The role of DNA content in shaping chromatin architecture and gene expression

Alfonso Carlos Barragán-Rosillo[1] [ID], Ricardo A. Chávez Montes[1,*] and Luis Herrera-Estrella[1,2,*] [ID]

[1]*Department of Plant and Soil Science, Institute of Genomics for Crop Abiotic Stress Tolerance, Texas Tech University, Lubbock, Texas, USA, and*

[2]*Unidad de Genómica Avanzada del Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional, Irapuato, Guanajuato, Mexico*

SUMMARY

Whole-genome duplication is an evolutionary force that drives speciation in all living kingdoms and is notably prevalent in plants. The evolutionary history of plants involved at least two genomic duplications that significantly expanded the plant morphology and physiology spectrum. Many important crops are polyploids, showing valuable features relative to morphological and stress response traits. After genome duplication, diploidization processes facilitate genomic adjustments to restore disomic inheritance. However, little is known about the chromatin changes triggered by nuclear DNA content alterations. Here, we report that synthetically induced genome duplication leads to chromatinization and significant changes in gene expression, resulting in a transcriptional landscape resembling a natural tetraploid. Interestingly, synthetic diploidization elicits only minor alterations in transcriptional activity and chromatin accessibility compared to the more pronounced effects of tetraploidization. We identified epigenetic factors, including specific histone variants, that showed increased expression following genome duplication and decreased expression after genome reduction. These changes may play a key role in the epigenetic mechanisms underlying the phenotypic complexity after tetraploidization in plants. Our findings shed light on the mechanisms that modulate chromatin accessibility remodeling and gene transcription regulation underlying plant genome adaptation in response to changes in genome size.

Keywords: *Arabidopsis thaliana*, ATAC-seq, chromatin accessibility, whole genome duplication, whole genome reduction, RNA-seq, gene transcription regulation.

BACKGROUND

Since Ohno provoked scientists with his hypothesis about evolution by gene duplication (Ohno, 1970), the study and the effect of polyploidization became essential topics of debate and research in evolutionary biology. Polyploidy, the condition of having multiple sets of chromosomes, is widespread among all kingdoms of life (Gallardo et al., 1999; Osada & Innan, 2008; Parks et al., 2018; Serres et al., 2009; Wolfe, 2015). Evolutionary studies in eukaryotic organisms suggest that polyploidization is essential in shaping phenotypic diversity and intriguingly seems more common in plants (Sierro et al., 2018; Wu et al., 2019). Some studies propose that at least two palaeopolyploidization events mark the evolution of the plant kingdom (DePamphilis et al., 2013; Wu et al., 2019). After polyploidization, re-diploidization starts as an adaptation process to restore bivalent chromosome pairing and disomic

inheritance (Li, McKibben, et al., 2021). The diploidization process comprehends genetic and epigenetic changes, including gene loss, gene silencing, gene subfunctionalization, and reconfiguration of chromosome architecture (Li, McKibben, et al., 2021; Wolfe, 2001; Zhang et al., 2021).

Most land plants are paleopolyploids, resulting from a diploidization process, suggesting that polyploidization and genome reduction can modulate phenotype plasticity. Furthermore, the selection of polyploid populations over their diploid relatives correlates with global climatic cataclysms and mass extinctions (Van de Peer et al., 2017), suggesting that stress responses, in general, are selective factors in the establishment of polyploid populations, supporting the assumption that synthetic polyploidization can be helpful for plant domestication and crop breeding (Salman-Minkov et al., 2016). Indeed, many important crops are paleopolyploids (Gardiner et al., 2018; Guo

et al., 1996) or recent polyploids (Cenci et al., 2019; Oustric et al., 2019; Yuan et al., 2019). Two different classes of polyploid organisms can emerge from a polyploidization event: (1) failure in genome reduction at meiosis can produce unreduced gametes, and the mix of two unreduced gametes from the same species generates autopolyploid individuals; and (2) fertilization between unreduced gametes from two different but related species will produce allopolyploid individuals. In crops, we have examples of both cases: potato (Spooner et al., 2010), strawberry (Yuan et al., 2019), and sugar cane (Zhang et al., 2018) are autopolyploids, while wheat (Ramírez-González et al., 2018), cotton (Yoo & Wendel, 2014), and canola (Chalhoub et al., 2014) are allopolyploids.

Despite the importance of whole-genome duplication (WGD) in plant evolution, the changes in chromatin architecture and gene expression modifications triggered by polyploidization still need to be characterized. Analysis using fluorescent markers indicates that centromeres are closer to each other in tetraploid cells than in diploids (Sas-Nowosielska & Bernas, 2016), suggesting genome heterochromatinization after WGD. Global transcriptional responses to WGD in Arabidopsis have been proposed to be tissue- and developmental-stage-dependent (Yu et al., 2010). The enrichment analysis of Gene Ontology (GO) categories for differentially expressed genes (DEGs) in tetraploid leaves, compared to diploid leaves, identified a significant overrepresentation of biological processes associated with hormonal signaling, stress response, light perception, ion transport, cell wall remodeling, and carbohydrate metabolism (Del Pozo & Ramirez-Parra, 2014). These results indicate that tetraploidy may promote greater physiological plasticity and metabolic efficiency, potentially enabling plants to better adapt to varying environmental conditions. Studies on Arabidopsis have consistently shown that WGD enhances tolerance to abiotic stresses such as high salinity and drought, interestingly associated with increased leaf potassium accumulation (Chao et al., 2013; Corneillie et al., 2019).

Speciation through diploidization after WGD involves three different genetic scenarios: loss of functions because of mutation (pseudogenization) (Chu et al., 2021), acquisition of new gene functions (neofunctionalization) (Teshima & Innan, 2008) with almost unaltered homologous loci, and transcription divergence in gene expression patterns between homologous loci (subfunctionalization) (Rastogi & Liberles, 2005). However, the genetic bases and molecular changes caused by whole genome reduction (WGR) still need to be better understood.

Here, we describe the consequences of WGD and WGR on chromatin accessibility and the global transcriptional landscape in *Arabidopsis thaliana* roots. We found that synthetic tetraploids acquire chromatin accessibility and gene expression features similar to those of natural tetraploids. At the same time, WGR promotes a discrete molecular response, suggesting that new diploids retain many features gained after genome doubling. We also identified 95 transcriptional regulators with enhanced transcription in tetraploids but with pre-existing chromatin accessibility in diploids that could activate changes in chromatin architecture and gene expression in the tetraploid state. Finally, we discovered 11 overexpressed and six silenced genes with epigenetic functions, suggesting a substantial role in establishing chromatin organization after WGD in Arabidopsis.

RESULTS

WGD promotes heterochromatinization

To characterize the molecular effects after WGD and WGR, we used the natural diploid Col-0 ecotype (hereafter named CC), the synthetic tetraploid Col-0 4x (CCCC), generated by colchicine treatment, the natural autotetraploid ecotype Wa-1 (WWWW), and the synthetic diploid Wa-1 2x (WW) obtained by chromosome elimination with a modified histone variant CENH3 (Ravi & Chan, 2010). We confirmed the ploidy level of our lines using flow cytometry (Figure S1). As anticipated, the diploid line CC shows peaks at the position of $2n$, $4n$, $8n$, and $16n$ nuclear content, whereas the synthetic tetraploid line CCCC lacks the $2n$ peak (indicated by the black arrow in Figure S1b), as does the natural tetraploid Wa-1. The peak at the $2n$ position in the derived line WW (red arrow in Figure S1b) confirms its diploid karyotype. Seeds from flow cytometry-verified plants were used in all subsequent experiments. We compared plant features like root architecture in 10-day-old seedlings to determine whether ploidy impacts plant development. We found tetraploids had more and longer lateral roots than diploid plants, independent of the ecotype (Figure S1a). After 30 days of growth, the root biomass in tetraploid plants was significantly higher than that of diploids (Figure S1d). We also confirmed the previously reported larger flower size phenotype of tetraploid plants (Figure S1c). These findings suggest that the differences associated with ploidy levels are mainly attributable to modifications in gene regulatory processes, which influence gene transcription and downstream molecular networks.

Given the importance of the root system in sessile organisms to nutrition acquisition and the phenotypic effect after tetraploidization, we selected roots as our tissue of interest. A hydroponic system was used to grow seedlings and determine chromosome accessibility and global gene expression (Figure S2a, see "Methods" section). We grew seedlings at 20°C in an 18-h light/6-h dark photoperiod. After 10 days of growth, seedlings were harvested, and the entire root system was dissected to isolate nuclei for ATAC-seq and whole root tissue for RNA-seq
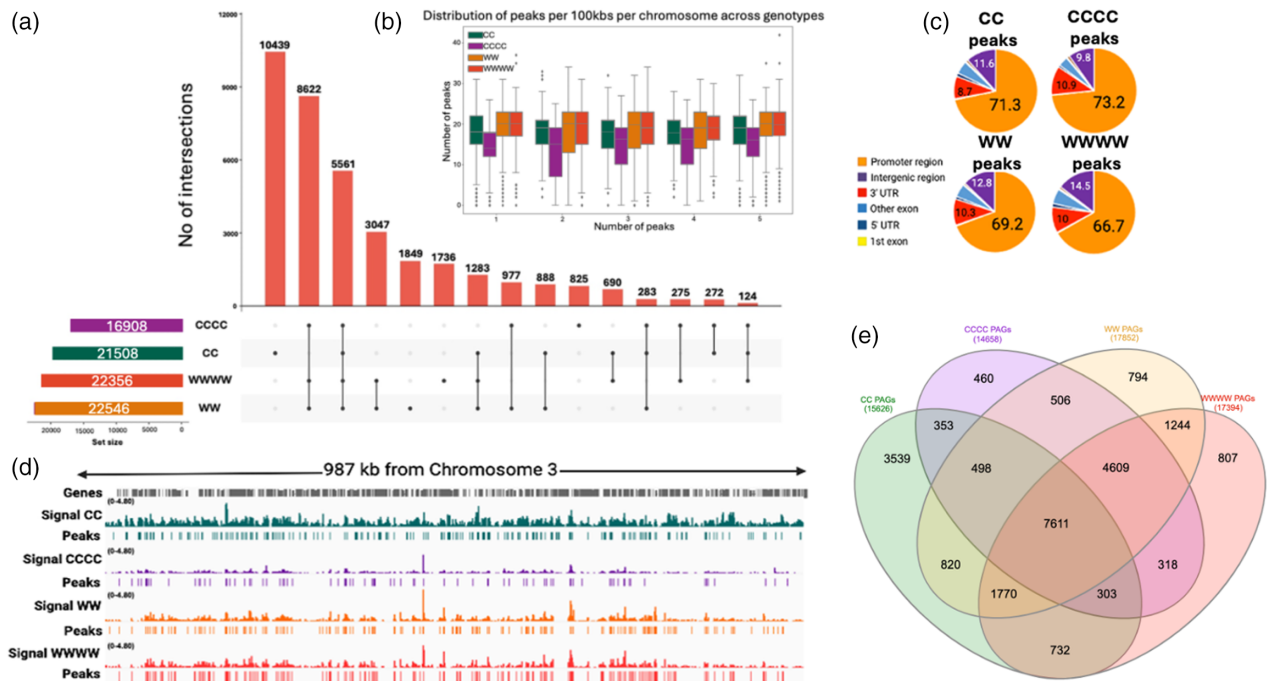
Figure 1. Genome accessibility after WGD and WGR.
(a) Upset plot with the number of shared and unique genomic accessible regions; most of the accessible regions detected were shared between genotypes.
(b) Bars plot of the number of peaks distribution per 100 kb by chromosome across genotypes; major changes in the number of peaks were observed after WGD; however, an incipient change was observed after WGR.
(c) Graphic pays with annotated peaks by genomic region; most of the open chromatin regions were detected at promoter regions.
(d) Genome browser display of accessibility signal (histogram) and peaks (rectangles); green represents CC, violet represents CCCC, orange represents WW, and red represents WWWW genotype.
(e) Venn diagram of CC, CCCC, WW, and WWWW PAGs. PAG, peak-associated gene; WGD, whole-genome duplication; WGR, whole-genome duplication.

experiments. Nuclei were isolated from two independent biological replicates to prepare ATAC-seq libraries (Figure S2a). Due to the low-quality and incomplete sequence of the Wa-1 genome, we aligned all samples to the well-curated and high-quality Col-0 reference genome of *Arabidopsis thaliana* to ensure accuracy and consistency in downstream analyses. We achieved an average alignment rate of 98.8% to the Col-0 Arabidopsis nuclear genome, indicating high-quality sequencing data with minimal off-target reads. The FRiP (Fragment Reads in Peaks) scores ranged from 20 to 50%, demonstrating that a substantial proportion of the sequenced fragments mapped to accessible chromatin regions, consistent with high-confidence peak calling. The number of fragment reads within peaks was high across the different ploidy levels analyzed. Specifically, we recovered 66 million fragment reads for the diploid sample (CC), 37.7 million fragment reads for the tetraploid sample (CCCC), 39.2 million fragment reads for the diploid wild-type (WW), and 36.6 million fragment reads for the tetraploid wild-type (WWWW) (Table S1). These results highlight a robust signal capture for chromatin accessibility and suggest that the data are well suited for downstream chromatin structure and regulatory element activity analyses. An MDS plot of

sequenced libraries showed a clear effect of WGD and WGR on global chromatin accessibility in Arabidopsis (Figure S2b). Next, we identified open chromatin regions (peaks) using MACS2 (Zhang et al., 2008) (see "Methods" section). We detected 36 871 high-confidence and unique accessibility regions across all samples (Figure 1a). Among all peaks detected, 21 508 were found in CC, 16 908 in CCCC, 22 546 in WW, and 22 356 in WWWW. Among these chromatin accessibility peaks, 10 439 were exclusive for CC (28.3% of the total unique peaks), and 5561were common in all genotypes (15% of the total unique peaks), whereas 1849 were exclusive to WW, 1736 to WWWW, and 825 to CCCC (Figure 1a; Table 1). Among accessibility peaks, 8622 were shared between CCCC, WWWW, and WW, showing similarities in chromatin organization level between both tetraploids i.e., CCCC and WWWW, and the new diploid WW (Figure 1a; Table 1).

To determine if genomic accessibility has a chromosome bias, we examined the number of peaks per 100 kb in each chromosome. We observed genomic accessibility differences between chromosomes after tetraploidization and diploidization (Figure 1b). To better understand the bias of genomic accessibility across all chromosomes, we determined the percentage of loss or gain peaks between

Table 1 Summary of peaks, PAGs, and DEGs resulted from the comparison between four different genotypes CC (Col-0), CCCC (Col-0 new tetraploid), WW (Wa-1 diploidized), and WWWW (Wa-1 old tetraploid)

Genotype	Peaks	PAGs	Comparison	Common PAGs	upPAGs	downPAGs	DEGs	upDEGs	downDEGs	PAGs/DEGs
CC	21 508	15 626	*CCCC* vs. *CC*	8765	1556	227	6937	3555	3382	2330
CCCC	16 908	14 658	*WWWW* vs. *CC*	10 416	1856	285	5792	2640	3152	2428
WW	22 546	17 852	*WW* vs. *WWWW*	15 234	1204	52	1015	259	756	503
WWWW	22 356	17 394	*WWWW* vs. *CCCC*	12 841	298	525	909	537	372	415
			WW vs. *CC*	10 699	2278	127	4012	2292	1720	1547

DEG, differentially expressed gene; PAG, peak-associated gene.

chromosomes among all the samples. We observe that tetraploidization generates the loss of chromatin accessibility by around 20%, but chromosome 2 shows chromatin accessibility decreased by around 28%, being the most affected chromosome after tetraploidization. After diploidization, we observe a gain in chromatin accessibility (Figure S3a,b). We also plotted the distribution of peaks per 100 kb per chromosome to compare chromatin accessibility across all genotypes. We observed that centromeres were the genomic regions with lower chromatin accessibility (Figure S3c,g). These results suggest that WGD in the synthetic tetraploid (CCCC) resulted in genome compaction compared to the CC diploid, while WGR promoted increased chromatin accessibility in the synthetic diploid (WW) concerning the natural tetraploid (WWWW).

