Small RNAs regulation and genomic harmony: insights into allopolyploid evolution in marsh orchids (*Dactylorhiza*)

Mimmi C. Eriksson[1,2,3*], Matthew Thornton[1], Emiliano Trucchi[1,4], Thomas M. Wolfe[1,5], Francisco Balao[1,6], Mikael Hedrén[7], Ovidiu Paun[1*]

[1]Botany and Biodiversity Research, University of Vienna, Rennweg 14, A-1030 Vienna, Austria. [2]Vienna Graduate School of Population Genetics, Veterinärplatz 1, A-1210 Vienna, Austria. [3]Department of Plant Biology, Swedish University of Agricultural Sciences, SE-750 07 Uppsala, Sweden. [4]Marche Polytechnic University, Ancona, Italy. [5]Institute of Forest Entomology, Forest Pathology and Forest Protection (IFFF), BOKU, Peter-Jordan-Straße 82/I, 1190 Vienna, Austria. [6]Departamento de Biologia Vegetal y Ecologia, University of Seville, Sevilla, Spain. [7]Department of Biology, Lund University, Sölvegatan 37, S-22362, Lund, Sweden.

*Authors for correspondence: mimmi.c.eriksson@gmail.com; ovidiu.paun@univie.ac.at

ORCID
Mimmi C. Eriksson 0000-0002-9548-8226
Emiliano Trucchi 0000-0002-1270-5273
Thomas M. Wolfe 0000-0002-1509-0147
Francisco Balao 0000-0003-2104-3846
Mikael Hedrén 0000-0001-8194-4617
Ovidiu Paun 0000-0002-8295-4937

Total word count	5458	Acknowledgements:	169
Summary	196	Author contributions:	90
Introduction:	1094	No. figures:	5
Materials and Methods:	1069	No. tables:	0
Results:	1426	No Supporting information:	19 (Figs S1-S10, Tables S1-S9)
Discussion:	1869		

SUMMARY

- Hybridization and polyploidy are major drivers of plant diversification, often accompanied by shifts in gene expression and genome composition. Small RNAs (smRNAs) are thought to influence such genomic changes, particularly through their interactions with transposable elements (TEs).
- We comparatively quantified smRNAs in established sibling allopolyploids *Dactylorhiza majalis* and *D. traunsteineri* and their diploid progenitors to assess how independent allopolyploidization events shaped smRNA landscapes.
- Despite independent origins, the allotetraploids exhibited substantial overlap in smRNA composition, including transgressive accumulation of smRNA near genes related to transcriptional regulation, cell division and stress response. Consistently, TE-associated 24 nt smRNAs more closely resembled the paternal and larger genome, while shorter smRNAs typically reflected the maternal and smaller genome. Nevertheless, distinct patterns were also evident: the older *D. majalis* showed greater accumulation of smRNAs near genes involved in transcriptional and translational regulation, while the younger *D. traunsteineri* displayed stronger non-additive patterns, suggesting ongoing resolution of post-polyploid meiotic and mitotic instability.
- Our results reveal both convergence and divergence in smRNA landscapes among independently formed allopolyploids. Our study highlights the central role of smRNAs in resolving genomic conflict, with possible implications for functional divergence and ecological innovation during polyploid evolution.

Key words

allopolyploidy, *Dactylorhiza*, gene regulation, marsh orchids, small RNAs, subgenome dominance, transposable elements

Plain-language summary

We studied two sibling marsh orchid species that arose through hybridization and genome doubling. By comparing their small RNA molecules and how these associate with genes, we found both shared and species-specific patterns. These differences reflect how the orchids' genomes have changed since formation, and may explain how repeated polyploidy events contribute to genetic and ecological diversity in plants.

INTRODUCTION

Allopolyploidization—the combination of whole-genome duplication (WGD) and interspecific hybridization—can be a potent evolutionary force, generating genetic diversity, gene redundancy, and buffering capacity that may contribute to evolutionary innovation and ecological diversification. However, the initial merger of divergent genomes into a single nucleus can destabilize genome regulation, triggering “genomic shock” (McClintock, 1984), on the background of a severe genetic bottleneck.

Small noncoding RNAs (smRNAs) have emerged as key mediators during post-polyploidization genome stabilization (e.g., Ha *et al.*, 2009; Shen *et al.*, 2017; Cavé-Radet *et al.*, 2020). These short RNA molecules suppress transposable elements (TEs) and fine-tune gene expression, promoting both genome integrity and adaptive responses. Specifically, repeat-associated or heterochromatic small-interfering RNAs (ra-siRNAs/het-siRNAs; Shen *et al.*, 2017) help maintain chromatin stability, while microRNAs (miRNAs) can modulate gene expression with phenotypic implications (Ha *et al.*, 2009; Ghani *et al.*, 2014; Jiang *et al.*, 2021). In addition, plant smRNAs orchestrate abiotic and biotic stress responses (Kaur *et al.*, 2021; Tiwari & Rajam, 2022; Zhan & Meyers, 2024), potentially facilitating adaptation to novel, non-parental environments. These dual roles make smRNAs integral to allopolyploid success.

In general, smRNAs operate through three main mechanisms: (1) post-transcriptional gene silencing (PTGS), binding to RNAs to repress translation or promote cleavage and degradation; (2) transcriptional gene silencing (TGS), directing epigenetic silencing at homologous genomic loci; and (3) splicing regulation, interacting with pre-mRNAs to influence splice-site selection (Borges & Martienssen, 2015; Wang & Chekanova, 2016). In plants, smRNAs are classified by biogenesis into miRNAs (20–22 nt long, hairpin-derived), and small-interfering RNAs (siRNAs, from double stranded RNA) (Borges & Martienssen, 2015; Chen & Rechavi, 2021), with siRNAs further divided into: (*i*) ra-siRNAs (24 nt and primarily involved in TGS), (*ii*) secondary siRNAs (~21–22nt long, which can function in both TGS and PTGS pathways), and (*iii*) natural antisense transcript siRNA (nat-siRNA, 21–24 nt, acting both in *cis* and *trans* via PTGS) (Borges & Martienssen, 2015; Chen a& Rechavi, 2021).

Recent reviews (e.g., Ariel & Manavella, 2021; Tossolini *et al.*, 2025) highlight the pivotal link between TE mobilization and smRNA evolution. TEs not only contribute regulatory capacity through enhancer-derived transcripts (e.g., in wheat; TE-initiated enhancer-like elements, Xie *et al.*, 2023), but also spur the emergence of smRNAs and long non-coding RNAs that

diversify gene regulation and support genome plasticity (Hassan *et al.*, 2024; Tossolini *et al.*, 2025).

WGD and environmental stress are known to lead to TE activation, resulting in accumulation of repeat-derived smRNAs, boosting both TGS and PTGS pathways (Ito *et al.*, 2011; Kenan-Eichler *et al.*, 2011; Marí-Ordóñez *et al.*, 2013; Fultz *et al.*, 2015; Xiao & Wang, 2025). For example, resynthesized *Arabidopsis* allotetraploids rapidly restore ra-siRNA populations after WGD, while miRNAs show non-additive expression patterns and play key roles in stabilizing hybrid genomes (Ha *et al.*, 2009). Similarly, early generations *Brassica napus* allotetraploids exhibit a surge in TE-derived 21 nt siRNAs, especially from Gypsy elements, helping to silence newly active TEs (Martinez Palacios *et al.*, 2019). Cavé-Radet *et al.* (2020, 2023) observed stronger miRNA responses in allopolyploid *Spartina* compared to homoploid hybrids, suggesting a more pronounced regulatory impact of WGD than hybridization alone (Cavé-Radet *et al.*, 2020). On the other hand, the divergence between progenitors determines the extent of genomic shock in wheat allotetraploids, including enhanced smRNA responses associated with TEs in hybrid crosses between parents with differing TE load, compared to those with similar TE content (Jiao *et al.*, 2018).

Despite growing interest in the role of smRNAs in allopolyploid evolution, most research has focused on synthetic or early-generation polyploids. To our knowledge, only Cavé-Radet *et al.*, (2020) have examined smRNAs in wild allopolyploids, albeit also in early stages (*Spartina* F_1 hybrids and *S. anglica*, ~170 years old). Thus, the long-term dynamics, regulatory roles and potential ecological relevance of smRNAs in more established allopolyploid lineages remain poorly explored.

In this study, we focus on two ecologically-distinct, established, sibling orchid allotetraploids, *Dactylorhiza majalis s.str.* (Rchb.) P.F.Hunt & Summerh. and *D. traunsteineri* (Saut. ex Rchb.) Soó, and their parents, to address the long-term fate and potential roles of smRNAs following allopolyploidization. The marsh orchids within *D. majalis s.l.* include several independently formed sibling allotetraploids, each likely with multiple origins involving overlapping stocks of parental populations (Brandrud *et al.*, 2020). The focal allotetraploids here share a complex history of repeated polyploidization and/or hybridization and back-crossing, but are today reasonably well separated in morphology and habitat preferences (Fig. 1a; Wolfe *et al.*, 2023). They are estimated to have formed approximately 104k (*D. majalis*) and 74k (*D. traunsteineri*) generations ago (Hawranek, 2021), from the diploid progenitors *D. fuchsii* (Druce) Soó (maternal, smaller genome) and *D. incarnata* (L.) Soó (paternal, larger genome) (Fig. 1a) (Aagard *et al.*, 2005; Brandrud *et al.*, 2020; Eriksson *et al.*, 2022). Transcriptomic comparisons of the diploid progenitors have revealed large-scale gene expression differences, importantly, including differences in smRNAs associated with

TEs and PTGS pathways (Balao *et al.*, 2017). Further, Eriksson *et al.* (2022) reported significant differences in both genome size and TE load between the two diploid progenitors.

Distinct ecological, spatial, morphological and physiological properties have been reported for the two sibling allopolyploids (Paun *et al.*, 2010, 2011; Balao *et al.*, 2016; Wolfe *et al.*, 2023). *D. majalis* is distributed across continental Europe, ranging from the Pyrenees to Scandinavia, with a broader ecological niche (i.e., mesic to moist meadows) and less geographically-driven genetic differentiation than *D. traunsteineri* which is found in disjunct marshes and fens in the Alps, Scandinavia and British Isles. Wolfe and colleagues (2023) reported physiological differences in photosynthesis and nutrient transport consistent with the contrasting ecological conditions preferred by the two allotetraploids. These appear to have helped the two allotetraploids adapt to different soil chemistry conditions, with *D. traunsteineri* restricted to wet, nutrient-poor soils, in particular with extremely low levels of available nitrate (Wolfe *et al.*, 2023). Furthermore, reciprocal transplant experiments have shown that much of the gene expression divergence between the sibling allopolyploids is environmentally plastic, and appeared to be associated to distinct, environmentally-driven, biotic communities in their roots (Emelianova *et al.*, 2025). In stark contrast, analysis of the genomes of several *Dactylorhiza* allopolyploids found relatively little difference between *D. majalis* and *D. traunsteineri* both in genome size and TE activity, suggesting a strong control of TEs following recurrent allopolyploidizations (Eriksson *et al.*, 2022).

In this study, we specifically investigate: (1) how smRNA abundance in the allopolyploids reflects parental patterns—whether it follows additive, dominant, or transgressive trends—and how consistent these patterns are across the sibling allopolyploids; (2) the extent and similarity of smRNA association with TEs in the two sibling allotetraploids; and (3) how variation in smRNA patterns may relate to distinct eco-physiological traits of *D. majalis* and *D. traunsteineri*.

