Polyploidy and the Proteome

Author names: Douglas E. Soltis^{1, 2, 3, 4}, Biswapriya B. Misra², Shengchen Shan^{1, 4}, Sixue Chen^{2, 3, 4, 5}, Pamela S. Soltis^{1, 3, 4}

¹ Florida Museum of Natural History, University of Florida, Gainesville, FL 32611, USA

² Department of Biology, University of Florida, Gainesville, FL 32611, USA

³ Genetics Institute, University of Florida, Gainesville, FL 32608, USA

⁴ Plant Molecular and Cellular Biology Program, University of Florida, Gainesville, FL 32610, USA

⁵ Interdisciplinary Center for Biotechnology Research, University of Florida, Gainesville, FL 32610, USA

Corresponding Authors:

Douglas E. Soltis and Pamela S. Soltis

Florida Museum of Natural History

University of Florida

Gainesville, FL 32611, USA

E-mail: dsoltis@ufl.edu, psoltis@flmnh.ufl.edu

ABSTRACT

Although major advances have been made during the past 20 years in our understanding of the genetic and genomic consequences of polyploidy, our knowledge of polyploidy and the proteome is in its infancy. One of our goals is to stimulate additional study, particularly broad-scale proteomic analyses of polyploids and their progenitors. Although it may be too early to generalize regarding the extent to which transcriptomic data are predictive of the proteome of polyploids, it is clear that the transcriptome does not always reflect the proteome. Despite limited data, important observations on the proteomes of polyploids are emerging. In some cases, proteomic profiles show qualitatively and/or quantitatively non-additive patterns, and proteomic novelty has been observed. Allopolyploids generally combine the parental contributions, but there is evidence of parental dominance of one contributing genome in some allopolyploids. Autopolyploids are typically qualitatively identical to but quantitatively different from their parents. There is also evidence of parental legacy at the proteomic level. Proteomes clearly provide insights into the consequences of genomic merger and doubling beyond what is obtained from genomic and/or transcriptomic data. Translating proteomic changes in polyploids to differences in morphology and physiology remains the holy grail of polyploidy - this daunting task of linking genotype to proteome to phenotype should emerge as a focus of polyploidy research in the next decade.

Keywords: Polyploid; Whole-Genome Duplication; Genome; Transcriptome; Proteomics; Proteome features.

1. Introduction

Polyploidy or whole-genome duplication (WGD) is widely recognized as a major evolutionary force throughout eukaryotes. Both recent and ancient polyploidy events are well documented in many eukaryotic lineages, including green plants (reviewed in [1–4]), vertebrates (e.g., [5,6]), and fungi [7]. Approximately 35% of vascular plant species may have originated via polyploidy [8]. Many ferns are polyploid [8,9], and all living seed plants share an ancient WGD event [10]. All angiosperms are of ancient polyploid ancestry [10,11], and numerous ancient WGDs are found across the angiosperm tree of life (e.g., [1,12,13]).

Repeated cycles of polyploidy result in a plethora of duplicated genes, with up to two-thirds of the genes in a “diploid” plant genome derived from ancient WGD events [14]. These duplicate genes may play a major role in the short-term success of polyploid populations and species as well as the success of polyploid lineages. Although not all polyploid events result in increased net diversification (e.g., [15,16]), some ancient WGD events are clearly associated with major bursts in diversification (e.g., origin of the angiosperms, commelinid monocots, core eudicots, Brassicales, and Compositae [16]). Moreover, these clades are characterized by morphological and/or chemical innovations [17].

Studies of natural polyploids and synthetics have revealed the dynamic nature of newly formed polyploid genomes. At microevolutionary scales, polyploidy generates significant genetic, genomic, and phenotypic novelty (e.g., [18–26]), but the connection between genotype and phenotype is particularly difficult to ascertain in polyploids.

Despite recent results that help elucidate the dynamics of gene copy number and gene expression in polyploid plants, how this diversity at the gene level is manifested on the transcriptome, translome, proteome, metabolome, and ultimately the phenotype and fitness is essentially unknown. Given breakthroughs in the analysis of the proteome and the key role of the proteome in modulating metabolism, we focus on the proteome of polyploids, particularly from the perspective of shifts in gene content, alterations in gene expression, and their further downstream consequences on the proteome.

Two general types of polyploids are typically recognized: autopolyploids, historically defined as polyploids involving the multiplication of one chromosome set, and allopolyploids, those resulting from the merger of structurally different chromosome sets. Kihara and Ono [27] first used the terms *autopolyploidy* (auto = “same”) and *allopolyploidy* (allo = “different”) to distinguish between these two types, and many authors have followed their lead (e.g., [28–32]). Most workers have focused on species boundaries, with autopolyploids representing polyploidy within a single species and allopolyploids derived from interspecific hybridization and chromosome doubling – importantly, auto- and allopolyploidy represent ends of a continuum of genetic, phylogenetic, and chromosomal divergence of the progenitor(s) [31,33]. Below we compare and contrast current understanding of –omics data in both auto- and allopolyploids.

2. Genomics and transcriptomics of polyploidy

2.1. Fates of genes following polyploidy

Following WGD, duplicated genes have several possible fates: both copies may be retained, one copy may be retained and the other copy lost or silenced, or one copy may be retained while the other gene copy diverges to take on a new function (e.g., [34–39]). Progress has been made in understanding these processes, but more research is needed. Significantly, gene turnover and genome evolution across diverse lineages appear to be non-random (e.g., [40,41]). At deep levels, several comparisons have shown that some genes are routinely retained as duplicated copies, whereas others are consistently returned to singleton status (e.g., [40–42]). Within Compositae, patterns of gene retention and loss are consistent across ancient polyploidy events [43]; studies of recently formed allopolyploids in this family suggest the same patterns in these young polyploids as those from ancient events that date to 40 million years ago (mya; [44]).

Patterns of gene retention in an allopolyploid may indicate a bias so that one parental genome copy may be more highly retained than the other copy, referred to as “genomic dominance” (Fig. 1) (e.g., [39,45,46]). The fractionation patterns recovered for the ancient polyploids *Arabidopsis thaliana* [47], *Brassica* species [48], and *Zea mays* [49,50], as well as *Gossypium* (cotton) allopolyploids [46], indicate the presence of a “dominant” parental genome. Such “biased fractionation” [51] describes the retention of more genes from one parental genome than the other.

This biased fractionation of polyploid genomes may be related to differences in gene expression, with the more highly expressed parental subgenome dominating the retained fraction of the other parental subgenome in the polyploid [52,53]. In *Zea mays*, one parental subgenome retains 70-90% of the parental genes, whereas only 40-60% of

the other putative subgenome is retained [52]. Significantly, genes from the “high-loss” maize subgenome are expressed at lower levels than those on the “low-loss” subgenome [52]. The strong correlation between decreased levels of gene expression and the patterns of biased gene loss in *Zea mays* suggests that genes are more frequently lost from the less-expressed subgenome than from the more highly expressed subgenome [54]. However, it remains to be determined if this correlation reflects an actual causal relationship, as other factors, such as differential TE activity, might be responsible. A key question remains: does the interaction between levels of gene expression and biased patterns of gene loss explain the observed patterns of gene loss in the genomes of ancient polyploids? Linked to this is a similar question at the protein level (see below).

2.2. Dosage balance

Duplicate genes resulting from WGD may be preserved at higher frequencies than duplicates arising from tandem duplication due to “dosage-balanced” gene retention—that is, genes that are part of stoichiometrically sensitive pathways, where loss of a gene copy would disrupt the function of the pathway, are more likely to be retained in duplicate [55–57]. Evidence for dosage-balanced gene retention and loss has been found in many eukaryotes (e.g., [40,43,44,58]).

The role of processes such as positive dosage and neo- and subfunctionalization (Fig. 2; [59]) in gene retention, fractionation, and diploidization remains largely unknown. Some data suggest that the forces acting to retain duplicates after WGD are not uniform through time; instead, relative dosage may be a strong force for preservation

very soon after WGD, but of minor importance for later preservation of duplicates [60].

The earliest stages of polyploid formation are critical both for determining whether or not the polyploid will establish [18,24,25] and for shaping its future evolution [61]. However, linking work on newly synthesized polyploids with effects such as changes in gene expression and epigenetic modifications [21,22,44,62] and longer-term genome changes [63] is challenging. Fractionation and related processes of diploidization are major drivers of genome evolution (e.g., [53]), but it is unclear how soon after polyploid formation future patterns of fractionation are established and what mechanisms are responsible.

2.3. Gene silencing and gene loss

Gene silencing plays an important role in regulating duplicate gene expression [21,34,39,64–67] and may be accomplished via genetic and/or epigenetic mechanisms. Gene silencing may occur through the accumulation of genetic changes in coding or regulatory regions [34,35,66,68] or via epigenetic mechanisms such as methylation. Apparent silencing may also occur through physical loss of one member of a duplicate gene pair, perhaps via translocations or non-reciprocal crossing-over [21,69,70]. Both gene silencing and homeolog loss are well documented in polyploids (e.g., [39,65,68]).

