research Council, Paterna, Spain.

* Lauren M. McIntyre corresponding author

Email: mcintyre@ufl.edu

ORCID:

0000-0003-4286-6489 Alison M Morse; 0000-0002-0077-3359 – Lauren M McIntyre; 0000-0001-8486-2801- Jeremy Newman; 0000-0001-5410-9026 Jessica Wedow; 0000-0002-3199-8999. Elizabeth Ainsworth; 0000-0001-6251-024X Andrew Leakey; 0000-0001-9597-311X Ana Conesa

none of the authors have any competing interests

Abstract

We examine the impact of sustained elevated ozone concentration on the leaf transcriptome of 5 diverse maize inbred genotypes, which vary in physiological sensitivity to ozone (B73, Mo17, Hp301, C123, NC338), using long reads to assemble transcripts and short reads to quantify expression of these transcripts. More than 99% of the long reads, 99% of the assembled transcripts, and 97% of the short reads map to both B73 and Mo17 reference genomes. Approximately 95% of the genes with assembled transcripts belong to known B73-Mo17 syntenic loci and 94% of genes with assembled transcripts are present in all temperate lines in the NAM pan-genome. While there is limited evidence for alternative splicing in response to ozone stress, there is a difference in the magnitude of differential expression among the 5 genotypes. The transcriptional response to sustained ozone stress in the ozone resistant B73 genotype (151 genes) was modest, while more than 3,300 genes were significantly differentially expressed in the more sensitive NC338 genotype. There is the potential for tandem duplication in 30% of genes with assembled transcripts, but there is no obvious association between potential tandem duplication and differential expression. Genes with a common response across the 5 genotypes (83 genes) were associated with photosynthesis, in particular photosystem I. The functional annotation of genes not differentially expressed in B73 but responsive in the other 4 genotypes (789) identifies reactive oxygen species. This suggests that B73 has a different response to long term ozone exposure than the other 4 genotypes. The relative magnitude of the genotypic response to ozone, and the enrichment analyses are consistent regardless of whether aligning short reads to: long read assembled transcripts; the B73 reference; the Mo17 reference. We find that prolonged ozone exposure directly impacts the photosynthetic machinery of the leaf.

Keywords

corn, stress response, climate change, Iso-seq, RNA-seq

Background

Tropospheric ozone is a prevalent, phytotoxic air pollutant, which significantly reduces global crop production (BURNEY and RAMANATHAN 2014; MCGRATH *et al.* 2015; MILLS *et al.* 2018). In the U.S., ozone pollution over the past 30 years was estimated to decrease maize yields by 10% (MCGRATH *et al.* 2015); a loss equivalent to the impact of aridity stress, temperature stress or nutrient stress (MILLS *et al.* 2018). Exposure of plants to elevated ozone concentration causes reactive oxygen species (ROS) production in cells and knowledge of the responses this elicits is expected to aid understanding of oxidative stress associated with other abiotic and biotic factors (CONKLIN and BARTH 2004; GADJEV *et al.* 2006). Loss of photosynthetic capacity associated with accelerated senescence is a key response of maize to elevated ozone and differs among genotypes (YENDREK *et al.* 2017a). Ozone stress increased the heritability of photosynthetic traits and strengthened genetic correlations among traits in maize with substantial differences among hybrids observed (CHOQUETTE *et al.* 2019).

In a range of species, presence/absence variation (PAV) of genes related to abiotic stress response have been observed (YAO *et al.* 2015; GOLICZ *et al.* 2016; JIN *et al.* 2016; MONTENEGRO *et al.* 2017; WANG *et al.* 2018b; ZHAO *et al.* 2018; GAO *et al.* 2019; HOOPES *et al.* 2019; LIU *et al.* 2020). In Amborella, PAV of genes have been linked primarily to responses to abiotic stress such as salt, water deprivation, and heat (HU *et al.* 2021). Gene families that have observed PAV associated with abiotic stress response include late embryonic abundant (LEA) genes (GOYAL *et al.* 2005; JIN *et al.* 2016), ROS signaling genes (MONTENEGRO *et al.* 2017), and genes related to the stay-green trait (QIAN *et al.* 2016). Alternative splicing (AS) has been reported to be associated with environmental stress response (MARRS and WALBOT 1997; EGAWA *et al.* 2006; LI *et al.* 2013; DING *et al.* 2014; CALIXTO *et al.* 2018). Divergence of AS, due to sequence divergence in splice sites or motifs, has been associated with stress response genes in Arabidopsis (WANG *et al.* 2019).

This study investigated the transcriptional response of maize to ozone stress in 5 diverse genotypes (B73, Mo17, Hp301, C123, NC338). Differences in a transcriptional response to ozone may be a consequence of PAV, differential splicing, and differential expression (DE). Large portions of grass genomes are syntenic (HULBERT *et al.* 1990; AHN and TANKSLEY 1993; MOORE *et al.* 1995; GALE and DEVOS 1998; FEUILLET and KELLER 2002). A recent *de novo*

assembly of the maize Mo17 genotype (SUN *et al.* 2018) and corresponding in-depth comparison to reference B73 genotype found only 122 genes present exclusively in one of the two genomes, and another pairwise comparison of B73 to a European inbred found ~400 genes were not shared (DARRACQ *et al.* 2018). Despite macrocollinearity, conservation of genes and gene order, there is evidence for genic rearrangements (reviewed in BENNETZEN 2005). Within maize, there is support for a pan-genome with documented divergence among genomes (SCHNABLE *et al.* 2009; LAI *et al.* 2010; HIRSCH *et al.* 2014), a megabase deletion (HAN *et al.* 2019), pervasive presence/absence variation (PAV) (SPRINGER *et al.* 2009; SWANSON-WAGNER *et al.* 2010), and exon shuffling (DOONER and WEIL 2007). In addition, a study of 503 inbred lines suggested that a "substantial portion of variation may lie outside the single reference genome for the species" (HIRSCH *et al.* 2014). A recent comprehensive evaluation of 26 maize genotypes (parents of the Nested Association Mapping Panel, NAM (YU *et al.* 2008)) finds evidence of a pan-genome with more than 100,000 genes and only approximately 30% shared among all 26 genotypes. B73 is reported to contain between 63% and 74% of all genes in the pan-genome (HUFFORD *et al.* 2021). At the heart of these discrepancies are varied views on how to account for nearly identical paralogs, tandem duplicates, and genes that have been duplicated through transposition. Whatever the nomenclature, it is important to consider that there are potential differences in dosage of highly similar transcripts and proteins between genotypes.

Differential expression in response to abiotic stress occurs in many plant species (reviewed in FUJITA *et al.* 2006), including maize (reviewed in SALIKA and RIFFAT 2021). Changes to expression following abiotic stresses related to climate change (e.g., heat, drought, elevated greenhouse gases, etc.) have also been observed (reviewed in AHUJA *et al.* 2010). Differential expression must be due to variation in cis-regulatory elements either directly affecting expression or indirectly leading to downstream effects (reviewed in WITTKOPP 2007). Cis-regulation induced by abiotic stress can include factors such as transposable element (e.g., ITO *et al.* 2013; MAKAREVITCH *et al.* 2015), epigenetic responses (e.g., VERHOEVEN *et al.* 2018; ENTRAMBASAGUAS *et al.* 2021; reviewed in KIM 2021) or other cis-regulatory factors (e.g., MARUYAMA *et al.* 2004; ZOU *et al.* 2011; RICCI *et al.* 2019).

Maize genotypes exist that show important phenotypic differences, including response to stress and ozone susceptibility (HULBERT *et al.* 1990; YENDREK *et al.* 2017a; AINSWORTH *et al.* 2020;

CHOQUETTE *et al.* 2020; WEDOW *et al.* 2021a). We seek to answer what the basis of the observed differences in ozone response among maize cultivars is, such as differences in the gene content (PAV), differences in splicing (AS), or differences in transcriptional regulation (DE)? We sequenced the leaf transcriptome of five diverse maize genotypes, B73, Mo17, C123, NC338 and Hp301, in ambient and elevated ozone conditions. While C123 and Mo17 are somewhat closely related, the other genotypes represent a diverse array of temperate genotypes including Hp301, a popcorn genotype (See Supplementary Figure 1 in LIU *et al.* 2003). These five genotypes are selected based on variation in physiological response to ozone in field experiments (CHOQUETTE *et al.* 2019) where B73 was the least sensitive of the 5 genotypes. We focus on the following questions: Is the expressed transcriptome part of the shared genome? Is there evidence for alternative splicing in response to stress? Is the transcriptional response to ozone similar among all 5 temperate genotypes? We used long read sequencing and an unbiased approach to assemble transcripts for each genotype and treatment (n=10) independently. These assembled transcripts were rigorously examined for diversity and then the union of the assembled transcriptomes was used as a transcriptional reference for quantitation of short read data from an experiment with twelve independent replicate plants for each genotype/condition (n=120).

Results

Do the leaf transcriptomes of different maize lines represent the shared genome, or, on the contrary, reveal genotype specific genes?

The majority of clusters are high quality with few technical artifacts. We use IsoSeq3, which does not require a reference genome (https://github.com/PacificBiosciences/IsoSeq3), to obtain high-quality polished clusters for each genotype/condition. Clustering is an intermediate step in the transcript assembly process (see methods). Clusters from each genotype/condition are mapped to the three available reference genomes (Figure 1, Supplementary Table 1, line 32). Despite genomic divergence (SUN *et al.* 2018), the cluster mapping rates for clusters from the Hp301, C123 and NC338 samples are similar to the B73 and Mo17 samples (Figure 1B) with 92.8-99.5% mapping at 95% identity or higher and at least 95% of the read length when mapped to the B73 and Mo17 references (Figure 1B, Supplementary Table 1, line 35).

The relatively few unmapped clusters are likely technical artifacts. There are ***12*** unmapped clusters in the B73 set of high-quality polished clusters mapped to the B73 genome: 5 in the ambient condition and 7 in the ozone condition. Only 2 of these 12 are greater than 1 kb in length. These are further examined by a BLAST (ALTSCHUL *et al.* 1990) against the B73 reference genome. Three of the 12 B73 clusters hit to the same small (61bp) region of the B73 genome while the remaining 9 sequences did not return a hit. These results indicate that these 12 unmapped clusters from B73 likely reflect technical artifacts. Similarly, for the 3 Mo17 samples, 50 high quality polished clusters are unmapped to the Mo17 Cau reference (14 in the ambient ozone library and 36 in the 2 elevated ozone libraries) (Supplementary Table 1, line 32). Similar to B73, BLAST results of partial hits from Mo17 clusters to the Mo17 genome contain mismatches and gaps. Taken together, this indicates that less than 1% of the polished clusters for these 4 samples are likely technical artifacts. BLAST results for unmapped clusters from Hp301, C123 and NC338 have similar results (https://github.com/McIntyre-Lab/papers/tree/master/nanni_maize_2022).