MATERIALS AND METHODS

Sampling

Adult specimens of each species of interest were transplanted from multiple localities across Europe to a common garden in Vienna, Austria. *Dactylorhiza* includes perennial plants that produce a new tuber yearly, which supports growth in the following year. Two years after transplantation, in the morning of 14 May 2014, leaf tissue at a similar developmental stage for all accessions was fixed in RNAlater (Sigma) and stored at -80°C until further processing. Our final sampling included four accessions of each diploid (*D. incarnata* and *D. fuchsii*), together with nine (*D. majalis*) and twelve (*D. traunsteineri*) individuals of each allotetraploid

(Supplementary Table S1). With few exceptions, each accession of each species originated from different localities across Europe. The diploid smRNA data has been previously investigated by Balao *et al.* (2017). Independent RNA isolations from the same tissue fixations of *D. majalis* and *D. traunsteineri* were used for RNA-seq analyses by Wolfe *et al.* (2023).

Library preparations and sequencing

Small RNA extraction, library preparation and sequencing for this experiment was previously described for diploid investigations (Balao *et al.*, 2017). In short, isolation of smRNA was performed with the mirVana miRNA Isolation Kit (Life Technologies). The smRNA-enriched RNA extracts were further purified by gel size selection on 15% TBE-Urea pre-casted gels (Life Technologies). The small RNA libraries were prepared with the NEBNext Multiplex Small RNA Library Prep Set for Illumina (New England Biolabs). Illumina sequencing as 50 bp single-end reads was performed at the Vienna BioCenter Core Facilities (VBCF; https://www.viennabiocenter.org/).

Read processing and quantifying

The libraries were demultiplexed using BamIndexDecoder v.1.03 (http://wtsi-npg.github.io/illumina2bam/#BamIndexDecoder). Raw reads were adapter-trimmed, size-selected for the 20-24 nt range, and then quality filtered using CLC Genomics Workbench v.8.0 (QIAGEN). Filtered reads were mapped to the *Dactylorhiza incarnata* reference genome v.1.0 (Wolfe *et al.*, 2023) with STAR v.2.7.3a (Dobin *et al.*, 2013), using --alignIntronMax 1 and --outFilterMismatchNoverLmax 0.05. A single-pass alignment approach was used, retaining multi-mapping reads below a threshold of 100 (--outFilterMulitmapNmax 100).

An annotation-free approach was initially used to identify genomic regions (genic and non-genic) that were potentially “targeted” by smRNAs, defined here as regions showing sequence homology with smRNAs and therefore indicating an association between smRNAs and the corresponding genomic sequence. Throughout this study, we use the term *“targeting”* in a purely operational sense to describe genomic regions showing sequence complementarity to smRNAs. This definition does not assume a direct causal or regulatory interaction between smRNAs and the corresponding genomic loci. Rather, it denotes potential sites of smRNA association inferred from sequence homology. While such associations often coincide with transcriptionally or post-transcriptionally regulated regions (e.g. in the case of heterochromatic siRNAs or miRNAs; Borges & Martienssen, 2015), they

may also reflect indirect or distal effects, or even passive sequence associations without immediate regulatory function.

After sorting bam files by reference coordinates with Samtools v.1.10 (Danacek *et al.*, 2011), smRNA read coverage was assessed for each individual in non-overlapping 100 bp windows across the genome using bamCoverage from deepTools2 v.3.5.1 (Ramirez *et al.*, 2016). Windows with a minimum coverage of ten reads in at least one sample were retained in a bed file, and adjacent windows were merged using bedtools merge v.2.29.2 (Quinlan & Hall, 2010). The resulting bed file of retained regions (i.e., peaks of smRNA "targeting") was converted to gff3 format using a custom python script. Counts were then obtained with featureCounts from Subread v.2.0.0 (Liao *et al.*, 2014), accounting for multiple mappers fractionally (-M --fraction), to include potential trans-associations. The full pipeline, scripts, raw count file, and peak gff3 file are available at https://github.com/mc-er/dact-sRNAs.

Differential smRNA "targeting" analysis

Differently "targeted" (DT) regions were identified following a similar approach to that used for detecting differentially expressed genes, using counts of smRNAs at peaks as input. Analyses were performed in R v.4.1.1 and the Bioconductor package EdgeR v.3.34.0 (Robinson *et al.*, 2010; McCarthy *et al.*, 2012). The count matrix obtained from featureCounts was filtered to retain genomic regions with a minimum smRNA mapping of one count per million (CPM) in at least 25% of samples within a group and a minimum of two samples overall. To account for differences in library size, data were normalised using the weighted trimmed mean of M-values (TMM) method. A quasi-likelihood negative binomial model was applied to identify DT regions across the genome, comparing either (i) the two sibling allotetraploids or (ii) each allotetraploid to its parental diploid. P-values were corrected for multiple testing using the Benjamini–Hochberg (BH) procedure, and only genomic regions passing an FDR < 0.05 were kept for further analysis.

Analyses of genomic regions around TEs were performed separately for reads of 20–23 nt, and for 24 nt reads (the latter representing potentially heterochromatic smRNAs, hetsiRNAs; Borges & Martienssen, 2015). For analyses of genic regions, all read lengths (20–24 nt) were combined into a single analysis per comparison.

Assessing genomic interactions

Differentially targeted peaks passing an FDR threshold of 0.05 were classified into three categories of genomic interactions (Supplementary Table S2): (1) additivity—the polyploid showed lower targeting relative to one diploid and higher targeting relative to the other, (2)

transgressive—the polyploid showed increased or decreased targeting relative to both diploids, and (3) dominance—the polyploid showed no significant difference relative to one diploid (the “dominant” parent) but differed significantly relative to the other diploid (illustrated on the Y-axis of Fig. 4c-f). Deviation from the stochastically expected number of shared peaks between the two tetraploids in each category was tested using a hypergeometric test implemented in the Bioconductor package SuperExactTest (Wang *et al.*, 2015).

Functional interpretation of differentially targeted genes

To gain insights into potential roles for smRNA associations in shaping the divergent ecological properties of the sibling allopolyploids, we performed gene ontology (GO) enrichment analysis of DT genes. Because the molecular consequences of smRNA association depend on the location of the targeted region, we analysed regions around promoters separately from those within exons or introns. Genes associated with smRNAs around promoters would be expected to be transcriptionally down-regulated or silenced, whereas associations within gene bodies may indicate post-transcriptional effects or altered splicing. GO enrichment analyses for genes potentially affected by DT peaks were performed in R v.4.1.1 using the Bioconductor package topGO v.2.44.0 (Alexa & Rahnenfuhre, 2021). Fisher’s exact tests were run using the “weight01” algorithm to test for over-representation of specific functional categories. P-values were adjusted for multiple testing using the FDR method, with significance set at FDR < 0.05.

RESULTS

smRNAs and their genomic targets

Sequencing produced an average 13.5 M reads per accession (±5.9 M SD). After size selection and quality filtering, an average of 5.8 M reads (±3.2 M SD) of 20–24 nt were retained (Supplementary Table S3). On average, 83.1% (±4.2% SD) of these reads mapped to the *D. incarnata* reference genome, with no apparent mapping bias across species (Supplementary Table S3). Read length distributions matched the expected profile for smRNAs (Borges & Martienssen, 2015), with modes at 21–22 nt and 24 nt in all species (Fig. 2a).

D. fuchsii exhibited higher CPM of 20–22 nt smRNAs compared to *D. incarnata*, with a significant difference observed only for 21 nt smRNAs (Fig 2a–c). The reverse trend was found for the 23–24 nt smRNAs, which were more abundant in *D. incarnata* (Fig. 2a). The two allotetraploids were more similar to each other than the diploids (Fig. 2a). For 21 nt smRNAs, both allotetraploids displayed significantly higher CPM than the paternal species

(*D. incarnata*), but significantly lower CPM than the maternal species (*D. fuchsii*) (Fig. 2a). A large proportion of 20–22 nt smRNAs mapped to exonic regions, while 24 nt smRNAs predominantly mapped to TEs and intergenic regions (Fig. 2b–c, Supplementary Fig. S1).

Our annotation-free approach identified 2,454,439 putative smRNA peaks across the reference genome. After filtering for CPM > 1 in at least two samples and 25% of samples per group, between 102,222 and 370,440 peaks were retained depending on the comparison. A multidimensional scaling (MDS) plot based on all 20–24 nt smRNAs across 212,434 peaks clearly separated the two diploid species, and distinguished diploids from allotetraploids, with the latter overlapping centrally (Fig. 1b). Within *D. traunsteineri* a geographical clustering was apparent (not shown).

smRNA patterns in ecologically divergent allopolyploids

We compared the two sibling allotetraploids using the combined 20–24 nt reads, retaining 152,503 peaks passing the CPM threshold (Supplementary Table S4). Among these, 2,106 DT peaks were identified between *D. traunsteineri* and *D. majalis*, of which approximately half (1,016 peaks) were associated with genes. Gene-associated peaks were defined as those overlapping a gene model (exons or introns) or within 1,000 bp upstream of a gene model (hereafter referred to as ‘promoters’; Supplementary Table S4).

Most genic DT peaks (762) were located within introns, and 82 overlapped multiple genic features (intron, exon, and/or promoters) (Supplementary Table S4). Across all genic categories, the older allotetraploid *D. majalis* displayed approximately three times as many peaks with high smRNA targeting compared to *D. traunsteineri* (Supplementary Table S4, Fig. 3a–c).

Comparison of DT peaks with previously identified differentially expressed (DE) genes from the same experiment (Wolfe *et al.*, 2023) revealed that most genes with DT peaks did not show differential expression between the two allopolyploids, consistent with potential post-transcritional effects or alternative splicing rather than direct transcriptional regulation (Fig. 3f, Supplementary Table S5).

GO enrichment analyses of genes associated with DT peaks between the allotetraploids revealed more enriched terms in *D. majalis* than *D. traunsteineri* (Fig. 5a, Supplementary Table S6). Most enriched terms were for genes targeted within gene bodies and related to transcription/translation regulation, transmembrane transport, response to biotic and abiotic stress, DNA repair, and cell cycle control (Supplementary Table S6).

smRNA targeting and genomic interactions

We further compared 20-24 nt smRNA patterns of each allotetraploid to those of the parental diploids. After filtering, 212,434 peaks were retained (Supplementary Table S4). Both allotetraploids showed broadly similar targeting patterns relative to the diploids, although the younger *D. traunsteineri* displayed greater divergence than *D. majalis* (Supplementary Table S4, Supplementary Fig. S2). Allotetraploids showed larger differences from the maternal parent (*D. fuchsii*) than from the paternal *D. incarnata* (Supplementary Table S4, Supplementary Fig. S2).

DT peaks were classified into four interaction types: additive, transgressive, and dominance towards either diploid. The majority of peaks showed no change relative to either diploid. Among DT peaks, non-additive patterns predominated, with dominance most frequent, particularly towards *D. incarnata* (Supplementary Table S7, Fig. 4a–f, Supplementary Fig. S3–S4). Notably, in intergenic regions lacking annotated TEs, a higher proportion of smRNA peaks showing dominance towards *D. fuchsii* than in other genomic contexts (Supplementary Table S7, Fig. 4d, Supplementary Fig. S3). Among transgressive peaks, both allotetraploids most often displayed higher smRNA targeting relative to both diploids (Supplementary Table S7, Fig. 4a, Fig. 4c–f, Supplementary Fig. S3–S4). A notable general pattern across interaction types was that a large proportion of DT peaks were shared between the sibling allotetraploids (Supplementary Table S7, Fig. 4b–c, Supplementary Fig. S3–S4), despite their independent origins (Brandrud *et al.*, 2020; Hawranek, 2021).

We also examined the distribution of interaction categories across genic regions (promoters, exons, introns). In *D. majalis*, nearly 75% promoter-associated DT peaks showed dominance towards *D. incarnata*, and almost 50% showed lower targeting than *D. fuchsii* (Supplementary Fig. S5). In contrast, in *D. traunsteineri* dominance toward *D. fuchsii* was more frequent across all genic regions (Supplementary Fig. S5).