Studies of recently formed natural polyploids and synthetic polyploids have also shown that gene silencing and loss can occur rapidly. Rapid gene loss, and the elimination of noncoding regions, occurred as early as the F₁ hybrid or the first amphiploid generation in synthetic wheat [71–74]; studies of synthetic polyploids in

Brassica napus (rapeseed, canola) also indicate rapid changes (e.g., [21]). In recently and repeatedly formed allotetraploids in *Tragopogon* (no more than 80 years, 40 generations old), both gene silencing and loss occur, but gene loss predominates. Furthermore, patterns of gene loss and silencing in these young polyploids have been repeated across polyploid populations of independent origin, as well as synthetics (e.g., [22,25,75]). Similar results are suggested for the naturally occurring recent allopolyploid *Senecio cambrensis* [24].

The regulation of gene expression via epigenetic mechanisms may result from various non-genetic changes, such as chromatin remodeling and methylation (e.g. [76,77]). Methylation has been documented as accompanying changes in gene expression in several polyploids (e.g., [18,78,79]). Epigenetic gene silencing is well documented in *Arabidopsis* polyploids, including synthetics (e.g., [63, 79–81]). For example, ~2.5% of the transcriptome of the allotetraploid *A. suecica* was silenced due to epigenetic DNA methylation [80].

One of the best examples of epigenetic gene silencing is nucleolar dominance, in which one set of parental ribosomal RNA (rRNA) genes is silenced in a hybrid or an allopolyploid [82–86]. In *A. suecica*, the rRNA genes of one parent (*A. thaliana*) were always silenced, with at least two generations required for complete silencing [86]. Similarly, allopolyploid *B. napus* has active rRNA genes from just one of its parents in vegetative tissues. In contrast, both parental copies are present in floral tissue, which indicates that the epigenetic changes are developmentally regulated [87].

2.4. Neofunctionalization and subfunctionalization

Genetic and/or regulatory changes may lead to functional diversification of the homoeologs, producing genes of new function (neofunctionalization) or with different copies expressed in different tissues or at different times (subfunctionalization) ([34,35,66,88–90]; reviewed in [38]; see Fig. 2). Functional diversification of duplicated genes can be difficult to document even in diploids; Prince and Pickett [66] have therefore predicted that this process will be rare. Neofunctionalization is not well understood and is therefore not well documented in polyploids. However, functional diversification of duplicate genes has been reported in maize, an ancient polyploid [65].

Organ-specific reciprocal silencing of duplicate genes – i.e. expression of one parental homeolog in ovules and of the other homeolog in other tissues – in natural and synthetic polyploids of *Gossypium* provides evidence for subfunctionalization [89]. Subfunctionalization was also documented for duplicate genes, perhaps resulting from ancient polyploidy, in *A. thaliana* [91]. Roulin et al. [92] examined ~18,000 duplicated genes across seven tissues in *Glycine max* (soybean), which has experienced two rounds of polyploidy, and found that ~50% of the paralogs showed subfunctionalization.

The cumulative effects of gene silencing and neo- and subfunctionalization in polyploids yield a transcriptome that is smaller than predicted based on the gene content of the polyploid genome. Thus, polyploids, in which genome downsizing [93] is well documented, may be expected to exhibit even greater downsizing of the transcriptome. In allopolyploid *Glycine dolichocarpa*, transcriptome downsizing is substantially greater than genome downsizing [78]. These processes that shape the transcriptome – is

downsizing random or directed with regard to the parental contributions? – thereby constrain the proteome to a subset of the possible outcomes that could not be predicted from the merger of two divergent progenitor genomes in an allopolyploid.

Novel patterns of gene expression may also occur in polyploids, with these shifts in gene expression possibly linked to novel phenotypes through changes in gene expression, including dosage effects and chromatin remodeling [94]. In comparisons between polyploids and their diploid progenitors, most of the novel transcripts appear to be transposons [64,74,95]. Novelty in gene expression, coupled with alternative splicing, which generates transcript variability, may provide a pool of transcripts that could yield novel proteins, perhaps through unique combinations of subunits or domains, possibly offsetting some of the constraints imposed on the proteome through the combined effects of homeolog loss, gene silencing, and transcriptome downsizing. Because, in a polyploid genome, some duplicate loci will exhibit gene loss, others gene silencing, and still others neo- or subfunctionalization, a polyploid individual will likely be a mosaic of different possible outcomes at the gene expression level. Thus, a polyploid species is expected to comprise an array of genetically, transcriptomically, and proteomically unique individuals.

3. Proteomics of polyploids

3.1. Proteomics technologies used in studies of polyploid plants

Understanding protein-level changes and regulation in polyploids will greatly enhance our knowledge of plant adaptation and function. Here, we briefly review the

techniques used in proteomic studies. Approaches based on two-dimensional gel electrophoresis (2-DE) and liquid chromatography (LC) are commonly used in modern proteomics [96].

Early work on *Brassica* proteomics used 2-DE to analyze the response of proteomes to allopolyploidy and autopolyploidy [97–99]. Use of differential in-gel electrophoresis (DIGE) allows comparison of proteins from two different samples in one gel [96] and was used in several polyploid proteomics experiments [100–102]. LC-based proteomics approaches such as isobaric tag for relative and absolute quantitation (iTRAQ) have gained popularity because of their high coverage, high throughput, and multiplexing capabilities [103–108]. While the outdated 2-DE and MALDI-based approaches were used to find only very abundant proteins, iTRAQ and LC-MS/MS technologies provide a more genome-wide assessment of proteomic differences. With advances in mass spectrometry, software, databases, and interpretation, we now have more tools to address the finer details of the information content of polyploid –omes.

Table 1 lists the proteomics technologies used, as well as the proteomic studies conducted, on polyploid plants. High-throughput proteomic studies have aided in quantitative profiling, analysis of dynamic post-translational modifications (PTMs), subcellular localization and compartmentalization, protein complexes, signaling pathways, and protein–protein interactions [109]. These advances contribute to the analysis of advanced mapping populations such as hybrid doubled haploid lines, near-isogenic lines, and recombinant inbred lines, which further verify the correlations between responsive quantitative trait loci and stress-tolerant phenotypes [109].

In contrast to transcriptome studies, coverage of the proteome is still far from complete due to the large diversity of proteoforms, their dynamics, and physicochemical properties. Quantitative methods are not mature; iTRAQ and Tandem Mass Tag® (TMT) suffer from diverse technical challenges such as ratio compression, while label-free proteomics requires more replicates and coverage. Data-independent acquisition (DIA)-based MS technology and parallel-reaction monitoring (PRM)-based targeted proteomics approaches are the future, which require concomitant establishment of a robust protein spectral database. These technical challenges have impeded comparable coverages at the RNA and protein levels.

3.2. Additive and non-additive protein expression in allopolyploids

Allopolyploidy can result in gene expression changes relative to the diploid parents that are organ-specific; these shifts may have effects at the protein level (e.g., [98,103,107,110,111]). Allopolyploidy can also result in rapid and reproducible genomic changes in the first generation of nascent polyploids, including elimination of DNA sequences, gene silencing, alteration of cytosine methylation, and activation of genes [112]. In several synthetic allopolyploids, a bias toward one genome has been reported [113–115], suggesting that genome changes might be influenced by nucleo-cytoplasmic interactions. However, in synthetic *Brassica napus*, the predominant pattern of gene expression was of paternal origin, and there was no bias in the cellular localization of non-additive proteins. Depending on the polyploid synthesis and on the level of analysis (DNA, transcript, protein), modifications range from what appears random to what has

been termed nearly deterministic [98]. Epigenetic changes such as histone modifications, chromatin remodeling, and DNA methylation are described in newly formed allopolyploids [98].

The combination of two divergent parental genomes in an allotetraploid may induce genome-wide non-additive gene regulation. In contrast, autopolyploidy may result in only small effects on gene regulation [116]. Furthermore, variation in the parental proteome may affect the differences observed in the polyploid proteome (e.g., [103,107]). Synthesized polyploids provide model systems because the exact progenitors of the synthetic polyploids are known. Proteins in synthetic polyploids can be compared with the presence and level of expression of these proteins in the progenitors; for example, the stress-related proteins were significantly enriched in up-regulated proteins as compared to the down-regulated proteins in *Brassica* hexaploids (*B. carinata* (4x) x *B. rapa* (2x)) [107].

Recently developed platforms are enabling new insights into the proteomes of polyploids, including *Triticum* (wheat; [117,118]), *B. napus* [98], *Gossypium* (cotton; [111]), *Tragopogon* [103], *A. suecica* [119], and others. General additivity of the parental profiles in the proteomes of several allopolyploids (*Triticum*, *Gossypium*, *Tragopogon*, *Arabidopsis*) [103,111,117,119] has been reported, but non-additive patterns occur for at least some proteins in some species (e.g., wheat [117] and cotton [111,118]) and in synthetic *B. napus* [98].