Using Chipseeker (Yu et al., 2015) (see "Methods" section), we determined the location of each peak relative to gene annotations. Most open chromatin regions, 66 to 73%, were in promoter regions, 9.8 to 14.5% were in intergenic regions, and 8.7 to 10.9% were in 3′ UTRs (Figure 1c). Although these results are globally similar among the different lines tested, there are differences in the location of the chromatin accessibility peaks between CC, CCCC, WW, and WWWW. A discrete 987 kb genomic region from chromosome 3 illustrates the diversity of genomic regions with different chromatin accessibility, showing common peaks among the genotypes, exclusive peaks per genotype, ecotype-dependent peaks, and ploidy-dependent peaks (Figure 1d).

WGD regulates genomic accessibility in genes related to stress and uridine metabolism

To examine the effect of a recent synthetic tetraploidization event on chromatin accessibility, we compared peak-associated genes (PAGs) between the CC and CCCC genotypes. We defined PAGs as the genes with the closest transcriptional start site to an accessible chromatin region. We detected 8765 PAGs shared between CC and CCCC, 6861 and 5893 exclusives to CC and CCCC (Figure 1e; Table 1). Gene ontology enrichment analysis of PAGs shared by CC and CCCC reflect the response to different stimuli, including responses to chemical and abiotic stimuli

(Table S2; File S1). The 6861 PAGs exclusive to CC, and thus lost in CCCC, were enriched in the 'DNA-binding transcription factor activity' and 'RNA-DNA hybrid ribonuclease activity' categories, among others (Table S3; File S1). The 5893 PAGs exclusive to CCCC, gained after polyploidization, were enriched in 'cellular process,' 'organic substance biosynthetic process,' and 'cellular biosynthetic process' categories, among others (Table S4; File S1).

We hypothesized that differences between CC and CCCC involve peaks exclusive to each genotype and common to both genotypes but with differential chromatin accessibility. Thus, we used csaw to identify differential chromatin accessibility regions (DARs) (Barragán-Rosillo et al., 2021; Lun & Smyth, 2015; Reske et al., 2020) (see "Methods" section). We identified, using Chipseeker (see "Methods" section), 1556 peak-associated genes with gain differential chromatin accessibility (upPAGs) and 227 peak-associated genes with reduced differential chromatin accessibility (downPAGs) in CCCC compared to CC (Table 1). In agreement with our hypothesis, 1102 (70%) of all upPAGs are part of the 8765 shared PAGs, but with higher chromatin accessibility in CCCC. These 1102 upPAGs were enriched in categories such as 'biological regulation,' 'response to stimulus,' 'developmental process,' 'response to chemical,' 'response to abiotic stimulus,' 'regulation of cellular process,' and 'regulation of metabolic process' (Table S5; File S1). These enrichments suggest that genome duplication may broadly rewire regulatory networks, enhancing the organism's ability to perceive and respond to environmental cues while modulating developmental and metabolic pathways. Further investigation into the functional roles of these upPAGs could provide deeper insights into how polyploidy contributes to increased phenotypic plasticity and adaptive potential.

Synthetic tetraploidization immediately acquires features present in a natural tetraploid

To determine the epi-transcriptomic changes caused by tetraploidization, we compared chromatin accessibility and gene expression between the synthetic tetraploid CCCC and the natural tetraploid WWWW with the natural diploid

CC as a reference. We found 17 394 PAGs in WWWW, 14 658 in CCCC, and 15 626 in CC (Figure 1e; Table 1; File S1). We found 966 PAGs exclusives to CCCC, which could represent temporary changes in chromatin accessibility during the early stages of a tetraploidization event or changes in chromatin accessibility that are Col-0 ecotype specific. We also found 2051 PAGs exclusive to WWWW, representing changes in chromatin accessibility regions in a natural tetraploid that could represent long-term adaptations or chromatin accessibility territories specific for the Wa-1 ecotype. The 4927 PAGs shared by CCCC and WWWW, absent in CC, are ecotype-independent and could represent the factual response to tetraploidization (Figure 1e). The 966 PAGs exclusive to CCCC were enriched in the 'triplet codon-amino acid adaptor activity' and 'molecular adaptor activity' categories. The 2051 PAGs exclusive to the WWWW were only in one enriched category, 'signaling receptor binding', showing that chromatin is remodeled after tetraploidization to maintain tight regulation of receptor binding as part of the adaptation process. The 4927 PAGs shared by CCCC and WWWW were enriched in GO categories such as 'pyrophosphatase activity,' 'hydrolase activity,' 'nitrogen compound metabolic process,' and 'organonitrogen compound biosynthetic process,' among others (Table S7; File S1). Our findings suggest that genome duplication in *Arabidopsis* leads to altered chromatin accessibility, particularly in nitrogen and phosphate metabolism genes. These changes reflect an adaptive molecular mechanism to balance the acquisition and utilization of these essential nutrients during developmental processes and metabolic demands.

Since changes in chromatin accessibility are often associated with transcriptional activation or repression of nearby genes (Barragán-Rosillo et al., 2021; Tsompana & Buck, 2014), we explored the differences in transcript abundance in response to alterations in genome ploidy and their correlation to the variations in chromatin accessibility regions in CC and CCCC.

Gene expression changes after WGD correlate with open chromatin regions

To determine the relationship between changes in chromatin accessibility and gene expression between diploids and tetraploids, we performed a transcriptomic analysis of RNA extracted from the roots of seedlings grown in the same conditions as those used for the ATAC-seq analysis. We used three biological replicates for each genotype (Table S8). RNA-seq reads from the different libraries were aligned to the Col-0 genome sequence, counted, normalized, and used to identify DEGs between CC and CCCC. Using htseq and DESeq2 (FDR ≤ 0.05, $\log_2$FC ≤ -1 or ≥ 1; see "Methods" section), we identified 6937 Differential Expressed Genes (DEGs) in the comparison between CCCC and CC, of which 3555 had higher transcript levels

(upDEGs) and 3382 had lower transcript levels (downDEGs) (Figure 2a; Table 1; File S2). upDEGs in CCCC were enriched in 'RNA modification,' 'response to ATP,' 'cellular response to DNA damage stimulus,' 'response to organophosphorus,' 'DNA repair,' 'regulation of root morphogenesis,' and 'chromatin silencing complex' GO categories, among others (Table S9; File S1). DownDEGs in CCCC were enriched in GO categories 'photosynthesis,' 'light reaction,' 'generation of precursor metabolites and energy,' 'thylakoid,' 'chloroplast,' 'plastid thylakoid,' 'plastid envelope,' and 'chloroplast thylakoid membrane' and 'obsolete protein-chromophore linkage' categories (Table S10; File S1). To identify the relationship between open chromatin regions and transcript levels after WGD, we determined overlaps between DEGs (6937 genes) and PAGs in CCCC (14 658 genes). We found that 3747 (54%) DEGs were also represented in the CCCC PAG list (Figure 2b). To identify the correlation between upDEGs responsive to a recent synthetic tetraploidization (CCCC versus CC) with chromatin accessibility gained after tetraploidization in CCCC, we compared the PAGs exclusive to CCCC with the upDEGs in CCCC. We found that among the 5893 PAGs present in CCCC but absent in CC, 820 were identified as upDEGs acquired after tetraploidization (Figure 2c) with GO enrichment categories related to photosynthesis and plastid organization among others (File S1). We also determined 597 downDEGs associated with acquired genome accessibility after tetraploidization (Figure 2d). GO enrichment analysis shows that genes differentially expressed associated with the gain in chromatin accessibility in CCCC are enriched in categories such as 'nitrogen compound metabolic process,' 'nucleic acid metabolic process,' 'RNA metabolic process,' 'pyrophosphatase activity,' among others (File S1). These results indicate that following WGD, the global transcriptional response associated with new open chromatin regions is primarily associated with upregulating genes related to nucleic acid and nitrogen metabolic processes. In contrast, the downregulation of genes in regions with increased chromatin accessibility after tetraploidization is primarily linked to photosynthesis.

It should be noted that the response to WGD observed in CCCC may differ from the transcriptomic landscape of the adapted natural tetraploid genome WWWW. We compared the natural diploid CC with the natural tetraploid WWWW to understand how an adaptive process affects the transcriptomic landscape after WGD. Differential gene expression between WWWW and CC should also reflect the effects of tetraploidization independently of the ecotype. We identified 5792 DEGs (Figure 2e; Table 1), of which 3813 (65%) were associated with open chromatin regions (Figure 2f); 2640 were upDEGs and 3152 downDEGs (Table 1). To identify upregulated genes (upDEGs) uniquely associated with chromatin accessibility in
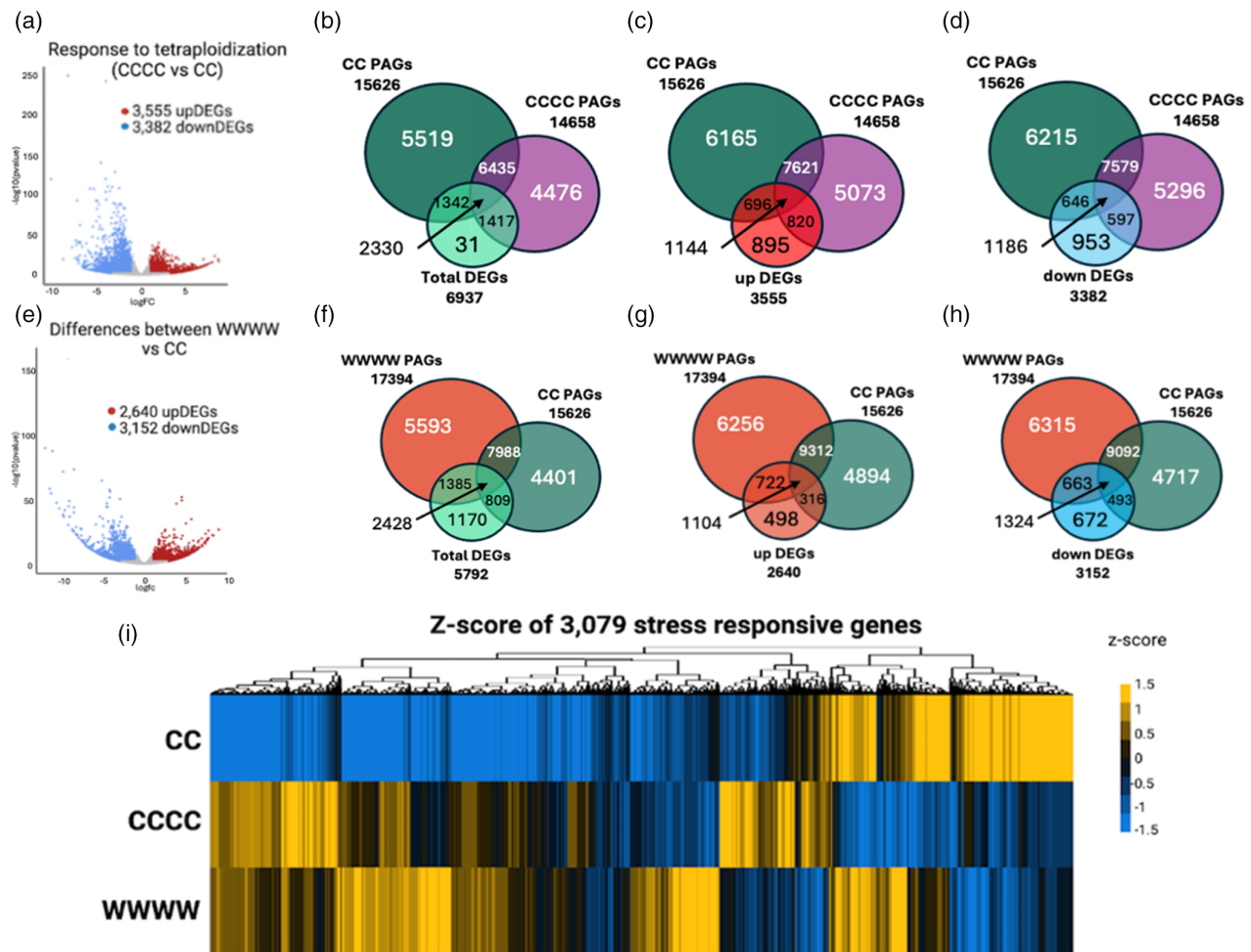
Figure 2. Transcriptional response to recent and old tetraploidization.
(a) Volcano plot of differentially expressed genes after tetraploidization, CCCC versus CC.
(b) Venn diagram of total DEGs, CC, CCCC PAGs.
(c) Venn diagram of upDEGs and CC, CCCC PAGs.
(d) Venn diagram of downDEGs and CC, CCCC PAGs.
(e) Volcano plot of differentially expressed genes in old tetraploidization, WWWW versus CC.
(f) Venn diagram of DEGs in response to old tetraploidization and WWWW, CC PAGs.
(g) Venn diagram of upDEGs of old tetraploidization and WWWW, CC PAGs.
(h) Venn diagram of downDEGs and WWWW, CC PAGs.
(i) Heat map of *z*-scores of 3079 stress-responsive Arabidopsis genes; both tetraploids show a similar transcriptional landscape. DEG, differentially expressed gene; PAG, peak-associated gene.

WWWW, we compared the 2640 upDEGs and 6978 exclusive PAGs in WWWW in contrast to CC. Specifically, we focused on the 4927 PAGs shared by tetraploids plus the 2051 exclusive PAGs found in the WWWW genotype (Figure 1e). Through this comparison, we identified 722 upDEGs linked with alterations in chromatin accessibility that are specific to the WWWW genotype (Figure 2g). The GO categories enriched in this set of genes were like those found for the recent tetraploidization, such as 'nitrogen compound metabolic process' and 'nucleic acid metabolic process,' however, we recovered other categories, such as 'response to stress' and 'response to abiotic stimulus'

(File S1), suggesting that after an adaptation process, tetraploidization can promote some responses that make them more competitive to their parental diploids. We also found 663 downDEGs that correlate with chromatin-accessible regions exclusive for WWWW (Figure 2h); these genes were enriched in GO categories 'secondary metabolic process,' 'photosynthesis, light harvesting in photosynthesis 1,' and 'response to light stimulus' reflecting similarities and differences between synthetic and natural tetraploids (File S1).

It has been suggested that genes with increased transcript levels in Arabidopsis tetraploids are associated with

stress responses. Therefore, we generated a heat map of gene expression *z*-scores from our RNA-seq data, which included 3079 genes in the Arabidopsis Stress Responsive Transcription Factor database (Sowdhamini et al., 2009). The data revealed that both tetraploids exhibited an increased transcript level of stress-responsive TFs compared to CC (Figure 2i), suggesting that the activation of stress-responsive genes is ecotype-independent. It occurs shortly after a WGD event and continues throughout the long-term stabilization of the polyploidization.