Potential copy number variation. We note that there are larger numbers of 'ignored' clusters (clusters that map but have low identity or low coverage) for Mo17, C123, NC338, and Hp301 mapping to B73 and for B73, C123, NC338, and Hp301 mapping to Mo17 compared to B73 mapping to B73 or Mo17 Mapping to Mo17(Supplementary Table 1 line 37). Some of these clusters are likely incomplete transcripts. Clusters less than 1000 nt (20%-40% of those ignored) suggest potential incomplete transcript sequences. Most of the low identity and low coverage clusters with at least 1,000 nt have BLAST hits to maize with mismatches and gaps (83-90%) (Supplementary Table 1 line 39, https://github.com/McIntyre-Lab/papers/tree/master/nanni_maize_2022) potentially suggesting genetic variation. The B73 clusters mapping with low identity to Mo17, but not B73, were examined to provide potential insight into the excess of clusters of low coverage in the Hp301, C123 and NC338 genomes. These clusters overall map well to B73 and have a similar length distribution as the clusters that map to both B73 and Mo17. We note that several of the best hits are to gene families (Supplementary Table 1 line 40) so we cannot preclude the possibility that there is ambiguity in assignment of clusters to genes due to variants among the genotypes that have diverged or the possibility that B73 and Mo17 have unique duplications for some of these

genes (tandem or dispersed). For completeness all cluster sequences are provided (https://github.com/McIntyre-Lab/papers/tree/master/nanni_maize_2022).

Assembled transcripts are in core or near-core genes when compared to the NAM pan-genome. We assemble transcripts from mapped clusters using Cupcake ToFU (https://github.com/Magdoll/cDNA_Cupcake). We match the junctions in sequential order (junction chain) of the resulting assembled transcripts to the existing reference annotations for B73 or Mo17 using SQANTI QC (TARDAGUILA *et al.* 2018). The B73 and Mo17 Yan references contain junction chains that match the majority of the assembled transcripts (Supplementary Figure 2B, Supplementary Table 1, lines 49 and 50). The short reads from all genotypes also map to both B73 and Mo17 genomes with only 1% to 3% unmapped (Supplementary Table 2).

A unique list of B73 v4 reference genes (n=15,055) representing the genes associated with the assembled transcripts across all samples (Supplementary Table 1, line 46) is compared to genes in the NAM pan-genome (HUFFORD *et al.* 2021). In this list there are 13,369 genes associated with a NAM pan-gene with 90.6% annotated as core (present in all parents of the NAM). An additional ~6.1% are near-core (present in all but 1 of the 26 NAM founders) and ~94.4% are present in all temperate NAM genotypes (HUFFORD *et al.* 2021). Of the 13,369 reference genes with associated assembled transcripts, 29 are annotated in the NAM population as private to B73 (present in only NAM B73). Of these 29 genes, 12 are expressed only in the B73 samples (Supplementary Table 4) and the other 17 genes are detected in our experiments of temperate lines not in the NAM collection.

Assembled transcriptome is 90%-94% B73-Mo17 syntenic. The high mapping rates to both reference genomes led us to hypothesize that the genes expressed in the leaf are part of the syntenic genome in B73-Mo17. To determine whether the assembled transcripts are from syntenic loci in B73 and Mo17, we evaluate B73-Mo17 synteny information using SynMap (HAUG-BALTZELL *et al.* 2017) and SynFind within CoGe (LYONS and FREELING 2008). Approximately 90 percent of the assembled transcripts from B73 and Mo17 map to B73-Mo17 syntenic loci. Using the syntenic list of B73-Mo17 loci from Sun et al. (SUN *et al.* 2018) shows a similar result (~94% mapping of assembled transcripts to syntenic loci).

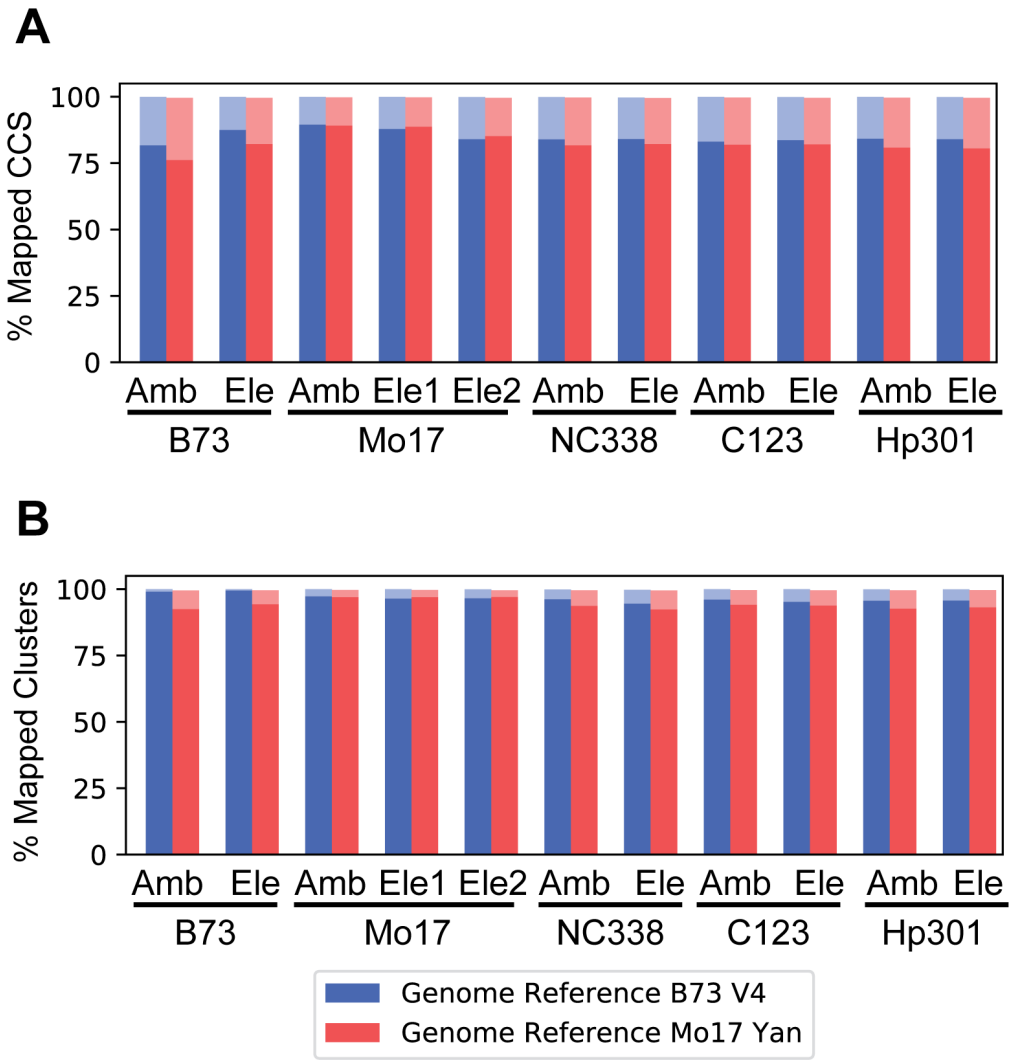

Figure 1. A) CCS long reads and B) Clusters map equally well to the B73 v4 or Mo17 Yan reference genomes for all 5 genotypes. The darker color (dark blue or dark red) indicates percent reads (A) or clusters (B) with alignment identity greater than or equal to 95% to the indicated reference sequence and alignment length greater than or equal to 95% of the read or cluster length. Percent reads or clusters with alignment identity less than 95% and/or alignment length less than 95% of the read or cluster length is indicated by a light blue or light red color. The percent mapped CCS long reads ranges from 99.4% to 99.9% and percent mapped clusters range from 99.5% to 99.99% (Supplementary Table 1). Mapping to annotated reference transcripts is also similar among all 5 genotypes (Supplementary Figure 3).

Assembled transcripts that map to intergenic regions ('novel' in SQANTI) or to more than one genic locus ('fusion' in SQANTI) are expressed in multiple genotypes. To determine whether

these loci are detected in the temperate genotypes in this study, we look at both short and long reads. Out of the 1,575 novel, fusion, or antisense loci, 928 are detected in all genotypes and 1,335 are detected in multiple genotypes (Supplementary Figure 4A, B). There are 3 novel loci expressed exclusively in B73. BLAST results of the transcripts within these loci indicate partial hits to B73, Mo17, and other maize genomes including putative transposase and retrotransposon genes (https://github.com/McIntyre-Lab/papers/tree/master/nanni_maize_2022).

There is potential ambiguity in gene identity among putative tandem duplicates and putative identical paralogs. Hoopes et al. (HOOPES *et al.* 2019) identified paralogous groups of genes. For the 12,604 annotated genes in the assembled leaf transcriptome, there are 1,032 paralogous groups containing 2,390 genes where each paralogous group contains a minimum of 2 genes. This demonstrates that transcripts are assembled separately for paralogs with sequence diversity. Approximately ~1% of maize genes have nearly identical paralogs (EMRICH *et al.* 2007), meaning they exhibit 99% similarity. In this case, long reads originating from nearly identical paralogs would not be separable. This is a problem with RNA-seq experiments in maize and is a general limitation in such studies. To further examine this issue, we use the annotation of putative tandem duplicates among the parents of the NAM population (HUFFORD *et al.* 2021). We find that 30% of the 12,604 annotated genes in this study have a potential tandem duplication in at least one of the parents of the NAM. This is an enrichment compared to the 16% of all genes reported as potential tandem duplicates in the NAM pan-genome (HUFFORD *et al.* 2021). The high number of potential duplicates and the ambiguity among paralogs is a limitation of this study and all studies of expression in this species (Supplementary File 5).

There is limited evidence for transcriptional presence/absence variation in this study. There is a distinction between evidence for expression and being able to conduct an analysis for differential expression (Figure 2). Differences in transcriptional presence/absence variation among maize inbred lines has been observed (BROHAMMER *et al.* 2018; HOOPES *et al.* 2019) and it has been hypothesized that presence/absence of transcription contributes to stress responses in maize (HOOPES *et al.* 2019). For genes with assembled transcripts (n=12,604), we have evidence of expression in all five genotypes for all but 81 genes (Figure 2A). Of these 81 annotated genes, 40 were previously identified as genes with potential PAV (BROHAMMER *et al.* 2018; HOOPES *et al.*

2019). It is possible that transcriptional PAV is present in these genes. These genes can be found in Supplementary File 5 (where *flag_detect_num_geno_gt0_lt5* = 1 and *flag_pav* = 1).

The ability to analyze the 12,604 annotated genes for expression is more variable across genotypes (Figure 2B). The majority of the genes (8951, 71%) are analyzable (TPM greater than 5 in at least 50% of replicates) in all 5 genotypes. Only 549 are analyzable in only one genotype. Expression is lower for the genes with fewer analyzable genotypes (Figure 3). Of the 549 genes analyzable in only one genotype, 473 are detected in all 5 genotypes. This finding is likely due to sampling variation (EMRICH *et al.* 2007).

Our long and short read data map to B73 and Mo17 references with ~94% of the genes previously identified as B73-Mo17 syntenic. There is some limited evidence for genes expressed in the leaf in C123, Hp301, or NC338 in either ambient or elevated ozone that are exclusive to those genotypes.