Transgressively targeted genes in both allotetraploids were enriched for similar functional categories, including stress response, transport, transcription/translation regulation, and cell cycle control, though largely involving different genes and pathways in each species (Fig. 5b–c, Supplementary Tables S8–S9). In *D. traunsteineri*, transgressive peaks were associated with enrichment for cell cycle regulation and stress response functions, and the respective genes were targeted within exons or introns (Fig. 5b, Supplementary Tables S8–S9). In *D. majalis*, transgressive peaks occurred mostly within introns, and enrichments were associated with regulation of gene function (Fig. 5c, Supplementary Tables S8).

smRNAs associations with TEs

For TE-focused analyses, reads were divided into 24 nt smRNAs (likely het-siRNAs and therefore expected to be more frequently associated with TEs; Fig2c, Supplementary Fig. S1) and 20-23 nt smRNAs. After CPM filtering, for the comparative analysis of *D. majalis* versus *D. traunsteineri* more 24 nt peaks (315,838) passed the threshold than 20–23 nt peaks (102,222) (Supplementary Table S4). However, DT analysis uncovered the opposite trend: 3,242 DT peaks were identified for the 20–23 nt smRNAs, compared to only 103 DT peaks for the 24 nt smRNAs (Supplementary Table S4, Fig. 3d–e) suggesting a consistent targeting of TEs by het-siRNAs across allopolyploids. Very few TE-associated peaks were DT between the allotetraploids in both smRNA size classes simultaneously (Supplementary Fig. S6). In both datasets, approximately half of the DT peaks overlapped annotated TEs (Supplementary Table S4). Many of these were also located within gene introns, while a smaller number were found in the vicinity of gene promoters (Supplementary Table S4).

Similar to the results for the genic regions, *D. majalis* exhibited more frequent higher smRNA targeting of TEs than *D. traunsteineri* among DT peaks, except for 20–23 nt peaks associated with TEs located around promoters, where *D. traunsteineri* showed higher targeting (Supplementary Table S4). The majority of TE-associated DT peaks were located within LTR retrotransposons, especially Ty3-*gypsy* elements (Supplementary Fig. S7–S8). Among smRNAs targeting Ty3-*gypsy* elements, a slightly higher number targeted non-chromoviruses than chromoviruses (Supplementary Fig. S8).

Comparisons of allotetraploids versus diploids showed similar trends. After filtering, 139,736 (20-23 nt) and 370,440 (24 nt) peaks were retained (Supplementary Table S4). As with genic targets, DT peaks were classified into additive, transgressive, and dominance towards either diploid. In both datasets, a significant number of DT peaks were found to be shared between *D. majalis* and *D. traunsteineri* across interaction categories ($p < 0.001$, hypergeometric test; Fig 3d-e, Supplementary Fig. S3, Fig. 4e–f, Supplementary Fig. S7, Supplementary Fig. S9–10).

Dominance patterns again predominated, with more peaks showing dominance towards either diploid than additive or transgressive patterns. For 20–23 nt smRNAs, slightly more peaks showed dominance toward the diploid with lower targeting, whereas for 24 nt smRNAs the opposite trend was observed (Fig. 4e–f, Supplementary Fig. S3, Supplementary Fig. S9). The contrast between dominant types (higher vs. lower targeting) was less pronounced for 20–23 nt but more distinct for 24 nt smRNAs (Fig. 4e–f, Supplementary Fig. S3, Supplementary Fig. S9–10). Three general patterns observed for smRNA-associated gene regions were also evident for TEs. First, for most additive peaks, the smaller genome of *D. fuchsii* exhibited higher smRNA targeting than the larger *D. incarnata* (Fig. 4e–f, Supplementary Fig. S3, Supplementary Fig. S9–10). Second, for

transgressive peaks, both allotetraploids generally showed elevated smRNA targeting than either diploid (Fig. 4e–f, Supplementary Fig. S3, Supplementary Fig. S7, Supplementary Fig. S9–10). Finally, *D. traunsteineri* displayed greater differences from the parental diploids than *D. majalis* (Fig. 4e–f, Supplementary Fig. S3, Supplementary Fig. S7).

DISCUSSION

smRNAs are key regulators of genome integrity and gene expression in polyploids, but most insights so far come from recent or resynthesized polyploids. In this study, we examine natural polyploids that have evolved over extended evolutionary timescales to provide a comprehensive overview of potential smRNA contributions to genomic and transcriptomic diversification following repeated allopolyploidization in *Dactylorhiza majalis* s.l. We analysed smRNA expression in two established sibling allotetraploids and their diploid progenitors, all grown under common garden conditions, thus minimizing environmental sources of variation. Our primary objectives were to comparatively characterize (*i*) patterns of genomic interaction in the allopolyploids and (*ii*) smRNA associations to TE control and gene activity.

Inferring smRNA *"targeting"* from sequence homology offers an informative yet indirect view of smRNA–genome interactions. Many plant smRNAs—particularly 24 nt siRNAs—act locally to reinforce transcriptional silencing of homologous TEs or nearby genes (Matzke & Mosher, 2014; Borges & Martienssen, 2015). Trans-acting and distal effects have nonetheless been documented (Lewsey *et al.*, 2016; Liu & Chen, 2018), though the prevalence and biological impact of such long-distance effects remain less well characterized. Therefore, our analyses emphasize patterns of smRNA association rather than assuming direct functional regulation. These patterns highlight potential regions of epigenetic remodeling and provide a comparative framework for understanding smRNA landscape evolution after allopolyploidization.

We acknowledge that the present design—limited to representative individuals of the two extant diploid progenitors—cannot encompass the full spectrum of genetic and regulatory polymorphism that may have existed at the time of allopolyploid formation. Given that both parental lineages likely accumulated substantial variation over tens of thousands of generations, some of the observed differences between the allopolyploids may partly reflect ancestral or subsequent divergence in the diploid gene pools. Accordingly, we interpret our results in terms of general and reproducible regulatory patterns rather than exact reconstructions of the parental regulatory states.

Consistent dominance in the allopolyploids towards the paternal, larger genome

Allopolyploidization merges divergent genomes, giving rise to novel genetic and regulatory interactions. Common outcomes include epistasis and non-additive genetic or epigenetic effects, including smRNA-mediated gene regulation and TE silencing. Our results reveal a consistent dominance of smRNA targeting aligned to the paternal, larger subgenome (*D. incarnata*), especially for TE-associated 24 nt smRNAs. This pattern was observed across both sibling allotetraploids (Fig. 4, Supplementary Fig. S5), and resulted in a comparatively small number of DT peaks between them. Similar non-additive smRNA responses following hybridization and/or whole-genome duplication have been documented in *Arabidopsis* (Ha *et al.*, 2009), *Brassica* (Fu *et al.*, 2016; Martinez Palacios *et al.*, 2019), and *Spartina* (Cavé-Radet *et al.*, 2020), though it has remained unclear whether such patterns remain consistent across longer evolutionary timescales.

Examining whether smRNA expression is additive, dominant, or transgressive allows us to understand how regulatory interactions from the parental genomes combine in allopolyploids. Additive patterns indicate intermediate regulation and minimal epistatic interactions, dominance reflects the stronger influence of one parental subgenome, and transgressive patterns reveal novel regulatory states beyond the parental range. These distinctions are conceptually important because they highlight mechanisms of genome stabilization, resolution of redundant gene copies, and the generation of potentially adaptive phenotypes that can enable allopolyploids to occupy ecological niches outside the parental range (Ha et al., 2009; Fu et al., 2016; Chen et al., 2020).

Given that *D. incarnata* has a larger genome and higher TE load than *D. fuchsii* (Eriksson *et al.*, 2022), our findings are consistent with prior reports showing that smRNA-mediated silencing preferentially targets the TE-rich subgenome, as observed in wheat (Woodhouse *et al.*, 2014) and *Brassica rapa* (Cheng *et al.*, 2016). In contrast, gene expression dominance in allopolyploids frequently shifts toward the subgenome with fewer TEs, a pattern that can arise rapidly after allopolyploidization (for example in *Mimulus*; Edger *et al.*, 2017) and persist across generations (Gaebelein *et al.*, 2019; Liu & Wang, 2023).

While our data indicate stronger dominance towards *D. incarnata*, some patterns may be influenced by reference genome bias, as we only had access to a *D. incarnata* reference genome. Accurate differentiation between homologous and homeologous targeting remains challenging due to the short length of smRNAs and their tolerance for mismatches in plants. Nonetheless, mapping success was comparable across samples (ca. 83%; Supplementary Table S3), and intergenic smRNA dominance towards *D. fuchsii* was also observed, suggesting that reference bias alone cannot account for our observations. Lineage-specific TE degradation or divergence in *D. fuchsii* may have reduced smRNA detection from this

genome. Notably, *D. incarnata* and *D. fuchsii* diverged approximately 5.5 million years ago (Brandrud *et al.*, 2020).

Divergent targeting of TEs by 24 nt and 20-23 nt smRNAs

As expected, 24 nt smRNAs—typically het-siRNAs—were strongly associated with TEs, particularly LTR retrotransposons (Fig. 2c, Supplementary Fig. S1). Surprisingly, the number of DT peaks for 24 nt smRNAs between the allotetraploids was low, suggesting that TE silencing proceeds in a relatively consistent manner across these independently evolved lineages (Fig.3e). In contrast, 20–23 nt smRNAs exhibited markedly greater divergence between the sibling allopolyploids (Fig. 3d), consistent with a more dynamic, lineage-specific regulatory role.

Previous studies have shown that TE-associated smRNA activity often peaks shortly after hybridization or whole-genome duplication, followed by stabilization via DNA methylation or other epigenetic mechanisms (Jiao *et al.*, 2018; Martinez Palacios *et al.*, 2019). In *Arabidopsis*, Ha *et al.*, (2009) proposed that allopolyploids exploit smRNAs inherited from the parents to buffer genomic shock and promote rapid stabilization. The comparatively low divergence in 24 nt smRNA targeting observed here may reflect the advanced age of these natural allotetraploids—approximately 74k (*D. traunsteineri*) and 104k generations (*D. majalis*) (Hawranek, 2021). Nonetheless, for the few DT peaks detected, smRNA abundance was higher in the older *D. majalis* (Fig. 3), possibly indicating continued regulatory refinement associated with its broader ecological range or a more advanced stage of genome diploidization.

Intriguingly, most smRNA differences associated with TE regulation in the ecologically divergent allotetraploids involved 20–23 nt smRNAs, which likely include miRNAs and nat-siRNAs. These classes are frequently implicated in responses to biotic and abiotic stress (Ito *et al.*, 2011; Marí-Ordóñez *et al.*, 2013; Creasey *et al.*, 2014; Lee *et al.*, 2020). Elevated expression of miRNAs (typically 20–22 nt long) has also been observed in resynthesized allopolyploids relative to diploid progenitors across several systems (Ha *et al.*, 2009; Kenan-Eichler *et al.*, 2011; Ghani *et al.*, 2014), underscoring the potential role of shorter smRNAs in stress adaptation and dynamic gene regulation in polyploids.

Despite *Ty1-copia* being the most abundant LTR-TE in *Dactylorhiza* genomes (Eriksson *et al.*, 2022), we observed more 20–23 nt smRNAs mapping to, and possibly derived from, *Ty3-gypsy* elements (Supplementary Figs. S6–S7), particularly non-chromovirus subgroups. Notably, *Ty3-gypsy* elements displayed contrasting patterns between the diploid progenitors: chromovirus subgroups were more prevalent in *D. fuchsii*, while non-chromoviruses were

more common in *D. incarnata* (Eriksson *et al.*, 2022). Elevated targeting of *Ty3-gypsy* elements by shorter smRNAs in allopolyploids may reflect subgenome conflict or TE activation following hybridization. A similar phenomenon occurs in *Gossypium*, where TEs originally in the higher-TE A-subgenome transposed into the D-subgenome post-allopolyploidization (Chen *et al.*, 2020), suggesting that transposition between subgenomes may contribute to regulatory instability and diversification.