Some studies have revealed that the proteome of the polyploid does not always match the results predicted from the transcriptome alone. Furthermore, novel proteins not

found in either parent may be produced (e.g., [26,103]). Other studies have revealed imbalanced levels of expression of the parental contributions in polyploids. The genetic redundancy within an allopolyploid genome could explain why functionally related proteins could display imbalanced levels of expression. Moreover, at the protein level, there is no evidence for the “genomic/transcriptomic shock” [120] observed in several recent polyploids (e.g., *Tragopogon*, *Spartina*, *Senecio*, *Brassica*; [18,22,24,121,122]). We summarize below published studies (see Table 1), noting that the goals varied from evolutionary to breeding. We begin with studies having broad proteomic evolutionary implications (e.g., additivity, novelty, parental dominance) and then end with studies that have breeding implications.

In resynthesized allotetraploids of *B. napus*, Marmagne et al. ([123]; see also [98,99]) compared protein and transcript accumulation. They observed that two-thirds of the genes had additive transcript levels and concluded that most of the differential protein regulation was not explained by transcriptional changes [123]. Instead, differential protein regulation is mainly governed by post-transcriptional modifications. Comparative proteomics of synthetic allopolyploids of *B. napus* and its diploid progenitors *B. rapa* and *B. oleracea* revealed hundreds of polypeptides that displayed additivity (i.e., mid-parent values) in the allopolyploids, while many of the proteins showed non-additivity [99]. Most of the protein isoforms (~55%) displayed “different” or “opposite” patterns of regulation in the polyploids, i.e., isoforms of the same protein showing both increase and decrease in the synthetic *B. napus* compared to the mid-parent value. A similar pattern was also seen in components of protein complexes in the allotetraploids [99].

iTRAQ LC–MS/MS analysis of the leaf proteomes of *B. napus* plants revealed 1976 proteins expressed during drought stress [124]. A total of 417 proteins showed significant changes in abundance, while 181 proteins underwent PTMs, and 49 were differentially expressed in drought-stressed plants. For phosphorylated or oxidized proteins, no correlation was observed between gene transcription and protein levels, highlighting that the PTM changes were not regulated at the transcription level. Comparative proteomic analysis using 2D gels of seedling leaves exposed to salt stress identified 44 protein spots that were differentially accumulated upon treatment [125]. Further large-scale analysis of the leaf proteome of allotetraploid *B. napus* compared with its diploid progenitors, *B. rapa* and *B. oleracea*, revealed that the tetraploid generally combined parental contributions, but many proteins expressed in the polyploid exhibited non-additive expression patterns [126]. Furthermore, leaf protein levels were extensively changed in synthetic *B. napus* compared to the parents, but the “housekeeping” proteins appeared to be unaltered. Kong et al. [126] detected proteomic changes over four generations of synthetics, attributed to gene silencing.

Proteomic changes in leaves of a *Brassica* synthetic hexaploid relative to its parents (*B. carinata*, 4x, and *B. rapa*, 2x) were investigated using iTRAQ LC-MS/MS, revealing that 452 proteins were differentially expressed between the hexaploid and its parents, and 100 proteins were non-additively expressed in the hexaploid, suggesting non-additive protein regulation following genomic merger and doubling [107]. Functional enrichment of proteins showed that proteins involved in the immune system and in response to stimuli were significantly enriched in the non-additive proteins,

providing a possible driving force for variation and adaptation in allopolyploids. These differentially expressed proteins showed expression-level dominance of the tetraploid progenitor of the hexaploid relative to the diploid progenitor.

The recently formed natural allopolyploid *Tragopogon mirus*, its diploid parents (*T. dubius* and *T. porrifolius*), a diploid interspecific F₁ hybrid, and synthetic *T. mirus* were analyzed using iTRAQ LC-MS/MS technology [103]. Of the 476 proteins identified, 408 displayed quantitative additivity of the parental profiles in *T. mirus*, while the remaining 68 proteins were non-additive: 32 protein expression changes were associated with hybridization and 22 with genome doubling, while 14 shifts occurred since the origin of *T. mirus* ca. 80 years ago. Included in these non-additive protein patterns are 6 proteins with novel expression in the F₁ and allopolyploid. In *T. mirus*, the impact of hybridization on the proteome is more important than polyploidization [103].

Proteomic analyses were conducted using leaves of both synthetic *Arabidopsis* autopolyploids and allotetraploids and their progenitors using iTRAQ LC-MS/MS [119]. In the allotetraploids derived from *A. thaliana* and *A. arenosa*, 97 (8.4%) and 95 (8.2%) of the proteins in F₁ and F₈ allotetraploids, respectively, showed non-additive patterns of abundance relative to the mid-parent values of the parents. These proteomic results were highly comparable to findings from a gene expression study based on microarrays [116] in that non-additively expressed proteins also showed transcript-level differences. Ng et al. [119] concluded that polyploidization may lead to high protein divergence between a polyploid and its parents as a result of both post-transcriptional and PTMs of proteins. We stress, however, that many non-additively accumulated proteins (61–62%) in the F₁

and F₈ allotetraploids were also differentially expressed in the parents [119], which may be indicative of parental legacy (see further discussion below).

A comparative analysis of cotton seed proteomes from the allopolyploid *G. hirsutum* (AD genome) and its A- and D-genome diploid progenitors indicated high degrees of divergence among the three proteomes; one-third of all protein forms were genome-specific [111]. Furthermore, higher proteomic similarity was seen between the allopolyploid and its D-genome donor than its A-genome donor, indicating unequal expression of proteins from diploid parental genomes in allopolyploids. Recently, using iTRAQ-LC-MS/MS, 5884 proteins were identified in cotton ovules [127]. Moreover, the expression of proteins involved in carboxylic acid metabolism, small-molecule metabolic processes, hormone regulation, and lipid metabolism was significantly enhanced in the ovules of wild *G. hirsutum*, with 2729 proteins preferentially accumulating at anthesis. Furthermore, this study integrated proteomics evidence with transcriptomic data sets to reveal that 2005 accumulated proteins correlated well with the RNA-Seq data sets, while the levels of 724 proteins did not correlate with their mRNA levels, implying that post-transcriptional or translational regulation was involved.

Similarly, the leaf soluble proteins from two citrus allotetraploids (mandarin + lime; mandarin + kumquat) and their diploid parents (mandarin, lime, and kumquat; *Citrus deliciosa*, *C. aurantifolia*, and *Fortunella margarita*, respectively) were analyzed using 2-DE. The allotetraploid protein patterns were closer to the mandarin parent than to the non-mandarin parents both qualitatively and quantitatively [128].

Cultivated banana varieties are triploids that originated from intra- and

interspecific hybridizations between subspecies of *Musa acuminata* (A-genome donor) and *M. balbisiana* (B-genome donor). Comparisons using 2-DE and 2D-DIGE with *de novo* MS/MS sequencing of the triploid varieties (AAA, AAB, ABB, and BBB) and their wild diploid parents indicated that the proteomic phenotype observed for allopolyploid banana does not always correspond to the expectations from the genomic standpoint [100], perhaps due to either gene deletion or gene silencing [129]. In what appears to be one of the first investigations of the role of paralogs at the protein level, Vanhove et al. [130] used the 2D-DIGE approach to examine the peptide for each paralog, finding that different paralogs have apparently evolved over time and do not behave in the same way when subjected to stress.

Proteins of etiolated seedlings from a synthetic allopolyploid line of wheat (*T. turgidum* sp. $\times$ *T. tauschii*) and its parents were compared using 2-DE; the patterns from the polyploids closely resembled a combination of their parents, except that one novel protein was observed. For most gene products investigated, the parental genomes acted independently of each other in the synthetic polyploids [117]. Islam et al. [118] concluded that the expression of proteins in hexaploid wheat was the product of complex interactions among the diploid genomes. For example, omega gliadin is currently controlled by the B-genome in hexaploid wheat, but was an initial product of interaction between the A- and B-genomes. In addition, some novel protein spots were detected, but the mechanism(s) generating them is unclear. These could indeed be novel proteins, or, alternatively, the actual donor genome of the hexaploid was not sampled at the diploid or tetraploid levels. Recently, using label-free gel-based proteomics approaches, the nuclear

proteomes of the grains of the two cultivated wheat species *T. aestivum* (hexaploid wheat; genomes A, B, and D) and *T. monococcum* (diploid wheat; genome A) were compared, with the identification of 528 proteins, of which 368 were common to both species [131].