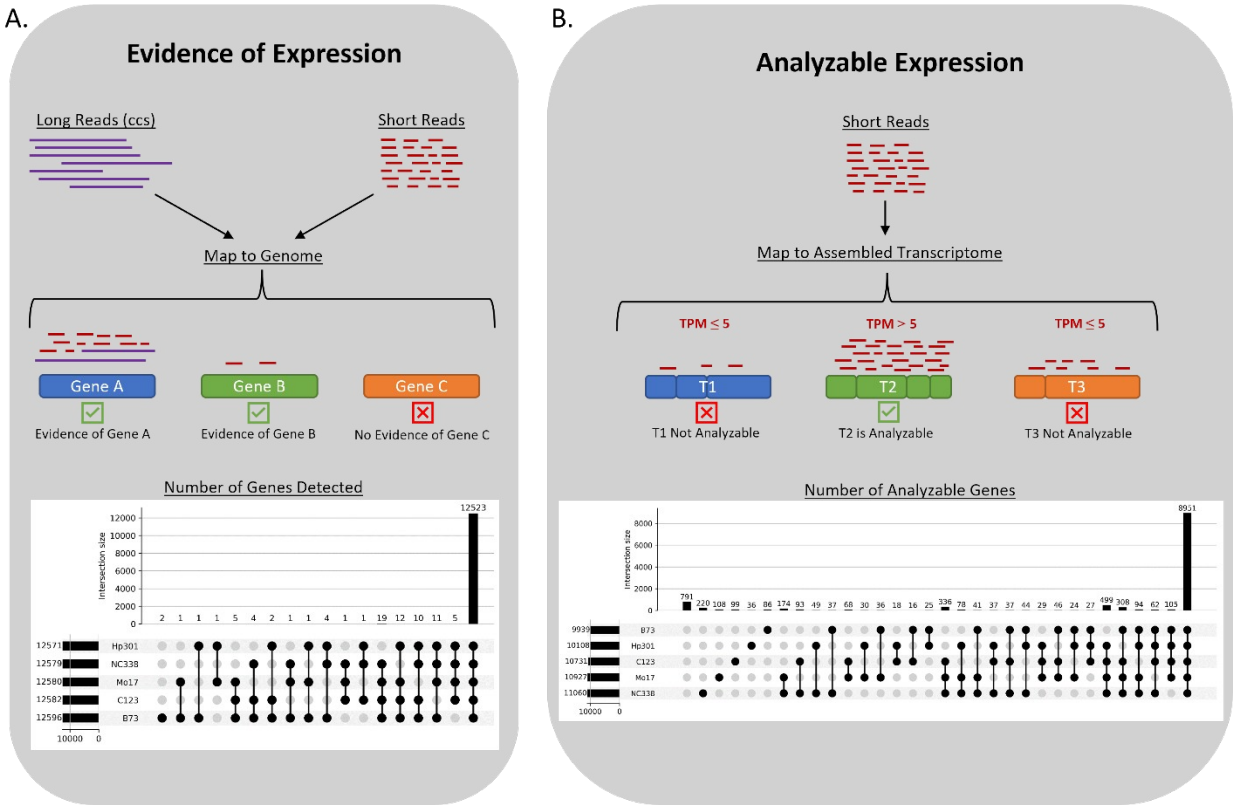

Figure 2. Genes with assembled transcripts have evidence of expression and are analyzable. (A) Evidence of expression, or detection, is defined as the presence of a single CCS long read or short read mapping to exonic gene regions. A gene is considered to have evidence for expression

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 genes with at least 1 assembled transcript (12,604). There are only 81 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 (77%) are analyzable in all five genotypes.

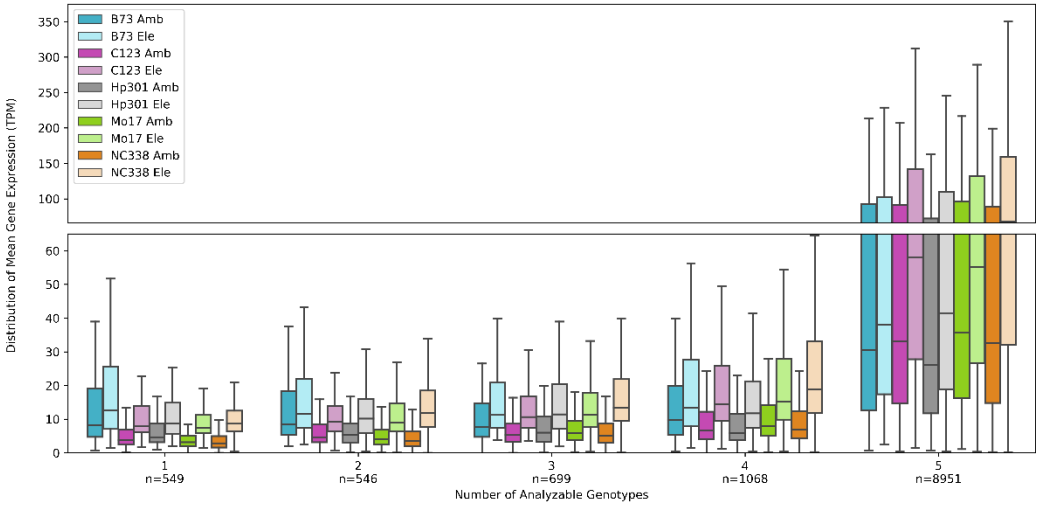

Figure 3. Expression as a function of the number of genotypes analyzable for differential expression. Mean gene expression (transcripts per million, TPM) for each genotype and treatment. B73 is blue, C123 is red, Hp301 is gray, Mo17 is green and NC338 is orange. Genes are analyzable if at least one transcript has a TPM greater than 5 in 50% of short-read replicates in either ambient or elevated ozone condition for that genotype. The number of genes for each group analyzable in 1 to 5 genotypes is indicated below each set of representative boxplots, totaling 11,813 genes. Most genes are analyzable in all 5 genotypes (8,951). For genes analyzable in fewer than 5 genotypes, expression values are relatively low. This indicates that it is unlikely that presence/absence variation is at play, and instead more likely that presence/absence is due to sampling variation.

Is there evidence for alternative splicing in response to stress?

The majority of genes in the assembled leaf transcriptome have a single transcript. The union of the individually assembled transcripts from all genotypes and conditions are combined into a single unique set of assembled transcripts using chaining (https://github.com/Magdoll/cDNA_Cupcake). The assembled transcripts are annotated using the B73 v4 (JIAO *et al.* 2017) reference annotation with SQANTI (TARDAGUILA *et al.* 2018)

(Supplementary Table 1, lines 62-71). After chaining there are 1,531 putative novel loci (187 of which are novel antisense loci of annotated genes). To remove redundancies, transcripts with the same splice junctions are consolidated and unspliced fragments in annotated genes with multi-exon reference genes are removed. Since transcripts with identical splice junction chains cannot be easily distinguished with RNA-seq, we selected the longest representative of each unique splicing pattern in the curated transcriptome. The curated assembled transcriptome has 31,388 assembled transcripts and 12,604 genes More than half (6,895 genes, ~54%) of the annotated genes have a single associated assembled transcript, and ~17% have more than 3 associated transcripts (Supplementary Figure 5B). The average number of transcripts per annotated gene is ~2.5 and the median is 1. The 9,451 genes in Wang et al. (WANG *et al.* 2018a) (~550,00 starting CCS reads in the leaf tissue dataset) represent 21% of all B73 v4 reference genes (Supplementary Figure 6A). Of these genes, 9% have 3 or more transcripts and 77% have one transcript. Starting with roughly 6 times the number of CCS reads as Wang et al. (WANG *et al.* 2018a), we find approximately twice as many assembled transcripts (31,388 vs. 13,580). Other tissues also showed a median of 1 transcript per gene (Supplementary Figure 6B-F).

Alternative splicing is not a major contributor to the maize leaf response to ozone stress. We estimate gene expression in alternative transcripts by mapping Illumina short read data from 120 individual plants (n=12 per experimental condition; n=24 per genotype) to the assembled leaf transcriptome using RSEM (LI and DEWEY 2011). We then evaluate differential isoform usage in the transcriptional response to ozone stress using tappAS (DE LA FUENTE *et al.* 2019). We assayed all genes with multiple transcripts in a particular genotype for differential isoform usage between ambient and elevated ozone conditions (Supplementary Figure 7). Despite reports of splicing as a response to stress in plants (IIDA *et al.* 2004; FILICHKIN *et al.* 2010; THATCHER *et al.* 2016; reviewed in LALOUM *et al.* 2018), only 59 annotated genes are identified as having major isoform switching in NC338 (< 2% of multi-transcript genes detected in NC338), and fewer than 35 genes (~1% or less of all multi-transcript genes) in each of the other 4 genotypes, with 1 gene (Zm00001d043613) having evidence for major isoform switching in B73 (Supplementary Table 5). Eighteen annotated genes are identified as having differential isoform usage with annotations linked to plant stress responses.

We note that there are quite a few annotated genes with differential isoform usage where the predicted coding regions of the transcripts are non-overlapping, leading to potentially distinct protein sequences. An example is shown in Supplementary Figure 8. This suggests that ozone stress response may differentially regulate the proteins produced by these genes, or that there are two similar genes in proximity, annotated as a single gene, that respond differently to ozone stress.