Genic smRNA targeting and possible regulatory implications

Polyploid formation introduces both meiotic/mitotic challenges and extensive genetic redundancy. Accurate differentiation between homologous and homeologous chromosomes is essential to ensure correct meiotic segregation. Preferential pairing of homologs has been observed in early stages of polyploid formation in *Arabidopsis suecica* (Pecinka *et al.*, 2011) and *Brassica napus* (Grandont *et al.*, 2014), although elevated homeolog mispairings has also been reported (Xiong *et al.*, 2021). Simultaneously, allopolyploids often undergo genetic diploidization, in which some redundant or misregulated gene copies may be silenced or lose function over time (Lee & Chen, 2001; Kashkush *et al.*, 2002; Koh *et al.*, 2010). Some duplicated copies can be retained and remain functional, for example if they confer adaptive benefits, become subfunctionalized, or acquire new functions via neofunctionalization (Adams *et al.*, 2003; Blanc & Wolfe, 2004; Duarte *et al.*, 2006; De Smet *et al.*, 2013).

In addition to silencing TEs, smRNAs—particularly 21–22 nt species—regulate gene expression via PTGS or TGS (Borges & Martienssen, 2015). We observed widespread genic smRNA targeting, especially in exons and 1 kb upstream of gene models (putative promoters), with 21 nt smRNAs contributing the most. While such targeting may mediate transcript degradation, alternative splicing, or cis-silencing (Guleria *et al.*, 2011; Moss, 2001), some associations may also have no functional consequence. Nevertheless, these patterns can reflect regulatory adjustments during early diploidization.

Substantial genic smRNA targeting was also detected in the diploids, especially *D. fuchsii*, likely reflecting stress responses in the common garden experiment conducted outdoors in Vienna, Austria—an area where only *D. incarnata* occurs naturally. Environmental mismatch may have triggered miRNA- and nat-siRNA-mediated regulation in *D. fuchsii*, consistent with previous reports of stress-induced smRNA upregulation (Zhang *et al.*, 2012; Mao *et al.*, 2021; Tiwari & Rajam, 2022).

Transgressive smRNA expression relative to the diploid progenitors may reflect how allopolyploids respond to genome duplication while also acquiring novel ecological features, potentially facilitating establishment in niches outside the parental range. We observed more

transgressive activity at peaks targeted by 20–23 nt smRNAs—compared to 24 nt—especially when the allotetraploids showed higher smRNA levels relative to both diploid parental species (Fig. 4e–f).

Genes transgressively targeted by smRNAs were enriched for functions related to mitosis, meiosis, gene regulation, stress responses, and transport (Fig. 5). While these functional categories were largely shared across the two allotetraploids, the specific genes and pathways involved differed, suggesting species-specific regulatory routes. In the younger *D. traunsteineri*, enrichment of smRNA targeting was observed for genes involved in cell cycle control, possibly reflecting ongoing meiotic stabilization. By contrast, the older *D. majalis* showed enrichments for genes involved in transcriptional and translational regulation, suggesting a later phase of polyploid evolution focused on dosage control and fine-tuning of gene expression.

Together, these observations support a dynamic evolution of smRNA-mediated regulation following allopolyploidization: an early phase dominated by stabilization of chromosome behavior during meiosis, together with refinement of gene redundancy and the emergence of functional divergence. Similar dynamics have been proposed in other systems including *Arabidopsis*, *Brassica*, and *Gossypium* (Ha *et al.*, 2009; Fu *et al.*, 2016; Chen *et al.*, 2020).

Conclusions and Future Directions

This study underscores the dual associations of smRNAs with genomic features in allopolyploids—both with TEs and genic regions—reflecting their potential roles in genome regulation. Despite their shared ancestry, *D. majalis* and *D. traunsteineri* exhibit clear lineage-specific patterns of smRNA activity, particularly in genic regions, suggesting divergent regulatory trajectories following allopolyploidization. Given their rapid action, diversity, and regulatory plasticity, smRNAs likely play central roles in both early and long-term genomic reorganization of allopolyploids. Our results support a model in which smRNAs mediate both immediate stabilization and longer-term functional divergence after genome merging. Future work should integrate smRNA expression profiles with homeolog-specific reference genomes and DNA methylation landscapes. These approaches would provide deeper insights into subgenome-specific regulation, the dynamics of smRNA–target interactions, and the evolutionary consequences of smRNA-mediated gene and TE regulation in natural allopolyploids.

ACKNOWLEDGEMENTS

This work was funded by the Austrian Science Fund (FWF) [grant DOI: 10.55776/Y661 awarded to OP and through the doctoral programme DOI: 10.55776/W1225, awarded to a faculty team including OP]. For open access purposes, the author has applied a CC BY public copyright license to any author accepted manuscript version arising from this submission. Faculty members, in particular Andreas Futschik, Robert Kofler and Christian Schlötterer, as well as the students and the scientific advisory board of the Vienna Graduate School of Population Genetics (www.popgen-vienna.at), Richard Bateman, and Ortrun Mittelsten Scheid are gratefully acknowledged for their valuable discussions and feedback throughout this work. We thank Juliane Baar, Daniela Paun and Maria Teresa Lorenzo for their support with laboratory procedures. Sequencing was performed at the Vienna BioCenter CoreFacilities (VBCF; https://www.viennabiocenter.org/). Computational resources were provided by the Life Science Compute Cluster (LiSC) of the University of Vienna. We also acknowledge Natural England and the Forestry Commission (UK), and regional county administrations in Austria, France, and Sweden for issuing the necessary collecting permits.

COMPETING INTERESTS

None declared.

AUTHOR CONTRIBUTIONS

Study conceived and designed by OP. Sampling and transplantation to the common garden conducted by OP, FB, MH, MCE and TMW. Read processing and mapping done by MT. Identification of putative smRNA peaks across the genome performed by MT and MCE. Data analyses of differentially targeted genes and TEs conducted by MCE. Analysis of overlapping differentially expressed and differentially targeted genes done by TMW. Interpretation of the results was undertaken by MCE. The manuscript was drafted by MCE with feedback from OP, and was revised and approved by all authors.

DATA AVAILABILITY

Newly reported raw Illumina smRNA-seq data are deposited on NCBI SRA (GenBank BioProject PRJNA317244, accessions SRR17191877–SRR17191897). Previously published Illumina RNA-seq and smRNA-seq are available from the same GenBank BioProject PRJNA317244. The pipeline, scripts and raw count file, together with the peak gff3 file are available at https://github.com/mc-er/dact-sRNAs.

REFERENCES

Adams KL, Cronn R, Percifield R, Wendel JF. 2003. Genes duplicated by polyploidy show unequal contributions to the transcriptome and organ-specific reciprocal silencing. *Proceedings of the National Academy of Sciences USA* **100**: 4649–4654. doi:10.1073/pnas.0630618100

Alexa A, Rahnenfuhrer J. 2021. topGO: Enrichment Analysis for Gene Ontology. R package version 2.44.0. [WWW document] URL https://bioconductor.org/packages/release/bioc/html/topGO.html. [accessed 15 May 2025].

Ariel FD, Manavella PA. 2021. When junk DNA turns functional: transposon-derived non-coding RNAs in plants. *Journal of Experimental Botany* **72**: 4132–4143. doi:10.1093/jxb/erab082

Balao F, Trucchi E, Wolfe TM, Hao BH, Lorenzo MT, Baar J, Paun O. 2017. Adaptive sequence evolution is driven by biotic stress in a pair of orchid species (*Dactylorhiza*) with distinct ecological optima. *Molecular Ecology* **26**: 3649–3662. doi:10.1111/mec.14123

Blanc G, Wolfe KH. 2004. Functional divergence of duplicated genes formed by polyploidy during *Arabidopsis* evolution. *Plant Cell* **16**: 1679–1691. doi:10.1105/tpc.021410

Borges F, Martienssen RA. 2015. The expanding world of small RNAs in plants. *Nature Reviews Molecular Cell Biology* **16**: 727–737. doi:10.1038/nrm4085

Brandrud MK, Baar J, Lorenzo MT, Athanasiadis A, Bateman RM, Chase MW, Hedrén M, Paun O. 2020. Phylogenomic relationships of diploids and the origins of allotetraploids in *Dactylorhiza*. *Systematic Biology* **69**: 91–109. doi:10.1093/sysbio/syz035

Carthew RW, Sontheimer EJ. 2009. Origins and mechanisms of miRNAs and siRNAs. *Cell* **136**: 642–665. doi:10.1016/j.cell.2009.01.035

Cavé-Radet A, Giraud D, Lima O, El Amrani A, Aïnouche ML, Salmon A. 2020. Evolution of small RNA expression following hybridization and allopolyploidization: insights from *Spartina* species (Poaceae, Chloridoideae). *Plant Molecular Biology* **102**: 55–72. doi:10.1007/s11103-019-00931-w

Cavé-Radet A, Salmon A, Tran Van Canh L, Ainouche M. 2023. Recent allopolyploidy alters *Spartina* microRNA expression in response to xenobiotic-induced stress. *Plant Molecular Biology* **111**: 309–328. doi:10.1007/s11103-022-01328-y

Chen X, Rechavi O. 2021. Plant and animal small RNA communications between cells and organisms. *Nature Reviews Molecular Cell Biology* **22**: 750–766. doi:10.1038/s41580-021-00425-y

Chen ZJ, Sreedasyam A, Ando A, Song Q, De Santiago LM, Hulse-Kemp AM, Schmutz J. 2020. Genomic diversifications of five *Gossypium* allopolyploid species and their impact on cotton improvement. *Nature Genetics* **52**: 525–533. doi:10.1038/s41588-020-0614-5

Cheng F, Sun C, Wu J, Schnable J, Woodhouse MR, Liang J, Wang X. 2016. Epigenetic regulation of subgenome dominance following whole genome triplication in *Brassica rapa. New Phytologist* **211**: 288–299. doi:10.1111/nph.13884

Creasey KM, Zhai J, Borges F, Van Ex F, Regulski M, Meyers BC, Martienssen RA. 2014. miRNAs trigger widespread epigenetically activated siRNAs from transposons in *Arabidopsis. Nature* 508: 411–415. doi:10.1038/nature13069

Danecek P, Auton A, Abecasis G, Albers CA, Banks E, Depristo MA, Handsaker RE, Lunter G, Marth GT, Sherry ST et al. 2011. The variant call format and VCFtools. *Bioinformatics* **27**: 2156–2158. doi:10.1093/bioinformatics/btr330.

De Smet R, Adams KL, Vandepoele K, Van Montagu MCE, Maere S, Van de Peer Y. 2013. Convergent gene loss following gene and genome duplications creates single-copy families in flowering plants. *Proceedings of the National Academy of Sciences, USA* **110**: 2898–2903. doi:10.1073/pnas.1300127110.

Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. 2013. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**: 15–21. doi:10.1093/bioinformatics/bts635.

