Reciprocal allopolyplid grasses (Festuca × Lolium) display stable patterns of genome dominance

Marek Glombik¹,², Dario Copetti³,⁴, Jan Bartos¹, Stepan Stoces¹, Zbigniew Zwierzykowski⁵, Tom Ruttink⁶, Jonathan F. Wendel⁷, Martin Duchoslav⁸, Jaroslav Dolezel¹, Bruno Studer³*, David Kopecky¹*

¹) Institute of Experimental Botany of the Czech Academy of Sciences, Centre of the Region Haná for Biotechnological and Agricultural Research, Šlechtitelná 31, 77900 Olomouc, Czech Republic

²) National Centre for Biomolecular Research, Faculty of Science, Masaryk University, Kotlářská 2, 61137 Brno, Czech Republic

³) Molecular Plant Breeding, Institute of Agricultural Sciences, ETH Zurich, Universitätsstrasse 2, 8092 Zurich, Switzerland

⁴) Department of Evolutionary Biology and Environmental Studies, University of Zurich, Zurich, Switzerland

⁵) Department of Environmental Stress Biology, Institute of Plant Genetics, Polish Academy of Sciences, Strzeszyńska 34, 60-479 Poznań, Poland

⁶) Flanders Research Institute for Agriculture, Fisheries and Food (ILVO), Plant Sciences Unit, Caritasstraat 39, 9090 Melle, Belgium

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/TPJ.15375

This article is protected by copyright. All rights reserved.
Running title: Genome dominance in plant allopolyploids

Key words: Gene expression, allopolyploidy, interspecific hybrids, \textit{cis/trans} regulation, genome dominance, homoeolog

*) authors for correspondence: \textbf{Bruno Studer}, Molecular Plant Breeding, Institute of Agricultural Sciences, ETH Zurich, Universitaetstrasse 2, 8092 Zurich, Switzerland brunostuder@usys.ethz.ch; \textbf{David Kopecky}, Institute of Experimental Botany of the Czech Academy of Sciences, Centre of the Region Haná for Biotechnological and Agricultural Research, Šlechtitelů 31, 77900 Olomouc, Czech Republic, kopecky@ueb.cas.cz
Summary

Allopolyploidization entailing the merger of two distinct genomes in a single hybrid organism, is an important process in plant evolution and a valuable tool in breeding programs. Newly established hybrids often experience massive genomic perturbations, including karyotype reshuffling and gene expression modifications. These phenomena may be asymmetric with respect to the two progenitors, with one of the parental genomes being ‘dominant’. Such ‘genome dominance’ can manifest in several ways, including biased homoeolog gene expression and expression level dominance. Here we employed a k-mer based approach to study gene expression in reciprocal *Festuca pratensis* Huds. × *Lolium multiflorum* Lam. allopolyploid grasses. Our study revealed significantly more genes where expression mimicked that of the *Lolium* parent compared to the *Festuca* parent. This genome dominance was heritable to successive generation and its direction was only slightly modified by environmental conditions and plant age. Our results suggest that *Lolium* genome dominance was at least partially caused by its more efficient trans-acting gene expression regulatory factors. Unraveling the mechanisms responsible for propagation of parent-specific traits in hybrid crops contributes to our understanding of allopolyploid genome evolution and opens a way to targeted breeding strategies.

Introduction

Interspecific hybridization and polyploidization are fundamental processes underlying the evolution of plant species. All angiosperms have experienced at least one round of whole genome doubling (WGD) during their evolutionary history, with multiple WGD episodes being commonplace during angiosperm diversification (Jiao et al., 2011; Wendel, 2015; Soltis and Soltis, 2016; Soltis et al., 2016; Van de Peer et al., 2017; Landis et al., 2018; Escudero and Wendel, 2020). The evolutionary success of polyploids is presumed to reflect the evolutionary genetic novelty acquired during polyploidization (Soltis and Soltis, 1993), although insights into the genomic basis of adaptive traits is only beginning to emerge (Nieto-Feliner et al., 2020). Two types of polyploids are widely recognized, auto- and allopolyploids, although these terms have several definitions and can represent endpoints of several continua (Wendel and Doyle, 2005; Doyle and Sherman-Broyles, 2017; Mason and Wendel, 2020). Under the taxonomic definition, autopolyploids arise from WGD within a species, whereas allopolyploids arise from WGD
accompanying the merger of two distinct species. Allopolyploidy often leads to heterosis, the manifestation of competitive advantage of the F1 over diploid progenitors (Comai, 2005), masks deleterious recessive alleles leading to increased mutational robustness (Madlung, 2013), and conveys increased stress tolerance through several mechanisms such as delayed reproduction, longer life span, and increased defense against herbivores and pathogens (Lohaus and Van de Peer, 2016).

Besides the key roles that interspecific hybridization and polyploidization play in plant evolution and diversification, these processes lay at the origin of many crops including wheat, oilseed rape, banana and cotton (Renny-Byfield and Wendel, 2014; Salmon-Minkov et al., 2016). Moreover, interspecific hybridization is frequently used in breeding programs to incorporate novel alleles into crops, often for beneficial characteristics such as tolerance or resistance to abiotic or biotic stresses. This approach has been used to develop hybrid grasses within the Festuca-Lolium complex. Festulolium (Festuca × Lolium) hybrids combine good palatability and digestibility, rapid establishment from seed, and competitiveness from ryegrasses (Lolium spp.) with abiotic and biotic stress tolerance from fescues (Festuca spp.) (Ghesquiere et al., 2010).

Despite the potential benefits of hybridization and polyploidization, many newly formed allopolyploids face survival challenges imposed by a diversity of genic incompatibilities and genomic instabilities resulting from the merger of differentiated genomes. These include numerous forms of chromosomal rearrangements as well as mitotic and meiotic abnormalities (Glombik et al., 2020). In addition to structural genomic changes, allopolyploidy is often accompanied by genome-wide gene expression alterations (Yoo et al., 2014), likely mediated by epigenetic responses to the novel genomic environment (Song and Chen, 2015; Ding and Chen, 2018). For instance, Yoo et al. (2013) monitored gene expression of over 50,000 genes in natural allopolyploid cotton, as well as in its putative progenitor species, their F1 interspecific hybrids, synthetic allopolyploids, and natural and domesticated cultivars. It was reported that (1) many duplicated genes collectively display expression level dominance (ELD), where the overall level of expression (for both homoeologs) is equivalent to the expression level of one of the two parents, with a bias favoring one parent, even though on a gene-by-gene basis this dominance could be reversed; and (2) genome-wide homoeolog expression bias (HEB), where the homoeolog from a dominant parental genome contributes to the expression of the particular gene significantly more than the homoeolog from a submissive parental genome. Indeed, polyploidy appears to be
Gene expression is transcriptionally regulated by cis-acting regulatory elements and trans-acting regulatory factors (Williams et al., 2007). Cis-acting elements have an allele-specific effect on the expression of the downstream gene, whereas trans-acting factors affect expression of other genes in the genome. In the case of allopolyploids, trans-acting factors may influence all homoeologous alleles of downstream-regulated genes in a hybrid nucleus, although not necessarily equivalently (Bottani et al., 2018). Cis-acting regulatory elements are believed to predominantly account for changes in gene expression in hybrids of genetically more divergent progenitors (e.g. interspecific hybrids), while trans-acting factors predominantly account for gene expression changes in hybrids of genetically less divergent progenitors (e.g. intraspecific hybrids) (Lemos et al., 2008; Wittkopp et al., 2008; Bell et al., 2013). One of the intriguing questions regarding cis- and trans-acting regulation in polyploids concerns its mechanistic relationship to the expression-level phenomena of ELD and HEB and how this relates to the phenotype. Bao et al. (2019) described a high level of trans-acting regulation accompanying the domestication of cotton, and proposed that enhanced trans-acting regulatory evolution may be both a general property of allopolyploidy and help to explain the evolutionary novelty accompanying allopolyploidization (Bottani et al., 2018; Hu and Wendel, 2019).

Crosses between the two forage grasses Festuca pratensis Huds. and Lolium multiflorum Lam. produce hybrids of considerable agricultural interest. These Festulolium hybrids display Lolium genome dominance in the form of gradual replacement of Festuca chromosomes by those of Lolium in subsequent generations (Kopecký et al., 2006; Zwierzykowski et al., 2006). Here, we investigate genome-wide transcriptome changes in the first two generations of reciprocal L. multiflorum × F. pratensis hybrids using RNA-sequencing (RNA-seq). K-mer based RNA-seq analysis was used to avoid any potential bias of the conventional implementation of mapping reads from hybrids to the reference sequence from one parental genome. In addition to exploring the consequences of genomic merger on patterns of gene expression in two successive generations (F1 and F2), we tested if the growth conditions and plant age may alter genome dominance at both
HEB and ELD levels. We further aim to provide insights into the mechanisms of ELD and HEB potentially applicable to other newly developed hybrid organisms.

Results

Development of reference-guided transcriptome assemblies and k-mer databases for Festuca and Lolium

Mapping parental RNA-seq data to the reference genome sequence assembly of the respective species and creating reference-guided transcriptome assemblies yielded 102,672 and 101,268 transcripts for *F. pratensis* and *L. multiflorum*, respectively. Filtering the transcriptomes for redundancy and longest isoform reduced this number to 39,764 transcripts for *F. pratensis* and 44,603 transcripts for *L. multiflorum*. Translated sequences of the same transcripts were then compared by reciprocal BLAST between the two species-specific reference-guided transcriptome assemblies, leading to the identification of 10,381 putative orthologous gene pairs between the two parental species. In order to examine possible reciprocal reference-read mapping bias, RNA-seq reads from both parental species were mapped to their own reference transcriptome (“self reference”) or to the reference transcriptome of the other parent (“non-self reference”). We observed a statistically significant difference in the number of genes displaying differential expression from two parental homoeologs (hypergeometric test, *p* << 0.001; Figure S1A, B). Using the *Festuca* reference transcriptome, 508 genes were significantly over-expressed in *Festuca* compared to *Lolium*, while 313 genes displayed significant over-expression in *Lolium* compared to *Festuca*. In contrast, using the *Lolium* reference transcriptome, 783 genes were over-expressed in *Lolium* and 419 genes were over-expressed in *Festuca*, indicating strong bias towards the “self reference” transcriptome (Figure S1A, B). Moreover, in both comparisons, less than half of the genes showing differential gene expression towards one of the parents (245 and 230 for F>L and L>F, respectively) were shared when mapped to the transcriptome reference of the other species (Table S1). These results indicate bias towards the “self reference” transcriptome which would subsequently lead to incorrect differential gene expression assessment of genome dominance on expression level in hybrid samples.

To reduce the bias in assessing differences in the sequences of *Festuca* and *Lolium* and the subsequent differential gene expression of their hybrid, a k-mer based approach was implemented. K-mers of length 31 bp were used because they have been previously been proven to be suitable
for identification of different allelic variants (Rahman et al., 2018; Voichek and Weigel, 2020), a large number of these is unique for each species in our RNA-Seq dataset, and it is the longest k-mer length that can effectively be represented on 64-bit machines. From RNA-seq reads from both parents, over 83 million k-mers of 31 bp length were identified. Such strings constituted two species-specific k-mer databases with over 20 million k-mers for each parental species. After the filtering steps (see methods), almost 165,000 pairs of k-mers between parents were identified including 50,362 unique and species-specific k-mers mapped to 10,381 orthologous gene pairs with unique gene assignment and corresponding to identified orthologous gene pairs (Table S2). Additional filtering steps were applied to address the fact that more than one k-mer of 31 bp length can originate from a single 100 bp read and thus, to remove redundancy. Despite the high reduction in the number of k-mers used for gene expression analysis, this condition was chosen to achieve accurate and unbiased results. This resulted in 26,494 k-mers (and their associated read counts) usable for determination of gene expression between 7,958 of Festuca and Lolium genes, and 5,790 of these genes were sufficiently expressed in all samples to be used for comparisons between parents and/or hybrids (Table S3).