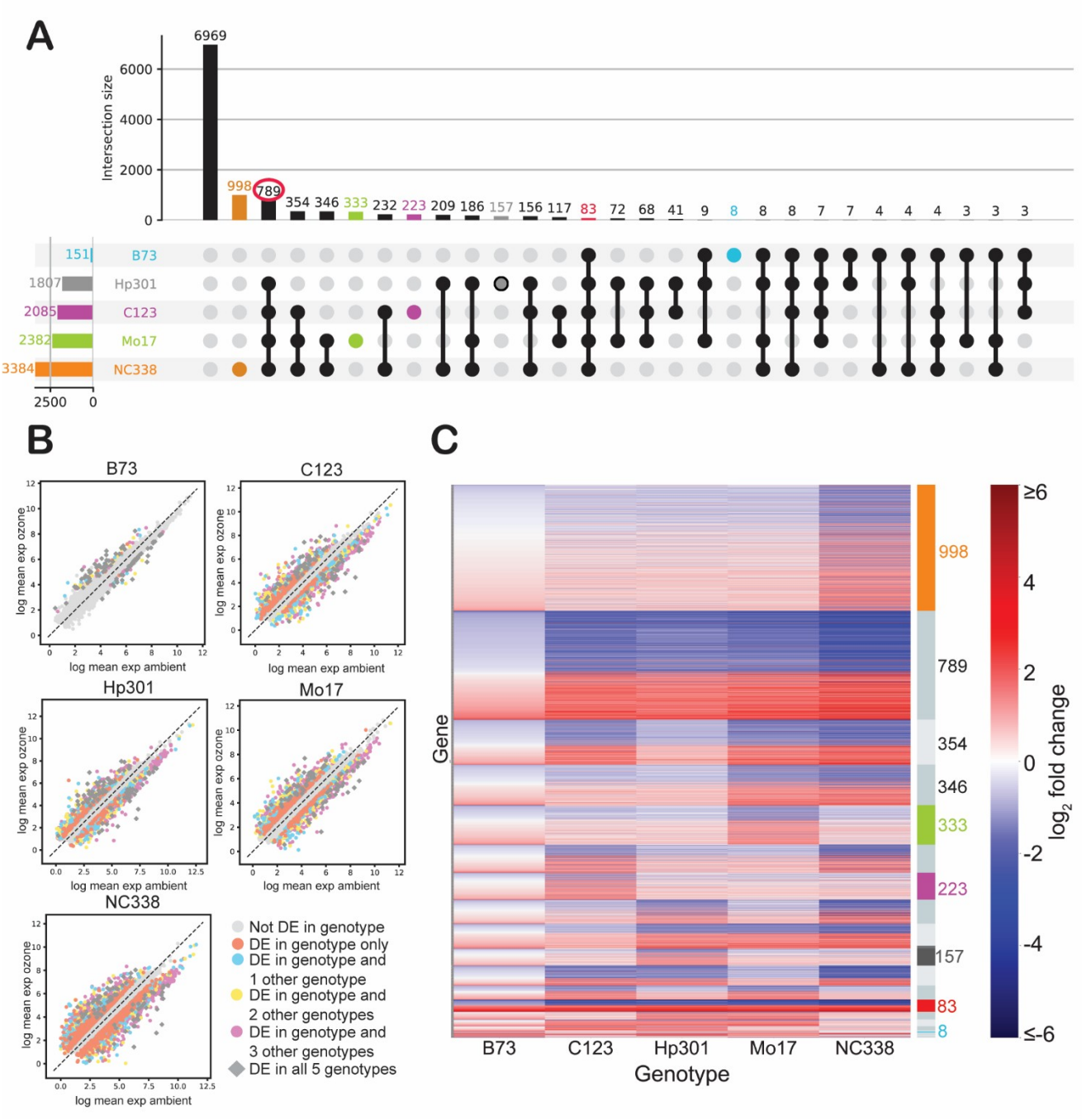

Figure 4. Differentially expressed genes. (A) Count of differential expression by genotype for all genes analyzed. Genes significantly differentially expressed in the labeled genotype left bar. The top histogram totals to the number of genes analyzed for differential expression (11,401) and the categories are mutually exclusive. For example, genes differentially expressed in NC338 only (998, orange). Relatively few genes are differentially expressed in all genotypes (83, red circle). More are differentially expressed in the 4 genotypes M017, C123, Hp301 and NC338. (B) For each genotype, a scatterplot is shown of the log2(TPM) ambient on the X axis and the log2(TPM) ozone on the Y axis. Differentially expressed (DE) genes in the indicated genotype are in pink, and in exactly one other genotype are blue, in two other genotypes are yellow, in 3 other genotypes are purple and in all 5 genotypes (gray diamonds). The largest fold changes are consistently apparent in those genes with the gray diamonds. Transcripts DE in only one genotype have smaller fold changes than those DE in more than one genotype (as seen by the orange points nearer the diagonal). (C) Genes significant for differential expression in at least one genotype (4,432). Rows are genes and columns are genotypes. Categories are ordered as in Panel A (matching color key is the right bar, numbers indicate the numbers of genes in each of the groups). Within each category in the right bar the rows are sorted by the B73 response, in order to provide a consistent reference point. The first (orange) group is significant for NC338 and the second group visualizes the 789 that are significant for the 4 genotypes M017, C123, Hp301 and NC338.

Is the transcriptional response to ozone similar among all 5 temperate genotypes?

Genotypes differ in the magnitude of the response to ozone stress. B73 shows a limited response to elevated ozone, with 151 differentially expressed genes, while NC338 shows evidence for more than 3,300 differentially expressed genes (Figure 4A). The number of differentially expressed genes is progressively greater in the other genotypes: Hp301 (1807 genes), C123 (2085 genes) and Mo17 (2382 genes). The number of genes differentially expressed in a genotype-specific manner followed the same pattern: B73 (8 genes), Hp301 (157 genes), C123 (223 genes), Mo17 (333 genes) and NC338 (998 genes) (Figure 4A). Across the four more ozone sensitive genotypes (non-B73) there are 789 differentially expressed genes. Eighty-three genes are identified as differentially expressed in all five genotypes (Supplementary File 4, flag_analyze_DE_all5), thus representing a common transcriptional response to elevated ozone concentration.

For genes significant in only one genotype, the estimated fold change is relatively modest, with the response to ozone in B73 substantially smaller in all comparisons (Figure 4B, 3C). In each genotype, enriched gene sets (SUBRAMANIAN *et al.* 2005) for the Gene Ontology biological

process category included photosynthesis (e. g., photosystem I, protein-chromophore linkages) (Supplementary Table 6, flag_analyze_DE_*genotype*).

Diverse genotypes show a common transcriptional response to ozone stress. Are differential responses to ozone stress associated with gene expression differences? The largest fold changes are found in genes significant in all genotypes (Figure 4B, Supplementary Figure 9,10), and the proportion of up- and down-regulated genes is balanced for all genotypes, including B73 and NC338 (Figure 4B, Supplementary Figure 11). The regulation of expression in response to ozone is in the same direction for 82 of the 83 transcripts differentially expressed in all 5 genotypes (Figure 4C, red). Zm00001d027525 (basic endochitinase B, EnsemblPlants) is the only transcript that is down-regulated in B73 under ozone conditions but up-regulated in the other 4 genotypes. Genes differentially expressed in all 5 genotypes are enriched for the following GO categories: biological process (heat shock response, photosynthesis (light harvesting in photosystem I, light reaction, response to red light, and protein-chromophore linkage)); cellular component category (photosystem I, photosystem II and plastoglobule); and molecular function category (chlorophyll binding and pigment binding) (Supplementary Table 5, *flag_analyze_DE_all5*).

In addition to photosynthesis, additional GO enrichment categories are found when genes significantly differentially expressed in the four more sensitive genotypes are analyzed (Figure 4A, n=789). These include GO biological processes related to metabolism (aromatic amino acid family biosynthetic process, carotenoid biosynthetic process, coenzyme biosynthetic process, glycine catabolic process, hydrogen peroxide catabolic process, oxylipin biosynthetic process, regulation of lipid metabolic process, unsaturated fatty acid biosynthetic process), additional photosynthesis terms (photosynthetic electron transport in photosystem I, photosystem I assembly, reductive pentose-phosphate cycle, triose phosphate transmembrane transport, response to red light, response to light stimulus, response to low light intensity stimulus), protein-chromophore linkage, iron-sulfur cluster assembly, and detection of biotic stimulus (Supplementary Table 6, *flag_analyze_DE_all_noB73*). The direction of regulation of the genes significant in these 4 more ozone sensitive genotypes is in the same direction for 788 of the 789 genes (~46% up-regulated, ~54% down-regulated) with Zm00001d037513 (germ-like protein1, EnsemblPlants) up-regulated in C123 and down-regulated in the other 3 genotypes (Figure 4C).

Chlorophyll content estimate from a spectrophotometer follows the same pattern as gene expression. The transcriptional response to ozone stress is largely attributable to genes involved in photosynthetic and respiratory functions in the leaf. These general effects are common across genotypes, but the magnitude of response varies. We hypothesize that the magnitude of the transcriptional response to ozone is correlated with changes in leaf chlorophyll content. To evaluate this hypothesis, we re-analyze the ambient and ozone data in Yendrek et al. (YENDREK *et al.* 2017b) for the maize genotypes in this study. In agreement with our transcriptional analysis, we find that chlorophyll levels are significantly lower under elevated ozone conditions for NC338 ($p<0.00001$), Mo17 ($p<0.01$), and C123 ($p=0.003$), while there is no significant difference in chlorophyll content in the ambient and elevated ozone for B73 (Figure 5).

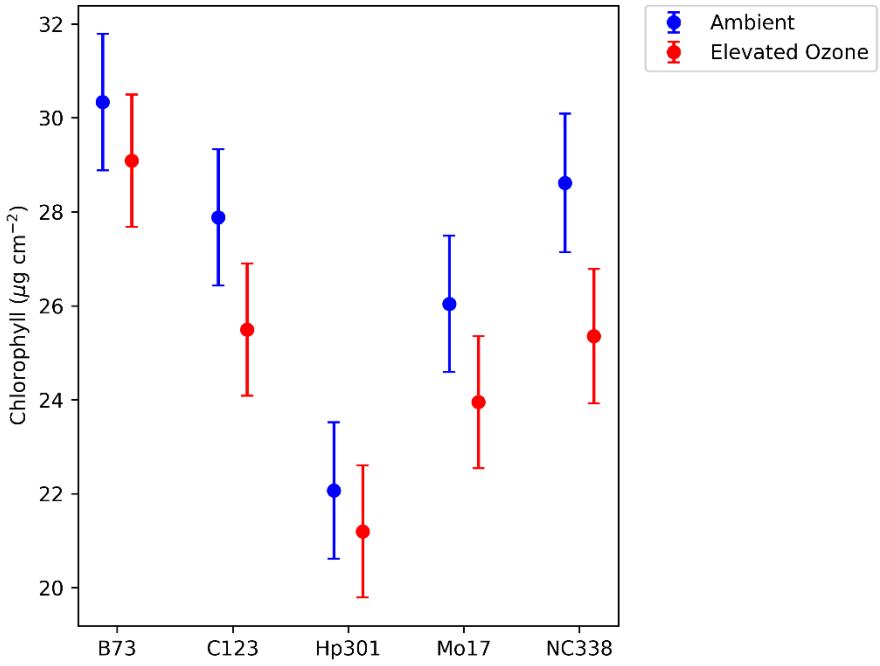

Figure 5: The effect of elevated ozone on the PLSR chlorophyll score for temperate genotypes B73, C123, Hp301, Mo17 and NC338. The chlorophyll content was estimated by high throughput spectrophotometry in Yendrek et al. (2017b). Shown here are effect sizes estimated from least squares means of the PLSR chlorophyll in elevated ozone relative to ambient ozone conditions. B73 has the highest levels of PLSR chlorophyll and is these levels are not significantly affected by ozone.

Results are robust to different short read mapping strategies. In short read data derived from gene families, there can be ambiguity as to the origin of these reads. In the literature, there is a general consensus that, overall, there is a reduction in bias when genes and transcripts unlikely to be expressed are eliminated from consideration in mapping (MATSUOKA *et al.* 2002; MCMULLEN *et al.* 2009; HIRSCH and SPRINGER 2017; NEWMAN *et al.* 2018; TARDAGUILA *et al.* 2018). However, it is possible that while there is overall bias reduction, for some genes true signal may be missed with this strategy. We conducted 3 additional differential expression analyses where we map short reads to the genome instead of the estimated transcriptome in order to evaluate whether this strategy impacted our findings (Supplementary Methods). We mapped all short reads to the B73 genome, and separately to the Mo17 Cau genome and for all analyzable genes performed differential expression analyses and a GO enrichment. In order to compare results, we restricted the GO enrichment to the set of genes with one-to-one B73-Mo17 syntenic gene pairs. We compared numbers of reads mapping for the one-to-one syntenic gene pairs and they are similar (Supplementary Figure 3). All inferences about differential expression are consistent with the results described above. In all cases, B73 has low levels of differential expression in response to ozone exposure and NC338 has orders of magnitude higher expression (See https://github.com/McIntyre-Lab/papers/tree/master/nanni_maize_2022). Many of the GO terms were identical among all 4 analyses, and variation among GO terms did not identify new potential mechanisms of action but rather pointed to the same overall processes. All results from all 4 analyses are provided, and the summary of the GO enrichment across all 4 analyses is also provided (See https://github.com/McIntyre-Lab/papers/tree/master/nanni_maize_2022).

In summary, our results indicate that the magnitude of the expression response to elevated ozone mirrors physiological differences among maize genotypes.

Discussion

Despite the diversity of the maize genotypes and the extensive literature on structural variation among maize genotypes, the long and short read data collected from leaf tissue in these experiments of temperate genotypes largely reflect a set of shared transcripts. Clustering of CCS long reads was carried out independently and without a reference, in order to avoid potential reference bias. A close examination of the CCS long reads, the clusters for each genotype, and the assembled transcriptome reveal that the overwhelming majority (~94%) of genes expressed

in the leaf are part of the annotated B73-Mo17 syntenic genome. The biological inferences of the differential expression analyses are robust to alternative mapping strategies.