Duarte JM, Cui L, Wall PK, Zhang Q, Zhang X, Leebens-Mack J, DePamphilis CW. 2006. Expression pattern shifts following duplication indicative of subfunctionalization and neofunctionalization in regulatory genes of *Arabidopsis. Molecular Biology and Evolution* **23**: 469–478. doi:10.1093/molbev/msj051

Edger PP, Smith R, McKain MR, Cooley AM, Vallejo-Marin M, Yuan Y, Puzey JR. 2017. Subgenome dominance in an interspecific hybrid, synthetic allopolyploid, and a 140-year-old naturally established neo-allopolyploid monkeyflower. *Plant Cell* **29**: 2150–2167. doi:10.1105/tpc.17.00010

Emelianova K, Hawranek A-S, Eriksson M, Wolfe TM, Paun O. 2025. Ecological divergence of sibling allopolyploid marsh orchids is associated with species specific

plasticity and distinct fungal communities. *Plant Journal* **121**: e70001. doi: 10.1111/tpj.70001

Eriksson MC, Mandáková T, McCann J, Temsch EM, Chase MW, Hedrén M, Weiss-Schneeweiss H, Paun O. 2022. Repeat dynamics across timescales: a perspective from sibling allotetraploid marsh orchids (*Dactylorhiza majalis* s.l.). *Molecular Biology and Evolution* **39**: msac167. doi:10.1093/molbev/msac167

Fedoroff NV. 2012. Transposable elements, epigenetics, and genome evolution. *Science* **338**: 758–77. doi:10.1126/science.338.6108.758

Fu Y, Xiao M, Yu H, Mason AS, Yin J, Li J, Fu D. 2016. Small RNA changes in synthetic *Brassica napus. Planta* **244**: 607–622. doi:10.1007/s00425-016-2529-z

Fultz D, Choudury SG, Slotkin RK. 2015. Silencing of active transposable elements in plants. *Current Opinion in Plant Biology* **27**: 67–76. doi:10.1016/j.pbi.2015.05.027

Gaebelein R, Schiessl SV, Samans B, Batley J, Mason AS. 2019. Inherited allelic variants and novel karyotype changes influence fertility and genome stability in *Brassica* allohexaploids. *New Phytologist* **223**: 965-978. doi:10.1111/nph.15804.

Ghani MA, Li J, Rao L, Raza MA, Cao L, Yu N, Chen L. 2014. The role of small RNAs in wide hybridisation and allopolyploidisation between *Brassica rapa* and *Brassica nigra*. *BMC Plant Biology* **14**: 272. doi:10.1186/s12870-014-0272-9

Grandont L, Cuñado N, Coriton O, Huteau V, Eber F, Chèvre AM, Jenczewski E. 2014. Homoeologous chromosome sorting and progression of meiotic recombination in *Brassica napus*: ploidy does matter! *Plant Cell* **26**: 1448–1463. doi:10.1105/tpc.114.122788

Guleria P, Mahajan M, Bhardwaj J, Yadav S. 2011. Plant small RNAs: biogenesis, mode of action and their roles in abiotic stresses. *Genomics, Proteomics & Bioinformatics* **9**: 183–199. doi:10.1016/S1672-0229(11)60022-9

Ha M, Lu J, Tian L, Ramachandran V, Kasschau KD, Chapman EJ, Chen ZJ. 2009. Small RNAs serve as a genetic buffer against genomic shock in *Arabidopsis* interspecific hybrids and allopolyploids. *Proceedings of the National Academy of Sciences USA* **106**: 17835–17840. doi:10.1073/pnas.0907003106

Hassan AH, Mokhtar MM, El Allali A. 2024. Transposable elements: multifunctional players in the plant genome. *Frontiers in Plant Science* **14**: 1330127. doi:10.3389/fpls.2023.1330127

Hawranek AS. 2021. *The impact of recurrent origins and gene flow on the genetic structure of allopolyploid marsh orchids (*Dactylorhiza*, Orchidaceae).* MSc thesis, University of Vienna, Vienna, Austria. doi:10.25365/thesis.70360

Ito H, Gaubert H, Bucher E, Mirouze M, Vaillant I, Paszkowski J. 2011. An siRNA pathway prevents transgenerational retrotransposition in plants subjected to stress. Nature 472: 115–120.

Jiang X, Song Q, Ye W, Chen ZJ. 2021. Concerted genomic and epigenomic changes accompany stabilization of *Arabidopsis* allopolyploids. *Nature Ecology & Evolution* **5**: 1382–1393. doi:10.1038/s41559-021-01523-y.

Jiao W, Yuan J, Jiang S, Liu Y, Wang L, Liu M, Wang J, Zhang L, Liu Y, Xia C et al. 2018. Asymmetrical changes of gene expression, small RNAs and chromatin in two resynthesized wheat allotetraploids. *Plant Journal* **93**: 828–842. doi:10.1111/tpj.13805.

Kashkush K, Feldman M, Levy AA. 2002. Gene loss, silencing and activation in a newly synthesized wheat allotetraploid. *Genetics* **160**: 1651–1659. doi:10.1093/genetics/160.4.1651.

Kaur R, Choudhury A, Chauhan S, Ghosh A, Tiwari R, Rajam MV. 2021. RNA interference and crop protection against biotic stresses. *Physiology and Molecular Biology of Plants* **27**: 2357–2373. doi:10.1007/s12298-021-01058-2.

Kenan-Eichler M, Leshkowitz D, Tal L, Noor E, Melamed-Bessudo C, Feldman M, Levy AA. 2011. Wheat hybridization and polyploidization results in deregulation of small RNAs. *Genetics* **188**: 263–272. doi:10.1534/genetics.111.128348.

Koh J, Soltis PS, Soltis DE. 2010. Homeolog loss and expression changes in natural populations of the recently and repeatedly formed allotetraploid *Tragopogon mirus* (Asteraceae). *BMC Genomics* **11**: 97. doi:10.1186/1471-2164-11-97.

Lee HS, Chen ZJ. 2001. Protein-coding genes are epigenetically regulated in *Arabidopsis* polyploids. *Proceedings of the National Academy of Sciences, USA* **98**: 6753–6758. doi:10.1073/pnas.121064898.

Lee SC, Ernst E, Berube B, Borges F, Parent J-S, Ledon P, Zuber H, Sperschneider J, Boccara M, Lalonde S et al. 2020. *Arabidopsis* retrotransposon virus-like particles and their regulation by epigenetically activated small RNA. *Genome Research* **30**: 576–588. doi:10.1101/gr.259044.119.

Lewsey MG, Hardcastle TJ, Melnyk CW, Molnar A, Valli A, Ulrich MA, Nery JR, Baulcombe DC, Ecker JR. 2016. Mobile small RNAs regulate genome-wide DNA methylation. *Proceedings of the National Academy of Sciences, USA* **113**: E801-E810.

Liao Y, Smyth GK, Shi W. 2014. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**: 923–930. doi:10.1093/bioinformatics/btt656.

Liu C, Wang YG. 2023. Does one subgenome become dominant in the formation and evolution of a polyploid? *Annals of Botany* **131**: 11–16. doi:10.1093/aob/mcac024.

Mao Y, Xu J, Wang Q, Li G, Tang X, Liu T, Zhang H, Wang W, Li T, Chen X et al. 2021. A natural antisense transcript acts as a negative regulator for the maize drought stress response gene ZmNAC48. *Journal of Experimental Botany* **72**: 2790–2806. doi:10.1093/jxb/erab023.

Marí-Ordóñez A, Marchais A, Etcheverry M, Martin A, Colot V, Voinnet O. 2013. Reconstructing de novo silencing of an active plant retrotransposon. *Nature Genetics* **45**: 1029–1033. doi:10.1038/ng.2703.

Martinez Palacios P, Jacquemot M-P, Tapie M, Rousselet A, Diop M, Remoué C, Demangeat G, Berges H, Pichon M, Marande W et al. 2019. Assessing the response of small RNA populations to allopolyploidy using resynthesized *Brassica napus* allotetraploids. *Molecular Biology and Evolution* **36**: 709–726. doi:10.1093/molbev/msz007.

Matzke MA, Mosher RA. 2014. RNA-directed DNA methylation: an epigenetic pathway of increasing complexity. *Nature Reviews Genetics* **15**: 394-408. doi:10.1038/nrg3683.

McCarthy DJ, Chen Y, Smyth GK. 2012. Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. *Nucleic Acids Research* **40**: 4288–4297. doi:10.1093/nar/gks042.

McClintock B. 1984. The significance of responses of the genome to challenge. *Science* **226**: 792–801. doi:10.1126/science.15739260.

Moss E. 2001. RNA interference: it's a small RNA world. *Current Biology* **11**: R772–R775. doi:10.1016/S0960-9822(01)00411-1.

Pecinka A, Fang W, Rehmsmeier M, Levy AA, Mittelsten Scheid O. 2011. Polyploidization increases meiotic recombination frequency in *Arabidopsis. BMC Biology* **9**: 24. doi:10.1186/1741-7007-10-33.

Poretti M, Praz CR, Meile L, Kälin C, Schaefer LK, Schläfli M, Widrig V, Sanchez-Vallet A, Wicker T, Parlange F et al. 2020. Domestication of high-copy transposons underlies the wheat small RNA response to an obligate pathogen. *Molecular Biology and Evolution* **37**: 839–848. doi:10.1093/molbev/msz272.

Quinlan AR, Hall IM. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**: 841–842. doi:10.1093/bioinformatics/btq033.

Ramírez F, Ryan DP, Grüning B, Bhardwaj V, Kilpert F, Richter AS, Heyne S, Dündar F, Manke T. 2016. deepTools2: a next generation web server for deep-sequencing data analysis. Nucleic Acids Research 44: W160–W165. doi:10.1093/nar/gkw257.

Robinson MD, McCarthy DJ, Smyth GK. 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26: 139–140. doi:10.1093/bioinformatics/btp616.

Shen Y, Sun S, Hua S, Shen E, Ye CY, Cai D, Timko MP, Zhu QH, Fan L, Xu L et al. 2017. Analysis of transcriptional and epigenetic changes in hybrid vigor of allopolyploid Brassica napus uncovers key roles for small RNAs. Plant Journal 91: 874–893. doi:10.1111/tpj.13605.

Tiwari R, Rajam MV. 2022. RNA- and miRNA-interference to enhance abiotic stress tolerance in plants. Journal of Plant Biochemistry and Biotechnology 31: 1–10. doi:10.1007/s13562-022-00770-9.

Tossolini I, Mencia R, Arce AL, Manavella PA. 2025. The genome awakens: transposon-mediated gene regulation. Trends in Plant Science. doi:10.1016/j.tplants.2025.02.005.

Wang HLV, Chekanova JA. 2016. Small RNAs: essential regulators of gene expression and defenses against environmental stresses in plants. Wiley Interdisciplinary Reviews: RNA 7: 356–381. doi:10.1002/wrna.1340.

Wang M, Zhao Y, Zhang B. 2015. Efficient test and visualization of multi-set intersections. *Scientific Reports* **5**: 16923. doi:10.1038/srep16923.

Wolfe TM, Balao F, Trucchi E, Bachmann G, Gu W, Baar J, Suda J, Paun O. 2023. Recurrent allopolyploidizations diversify ecophysiological traits in marsh orchids (*Dactylorhiza majalis* s.l.). *Molecular Ecology* **32**: 4777–4790. doi:10.1111/mec.17070.

Xiao Y, Wang J. 2025. Understanding the regulation activities of transposons in driving the variation and evolution of polyploid plant genome. *Plants* **14**: 1160. doi:10.3390/plants14081160.

Xie Y, Ying S, Li Z, Zhang Y, Zhu J, Zhang J, Li X, Liu H, Wang Y, Liu Y et al. 2023. Transposable element-initiated enhancer-like elements generate the subgenome-biased spike specificity of polyploid wheat. *Nature Communications* **14**: 7465. doi:10.1038/s41467-023-42771-9.