To reveal the efficacy of the k-mer based approach, Spearman’s correlations were calculated between three read count datasets (7,958 genes): 1, generated by the k-mer based approach; 2, generated by mapping to the reference transcriptome of the corresponding parent “self reference”; and 3, generated by mapping to the reference transcriptome of the other parent, “non-self reference”) for each species. For Festuca, the average Spearman’s correlations between the k-mer counts and the “self reference” was 0.95; between k-mer counts and the “non-self reference” was 0.94; and between the “self reference” and the “non-self reference” was 0.92. Similarly, for Lolium the average Spearman’s correlation between the k-mer counts and the “self reference” was 0.96; between k-mer counts and the “non-self reference” was 0.92; and between the “self reference” and the “non-self reference” was 0.95 (Figure S1C, D). These results indicate that our k-mer approach is a suitable strategy for hybrid gene expression analysis and is not biased by differential mapping efficacy to a reference sequence of either parent.

Homoeolog Expression Bias (HEB)
The analysis of HEB reflects the relative contribution of two parental homoeologs to expression of a particular gene, either on a gene-by-gene basis or on a genome-wide scale. In total, we were able to monitor homoeolog-specific gene expression of 5,790 genes described above in F1 and F2 generations of Festuca × Lolium reciprocal hybrids (Figure 1). Of those, 5,224 (90.2%) and 4,876 (82.5%) showed equivalent expression of both parental homoeologs in the three plants of Festuca × Lolium and three plants of Lolium × Festuca hybrids, respectively. A slightly higher number of genes (287 and 530 of 5,790 genes) displayed higher expression of the Festuca homoeolog, compared to the genes showing higher expression of the Lolium homoeolog (279 and 384 of 5,790) in F1 Festuca × Lolium and Lolium × Festuca hybrids, respectively (Figure 2).

The comparison of the relative contribution of parental homoeologs in hybrids with the expression levels in parents revealed that the transmission of parental expression levels accounted for 4,443 (76.7%) and 4,239 (73.0%) genes in Festuca × Lolium and Lolium × Festuca hybrids, respectively (Table 1). Additionally, 985 (17.0%) and 907 (15.7%) homoeolog pairs with significant differential expression level between the parents displayed non-biased expression in Festuca × Lolium and Lolium × Festuca hybrids, respectively. Another 349 (6.0%) and 619 (10.7%) genes had the same expression in parents, but displayed preferential expression from one parental homoeolog in Festuca × Lolium and Lolium × Festuca hybrids, respectively. There was a slightly higher number of genes for which the Festuca homoeolog was preferentially expressed than those for which the Lolium homoeolog was preferentially expressed in reciprocal hybrids. A low number of genes (13 and 39) reverted the expression bias observed in the parents (Table 1).

Hybrids were also searched for homoeolog-specific silencing. Homoeologs were considered silenced when the average CPM (count per million) value from all samples was < 1. Overall, only a few genes out of the 5,790 genes exhibited silencing of one homoeolog, but the pattern was the same in reciprocal hybrids: a higher number of genes was observed with a silenced Lolium homoeolog (7 and 40) than with a silenced Festuca homoeolog (1 and 4) (Figure 3).

Overall, HEB analysis revealed that both homoeologs contributed equally to gene expression in hybrids for the majority of genes (> 84% of 5,790 genes). The number of genes displaying preferential expression of the homoeolog of one parent was almost the same for over-expression of both Lolium and Festuca homoeologs, with only slightly more genes preferentially expressing the Festuca than the Lolium homoeolog. Parental legacy (one parent has a higher expression of a particular gene than the other parent, with the hybrid mirroring this difference) was observed for
70% or more of the genes studied. Moreover, it seems that the direction of the cross plays only a minimal role in determining homoeolog expression bias.

**Expression Level Dominance (ELD)**

A comparison of the total expression level of homoeologous gene pairs in hybrids to the expression level of their two parents enabled assignment of 4,851-4,905 genes to one of the 13 ELD categories previously described by Rapp et al. (2009). For the remaining 885-939 of genes in the reciprocal hybrids, it was not possible to assign unambiguously a particular gene to a category; such genes were classified as ‘unassigned’. From the assigned genes, 2,811 (57.7%) and 2,629 (54.2%) showed no difference in expression between both parents and between parents and hybrids (category 0) in three plants of *Festuca × Lolium* and three plants of *Lolium × Festuca* hybrids, respectively. Additivity, a category where parents differ in gene expression level and the hybrid displays a mid-parent level of expression, accounted for 272 (5.6%) and 274 (5.7%) genes. Transgressive expression was observed for 981 (20.2%) and 1,173 genes (24.1%) in the hybrid progeny with down-regulation being more frequent than up-regulation. Expression level dominance, where parents differ in the level of expression and the hybrid displays the expression level of one parent, was observed in 805 (16.6%) and 775 (15.9%) genes in *Festuca × Lolium* and *Lolium × Festuca* hybrids, respectively. Significantly more genes were assigned to *Lolium* expression dominance categories, where the hybrid displays the same expression level as the *Lolium* parent (category II and XI) as compared to the number of genes assigned to *Festuca* expression dominance (categories IV and IX) in *Festuca × Lolium* and *Lolium × Festuca* hybrids (489 vs. 316 and 434 vs. 341, respectively; Table 2, Table S4) using a paired sample sign test (p < 0.004). The results were highly consistent between reciprocal hybrids with 82.3% overlap of genes assigned to particular categories, suggesting that *Lolium* expression dominance is established immediately after hybridization with no effect of the cross direction.

**Transmission of the gene expression patterns to the F2 generation**

The investigation whether the gene expression changes associated with hybridization are heritable revealed that there were more genes significantly differentially expressed from one or the other
homoeolog (HEB) in three plants of F2 generation of *Festuca × Lolium* (1,016 compared to 566 genes in F1 hybrids). This homoeolog expression was slightly more biased towards the *Festuca* homoeologs, because 581 (10.0%) genes were preferentially expressed from the *Festuca* homoeolog compared to 435 (7.5%) genes expressed preferentially from the *Lolium* homoeolog (Figure 2). Considering expression differences in parents, the number of genes displaying parental legacy decreased only slightly, from 4,443 (76.7%) to 4,253 (73.5%). The most marked change was the increase in the number of genes expressed at the same level in parents, but preferentially from one homoeolog in hybrids (from 349 in F1 hybrids to 661 in F2 generation). This novel bias was more towards *Festuca* homoeologs as compared to *Lolium* homoeologs (187 vs. 162 in F1 hybrids as compared to 380 vs. 281 in F2; Table 1).

At the overall expression level (ELD), there was the same tendency towards higher number of genes with *Lolium* expression dominance as compared to *Festuca* expression dominance in both the F1 and F2 generations (489 vs. 316 and 494 vs. 283, respectively) with statistically significant difference using a paired sample sign test (p < 0.004). The overlap of genes in different categories between F1 F×L and F2 F×L generations was 60.3%. Interestingly, there were almost the same frequencies of the genes in the ELD categories, with a slight increase (from 542 to 741 genes) in transgressively down-regulated genes and a decrease in the number of transgressively up-regulated genes (from 439 to 341 genes). The number of genes displaying *Lolium* expression dominance was almost identical in both the F1 and F2 generations (489 and 494), while the number of genes with *Festuca* expression dominance slightly decreased from 316 (6.5%) to 283 (5.8%; Table 2, Table S4). These results indicate that the established *Lolium* genome dominance at ELD level was transmitted to the successive F2 generation.

**Gene expression changes caused by different plant age and growth conditions**

To test whether the *Lolium* genome dominance could be altered in hybrids by different growth conditions and plant age, the F1 and F2 hybrids were re-analyzed after four years of cultivation and also after cold treatment (4°C for three weeks). After four years of cultivation HEB and ELD were only slightly altered. A higher number of genes with equal contribution of both homoeologs to gene expression was found early after the establishment of F1 generation than after four years of cultivation (5,224 vs. 4,811 genes; p < 0.001 based on the hypergeometric test). However, these
differences were not statistically significant in F2 (4,774 vs. 4,806 genes; Figure 2). Consequently, the number of genes with equal parental expression that became biased in the hybrids increased in F1 from 349 to 680 genes (p < 0.001), but decreased slightly in F2 from 661 to 627 genes (not statistically significant). In terms of ELD, the most significant change after four years of cultivation was an increase in the number of genes with transgressive up-regulation in both F1 and F2 (Table S4). The number of genes displaying *Lolium* expression dominance remained almost the same (489 vs. 475 genes in F1 and 494 vs. 496 genes in F2; not statistically significant in both cases), similar to those displaying *Festuca* expression dominance (316 vs. 299 genes in F1 and 283 vs. 288 genes in F2; not statistically significant in both cases; Table 2). Our analysis further revealed that the four years of cultivation slightly stabilized gene expression; the overlap of genes in categories between F1 and F2 increased from 60.3% to 67.5%. However, the overlaps for the *Lolium* expression dominance and *Festuca* expression dominance categories between the samples were always significant based on the hypergeometric test (p < 0.001) (Figure S2).

The cold stress treatment was only performed after four years of cultivation, meaning that gene expression changes might reflect either or both the length of cultivation and cold treatment. Thus, the expression data was compared not only to that of the parents, but also to hybrids after four years of cultivation without cold treatment and only differentially expressed genes identified between plants after four years of cultivation and their corresponding clones after cold treatment were considered. We identified 1,176 (518 up-regulated and 658 down-regulated) and 754 (263 up-regulated and 491 down-regulated) differentially expressed genes in F1 and F2 generation, respectively, with 468 being identified in both generations. Of these 468 genes, 464 (99.1%) shared the same gene expression pattern, with 161 (34.4%) being up-regulated and 307 (64.7%) down-regulated after cold stress treatment compared to control conditions.

HEB analysis of F1 and F2 hybrids after cold stress treatment showed a trend similar to other samples (F1 and the F2 generation early after establishment and after four years of cultivation). Similar to the hybrids after four years of cultivation, most genes had equal contribution from both homoeologs (4,581 and 4,662 genes in F1 and F2 generation, respectively). For genes with unequal contribution of homoeologs, there were only slightly more genes with preferential expression from the *Festuca* homoeolog (633 and 613 genes in F1 and F2, respectively) compared to those with preferential expression from the *Lolium* homoeolog (576 and 515 genes in F1 and F2, respectively), but in both cases these differences were small. Slightly higher frequencies were
revealed in the subset of 468 genes differentially expressed after cold stress in both generations (145 genes in F1 and 139 genes in F2; Figure 4), with some asymmetry favoring the expression of *Festuca* homoeologs.