We then recognized the similarities and differences in gene expression associated with genome accessibility between synthetic and natural tetraploids. We found 3747 and 3813 DEGs associated with PAGs (DEGPAGs) in CCCC and WWWW, respectively (Figure 2b,f). We found 2428 DEGPAGs shared by both tetraploids, potentially representing the ecotype-independent loci affected by tetraploidization. Some of these genes have higher expression in WWWW than CCCC and vice versa, suggesting that changes in accessibility and gene expression in tetraploids are modulated during the stabilization of the WGD event over time. This indicates that a genomic duplication causes a wide range of genomic changes, some occurring early after the WGD event, which become exacerbated or attenuated over time. A GO categories enrichment analysis of the 2428 DEGPAGs shared by the two tetraploids showed that the transcriptional response associated with changes in chromatin accessibility correlated with WGD includes response to abiotic stimulus, metabolic processes, and photosynthesis categories (Table S11; File S1).

WGR maintains the transcriptional profile of tetraploids

Our results showed that tetraploidization activates changes in chromatin accessibility and gene expression, but can the reverse process occur during the diploidization of a tetraploid? To examine the effect of a recent synthetic diploidization of a natural tetraploid, we compared the analyzed peaks of open chromatin in WWWW and WW and their correlation with changes in gene expression. We found that the majority of PAGs were shared by the two genotypes (15 234 out of over 17 000 PAGs for each genotype) (Table 1), 2160 were WWWW-specific, representing regions of open chromatin that were lost during the WGR process, and 2618 PAGs were only present in WW that represent open chromatin regions gained during the diploidization process (Figure 1e). WWWW exclusive PAGs were enriched in only two GO categories 'catalytic activity on DNA,' and "exonuclease activity" (Table S12; File S1). Intriguingly, no enriched GO categories could be found for the 2618 PAGS exclusive to WW. To better understand the effect of WGR at the gene expression level, we performed an RNA-seq assay on the new diploid WW and compared it to the WWWW transcriptome. We found 1015 DEGs, 503 shared by both genotypes, 95 exclusives to WW, and 66

exclusives to WWWW (Figure 3a,b; Table 1); from those, 259 were upDEGs and 756 downDEGs in the WW versus WWWW comparison (Figure 3a–c; Table 1; File S2). The list of upDEGs was enriched in categories such as 'xyloglucan: xyloglucosyl transferase activity,' '*O*-methyltransferase activity,' 'transmembrane transporter activity,' 'secondary metabolic process,' 'phenylpropanoid biosynthesis,' and 'Glucosinolate biosynthesis (from aromatic am*ino acid*)' (Table S13; File S1). DownDEGs were enriched categories such as 'photosynthesis,' 'metabolic pathway,' 'Carbon fixation,' 'carbon metabolism,' and 'pentose phosphate pathway' and several others correlated to responses to different types of stresses (Table S14; File S1). To determine the role of genome accessibility in differential gene expression after genome reduction, we compared PAGs with upDEGs and downDEGs. We found that 159 out of 259 upDEGs in WW were linked to at least one chromatin-accessible region (Figure 3c), and 439 of 756 downDEGs were also associated with an open chromatin zone in WW (Figure 3d). Since DNA metabolism and stress categories were enriched during the tetraploidization event, the categories enriched in the WW comparison suggest that the WGR process only partially loses the features acquired after WGD.

We performed a comparative analysis to identify genes whose expression correlates positively with genome size – upregulated following tetraploidization and downregulated after diploidization. Specifically, we intersected the upregulated differentially expressed genes (upDEGs) identified after WGD (WWWW versus CC) with the downregulated differentially expressed genes (downDEGs) observed after whole-genome reduction (WGR, WW versus WWWW). This approach highlights candidate genes whose expression levels are directly associated with genomic alterations and polyploidization. We found 90 genes that were upregulated in WWWW and downregulated in WW (Figure 3e). Six GO categories were enriched for this set of genes, all relative to "nitrile" (Table S15; File S1). A heat map of *z*-scores showed that the expression of these 90 genes increases in both tetraploids WWWW and CCCC (Figure 3f). To better visualize the magnitude of the changes in gene expression, we generated a heat map of the log_2 fold change of these 90 genes. We observed a similar pattern; an incipient activation of these 90 genes was observed after recent tetraploidization, and a clearer activation was observed in old tetraploids compared to the response to diploidization (Figure 3g). A raincloud plot (half violin, boxplot, and scatter plot) reflects graphically how those genes respond depending on genome ploidy (Figure 3h). We found that 66 out of the 90 genes (File S1) were associated with open chromatin regions.

To get a general view of the transcriptional profile, we generated *z*-score heat maps of DEGs specific to each tetraploid comparison: 2928 for recent tetraploidization
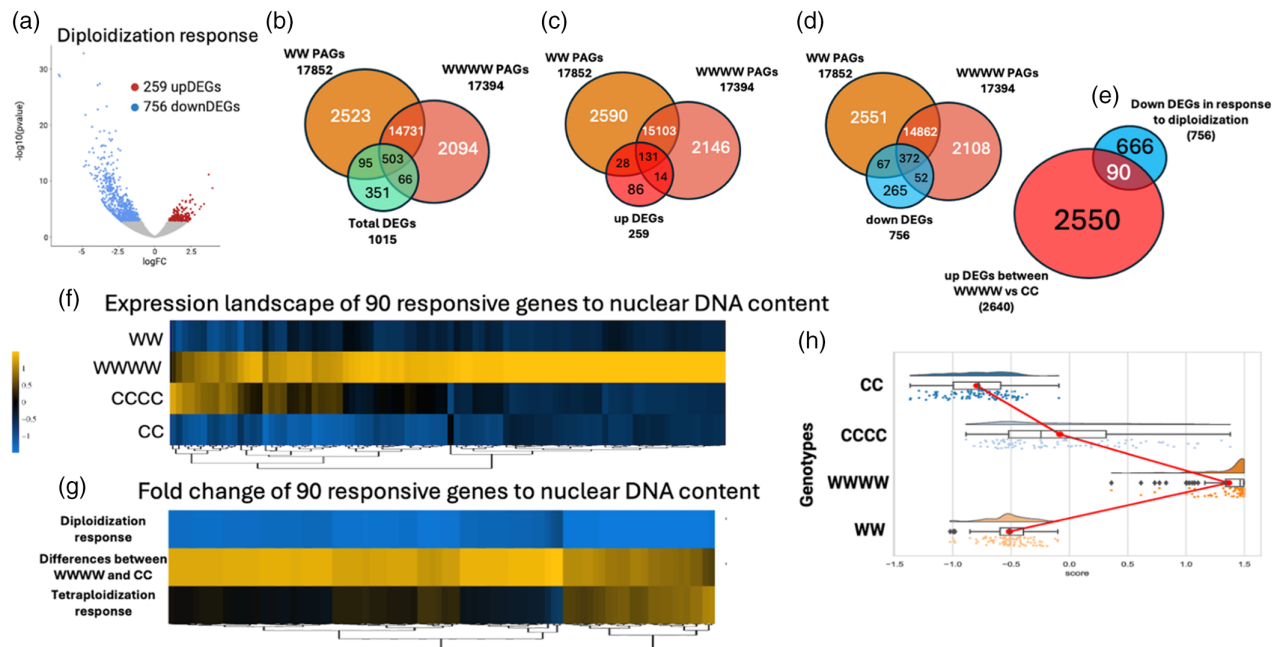
Figure 3. Genomic response to diploidization.
(a) Volcano plot of differentially expressed genes in response to diploidization.
(b) Venn diagram between total DEGs and WWWW, WW PAGs.
(c) Venn diagram between upDEGs and WWWW, WW PAGs.
(d) Venn diagram between downDEGs and WWWW, WW PAGs.
(e) Venn diagram of downDEGs responsive to diploidization and upDEGs responsive to old tetraploidization.
(f) Heat map of *z*-scores of 90 genes overlapped, showing similarities between diploid and tetraploid lines.
(g) Heat map of change index of 90 overlapped genes in (f), showing that new tetraploidization acquires some of the tetraploid transcriptomic features.
(h) Raincloud plot of 90 overlapped *z*-scores of genes in (f) of all genotypes, showing graphically the transcriptional dynamics relative to changes in genome size. The Raincloud plot includes half of the violin plot, box plot, point events, and the median value union between our four genotypes with the red line. DEG, differentially expressed gene; PAG, peak-associated gene.

response, 1783 for an old tetraploidization, and one for 4009 DEGs shared in response to tetraploidization (Figure 4a,c). The heat maps of the shared DEGs showed that (i) the transcriptional profile of CCCC is more like those of WWWW and WW than CC, and (ii) the WW profile is more like WWWW than CC (Figure 4c). The relationship between differentially expressed transcription factors and their post-tetraploidization role requires further understanding, as it is already well established that genomic accessibility is closely associated with DNA-binding sites for transcription factors (Sijacic et al., 2018).

The phenomenological effects of WGD correlate with the expression level of specific TFs

We hypothesized that transcription factors and the epigenetic system regulate the transcriptional response to synthetic tetraploidization (CCCC versus CC). To test this hypothesis, we employed a two-step approach. First, we used the DIANE algorithm (Cassan et al., 2021) to cluster DEGs responsive to tetraploidization. We identified nine clusters based on their transcriptional profiles. Cluster 1, consisting of 242 genes, exhibited high expression levels in CC but low in CCCC, WW, and WWWW; the enriched GO terms for cluster 1 were related to photosynthesis (Figure 5a; File S3). Cluster 2, containing 743 genes, had high expression levels in CC but low expression in the other genotypes with no enriched GO category relative to. Cluster 3, comprising 419 genes with higher expression in CC and low in CCCC, WW and WWWW; this cluster was enriched in categories related to photosynthesis and plastid organization GO (Figure 5a; File S3). Cluster 4, with 570 genes, showed low expression in CC, higher expression in CCCC, and low expression in WW and WWWW but higher than in CC; this cluster was enriched in such as 'Ribosome biogenesis' and 'Ribosome assembly' GO categories (Figure 5a; File S3). Clusters 5 and 6 included 564 and 2789 genes, respectively, exhibiting low expression in CC and higher in CCCC, WW, and WWWW. Cluster 5 did not exhibit any enrichment in a specific category, whereas cluster 6 was enriched in 'DNA repair' and 'Cellular response to DNA damage' (Figure 5a; File S3). Cluster 7, containing 496 genes, had high expression in CC, low in CCCC and WW, and high in WWWW, and was enriched in GO categories relatives to light intensity response and
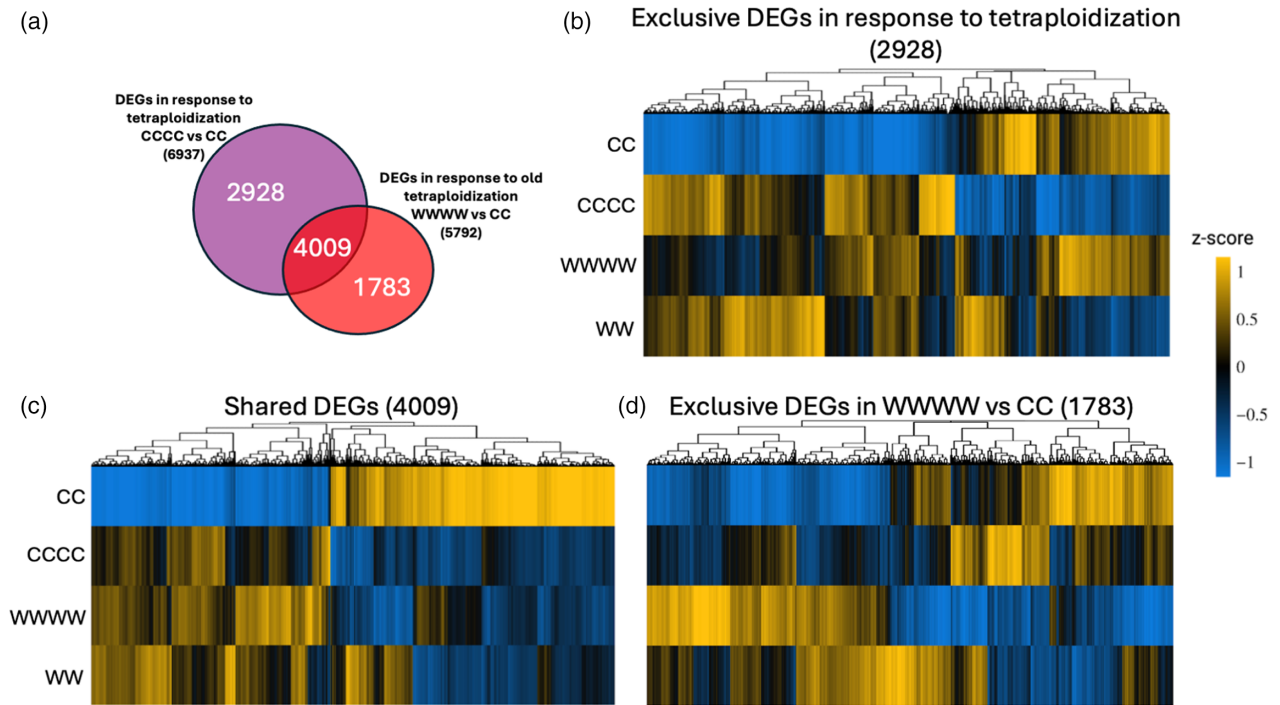
Figure 4. WGR plants reflect similarities with tetraploid plants.
(a) Venn diagram overlapped DEGs in recent and ancient tetraploidization.
(b) Heat map of *z*-scores of 2928, acquired at the WGD event but lost over time; the transcriptional landscape shows similarities between CCCC, WWWW, and WW.
(c) The Heat map of 4009 *z*-scores represents the responsive DEGs acquired at the WGD event and is still present after the adaptation process in WWWW, showing similarities at the transcriptional level between CCCC, WWWW, and WW.
(d) The Heat map of 1783 DEGs acquired after the adaptation response is ecotype-dependent. DEG, differentially expressed gene; WGD, whole-genome duplication; WGR, whole-genome duplication.

starch biosynthetic processes (Figure 5a; File S3). Cluster 8 showed high expression in CC and low in CCCC, WW, and WWWW, but slightly higher in WW, similar to cluster 2; cluster 8 did not exhibit any enrichment category. Lastly, cluster 9, with 462 genes, showed high expression levels in CC, but low expression, although not zero, in the other genotypes. These genes were enriched in categories related to cell wall modification and water transport among different categories (Figure 5a; File S3).