An octoploid salt-tolerant amphiploid genotype was previously generated by combining hexaploid bread wheat variety Chinese Spring with the diploid wild wheatgrass, *Lophopyrum elongatum*, which is adapted to salt marshes. Analysis using a proteomic workflow of 2D-DIGE and MALDI-TOF/TOF MS to compare the protein profiles of isolated mitochondrial samples from roots and shoots of both genotypes after control or salt treatment indicated a clear enrichment of enzymes involved in detoxification of reactive oxygen species in both genotypes [101]. Another 2D-DIGE study involving cold treatment (4 °C) in crowns of the frost-tolerant winter wheat cultivar Samanta and the frost-sensitive spring wheat cultivar Sandra reported 386 differentially abundant protein spots [102]. Fifty-eight protein spots selected for MALDI-TOF/TOF identification resulted in identification of 36 proteins that showed an increased relative abundance upon cold treatment. These included proteins involved in carbohydrate catabolism (glycolysis enzymes), redox metabolism (thioredoxin-dependent peroxidase), chaperones, and defense-related proteins (including a protein similar to thaumatin). Proteins exhibiting a cold-induced increase in expression in the winter wheat cultivar included those involved in regulation of stress response and development (germin E, lectin VER2), while proteins showing a cold-induced increase in expression in the spring wheat cultivar included those involved in restoration of cell division and plant growth (eIF5A2, glycine-rich RNA-binding protein, adenine phosphoribosyltransferase). Thus,

the proteomes of spring and winter wheat respond differently to cold acclimation.

In allohexaploid *Camelina sativa*, a new Brassicaceae oil crop, study of the seed proteome of plants overexpressing the *Arabidopsis* G-protein γ subunit 3 (AGG3) led to the identification of 1532 proteins [108]. These differentially expressed seed proteins were involved in primary and secondary metabolism, nucleic acid and amino acid metabolism, and abscisic acid-related responses, corroborating the physiological effects of AGG3 overexpression in providing protection against oxidative and heavy metal stress conditions. Recent investigations of the proteome of the latex of the allotetraploid rubber tree (*Hevea brasiliensis*) [132,133] found that the proteomes of rubber trees susceptible to or tolerant of *Phytophthora* sp. had comparable protein counts [132]. The differentially abundant proteins from latex of the resistant strains indicated enrichment of defense-related proteins, e.g., beta-1,3-glucanase, chitinase, and lectin; these results were also supported by SDS-PAGE and 2-D Western blot analysis. These efforts in proteomics have benefited from the availability of a reference rubber tree genome [134], as is the case for many such proteomics studies. Nonetheless, previously, post-germination changes in seed proteomes [135], rubber particles [136,137], and proteins in subcellular fractions of latex [138] were also captured using mostly 2D-PAGE methods.

3.3. Proteomics of autopolyploids

Autopolyploids have long been overlooked in studies of the genetic consequences of polyploidy (e.g., [26]). It is noteworthy, therefore, that several proteomic investigations have been conducted on autopolyploids. Perhaps not surprisingly, most

analyses have shown few or no differences between diploids and autotetraploids, while a few have revealed clear proteomic differences between autopolyploids and parents.

In haploid, diploid, and autotetraploid *Brassica oleracea*, no clear proteomic differences were observed among ploidal levels. Inter-organ (such as leaf vs. stem) and genotype differences were the only sources of protein-level changes [97]. Similarly, very few proteomic differences were observed between diploid and autotetraploid *A. thaliana* (~6.8% differences; [119]). However, functional categories of abiotic and biotic stresses were significantly enriched in the *A. thaliana* autotetraploid compared to the diploid [119]. Genome doubling had only a small effect on gene regulation in autotetraploid *A. thaliana* and did not induce genome-wide non-additive regulation [110]. In *Musa*, *M. acuminata* (A-genome donor to cultivated bananas; see above) and triploid variety AAA have nearly identical proteomic patterns [100].

Some autotetraploids have proteomic profiles distinct from those of their diploid parents. The leaf proteome of a diploid cultivar of cassava (*Manihot esculenta*) was clearly differentiated from its autotetraploid derivative [139]. Significant differences in photosynthetic activity, cyanide content, and resistance to salt stress between diploid and autotetraploid plants were closely linked to their proteomic profiles. Differences between diploids and autotetraploids are particularly evident for stress response proteins, as noted above for *A. thaliana* [119]. In proteomic comparisons between diploid and autotetraploid mandarins, genetic factors have a predominant effect on salt response, although a common stress response independent of genotype was also observed, indicating selection of a specific set of proteins in the polyploidization process [140].

Likewise, differences were observed in leaf proteins of diploid and autotetraploid black locust (*Robinia pseudoacacia* L.) under salt stress [141]. Differentially expressed proteins between the tetraploid and diploid are involved in photosynthesis, stress and defense, energy, metabolism, transcription/translation, and transport. Diploid plants of *R. pseudoacacia* experienced greater negative effects than those of tetraploid plants under the stress conditions employed [141]. In autopolyploid sugarcane [142], a label-free quantitative proteomics approach revealed several differentially expressed proteins between control and salt-treated plants, but there was no comparison across ploidal levels [143].

3.4 Homeolog-specific protein expression

Homeolog loss and silencing have been commonly documented in allopolyploids based on genomic and transcriptomic studies (as reviewed above). Only a few studies pertinent to this topic have been conducted, however, at the proteomic level, but they provide evidence for homeolog-specific expression. In upland cotton (*G. hirsutum* L.), based on two complementary proteomic approaches, fiber proteomes were examined for four representative stages of fiber development [105]. These studies revealed unequal contributions of homeologous proteins for over one-third of the fiber proteome, but the overall expression was balanced with respect to the genome of origin [105]. Furthermore, about 30% of the proteins were differentially expressed during fiber development in both wild and domesticated cotton. In addition, the expression levels were altered by domestication.

Comparing the naturally occurring allopolyploid *T. mirus* and its diploid parents (*T. dubius*, *T. porrifolius*), as well as a diploid F₁ hybrid and synthetic *T. mirus*, Koh et al. [103] noted homeolog-specific protein expression. For example, trimeric dihydrolipoamide succinyltransferase and an unidentified conserved protein have nine and two amino acid substitutions between *T. dubius* and *T. porrifolius*, respectively, and the expressed proteins in *T. mirus* come from only the *T. dubius* proteome. Moreover, these cases both result from silencing of the *T. porrifolius* gene so that no transcript from *T. porrifolius* is produced. The paternal *T. dubius* homeologs were more often up-regulated than the *T. pratensis* homeologs (77% vs. 23%, respectively). In addition, the up-regulation of the *T. dubius*-derived fructose-bisphosphate aldolase is a result of the absence of the *T. porrifolius*-derived proteins following silencing of the *T. porrifolius* homeolog in the synthetic polyploid [144]. These two cases of homeolog-specific protein expression in *T. mirus* suggest that silencing was not associated with hybridization, but occurred subsequent to both hybridization and polyploidization—as a result of subsequent evolution post-polyploidization. The diploid F₁ hybrid and synthetic polyploid have both parental homeologs in their genomic DNA, and both parental forms are also present at the transcript and protein levels, but subsequent changes were observed in the natural polyploid plants.

Homeolog-specific expression favoring the same parent at multiple protein-coding genes will yield expression-level dominance of one parent over the other. Expression-level dominance is frequently observed in the transcriptome, as reviewed in Yoo et al. [39]. However, because proteomic studies of polyploids are still in their

infancy, it is unclear if these results from the transcriptome correspond to the proteome (see below). However, several examples of expression-level dominance at the proteomic stage are apparent. For example, the synthetic *Brassica* hexaploid noted above showed expression-level dominance toward its tetraploid progenitor (*B. carinata*) [107].

3.5. Protein post-translational modifications (PTMs)

Protein PTMs play important roles in many processes of biological function, including signal transduction and metabolism, and PTM analysis has become an important area of modern proteomics. Only one study has reported protein PTMs that have led to novel expression in polyploids: synthetic and natural polyploids of *Tragopogon* exhibit six such PTMs [103]. Since the PTMs also occurred in the diploid F₁ hybrid, the expression changes were caused by hybridization rather than polyploidization. Further research is needed to clarify the detailed physiological and evolutionary significance of protein PTMs in polyploids.

Histone acetylation/deacetylation and methylation play important roles in regulating gene expression in polyploids and their parents. In *Arabidopsis* polyploids, histone modifications were linked to phenotypic changes in flowering time [145] and chlorophyll and sugar contents [146]. In *Brassica* species, the levels of histone methylation differed between the polyploids (*B. juncea* and *B. carinata*) and diploid (*B. nigra*) [147]. The role of PTMs in plant stress modulation is further evidenced by the fact that out of 49 differential PTMs captured, 46 were oxidations, and three were phosphorylations in a drought-responsive leaf proteome [124]. Thus, PTM discovery in

polyploid species and assignment of their roles in plant physiological responses have enormous potential in providing insights into polyploid evolution and adaptation.