ELD analysis showed that the most significant change in gene expression after cold stress was the increase in the number of genes displaying transgressive down-regulation and transgressive up-regulation. The frequency of genes displaying expression level dominance remained almost the same as compared to plants after four years of cultivation without cold stress, but there was a slight increase in the frequency of genes with *Festuca* expression dominance (from 299 to 328 genes and from 288 to 317 genes in F1 and F2 generations, respectively), and a slight decrease in the frequency of genes with *Lolium* expression dominance (from 475 to 446 genes and from 496 to 455 genes in F1 and F2 generations, respectively). Considering only genes with significant differential expression between plants exposed to cold stress and from controlled conditions, the pattern was reverted: there were slightly more genes exhibiting *Festuca* expression dominance as compared to those with *Lolium* expression dominance (86 vs. 79 genes and 61 vs. 57 genes in F1 and F2 generation, respectively; Table 3, Table S5). Thus, even though the pattern of *Lolium* genome dominance remained almost unchanged after cold treatment, the expression of *Festuca* homoeologs is predominantly being modified in genes exhibiting the highest expression level changes. Moreover, we revealed that cold stress stabilized gene expression: the overlap of genes in ELD categories between F1 and F2 generations increased from 60.3% in young hybrids to 67.5% in hybrids after four years of cultivation to 83.1% in hybrids after cold stress and the overlaps of *Lolium* expression dominance and *Festuca* expression dominance categorized genes showed to be significant (hypergeometric test, p < 0.001).

**Cis and trans regulatory divergence**

To estimate the effects of *cis* and *trans* regulatory divergence on changes in gene expression in hybrids, the approach of McManus et al. (2010) was used, in which all 5790 genes were assigned to one of the seven categories: *cis*-only, *trans*-only, *cis* + *trans*, *cis* × *trans*, compensatory, conserved and ambiguous (see Methods).

In F1 *Festuca* × *Lolium* and *Lolium* × *Festuca* hybrids, a majority of genes could be classified as conserved (2,487 and 1,595 genes) and ambiguous (2,928 and 3,496 genes), with slightly more
genes regulated by \textit{cis}-only (159 and 167 genes) vs. \textit{trans}-only (113 and 155 genes, respectively). When both \textit{cis} and \textit{trans} regulatory changes occurred, they favored the expression of the same homoeolog (\textit{cis} + \textit{trans}) rather than the opposite homoeolog (\textit{cis} \times \textit{trans}) (Table S6). In F2 generation of \textit{Festuca} \times \textit{Lolium} hybrids, there was a reduction of genes classified as conserved and an increase in all other categories, especially \textit{cis} \times \textit{trans}, \textit{cis} + \textit{trans} and compensatory. Similarly, the most obvious change in the hybrids after four years of cultivation was an increased number of genes in compensatory, \textit{cis} + \textit{trans} and \textit{cis} \times \textit{trans} categories (Table S6). On the other hand, cold stress treatment reduced the number of genes in all categories except the category conserved. The analysis of 468 genes up- or down-regulated after cold stress in both F1 and F2 generations revealed higher percentage of genes in the categories with \textit{trans} regulatory divergence: \textit{trans}-only, \textit{cis} + \textit{trans}, \textit{cis} \times \textit{trans} and compensatory.

The relationship between HEB and ELD

To determine how homoeolog expression bias (HEB) gives rise to ELD, the expression of both homoeologs in hybrids was separately compared to the expression level in their parents. Strikingly, for both \textit{Lolium} expression level dominance (category II) and \textit{Festuca} expression level dominance (category IV), up-regulation of the homoeolog of the submissive parent was observed in a majority of the genes (Table 4). However, the number of up-regulated homoeologs of the submissive parent differed between \textit{Lolium} and \textit{Festuca} expression dominance. In category II (\textit{Lolium} expression dominance), the \textit{Festuca} homoeolog was up-regulated in 136 (91.0%), 160 (87.9%) and 155 (83.8%) genes, both homoeologs were up-regulated in 11 (7.3%), 12 (6.6%) and 21 (11.4%) genes and the \textit{Lolium} homoeolog was up-regulated in three (2.0%), six (3.3%) and five (2.7%) genes in \textit{Festuca} \times \textit{Lolium}, \textit{Lolium} \times \textit{Festuca} and F2 generation of \textit{Festuca} \times \textit{Lolium} hybrids, respectively. On the other hand, in category IV (\textit{Festuca} expression dominance), the \textit{Lolium} homoeolog was up-regulated in 46 (70.8%), 52 (76.9%), and 52 (76.5%) genes, both homoeologs were up-regulated in three (4.6%), four (5.1%), and one (1.5%) genes and the \textit{Festuca} homoeolog was up-regulated in 16 (24.6%), 10 (12.8%), and 9 (13.2%) genes in the \textit{Festuca} \times \textit{Lolium}, \textit{Lolium} \times \textit{Festuca} and F2 generation of \textit{Festuca} \times \textit{Lolium} hybrids, respectively (Table 4). The same trend was observed in F1 and F2 generations after four years of cultivation, independent on whether the hybrids were exposed to cold stress or not (Table S7). Reciprocally, that is, in case of the expression level dominance with the expression level in hybrids being equal to the parent
with lower expression (categories XI and IX), the *Festuca* homoeolog was more efficiently downregulated than the *Lolium* homoeolog: 12 (9.2%) vs. 8 (7.5%), 27 (20.3%) vs. 24 (17.2%) and 28 (19.8%) vs. 23 (15.1%) genes in the F1 *Festuca × Lolium*, F1 *Lolium × Festuca* and F2 *Festuca × Lolium* hybrids, respectively (Table 4). Taken together, these data show that *Festuca* homoeologs are more prone to up- and down-regulation than *Lolium* homoeologs in all comparisons.

**Biological processes associated with gene sets displaying expression level dominance**

To explore potential functional consequences of ELD we focused on genes displaying *Festuca* expression dominance and *Lolium* expression dominance. In total, 678 and 886 genes displaying *Festuca* expression dominance and *Lolium* expression dominance were identified, respectively, in both reciprocal hybrids, in F1 and F2 hybrids early after establishment and in F1 and F2 hybrids after four years of cultivation (F1 F×L, F2 F×L, F1 L×F, F1 F×L aging and F2 F×L aging). GO terms were transferred from the functional annotation of *L. perenne* orthologues to our *L. multiflorum* reference transcriptome, and significantly enriched GO terms (p < 0.01) were identified per ELD group (Table S8). We paid special attention to up- or down-regulated genes after cold stress shared between F1 and F2 generations of hybrids. Down-regulated gene sets displayed enrichment of GO terms concerning: cell wall organization or biogenesis (GO:0071554) with several transcripts encoding cellulose synthase A catalytic subunit 3, pectinesterase 31, protein cellulose synthase interactive 1, expansin-A2; secondary metabolic process (GO:0019748) containing several transcripts encoding isoflavone reductase, cinnamyl alcohol dehydrogenase, hydroxycinnamoyltransferase 2, cinnamoyl CoA reductase; cuticle development (GO:0042335) with two transcripts encoding 3-ketoacyl-CoA synthase; and water transport (GO:0006833) comprised of three transcripts, all encoding PIP type aquaporins. The set of up-regulated genes displayed GO term enrichment of the categories: response to cold (GO:009409) with transcripts encoding cysteine proteinase inhibitor, beta-amylase 3, CBL-interacting protein kinase 19 and 9, arginine decarboxylase 1, Phytochrome B, diacylglycerol kinase 2, WRKY27 transcription factor, HVA22-like protein e, and GIGANTEA; and polyamine metabolic process (GO:0006595) containing three transcripts encoding N-carbamoylputrescine amidase, arginine decarboxylase 1, organelle RRM domain-containing protein 2, mitochondrial.
Discussion

A novel k-mer approach for ELD and HEB analysis in hybrids using RNA-seq

Currently available bioinformatic approaches to study ELD and HEB in hybrids are based on RNA-seq and on mapping sequence reads to a reference genome/transcriptome (usually from one of the parents) followed by variant calling to identify loci suitable for discrimination of homoeologous reads. However, such approaches may introduce bias due to a different efficiency of mapping reads from two parental genomes to a single reference genome/transcriptome (Degner et al., 2009; Satya et al., 2012; Liu et al., 2014). Here we found that using the reference transcriptome sequence from one parental species significantly affected the results and introduced a bias in genome dominance assessment (Figure S1). Several tools like Sailfish and Kallisto have been developed to minimize this bias by focusing more on k-mer composition of reads and references, which is subsequently used to accurately and quickly assign specific reads to specific genes/transcripts (Patro et al., 2014; Bray et al., 2016). However, these tools only report the transcript abundancy, or in case of Kallisto pseudomapping also bam-like files which then may be processed for variant calling. Moreover, as shown in our results, Kallisto may still be influenced by the reference transcriptome in more distant species (Figure S1; Table S1). Only recently, GWAS analysis of *Escherichia coli* and *Arabidopsis thaliana* presented a k-mer based approach for inferring sequence-phenotype associations directly from sequenced reads (Rahman et al., 2018; Voichek and Weigel, 2020). These studies show that using k-mers directly derived from the sequencing data largely agrees with the results obtained using a standard approach in GWAS studies of the single species. Our results indicate that the k-mer based approach for inferring differences between two species avoids any bias caused by mapping reads from hybrids to a reference genome/transcriptome from one of the parental species of the hybrid (Figure S1). Thus, using k-mers may be a useful approach for quantifying homoeolog-specific gene expression in allopolyploids.

Gene expression alterations induced by genome merger

Immediately after merging two different genomes via interspecific hybridization, multiple genome and epigenome modifications are initiated, often collectively referred to as “genomic shock”
One of the consequences is a genome dominance, whereby one of the two co-resident genomes becomes “dominant”. Several distinct phenomena are encompassed by this concept, including preferential gene expression favoring the dominant genome. We note that genome dominance does not mean ‘complete’ dominance, as the absolute silencing of the submissive genome has not been observed in plant hybrids or allopolyploids. Often some genes in the submissive genome are more highly expressed than their homoeologous counterparts in the dominant genome (Edger et al., 2017).

Although differential homoeolog expression appears to be ubiquitous at the onset of polyploidy, it likely takes some time for a homoeolog to be completely silenced. Studies in cotton identified 180 and 191 genes (0.71% and 0.75% of all genes screened), where one homoeolog was silenced in natural allopolyploids and only 14 such genes were silenced in synthetically developed allopolyploids (Yoo et al., 2013). Similarly, we found that only 8 and 44 genes showed homoeolog silencing in Festuca × Lolium and Lolium × Festuca hybrids, respectively (Figure 3). The general lack of genes commonly displaying silencing in various hybrids may indicate that this process is stochastic and prone to variation between progenies from the same cross (Buggs et al., 2009; Soltis et al., 2012; Boatwright et al., 2018; our results). The low frequency of rapid silencing suggests a retention of a majority of homoeologs in allopolyploids, which then potentially serve as a reservoir of genetic variation that may become valuable during subsequent stages of polyploid evolution and diversification (Adams et al., 2003; Nieto-Feliner et al., 2020).

Two aspects of genome dominance have been commonly defined and distinguished: homoeolog expression bias (HEB) and expression level dominance (ELD) (Grover et al., 2012; Yoo et al., 2013). In some hybrids, the genome dominating in the homoeolog-specific expression (HEB) is not the same as the genome dominating at the ELD level. In fish hybrids Megalobrama amblycephala × Culter alburnus, homoeologs of the M. amblycephala genome are more frequently expressed than their counterparts (HEB), while the overall expression in hybrids more frequently mimic the expression level of the C. alburnus genome (ELD; Ren et al., 2019). Similarly, we also found more genes with higher expression of the Festuca homoeolog than genes with higher expression of Lolium homoeologs, while overall gene expression levels in hybrids reflect more frequently the expression of the Lolium parent in all analyzed samples of hybrids, including those after four years of cultivation and after cold stress treatment (Figure 2, Figure 4, Table 2 and Table 3). In a genome dominance context, we note that the unequal contribution of

(McClintock, 1984).
both homoeologs to the overall expression of the particular gene (HEB) is not necessarily evolutionarily meaningful. To the extent that homoeologs are functionally equivalent and that total expression is important to establish a given phenotype (at any level), ELD is a key feature of polyploidy and genome dominance (Nieto-Feliner et al., 2020).