To recognize the TFs that might initiate and regulate the transcriptional and chromatin architecture changes after tetraploidization (i.e., CCCC versus CC), we use the set of 1717 genes from the PlantTFDB platform. We select those with higher expression in CCCC than CC but associated with pre-existing genome accessibility in CC (Figure 5b). In other words, TFs that increase expression levels during tetraploidization but do not require changes in chromatin accessibility to increase their expression and that could activate the expression of genes necessary for chromatin remodeling or for recruiting the chromatin remodeling machinery to alter chromatin architecture and gene expression after the WGD event. We identified 95 TFs with higher transcript levels in CCCC than CC but with similar chromatin accessibility in

both genotypes (upDEGPAGTFs) (Figure 5c). Then, we built a transcriptional regulatory network comprising 15 distinct modules, including 894 genes and 19 regulators (Figure 5d; File S4). To better understand the regulatory effect of these 95 upDEGPAGTFs, we delve into the major modules that shape the gene regulatory network (Figure 5d). The first module consists of two nodes. Node 1 is a super node containing 12 different transcription factors, including *BASIC PENTACYSTEINE 6, MYB DOMAIN PROTEIN 112, TCP DOMAIN PROTEIN 10, HEAT SHOCK TRANSCRIPTION FACTOR A3, PHYTOCLOCK 1, AT2G41710, WRKY DNA-BINDING PROTEIN 23, HEAT SHOCK TRANSCRIPTION FACTOR B2A, EIN2 NUCLEAR ASSOCIATED PROTEIN 1, HOMEOBOX PROTEIN 16, WUSCHEL RELATED HOMEOBOX 13,* and *CONSTANS-LIKE 9*, potentially regulating 77 genes (File S4). The second node is represented by *AT5G25475*, an AP2/B3-like transcription factor that regulates 37 genes (File S4). The second module is enriched with two terms, 'response to heat' and 'protein folding,' and comprises two primary nodes. The first node is a super node that includes 11 different gene regulators, including *PEACOCK 1, MYB DOMAIN PROTEIN 56, WRKY DNA-BINDING PROTEIN 51, AT5G05790, BASIC*
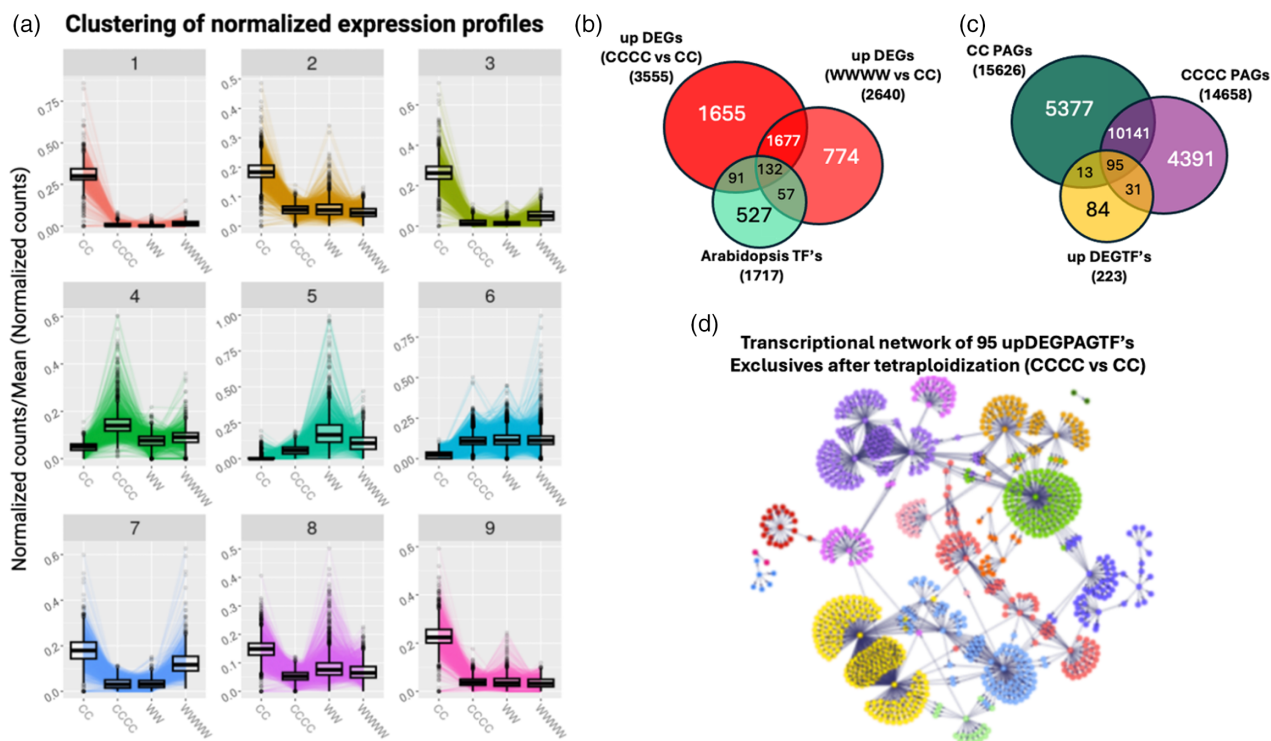
Figure 5. Clustering and networks.
(a) Clusters of normalized expression profiles. Nine different profiles were obtained using the Diane algorithm.
(b) Venn diagram included upDEGs in recent tetraploidization, upDEGs in ancient tetraploidization, and Arabidopsis transcription factors.
(c) Venn diagram included upDEG TFs in recent tetraploidization, CCCC PAGs, and CC PAGs.
(d) Transcriptional network of 95 upDEG PAG TFs responsive to recent tetraploidization. DEG, differentially expressed gene; PAG, peak-associated gene; TF, transcription factor.

LEUCINE-ZIPPER 8, ABERRANT TESTA SHAPE, GATA17-LIKE, CYTOKININ RESPONSE FACTOR 2, CYTOKININ RESPONSE FACTOR 10, ASYMMETRIC LEAVES 2-LIKE 16, and *TRNA METHYLTRANSFERASE 2B* that target 153 genes (File S4). The second node is represented by *JASMONATE ASSOCIATED MYC2 LIKE 3* and *NAC DOMAIN CONTAINING PROTEIN 17*, which target 127 genes (File S4).

Module three consists of three distinct nodes that regulate specific sets of genes. The first node is composed of five gene regulators, *BIG PETAL P, KNOTTED-LIKE HOMEOBOX OF ARABIDOPSIS THALIANA 7, OCS-ELEMENT BINDING FACTOR 5, ARABIDOPSIS 6B-INTERACTING PROTEIN 1-LIKE 2*, and *ATSTKL1*, that target and regulate 35 genes (File S4). The second node comprises *GATA TRANSCRIPTION FACTOR 24* and regulates 27 genes (File S4). In contrast, the third node consists of three genes, *AT3G24490, AT2G20280*, and *NUCLEAR FACTOR Y*, that regulate 36 genes (File S4). Despite the association of several transcription factors with the tetraploid-relative phenotype, their precise relationship in modulating the epigenetic response following genome doubling remains unclear.

The epigenetic system is an important player after WGD

Ploidy alterations, such as transitions between diploidy and polyploidy, introduce profound changes to cellular and molecular processes, driving plants' developmental innovation and environmental adaptability. *Arabidopsis thaliana*, a model plant organism, offers an ideal system for unraveling the intricate molecular interactions that underpin these responses. Using the epigenetic factors mentioned above as nodes, we infer two different subnetworks using a previously established Arabidopsis root transcriptional regulatory network (Montes et al., 2014). Figure 6 highlights the cascading effects of ploidy changes on gene regulatory networks, revealing broad transcriptional reprogramming (Figure 6a) and pinpointing specific epigenetic mechanisms mediated principally by histone variants (Figure 6b). In Figure 6(a), we see an abstract of an extensive transcriptional network; in the center of this intricate network are important regulators, such as VRN2 (*REDUCED VERNALIZATION RESPONSE 2*), DET1 (*DE-ETIOLATED 1*), CHR9 (*CHROMATIN REMODELING 9*), EMF2 (*EMBRYONIC FLOWER 2*), JMJ14 (*JUMONJI 14*), FRG5 (*RING-HELICASE LIKE 5*), AGO2 (*ARGONAUTE 2*), and
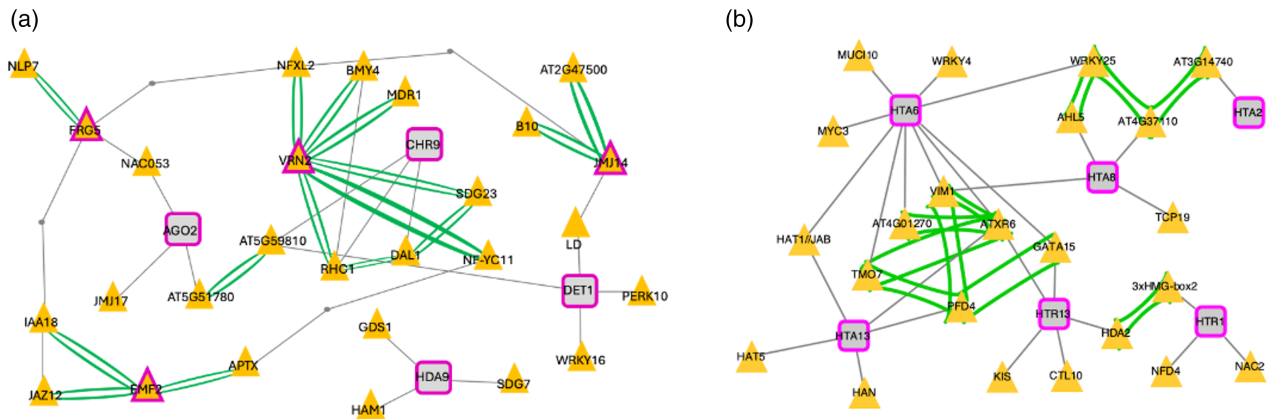
(a)

(b)

Figure 6. Transcriptional subnetworks.
(a) Subnetwork of overexpressed epigenetic factors after WGD.
(b) Subnetwork with six silenced histone variants found after WGD. WGD, whole-genome duplication.

HDA9 (*HISTONE DEACETYLASE 9*), which serve as pivotal hubs, influencing the expression of multiple downstream genes. The edges connecting these nodes vary in nature, with green edges showing reported physical interactions. These visual captures connect biological processes affected by ploidy changes, including gene silencing, chromatin remodeling, and transcriptional regulation. Central hubs in the network, composed of key transcription factors and chromatin regulators, likely coordinate large-scale gene expression changes driven by gene dosage variations and altered genomic architecture inherent to polyploidy. Each cluster within the global network appears to represent a distinct regulatory module (Table S16). For example, the network in Figure 6(a) contains eight nodes, represented by *VRN2*, *DET1*, *CHR9*, *FRG5*, *AGO2*, *EMF2*, *HDA9*, and *JMJ14*. This intricate network shows the interaction of *VRN2*, which plays an important role in regulating flowering by vernalization and cold response, is a nuclear-localizing zinc finger protein important in the Polycomb Repressive Complex 2 (PRC2) (de Lucas et al., 2016; Diallo et al., 2010) with *NFXL2*, *BMY4*, and *MDR1*, involved in salt response (Lisso et al., 2006), growth regulation (Reinhold et al., 2011) and auxin transport (Liu et al., 2022) respectively. The *DET1* node interacts principally with *WRKY16*, a transcription factor reported as an important gene to trigger immunity in Arabidopsis (Birkenbihl et al., 2018), *LD* (*LUMINIDEPENDENS*), important to the positive regulation of flowering (Qi et al., 2022) and *PERK10* (*PROLINE-RICH EXTENSIN-LIKE RECEPTOR KINASE 10*), relative to signaling and negative regulation of root growth (Humphrey et al., 2015). *CHR9*, a SWI2 chromatin remodeler gene family (Shaked et al., 2006; Yang et al., 2018) interacts with multiple proteins, including *RHC1*, involved in the response to drought stress (Cho et al., 2022). The cluster in the bottom left is dominated by *EMF2* (*EMBRYONIC FLOWER 2*), a Polycomb group protein involved in

chromatin-based repression; EMF2 is a critical component of the PRC2, which regulates gene expression by depositing the repressive histone mark H3K27me3 (del Olmo et al., 2016; Xiao et al., 2017). In the context of ploidy changes, EMF2 probably acts to restrict the expression of developmental genes that could otherwise disrupt proper organogenesis and flowering (de Lucas et al., 2016; Gu et al., 2014; He et al., 2012), maintaining a balance between repression and activation of key developmental pathways (Zhu et al., 2024) *JAZ12* (*JASMONATE-ZIM-DOMAIN PROTEIN 12*), involved in the jasmonate pathway and responsive to biotic stress (de Torres et al., 2016), *IAA18* (*INDOLE-3-ACETIC ACID INDUCIBLE 18*), important to regulate the pattern of root organogenesis (Bustillo-Avendaño et al., 2018) *APTX* (*APRATAXIN-LIKE*), involved in modulating axillary meristem formation (Yang et al., 2012), are the most important interactors of *EMF2* (Wen et al., 2008). *JMJ14* dominates another cluster, emphasizing its role in histone demethylation (Wang et al., 2023); its principal interactors (Huang et al., 2017) *LD* and *AT2G47500* are involved in flowering and cell cycle regulation respectively (Qi et al., 2022; Vanstraelen et al., 2006). Another cluster is controlled by *FRG5*, an SNF2 chromatin remodeler required for DNA methylation (Groth et al., 2014); their functionally principal related gene is *NLP7* (*NIN LIKE PROTEIN 7*), involved in the nitrogen sensing and transport pathway (Shanks et al., 2024; Zhang et al., 2024). Another key regulator visible in the network is *AGO2* (*ARGONAUTE 2*), a core component of the RNA-induced silencing complex (RISC) that fine-tunes gene expression during stress responses and developmental transitions (Allen et al., 2005; Liu et al., 2023; Ma et al., 2015; Stroud et al., 2013; Zhang et al., 2014); at the edges connecting this node are genes like (Gaudinier et al., 2018) *NAC053*, responsive to drought stress and important for the regulation of reactive oxygen species generation (Lee & Park, 2012), and *JMJ17*

(*JUMONJI DOMAIN-CONTAINING PROTEIN 17*), important to trigger H3K4 demethylation and promote dehydration stress response (Huang et al., 2019). In the context of ploidy changes, AGO2 may modulate the expression of genes related to genome stability, a common consequence of WGD (Comai, 2005). Finally, the *HDA9* node, around a gene encoding a chromatin remodeler regulating the response to drought stress (Baek et al., 2020), is associated with *HAM1* (*HISTONE ACETYLTRANSFERASE OF THE MYST FAMILY 1*), an acetyltransferase gene related to transcription regulation and plant development (Wu et al., 2023) *GDS1* (*GROWTH, DEVELOPMENT AND SPLICING 1*), an important molecular player induced by nitrogen starvation (Fan et al., 2023) and *SDG7* (*HISTONE-LYSINE N-METHYLTRANSFERASE ASHH3*), a methyltransferase involved in the vernalization response (Lee et al., 2015). Together, the global network paints a picture of a highly coordinated system where epigenetic mechanisms regulate fine-tuned gene expression in response to changes in genome content.