The pattern of the expression response to elevated ozone is very similar in loci differentially expressed in all 5 genotypes (82/83). Loci differentially expressed in NC338, Hp301, Mo17 and C123 also display the same pattern in over/under expression among these 4 genotypes (788/789), even as the magnitude of expression in response to elevated ozone is dramatically different, with NC338 having the strongest response. We see no evidence for structural variation (or tandem duplication) in the transcriptional response to ozone but cannot rule out variation in *cis* regulation of these genes, which may well be influenced by adjacent variable genomic content (STUPAR and SPRINGER 2006; DOONER and WEIL 2007; HAWKINS *et al.* 2014; MAKAREVITCH *et al.* 2015; HIRSCH and SPRINGER 2017; STEIGE *et al.* 2017; ROESSLER *et al.* 2018).

There are 83 differentially expressed genes in response to elevated ozone in all 5 maize genotypes. Decreased expression of genes involved in light harvesting dominated the conserved transcriptional response. These include chlorophyll biosynthesis genes, chlorophyll a/b binding proteins, and subunits of photosynthetic proteins. Down-regulation of photosynthesis is a common transcriptional response to ozone (XU *et al.* 2015), and chlorophyll biosynthetic genes are also reduced in an ozone sensitive Medicago accession (PUCKETTE *et al.* 2008). In an ozone sensitive soybean genotype, decreased photosystem II gene expression is observed after 24 hrs of ozone exposure (WALDECK *et al.* 2017). The sensitivity of photosynthesis to ozone is clear across both C_3 and C_4 crops, and remains a critical target for improving ozone tolerance (EMBERSON *et al.* 2018). Perhaps to combat damage to photosynthetic proteins, increased expression of heat shock proteins is another common response of the 5 genotypes to elevated ozone.

The genotypes assayed in this study differed in photosynthetic capacity under elevated ozone in the field (YENDREK *et al.* 2017a) and they differ in transcriptional responses to ozone stress. Allelic variation in transcriptional response to stress has been documented (GUO *et al.* 2004). Over 20 times more genes are differentially expressed in NC338 in ambient and elevated ozone compared B73, and the transcriptional response of B73 is muted relative to the other genotypes. In a companion field experiment, there is a greater decline in chlorophyll content associated with accelerated senescence in inbreds C123 and NC338 as compared to B73 (YENDREK *et al.*

2017b). The metabolite profile of B73 was not responsive to elevated ozone in the field (WEDOW *et al.* 2021b). Additionally, hybrids containing parents Hp301 and NC338 showed greater reductions in photosynthesis under elevated ozone in the field (CHOQUETTE *et al.* 2019). Ozone tolerance has been associated with a dampened transcriptional response in genotypes of other species including Arabidopsis (XU *et al.* 2015), soybean (BURTON *et al.* 2016; WALDECK *et al.* 2017) and Medicago (PUCKETTE *et al.* 2008). A dampened transcriptional response to other abiotic stresses including drought (FRACASSO *et al.* 2016), cold stress (DA MAIA *et al.* 2017) and salt stress (YANG *et al.* 2017b) was observed in tolerant genotypes.

These results are consistent with oxidative stress tolerance being associated with a weaker response to a reactive oxygen species signal. Notably, B73 has equivalent, or greater, stomatal conductance than the more sensitive maize genotypes studied here (YENDREK *et al.* 2017b). This means that the flux of ozone into the leaf mesophyll would tend to be greatest in B73. That an equal or greater ozone dose elicited a weaker transcriptional response, suggests that genetic variation in downstream processes is key to ozone tolerance; i.e., the processes that either sense ROS in the apoplast, transduce signals, produce the secondary burst of ROS in the cytoplasm, or respond to the secondary ROS burst (AINSWORTH 2017). For example, GO enrichment analysis (Supplementary Table 6), indicated differential expression of genes related to *cellular response to hydrogen peroxide*, *detection of biotic stimulus*, *hydrogen peroxide catabolic process*, and *regulation of response to biotic stimulus* in the ozone sensitive genotypes, but not B73. At the same time, there is differential gene expression related to *response to salicylic acid* in B73 that is not observed in the more ozone sensitive genotypes. The strength of ozone impacts on Arabidopsis varies with the strength of ethylene-signaling components of the pathway when modulated genetically or chemically (RAO *et al.* 2002). The transcriptional data presented here suggest a similar strategy might be relevant in maize and could assist in the identification of potential targets for manipulation.

Conclusions:

The long and short read data collected from leaf tissue in these experiments of temperate genotypes (B73, Mo17, Nc338,C123, Hp301) largely reflect a set of shared transcripts. While the overall response to ozone is largely consistent in direction, the extent of the response is dramatically different among the genotypes, ranging from 150 genes significantly differentially

expressed (B73) to more than 3,000 genes significantly differentially expressed (NC338). B73 has a dampened transcriptional response, suggesting potential fundamental differences compared to the other 4 genotypes.

We found that all genotypes show regulatory responses in genes related to photosynthesis. Our findings show that ozone directly impacts the photosynthetic machinery of the leaf, and a potential key element to reduction in sensitivity to ozone is a controlled response to the presence of ozone.

Methods:

Experimental design and sample collection: Twenty-four plants from each genotype were grown in 5.7 L pots in LC1 mix (Sun Gro Horticulture Distribution Inc., Bellevue, WA, USA) in six growth chambers (Growth Chamber, Chagrin Falls, OH, USA) set to maintain a 15 hr day at 25 °C and photosynthetic photon flux (PPF) of ~350 µmol m^{-2} s^{-1}, 9 hr night at 21 °C, and a relative humidity of 60%. We grew 120 plants in the six growth chambers, with three chambers maintained at low ozone conditions (< 10 ppb) and three at elevated ozone conditions (100 ppb). Ozone is generated using a variable output UV-C light bulb ballast (HVAC 560 ozone generator, Crystal Air, Langley, Canada) and plants are fumigated with 100 ppb ozone for 6 hrs per day from emergence to sampling. Four plants of each of the five genotypes, B73, C124, Hp301, NC338, Mo17 were randomly assigned to each chamber for a total of twenty-four plants per genotype. Plants were fertilized (Osmocote Blend 19-5-8) at the start of the experiment. Twenty-six days after planting, leaf tissue from the 5^{th} leaf is sampled directly into liquid N_2 for a total of 120 independent leaf samples (5 genotypes x 4 plants per genotype x 2 treatment conditions x 3 growth chambers per treatment) for a total of 12 replicates per genotype and ozone treatment. Leaf samples were sent to University of Florida where leaf punches were collected in a cold room on dry ice. A separate sample from a single plant from each genotype and ozone treatment was collected for long read sequencing.

Long-read library preparation and sequencing: Leaf tissue (~175 mg) from a single individual plant of each genotype/treatment combination (n=10) was collected and snap frozen in liquid nitrogen. Total RNA was isolated independently for each of the 10 samples (5 genotypes x 2 treatments) followed by full-length cDNA synthesis using the SMARTer PCR cDNA Synthesis kit and single molecule SMRTbell library construction followed by PCR optimization and ELF

size selection. One library (Mo17 elevated) was selected to calibrate the run, and then all ten libraries sequenced on the PacBio SEQUEL platform using 1 LR SMRT cell per sample and SMRT Link 6.0.0 chemistry to generate raw subreads for each sample by the University of Florida Interdisciplinary Center for Biotechnology Research (ICBR). Data (subread.bam files) have been deposited at the SRA (BioProject accession PRJNA604929).

RNA-seq library preparation and sequencing: Approximately 20mg of flash frozen leaf tissue from each sample was placed into individual mini tubes in a 96-tube plate format (Axygen MTS-11-C-R). Sample freeze-drying, mRNA purification, cDNA synthesis and library preparation were carried out by Rapid Genomics (Gainesville, FL) to generate an individual dual indexed library per sample. Individual libraries were quantified and pooled to generate equimolar samples for sequencing. The final pooled library was sequenced on 3 lanes of an Illumina HiSeq 3000 at Rapid Genomics and 1 lane of Illumina Novaseq at Novogene (Sacramento, CA). Data have been deposited at the SRA (BioProject accession PRJNA604929).

Mapping short-reads to available reference genomes: Adapter trimmed unique short reads are mapped to B73 (Ensembl version 41, ftp://ftp.ensemblgenomes.org/pub/plants/release-41/fasta/zea_mays/, (JIAO *et al.* 2017)), Mo17 Yan (https://ftp.maizegdb.org/MaizeGDB/FTP/Mo17-YAN/, (YANG *et al.* 2017a)) and Mo17 Cau (https://ftp.maizegdb.org/MaizeGDB/FTP/Mo17-CAU/, (SUN *et al.* 2018)) using BWA-mem with default parameters and '-M' to mark short split reads as secondary alignment (v0.7.17, (LI 2013)). Unique primary alignments with an alignment score of at least 30 (default of BWA-mem) are selected for further analysis. The maximum percentage of unmapped reads was 3.8% therefore more than 96% of the reads mapped to all three genomes (Supplementary Table 2). On average, 98% of reads across all samples mapped to the B73 v4 reference genome. The sample with the lowest mapping proportion of reads of all samples was a NC338 elevated ozone sample (https://github.com/McIntyre-Lab/papers/tree/master/nanni_maize_2022)

Long-read pre-processing: The IsoSeq3 pipeline (v3.0.0, https://github.com/PacificBiosciences/IsoSeq3) was used to process subreads into circular consensus reads (CCS (v3.1.0)). CCS for each sample had, on average, approximately 3 sub-reads, indicating successful multiple passes during sequencing and enabling error correction. Primers were removed from CCS using lima (v1.7.1). IsoSeq3 cluster was used to trim polyA

tails, remove false concatemers, and cluster full-length reads by sequence similarity. Consensus ('polished') sequences for each cluster were generated using IsoSeq3 polish. The expected error rate after this step was less than < 2% (WEIRATHER *et al.* 2017). Up to this point no genome references were used.

Mapping CCS: CCS, after primer trimming, were mapped to the three available reference genomes using minimap2 (v2.12, (LI 2018)) and parameters recommended for PacBio IsoSeq processing (-ax splice -uf --secondary=no -C5, https://github.com/Magdoll/cDNA_Cupcake/wiki/Best-practice-for-aligning-Iso-Seq-to-reference-genome:-minimap2,-deSALT,-GMAP,-STAR,-BLAT/24228b7e9c329268b0fdc85226097129d9c44a82) (Supplementary Figure 12). Mapping rates exceeded 99% (Supplementary Table 1, lines 23 and 32). Unmapped reads were compared to the NCBI nucleotide non-redundant database using BLAST (ALTSCHUL *et al.* 1990). Greater than 47% (3,613 out of 7,637) of unmapped reads had no BLAST hit, the remaining reads resulted in partial hits, many with numerous mismatches and gaps.

Mapping clusters: Although there were similar numbers of CCS reads across samples (Supplementary Table 1, line 8), the number of clusters was more variable: 24,000 to 44,000 clusters (Supplementary Table 1, line 23). The polished cluster sequences were mapped to the three available references using minimap2 (v2.12, (LI 2018)) and parameters recommended for PacBio IsoSeq3 processing (-ax splice -uf --secondary=no -C5). Unmapped clusters were binned according to length and unmapped clusters larger than 1 kb were evaluated by BLAST (ALTSCHUL *et al.* 1990).