**Predetermination, heritability, and evolution of genome dominance**

Genome dominance at both the HEB and ELD levels has been observed in many ancient or relatively recent allopolyploids, including maize, *Arabidopsis*, cotton, bread wheat and *Tragopogon* (reviewed in Bird et al., 2018), while there are exceptions displaying little dominance (at the HEB level), such as *Capsella bursa-pastoris* and the Ethiopian cereal teff (Douglas et al., 2015; VanBuren et al., 2020). It seems likely that the genomic conditions for dominance are mechanistically established immediately after merging of two distinct genomes in a single nucleus, and in the absence of selection (Adams et al., 2003; Flagel et al., 2008). Despite the scarcity of studies analyzing genomic responses during the first few generations after genome merging, our observations suggest that it appears to be mechanistically deterministic, and imply that genome dominance (both HEB and ELD) is heritable. We are aware of only a few studies in which this idea has been formally tested. In fish hybrids, *Megalobrama amblycephala × Culter alburnus*, the number of genes in the expression dominance categories, together with additivity, decreased between the F1 and F3 generations (Ren et al., 2019), suggesting a gradual weakening of parental effects in hybrids. Here, we observed no significant reduction of the number of the genes displaying either *Lolium* expression dominance or *Festuca* expression dominance in the F2 generation (489 vs. 494 and 316 vs. 283 genes, respectively), demonstrating that these alterations of gene expression level are heritable.

To the extent that genome dominance in F1 hybrids is mechanistically specified, one may assume that genome dominance is predetermined. We found that *Festuca*-genome HEB and *Lolium*-genome ELD were displayed in all analyzed hybrid plants, independent of the direction of the cross (Table 2). Alger and Edger (2020) hypothesized that environmental conditions might influence which genome becomes dominant in certain hybrids and allopolyploids. However, we observed that environmental or physiological conditions did not revert the established genome dominance. The number of genes displaying *Lolium* expression dominance and *Festuca*
expression dominance (ELD) did not change significantly after the hybrids were cultivated for four years and similar results were found after cold stress treatment (Table 3). Taken together, our data suggest that genome dominance is predetermined and independent from the direction of the cross, parental genotypes and cold stress conditions.

Over the longer term it is expected that homoeologous gene expression, including ELD, will be subjected to evolutionary forces, including those imposed by selection and drift and in the context of genome fractionation. Yoo et al. (2014) showed that older allopolyploids exhibit more ELD relative to neo-allopolyploids. In contrast, Wu et al. (2018) identified about one-third of the genes displaying ELD in newly resynthesized *B. napus*, but only one-fifth of the genes retained ELD in natural allotetraploid *B. napus* after ~7,500 years of independent evolution and domestication (Li et al., 2020). This suggests that at least in some cases the consequences of the genomic shock experienced by the newly established hybrids might be alleviated during subsequent evolutionary processes and that gene expression may gradually become balanced and stabilized. Similarly, comparison of F1 hybrids, synthetically developed allopolyploids and naturally occurring allopolyploids and domesticated cultivars revealed that the homoeolog expression bias can influence increasingly more genes in successive generations (Yoo et al., 2013; Edger et al., 2017).

Generally, older allopolyploids such as cotton and allopolyploid *Brachypodium* seem to display a slightly lower frequency of genes with parental legacy (about 63%-65%) compared to young (~100 years old *Tragopogon*; 69%) or newly established allopolyploids (*Festuca × Lolium* hybrids; 73%-77%) (Yoo et al., 2013; Takahagi et al., 2018; Shan et al., 2020; Table 1). Such balancing and stabilizing gene expression patterns might also be revealed with the ontogeny of the hybrid plant. We found that the overlap among the genes in different gene expression categories between F1 and F2 hybrids is higher after four years of cultivation compared to newly established hybrids (60.3% vs. 67.5%, Table S4). Collectively, these results reveal the complex temporal and evolutionary contexts that can generate diverse homoeologous expression patterns.

**Mechanisms underlying genome dominance**

Our knowledge regarding the mechanisms responsible for HEB and ELD remains vague. Transposable elements (TEs) have been proposed to play an important role, due to frequent repression of adjacent genes (Hollister and Gaut, 2009; Freeling et al., 2012; Edger et al., 2017,
Similarly, gene expression can be affected by small RNAs (sRNAs), including microRNAs (miRNAs) and small interfering RNAs (siRNAs) (Ng et al., 2012).

In addition to the possible role of TEs and sRNAs in uneven contribution of homoeologs to gene expression in hybrids, mismatches between effectors (transcription factors) and their target genes may generate expression alterations (Bottani et al., 2018; Hu and Wendel, 2019). The mismatches may arise from the divergence of the relevant cis- and/or trans-acting regulatory machinery. Cis changes are allele-specific, affecting expression of the allele linked to the changes and refer mainly to changes in the promoter region of a gene. In contrast, trans changes are diffusible, potentially effecting expression of both parental alleles, due to changes that affect the timing, level or activity of the transcription factors or other regulators controlling expression of TFs or their targets (McManus et al., 2010). We investigated if and how the changes in cis-acting and/or trans-acting regulation can modify gene expression and contribute to genome dominance. In theory, if there is a higher expression of a gene in parent A than in parent B and the hybrid displays the same gene expression level as the parent A, then, there can be either up-regulation of the A homoeolog, up-regulation of B homoeolog or up-regulation of both. We found that up-regulation of homoeolog from a submissive parent is the most likely scenario for both Festuca and Lolium expression dominance in Festuca × Lolium hybrids. However, it seems that the Lolium transcriptome machinery can modulate (activate or repress) the expression of Festuca homoeologs more effectively than vice versa. The higher efficiency of Lolium regulators compared to those of Festuca was found for both up- and down-regulations of the counterpart homoeolog (Table 4). The results are highly consistent, as the same pattern was observed in the F2 generation, both F1 and F2 generations after four years of cultivation and in reciprocal crosses. Similar results were observed in a homoploid F1 hybrid and synthetic and two natural allopolyploids of cotton and B. napus, where up-regulation of homoeologs from the submissive genome were more frequent than the opposite situation (Yoo et al., 2013; Li et al., 2020). Much of this behavior may reflect a difference in parental spectrum of transcription factors, their concentrations and/or their affinities for targeted gene promoters (Bottani et al., 2018). From a practical standpoint, an increased understanding of the relative efficiency of transcription factors in parents may enable future harnessing of this information to manipulate genome dominance in hybrids for agronomic purposes.
Conclusions

Here we used a k-mer based RNA-seq approach to provide insights into molecular mechanisms underlying genome dominance in *Festuca-Lolium* allopolyploids. We showed that the expression alterations at the onset of allopolyploidy are heritable and to a large extent independent of plant age or environmental conditions. We also demonstrated that gene expression level dominance mainly reflects up- or down-regulation of homoeologs of the submissive genome. We speculate that besides other factors, this might be at least partially caused by the divergence of trans-acting regulatory factors between the progenitors, and their cascading effects on downstream gene expression. Future insights are likely to derive from studies on dynamics of transcription factor concentrations, affinities, binding kinetics, and changes in chromatin architecture resulting from hybridization and whole genome doubling. Understanding the molecular basis of genome dominance will have practical implications with respect to plant breeding. Enhanced understanding of the molecular genetic underpinnings of genome dominance will increase predictive power with respect to parental combinations that might yield desired phenotypes.

Methods

Plant material

Synthetically derived autotetraploid *F. pratensis* cultivar ‘Westa’ and synthetically derived autotetraploid *L. multiflorum* cultivar ‘Mitos’ were crossed to produce several F1 hybrids. Out of these, three independent F1 hybrid *Festuca ♀ × Lolium ♂* (F1 F×L) plants and three independent reciprocal *Lolium ♀ × Festuca ♂* (F1 L×F) plants were randomly selected. One plant of the F1 F×L hybrid used for the analysis was found to be self-compatible and three siblings of an F2 generation were selected (F2 F×L; Figure 1). Nine plants (three plants of F×L, three plants of F1 L×F and three plants of F2 generation of F×L) were used in this study. The genomic composition of all hybrid plants used in this study was verified by genomic in situ hybridization (GISH). All plants contained an equal proportion of parental genomes with 14 chromosomes each from *F. pratensis* and *L. multiflorum* (Figure 1). All plant material was grown in a cultivation chamber (Weiss-Gallenkamp, Loughborough, United Kingdom). Plants were cut to approximately 5 cm height and grown for three more weeks in the cultivation chamber (day: 14 h, 24 °C, 50% relative humidity, 20.000 lux; night: 10 h, 20 °C, 50% relative humidity). From each plant, the basal parts of a shoot and part of the root system (about 1 cm of each organ) were sampled for transcriptome...
sequencing. After sampling tissue for RNA extraction (tissue frozen in liquid nitrogen), remaining parts of each plant were separated into two clones and grown in a greenhouse. The sampling of F1 F×L and F2 F×L was repeated after four years. One clone of each hybrid plant was grown under the same conditions as the experiment above, and the second clone was grown under cold stress conditions (day: 8 h, 4 °C, 85% relative humidity, 20.000 lux; night: 16 h, 4 °C, 85% relative humidity), both for three weeks in the cultivation chambers. Three biological replicates (the tissues from different ramets of the same plant) from each plant were sampled.

**RNA extraction and sequencing**

RNA was extracted from 100 mg of tissue using the RNeasy Plant Mini Kit (Qiagen, Inc., Valencia, CA, USA) according to the manufacturer’s instructions. The quality of total RNA was checked using an RNA Pico 6000 chip on a Bioanalyzer 2100 (Agilent Technologies, Inc., Santa Clara, CA, USA) and only samples with RNA integrity number (RIN) > 7 were sequenced. RNA-seq was performed at Istituto di Genomica Applicata (IGA Technology Services, Udine, Italy). Libraries were developed with the TruSeq RNA Sample Prep kit v2 according to the manufacturer’s instructions (Illumina Inc., San Diego, CA, USA). Indexed adapters were ligated to the complementary DNA and 200 ± 25 bp fragments were gel purified and PCR-amplified. Libraries were quantified using a Bioanalyzer 2100 (Agilent Technologies), pooled in equimolar amounts for sequencing using Illumina HiSeq2000 or HiSeq2500 instruments to produce 100 or 125 bp paired-end reads..

**Analysis of RNA-seq data**

Reads of the two parents were first quality-trimmed (LEADING:3 TRAILING:3 SLIDINGWINDOW:4:15 MINLEN:50) using Trimmomatic ver. 0.36 (Bolger et al., 2014). Reference-guided transcriptome assembly for both parental species and identification of orthologous gene pairs between the parental species was performed. Quality-trimmed reads were mapped to the scaffolds of reference genome assemblies of the respective parental species with STAR ver. 2.4.1 (Dobin et al., 2013). The assemblies of *F. pratensis* cv. ‘Columbus’ and *L. multiflorum* cv. ‘Rabiosa’ were kindly provided by Dr. Christian S. Jensen (DLF Seeds A/S,
Denmark) and prof. Bruno Studer (ETH, Switzerland), respectively. Reference transcriptomes for *Festuca* and *Lolium* were then produced using StringTie ver. 2.1.0 (Pertea et al., 2015). To reduce redundancy within the StringTie transcriptomes, cd-hit ver. 4.6.1 (Fu et al., 2012) was used for clustering of transcripts, and only the longest isoform of each transcript was retained. Putative orthologous gene pairs used for subsequent analysis were identified by best reciprocal BLAST hits (BLAST+ ver. 2.10.1; e-value <1e-4) (Altschul et al., 1990) between protein sequences translated from *Festuca* and *Lolium* transcriptomes (e-value <1e-4).