In contrast, Figure 6(b) dives deeper, focusing on a smaller subnetwork anchored by histone variants HTA8, HTA6, HTA13, and HTA2. The genes encoding these variants, marked by magenta squares, emerge as key players in a tightly interconnected regulatory framework (Table S17). This subnetwork reveals how histone dynamics influence transcriptional regulation by interacting with specific TFs and their downstream targets. Central nodes such as *HTA2* and *HTA13*, which encode histone *H2A* variants, reflect the critical role of chromatin remodeling in responding to ploidy changes (Allfrey et al., 1964; Jamge et al., 2023). Histone variants influence nucleosome structure and DNA accessibility, thereby controlling which genes are activated or silenced (Pratx et al., 2023; Yelagandula et al., 2014; Zilberman et al., 2008). These histone variants likely interact with transcription factors (TFs) such as WRKY25, TMO7 (TARGET OF MONOPTEROS 7), and GATA15 (GATA TRANSCRIPTION FACTOR 15). In tetraploidization, WRKY25 appears critical in mediating responses to abiotic stresses, including dehydration and salt stress, particularly in roots. This role may extend to stress memory mechanisms, enabling tetraploid plants to survive in hostile conditions (Deolu-Ajayi et al., 2019; Ding et al., 2013); *TMO7*, on the other hand, is potentially involved in root-specific signaling pathways, where it regulates secondary root initiation. This function could be crucial for enhancing nutrient acquisition efficiency in tetraploid plants, which require more robust root systems to sustain their increased genome size and metabolic demands (Lu et al., 2018; Schlereth et al., 2010). *GATA15*, a zinc finger transcription factor, has been implicated in nitrogen metabolism; *GATA15* may act as a key regulator of nitrogen acquisition in tetraploids, contributing to improved nutrient assimilation. Additionally, it might influence

circadian rhythms, thereby modulating early flowering, a trait often observed following tetraploidization (Manfield et al., 2007; Peng et al., 2007).

DISCUSSION

In polyploid organisms, additional copies of all genes increase functional redundancy, providing a genetic buffer to compensate for loss-of-function mutations and providing great flexibility for neo-and sub-functionalization of genes to drive adaptation and speciation through evolution. Additional loci can also interact, leading to complex phenotypic variations (Comai, 2005; Flagel & Wendel, 2010; Ohno, 1970; Osborn et al., 2003), but due to increased genetic complexity, polyploidization enables the organism to recognize and respond to a broader range of stimuli. However, tetraploidization can also hamper the ability of regulatory elements to bind to their target genes, leading to changes in gene expression and contributing to genome instability and genetic disorders (Bailey et al., 2002). Increasing evidence suggests that epigenetic processes, such as changes in DNA methylation patterns, play critical roles after genome ploidy changes (Alexandre et al., 2018; Comai, 2005; Jordan et al., 2020; Osborn et al., 2003; Wang et al., 2021; Wendel, 2000; Yen et al., 2017). We used whole root tissue to identify genome-wide chromatin accessibility differences between diploid and tetraploid Arabidopsis accessions. This approach will likely mainly reveal chromatin accessibility regions common to the different root cell types or regions of chromatin accessibility that drastically differ between the same cell type of two distinct genotypes. Single-cell approaches will be necessary to determine if ploidy levels cause cell type-specific changes in chromatin accessibility. It is possible that artificial tetraploidization can cause stochastic changes in chromatin accessibility that are not correlated with ploidy levels. However, two lines of evidence suggest that these changes are at least not all stochastic: (i) It has been reported that colchicine-induced tetraploidization often leads to highly stable phenotypes (Fischer et al., 2022; Parra-Nunez et al., 2024) Our results revealed a high percentage of chromatin accessibility peaks present in both synthetic and natural tetraploids that are absent in a natural diploids, suggesting that both phenotypic and epigenetic changes are not stochastic; and (ii) the similarities between our data and previously reported data (Fischer et al., 2022). We compared published root transcriptomic data from colchicine-induced tetraploids of Arabidopsis Col-0 plants with our transcriptomic data. As expected, a heatmap of gene counts from the three published tetraploid samples, our three tetraploids, and one of our diploids demonstrated high similarity among tetraploid samples (Figure S4a). Following data normalization (Figure S4b), principal component analysis (PCA) further revealed significant differences between diploid and tetraploid

transcriptomic profiles (Figure S4c), supporting the reliability of our data. This finding further supports the use of a hydroponic system for growing the plants in this study, as the high similarity between our data and that reported by Fischer et al. – obtained using plants grown on agar media – indicates that hydroponic growth conditions can reliably reproduce comparable physiological and molecular responses. This consistency validates hydroponic systems as a robust alternative for experimental strategies, enabling controlled nutrient delivery and environmental conditions while maintaining data reproducibility across different growth substrates. This study showed that chromatin accessibility is an important component of the genomic response to WGD. In agreement with the notion that chromatin accessibility should decrease after WGD due to the reduced nuclear volume and closer proximity of centromeres in tetraploid cells compared to diploid cells (Corneillie et al., 2019; Sas-Nowosielska & Bernas, 2016; Tsukaya, 2013), we observed a loss of genomic accessible regions in a synthetic tetraploid CCCC and a gain of chromatin accessibility after WGR in WW compared to the natural tetraploid WWWW.

This work is based on the overlapped analysis of the large databases to decipher open chromatin regions and gene transcription associations; a hypergeometric test was performed to any comparison using Venn diagrams to determine the significance of every contrast (Table S18). We identified 24 364 unique PAGs across all genotypes, of which 31% (7611) were ploidy and genotype-independent (Figure 1e). These conserved PAGs were enriched in response categories, such as 'response to stimulus and 'response to chemical,' response to stress,' and 'response to abiotic stimulus' among others (File S1). This aligns well with a previous report showing that genes exhibiting high genomic accessibility were primarily associated with stress responses (Monroe et al., 2022). We found that only 14% in CC, 1.8% in CCCC, 3.3% in WWWW, and 3.2% in WW of PAGs were line-specific (Figure 1e), suggesting a minor ecotype-specific effect on chromatin accessibility after synthetic tetraploidization and diploidization. We also found 4927 PAGs shared by the CCCC and WWWW that are absent in CC and represent the loci that gained chromatin accessibility in the tetraploid state (Figure 1e), of which 2040 had changes in gene expression.

Previous reports have indicated a lack of significant correlation between chromatin accessibility and gene expression (Kiani et al., 2022; Tsompana & Buck, 2014). In this study, we found that over 50% of the DEGs in tetraploids compared to diploids were associated with an open chromatin region (Figure 2c,g), highlighting the critical role of chromatin architecture in regulating gene expression following WGD and WGR (Figures 2 and 3). Additionally, it has been reported that alteration in gene expression in Arabidopsis occurring after WGD could be tissue or cell-type-specific (Del Pozo & Ramirez-Parra, 2014). Our study shows clear resemblances in gene ontology (GO) enrichment categories between our root data and that previously published for leaves. We detected GO enrichment categories in the DEGs of CCCC and WWWW roots that exhibited similarities to salt and drought stress response genes previously reported in the literature for shoots of Arabidopsis tetraploids (Del Pozo & Ramirez-Parra, 2014). We also identified 1809 upDEGs shared by both tetraploids that are conserved between a synthetic and a natural polyploids (Figure S5). The enriched categories observed in these DEGs included 'nitrogen compound metabolic process,' 'nucleic acid metabolic process,' 'inositol phosphate metabolism,' 'phosphate metabolic process,' 'response to stress,' 'chromatin organization,' and 'chromatin remodeling' (Table S19), reflecting that chromatin organization among nitrogen and phosphate metabolism-related genes are important adaptive traits gained in tetraploids following WGD. Our results indicate that Arabidopsis chromatin remodeling and changes in gene expression associated with a WGD event are established soon after tetraploidization, stably maintained over time, and retained after synthetic diploidization. Following this idea, the relation of transcriptomic libraries of CC, CCCC, WWWW, and WW in a PCA plot suggests that the transcriptional changes that appear after WGD in C are maintained in WWWW and still largely present in WW (Figure S6). Moreover, of the 17 394 PAGs detected in WWWW, 15234 are conserved in WW, suggesting that open chromatin associated with changes in gene expression in the tetraploid are largely maintained after genome reduction. At the experimental level, it is important to determine the relation between PAGs and DEGs to help decipher the regulatory networks involved in intergenic long-distance regulation in open chromatin regions. Our results show that the alterations in gene expression activated upon WGD are not, at least not rapidly, reverted after diploidization.

We also found 66 genes positively correlated with the ploidy level, i.e., increased expression in tetraploids and decreased expression in diploids. These 66 genes include *GLUCOSE-6-PHOSPHATE/PHOSPHATE TRANSLOCATOR 2*, involved in carbon exchange across the chloroplast envelope (Weise et al., 2019), which has been reported to promote higher rates of photosynthesis (Dyson et al., 2015), *NITRILASE 1*, an enzyme involved in nitrogen metabolism and a positive regulator of flowering (Yang et al., 2022), *NITRATE TRANSPORTER 1.5*, involved in nitrate transport (Watanabe et al., 2020) and potassium homeostasis via the ethylene pathway (Chen et al., 2021), *NITRATE TRANSPORTER 1.7*, a nitrate transmembrane transport responsible for nitrate remobilization, *ALDEHYDE DEHYDROGENASE 7B4*, a gene responsive to abscisic acid, desiccation, and salt stress (Huang et al., 2018; Zhao et al., 2017), and *WALL-ASSOCIATED KINASE 2*, a cell

wall-associated kinase involved in cell expansion and salt stress responses (Kohorn et al., 2014; Renault et al., 2013). The higher expression of these genes in tetraploids could explain, at least in part, some of the phenotypes described in tetraploids, including larger cells, increased salt-drought tolerance, and increased potassium leaf content.

After clustering the transcriptomic response to tetraploidization, we determined nine clusters with different transcriptional profiles (Figure 5) to infer a functional network based on 95 upDEGPAGTFs in CCCC, with pre-existing chromatin accessibility in CC. We identified 15 modules within this network, each with a varying number of nodes (TFs). The analysis of these 15 modules suggests that a limited number of TFs could orchestrate changes in chromatin accessibility during tetraploidization and directly or indirectly influence the observed phenotype, including stress response, regulation of flowering time, cell size, DNA repair, nutrient deficiency, and regulation of plant hormone signaling. However, the question of whether overexpressing a few TFs can activate the epigenetic system to establish the tetraploid phenotype remains to be tested. Nevertheless, we found clusters 6 and 9 particularly intriguing as they encompassed most genes encoding epigenetic factors whose expression is altered by genome doubling. Using these epigenetic factors as nodes, we infer two different subnetworks using a previously established Arabidopsis root transcriptional regulatory network (Montes et al., 2014) (Figure 6a,b). We identified seven modules with 31 transcription factors (Table S17) as principal neighbors of the epigenetic players with higher expression in tetraploids. This suggests their role as primary orchestrators of the transcriptional changes activated during tetraploidization. These transcription factors participate in biological processes related to the tetraploid-acquired phenotype, such as *NFXL2*, *NTL4*, and *JMJ14* involved in salt stress response (Lisso et al., 2006), drought recovery (Ding et al., 2013), and flowering regulation (Rodrigues & Dolde, 2021), respectively.

On the other hand, in cluster 9, we identified distinct histone variants that exhibit transcriptional decay in synthetic CCCC tetraploids: *H3.1*, *H2A.W.6*, *H2A.2*, *H2A.13*, and *H2A.Z*. The transcriptional effects of these histone variants are expected to have a broad and complex impact on the tetraploid Arabidopsis genome. These effects may include changes in chromatin architecture, disruptions in gene regulation, and increased genomic instability. Histone variants are involved in several functional pathways: H3.1 and H2A.W.6 are linked to heterochromatin formation and maintenance of genome stability, while H2A.2, H2A.Z, and H2A.13 are associated with chromatin remodeling and stress responses. Changes in the expression of these histone variant genes could disrupt several epigenetic pathways, resulting in improper chromatin structure, altered

gene expression, and compromised genome integrity. Further experimental studies are needed to validate these hypotheses and clarify the specific molecular mechanisms by which these histone variants influence genome function after ploidy changes (Jamge et al., 2023; Probst et al., 2020; Wang et al., 2018). Recent reports show that histone variants are pivotal in determining the definition and distribution of chromatin states in Arabidopsis (Jamge et al., 2023). These histone variants exhibit a strong correlation with 28 transcription factors, including *HD2A3*, which is involved in drought response and cell proliferation and differentiation (Tahir et al., 2022); *GATA1*, responsive to temperature stimulus (Manfield et al., 2007); *ATXR6*, a member of the Trithorax group of proteins that modulate genome stability (Potok et al., 2022); *MYC3*, involved in the jasmonic acid response (Wilkinson et al., 2023); and *WRKY4*, involved in response to salt stress (Li, Li, & Jiang, 2021). While the precise molecular mechanism underlying the epigenetic response to tetraploidization remains elusive, further investigation should be made to highlight the function of histone variant deposition after tetraploidization in shaping chromatin architecture.

In summary, our data demonstrated that tetraploidization induces significant changes in chromosome accessibility and suggests that impaired deposition of histone variants facilitates chromatin heterochromatinization, promoting DNA accessibility and enhancing transcription in genes relative to stress response and uridine metabolism principally (Figure 7). Nevertheless, through whole-genome reduction, chromatin changes observed in tetraploids are maintained in a synthetic diploid, indicating that new diploids might retain specific characteristics gained after genome doubling (Figure 7). Therefore, our results offer the first functional genomic framework to understand better how plant genome doubling drives a myriad of responses, including salt and drought tolerance (Del Pozo & Ramirez-Parra, 2014) and increased potassium content in leaves (Chao et al., 2013). Tetraploidization has played a critical evolutionary role, with polyploid genomes being preferentially selected over their diploid progenitors during mass extinction events that caused widespread species loss. Notably, polyploidy is common in many crops, such as cotton (an allotetraploid) and wheat (an allohexaploid), due to human selection for better performance during plant breeding. Our research highlights immediate transcriptional changes following genome duplication in Arabidopsis, particularly in stress-responsive genes (Figure 2i). These findings suggest that leveraging genome content changes could provide a testable strategy to enhance crop breeding, fostering the development of resilient crops to face and tolerate abiotic stresses, including nutrient scarcity, drought, and salinity, under the pressures of climate change.
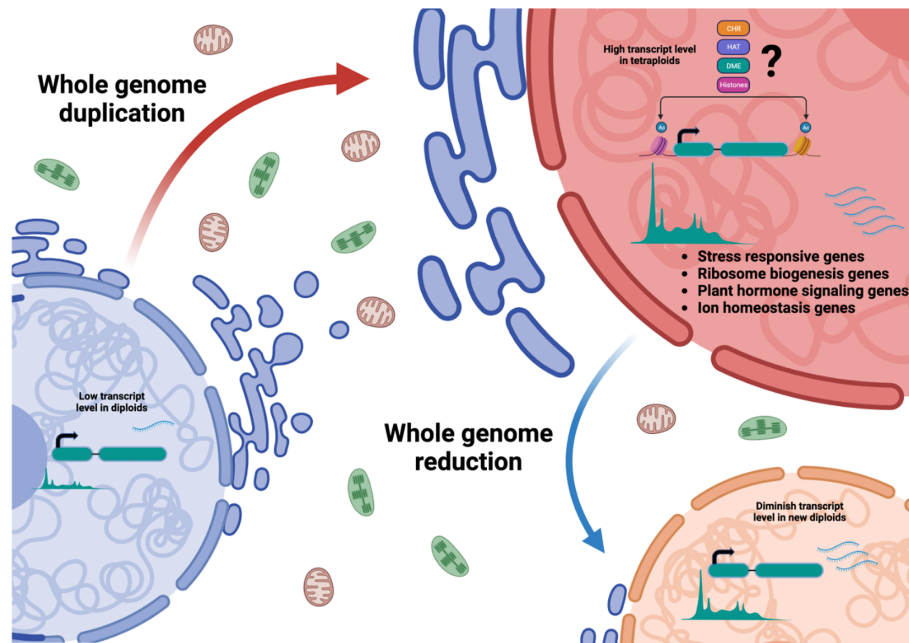
Figure 7. The model.
Our model reflects that tetraploidization leads to heterochromatinization and enhances stress-responsive gene transcription. The presence of the epigenetic factors in new and old tetraploids suggests that after tetraploidization, the genome structure suffers several modifications affecting gene transcription. Moreover, through whole-genome reduction, molecular changes are maintained, suggesting that new diploids still retain certain molecular features obtained after genome doubling.