Xiong Z, Gaeta RT, Edger PP, Cao Y, Zhao K, Zhang S, Wan T, Liu Y, Coate JE, Pires JC. 2021. Chromosome inheritance and meiotic stability in allopolyploid *Brassica napus. G3: Genes, Genomes, Genetics* **11**: jkaa011. doi:10.1093/g3journal/jkaa011.

Liu L, Chen X. 2018. Intercellular and systemic trafficking of RNAs in plants. *Nature Plants* **4**: 869–878. doi:10.1038/s41477-018-0288-5.

Zhan J, Meyers BC. 2024. Plant small RNAs: their biogenesis, regulatory roles, and functions. *Annual Review of Plant Biology* **75**: 21–51. doi:10.1146/annurev-arplant-070122-035226.

Zhang X, Xia J, Lii YE, Barrera-Figueroa BE, Zhou X, Gao S, Lu L, Niu D, Chen Z, Liang W et al. 2012. Genome-wide analysis of plant nat-siRNAs reveals insights into their distribution, biogenesis and function. *Genome Biology* **13**: R20. doi:10.1186/gb-2012-13-3-r20.

FIGURE LEGENDS

Figure 1. Study system comprising two diploids (*Dactylorhiza fuchsii* and *D. incarnata*) and their sibling allotetraploids (*D majalis* and *D. traunsteineri*). **(a)** Overview of the four species. Coloured circles and their relative positions illustrate population densities. Cell icons indicate difference in genome size between the diploids, while gender symbols show the direction of hybridization leading to the formation of the allopolyploids. Additional symbols denote broad ecological preferences. **(b)** Multidimensional scaling (MDS) plot based on small RNA count data. Colours correspond to the species frames shown in (a); triangles indicate diploid accessions, and circles, allotetraploid accessions.

Figure 2. Small RNA read length distributions across species. **(a)** Mean counts per million (CPM) for each read length; error bars represent the standard error of the mean. Horizontal lines indicate significant differences between species, with asterisks denoting significance levels: * P < 0.05, ** P < 0.01, *** P < 0.001. **(b, c)** Proportions of reads by species and read length mapped to genomic features: (b) exons and (c) transposable elements. fuc*, D. fuchsii*; inc, *D. incarnata*; maj, *D. majalis*; tra, *D. traunsteineri*.

Figure 3. MA plots showing differences in small RNA (smRNA) "targeting" between *Dactylorhiza majalis* and *D. traunsteineri*. The central dashed line marks logFC = 0, while dotted lines indicate logFC = -1.5 and 1.5. Negative logFC values represent higher smRNA targeting in *D. majalis* relative to *D. traunsteineri*, whereas positive logFC values indicate higher targeting in *D. traunsteineri*. **(a–c)** Peaks differently targeted in genic regions by 20–24 nt smRNAs: **(a)** promoter/1,000 bp upstream gene models, **(b)** exons and **(c)** introns. **(d)** Transposable elements (TEs) targeted by 20–23 nt smRNAs. **(e)** TEs targeted by 24 nt smRNAs. **(f)** MA plot of RNA-seq data from Wolfe *et al.* (2023), showing genes differentially expressed (DE) and/or DT by smRNAs as indicated in the legend.

Figure 4. Patterns of genomic interactions in smRNA "targeting". **(a–b)** Comparison of logFC values for *Dactylorhiza majalis* relative to either diploid parent (*D. fuchsii* on the Y-axis, *D. incarnata* on the X-axis) for all peaks (20–24 nt smRNAs) in annotated genes. **(a)** Colored symbols indicate additive and transgressive patterns; **(b)** shows dominance toward either diploid, as indicated in the legend. **(c-f)** Number of differentially targeted (DT) peaks in each type of genomic interaction, with peaks in *D. traunsteineri* shown in blue and *D. majalis* in red. Striped bars denote DT peaks shared between the allotetraploids. The annotation columns along the Y-axis represent the interaction type (colored bars, according to legend), while dots connected by lines show the direction of the interaction. Letters above each column indicate I - *D. incarnata*, P - polyploid, and F - *D. fuchsii.* **(c)** 20–24 nt smRNAs

associated with genic regions, **(d)** 20–24 nt smRNA peaks outside annotated genic and TE regions, **(e)** 20–23 nt smRNAs in TEs, and **(f)** 24 nt smRNAs in TEs.

Figure 5. Summary of representative functional enrichments for **(a)** the allopolyploid comparison, **(b)** transgressive smRNA "targeting" of genic regions in *Dactylorhiza traunsteineri*, and **(c)** transgressive targeting in *D. majalis*. Bar length represents the adjusted significance of each enriched GO term. Full results are provided in Supplementary Tables S6, S8 and S9. Bar color indicates the z-score, computed from the logFC values of the corresponding targeting, as indicated in the legends. Arrows on bars in panels (b) and (c) mark GO terms shared between the two comparisons. Symbols next to the GO terms denote the broad functional categories (see legend), with numbers indicating the count of enriched genes within each category. Letters to the right of each plot specify the genetic region targeted by smRNAs; p - 'promotor', e - exon, i - intron. Abbreviations used in GO terms: + positive, - negative, DS double strand, pol polymerase, ts transcription.

SUPPLEMENTARY TABLES LEGENDS

Table S1. Sample information. Collector abbreviations are FB, Francisco Balao; MC, Mark Chase; MF, Michael Fay; MH, Mikael Hedren; and OP, Ovidiu Paun. Note: the smRNA-seq data has been collected within a common garden setup; here only the population of origin of the plants is indicated.

Table S2. Criteria for categories of genomic interactions. Abbreviations in the column names are as follows: 4x, allotetraploid; fuc - *D. fuchsii*; inc - *D. incarnata*. Abbreviations in the column Direction are as follows: F, *D. fuchsii*; I, *D. incarnata*; P, polyploid (i.e. *D. majalis* or *D. traunsteineri*).

Table S3. Mapping statistics for each sample. Accession IDs ending with ‘r' indicate technical replicates.

Table S4. Number of smRNA targeting peaks in differential targeting (DT) analysis. Subscripts in the column names are as follow: A, all analyzed DT peaks; all, significant DT peaks across the whole genome; TE, significant DT peaks in annotated transposable elements; G, significant DT peaks in annotated genes; Gp, significant DT peaks 1k upstreams of annotated genes; Gi, significant DT peaks in introns of annotated genes; Ge, significant DT peaks in exons of annotated genes; Multiple regions, DT peaks spanning the border between two different genic features (exon, intron or promotor). Significant = adjusted p-value <=0.05. Abbreviations in the table are as follows: 4X, polyploid; 2x, diploid; maj, *D.majalis*; tra, *D.traunsteineri*; fuc - *D. fuchsii*; inc - *D. incarnata*.

Table S5. List of genes both differentially expressed between the allotetraploids in Wolfe *et al.*, (2023) and differentially targeted by smRNAs (this study). smRNA, small RNA values from this study; RNA, RNA-seq values from Wolfe et al. (2023); GI, Genomic Interaction; maj, *D. majalis*; tra, *D. traunsteineri*.

Table S6. Enrichment results for genes differentially targeted between the polyploids (D. majalis and D. traunsteineri). N_peaks, number of peaks in GO term; N_genes, number of genes in GO term; adj_pval, adjusted p-value from topGO; Z-score, normalized value based on number of genes and the average log fold change value.

Table S7. Number of peaks for each defined genomic interaction (additive, transgressive or dominant towards either diploid). Regions mean: all - genome wide, genes - around annotated genes, TEs - around annotated TEs, other regions - outside annotated TEs and genes. Abbreviations in the column names are as follows: maj - number of peaks in *D.*

majalis, tra - number of peaks in *D. traunsteineri*, sig. Significance level for hypergeometric test done for shared. * < 0.05, ** < 0.01, *** < 0.001.

Table S8. Enrichment results for genes transgressivly differentially targeted between the polyploid (D. majalis) and either of the diploids (D. fuchsii and D. incarnata). N_peaks, number of peaks in GO term; N_genes, number of genes in GO term; adj_pval, adjusted p-value from topGO; Z-score, normalized value based on number of genes and the average log fold change value.

Table S9. Enrichment results for genes transgressivly differentially targeted between the polyploid (D. traunsteineri) and either of the diploids (D. fuchsii and D. incarnata). N_peaks, number of peaks in GO term; N_genes, number of genes in GO term; adj_pval, adjusted p-value from topGO; Z-score, normalized value based on number of genes and the average log fold change value.

SUPPLEMENTARY FIGURES LEGENDS

Figure S1. Distribution of normalised reads counts for different smRNA lengths over genomic regions (promoter/1000 bp upstream, exons, introns and intergenic) in the four tested species. fuc = *D. fuchsii*, inc = *D. incarnata*, maj = *D. majalis*, tra = *D. traunsteineri*.

Figure S2. Number of differentially targeted peaks in genes and TEs for each comparison between a polyploid (4x) and a diploid (2x) for all three data sets, 20-24 nt 20-23 nt and 24 nt smRNAs. The top panel contains the number of peaks where the polyploid is over regulated compared to the diploid, whereas the bottom panel contains the number of peaks where the diploid is over targeted. The coloured boxed below the figure represents the comparison shown in the corresponding column.

Figure S3. Number of DT peaks in all datasets 20−24 nt, 20−23 nt and 24 nt smRNAs for each genomic interaction (additive, transgressing or dominant to either diploid). The top row contains DT peaks in genes, the middle row shows DT peaks in TEs and the bottom row shows peaks found outside annotated genic and TE regions, e.g. intergenic. Colours represent, red - *D. majalis*, blue - *D. traunsteineri* and striped pattern - number of DT peaks showing the same genomic interaction in both *D. majalis* and *D. traunsteineri*. The two annotation columns along the Y−axis represent the type (coloured bars according to legend) of genomic interaction Fig. 4, while dots connected with lines show the direction. Letters above each column of dots represent **F** - *D. fuchsii*, **P** - polyploid and **I** - *D. incarnata*.

Figure S4. Patterns of genomic interactions comparing logFC values for *D. majalis* (left) and *D. traunsteineri* (right) towards either diploid, *D. fuchsii* on the Y-axis and *D. incarnata* on the X-axis for all peaks (i.e., 20−24 nt smRNAs) found in annotated genes. (a-b) shows with coloured symbols transgressive and additive patterns; whereas (c-d) is dominant to either diploid, according to the legend.

Figure S5 Proportions of peaks within each genomic interaction (additive, transgressing or dominant to either diploid) for each genomic region (1000 bp upstream/promoter, exon, intron). Colours are according to legend. Letters in legend represent **F** - *D. fuchsii*, **p** - polyploid and **I** - *D. incarnata* while symbols represent level of smRNA between the diploids and polyploid.