For conventional analysis (using mapping reads to reference transcriptome from one parent), parental RNA-seq reads were pseudo-aligned separately to the reference transcriptomes of both parental species with kallisto ver. 0.44.0 (Bray et al., 2016). Abundances for uniquely assigned reads were loaded to the R environment ver. 3.4.4 (R Core Team, 2018). Differential gene expression analysis was assessed from generated read count data using the edgeR package (Robinson et al., 2010). The transcript abundance of each gene was estimated based on counts per million reads mapped (CPM) value. At first, only genes with abundance > 1 CPM in each biological replicate were considered for the analysis. Then, genes with a log2 fold change value > 1 and false discovery rate (FDR) < 0.05 were defined as differentially expressed. The results indicated that there was a statistically significant difference in the ratio of genes with homoeolog expression bias between two reference transcriptomes used for read mapping (hypergeometric test, p << 0.001; Figure S1A, B). Therefore, a k-mer based analysis approach was implemented to reduce the potential bias introduced by conventional analysis using read mapping to a respective reference transcriptome selected for the analysis.

Quality trimmed reads were used to create species-specific k-mer (k = 31) databases with KMC (K-mer Counter) ver. 3.0.0 (Kokot et al., 2017) and differing in one nucleotide with vmatch ver. 2.3.1 (Kurtz, 2003). Using bash, awk and vmatch, the database entries were filtered to: (1) only retain k-mers with an abundance of > 20 counts (sum of all three replicates) to filter out potential sequencing errors, (2) only retain k-mers with polymorphism in the central position of the 31-bp sequence (nucleotide 16) that discriminates between parents to minimize possible identification of false positive variants in highly polymorphic regions, (3) exclude k-mers having more than one polymorphism and (4) exclude heterozygosity for the discriminating nucleotide in one or both parents (by retaining only within-species homozygous k-mers). Subsequently, k-mer databases were aligned with BLASTN (word size = 15) against the reference transcriptomes to assign each
k-mer to its respective transcript, also recording the location on the transcript sequence. To avoid measuring expression values between paralogous genes, k-mers with a match to more than one transcript were removed. Because the k-mer data (31bp) was generated from the reads of 100 bp, a single read can yield multiple k-mers. Therefore, additional filtering steps were applied (see below).

First, the central position of each k-mer in a transcript was treated as a SNP that discriminated between parents. Using the biostar214299 utility of jvarkit (Lindenbaum, 2015), RNA-seq reads of both parents that mapped to their corresponding reference transcriptome and were spanning the k-mer derived SNP positions were extracted and parsed into four groups: (1) REF, where reads spanning 1 or more SNPs contain the same variant as the reference on the SNP position; (2) ALT, where reads spanning 1 or more SNPs contain the variant corresponding to the other parent; (3) AMBIGUOUS, where reads spanning 2 or more SNPs contain both the variant as the reference and the variant corresponding to the other parent; and (4) UNKNOWN, where reads spanning 1 or more SNPs contain variant(s) not corresponding to either parents. Only k-mers originating from reads in the REF group were retained. Counts of full-length k-mers were produced using kat (K-mer Analysis Toolkit) ver. 2.3.4 (Mapleson et al., 2017). Since k-mers were directly counted from RNA-seq data, reads containing only part of a k-mer sequence were not used. Thus only full-length k-mer sequence present in a read sequence and reference target were used in the subsequent analysis. The counts were aggregated and the mean count was calculated for each SNP cluster in case that more SNPs discriminating the parents were present in a single read. Then, the mean k-mer counts per read per transcript were summed to obtain a count representing the expression level per gene.

To validate the usefulness of our strategy, the k-mer based method we developed was compared with the conventional method of mapping reads to the reference transcriptome of one or the other parental species. Only SNPs from k-mers passing all filtering steps were used for this comparison. Once mapped to both transcriptomes, the parental RNA-seq reads spanning these SNPs were used to generate read count data for the conventional approach. The counts obtained from the two approaches were normalized to CPM and used to create a heatmap of Spearman’s correlation among parental gene expression counts generated by different approaches.
Expression Level Dominance (ELD) and Homoeolog Expression Bias (HEB) analysis

For the ELD analysis, total expression of a homoeologous gene pair in the hybrid was compared to the expression of both parents. Genes were then classified into 13 categories based on their expression level according to Rapp et al. (2009) and Yoo et al. (2013). To determine HEB, the expression levels of both homoeologs in hybrids were compared. All homoeologs showing average CPM <1 in hybrids were classified as silenced.

Cis and trans regulatory divergence analysis

The approach of McManus et al. (2010) was used to quantify cis and trans regulatory divergence. K-mer count-based expression levels of parents and hybrids were analyzed for evidence of differential expression using the binomial exact test followed by FDR analysis (FDR < 0.05). To moderate the significance of differential expression, fold change was not used. Three types of differential expression were evaluated for each gene: significant differential expression between homoeologs in parents (Lolium/Festuca; P), significant differential expression between the homoeologs in hybrids (Lolium/Festuca; H), and significant differential expression of each particular homoeolog between the parental (tested separately for each parent) and hybrid data (T). All genes were then assigned to one of seven regulatory categories according to McManus et al. (2010):

- **cis-only**: Significant differential expression in P and H. Not significant T.
- **trans-only**: Significant differential expression in P, but not H. Significant T.
- **cis + trans**: Significant differential expression in P and H. Significant T. The log2 fold change ratios of these genes have the same sign (P>0 and H>0, or P<0 and H<0), the regulatory divergence is in favor of expression of the same homoeolog.
- **cis × trans**: Significant differential expression in P and H. Significant T. The log2 fold change ratios of these genes have the opposite sign (P>0 and H<0, or P<0 and H>0), the regulatory divergence favors the expression of opposite homoeologs.
• Compensatory: Significant differential expression in H, but not P. Significant T. Regulatory divergences compensate each other, resulting in no expression difference between the two parental genomes.

• Conserved: No significant differential expression in P or H. Not significant T. Regulation of these genes is conserved.

• Ambiguous: All the other patterns of significance tests, which have no clear biological interpretation.

Gene ontology enrichment analysis

Genes were annotated using the gene ontology (GO) for Lolium perenne, downloaded from Monocots PLAZA 4.5 (https://bioinformatics.psb.ugent.be/plaza versions/plaza_v4_5_monocots/; based on reciprocal BLAST comparison with L. multiflorum. (Blanco-Pastior et al. 2021). The TopGO Bioconductor package in R was used for GO analysis. The Fisher’s test implemented in TopGO was used to identify enriched GO terms per comparison and all enriched GO terms with p < 0.01 were selected.

Data availability statement

RNA-seq data from this article can be found in the SRA data archive under accession number PRJNA685345.

Acknowledgements

This research was funded by the Czech Science Foundation (grant awards 20-10019S) and by the European Regional Development Fund OPVVV project “Plants as a tool for sustainable development” number CZ.02.1.01/0.0/0.0/16_019/0000827 supporting Excellent Research at CRH. We would like to express our thanks to Prof. Adam J. Lukaszewski for critical reading and valuable comments. The authors declare no conflict of interest.
Author contributions

Conceptualization MG, JB, JFW, TR, JD, BS, DK; Methodology MG, DC, JB, SS, TR, BS, DK; Formal analysis MG, DC, SS, MD; Investigation MG, DC, SS; Resources ZZ, DK; Data Curation MG; Writing – Original Draft MG, DK; Writing – Review and Editing TR, JB, JFW, JD, BS; Visualization MG, TR, DK; Supervision JB, JFW, BS, DK; Project administration DK, Funding acquisition JD, DK.

Conflict of interest

The authors declare no conflict of interest.

Supporting material legend

Table S1. Differential gene expression statistics.

Table S2. Statistics of species specific k-mer databases.

Table S3. ELD, HEB and cis/trans regulatory profiles of genes expressed in all analyzed hybrids.

Table S4. Expression level dominance (ELD) in reciprocal Festuca × Lolium hybrids. ELD scored in F1 generation of Festuca × Lolium (F1 F×L) and Lolium × Festuca (F1 L×F) hybrids, F2 generation of Festuca × Lolium (F2 F×L), F1 and F2 generations of Festuca × Lolium hybrids after four years of cultivation (F1 F×L aging and F2 F×L aging, respectively), and the overlaps of gene distribution in 13 categories. Roman numerals mark the categories as in Rapp et al. (2009), with schematic expression level graphs describing the differential gene expression profiles for
*Festuca* (F), hybrid (H) and *Lolium* (L). LED – *Lolium* expression dominance, FED – *Festuca* expression dominance.

**Table S5. Expression level dominance (ELD) after cold stress in *Festuca × Lolium* reciprocal hybrids.** ELD observed in F1 (F1 F×L) and F2 (F2 F×L) generations of *Festuca × Lolium* hybrids after cold stress, and the overlaps of gene distribution in ELD categories describing possible gene expression patterns in hybrids as compared to parents. Roman numerals mark 13 categories as in Rapp et al. (2009), with schematic expression level graphs describing the differential gene expression profiles for *Festuca* (F), hybrid (H) and *Lolium* (L). LED – *Lolium* expression dominance, FED – *Festuca* expression dominance.

**Table S6.** Numbers of genes in different cis/trans regulatory categories.

**Table S7.** Relationship between ELD and HEB in F1 and F2 generations of *Festuca × Lolium* hybrids after four years of cultivation (F1 F×L aging and F2 F×L aging, respectively), and these hybrids after cold stress treatment. The graphs at the top represent categories of ELD for *Lolium* expression dominance (II and XI) and *Festuca* expression dominance (IV and IX), as in Table 2.

**Table S8. Gene ontology (GO) analysis.** GO terms are provided for the genes with expression level dominance – *Lolium* expression dominance (categories II and XI) and *Festuca* expression dominance (categories IV and IX).

**Figure S1. Reference bias and correlation between conventional and k-mer based quantification of gene expression.** Differentially expressed genes (DEG) between parents when parental RNA-seq reads were mapped to the reference transcriptome of *Festuca* (a) and *Lolium* (b). Three categories of relative homoeolog expression were defined: L=F (not significantly differentially expressed), L>F (*Lolium* homoeolog is significantly more expressed than *Festuca* homoeolog) and F>L (*Festuca* homoeolog is significantly more expressed than *Lolium* homoeolog).
homoeolog). (c,d) Correlation matrix of read count data of parents generated by different approaches for each replicate: “k-mer”: DEG data generated by k-mer approach, “self reference”: DEG data generated by mapping to the reference transcriptome of the corresponding parent, “non-self reference”: DEG data generated by mapping to the reference transcriptome of the other parent. (c) correlation for three *F. pratensis* replicates; (d) correlation for three *L. multiflorum* replicates.