METHODS

Plant materials and growth conditions

Wild-type Arabidopsis (*Arabidopsis thaliana*) accession Columbia-0 (Col-0) was used for all experiments; the synthetic Col-0 tetraploid, the natural tetraploid Wa-1, and synthetic Wa-1 haploid were donated by Luca Comai (California University Davis). Seeds were surface sterilized with 90% (v/v) ethanol for 5 min and 50% (v/v) bleach solution for 5 min before four washes with sterile distilled water. Seedlings were grown using a hydroponic system with $0.1 \times$ Murashige and Skoog (MS) medium. Complete root systems for all four genotypes were analyzed 10 days after germination (DAG). Seedlings were grown at 20°C in an 18-h light/6-h dark photoperiod.

Nuclei isolation

For nuclei isolation, we followed the method of Bajic et al. (2018) with some modifications: the extraction buffer consisted of 15 mm Tris–HCl pH 7.5, 20 mm NaCl, 80 mm KCl, 0.2% Triton X-100, and 5 mm β-mercaptoethanol. Nuclei suspensions were filtered using a 30-μm mesh; to eliminate mitochondria and chloroplasts, we used sucrose sedimentation buffer and centrifuged at 600 ***g*** or 4,500 rpm for 20 min: 20 mm Tris–HCl pH 8, 0.2 mm $MgCl_2$, 2 mm EDTA, 0.2%Triton X-100, 15 mm β-mercaptoethanol, and 1.7 m sucrose.

ATAC-seq library preparation and sequencing

For ATAC-seq assays, two replicates per sample were processed using between 80 000 and 120 000 nuclei, as determined by flow cytometry. Chromatin was digested by Tn5-mediated tagmentation and adapter incorporation, according to the manufacturer's protocol (Nextera DNA sample preparation kit, Illumina®) at 37°C

for 30 min; each library was amplified for 12–15 cycles according to the published protocol (Maher et al., 2018). The quality of the libraries was assessed by a DNA-based fluorometric assay and by electrophoresis. Samples were sequenced on a HiSeq2500 Illumina sequencer system as paired-end reads of 2×50 bp.

RNA extraction

Harvested whole root systems were frozen in liquid nitrogen and ground to a fine powder. Total RNA was isolated using TRIzol reagent (Invitrogen) according to the manufacturer's instructions. Novogene, Inc. (Sacramento, CA, USA) generated mRNA-seq libraries.

RNA-seq analysis

Adapter sequences and low-quality reads were removed from raw reads with trimGalore v0.6.4 (https://github.com/FelixKrueger/TrimGalore). Mapping of reads to the genome and gene counts were performed using RNA-STAR v2.7.5b (Dobin et al., 2013) and Galaxy (Afgan et al., 2018) through the usegalaxy.eu server and fragment read counts over genes were obtained using htseq-count v0.9.1+galaxy1 (Anders et al., 2015). Differential gene expression analysis was performed using the DESeq2 package in R (Love et al., 2014). Gene Ontology (GO) enrichment analysis and cluster analysis by biological process were performed using g:profiler (Raudvere et al., 2019) and DIANE algorithm (DIANE). Heatmaps of DEGs were constructed following published bioinformatics methods (Ojeda-Rivera et al., 2020).

ATAC-seq bioinformatic analysis

Trimming of adapter sequences and removing low-quality reads from raw reads were performed using trimGalore v0.6.4

(https://github.com/FelixKrueger/TrimGalore). Clean reads were then aligned to the Arabidopsis TAIR10 release 43 reference genome using Bowtie2 v2.4.5 (Langmead & Salzberg, 2012) with options -k 10 -very sensitive. PCR duplicates were marked with sambamba-markdup v0.7.0 (Tarasov et al., 2015); all steps up to this point in the analysis were automated using snakePipes (Bhardwaj et al., 2019). PCR duplicates and reads mapping to the organellar genomes were removed with samtools v1.10 (Li et al., 2009). Quality control of filtered mapped data was performed using ATACseqQC v1.10.4 (Ou et al., 2018).

Peaks were called for each replicate using the find Peaks function within the MACS2 suite v2.1.1.20160309 (Heinz et al., 2010) with the following parameters: Effective genome size: 135 000 000, Build the shifting model, Set lower mold bound = 5, Set upper mold bound = 50, Bond width for picking regions to compute fragment size = 300, Peak detection based in = $q$value to detect false positives, minimum FDR (q-value) cut-off = 0.05, how many duplicate tags at the exact same location are allowed = 1. Peak datasets were annotated to the transcription start site (TSS) of the nearest gene using ChipSeeker v1.22.1 with org.At.tair.db and TxDb. Athaliana.BioMart.plantsmart28 Bioconductor packages (Carlson, 2017; 2020; Yu et al., 2015). Promoters were defined as spanning 1000 bp of sequence upstream of the TSS and 400 bp downstream of the TSS. GO categories and cluster analysis by biological process were analyzed using g:profiler (Raudvere et al., 2019). Signal visualization files and images were generated using deepTools v.3.5.0 (Ramírez et al., 2016). MultiBamSummary scaling factors were used to generate bigwig files with bamCoverage. Overlap between genomic regions was determined using Intervene 0.6.4 (Khan & Mathelier, 2017).

Differential accessibility analysis

Differential accessibility was determined using the R package csaw (Lun & Smyth, 2014). BAM alignment files from our ATAC-seq data, pre-processed following the pipeline from Reske and colleagues (Reske et al., 2020), were used as input. Differential accessibility was calculated for regions defined by our PAGs gene list (File S1) with the parameters logCPM of peaks higher than −3; TMM normalization factor using 10 000 bases wide bins; contrast CCCC-CC, i.e., CC as the reference. Regions with an FDR of 0.05 or lower were considered differential. Differential regions were then analyzed using ChIPseeker (Yu et al., 2015). Our csaw R script is available at https://github.com/ricardo-aaron/polyploidy.

AUTHOR CONTRIBUTIONS

ACB-R and LH-E designed the experiments. ACB-R performed the experiments. ACB-R and RACM carried out the bioinformatic analysis. ACB-R, RACM, and LH-E analyzed the data. ACB-R, RACM, and LH-E wrote the manuscript. All authors read and approved the manuscript.

ACKNOWLEDGMENTS

CONFLICT OF INTEREST

The authors declare that they have no competing interests.

DATA AVAILABILITY STATEMENT

The datasets generated in this study are available at the NCBI GEO repository under accession number: PRJNA 1200049. Reviewer link: https://dataview.ncbi.nlm.nih.gov/object/PRJNA1200049?reviewer=c4pnoonj3scb92kn46lgfu6j4b.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article.

Figure S1. Phenotype and ploidy measurement after WGD and WGR.
Figure S2. Experimental design, MDS Plot.
Figure S3. Genome accessibility analysis across chromosomes.
Figure S4. Commons and uniques PAGs between genotypes.
Figure S5. upDEGs recent versus old tetraploidization hypothesis.
Figure S6. PCA transcriptomic response.
Table S1. QC of ATAC-seq experiment.
Table S2. Top 10 gene ontology enrichment category of 8765 shared PAGs between CC and CCCC.
Table S3. Gene ontology enrichment of 6861 unique PAGs in CC.
Table S4. Top 10 gene ontology enrichment of 5893 unique PAGs in CCCC.
Table S5. Top 10 gene ontology enrichment category of 1102 genes related to upDARs in response to tetraploidization.
Table S6. Gene ontology enrichment category of 227 genes relatives to downDARs in response to tetraploidization.
Table S7. Top 10 gene ontology enrichment category of 4927 commons PAGs between CCCC and WWWW.
Table S8. QC of RNA-seq experiment alignments.
Table S9. Gene ontology enrichment category of 3555 upDEGs in response to tetraploidization.
Table S10. Top 10 gene ontology enrichment category of 3382 downDEGs in response to tetraploidization.
Table S11. Top 10 gene ontology enrichment category of 2040 DEGs with chromatin accessibility in CCCC and WWWW.
Table S12. Gene ontology enrichment of 2160 PAGs loss after diploidization.
Table S13. Top 10 gene ontology enrichment category of 259 upDEGs responsive to diploidization.
Table S14. Top 10 gene ontology enrichment category of 756 downDEGs responsive to diploidization.
Table S15. Gene ontology enrichment category of 90 downDEGs in diploidization response and upDEGs after diploidization.
Table S16. Eleven epigenetic factors in CCCC and their associated regulators.

Table S17. Associated regulators with six histone variants.

Table S18. Hypergeometric test to determine the statistical significance of Venn diagrams overlaps.

Table S19. Top 10 gene ontology enrichment category of 1,809 shared upDEGs in tetraploids taken CC as a reference.

File S1. PAGs and GO enrichment.

File S2. Differential expression.

File S3. Transcriptional clusters.

File S4. Transcriptional network 95 TF.

REFERENCES

Afgan, E., Baker, D., Batut, B., Van Den Beek, M., Bouvier, D., Ech, M. *et al.* (2018) The galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. *Nucleic Acids Research*, **46**(W1), W537–W544.

Alexandre, C.M., Urton, J.R., Jean-Baptiste, K., Huddleston, J., Dorrity, M.W., Cuperus, J.T. *et al.* (2018) Complex relationships between chromatin accessibility, sequence divergence, and gene expression in *Arabidopsis thaliana*. *Molecular Biology and Evolution*, **35**(4), 837–854.

Allen, E., Xie, Z., Gustafson, A.M. & Carrington, J.C. (2005) microRNA-directed phasing during trans-acting siRNA biogenesis in plants. *Cell*, **121**(2), 207–221.

Allfrey, V.G., Faulkner, R. & Ms, A.E. (1964) Acetylation and methylation of histones and their possible role in the. *Proceedings of the National Academy of Sciences of the United States of America*, **51**(1938), 786–794.

Anders, S., Pyl, P.T. & Huber, W. (2015) HTSeq-a python framework to work with high-throughput sequencing data. *Bioinformatics*, **31**(2), 166–169.

Baek, D., Shin, G., Kim, M.C., Shen, M., Lee, S.Y. & Yun, D.J. (2020) Histone deacetylase HDA9 with ABI4 contributes to abscisic acid homeostasis in drought stress response. *Frontiers in Plant Science*, **11**, 1–12.

Bailey, J.A., Gu, Z., Clark, R.A., Reinert, K., Samonte, R.V., Schwartz, S. *et al.* (2002) Recent segmental duplications in the human genome. *Science (1979)*, **297**(5583), 1003–1007.

Bajic, M., Maher, K.A. & Deal, R.B. (2018) Identification of open chromatin regions in plant genomes using ATAC-seq. *Plant Chromatin Dynamics*, **1675**, 183–201. Available from: https://doi.org/10.1007/978-1-4939-7318-7_12

Barragán-Rosillo, A.C., Peralta-Alvarez, C.A., Ojeda-Rivera, J.O., Arzate-Mejía, R.G., Recillas-Targa, F. & Herrera-Estrella, L. (2021) Genome accessibility dynamics in response to phosphate limitation is controlled by the PHR1 family of transcription factors in Arabidopsis. *Proceedings of the National Academy of Sciences of the United States of America*, **118**(33), e2107558118.

Bhardwaj, V., Heyne, S., Sikora, K., Rabbani, L., Rauer, M., Kilpert, F. *et al.* (2019) SnakePipes: facilitating flexible, scalable and integrative epigenomic analysis. *Bioinformatics*, **35**(22), 4757–4759.

Birkenbihl, R.P., Kracher, B., Ross, A., Kramer, K., Finkemeier, I. & Somssich, I.E. (2018) Principles and characteristics of the Arabidopsis WRKY regulatory network during early MAMP-triggered immunity. *The Plant Journal.*, **96**(3), 487–502.

Bustillo-Avendaño, E., Ibáñez, S., Sanz, O., Sousa Barros, J.A., Gude, I., Perianez-Rodriguez, J. *et al.* (2018) Regulation of hormonal control, cell reprogramming, and patterning during de novo root organogenesis. *Plant Physiology*, **176**(2), 1709–1727.

Carlson, M. (2020) TxDb.Athaliana.BioMart.plantsmart28. 1–2.

Carlson, M. (2017) org.At.tair.db. Bioconductor. Available from: https://doi.org/10.18129/B9.BIOC.ORG.AT.TAIR.DB

Cassan, O. & Lèbre, S. & Martin, A. (2021) Inferring and analyzing gene regulatory networks from multi-factorial expression data: a complete and interactive suite. *BMC Genomics*, **22**(1), 387. Available from: https://doi.org/10.1186/s12864-021-07659-2

Cenci, A., Hueber, Y., Zorrilla-Fontanesi, Y., Van Wesemael, J., Kissel, E., Gislard, M. *et al.* (2019) Effect of paleopolyploidy and allopolyploidy on gene expression in banana. *BMC Genomics*, **20**(1), 1–12.

Chalhoub, B., Denoeud, F., Liu, S., Parkin, I.A.P., Tang, H., Wang, X. *et al.* (2014) Early allopolyploid evolution in the post-neolithic *Brassica napus* oilseed genome. *Science (1979)*, **345**(6199), 950–953.

Chao, D.Y., Dilkes, B., Luo, H., Douglas, A., Yakubova, E., Lahner, B. *et al.* (2013) Polyploids exhibit higher potassium uptake and salinity tolerance in Arabidopsis. *Science (1979)*, **341**(6146), 658–659.

Chen, H., Zhang, Q., Wang, X., Zhang, J., Ismail, A.M. & Zhang, Z. (2021) Nitrogen form-mediated ethylene signal regulates root-to-shoot K+ translocation via NRT1.5. *Plant, Cell & Environment*, **44**(12), 3576–3588.

Cho, N.H., Woo, O.G., Kim, E.Y., Park, K., Seo, D.H., Yu, S.G. *et al.* (2022) E3 ligase AtAIRP5/GARU regulates drought stress response by stimulating SERINE CARBOXYPEPTIDASE-LIKE1 turnover. *Plant Physiology*, **190**(1), 898–919.

Chu, X., Li, S., Wang, S., Luo, D. & Luo, H. (2021) Gene loss through pseudogenization contributes to the ecological diversification of a generalist Roseobacter lineage. *ISME Journal*, **15**(2), 489–502.

Comai, L. (2005) The advantages and disadvantages of being polyploid. *Nature Reviews. Genetics*, **6**(11), 836–846.

Corneillie, S., De Storme, N., Van Acker, R., Fangel, J.U., De Bruyne, M., De Rycke, R. *et al.* (2019) Polyploidy affects plant growth and alters cell wall composition. *Plant Physiology*, **179**(1), 74–87.

de Lucas, M., Pu, L., Turco, G., Gaudinier, A., Morao, A.K., Harashima, H. *et al.* (2016) Transcriptional regulation of Arabidopsis Polycomb repressive complex 2 coordinates cell-type proliferation and differentiation. *The Plant Cell*, **28**(10), 2616–2631.

de Torres, Z.M., Zhai, B., Jayaraman, S., Eleftheriadou, G., Winsbury, R., Yang, R. *et al.* (2016) Novel JAZ co-operativity and unexpected JA dynamics underpin *Arabidopsis* defence responses to *Pseudomonas syringae* infection. *The New Phytologist*, **209**(3), 1120–1134.

del Olmo, I., López, J.A., Vázquez, J., Raynaud, C., Piñeiro, M. & Jarillo, J.A. (2016) Arabidopsis DNA polymerase ε recruits components of Polycomb repressor complex to mediate epigenetic gene silencing. *Nucleic Acids Research*, **44**(12), 5597–5614.