Assembled Transcripts for each sample: For clusters that map to each of the three references (Supplementary Table 1, line 24), we used the IsoSeq3 supporting script in Cupcake ToFU (version 20180629, https://github.com/Magdoll/cDNA_Cupcake, *collapse_isoforms_by_sam.py*) to identify identical isoforms within each library and collapse them into a single unique assembled transcript (Supplementary Table 1, line 27).

SQANTI QC: SQANTI QC was used to match the assembled transcripts to an annotated reference transcript (TARDAGUILA *et al.* 2018) and to classify the resulting splice junction pattern. A full-splice match (FSM) is a transcript where all splice junctions in the assembled transcript match exactly to an annotated reference transcript. Incomplete-splice matches (ISM),

have junctions in the same order as a reference transcript, but may be missing some junctions on the 5' and/or 3' end of the transcript. Transcripts that contained novel combinations of annotated junctions, novel in catalog (NIC), and those with new junctions within an annotated gene, novel not in catalog (NNC) were also identified. In addition, novel transcripts within potentially genic areas were identified (TARDAGUILA *et al.* 2018). SQANTI QC was run on CCS long reads (Supplementary Table 1, line 17-20) assembled transcripts (Supplementary Table 1, line 44-52) and the assembled leaf transcriptome (Supplementary Table 1, line 62-71).

Evaluation of filtered clusters: Clusters that did not pass Cupcake ToFU filtering process ('ignored') due to low identity (< 95%) or low coverage (< 99%) alignments to the B73 reference genome were potential candidates for genes with sequence divergence compared to the B73 reference. Alternatively, they may be paralogous genes in the non-B73 genotypes with PAV. Ignored clusters greater than 1 kb (of sufficient length to suggest a complete transcript) were evaluated using BLAST BLAST (ALTSCHUL *et al.* 1990) against the NCBI non-redundant nucleotide database (Supplementary Table 1, line 37-40). Best hits were defined as the BLAST hit with the largest bitscore and smallest e-value for each filtered cluster (https://github.com/McIntyre-Lab/papers/tree/master/nanni_maize_2022). Best hits were further assessed for subject names associated with gene families ("family" or "kinase" in name). Ignored clusters from C123, Hp301, and NC338 were also compared to the assembled transcripts from the B73 and Mo17 samples using BLAST (https://github.com/McIntyre-Lab/papers/tree/master/nanni_maize_2022).

Comparison of assembled transcripts to NAM pan-gene classifications. A recent assembly and annotation of 26 maize lines from the nested association mapping (NAM) population (including B73 and Hp301) identified pan-gene associations and classified B73 genes as "core" (in all 26 lines), "near-core" (in 24 of the 26 lines), "dispensable" (in 2 to 23 lines), and "private" (only in a single line) (HUFFORD *et al.* 2021). We used the Ensembl B73 v4 reference annotation which required a conversion to the NAM associated v5 gene identifiers to compare pan-gene classifications. A file from maizeGDB (https://maizegdb.org/)(WOODHOUSE *et al.* 2021) was used to link v5 genes to each v4 gene (https://download.maizegdb.org/Pan-genes/B73_gene_xref/B73v4_to_B73v5.tsv). A unique list of B73 v4 reference genes representing the genes associated with the assembled transcripts

across all samples was compared to the pan-genes in the NAM population. Additionally, the 12,604 annotated genes in the assembled leaf transcriptome were linked to the NAM pan-gene classifications.

Chaining multiple samples to identify redundancy: Cupcake ToFU IsoSeq3 supporting scripts (*summarize_sample_GFF_junctions.py*, *scrub_sample_GFF_junctions.py* and *chain_samples.py*) were used to create a non-redundant union of assembled transcripts. The clustering and chaining protocol was carried out on unique assembled transcripts mapped to the B73 or Mo17 Yan references separately (Supplementary Figure 13), yielding two chained sets of assembled transcripts (Supplementary Table 1). The B73 and Mo17 Yan assembled transcriptomes had similar numbers of transcripts (79,620 and 76,980 respectively; Supplementary Table 1, line 60). Transcript quality control was performed with SQANTI QC and filter functions (TARDAGUILA *et al.* 2018) and included identification of transcripts with known technological artifacts such reverse transcriptase switching (TARDAGUILA *et al.* 2018). There was a slightly higher number of annotated genes (Supplementary Table 1 line 65) and transcripts with annotated junction chains (Supplementary Table 1 line 68-69) when the assembled transcriptomes were mapped to B73 v4 as compared to Mo17 Yan. There were fewer novel loci identified when the assembled transcriptomes were mapped to B73 v4 as compared to Mo17 Yan (Supplementary Table 1 line 69).

Synteny: SynMap (HAUG-BALTZELL *et al.* 2017) and SynFind within CoGe (LYONS and FREELING 2008) were used to identify collinear gene-sets between the B73 v4 reference and the Mo17 Cau reference (Margaret Woodhouse, maizeGDB). We also verified syntenic genes according to (SUN *et al.* 2018). Information in Supplementary File 5.

Novel genes: Assembled transcripts associated with putative novel loci were quantified using HTSeq (v 0.11.2, (ANDERS *et al.* 2015)). Novel loci show evidence for expression if at least one read is detected in at least one replicate of either treatment (Supplementary Figure 4).

Annotation of assembled transcripts: FASTA files of the assembled transcripts and the associated translated sequences were used as queries to obtain putative functional, structural or signaling motif information from InterProScan (version 5.29-68.0, (QUEVILLON *et al.* 2005)), TMHMM (version 2.0, (SONNHAMMER *et al.* 1998; KROGH *et al.* 2001)) and SignalP (version 4.1, (NIELSEN *et al.* 1997; NIELSEN 2017)) for tappAS (DE LA FUENTE *et al.* 2019). Translated

transcripts were linked to proteins in public databases via sequence similarity (Supplementary Table 9). Repetitive regions were annotated using RepeatMasker (version 4.0.5, http://www.repeatmasker.org) and putative RNA regulatory sequences identified using ScanForMotif (BISWAS and BROWN 2014). Default parameters were used. We collated all the above annotation into a GFF-like file (Supplementary File 1).

Evidence for expression: The list of genes with assembled transcripts were evaluated for expression using HTSeq (version 0.11.2, (ANDERS *et al.* 2015)) with CCS primer trimmed reads and adapter trimmed unique short reads for each sample separately. Evidence of expression for each gene in each sample is defined as at least 1 read (short or long) mapping to the gene (Figure 2A).

Secondary Data: The raw B73 PacBio data (.bas.h5) were obtained from Wang et al. (WANG *et al.* 2018a) and consist of the following tissue types: leaf, silk, pericarp, bract, shoot and seedling (Supplementary Figure 6). All PacBio data were processed as described above. Clustering and polishing of trimmed reads yielded 38,741 high quality transcripts in the leaf.

Quantification of the assembled transcriptome: Paired-end Illumina short-read data was used for quantification of the B73 assembled leaf transcriptome. FASTQ files were trimmed of adapter sequence (cutadapt, version 2.1, (MARTIN 2011)) and merged if R1 and R2 reads overlapped (*bbmerge.py* from BBMap version 38.44, (BUSHNELL *et al.* 2017)). Duplicate reads were removed. Gene and transcript expression values estimated using RSEM with STAR (LI and DEWEY 2011; DOBIN *et al.* 2013) against the B73 v4 genome with the B73 assembled transcriptome GTF file (Supplementary File 2). Merged (single-end) and unmerged (paired-end) reads were aligned separately. TPM (transcripts per kilobase million) values are output for genes and transcripts. Single-end (R1-R2 merged reads) and paired-end counts were each assumed to represent 1 read and TPM were summed. Transcripts are analyzable for differential expression if their TPM values are greater than 5 in 50% of the replicates in either ambient or elevated ozone condition (Figure 2B, Supplementary File 3, flag *_genotype_condition*).

Differential expression: For each genotype, analyzable transcripts in either ambient or elevated ozone conditions were evaluated for differential expression and differential splicing using tappAS (version 1.0.6, (DE LA FUENTE *et al.* 2019)) at FDR level (BENJAMINI and HOCHBERG 1995) of 0.05 (Supplementary Figure 14). GO enrichment using Fisher's exact test and an FDR

threshold of 0.05 is performed (Supplementary Table 6).Analysis of chlorophyll data: We re-analyzed the PLSR predicted chlorophyll values inform the Illinois FACE (Free Air Concentration Enrichment, https://www.igb.illinois.edu/soyface/) site under ambient and elevated ozone conditions (Yendrek et al. 2017b) Each ringpair was divided into 5 sub-blocks and genotypes were planted in each sub-block. The difference in chlorophyll (as measured by the PLSR model) between ozone and ambient conditions was tested in a mixed effects model with a random effect of ring-pair and fixed effects of genotype and ozone treatment. The model also included an effect for potential excess ozone application in one of the sub-blocks due to a prevailing wind.

Acknowledgments

NSF Plant Genome Research Program (PGR-1238030; ADBL, LMM, and EAA) MCA-PGR: Genetic and genomic approaches to understand and improve maize responses to ozone NIH R01GM128193 (LMM), R03CA222444 (AC, LMM) USDA SoyFACE Global Change Research. Project Number: 5012-21000-030-17-S. Brad Barbazuk for helpful discussion, Margaret Woodhouse (with maizeGDB for synteny file), Junping Shi for providing the synteny list from Sun et al. 2018. Deborah Morse and James Resnick for assistance with sample processing. Francisco Pardo-Palacios and Pedro Salguero Garcia for help with maize functional annotation. Dr. Bo Wang, Peter van Buren and Dr. Doreen Ware for graciously sharing maize PacBio raw data. Three anonymous reviewers for excellent criticism.

Data Availability

Raw data (BAM files for the PacBio data and FASTQ files for the RNA-seq data) are deposited under SRA BioProject accession PRJNA604929. All processed data are attached to this manuscript, including the GTF file for the transcriptome (Supplementary File 2), quantified expression data for analysis (Supplementary File 3) and annotations for the transcriptome (Supplementary File 1). All scripts and full detailed documentation for all the analyses are posted on github (https://github.com/McIntyre-Lab/papers/tree/master/nanni_maize_2022). All data files including all FASTA files used for BLAST analyses; all BLAST results, all results from alternate differential expression analyses, and all enrichment results from these alternative

analyses are deposited on zenodo (link- we will create this deposit when the final content is approved, these deposits are not easily updated).

Author contributions

Adalena V. Nanni developed all analysis plans, analyzed all short-read data, contributed to analysis of long read data, annotated the maize transcriptomes, managed all references, designed and contributed to figures, contributed to the writing of the paper. Worked on the github page.

Alison M. Morse designed and executed the analysis performed all long-read analyses, interpreted the findings, designed and contributed to figures, contributed to the writing of the paper and is responsible for the github page.

Jeremy R. B. Newman participated in analysis, designed and contributed to figures, contributed to the writing of the paper.

Nicole E. Choquette identified genotypes for testing, performed the growth chamber experiment and collected tissue.

Jessica M. Wedow identified genotypes for testing, performed the growth chamber experiment and collected tissue.

Zihao Liu created the transcriptome visualization plots.

Andrew D. B. Leakey contributed to the overall design of the umbrella project (PGR-1238030), interpreted the findings, designed figures, interpreted the findings, contributed to the writing of the paper.