**Figure S2. Overlaps in gene sets showing expression level dominance.** Overlaps produced between the F1 generation of *Festuca × Lolium* (F1 F×L) and *Lolium × Festuca* (F1 L×F) hybrids, the F2 generation of *Festuca × Lolium* (F2 F×L), F1 and F2 generations of *Festuca × Lolium* hybrids after four years of cultivation (F1 F×L aging and F2 F×L aging, respectively) for *Festuca* (categories IV and IX) and *Lolium* (categories II and XI) expression dominance. FED – *Festuca* expression dominance, LED – *Lolium* expression dominance.

**References**

Adams, K.L., Cronn, R., Percifield, R., and Wendel, J.F. (2003). Genes duplicated by polyploidy show unequal contributions to the transcriptome and organ-specific reciprocal silencing. Proceedings of the National Academy of Sciences of the United States of America 100:4649-4654. 10.1073/pnas.0630618100.

Alger, E.I., and Edger, P.P. (2020). One subgenome to rule them all: underlying mechanisms of subgenome dominance. Current Opinion in Plant Biology 54:108-113. 10.1016/j.pbi.2020.03.004.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990). Basic local alignment search tool. Journal of Molecular Biology 215:403-410. 10.1006/jmbi.1990.9999.

Bao, Y., Hu, G.J., Grover, C.E., Conover, J., Yuan, D.J., and Wendel, J.F. (2019). Unraveling cis and trans regulatory evolution during cotton domestication. Nature Communications 1010.1038/s41467-019-13386-w.

Bell, G.D.M., Kane, N.C., Rieseberg, L.H., and Adams, K.L. (2013). RNA-Seq Analysis of Allele-Specific Expression, Hybrid Effects, and Regulatory Divergence in Hybrids Compared with Their Parents from Natural Populations. Genome Biology and Evolution 5:1309-1323. 10.1093/gbe/evt072.
Bird, K.A., VanBuren, R., Puzey, J.R., and Edger, P.P. (2018). The causes and consequences of subgenome dominance in hybrids and recent polyploids. New Phytologist 220:87-93. 10.1111/nph.15256.

Blanco-Pastor, J.L., Barre, P., Keep, T., Ledauphin, T. Escobar-Gutiérrez, A., Roschanski, A.M., Willner, E., Dehmer, K.J., Hegarty, M., Muylle, H., Veeckman, E., Vandepoele, K., Ruttink, T., Roldán-Ruiz, I., Manel, S., and Sampoux, J.P. (2021). Canonical correlations reveal adaptive loci and phenotypic responses to climate in perennial ryegrass. Molecular Ecology Resources 21:849-870. 10.1111/1755-0998.13289.

Boatwright, J.L., McIntyre, L.M., Morse, A.M., Chen, S.X., Yoo, M.J., Koh, J., Soltis, P.S., Soltis, D.E., and Barbazuk, W.B. (2018). A Robust Methodology for Assessing Differential Homeolog Contributions to the Transcriptomes of Allopolyploids. Genetics 210:883-894. 10.1534/genetics.118.301564.

Bottani, S., Zabet, N.R., Wendel, J.F., and Veitia, R.A. (2018). Gene Expression Dominance in Allopolyploids: Hypotheses and Models. Trends in Plant Science 23:393-402. 10.1016/j.tplants.2018.01.002.

Bray, N.L., Pimentel, H., Melsted, P., and Pachter, L. (2016). Near-optimal probabilistic RNA-seq quantification. Nature Biotechnology 34:525-527. 10.1038/nbt.3519.

Buggs, R.J.A., Doust, A.N., Tate, J.A., Koh, J., Soltis, K., Feltus, F.A., Paterson, A.H., Soltis, P.S., and Soltis, D.E. (2009). Gene loss and silencing in Tragopogon miscellus (Asteraceae): comparison of natural and synthetic allotetraploids. Heredity 103:73-81. 10.1038/hdy.2009.24.

Comai, L. (2005). The advantages and disadvantages of being polyploid. Nature Reviews Genetics 6:836-846. 10.1038/nrg1711.

Degner, J.F., Marioni, J.C., Pai, A.A., Pickrell, J.K., Nkadori, E., Gilad, A., and Pritchard, J.K. (2009). Effect of read-mapping biases on detecting allele-specific expression from RNA-sequencing data. Bioinformatics 25:3207-3212. 10.1093/bioinformatics/btp579.

Ding, M.Q., and Chen, Z.J. (2018). Epigenetic perspectives on the evolution and domestication of polyploid plant and crops. Current Opinion in Plant Biology 42:37-48. 10.1016/j.pbi.2018.02.003.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29:15-21. 10.1093/bioinformatics/bts635.

Douglas, G.M., Gos, G., Steige, K.A., Salcedo, A., Holm, K., Josephs, E.B., Arunkumar, R.,
Agren, J.A., Hazzouri, K.M., Wang, W., et al. (2015). Hybrid origins and the earliest stages of diploidization in the highly successful recent polyploid Capsella bursa-pastoris. Proceedings of the National Academy of Sciences of the United States of America 112:2806-2811. 10.1073/pnas.1412277112.

Doyle, J.J., and Sherman-Broyles, S. (2017). Double trouble: taxonomy and definitions of polyploidy. New Phytologist 213:487-493. 10.1111/nph.14276.

Edger, P.P., McKain, M.R., Bird, K.A., and VanBuren, R. (2018). Subgenome assignment in allopolyploids: challenges and future directions. Current Opinion in Plant Biology 42:76-80. 10.1016/j.pbi.2018.03.006.

Edger, P.P., Smith, R., McKain, M.R., Cooley, A.M., Vallejo-Marin, M., Yuan, Y.W., Bewick, A.J., Ji, L.X., Platts, A.E., Bowman, M.J., et al. (2017). Subgenome Dominance in an Interspecific Hybrid, Synthetic Allopolyploid, and a 140-Year-Old Naturally Established Neo-Allopolyploid Monkeyflower. Plant Cell 29:2150-2167. 10.1105/tpc.17.00010.

Escudero, M., and Wendel, J.F. (2020). The grand sweep of chromosomal evolution in angiosperms. New Phytologist 228:805-808. 10.1111/nph.16802.

Flagel, L., Udall, J., Nettleton, D., and Wendel, J. (2008). Duplicate gene expression in allopolyploid Gossypium reveals two temporally distinct phases of expression evolution. Bmc Biology 610.1186/1741-7007-6-16.

Freeling, M., Woodhouse, M.R., Subramaniam, S., Turco, G., Lisch, D., and Schnable, J.C. (2012). Fractionation mutagenesis and similar consequences of mechanisms removing dispensable or less-expressed DNA in plants. Current Opinion in Plant Biology 15:131-139. 10.1016/j.pbi.2012.01.015.

Fu, L.M., Niu, B.F., Zhu, Z.W., Wu, S.T., and Li, W.Z. (2012). CD-HIT: accelerated for clustering the next-generation sequencing data. Bioinformatics 28:3150-3152. 10.1093/bioinformatics/bts565.

Ghesquiere, M., Humphreys, M.W., and Zwierzykowski, Z. (2010). Festulolium. Fodder Crops and Amenity Grasses 5:293-316. 10.1007/978-1-4419-0760-8_12.

Glombik, M., Bacoovsky, V., Hobza, R., and Kopecky, D. (2020). Competition of Parental Genomes in Plant Hybrids. Frontiers in Plant Science 1110.3389/fpls.2020.00200.

Grover, C.E., Gallagher, J.P., Szkadkowski, E.P., Yoo, M.J., Flagel, L.E., and Wendel, J.F. (2012). Homoeolog expression bias and expression level dominance in allopolyploids. New Phytologist 196:966-971. 10.1111/j.1469-8137.2012.04365.x.
Hollister, J.D., and Gaut, B.S. (2009). Epigenetic silencing of transposable elements: A trade-off between reduced transposition and deleterious effects on neighboring gene expression. Genome Research 19:1419-1428. 10.1101/gr.091678.109.

Hu, G.J., and Wendel, J.F. (2019). Cis-trans controls and regulatory novelty accompanying allopolyploidization. New Phytologist 221:1691-1700. 10.1111/nph.15515.

Jiao, Y.N., Wickett, N.J., Ayyampalayam, S., Chanderbali, A.S., Landherr, L., Ralph, P.E., Tomsho, L.P., Hu, Y., Liang, H.Y., Soltis, P.S., et al. (2011). Ancestral polyploidy in seed plants and angiosperms. Nature 473:97-U113. 10.1038/nature09916.

Kokot, M., Dlugosz, M., and Deorowicz, S. (2017). KMC 3: counting and manipulating k-mer statistics. Bioinformatics 33:2759-2761. 10.1093/bioinformatics/btx304.

Kopecky, D., Loureiro, J., Zwierzykowski, Z., Ghesquiere, M., and Dolezel, J. (2006). Genome constitution and evolution in Lolium x Festuca hybrid cultivars (Festulolium). Theoretical and Applied Genetics 113:731-742. 10.1007/s00122-006-0341-z.

Kurtz, S. (2003). The Vmatch large scale sequence analysis software. Ref Type: Computer Program 412:297.

Landis, J.B., Soltis, D.E., Li, Z., Marx, H.E., Barker, M.S., Tank, D.C., and Soltis, P.S. (2018). Impact of whole-genome duplication events on diversification rates in angiosperms. American Journal of Botany 105:348-363. 10.1002/ajb2.1060.

Lemos, B., Araripe, L.O., Fontanillas, P., and Hartl, D.L. (2008). Dominance and the evolutionary accumulation of cis- and trans-effects on gene expression. Proceedings of the National Academy of Sciences of the United States of America 105:14471-14476. 10.1073/pnas.0805160105.

Li, M.D., Wang, R.H., Wu, X.M., and Wang, J.B. (2020). Homoeolog expression bias and expression level dominance (ELD) in four tissues of natural allotetraploid Brassica napus. Bmc Genomics 2110.1186/s12864-020-6747-1.

Lindenaubm, P. (2015). JVarkit: java-based utilities for Bioinformatics. figshare. http://dx.doi.org/10.6084/m9.figshare.1425030

Liu, Z., Yang, J., Xu, H., Li, C., Wang, Z., Li, Y., Dong, X., and Li, Y. (2014). Comparing computational methods for identification of allele-specific expression based on next generation sequencing data. Genetic Epidemiology 38:591-598. 10.1002/gepi.21846.