4. Interpreting –omics patterns in polyploids

4.1. Comparing the transcriptome and proteome

Several steps of regulation occur between the mRNA and protein levels in cells, including RNA processing, alternative splicing, RNA stability control, protein translation, protein modifications, and protein cleavage/degradation. Therefore, mRNA levels may not correlate well with protein levels [148]. Technological advances in high-throughput transcriptomics and quantitative proteomics have made it possible to test this relationship in diverse organisms, including bacteria, yeast, worms, flies, mammals, and plants [124,149–152]. These studies have clearly shown that mRNA levels are often inconsistent with the abundance of the corresponding proteins in cells. In a study of ~6000 genes in yeast, the abundance of a protein for a given mRNA level may vary by 30-fold and vice versa, and the overall correlation coefficient obtained between the mRNA and protein levels was ~0.4 [150].

Only a few studies of plants have compared the proteome and transcriptome in polyploids, with varying results. Hence, it is too early to make firm conclusions. On the one hand, the proteomic results for both allo- and autopolyploids in *Arabidopsis* were highly comparable to results from an expression study based on microarray data [119]—differentially or non-additively expressed proteins also showed transcript-level differences [119]. In contrast, comparison of transcript levels to protein levels in

resynthesized *B. napus* revealed important differences—two-thirds of non-additive proteins had additive transcript levels [123]. The results suggest that differential protein regulation is mainly the result of post-transcriptional modifications [123]. In tetraploid *B. napus*, comparison of protein expression changes in response to drought treatment with those of transcription changes revealed a high correlation between transcript and protein levels in both the direction of change and degree of expression ($r = 0.65$, $p\text{-value} = 5.40\text{e-}11$; [124]). In a comparison of the phosphorylated or oxidized proteins in *Tragopogon*, no correlation was observed between the transcript and protein levels, highlighting the importance of PTMs and the need for analysis at the protein level [103].

It is clear from these studies that the transcriptome and proteome may differ, and for a number of reasons. First, mRNA levels are not always proportional to the amount of protein actually produced (or required) [153], and underestimation of either mRNA and/or protein abundance may further contribute to the poor correlation between the transcriptome and proteome in polyploids [39]. In yeast, gene duplication is preferential among genes whose products exhibit lesser protein–protein interactions [154], and this general pattern is expected to occur in angiosperms as well [55,56]. In addition, dosage compensation at transcriptional levels is well known, and any one protein or mRNA can certainly result in modulation in either monosomic or trisomic lines [155]. As mentioned above, protein PTM is another important factor.

Although label-free proteomic approaches, peak integration, and spectral counting are techniques that hold promise for high-throughput protein quantitation, their application is challenging for polyploids because polyploids have multiple paralogous

genes that may frequently have greater than 90% identity at the amino acid level [156]. Paralogy remains a challenge in teasing apart proteome vs. transcriptome and making meaningful comparisons between polyploids and their parents.

Although mRNA data are useful in analyzing transcriptional regulation and identifying potential protein candidates for further investigation, they may not reflect the levels of proteins, which contain the functional information of the biochemical processes underlying cellular functions and phenotypes. It is essential that we extend from studies of the genomics and transcriptomics of polyploid evolution into proteomics, which can provide fundamental protein-level information connecting genotypes and phenotypes.

Also holding great potential is the ability to extend from reference genome sequence of an allopolyploid to identification of homeolog loss to investigation of the proteome—a reference genome affords the opportunity to distinguish individual gene loss or silencing events from segmental gene loss (e.g., due to chromosome restructuring). The recent published genome of *Brassica napus* [157] provides an example for potential application of such methods to diverse species and future studies of the proteome.

4.2. *The translome*

The translome may be a better proxy for the proteome than the transcriptome appears to be. Coate et al. [158] reported the first translation-level study in an allotetraploid (*Glycine dolichocarpa*) through polysome profiling. Globally, the allopolyploid translome is closely correlated with the transcriptome ($r^2 \geq 0.89$). However, about 25% of the transcriptome is translationally regulated. Importantly, Coate

et al. [158] showed that translational regulation preferentially targets those genes involved in transcription, as well as translation and photosynthesis. Translational regulation also causes regional and possibly whole-chromosome shifts in expression bias between homeologs. Furthermore, translational regulation reduces transcriptional differences between the polyploid and its diploid progenitors, possibly attenuating misregulation resulting from hybridization and polyploidy. Finally, translational regulation appears to correlate positively with the long-term retention of homeologs from a paleopolyploidy event, a result that suggests that this regulation plays a significant role in polyploid evolution [158]. These results have major implications, but their generality remains to be evaluated; similar analyses of the translome are needed for other systems.

4.3. Protein novelty

Alterations to polyploid genomes may also generate novel features and unique combinations of the parental features such that every individual in a polyploid population or species may be unique [26]. Extension of this genetic, genomic, and transcriptomic novelty to the protein and biochemical levels may lead to novel phenotypes. However, the presence of novel proteins in polyploids has received little attention. Koh et al. [103] reported protein PTM isoforms in *Tragopogon* polyploids, but not in the diploid parents. Islam et al. [118] noted some novel proteins in hexaploid wheat, but, as reviewed above, it is unclear if these are truly novel or are the result of a sampling bias, as the actual parental genomes of the hexaploid were not sampled. This lack of parental information in

analyses of polyploids – whether because the parental species are uncertain or possibly extinct – is a recurring problem in the study of polyploids.

Another potential source of proteomic novelty is alternative splicing (AS). AS can play a major role in generating transcriptomic variation by producing multiple isoforms of a transcript from a single gene (e.g., [159–161]) and can thereby influence protein profiles as well. The number of isoforms generated via AS in the few model plants studied in detail to date is extremely high. The genome-wide frequency of AS in *A. thaliana* was estimated to be 61% [162]. However, very few studies of gene expression and AS have focused on polyploids, and there are no data at this point on the proteomic consequences of AS in polyploids. This is another rich area of research in need of investigation.

4.4. Parental legacy

An important consideration in comparing allopolyploids to their parents for any attribute is parental legacy—the extent to which features in a polyploid reflect interspecific differences between the parents. In the case of gene expression, parental legacy refers to differences in expression between duplicate genes present in an allopolyploid due to a legacy of expression differences between the diploid progenitor species [23].

A classic example at the protein level involves the work of Roose and Gottlieb [163,164], who found greater expression of the *T. dubius* parental homeolog than the *T. pratensis* parental homeolog of alcohol dehydrogenase (ADH) in the recently formed allotetraploid *T. miscellus*. Roose and Gottlieb [164] quantified the protein expression of

ADH3 in *T. miscellus* and its diploid parents and demonstrated that there was actually twice as much ADH3 protein in *T. pratensis* as in the other parent, *T. dubius*. Hence, the pattern of differential expression in the allopolyploid was not due to polyploidy or hybridization, but reflected parental legacy (reviewed in [23]). Caution is therefore required when interpreting transcriptome patterns so that differences between homeologous copies are not interpreted as divergence since polyploid formation when they are in fact simply mirroring the contributions of the progenitor species. To avoid confounding post-polyploidization divergence with parental legacy as an explanation for differences in gene expression, comparisons should include the parental species as well as the polyploid. Similar issues of parental legacy are likely to contribute to erroneous interpretations of protein expression as well.

Broad studies of parental legacy at the proteomic level have not been explicitly conducted. However, a few studies point to possible evidence of parental legacy as the explanation for protein profiles in polyploids relative to their diploid progenitors. Variation in protein expression between two plants of *T. mirus* reflects variation detected between individuals of one of the parental species, *T. dubius* [103]. If only one individual of each species had been sampled, the erroneous conclusion of divergence of protein profile in the tetraploid relative to its diploid parents might have been reached. Likewise, synthetic *Arabidopsis* allotetraploids (*A. thaliana* x *A. arenosa*) exhibit non-additive protein profiles, but ~61–62% of the non-additively accumulated proteins were also differentially expressed between the progenitors [119]. How many of these apparent expression differences in the allotetraploid actually represent the legacy of differences between the parents is unclear.

5. Conclusions and future prospects

Proteomes provide insights into the consequences of genomic merger and doubling beyond what is obtained from genomic and/or transcriptomic data. However, it is evident that our understanding of the consequences of polyploidy on the proteome remains in its infancy. Many proteomic studies conducted to date involve a limited number of proteins, making it difficult to extrapolate to the proteome as a whole. Our major goal, therefore, is to stimulate further study at the proteomic level, particularly broad-scale analyses that survey large swaths of the proteome using state-of-the-art proteomics technologies (e.g., data independent acquisition mass spectrometry).