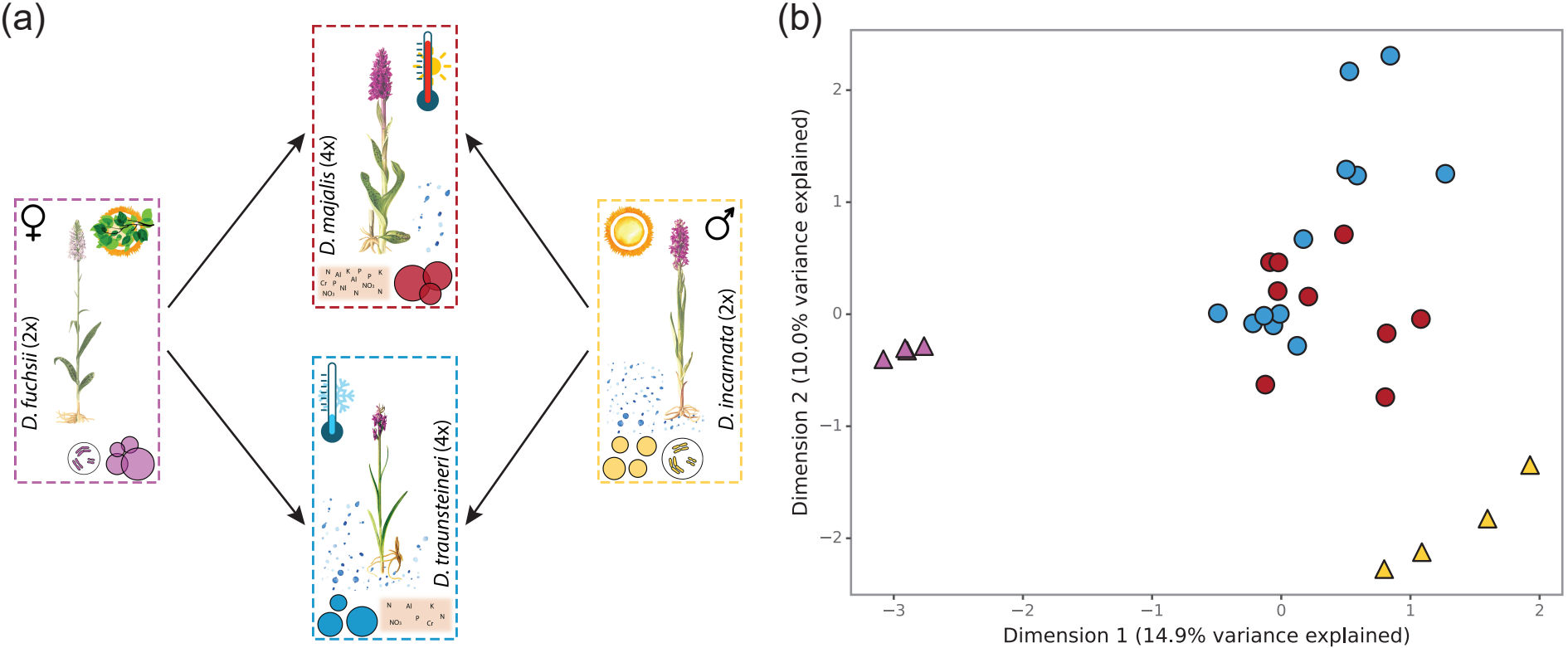
Figure S6 MA plots showing differences between *D. majalis* and *D. traunsteineri* for TEs types grouped together, the central dashed line marks 0 while the dotted lines mark logFC -1.5 and 1.5. Negative logFC values represent over-targeting in *D. majalis* as compared to *D. traunsteineri*, whereas positive logFC values indicate over-targeting in *D. traunsteineri*. (a) and (b) show 20−23 nt smRNAs while (c) and (d) show 24 nt smRNAs. (a) and (c) show DT peaks specific for each smRNA group, i.e. 20−23 nt or 24 nt, while (b) and (d) show DT peaks for both 20−23 nt and 24 nt smRNAs.

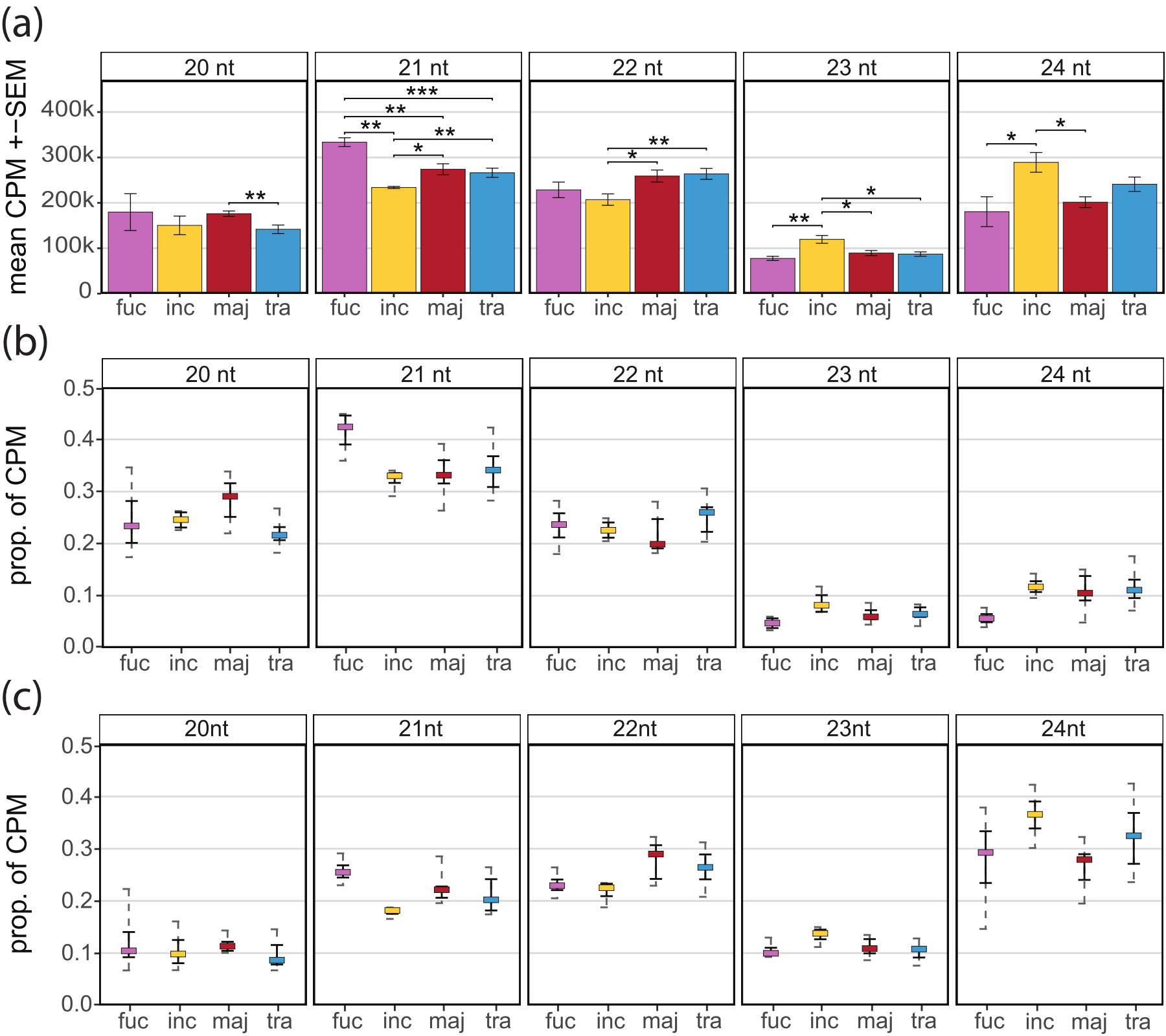
Figure S7 MA plots showing differences between *D. majalis* and *D. traunsteineri* for TEs, the central dashed line marks 0 while the dotted lines mark logFC -1.5 and 1.5. Negative logFC values represent over-targeting in *D. majalis* as compared to *D. traunsteineri*, whereas positive logFC values indicate over-targeting in *D. traunsteineri*. Colours correspond to different types of TEs, indicated by column headers. Rows correspond to smRNAs, where the 1st and 2nd row show 20−23 nt smRNAs and 3rd and 4th show 24 nt smRNAs. The 2nd and 3rd row show peaks that are DT for both 20−23 nt and 24 nt smRNAs.

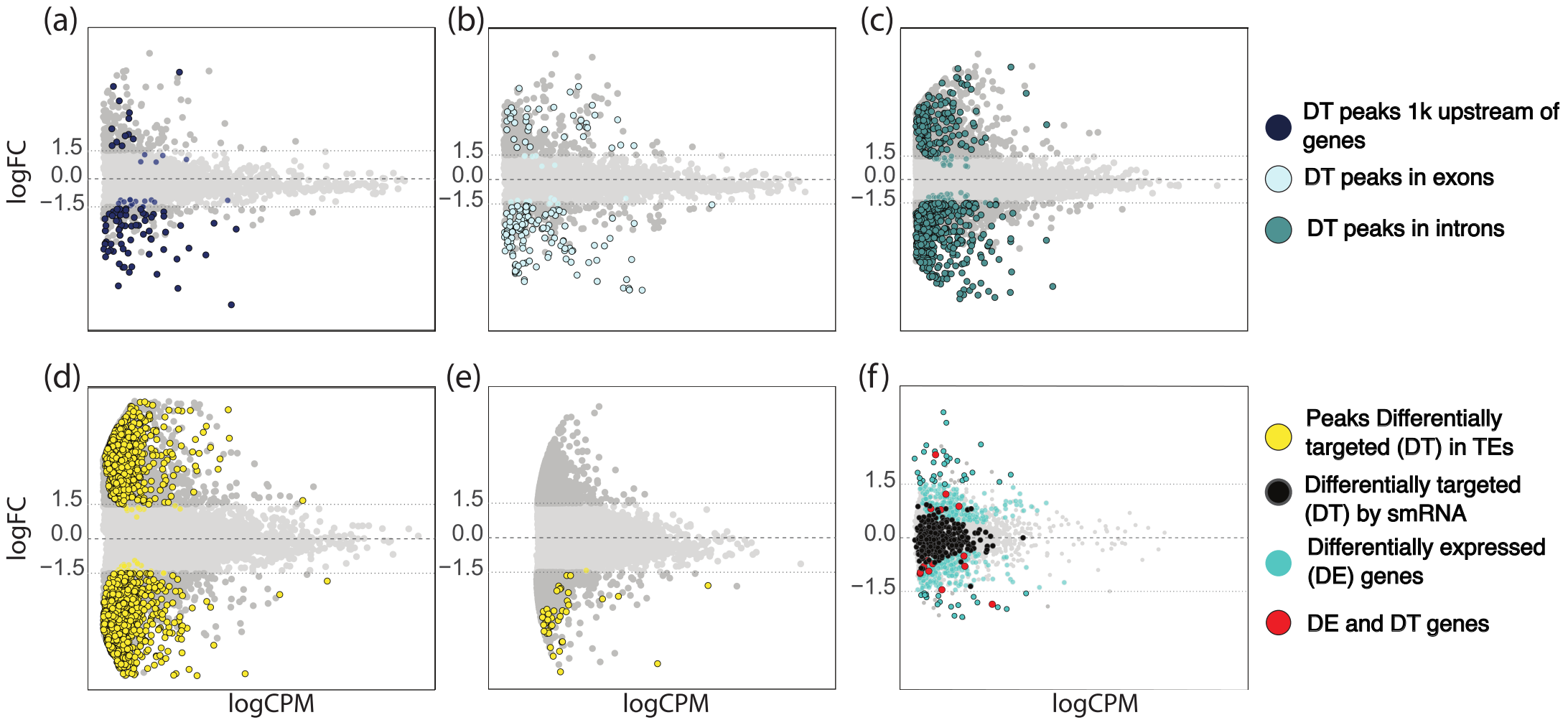
Figure S8 Count of DT peaks between *D. majalis* and *D. traunsteineri*, the coloured points of the plots in the 1st row (i.e. 20−23 nt smRNAs) in Fig. S7. Here Ty3-gypsy has been divided into the next annotation level.