Lohaus, R., and Van de Peer, Y. (2016). Of dups and dinos: evolution at the K/Pg boundary. Current Opinion in Plant Biology 30:62-69. 10.1016/j.pbi.2016.01.006.
Madlung, A. (2013). Polyploidy and its effect on evolutionary success: old questions revisited with new tools. Heredity 110:99-104. 10.1038/hdy.2012.79.
Mapleson, D., Garcia Accinelli, G., Kettleborough, G., Wright, J., and Clavijo, B.J. (2017). KAT: a K-mer analysis toolkit to quality control NGS datasets and genome assemblies. Bioinformatics 33:574-576. 10.1093/bioinformatics/btw663.
Mason, A.S., and Wendel, J.F. (2020). Homoeologous Exchanges, Segmental Allopolyploidy, and Polyploid Genome Evolution. Frontiers in Genetics 1110.3389/fgene.2020.01014.
McClintock, B. (1984). The significance of responses of the genome to challenge. Science 226:792-801. 10.1126/science.15739260.
McManus, C.J., Coolon, J.D., Duff, M.O., Eipper-Mains, J., Graveley, B.R., and Wittkopp, P.J. (2010). Regulatory divergence in Drosophila revealed by mRNA-seq. Genome Research 20:816-825. 10.1101/gr.102491.109.
Ng, D.W.K., Lu, J., and Chen, Z.J. (2012). Big roles for small RNAs in polyploidy, hybrid vigor, and hybrid incompatibility. Current Opinion in Plant Biology 15:154-161. 10.1016/j.pbi.2012.01.007.
Nieto-Feliner, G.N., Casacuberta, J., and Wendel, J.F. (2020). Genomics of Evolutionary Novelty in Hybrids and Polyploids. Frontiers in Genetics 1110.3389/fgene.2020.00792.
Patro, R., Mount, S.M., and Kingsford, C. (2014). Sailfish enables alignment-free isoform quantification from RNA-seq reads using lightweight algorithms. Nature Biotechnology 32:462-U174. 10.1038/nbt.2862.
Pertea, M., Pertea, G.M., Antonescu, C.M., Chang, T.C., Mendell, J.T., and Salzberg, S.L. (2015). StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. Nature Biotechnology 33:290-+. 10.1038/nbt.3122.
Rahman, A., Hallgrímsdóttir, I., Eisen, M., and Pachter, L. (2018). Association mapping from sequencing reads using k-mers. Elife 7:e32920. 10.7554/eLife.32920.
Rapp, R.A., Udall, J.A., and Wendel, J.F. (2009). Genomic expression dominance in allopolyploids. Bmc Biology 710.1186/1741-7007-7-18.
Ren, L., Li, W.H., Qin, Q.B., Dai, H., Han, F.M., Xiao, J., Gao, X., Cui, J.L., Wu, C., Yan, X.J., et al. (2019). The subgenomes show asymmetric expression of alleles in hybrid lineages of Megalobrama amblycephala x Culter alburnus. Genome Research 29:1805-1815. 10.1101/gr.249805.119.
Renny-Byfield, S., and Wendel, J.F. (2014). Doubling down on genomes: Polyploidy and crop
Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26:139-140. 10.1093/bioinformatics/btp616.

Salman-Minkov, A., Sabath, N., and Mayrose, I. (2016). Whole-genome duplication as a key factor in crop domestication. Nature Plants 210.1038/nplants.2016.115.

Sotis, D.E., and Soltis, P.S. (1993). Molecular data and the dynamic nature of polyploidy. Critical Reviews in Plant Sciences 12:243-273. 10.1080/713608048.

Sotis, D.E., Visger, C.J., Marchant, D.B., and Soltis, P.S. (2016). Polyploidy: Pitfalls and paths to a paradigm. American Journal of Botany 103:1146-1166. 10.3732/ajb.1500501.

Soltis, D.E., and Soltis, P.S. (1993). Molecular data and the dynamic nature of polyploidy. Critical Reviews in Plant Sciences 12:243-273. 10.1080/713608048.

Sotis, D.E., Visger, C.J., Marchant, D.B., and Soltis, P.S. (2016). Polyploidy: Pitfalls and paths to a paradigm. American Journal of Botany 103:1146-1166. 10.3732/ajb.1500501.

Soltis, D.E., Mavrodiev, E.V., Meyers, S.C., Severns, P.M., Zhang, L.J., Gitzendanner, M.A., Ayers, T., Chester, M., and Soltis, P.S. (2012). Additional origins of Ownbey's Tragopogon mirus. Botanical Journal of the Linnean Society 169:297-311. 10.1111/j.1095-8339.2012.01244.x.

Soltis, P.S., and Soltis, D.E. (2016). Ancient WGD events as drivers of key innovations in angiosperms. Current Opinion in Plant Biology 30:159-165. 10.1016/j.pbi.2016.03.015.

Song, Q.X., and Chen, Z.J. (2015). Epigenetic and developmental regulation in plant polyploids. Current Opinion in Plant Biology 24:101-109. 10.1016/j.pbi.2015.02.007.

Takahagi, K., Inoue, K., Shimizu, M., Uchara-Yamaguchi, Y., Onda, Y., and Mochida, K. (2018). Homoeolog-specific activation of genes for heat acclimation in the allopolyploid grass Brachypodium hybridum. Gigascience 710.1093/gigascience/giy020.

Van De Peer, Y., Mizrachi, E., and Marchal, K. (2017). The evolutionary significance of polyploidy. Nature Reviews Genetics 18:411-424. 10.1038/nrg.2017.26.

VanBuren, R., Wai, C.M., Wang, X.W., Pardo, J., Yocca, A.E., Wang, H., Chaluvadi, S.R., Han, G.M., Bryant, D., Edger, P.P., et al. (2020). Exceptional subgenome stability and functional divergence in the allotetraploid Ethiopian cereal teff. Nature Communications 1110.1038/s41467-020-14724-z.
Voichek, Y., and Weigel, D. (2020). Identifying genetic variants underlying phenotypic variation in plants without complete genomes. Nature Genetics 52:534-+. 10.1038/s41588-020-0612-7.

W Wendel, J.F. (2015). The wondrous cycles of polyploidy in plants. American Journal of Botany 102:1753-1756. 10.3732/ajb.1500320.

W Wendel, J., Doyle, J.J. (2005). Polyploidy and evolution in plants. Plant diversity and evolution: genotypic and phenotypic variation in higher plants 97.

Williams, R.B.H., Chan, E.K.F., Cowley, M.J., and Little, P.F.R. (2007). The influence of genetic variation on gene expression. Genome Research 17:1707-1716. 10.1101/gr.6981507.

Wittkopp, P.J., Haerum, B.K., and Clark, A.G. (2008). Regulatory changes underlying expression differences within and between Drosophila species. Nature Genetics 40:346-350. 10.1038/ng.77.

Wu, J., Lin, L., Xu, M.L., Chen, P.P., Liu, D.X., Sun, Q.F., Ran, L.P., and Wang, Y.P. (2018). Homoeolog expression bias and expression level dominance in resynthesized allopolyplloid Brassica napus. Bmc Genomics 1910.1186/s12864-018-4966-5.

Yoo, M.J., Szadkowski, E., and Wendel, J.F. (2013). Homoeolog expression bias and expression level dominance in allopolyplloid cotton. Heredity 110:171-180. 10.1038/hdy.2012.94.

Yoo, M.J., Liu, X.X., Pires, J.C., Soltis, P.S., and Soltis, D.E. (2014). Nonadditive Gene Expression in Polyploids. Annual Review of Genetics, Vol 48 48:485-517. 10.1146/annurev-genet-120213-092159.

Zwierzykowski, Z., Kosmala, A., Zwierzykowska, E., Jones, N., Joks, W., and Bocianowski, J. (2006). Genome balance in six successive generations of the allotetraploid Festuca pratensis x Lolium perenne. Theoretical and Applied Genetics 113:539-547. 10.1007/s00122-006-0322-2.

Figure legends

Figure 1. Experimental design. The initial cross involved tetraploid L. multiflorum and tetraploid F. pratensis. Three and three independent reciprocal F1 hybrids were obtained and used in the study. One of these plants (Festuca × Lolium) was self-fertile and produced three plants of F2
generation without cross-pollination with other hybrids. All nine hybrid plants together with parents were used for the analysis of gene expression. Each plant was represented by three biological replicates for RNA-seq (tissues sampled from different ramets of the same plant). All plants were cytogenetically verified (representative karyotypes are given in the figure). WGD – Whole Genome Doubling.

Figure 2. Homoeologous expression bias (HEB) in reciprocal Festuca × Lolium hybrids. HEB in the F1 generation of Festuca × Lolium (F1 F×L) and Lolium × Festuca (F1 L×F) hybrids, the F2 generation of Festuca × Lolium (F2 F×L) and F1 and F2 generations of Festuca × Lolium hybrids after four years of cultivation (F1 F×L aging and F2 F×L aging, respectively). Homoeologous expression bias was computed for 5,790 genes. Three categories of relative homoeolog expression were defined: L=F (the expression of both homoeologs is not significantly different), L>F (Lolium homoeolog is significantly more expressed than Festuca homoeolog) and F>L (Festuca homoeolog is significantly more expressed than Lolium homoeolog).

Figure 3. Number of genes with homoeolog-specific silencing. Silencing of homoeologs observed in the F1 generation of Festuca × Lolium (F1 F×L) and Lolium × Festuca (F1 L×F) hybrids, the F2 generation of Festuca × Lolium (F2 F×L) and the F1 and F2 generations of Festuca × Lolium hybrids after four years of cultivation (F1 F×L aging and F2 F×L aging, respectively). Homoeologs were considered silenced when the average CPM value in all samples (three plants, each with three biological replicates) was < 1. L = Lolium multiflorum homoeolog; F = Festuca pratensis homoeolog.
Figure 4. Homoeologous expression bias (HEB) in reciprocal Festuca × Lolium hybrids after cold stress. Significant difference in homoeolog expression was computed for all 5790 genes and for the 468 genes differentially expressed after cold stress in both F1 and F2 generations. Three categories were defined: L=F (not significantly differentially expressed), L>F (Lolium homoeolog is significantly more expressed than Festuca homoeolog) and F>L (Festuca homoeolog is significantly more expressed than Lolium homoeolog).

Table legends

Table 1. Homoeolog expression bias in Festulolium hybrids. HEB analyzed in various F1 (F1 L×F, F1 F×L, F2 F×L, and F1 F×L (F1 F×L aging) and F2 (F2 F×L aging) hybrids after four years of cultivation) compared with the ratio of gene expression levels in the parents. L=F denotes equal expression; L>F and L<F denote Lolium-biased and Festuca-biased expression, respectively.

Table 2. Expression level dominance in Festulolium hybrids. ELD scored in F1 generation of Festuca × Lolium (F1 F×L) and Lolium × Festuca (F1 L×F) hybrids, F2 generation of Festuca × Lolium (F2 F×L), F1 and F2 generations of Festuca × Lolium hybrids after four years of cultivation (F1 F×L aging and F2 F×L aging, respectively), and the overlaps of gene distribution in the ELD categories. Roman numerals mark the ELD categories as in Rapp et al. (2009), with schematic expression level graphs describing the differential gene expression profiles for Festuca (F), hybrid (H) and Lolium (L).

Table 3. Expression level dominance after cold stress. ELD observed in F1 (F1 F×L) and F2 (F2 F×L) generations of Festuca × Lolium hybrids after cold stress, and the overlaps of gene distribution in ELD categories describing possible gene expression patterns in hybrids as compared to parents. Roman numerals mark the ELD categories as in Rapp et al. (2009), with schematic expression level graphs describing the differential gene expression profiles for Festuca (F), hybrid (H) and Lolium (L).