Del Pozo, J.C. & Ramirez-Parra, E. (2014) Deciphering the molecular bases for drought tolerance in Arabidopsis autotetraploids. *Plant, Cell & Environment*, **37**(12), 2722–2737.

Deolu-Ajayi, A.O., Meyer, A.J., Haring, M.A., Julkowska, M.M. & Testerink, C. (2019) Genetic loci associated with early salt stress responses of roots. *iScience*, **21**, 458–473.

DePamphilis, C.W., Palmer, J.D., Rounsley, S., Sankoff, D., Schuster, S.C., Ammiraju, J.S.S. *et al.* (2013) The *Amborella* genome and the evolution of flowering plants. *Science (1979)*, **342**(6165), 1241089.

Diallo, A., Kane, N., Agharbaoui, Z., Badawi, M. & Sarhan, F. (2010) Heterologous expression of wheat VERNALIZATION 2 (TaVRN2) gene in Arabidopsis delays flowering and enhances freezing tolerance. *PLoS One*, **5**(1), e8690.

Ding, Y., Liu, N., Virlouvet, L., Riethoven, J.J., Fromm, M. & Avramova, Z. (2013) Four distinct types of dehydration stress memory genes in *Arabidopsis thaliana*. *BMC Plant Biology*, **13**(1), 229.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S. *et al.* (2013) STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*, **29**(1), 15–21.

Dyson, B.C., Allwood, J.W., Feil, R., Xu, Y., Miller, M., Bowsher, C.G. *et al.* (2015) Acclimation of metabolism to light in *Arabidopsis thaliana*: the glucose 6-phosphate/phosphate translocator GPT2 directs metabolic acclimation. *Plant, Cell & Environment*, **38**(7), 1404–1417.

Fan, H., Quan, S., Ye, Q., Zhang, L., Liu, W., Zhu, N. *et al.* (2023) A molecular framework underlying low-nitrogen-induced early leaf senescence in *Arabidopsis thaliana*. *Molecular Plant*, **16**(4), 756–774.

Fischer, S., Flis, P., Zhao, F.J. & Salt, D.E. (2022) Transcriptional network underpinning ploidy-related elevated leaf potassium in neo-tetraploids. *Plant Physiology*, **190**(3), 1715–1730.

Flagel, L.E. & Wendel, J.F. (2010) Evolutionary rate variation, genomic dominance and duplicate gene expression evolution during allotetraploid cotton speciation. *New Phytologist*, **186**(1), 184–193.

Gallardo, M.H., Bickham, J.W., Honeycutt, R.L., Ojeda, R.A. & Köhler, N. (1999) Discovery of tetraploidy in a mammal. *Nature*, **401**(6751), 341.

Gardiner, L.J., Joynson, R., Omony, J., Rusholme-Pilcher, R., Olohan, L., Lang, D. *et al.* (2018) Hidden variation in polyploid wheat drives local adaptation. *Genome Research*, **28**(9), 1319–1332.

Gaudinier, A., Rodriguez-Medina, J., Zhang, L., Olson, A., Liseron-Monfils, C., Bågman, A.M. *et al.* (2018) Transcriptional regulation of nitrogen-associated metabolism and growth. *Nature*, **563**(7730), 259–264.

Groth, M., **Stroud, H.**, **Feng, S.**, **Greenberg, M.V.C.**, **Vashisht, A.A.**, **Wohlschlegel, J.A.** *et al.* (2014) SNF2 chromatin remodeler-family proteins FRG1 and -2 are required for RNA-directed DNA methylation. *Proceedings of the National Academy of Sciences of the United States of America*, **111**(49), 17666–17671.

Gu, X., **Xu, T.** & **He, Y.** (2014) A histone H3 Lysine-27 methyltransferase complex represses lateral root formation in *Arabidopsis thaliana*. *Molecular Plant*, **7**(6), 977–988.

Guo, M., **Davis, D.** & **Birchler, J.A.** (1996) Dosage effects on gene expression in a maize ploidy series. *Genetics*, **142**(4), 1349–1355.

He, C., **Chen, X.**, **Huang, H.** & **Xu, L.** (2012) Reprogramming of H3K27me3 is critical for acquisition of pluripotency from cultured Arabidopsis tissues. *PLoS Genetics*, **8**(8), e1002911.

Heinz, S., **Benner, C.**, **Spann, N.**, **Bertolino, E.**, **Lin, Y.C.**, **Laslo, P.** *et al.* (2010) Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Molecular Cell*, **38**(4), 576–589.

Huang, F., **Zhu, Q.**, **Zhu, A.**, **Wu, X.**, **Xie, L.**, **Wu, X.** *et al.* (2017) Mutants in the imprinted *PICKLE RELATED 2* gene suppress seed abortion of *fertilization independent seed* class mutants and paternal excess interploidy crosses in Arabidopsis. *The Plant Journal*, **90**(2), 383–395.

Huang, K.C., **Lin, W.C.** & **Cheng, W.H.** (2018) Salt hypersensitive mutant 9, a nucleolar APUM23 protein, is essential for salt sensitivity in association with the ABA signaling pathway in Arabidopsis. *BMC Plant Biology*, **18**(1), 1–21.

Huang, S., **Zhang, A.**, **Jin, J.B.**, **Zhao, B.**, **Wang, T.**, **Wu, Y.** *et al.* (2019) *Arabidopsis* histone H3K4 demethylase JMJ 17 functions in dehydration stress response. *The New Phytologist*, **223**(3), 1372–1387.

Humphrey, T.V., **Haasen, K.E.**, **Aldea-Brydges, M.G.**, **Sun, H.**, **Zayed, Y.**, **Indriolo, E.** *et al.* (2015) PERK–KIPK–KCBP signalling negatively regulates root growth in *Arabidopsis thaliana*. *Journal of Experimental Botany*, **66**(1), 71–83.

Jamge, B., **Lorković, Z.J.**, **Axelsson, E.**, **Osakabe, A.**, **Shukla, V.**, **Yelagandula, R.** *et al.* (2023) Histone variants shape the chromatin states in Arabidopsis. *bioRxiv* [Internet]. 2023.03.08.531698. Available from: https://doi.org/10.1101/2023.03.08.531698v1

Jordan, K.W., **He, F.**, **De Soto, M.F.**, **Akhunova, A.** & **Akhunov, E.** (2020) Differential chromatin accessibility landscape reveals structural and functional features of the allopolyploid wheat chromosomes. *Genome Biology*, **21**(1), 1–30.

Khan, A. & **Mathelier, A.** (2017) Intervene: a tool for intersection and visualization of multiple gene or genomic region sets. *BMC Bioinformatics*, **18**(1), 1–8.

Kiani, K., **Sanford, E.M.**, **Goyal, Y.** & **Raj, A.** (2022) Changes in chromatin accessibility are not concordant with transcriptional changes for single-factor perturbations. *Molecular Systems Biology*, **18**(9), 1–17.

Kohorn, B.D., **Kohorn, S.L.**, **Saba, N.J.** & **Martinez, V.M.** (2014) Requirement for pectin methyl esterase and preference for fragmented over native pectins for wall-associated kinase-activated, EDS1/PAD4-dependent stress response in arabidopsis. *Journal of Biological Chemistry*, **289**(27), 18978–18986.

Langmead, B. & **Salzberg, S.L.** (2012) Fast gapped-read alignment with bowtie 2. *Nature Methods*, **9**(4), 357–359.

Lee, J., **Yun, J.Y.**, **Zhao, W.**, **Shen, W.H.** & **Amasino, R.M.** (2015) A methyltransferase required for proper timing of the vernalization response in *Arabidopsis*. *Proceedings of the National Academy of Sciences of the United States of America*, **112**(7), 2269–2274.

Lee, S. & **Park, C.M.** (2012) Regulation of reactive oxygen species generation under drought conditions in Arabidopsis. *Plant Signaling & Behavior*, **7**(6), 599–601.

Li, H., **Handsaker, B.**, **Wysoker, A.**, **Fennell, T.**, **Ruan, J.**, **Homer, N.** *et al.* (2009) The sequence alignment/map format and SAMtools. *Bioinformatics*, **25**(16), 2078–2079.

Li, P., **Li, X.** & **Jiang, M.** (2021) CRISPR/Cas9-mediated mutagenesis of WRKY3 and WRKY4 function decreases salt and me-JA stress tolerance in *Arabidopsis thaliana*. *Molecular Biology Reports*, **48**(8), 5821–5832.

Li, Z., **McKibben, M.T.W.**, **Finch, G.S.**, **Blischak, P.D.**, **Sutherland, B.L.** & **Barker, M.S.** (2021) Patterns and processes of diploidization in land plants. *Annual Review of Plant Biology*, **72**, 387–410.

Lisso, J., **Altmann, T.** & **Müssig, C.** (2006) The AtNFXL1 gene encodes a NF-X1 type zinc finger protein required for growth under salt stress. *FEBS Letters*, **580**(20), 4851–4856.

Liu, J., **Ghelli, R.**, **Cardarelli, M.** & **Geisler, M.** (2022) Arabidopsis TWISTED DWARF1 regulates stamen elongation by differential activation of ABCB1,19-mediated auxin transport. *Journal of Experimental Botany*, **73**(14), 4818–4831.

Liu, L., **Wang, H.**, **Fu, Y.**, **Tang, W.**, **Zhao, P.**, **Ren, Y.** *et al.* (2023) Turnip crinkle virus-encoded suppressor of RNA silencing interacts with *Arabidopsis* SGS3 to enhance virus infection. *Molecular Plant Pathology*, **24**(2), 154–166.

Love, M.I., **Huber, W.** & **Anders, S.** (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, **15**(12), 550.

Lu, K.J., **De Rybel, B.**, **van Mourik, H.** & **Weijers, D.** (2018) Regulation of intercellular TARGET OF MONOPTEROS 7 protein transport in the *Arabidopsis* root. *Development*, **145**(2), 2960.

Lun, A.T.L. & **Smyth, G.K.** (2014) De novo detection of differentially bound regions for ChIP-seq data using peaks and windows: controlling error rates correctly. *Nucleic Acids Research*, **42**(11), e95.

Lun, A.T.L. & **Smyth, G.K.** (2015) Csaw: a bioconductor package for differential binding analysis of ChIP-seq data using sliding windows. *Nucleic Acids Research*, **44**(5), e45.

Ma, X., **Nicole, M.C.**, **Meteignier, L.V.**, **Hong, N.**, **Wang, G.** & **Moffett, P.** (2015) Different roles for RNA silencing and RNA processing components in virus recovery and virus-induced gene silencing in plants. *Journal of Experimental Botany*, **66**(3), 919–932.

Maher, K.A., **Bajic, M.**, **Kajala, K.**, **Reynoso, M.**, **Pauluzzi, G.**, **West, D.A.** *et al.* (2018) Profiling of accessible chromatin regions across multiple plant species and cell types reveals common gene regulatory principles and new control modules. *The Plant Cell*, **30**(1), 15–36.

Manfield, I.W., **Devlin, P.F.**, **Jen, C.H.**, **Westhead, D.R.** & **Gilmartin, P.M.** (2007) Conservation, convergence, and divergence of light-responsive, circadian-regulated, and tissue-specific expression patterns during evolution of the Arabidopsis GATA gene family. *Plant Physiology*, **143**(2), 941–958.

Monroe, J.G., **Srikant, T.**, **Carbonell-Bejerano, P.**, **Becker, C.**, **Lensink, M.**, **Exposito-Alonso, M.** *et al.* (2022) Mutation bias reflects natural selection in *Arabidopsis thaliana*. *Nature*, **602**(7895), 101–105. Available from: https://doi.org/10.1038/s41586-021-04269-6

Montes, R.A.C., **Coello, G.**, **González-Aguilera, K.L.**, **Marsch-Martínez, N.**, **de Folter, S.** & **Alvarez-Buylla, E.R.** (2014) ARACNe-based inference, using curated microarray data, of *Arabidopsis thaliana* root transcriptional regulatory networks. *BMC Plant Biology*, **14**(1), 97. Available from: https://doi.org/10.1186/1471-2229-14-97

Ohno, S. (1970) *Evolution by gene duplication*. New York: Springer.

Ojeda-Rivera, J.O., **Oropeza-Aburto, A.** & **Herrera-Estrella, L.** (2020) Dissection of root transcriptional responses to low pH, aluminum toxicity and iron excess under pi-limiting conditions in Arabidopsis wild-type and stop1 seedlings. *Frontiers in Plant Science*, **11**, 1–16.

Osada, N. & **Innan, H.** (2008) Duplication and gene conversion in the *Drosophila melanogaster* genome. *PLoS Genetics*, **4**(12), e1000305.

Osborn, T.C., **Chris Pires, J.**, **Birchler, J.A.**, **Auger, D.L.**, **Chen, Z.J.**, **Lee, H.S.** *et al.* (2003) Understanding mechanisms of novel gene expression in polyploids. *Trends in Genetics*, **19**(3), 141–147.

Ou, J., **Liu, H.**, **Yu, J.**, **Kelliher, M.A.**, **Castilla, L.H.**, **Lawson, N.D.** *et al.* (2018) ATCSeqQC. *BMC Genomics*, **19**(1), 169.

Oustric, J., **Morillon, R.**, **Luro, F.**, **Herbette, S.**, **Martin, P.**, **Giannettini, J.** *et al.* (2019) Nutrient deficiency tolerance in citrus is dependent on genotype or ploidy level. *Frontiers in Plant Science*, **10**, 127. Available from: https://doi.org/10.3389/fpls.2019.00127

Parks, M.B., **Nakov, T.**, **Ruck, E.C.**, **Wickett, N.J.** & **Alverson, A.J.** (2018) Phylogenomics reveals an extensive history of genome duplication in diatoms (Bacillariophyta). *American Journal of Botany*, **105**(3), 330–347.

Parra-Nunez, P., **Fernández-Jiménez, N.**, **Pachon-Penalba, M.**, **Sanchez-Moran, E.**, **Pradillo, M.** & **Santos, J.L.** (2024) Synthetically induced *Arabidopsis thaliana* autotetraploids provide insights into the analysis of meiotic mutants with altered crossover frequency. *The New Phytologist*, **241**(1), 197–208.

Peng, M., **Bi, Y.M.**, **Zhu, T.** & **Rothstein, S.J.** (2007) Genome-wide analysis of Arabidopsis responsive transcriptome to nitrogen limitation and its regulation by the ubiquitin ligase gene NLA. *Plant Molecular Biology*, **65**(6), 775–797.

Potok, M.E., Zhong, Z., Picard, C.L., Liu, Q., Do, T., Jacobsen, C.E. *et al.* (2022) The role of ATXR6 expression in modulating genome stability and transposable element repression in Arabidopsis. *Proceedings of the National Academy of Sciences of the United States of America*, **119**(3), e2115570119.