Accumulating evidence from transcriptome analyses reveals expression bias between the parental genomes in many allopolyploids. Numerous processes, from gene loss to gene silencing to regulation of an entire parental transcriptome in a polyploid, make the transcriptome an imperfect reflection of the genome. In some cases, proteomic profiles show qualitatively and/or quantitatively non-additive patterns; in others, novel transcripts have been detected in polyploids, and several studies indicate novelty at the proteomic level. Moreover, the concept of parental legacy has emerged as an important facet at the level of gene expression and should be considered in proteomic analyses. Even as proteomic studies of allopolyploids are beginning to reveal elements of the consequences of genome merger and duplication on the proteome, the effects of genome duplication alone, as in autopolyploids, remain understudied.

Much additional work remains to be done. Rigorous comparisons at the genomic,

transcriptomic and proteomic levels in multiple genotypes using multiple replications are badly needed—no such investigation has been conducted in any system. We also need more studies to tease apart the immediate evolutionary consequences of genome doubling on the proteome (using natural systems) and the longer-term impact. Another major goal for the next decade should be to translate the proteomic changes in polyploids to the phenotype through analyses of the molecular basis underlying differences in morphology, physiology, and fitness.

Acknowledgments

The authors thank Drs. Jin Koh and Mi-Jeong Yoo for helpful discussions and two reviewers for their helpful input. The -omics research on polyploidy was supported by University of Florida Research Opportunity Seed Fund Grant to (00095031) to S. Chen, D. E. Soltis, W. B. Barbazuk, P. S. Soltis, and L. McIntyre. We also acknowledge NSF grant DEB-1146065.

Table 1. Table 1. List of proteomics studies in plant polyploids (NA: Not Available, ND: Not Done, ToF: Time of Flight, MS: mass spectrometry; MS/MS: tandem mass spectrometry, ESI: electron spray ionization, LC: liquid chromatography; iTRAQ: Isobaric tags for relative and absolute quantitation)

Species	Ploidy	Approach	Total proteins/ spots	Differential proteins/spots	MS Instrument	Key findings	References
<i>Brassica oleracea</i>	Autoploid	2-DE	1060-1194 (spots)	4-28 (spots)	ND	Little contribution of autoploidy to proteomes of green tissues	[97]
<i>B. oleracea</i> var. <i>botrytis</i> , <i>B. rapa</i>	Allopoloid	2-DE	>1600 (spots)	519-583 (additive polypeptides)	LCQ deca xp (ThermoFinnigan)	Nonadditive stem and root proteomes in the synthetic amphiploids (allotetraploids)	[98]
<i>B. napus</i>	Allopoloid	2-DE	583 (spots)	205-335 (non-additive spots)	LCQ deca xp (ThermoFinnigan)	Protein function and subcellular localization do not reflect gene regulation of proteome	[99]
<i>B. napus</i> , <i>B. rapa</i> , <i>B. oleracea</i>	Allopoloid	2-DE	304-340 (spots)	18	MALDI-TOF (AB Sciex)	Allopoloidization led to extensive proteome changes, but the house-keeping proteins were not disturbed	[126]
<i>B. napus</i>	Allopoloid	iTRAQ LC- MS/MS	1976	417	Hybrid quadrupole Orbitrap (Q Exactive) (Thermo Fisher Scientific,)	Drought stress responsive <i>B. napus</i> leaf proteomic and PTM changes.	[124]
<i>B. carinata</i> , <i>B. rapa</i> and a synthetic hexaploid	Allopoloid	iTRAQ LC- MS/MS	2044	452	Q-Exactive Orbitrap (Thermo Scientific)	There were expression level dominance of one parental expression and an expression level dominance bias towards tetraploid progenitor.	[107]
<i>Triticum turgidum</i> sp. $\times$ <i>T. tauschii</i> sp.	Allopoloid	2-DE	NA	8	ND	Parental genome expression favored in synthetic polyploids	[117]
<i>T. aestivum</i>	Allopoloid	2-DE	NA	NA	QTOF-MS/MS (Waters)	Genome interaction in proteome expression	[118]
<i>Gossypium hirsutum</i>	Allopoloid	2-DE	854	78-108 across sub-genome	LTQ XL MS (Thermo Fisher)	Biased accumulation of seed storage proteins	[111]
<i>G. hirsutum</i>	Allopoloid	2-DE, iTRAQ LC-MS/MS	1223	240 iTRAQ proteins and	QSTAR Elite MS (AB Sciex)	Domestication was accompanied by a doubling of protein developmental	[105]

				293 2D spots		dynamics	
<i>G. hirsutum</i>	Allopoloid	2-DE, iTRAQ LC-MS/MS	5884	2729	Triple TOF 5600 system (AB Sciex)	Cotton ovule proteome involved in fiber initiation.	[127]
<i>Tragopogon mirus</i> , <i>T. dubius</i> , <i>T. porrifolius</i>	Allopoloid	iTRAQ LC-MS/MS	476	68	QSTAR Elite MS (AB Sciex)	Impact of hybridization on the proteome is more important than polyploidization	[103]
<i>Lophopyrum elongatum</i> $\times$ <i>wheat</i>	Allopoloid	2D-DIGE	55 (mitochondrial proteins)	15	MALDI TOF/TOF (Bruker)	Role of mitochondrial proteins in salinity stress tolerance in synthetic allopoloids	[101]
<i>T. aestivum</i>	Allopoloid	2D-DIGE	NA	386	MALDI-TOF/TOF 5800 (AB Sciex)	Similarities in stress adaptation and differences in regulatory processes	[102]
<i>T. aestivum</i> , <i>T. monococcum</i>	Allopoloid	Label free, LC—MS/MS	528	NA	Q-Exactive (Thermo Electron)	Nuclear protein of the grain of cultivated wheat species	[131]
<i>Citrus deliciosa</i> , <i>C. aurantifolia</i> , <i>Fortunella margarita</i>	Allopoloid	2-DE	434-604	40%	NA	Allotetraploid hybrid maps were found closer to that of their mandarin parent than to those of their non- mandarin parents	[128]
<i>Musa acuminata</i> , <i>M. balbisiana</i>	Allopoloid	2D-DIGE	NA	NA	MALDI-TOF/TOF 4800 (AB Sciex)	Proteotype observed for allopoloid banana does not always correspond to the expectation from the genomic standpoint, perhaps due to either gene deletion or gene silencing	[100]
<i>M. acuminata</i>	Allopoloid	2D-DIGE 2DE	NA	NA	Q-Exactive Orbitrap (Thermo Scientific)	Characterization of the HSP70 family during osmotic stress	[130]
<i>Fragaria</i> $\times$ <i>ananassa</i>	Allopoloid	2D-DIGE 1-D SDS-PAGE	520-622 (spots)	NA	ESI-IT-MS/MS (Agilent)	Constitutive and differential proteins across different genotypes involved in processes controlling major fruit quality traits were revealed	[165]
<i>Camelina sativa</i>	Allopoloid	iTRAQ-LC-MS/MS	1532	NA	Triple TOF 5600 mass spectrometer (AB Sciex)	Seed proteome, unravelling the effect of overexpression of the <i>Arabidopsis</i> G-protein γ subunit 3 (AGG3)	[108]
<i>Hevea brasiliensis</i>	Allopoloid	SDS-PAGE, 2D, nano-LC-ESI-MS/MS	440–520 (spots)	34 (spots)	ND	Differential latex serum proteomes in <i>Phytophthora</i> tolerant clone BPM24 and susceptible clone RRIM600	[132]
<i>H. brasiliensis</i>	Allopoloid	iTRAQ-	1499	101	TripleTOF 5600 (AB	Differential latex serum proteome in	[133]

		coupled to 2D- LC-MS/MS			Sciex)	ethylene and methyl jasmonate treated rubber tree barks.	
<i>Manihot esculenta</i>	Autoploid	2-DE	300 (spots)	52	MALDI-TOF Ultraflex (Bruker)	Leaf proteome of a diploid cultivar of was clearly differentiated from that of its autotetraploid derivative	[139]
<i>Arabidopsis suecica</i> , <i>A. arenosa</i>	Autoploid	iTRAQ LC- MS/MS	2083	95-203	LTQ MS (Thermo Fisher)	High protein divergence between species and rapid changes in post- transcriptional regulation and PTMs	[119]
<i>Saccharum officinarum</i>	Autoploid	2-DE	NA	83	ESI-Q-ToF	Root proteome responses to salt stress in sugar cane varieties	[142]
<i>S. officinarum</i>	Autoploid	2-DE, label free	752	14-23	SYNAPT G2-S HDMS (Waters)	Embryogenic and non-embryogenic callus during somatic embryogenesis	[143]
<i>Robinia pseudoacacia</i>	Autoploid	2 DE	NA	58	MALDI-TOF Autoflex (Bruker), MALDI-TOF/TOF 4700 (AB Sciex)	Diploid plants experienced greater negative effects than those of tetraploid plants under stress conditions	[165]

Figure legends

Fig. 1. Following WGD, gene loss leads to a “diploid” genome via fractionation. Most retained genes may trace to one parent, which is consistent with a process of genomic dominance.