Table 4. Relationship between ELD and HEB. Relationships observed in F1 of Festuca × Lolium (F1 F×L) and Lolium × Festuca (F1 L×F) and F2 generation of Festuca × Lolium (F2 F×L) hybrids. The graphs at the top represent categories of ELD for Lolium expression dominance (II and XI) and Festuca expression dominance (IV and IX), as in table 2.
| Expression in parents | Expression in progeny | Phenomena          | F1 LxF | F1 FxL | F2 FxL | F1 FxL aging | F2 FxL aging |
|-----------------------|----------------------|--------------------|--------|--------|--------|--------------|--------------|
| L=F                   | L=F                  | Parental legacy    | 3969   | 4239   | 3927   | 3908         | 3961         |
|                       |                      |                    | (68.55 %) | (73.21 %) | (67.82 %) | (67.5 %)   | (68.41 %)   |
| L>F                   | L>F                  | Parental legacy    | 126    | 110    | 139    | 150          | 145          |
|                       |                      |                    | (2.18 %) | (1.9 %) | (2.4 %) | (2.59 %)     | (2.5 %)     |
| L<F                   | L<F                  | Parental legacy    | 130    | 94     | 187    | 121          | 177          |
|                       |                      |                    | (2.25 %) | (1.62 %) | (3.23 %) | (2.09 %)    | (3.06 %)    |
| L<F                   | L=F                  | Bias lost in hybrids | 416    | 461    | 360    | 428          | 366          |
|                       |                      |                    | (7.18 %) | (7.96 %) | (6.22 %) | (7.39 %)    | (6.32 %)    |
| L>F                   | L=F                  | Bias lost in hybrids | 491    | 524    | 487    | 475          | 479          |
|                       |                      |                    | (8.48 %) | (9.05 %) | (8.41 %) | (8.2 %)     | (8.27 %)    |
| L=F                   | L>F                  | Novel bias in hybrids | 242    | 162    | 281    | 318          | 276          |
|                       |                      |                    | (4.18 %) | (2.8 %) | (4.85 %) | (5.49 %)    | (4.77 %)    |
| L=F                   | L<F                  | Novel bias in hybrids | 377    | 187    | 380    | 362          | 351          |
|                       |                      |                    | (6.51 %) | (3.23 %) | (6.56 %) | (6.25 %)    | (6.06 %)    |
| L<F                   | L>F                  | Bias reverted in hybrids | 16     | 7      | 15     | 13           | 19           |
|                       |                      |                    | (0.28 %) | (0.12 %) | (0.26 %) | (0.22 %)    | (0.33 %)    |
| L>F                   | L<F                  | Bias reverted in hybrids | 23     | 6      | 14     | 15           | 16           |
|                       |                      |                    | (0.4 %) | (0.1 %) | (0.24 %) | (0.26 %)    | (0.28 %)    |
| Total number of genes |                      |                    | 5790   | 5790   | 5790   | 5790         | 5790         |
|                       |                      |                    | (100 %) | (100 %) | (100 %) | (100 %)     | (100 %)     |
| Category     | II   | XI   | IV   | IX   | Total assigned genes | Unassigned | Total analyzed genes |
|--------------|------|------|------|------|-----------------------|------------|----------------------|
| **F1 LxF**   | 234  | 200  | 155  | 186  | 4869                  | 921        | 5790                 |
|              |      |      |      |      | (100 %)               |            |                      |
|              | 434  |      |      |      | (8.9 %)               | -          | -                    |
|              |      |      |      |      |                       | -          | -                    |
|              |      |      |      |      |                       | -          | -                    |
| **F1 FxL**   | 256  | 233  | 146  | 170  | 4895                  | 895        | 5790                 |
|              |      |      |      |      | (100 %)               |            |                      |
|              | 489  |      |      |      | (10.1 %)              |            |                      |
|              |      |      |      |      |                       | -          | -                    |
|              |      |      |      |      |                       | -          | -                    |
| **F2 FxL**   | 255  | 239  | 135  | 148  | 4851                  | 939        | 5790                 |
|              |      |      |      |      | (100 %)               |            |                      |
|              | 494  |      |      |      | (10.1 %)              |            |                      |
|              |      |      |      |      |                       | -          | -                    |
|              |      |      |      |      |                       | -          | -                    |
| **F1 FxL aging** | 237 | 238  | 141  | 158  | 4862                  | 928        | 5790                 |
|              |      |      |      |      | (100 %)               |            |                      |
|              | 475  |      |      |      | (9.8 %)               |            |                      |
|              |      |      |      |      |                       | -          | -                    |
|              |      |      |      |      |                       | -          | -                    |
| **F2 FxL aging** | 289 | 207  | 165  | 123  | 4905                  | 885        | 5790                 |
|              |      |      |      |      | (100 %)               |            |                      |
|              | 496  |      |      |      | (10.1 %)              |            |                      |
|              |      |      |      |      |                       | -          | -                    |
|              |      |      |      |      |                       | -          | -                    |
| **Overlap F1, F2 FxL** | 140 | 136  | 65   | 70   | 2944                  | 328        | 5790                 |
|              |      |      |      |      | (60.3 %)              |            |                      |
|              | 276  |      |      |      | (5.7 %)               |            |                      |
|              |      |      |      |      |                       | -          | -                    |
|              |      |      |      |      |                       | -          | -                    |
| **Overlap F1 FxL, F1 LxF** | 193 | 167  | 114  | 140  | 3999                  | 527        | 5790                 |
|              |      |      |      |      | (82.3 %)              |            |                      |
|              | 360  |      |      |      | (7.4 %)               |            |                      |
|              |      |      |      |      |                       | -          | -                    |
|              |      |      |      |      |                       | -          | -                    |
| **Overlap F1, F2 FxL aging** | 191 | 144  | 90   | 86   | 3298                  | 297        | 5790                 |
|              |      |      |      |      | (67.5 %)              |            |                      |
|              | 335  |      |      |      | (6.9 %)               |            |                      |
|              |      |      |      |      |                       | -          | -                    |
|              |      |      |      |      |                       | -          | -                    |
| Category            | II  | IV  | IX  | XI  | Total assigned genes | Unassigned | Total analyzed genes |
|---------------------|-----|-----|-----|-----|-----------------------|------------|----------------------|
| F1 FxL stress       | 249 | 197 | 167 | 161 | 4890 (100 %)          | 900        | 5790                 |
|                     | 446 |     |     |     | 328 (7.7 %)           |            |                      |
| F2 FxL stress       | 256 | 199 | 161 | 156 | 4914 (100 %)          | 876        | 5790                 |
|                     | 455 |     |     |     | 317 (6.5 %)           |            |                      |
| F1 FxL stress - aging | 44  | 35  | 44  | 42  | 996 (100 %)           | 180        | 1176                 |
|                     | 79  |     |     |     | 86 (8.6 %)            |            |                      |
| F2 FxL stress - aging | 27  | 35  | 29  | 32  | 653 (100 %)           | 101        | 754                  |
|                     | 57  |     |     |     | 61 (9.3 %)            |            |                      |
| Overlap F1, F2 FxL stress | 206 | 163 | 131 | 137 | 4172 (85.1 %)         | 533        | 5790                 |
|                     | 369 |     |     |     | 263 (5.4 %)           |            |                      |
| ID in both F1 stress | 14  | 18  | 17  | 16  | 405 (100 %)           | 63         | 468                  |
|                     | 32  |     |     |     | 33 (8.2 %)            |            |                      |
| ID in both F2 stress | 13  | 22  | 18  | 17  | 411 (100 %)           | 57         | 468                  |
|                     | 35  |     |     |     | 35 (8.5 %)            |            |                      |
| Overlap ID in both F1, F2 FxL stress | 9   | 17  | 14  | 12  | 339 (83.1 %)          | 29         | 368                  |
|                     | 26  |     |     |     | 26 (6.5 %)            |            |                      |
| Homeolog regulation | II | IV | IX | XI |
|---------------------|----|----|----|----|
|                     | F1 | F2 | F1 | F1 | F1 | F1 | F2 | F1 | F2 |
| Population          | FxL | FxL | LxF | FxL | FxL | LxF | FxL | FxL | LxF |
| Both up-regulated   | 11 (7.3 %) | 21 (11.4 %) | 12 (6.6 %) | 3 (4.6 %) | 1 (1.5 %) | 4 (5.1 %) | 0 | 0 | 0 |
| F up-regulated      | 136 (81.9 %) | 155 (83.8 %) | 160 (87.9 %) | 16 (24.6 %) | 0 (0 %) | 10 (12.8 %) | 0 | 0 | 0 |
| L up-regulated      | 3 (2 %) | 5 (2.7 %) | 6 (3.3 %) | 46 (70.8 %) | 52 (76.5 %) | 60 (76.9 %) | 0 | 0 | 0 |
| F up- and L down-regulated | 0 | 3 (1.6 %) | 3 (1.6 %) | 0 | 0 | 2 (2.6 %) | 0 | 3 (2.6 %) | 6 (5.1 %) |
| Both down-regulated | 0 | 0 | 0 | 4 (4.6 %) | 13 (11.2 %) | 13 (11 %) | 7 (4.4 %) | 18 (9.7 %) | 15 (9.6 %) |
| F down-regulated    | 0 | 0 | 1 (0.5 %) | 0 | 0 | 0 | 4 (4.6 %) | 10 (8.6 %) | 11 (9.3 %) |
| L down-regulated    | 0 | 0 | 0 | 0 | 6 (8.8 %) | 1 (1.3 %) | 80 (90.9 %) | 90 (77.6 %) | 88 (74.6 %) |
| F down- and L up-regulated | 0 | 1 (0.5 %) | 0 | 0 | 0 | 1 (1.3 %) | 0 | 0 | 2 (1 %) |
| Not significant     | 106 | 70 | 52 | 81 | 67 | 77 | 82 | 32 | 68 |
| Total number of genes | 256 | 255 | 234 | 146 | 135 | 155 | 170 | 148 | 186 | 233 | 239 | 200 |
F. pratensis
2n=2x=14

L. multiforum
2n=2x=14

F. pratensis
2n=4x=28 female parent

L. multiforum
2n=4x=28 male parent

Interspecific hybridization

F1 FxL
2n=4x=28 (14+14)
(3 plants)

4-year cultivation
4-year cultivation +
cold stress

F1 FxL aging
2n=4x=28 (14+14)
(3 plants)

Self-fertilization of one F1 plant

F2 FxL
2n=4x=28 (14+14)
(3 plants)

4-year cultivation
4-year cultivation +
cold stress

F2 FxL aging
2n=4x=28 (14+14)
(3 plants)
|                | early after establishment | 4 year aging |
|----------------|---------------------------|--------------|
| **F1 LxF**     | 530 9.2%                  | 498 8.6%     |
| **F1 FxL**     | 287 5.0%                  | 544 9.4%     |
| **F2 FxL**     | 581 10.0%                 |              |
| **L>F**        |                           |              |
| 384 6.6%       | 279 4.8%                  | 481 8.3%     |
| 6.6%           | 4.8%                      | 7.6%         |
| **L=F**        |                           |              |
| 4876 84.2%     | 5224 90.2%                | 4811 83.1%   |
| 4876           | 5224                      | 4811         |
| 84.2%          | 90.2%                     | 83.1%        |
| **L=F**        |                           |              |
| 4774 82.5%     | 498 83.0%                 |              |
| 4774           | 498                       |
| 82.5%          | 83.0%                     |              |
|                  | early after establishment | 4 year aging |
|------------------|----------------------------|--------------|
|                 | F1 LxF                     | F1 FxL       |
| genes with      | 40                         | 7            |
| silenced        |                            |              |
| Lolium allele   |                            | 26           |
|                 |                            | 4            |
| genes with      |                            |              |
| silenced        |                            | 13           |
| Festuca allele  |                            | 2            |
|                 |                            | 19           |

This article is protected by copyright. All rights reserved.
| stress (468) | stress | L=F | L>F | F>L |
|--------------|--------|-----|-----|-----|
| F1 LxF       | 633    | 613 | 662 | 323 |
| F2 FxL       | 576    | 515 | 466 | 329 |

- **F>L**: 633 (10.9%), 613 (10.6%), 81 (17.3%), 78 (16.7%)
- **L>F**: 576 (9.9%), 515 (8.9%), 64 (13.7%), 61 (13.0%)
- **L=F**: 4581 (79.1%), 4662 (80.5%), 323 (69.0%), 329 (70.3%)

This article is protected by copyright. All rights reserved