Pratx, L., Wendering, P., Kappel, C., Nikoloski, Z. & Bäurle, I. (2023) Histone retention preserves epigenetic marks during heat stress-induced transcriptional memory in plants. *The EMBO Journal*, **42**(24), e113595.

Probst, A.V., Desvoyes, B. & Gutierrez, C. (2020) Similar yet critically different: the distribution, dynamics and function of histone variants. *Journal of Experimental Botany*, **71**(17), 5191–5204.

Qi, P.L., Zhou, H.R., Zhao, Q.Q., Feng, C., Ning, Y.Q., Su, Y.N. *et al.* (2022) Characterization of an autonomous pathway complex that promotes flowering in *Arabidopsis*. *Nucleic Acids Research*, **50**(13), 7380–7395.

Ramírez, F., Ryan, D.P., Grüning, B., Bhardwaj, V., Kilpert, F., Richter, A.S. *et al.* (2016) deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Research*, **44**(W1), W160–W165.

Ramírez-González, R.H., Borrill, P., Lang, D., Harrington, S.A., Brinton, J., Venturini, L. *et al.* (2018) The transcriptional landscape of polyploid wheat. *Science (1979)*, **361**(6403), eaar6089.

Rastogi, S. & Liberles, D.A. (2005) Subfunctionalization of duplicated genes as a transition state to neofunctionalization. *BMC Evolutionary Biology*, **5**(28), 28. Available from: https://doi.org/10.1186/1471-2148-5-28

Raudvere, U., Kolberg, L., Kuzmin, I., Arak, T., Adler, P., Peterson, H. *et al.* (2019) G:profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update). *Nucleic Acids Research*, **47**(W1), W191–W198.

Ravi, M. & Chan, S.W.L. (2010) Haploid plants produced by centromere-mediated genome elimination. *Nature*, **464**(7288), 615–618.

Reinhold, H., Soyk, S., Šimková, K., Hostettler, C., Marafino, J., Mainiero, S. *et al.* (2011) β-Amylase–like proteins function as transcription factors in *Arabidopsis*, controlling shoot growth and development. *The Plant Cell*, **23**(4), 1391–1403.

Renault, H., El Amrani, A., Berger, A., Mouille, G., Soubigou-Taconnat, L., Bouchereau, A. *et al.* (2013) γ-Aminobutyric acid transaminase deficiency impairs central carbon metabolism and leads to cell wall defects during salt stress in Arabidopsis roots. *Plant, Cell & Environment*, **36**(5), 1009–1018.

Reske, J.J., Wilson, M.R. & Chandler, R.L. (2020) ATAC-seq normalization method can significantly affect differential accessibility analysis and interpretation. *Epigenetics & Chromatin*, **13**(1), 1–17.

Rodrigues, V.L. & Dolde, U.S. (2021) A microprotein repressor complex in the shoot meristem controls the transition to flowering. *Plant Physiology*, **187**(1), 187–202. Available from: https://doi.org/10.1093/plphys/kiab235

Salman-Minkov, A., Sabath, N. & Mayrose, I. (2016) Whole-genome duplication as a key factor in crop domestication. *Nature Plants*, **2**(8), 1–4.

Sas-Nowosielska, H. & Bernas, T. (2016) Spatial relationship between chromosomal domains in diploid and autotetraploid *Arabidopsis thaliana* nuclei. *Nucleus*, **7**(2), 216–231.

Schlereth, A., Möller, B., Liu, W., Kientz, M., Flipse, J., Rademacher, E.H. *et al.* (2010) MONOPTEROS controls embryonic root initiation by regulating a mobile transcription factor. *Nature*, **464**(7290), 913–916.

Serres, M.H., Kerr, A.R.W., McCormack, T.J. & Riley, M. (2009) Evolution by leaps: gene duplication in bacteria. *Biology Direct*, **4**, 1–17.

Shaked, H., Avivi-Ragolsky, N. & Levy, A.A. (2006) Involvement of the Arabidopsis *SWI2/SNF2* chromatin remodeling gene family in DNA damage response and recombination. *Genetics*, **173**(2), 985–994.

Shanks, C.M., Rothkegel, K., Brooks, M.D., Cheng, C.Y., Alvarez, J.M., Ruffel, S. *et al.* (2024) Nitrogen sensing and regulatory networks: it's about time and space. *The Plant Cell*, **36**(5), 1482–1503.

Sierro, N., Battey, J.N.D., Bovet, L., Liedschulte, V., Ouadi, S., Thomas, J. *et al.* (2018) The impact of genome evolution on the allotetraploid *Nicotiana rustica* – an intriguing story of enhanced alkaloid production. *BMC Genomics*, **19**(1), 1–18.

Sijacic, P., Bajic, M., McKinney, E.C., Meagher, R.B. & Deal, R.B. (2018) Changes in chromatin accessibility between Arabidopsis stem cells and mesophyll cells illuminate cell type-specific transcription factor networks. *Plant Journal*, **94**(2), 215–231.

Sowdhamini, R., Shameer, K., Ambika, S., Varghese, S.M., Karaba, N. & Udayakumar, M. (2009) stifdb arabidopsis stress responsive transcription factor database. *International Journal of Plant Genomics*, **2009**, 583429.

Spooner, D.M., Gavrilenko, T., Jansky, S.H., Ovchinnikova, A., Krylova, E., Knapp, S. *et al.* (2010) Ecogeography of ploidy variation in cultivated potato (*Solanum* sect. Petota). *American Journal of Botany*, **97**(12), 2049–2060.

Stroud, H., Greenberg, M.V.C., Feng, S., Bernatavichute, Y.V. & Jacobsen, S.E. (2013) Comprehensive analysis of silencing mutants reveals complex regulation of the Arabidopsis methylome. *Cell*, **152**(1–2), 352–364.

Tahir, M.S., Karagiannis, J. & Tian, L. (2022) HD2A and HD2C co-regulate drought stress response by modulating stomatal closure and root growth in Arabidopsis. *Frontiers in Plant Science*, **13**, 1062722.

Tarasov, A., Vilella, A.J., Cuppen, E., Nijman, I.J. & Prins, P. (2015) Sambamba: fast processing of NGS alignment formats. *Bioinformatics*, **31**(12), 2032–2034.

Teshima, K.M. & Innan, H. (2008) Neofunctionalization of duplicated genes under the pressure of gene conversion. *Genetics*, **178**(3), 1385–1398.

Tsompana, M. & Buck, M.J. (2014) Chromatin accessibility: a window into the genome. *Epigenetics & Chromatin*, **7**(1), 1–16.

Tsukaya, H. (2013) Does ploidy level directly control cell size? Counterevidence from Arabidopsis genetics. *PLoS One*, **8**(12), 1–7.

Van de Peer, Y., Mizrachi, E. & Marchal, K. (2017) The evolutionary significance of polyploidy. *Nature Reviews. Genetics*, **18**, 411–424.

Vanstraelen, M., Inzé, D. & Geelen, D. (2006) Mitosis-specific kinesins in Arabidopsis. *Trends in Plant Science*, **11**(4), 167–175.

Wang, L., Cao, S., Wang, P., Lu, K., Song, Q., Zhao, F.J. *et al.* (2021) DNA hypomethylation in tetraploid rice potentiates stress-responsive gene expression for salt tolerance. *Proceedings of the National Academy of Sciences of the United States of America*, **118**(13), e2023981118. Available from: https://doi.org/10.1073/pnas.2023981118

Wang, M., Zhong, Z., Gallego-Bartolomé, J., Feng, S., Shih, Y.H., Liu, M. *et al.* (2023) Arabidopsis TRB proteins function in H3K4me3 demethylation by recruiting JMJ14. *bioRxiv* [Internet]. Available from: https://doi.org/10.1101/2023.01.10.523456

Wang, Y., Long, H., Yu, J., Dong, L., Wassef, M., Zhuo, B. *et al.* (2018) Histone variants H2A.Z and H3.3 coordinately regulate PRC2-dependent H3K27me3 deposition and gene expression regulation in mES cells. *BMC Biology*, **16**, 107.

Watanabe, S., Takahashi, N., Kanno, Y., Suzukia, H., Aoi, Y., Takeda-Kamiya, N. *et al.* (2020) The Arabidopsis NRT1/PTR FAMILY protein NPF7.3/NRT1.5 is an indole-3-butyric acid transporter involved in root gravitropism. *Proceedings of the National Academy of Sciences of the United States of America*, **117**(49), 31500–31509. Available from: https://doi.org/10.1073/pnas.2013305117

Weise, S.E., Liu, T., Childs, K.L., Preiser, A.L., Katulski, H.M., Perrin-Porzondek, C. *et al.* (2019) Transcriptional regulation of the glucose-6-phosphate/phosphate translocator 2 is related to carbon exchange across the chloroplast envelope. *Frontiers in Plant Science*, **10**, 1–13.

Wen, R., Torres-Acosta, J.A., Pastushok, L., Lai, X., Pelzer, L., Wang, H. *et al.* (2008) *Arabidopsis UEV1D* promotes Lysine-63-linked polyubiquitination and is involved in DNA damage response. *The Plant Cell*, **20**(1), 213–227.

Wendel, J.F. (2000) Genome evolution in polyploids. *Plant Molecular Biology*, **42**(1), 225–249.

Wilkinson, S.W., Hannah Parker, A., Muench, A., Wilson, R.S., Hooshmand, K., Henderson, M.A. *et al.* (2023) Long-lasting memory of jasmonic acid-dependent immunity requires DNA demethylation and ARGONAUTE1. *Nature Plants*, **9**(1), 81–95.

Wolfe, K.H. (2001) Yesterday's polyploids and the mystery of diploidization. *Nature Reviews. Genetics*, **2**(5), 333–341.

Wolfe, K.H. (2015) Origin of the yeast whole-genome duplication. *PLoS Biology*, **13**(8), 1–7.

Wu, C.J., Yuan, D.Y., Liu, Z.Z., Xu, X., Wei, L., Cai, X.W. *et al.* (2023) Conserved and plant-specific histone acetyltransferase complexes cooperate to regulate gene transcription and plant development. *Nature Plants*, **9**(3), 442–459.

Wu, S., Han, B. & Jiao, Y. (2019) Genetic contribution of paleopolyploidy to adaptive evolution in angiosperms. *Molecular Plant*, **13**, 59–71.

Xiao, J., Jin, R., Yu, X., Shen, M., Wagner, J.D., Pai, A. *et al.* (2017) Cis and trans determinants of epigenetic silencing by Polycomb repressive complex 2 in Arabidopsis. *Nature Genetics*, **49**(10), 1546–1552.

Yang, D.L., Zhang, G., Wang, L., Li, J., Xu, D., Di, C. *et al.* (2018) Four putative SWI2/SNF2 chromatin remodelers have dual roles in regulating DNA methylation in Arabidopsis. *Cell Discovery*, **4**(1), 55.

Yang, F., **Wang, Q.**, **Schmitz, G.**, **Müller, D.** & **Theres, K.** (2012) The bHLH protein ROX acts in concert with RAX1 and LAS to modulate axillary meristem formation in Arabidopsis. *The Plant Journal*, **71**(1), 61–70.

Yang, S., **Zhang, T.**, **Wang, Z.**, **Zhao, X.**, **Li, R.** & **Li, J.** (2022) Nitrilases NIT1/2/3 positively regulate flowering by inhibiting MAF4 expression in Arabidopsis. *Frontiers in Plant Science*, **13**, 889460.

Yelagandula, R., **Stroud, H.**, **Holec, S.**, **Zhou, K.**, **Feng, S.**, **Zhong, X.** *et al.* (2014) The histone variant H2A.W defines heterochromatin and promotes chromatin condensation in Arabidopsis. *Cell*, **158**(1), 98–109.

Yen, M.R., **Suen, D.F.**, **Hsu, F.M.**, **Tsai, Y.H.**, **Fu, H.**, **Schmidt, W.** *et al.* (2017) Deubiquitinating enzyme OTU5 contributes to DNA methylation patterns and is critical for phosphate nutrition signals. *Plant Physiology*, **175**(4), 1826–1838.

Yoo, M.J. & **Wendel, J.F.** (2014) Comparative evolutionary and developmental dynamics of the cotton (*Gossypium hirsutum*) fiber transcriptome. *PLoS Genetics*, **10**(1), e1004073.

Yu, G., **Wang, L.G.** & **He, Q.Y.** (2015) ChIP seeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization. *Bioinformatics*, **31**(14), 2382–2383.

Yu, Z., **Haberer, G.**, **Matthes, M.**, **Rattei, T.**, **Mayer, K.F.X.**, **Gierl, A.** *et al.* (2010) Impact of natural genetic variation on the transcriptome of autotetraploid *Arabidopsis thaliana*. *Proceedings of the National Academy of Sciences of the United States of America*, **107**(41), 17809–17814.

Yuan, H., **Yu, H.**, **Huang, T.**, **Shen, X.**, **Xia, J.**, **Pang, F.** *et al.* (2019) The complexity of the *Fragaria* × *ananassa* (octoploid) transcriptome by single-molecule long-read sequencing. *Horticulture Research*, **6**(1), 46.

Zhang, J., **Zhang, X.**, **Tang, H.**, **Zhang, Q.**, **Hua, X.**, **Ma, X.** *et al.* (2018) Allele-defined genome of the autopolyploid sugarcane *Saccharum spontaneum* L. *Nature Genetics*, **50**(11), 1565–1573. Available from: https://doi.org/10.1038/s41588-018-0237-2

Zhang, X., **Niu, D.**, **Carbonell, A.**, **Wang, A.**, **Lee, A.**, **Tun, V.** *et al.* (2014) ARGONAUTE PIWI domain and microRNA duplex structure regulate small RNA sorting in Arabidopsis. *Nature Communications*, **5**(1), 5468.

Zhang, X., **Zhang, Q.**, **Gao, N.**, **Liu, M.**, **Zhang, C.**, **Luo, J.** *et al.* (2024) Nitrate transporters and mechanisms of nitrate signal transduction in *Arabidopsis* and rice. *Physiologia Plantarum*, **176**(4), e14486.

Zhang, Y., **Liu, T.**, **Meyer, C.A.**, **Eeckhoute, J.**, **Johnson, D.S.**, **Bernstein, B.E.** *et al.* (2008) Model-based analysis of ChIP-seq (MACS). *Genome Biology*, **9**(9), R137.

Zhang, Y., **Shen, Q.**, **Leng, L.**, **Zhang, D.**, **Chen, S.**, **Shi, Y.** *et al.* (2021) Incipient diploidization of the medicinal plant perilla within 10,000 years. *Nature Communications*, **12**(1), 5508.

Zhao, J., **Missihoun, T.D.** & **Bartels, D.** (2017) The role of Arabidopsis aldehyde dehydrogenase genes in response to high temperature and stress combinations. *Journal of Experimental Botany*, **68**(15), 4295–4308.

Zhu, Z., **Krall, L.**, **Li, Z.**, **Xi, L.**, **Luo, H.**, **Li, S.** *et al.* (2024) Transceptor NRT1.1 and receptor-kinase QSK1 complex controls PM H+-ATPase activity under low nitrate. *Current Biology*, **34**(7), 1479–1491.e6.

Zilberman, D., **Coleman-Derr, D.**, **Ballinger, T.** & **Henikoff, S.** (2008) Histone H2A.Z and DNA methylation are mutually antagonistic chromatin marks. *Nature*, **456**(7218), 125–129.