Fig. 2. Neofunctionalization vs. subfunctionalization. In neofunctionalization, duplicates take on new functions (top: gray vs. blue boxes). In subfunctionalization, duplicates may be expressed in different tissues (bottom).

References:

- [1] D.E. Soltis, V. a. Albert, J. Leebens-Mack, C.D. Bell, A.H. Paterson, C. Zheng, D. Sankoff, C.W. DePamphilis, P.K. Wall, P.S. Soltis, Polyploidy and angiosperm diversification, *Am. J. Bot.* 96 (2009) 336–348.
- [2] Y. Van de Peer, S. Maere, A. Meyer, The evolutionary significance of ancient genome duplications, *Nat. Rev. Genet.* 10 (2009) 725–32.
- [3] Y. Van de Peer, S. Maere, A. Meyer, 2R or not 2R is not the question anymore, *Nat. Rev. Genet.* 11 (2010) 166.
- [4] B.C. Husband, S.J. Baldwin, J. Suda, The incidence of polyploidy in natural plant populations: major patterns and evolutionary processes, in: J. Greilhuber, J. Dolezel, J.F. Wendel (Eds.), *Plant genome diversity volume 2*, Springer, Vienna, 2013, pp. 255–276.

- [5] C. Cañestro, Two rounds of whole-genome duplication: evidence and impact on the evolution of vertebrate innovations, in: P.S. Soltis, D.E. Soltis (Eds.), *Polyploidy and genome evolution*, Springer, New York, 2012, pp. 309–339.
- [6] I. Braasch, J. Postlethwait, Polyploidy in fish and the teleost genome duplication, in: P.S. Soltis, D.E. Soltis (Eds.), *Polyploidy and genome evolution*, Springer, New York, 2012, pp. 341–383.
- [7] C. Hudson, G. Conant, Yeast as a window into changes in genome complexity due to polyploidization, in: P.S. Soltis, D.E. Soltis (Eds.), *Polyploidy and genome evolution*, Springer, New York, 2012, pp. 293–308.
- [8] T.E. Wood, N. Takebayashi, M.S. Barker, I. Mayrose, P.B. Greenspoon, L.H. Rieseberg, The frequency of polyploid speciation in vascular plants, *Proc. Natl. Acad. Sci.* 106 (2009) 13875–13879.
- [9] M.S. Barker, P.G. Wolf, Unfurling fern biology in the genomics age, *Bioscience* 60 (2010) 177–185.
- [10] Y. Jiao, N.J. Wickett, S. Ayyampalayam, A.S. Chanderbali, L. Landherr, P.E. Ralph, L.P. Tomsho, Y. Hu, H. Liang, P.S. Soltis, D.E. Soltis, S.W. Clifton, S.E. Schlarbaum, S.C. Schuster, H. Ma, J. Leebens-Mack, C.W. dePamphilis, Ancestral polyploidy in seed plants and angiosperms, *Nature* 473 (2011) 97–100.
- [11] *Amborella* Genome Project, The *Amborella* genome and the evolution of flowering plants, *Science* 342 (2013) 1241089.
- [12] L. Cui, P.K. Wall, J.H. Leebens-Mack, B.G. Lindsay, D.E. Soltis, J.J. Doyle, P.S. Soltis, J.E. Carlson, K. Arumuganathan, A. Barakat, V. a. Albert, H. Ma, C.W.

- DePamphilis, Widespread genome duplications throughout the history of flowering plants, *Genome Res.* 16 (2006) 738–749.
- [13] K. Vanneste, S. Maere, Y. Van de Peer, Tangled up in two: a burst of genome duplications at the end of the Cretaceous and the consequences for plant evolution, *Philos. Philos. Trans. R. Soc. London B Biol. Sci.* 369 (2014) 20130353.
- [14] M. Barker, G. Baute, S.L. Liu, Duplications and turnover in plant genomes, in: J.F. Wendel, J. Greilhuber, J. Dolezel, I.J. Leitch (Eds.), *Plant genome diversity volume 1*, Springer, Vienna (2012) pp. 155-169.
- [15] M.C. Estep, M.R. McKain, D. Vela Diaz, J. Zhong, J.G. Hodge, T.R. Hodkinson, D.J. Layton, S.T. Malcomber, R. Pasquet, E. a Kellogg, Allopolyploidy, diversification, and the Miocene grassland expansion, *Proc. Natl. Acad. Sci.* 111 (2014) 15149–15154.
- [16] D.C. Tank, J.M. Eastman, M.W. Pennell, P.S. Soltis, D.E. Soltis, C.E. Hinchliff, J.W. Brown, E.B. Sessa, L.J. Harmon, Nested radiations and the pulse of angiosperm diversification : increased diversification rates often follow whole genome duplications, *New Phytol.* 207 (2015) 454–467.
- [17] P.S. Solis, D.E. Soltis, Ancient WGD events as drivers of key innovations in angiosperms. *Curr. Opin. Genet. Dev.* (in press).
- [18] M. Ainouche, H. Chelaifa, J. Ferreira, S. Bellot, A. Ainouche, A. Salmon, Polyploid evolution in *Spartina*: dealing with highly redundant hybrid genomes, in: P.S. Soltis, D.E. Soltis (Eds.), *Polyploidy and genome evolution*, Springer, New York, 2012, pp. 225-243.

- [19] P.S. Soltis, D.E. Soltis, The role of genetic and genomic attributes in the success of polyploids, *Proc. Natl. Acad. Sci.* 97 (2000) 7051–7057.
- [20] P.S. Soltis, D.E. Soltis, The role of hybridization in plant speciation, *Annu. Rev. Plant Biol.* 60 (2009) 561–588.
- [21] R.T. Gaeta, J.C. Pires, F. Iniguez-Luy, E. Leon, T.C. Osborn, Genomic changes in resynthesized *Brassica napus* and their effect on gene expression and phenotype, *Plant Cell* 19 (2007) 3403–3417.
- [22] R.J.A. Buggs, L. Zhang, N. Miles, J. a Tate, L. Gao, W. Wei, P.S. Schnable, W.B. Barbazuk, P.S. Soltis, D.E. Soltis, Transcriptomic shock generates evolutionary novelty in a newly formed, natural allopolyploid plant, *Curr. Biol.* 21 (2011) 551–556.
- [23] R.J.A. Buggs, J.F. Wendel, J.J. Doyle, D.E. Soltis, P.S. Soltis, J.E. Coate, The legacy of diploid progenitors in allopolyploid gene expression patterns, *Philos. Trans. R. Soc. London B Biol. Sci.* 369 (2014) 20130354.
- [24] M.J. Hegarty, R.J. Abbott, S.J. Hiscock, Allopolyploid speciation in action: the origins and evolution of *Senecio cambrensis*, in: P.S. Soltis, D.E. Soltis (Eds.), *Polyploidy and genome evolution*, Springer, New York, 2012, pp. 245–270.
- [25] D.E. Soltis, R.J.A. Buggs, W.B. Barbazuk, S. Chamala, M. Chester, J.P. Gallagher, P.S. Schnable, P.S. Soltis, The early stages of polyploidy: rapid and repeated evolution in *Tragopogon*, in: P.S. Soltis, D.E. Soltis (Eds.), *Polyploidy and genome evolution*, Springer, New York, 2012, pp. 271–292.
- [26] D.E. Soltis, C.J. Visger, P.S. Soltis, The polyploidy revolution then...and now:

- Stebbins revisited, *Am. J. Bot.* 101 (2014) 1057–1078.
- [27] H. Kihara, T. Ono, Chromosomenzahlen und systematische Gruppierung der *Rumex*-Arten, *Z. Zellf. Mikroskop. Anat.* 4 (1926) 475–481.
- [28] A. Müntzing, The evolutionary significance of autopolyploidy, *Hereditas* 21 (1936) 363–378.
- [29] C.D. Darlington, Recent advances in cytology, second ed., P. Blackiston's, Philadelphia, 1937.
- [30] J. Clausen, D.D. Keck, W.M. Hiesey, Experimental studies on the nature of species. II. Plant evolution through amphiploidy, with examples from the *Madiinae*, Carnegie Institute of Washington. Washington DC, 1945.
- [31] C.L. Stebbins Jr, Variation and evolution in plants, Oxford Univ. Press, New York, 1950.
- [32] V. Grant, Plant speciation, second ed., Columbia Univ. Press, New York, 1981.
- [33] W. Lewis, Polyploidy in angiosperms: dicotyledons, in: W. Lewis (Eds.), Polyploidy, Plenum Press, New York, 1980, pp. 241–268.
- [34] S. Ohno, Evolution by gene duplication, Springer, New York, 1970.
- [35] M. Lynch, J.S. Conery, The evolutionary fate and consequences of duplicate genes, *Science*, 290 (2000) 1151–1155.
- [36] J.J. Doyle, L.E. Flagel, A.H. Paterson, R.A. Rapp, D.E. Soltis, P.S. Soltis, J.F.