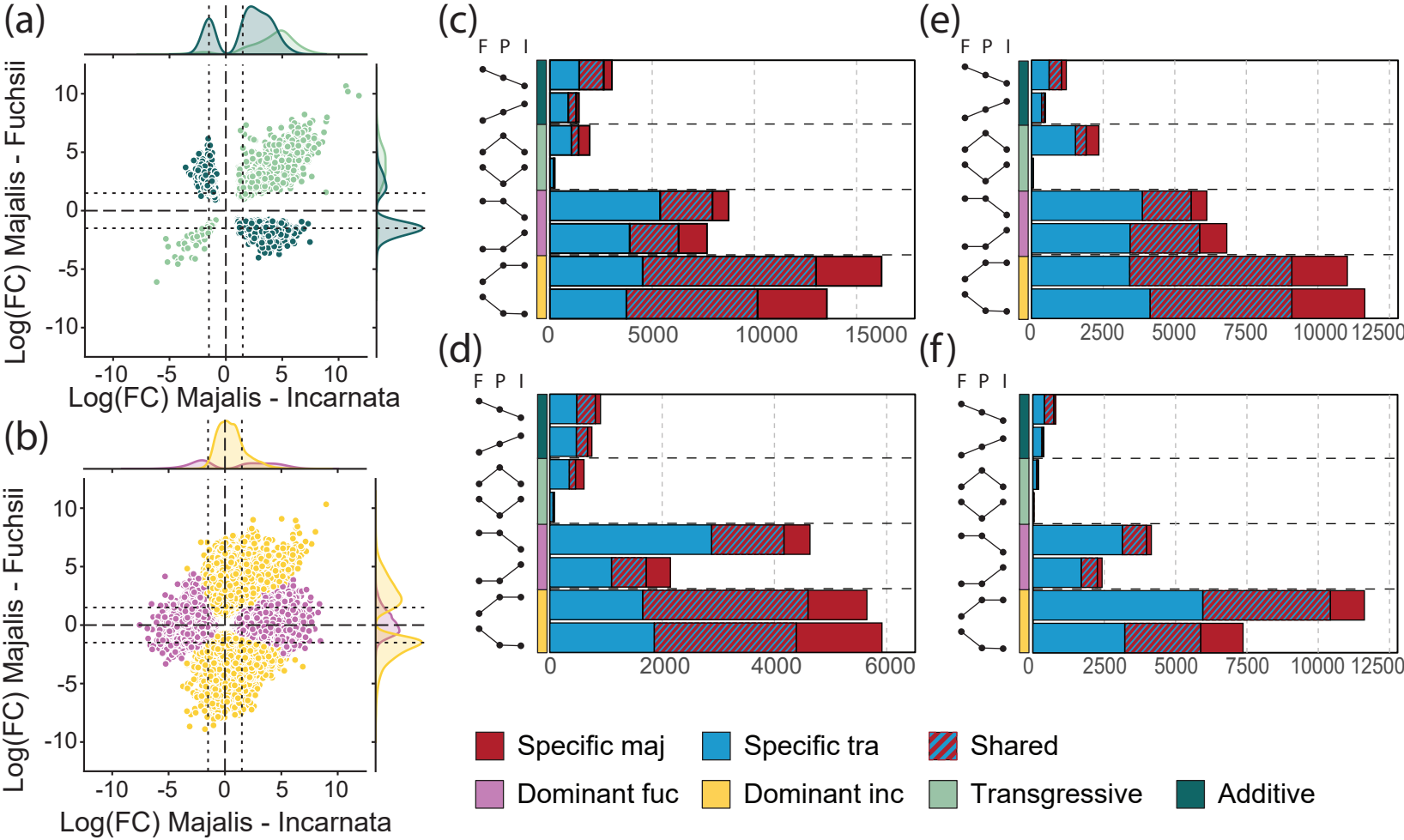
Figure S9 Patterns of genomic interactions comparing logFC values for *D. majalis* (left) and *D. traunsteineri* (right) towards either diploid, *D. fuchsii* on the Y-axis and *D. incarnata* on the X-axis for 20-23 nt smRNAs peaks found in annotated TEs. Top row shows with coloured symbols transgressive and additive patterns; whereas bottom row shows dominant to either diploid.

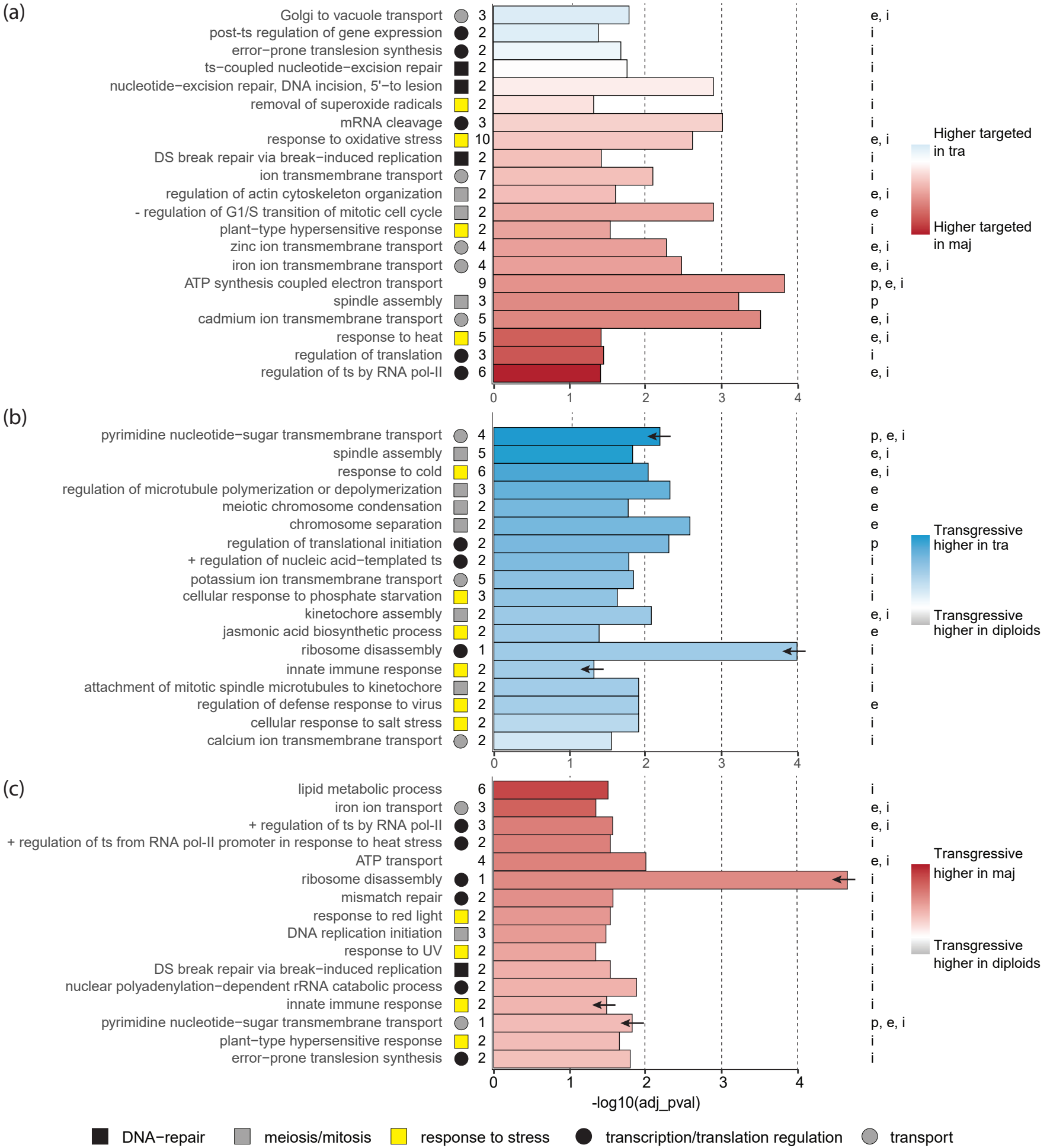
Figure S10 Patterns of genomic interactions comparing logFC values for *D. majalis* (left) and *D. traunsteineri* (right) towards either diploid, *D. fuchsii* on the Y-axis and *D. incarnata* on the X-axis for 24 nt smRNAs peaks found in annotated TEs. Top row shows with coloured symbols transgressive and additive patterns; whereas bottom row shows dominant to either diploid.


(a)
♀
D. fuchsii (2x)
D. majalis (4x)
D. traunsteineri (4x)
D. incarnata (2x)
♂
(b)
Dimension 2 (10.0% variance explained)
Dimension 1 (14.9% variance explained)
2
1
0
−1
−2
−3
−2
−1
0
1
2


(a)
mean CPM +−SEM
20 nt
21 nt
22 nt
23 nt
24 nt
400k
300k
200k
100k
0
fuc
inc
maj
tra
(b)
prop. of CPM
0.5
0.4
0.3
0.2
0.1
0.0
(c)
20nt
21nt
22nt
23nt
24nt


(a)
(b)
(c)
(d)
(e)
(f)
logFC
logCPM
1.5
0.0
−1.5
DT peaks 1k upstream of genes
DT peaks in exons
DT peaks in introns
Peaks Differentially targeted (DT) in TEs
Differentially targeted (DT) by smRNA
Differentially expressed (DE) genes
DE and DT genes


(a)
Log(FC) Majalis - Fuchsii
Log(FC) Majalis - Incarnata
(b)
Log(FC) Majalis - Fuchsii
Log(FC) Majalis - Incarnata
(c)
(d)
(e)
(f)
F P I
Specific maj
Specific tra
Shared
Dominant fuc
Dominant inc
Transgressive
Additive


(a)
Golgi to vacuole transport 3 e, i
post-ts regulation of gene expression 2 i
error−prone translesion synthesis 2 i
ts−coupled nucleotide−excision repair 2 i
nucleotide−excision repair, DNA incision, 5'−to lesion 2 i
removal of superoxide radicals 2 i
mRNA cleavage 3 i
response to oxidative stress 10 e, i
DS break repair via break−induced replication 2 i
ion transmembrane transport 7 i
regulation of actin cytoskeleton organization 2 e, i
- regulation of G1/S transition of mitotic cell cycle 2 e
plant−type hypersensitive response 2 i
zinc ion transmembrane transport 4 e, i
iron ion transmembrane transport 4 e, i
ATP synthesis coupled electron transport 9 p, e, i
spindle assembly 3 p
cadmium ion transmembrane transport 5 e, i
response to heat 5 e, i
regulation of translation 3 i
regulation of ts by RNA pol-II 6 e, i
0 1 2 3 4
Higher targeted in tra
Higher targeted in maj
(b)
pyrimidine nucleotide−sugar transmembrane transport 4 p, e, i
spindle assembly 5 e, i
response to cold 6 e, i
regulation of microtubule polymerization or depolymerization 3 e
meiotic chromosome condensation 2 e
chromosome separation 2 e
regulation of translational initiation 2 p
+ regulation of nucleic acid−templated ts 2 i
potassium ion transmembrane transport 5 i
cellular response to phosphate starvation 3 i
kinetochore assembly 2 e, i
jasmonic acid biosynthetic process 2 e
ribosome disassembly 1 i
innate immune response 2 i
attachment of mitotic spindle microtubules to kinetochore 2 i
regulation of defense response to virus 2 e
cellular response to salt stress 2 i
calcium ion transmembrane transport 2 i
0 1 2 3 4
Transgressive higher in tra
Transgressive higher in diploids
(c)
lipid metabolic process 6 i
iron ion transport 3 e, i
+ regulation of ts by RNA pol-II 3 e, i
+ regulation of ts from RNA pol-II promoter in response to heat stress 2 i
ATP transport 4 e, i
ribosome disassembly 1 i
mismatch repair 2 i
response to red light 2 i
DNA replication initiation 3 i
response to UV 2 i
DS break repair via break−induced replication 2 i
nuclear polyadenylation−dependent rRNA catabolic process 2 i
innate immune response 2 i
pyrimidine nucleotide−sugar transmembrane transport 1 p, e, i
plant−type hypersensitive response 2 i
error−prone translesion synthesis 2 i
0 1 2 3 4
-log10(adj_pval)
Transgressive higher in maj
Transgressive higher in diploids
DNA−repair
meiosis/mitosis
response to stress
transcription/translation regulation
transport