Ana Conesa contributed to the analyses of long reads, interpreted the findings, contributed to the writing of the paper.

Elizabeth A. Ainsworth contributed to the overall design of the umbrella project (PGR-1238030), designed this particular experiment, identified genotypes for testing, supervised tissue collection, participated in analysis, designed figures, interpreted the findings, contributed to the writing of the paper.

Lauren M. McIntyre contributed to the overall design of the umbrella project (PGR-1238030), designed this particular experiment, designed and supervised the analysis, interpreted the findings, contributed to the writing of the paper.

Ahn, S., and S. D. Tanksley, 1993 Comparative linkage maps of the rice and maize genomes. Proceedings of the National Academy of Sciences of the United States of America 90**:** 7980-7984.

Ahuja, I., R. C. H. de Vos, A. M. Bones and R. D. Hall, 2010 Plant molecular stress responses face climate change. Trends in Plant Science 15**:** 664-674.

Ainsworth, E. A., 2017 Understanding and improving global crop response to ozone pollution. Plant Journal 90**:** 886-897.

Ainsworth, E. A., P. Lemonnier and J. M. Wedow, 2020 The influence of rising tropospheric carbon dioxide and ozone on plant productivity. Plant Biology 22**:** 5-11.

Altschul, S. F., W. Gish, W. Miller, E. W. Myers and D. J. Lipman, 1990 Basic local alignment search tool. Journal of Molecular Biology 215**:** 403-410.

Anders, S., P. T. Pyl and W. Huber, 2015 HTSeq-a Python framework to work with high-throughput sequencing data. Bioinformatics 31**:** 166-169.

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.

Bennetzen, J. L., 2005 Transposable elements, gene creation and genome rearrangement in flowering plants. Current Opinion in Genetics & Development 15**:** 621-627.

Biswas, A., and C. M. Brown, 2014 Scan for Motifs: a webserver for the analysis of post-transcriptional regulatory elements in the 3′ untranslated regions (3′ UTRs) of mRNAs. BMC Bioinformatics 15**:** 174.

Brohammer, A. B., T. J. Y. Kono, N. M. Springer, S. E. McGaugh and C. N. Hirsch, 2018 The limited role of differential fractionation in genome content variation and function in maize (Zea mays L.) inbred lines. Plant Journal 93**:** 131-141.

Burney, J., and V. Ramanathan, 2014 Recent climate and air pollution impacts on Indian agriculture. Proc Natl Acad Sci U S A 111**:** 16319-16324.

Burton, A. L., K. O. Burkey, T. E. Carter, J. Orf and P. B. Cregan, 2016 Phenotypic variation and identification of quantitative trait loci for ozone tolerance in a Fiskeby III × Mandarin (Ottawa) soybean population. Theor Appl Genet 129**:** 1113-1125.

Bushnell, B., J. Rood and E. Singer, 2017 BBMerge - Accurate paired shotgun read merging via overlap. Plos One 12.

Calixto, C. P. G., W. B. Guo, A. B. James, N. A. Tzioutziou, J. C. Entizne *et al.*, 2018 Rapid and Dynamic Alternative Splicing Impacts the Arabidopsis Cold Response Transcriptome. Plant Cell 30**:** 1424-1444.

Choquette, N., F. Ogut, T. Wertin, C. Montes, C. Sorgini *et al.*, 2019 Uncovering hidden genetic variation in photosynthesis of field-grown maize under ozone pollution. Global Change Biology 25**:** 4327-4338.

Choquette, N. E., E. A. Ainsworth, W. Bezodis and A. P. Cavanagh, 2020 Ozone tolerant maize hybrids maintain Rubisco content and activity during long-term exposure in the field. Plant Cell and Environment 43**:** 3033-3047.

Conklin, P. L., and C. Barth, 2004 Ascorbic acid, a familiar small molecule intertwined in the response of plants to ozone, pathogens, and the onset of senescence. Plant Cell and Environment 27**:** 959-970.

da Maia, L. C., P. R. B. Cadore, L. C. Benitez, R. Danielowski, E. J. B. Braga *et al.*, 2017 Transcriptome profiling of rice seedlings under cold stress. Functional Plant Biology 44**:** 419-429.

Darracq, A., C. Vitte, S. Nicolas, J. Duarte, J.-P. Pichon *et al.*, 2018 Sequence analysis of European maize inbred line F2 provides new insights into molecular and chromosomal characteristics of presence/absence variants. Bmc Genomics 19.

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.

Ding, F., P. Cui, Z. Y. Wang, S. D. Zhang, S. Ali *et al.*, 2014 Genome-wide analysis of alternative splicing of pre-mRNA under salt stress in Arabidopsis. Bmc Genomics 15.

Dobin, A., C. A. Davis, F. Schlesinger, J. Drenkow, C. Zaleski *et al.*, 2013 STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29**:** 15-21.

Dooner, H. K., and C. F. Weil, 2007 Give-and-take: interactions between DNA transposons and their host plant genomes. Current Opinion in Genetics & Development 17**:** 486-492.

Egawa, C., F. Kobayashi, M. Ishibashi, T. Nakamura, C. Nakamura *et al.*, 2006 Differential regulation of transcript accumulation and alternative splicing of a DREB2 homolog under abiotic stress conditions in common wheat. Genes & Genetic Systems 81**:** 77-91.

Emberson, L. D., H. Pleijel, E. A. Ainsworth, M. van den Berg, W. Ren *et al.*, 2018 Ozone effects on crops and consideration in crop models. European Journal of Agronomy 100**:** 19-34.

Emrich, S. J., L. Li, T. J. Wen, M. D. Yandeau-Nelson, Y. Fu *et al.*, 2007 Nearly identical paralogs: Implications for maize (Zea mays L.) genome evolution. Genetics 175**:** 429-439.

Entrambasaguas, L., M. Ruocco, K. J. F. Verhoeven, G. Procaccini and L. Marin-Guirao, 2021 Gene body DNA methylation in seagrasses: inter- and intraspecific differences and interaction with transcriptome plasticity under heat stress. Scientific Reports 11.

Feuillet, C., and B. Keller, 2002 Comparative Genomics in the grass family: Molecular characterization of grass genome structure and evolution. Annals of Botany 89**:** 3-10.

Filichkin, S. A., H. D. Priest, S. A. Givan, R. K. Shen, D. W. Bryant *et al.*, 2010 Genome-wide mapping of alternative splicing in Arabidopsis thaliana. Genome Research 20**:** 45-58.

Fracasso, A., L. M. Trindade and S. Amaducci, 2016 Drought stress tolerance strategies revealed by RNA-Seq in two sorghum genotypes with contrasting WUE. BMC Plant Biology 16**:** 115.

Fujita, M., Y. Fujita, Y. Noutoshi, F. Takahashi, Y. Narusaka *et al.*, 2006 Crosstalk between abiotic and biotic stress responses: a current view from the points of convergence in the stress signaling networks. Current Opinion in Plant Biology 9**:** 436-442.

Gadjev, I., S. Vanderauwera, T. S. Gechev, C. Laloi, I. N. Minkov *et al.*, 2006 Transcriptomic footprints disclose specificity of reactive oxygen species signaling in Arabidopsis. Plant Physiology 141**:** 436-445.

Gale, M., and K. Devos, 1998 Comparative genetics in the grasses. Proceedings of the National Academy of Sciences of the United States of America 95**:** 1971-1974.

Gao, L., I. Gonda, H. H. Sun, Q. Y. Ma, K. Bao *et al.*, 2019 The tomato pan-genome uncovers new genes and a rare allele regulating fruit flavor. Nature Genetics 51**:** 1044-+.

Golicz, A. A., P. E. Bayer, G. C. Barker, P. P. Edger, H. Kim *et al.*, 2016 The pangenome of an agronomically important crop plant Brassica oleracea. Nature Communications 7.

Goyal, K., L. J. Walton and A. Tunnacliffe, 2005 LEA proteins prevent protein aggregation due to water stress. Biochemical Journal 388**:** 151-157.

Guo, M., M. A. Rupe, C. Zinselmeier, J. Habben, B. A. Bowen *et al.*, 2004 Allelic variation of gene expression in maize hybrids. The Plant cell 16**:** 1707-1716.

Han, X., Y. Qin, F. Yu, X. Ren, Z. Zhang *et al.*, 2019 A megabase-scale deletion is associated with phenotypic variation of multiple traits in maize. Genetics 211**:** 305-316.

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.

Hawkins, J. S., V. Delgado, L. Feng, M. Carlise, H. K. Dooner *et al.*, 2014 Variation in allelic expression associated with a recombination hotspot in Zea mays. The Plant Journal 79**:** 375-384.

Hirsch, C. D., and N. M. Springer, 2017 Transposable element influences on gene expression in plants. Biochim Biophys Acta Gene Regul Mech 1860**:** 157-165.

Hirsch, C. N., J. M. Foerster, J. M. Johnson, R. S. Sekhon, G. Muttoni *et al.*, 2014 Insights into the maize pan-genome and pan-transcriptome. Plant Cell 26**:** 121-135.

Hoopes, G. M., J. P. Hamilton, J. C. Wood, E. Esteban, A. Pasha *et al.*, 2019 An updated gene atlas for maize reveals organ-specific and stress-induced genes. Plant Journal 97**:** 1154-1167.

Hu, H., A. Scheben, B. Verpaalen, S. Tirnaz, P. E. Bayer *et al.*, 2021 Amborella gene presence/absence variation is associated with abiotic stress responses that may contribute to environmental adaptation. The New phytologist.

Hufford, M. B., A. S. Seetharam, M. R. Woodhouse, K. M. Chougule, S. J. Ou *et al.*, 2021 De novo assembly, annotation, and comparative analysis of 26 diverse maize genomes. Science 373**:** 655-+.

Hulbert, S. H., T. E. Richter, J. D. Axtell and J. L. Bennetzen, 1990 Genetics-mapping and characterization of sorghum and related crops by means of maize DNA probes. Proceedings of the National Academy of Sciences of the United States of America 87**:** 4251-4255.

Iida, K., M. Seki, T. Sakurai, M. Satou, K. Akiyama *et al.*, 2004 Genome-wide analysis of alternative pre-mRNA splicing in Arabidopsis thaliana based on full-length cDNA sequences. Nucleic Acids Research 32**:** 5096-5103.

Ito, H., T. Yoshida, S. Tsukahara and A. Kawabe, 2013 Evolution of the ONSEN retrotransposon family activated upon heat stress in Brassicaceae. Gene 518**:** 256-261.

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

Jin, M. L., H. J. Liu, C. He, J. J. Fu, Y. J. Xiao *et al.*, 2016 Maize pan-transcriptome provides novel insights into genome complexity and quantitative trait variation. Scientific Reports 6.

Kim, J. H., 2021 Multifaceted Chromatin Structure and Transcription Changes in Plant Stress Response. International Journal of Molecular Sciences 22.

Krogh, A., B. Larsson, G. von Heijne and E. L. Sonnhammer, 2001 Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. J Mol Biol 305**:** 567-580.

Lai, J., R. Li, X. Xu, W. Jin, M. Xu *et al.*, 2010 Genome-wide patterns of genetic variation among elite maize inbred lines. Nature Genetics 42**:** 1027-+.

Laloum, T., G. Martin and P. Duque, 2018 Alternative Splicing Control of Abiotic Stress Responses. Trends in Plant Science 23**:** 140-150.

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.

Li, H., 2018 Minimap2: pairwise alignment for nucleotide sequences. Bioinformatics 34**:** 3094-3100.

Li, W. F., W. D. Lin, P. Ray, P. Lan and W. Schmidt, 2013 Genome-Wide Detection of Condition-Sensitive Alternative Splicing in Arabidopsis Roots. Plant Physiology 162**:** 1750-1763.

Liu, K., M. Goodman, S. Muse, J. S. Smith, E. Buckler *et al.*, 2003 Genetic structure and diversity among maize inbred lines as inferred from DNA microsatellites. Genetics 165**:** 2117-2128.

Liu, Y. C., H. L. Du, P. C. Li, Y. T. Shen, H. Peng *et al.*, 2020 Pan-Genome of Wild and Cultivated Soybeans. Cell 182**:** 162-+.

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

Makarevitch, I., A. J. Waters, P. T. West, M. Stitzer, C. N. Hirsch *et al.*, 2015 Transposable Elements Contribute to Activation of Maize Genes in Response to Abiotic Stress. Plos Genetics 11.

Marrs, K. A., and V. Walbot, 1997 Expression and RNA splicing of the maize glutathione S-transferase Bronze2 gene is regulated by cadmium and other stresses. Plant Physiology 113**:** 93-102.

Martin, M., 2011 Cutadapt removes adapter sequences from high-throughput sequencing reads. 2011 17**:** 3.

Maruyama, K., Y. Sakuma, M. Kasuga, Y. Ito, M. Seki *et al.*, 2004 Identification of cold-inducible downstream genes of the Arabidopsis DREB1A/CBF3 transcriptional factor using two microarray systems. Plant Journal 38**:** 982-993.

Matsuoka, Y., Y. Vigouroux, M. M. Goodman, J. Sanchez G, E. Buckler *et al.*, 2002 A single domestication for maize shown by multilocus microsatellite genotyping. Proc Natl Acad Sci U S A 99**:** 6080-6084.

McGrath, J. M., A. M. Betzelberger, S. Wang, E. Shook, X. G. Zhu *et al.*, 2015 An analysis of ozone damage to historical maize and soybean yields in the United States. Proc Natl Acad Sci U S A 112**:** 14390-14395.

McMullen, M. D., S. Kresovich, H. S. Villeda, P. Bradbury, H. Li *et al.*, 2009 Genetic properties of the maize nested association mapping population. Science 325**:** 737-740.

Mills, G., K. Sharps, D. Simpson, H. Pleijel, M. Broberg *et al.*, 2018 Ozone pollution will compromise efforts to increase global wheat production. Glob Chang Biol 24**:** 3560-3574.

Montenegro, J. D., A. A. Golicz, P. E. Bayer, B. Hurgobin, H. Lee *et al.*, 2017 The pangenome of hexaploid bread wheat. Plant Journal 90**:** 1007-1013.

Moore, G., K. M. Devos, Z. Wang and M. D. Gale, 1995 Cereal genome evolution - grasses, line up and form a circle. Current Biology 5**:** 737-739.

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.

Nielsen, H., 2017 Predicting Secretory Proteins with SignalP, pp. 59-73 in *Protein Function Prediction: Methods and Protocols*, edited by D. Kihara. Springer New York, New York, NY.

Nielsen, H., J. Engelbrecht, S. Brunak and G. von Heijne, 1997 Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. Protein Eng 10**:** 1-6.

Puckette, M. C., Y. Tang and R. Mahalingam, 2008 Transcriptomic changes induced by acute ozone in resistant and sensitive Medicago truncatula accessions. BMC Plant Biol 8**:** 46.

Qian, L. W., K. Voss-Fels, Y. X. Cui, H. U. Jan, B. Samans *et al.*, 2016 Deletion of a Stay-Green Gene Associates with Adaptive Selection in Brassica napus. Molecular Plant 9**:** 1559-1569.

Quevillon, E., V. Silventoinen, S. Pillai, N. Harte, N. Mulder *et al.*, 2005 InterProScan: protein domains identifier. Nucleic acids research 33**:** W116-W120.

Rao, M. V., H. Lee and K. R. Davis, 2002 Ozone-induced ethylene production is dependent on salicylic acid, and both salicylic acid and ethylene act in concert to regulate ozone-induced cell death. Plant Journal 32**:** 447-456.

Ricci, W. A., Z. F. Lu, L. X. Ji, A. P. Marand, C. L. Ethridge *et al.*, 2019 Widespread long-range cis-regulatory elements in the maize genome. Nature Plants 5**:** 1237-1249.

Roessler, K., A. Bousios, E. Meca and B. S. Gaut, 2018 Modeling interactions between transposable elements and the plant epigenetic response: a surprising reliance on element retention. Genome Biol Evol 10**:** 803-815.

Salika, R., and J. Riffat, 2021 Abiotic stress responses in maize: a review. Acta Physiologiae Plantarum 43.

Schnable, P. S., D. Ware, R. S. Fulton, J. C. Stein, F. S. Wei *et al.*, 2009 The B73 Maize Genome: Complexity, Diversity, and Dynamics. Science 326**:** 1112-1115.

Sonnhammer, E. L., G. von Heijne and A. Krogh, 1998 A hidden Markov model for predicting transmembrane helices in protein sequences. Proc Int Conf Intell Syst Mol Biol 6**:** 175-182.

Springer, N. M., K. Ying, Y. Fu, T. Ji, C.-T. Yeh *et al.*, 2009 Maize Inbreds Exhibit High Levels of Copy Number Variation (CNV) and Presence/Absence Variation (PAV) in Genome Content. Plos Genetics 5.

Steige, K. A., B. Laenen, J. Reimegard, D. G. Scofield and T. Slotte, 2017 Genomic analysis reveals major determinants of cis-regulatory variation in Capsella grandiflora. Proceedings of the National Academy of Sciences of the United States of America 114**:** 1087-1092.

Stupar, R. M., and N. M. Springer, 2006 Cis-transcriptional variation in maize inbred lines B73 and Mo17 leads to additive expression patterns in the F1 hybrid. Genetics 173**:** 2199-2210.

Subramanian, A., P. Tamayo, V. K. Mootha, S. Mukherjee, B. L. Ebert *et al.*, 2005 Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. Proceedings of the National Academy of Sciences 102**:** 15545-15550.

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-+.

Swanson-Wagner, R., S. Eichten, S. Kumari, P. Tiffin, J. Stein *et al.*, 2010 Pervasive gene content variation and copy number variation in maize and its undomesticated progenitor. Genome Research 20**:** 1689-1699.

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.

Thatcher, S. R., O. N. Danilevskaya, X. Meng, M. Beatty, G. Zastrow-Hayes *et al.*, 2016 Genome-Wide Analysis of Alternative Splicing during Development and Drought Stress in Maize. Plant Physiology 170**:** 586-599.

Verhoeven, K. J. F., E. H. Verbon, T. P. van Gurp, C. Oplaat, J. F. de Carvalho *et al.*, 2018 Intergenerational environmental effects: functional signals in offspring transcriptomes and metabolomes after parental jasmonic acid treatment in apomictic dandelion. New Phytologist 217**:** 871-882.

Waldeck, N., K. Burkey, T. Carter, D. Dickey, Q. Song *et al.*, 2017 RNA-Seq study reveals genetic responses of diverse wild soybean accessions to increased ozone levels. BMC Genomics 18**:** 498.

Wang, B., M. Regulski, E. Tseng, A. Olson, S. Goodwin *et al.*, 2018a A comparative transcriptional landscape of maize and sorghum obtained by single-molecule sequencing. Genome Res 28**:** 921-932.

Wang, W. S., R. Mauleon, Z. Q. Hu, D. Chebotarov, S. S. Tai *et al.*, 2018b Genomic variation in 3,010 diverse accessions of Asian cultivated rice. Nature 557**:** 43-+.

Wang, X. C., M. Yang, D. Q. Ren, W. Terzaghi, X. W. Deng *et al.*, 2019 Cis-regulated alternative splicing divergence and its potential contribution to environmental responses in Arabidopsis. Plant Journal 97**:** 555-570.

Wedow, J. M., E. A. Ainsworth and S. Li, 2021a Plant biochemistry influences tropospheric ozone formation, destruction, deposition, and response. Trends in biochemical sciences.

Wedow, J. M., C. H. Burroughs, L. R. Acosta, A. D. B. Leakey and E. A. Ainsworth, 2021b Age-dependent increase in alpha-tocopherol and phytosterols in maize leaves exposed to elevated ozone pollution. Plant Direct 5.

Weirather, J. L., M. de Cesare, Y. Wang, P. Piazza, V. Sebastiano *et al.*, 2017 Comprehensive comparison of Pacific Biosciences and Oxford Nanopore Technologies and their applications to transcriptome analysis. F1000Research 6**:** 100-100.

Wittkopp, P. J., 2007 Variable gene expression in eukaryotes: a network perspective. Journal of Experimental Biology 210**:** 1567-1575.

Woodhouse, M. R., E. K. Cannon, J. L. Portwood, L. C. Harper, J. M. Gardiner *et al.*, 2021 A pan-genomic approach to genome databases using maize as a model system. Bmc Plant Biology 21.

Xu, E., L. Vaahtera, H. Hõrak, D. K. Hincha, A. G. Heyer *et al.*, 2015 Quantitative trait loci mapping and transcriptome analysis reveal candidate genes regulating the response to ozone in Arabidopsis thaliana. Plant Cell Environ 38**:** 1418-1433.

Yang, N., X. Xu, R. Wang, W. Peng, L. Cai *et al.*, 2017a Contributions of Zea mays subspecies mexicana haplotypes to modern maize. Nature Communications 8.

Yang, Z., R. Lu, Z. Dai, A. Yan, Q. Tang *et al.*, 2017b Salt-stress response mechanisms using de novo transcriptome sequencing of salt-tolerant and sensitive corchorus spp. genotypes. Genes 8**:** 226.

Yao, W., G. W. Li, H. Zhao, G. W. Wang, X. M. Lian *et al.*, 2015 Exploring the rice dispensable genome using a metagenome-like assembly strategy. Genome Biology 16.

Yendrek, C., G. Erice, C. Montes, T. Tomaz, C. Sorgini *et al.*, 2017a Elevated ozone reduces photosynthetic carbon gain by accelerating leaf senescence of inbred and hybrid maize in a genotype-specific manner. Plant Cell and Environment 40**:** 3088-3100.

Yendrek, C. R., T. Tomaz, C. M. Montes, Y. Y. Cao, A. M. Morse *et al.*, 2017b High-Throughput Phenotyping of Maize Leaf Physiological and Biochemical Traits Using Hyperspectral Reflectance. Plant Physiology 173**:** 614-626.

Yu, J. M., J. B. Holland, M. D. McMullen and E. S. Buckler, 2008 Genetic design and statistical power of nested association mapping in maize. Genetics 178**:** 539-551.

Zhao, Q., Q. Feng, H. Y. Lu, Y. Li, A. Wang *et al.*, 2018 Pan-genome analysis highlights the extent of genomic variation in cultivated and wild rice. Nature Genetics 50**:** 279-+.

Zou, C., K. Sun, J. D. Mackaluso, A. E. Seddon, R. Jin *et al.*, 2011 Cis-regulatory code of stress-responsive transcription in Arabidopsis thaliana. Proceedings of the National Academy of Sciences of the United States of America 108**:** 14992-14997.