- Wendel, Evolutionary genetics of genome merger and doubling in plants, *Annu. Rev. Genet.* 42 (2008) 443–461.
- [37] S. Jackson, Z.J. Chen, Genomic and expression plasticity of polyploidy, *Curr. Opin. Plant Biol.* 13 (2010) 153–159.
- [38] Y. Wang, X. Wang, A.H. Paterson, Genome and gene duplications and gene expression divergence: a view from plants, *Ann. N. Y. Acad. Sci.*, 1256 (2012) 1–14.
- [39] M.J. Yoo, X. Liu, J.C. Pires, P.S. Soltis, D.E. Soltis, Nonadditive gene expression in polyploids, *Annu. Rev. Genet.* 48 (2014) 485–517.
- [40] A.H. Paterson, B. a. Chapman, J.C. Kissinger, J.E. Bowers, F. a. Feltus, J.C. Estill, Many gene and domain families have convergent fates following independent whole-genome duplication events in *Arabidopsis*, *Oryza*, *Saccharomyces* and *Tetraodon*, *Trends Genet.* 22 (2006) 597–602.
- [41] R. De Smet, K.L. Adams, K. Vandepoele, M.C.E. Van Montagu, S. Maere, Y. Van de Peer, Convergent gene loss following gene and genome duplications creates single-copy families in flowering plants, *Proc. Natl. Acad. Sci.* 110 (2013) 2898–2903.
- [42] B.A. Chapman, J.E. Bowers, F. a Feltus, A.H. Paterson, Buffering of crucial functions by paleologous duplicated genes may contribute cyclicity to angiosperm genome duplication, *Proc. Natl. Acad. Sci.* 103 (2006) 2730–2735.
- [43] M.S. Barker, N.C. Kane, M. Matvienko, A. Kozik, R.W. Michelmore, S.J. Knapp, L.H. Rieseberg, Multiple paleopolyploidizations during the evolution of the

- compositae reveal parallel patterns of duplicate gene retention after millions of years, *Mol. Biol. Evol.* 25 (2008) 2445–2455.
- [44] R.J.A. Buggs, S. Chamala, W. Wu, J. a. Tate, P.S. Schnable, D.E. Soltis, P.S. Soltis, W.B. Barbazuk, Rapid, repeated, and clustered loss of duplicate genes in allopolyploid plant populations of independent origin, *Curr. Biol.* 22 (2012) 248–252.
- [45] B. Chaudhary, L. Flagel, R.M. Stupar, J. a. Udall, N. Verma, N.M. Springer, J.F. Wendel, Reciprocal silencing, transcriptional bias and functional divergence of homeologs in polyploid cotton (*Gossypium*), *Genetics* 182 (2009) 503–517.
- [46] J. Wendel, L. Flagel, K. Adams, Jeans, genes, and genomes: cotton as a model for studying polyploidy, in: P.S. Soltis, D.E. Soltis (Eds.), *Polyploidy and genome evolution*, Springer, New York, 2012, pp. 181–207.
- [47] J.E. Bowers, B.A. Chapman, J. Rong, A.H. Paterson, Unravelling angiosperm genome evolution by phylogenetic analysis of chromosomal duplication events, *Nature* 422 (2003) 433–438.
- [48] X. Wang, H. Wang, J. Wang, R. Sun, J. Wu, S. Liu, Y. Bai, J.-H. Mun, I. Bancroft, F. Cheng, S. Huang, X. Li, W. Hua, J. Wang, X. Wang, M. Freeling, J.C. Pires, A.H. Paterson, B. Chalhoub, B. Wang, et al., The genome of the mesopolyploid crop species *Brassica rapa*, *Nat. Genet.* 43 (2011) 1035–1039.
- [49] B.S. Gaut, J.F. Doebley, DNA sequence evidence for the segmental allotetraploid origin of maize, *Proc. Natl. Acad. Sci.* 94 (1997) 6809–6814.
- [50] A.H. Paterson, Z. Wang, J. Li, H. Tang, Ancient and recent polyploidy in the

- monocots, in: P.S. Soltis, D.E. Soltis (Eds.), *Polyploidy and genome evolution*, Springer, New York, 2012, pp. 93–108.
- [51] B.C. Thomas, B. Pedersen, M. Freeling, Following tetraploidy in an *Arabidopsis* ancestor, genes were removed preferentially from one homeolog leaving clusters enriched in dose-sensitive genes, *Genome Res.* 16 (2006) 934–946.
- [52] J.C. Schnable, N.M. Springer, M. Freeling, Differentiation of the maize subgenomes by genome dominance and both ancient and ongoing gene loss, *Proc. Natl. Acad. Sci.* 108 (2011) 4069–4074.
- [53] M. Freeling, M.R. Woodhouse, S. Subramaniam, G. Turco, D. Lisch, J.C. Schnable, Fractionation mutagenesis and similar consequences of mechanisms removing dispensable or less-expressed DNA in plants, *Curr. Opin. Plant Biol.* 15 (2012) 131–139.
- [54] J.C. Schnable, M. Freeling, Genes identified by visible mutant phenotypes show increased bias toward one of two subgenomes of maize, *PLoS One*, 6 (2011) e17855.
- [55] J.A. Birchler, R.A. Veitia, The gene balance hypothesis: from classical genetics to modern genomics, *Plant Cell* 19 (2007) 395–402.
- [56] J.A. Birchler, R.A. Veitia, Gene balance hypothesis: Connecting issues of dosage sensitivity across biological disciplines, *Proc. Natl. Acad. Sci.* 109 (2012) 14746–14753.
- [57] M. Freeling, Bias in plant gene content following different sorts of duplication: tandem, whole-genome, segmental, or by transposition, *Annu. Rev. Plant Biol.* 60

- (2009) 433–453.
- [58] S. Maere, S. De Bodt, J. Raes, T. Casneuf, M. Van Montagu, M. Kuiper, Y. Van de Peer, Modeling gene and genome duplications in eukaryotes, *Proc. Natl. Acad. Sci.* 102 (2005) 5454–5459.
 - [59] M.W. Hahn, Distinguishing among evolutionary models for the maintenance of gene duplicates, *J. Hered.* 100 (2009) 605–617.
 - [60] M. Bekaert, P.P. Edger, J.C. Pires, G.C. Conant, Two-phase resolution of polyploidy in the *Arabidopsis* metabolic network gives rise to relative and absolute dosage constraints, *Plant Cell* 23 (2011) 1719–1728.
 - [61] H. Tang, M.R. Woodhouse, F. Cheng, J.C. Schnable, B.S. Pedersen, G. Conant, X. Wang, M. Freeling, J.C. Pires, Altered patterns of fractionation and exon deletions in *Brassica rapa* support a two-step model of paleohexaploidy, *Genetics* 190 (2012) 1563–1574.
 - [62] R.J.A. Buggs, N.M. Elliott, L. Zhang, J. Koh, L.F. Viccini, D.E. Soltis, P.S. Soltis, Tissue-specific silencing of homoeologs in natural populations of the recent allopolyploid *Tragopogon mirus*, *New Phytol.* 186 (2010) 175–183.
 - [63] J.S. Taylor, J. Raes, Duplication and divergence: the evolution of new genes and old ideas, *Annu. Rev. Genet.* 38 (2004) 615–643.
 - [64] L. Comai, Phenotypic instability and rapid gene silencing in newly formed *Arabidopsis* allotetraploids, *Plant Cell* 12 (2000) 1551–1568.
 - [65] J.F. Wendel, Genome evolution in polyploids, *Plant Mol. Biol.* 42 (2000) 225–249.
 - [66] V.E. Prince, F.B. Pickett, Splitting pairs: the diverging fates of duplicated genes,

- Nat. Rev. Genet. 3 (2002) 827–837.
- [67] V. Chagué, J. Just, I. Mestiri, S. Balzergue, A.M. Tanguy, C. Huneau, V. Huteau, H. Belcram, O. Coriton, J. Jahier, B. Chalhoub, Genome-wide gene expression changes in genetically stable synthetic and natural wheat allohexaploids, *New Phytol.* 187 (2010) 1181–1194.
- [68] J.A. Tate, D.E. Soltis, P.S. Soltis, Polyploidy in plants, *Evol. Genome* (2005) 371–426.
- [69] Z. Xiong, R.T. Gaeta, J.C. Pires, Homoeologous shuffling and chromosome compensation maintain genome balance in resynthesized allopolyploid *Brassica napus*, *Proc. Natl. Acad. Sci.* 108 (2011) 7908–7913.
- [70] M. Chester, J.P. Gallagher, V.V. Symonds, A.V. Cruz da Silva, E.V. Mavrodiev, A.R. Leitch, P.S. Soltis, D.E. Soltis, Extensive chromosomal variation in a recently formed natural allopolyploid species, *Tragopogon miscellus* (Asteraceae), *Proc. Natl. Acad. Sci.* 109 (2012) 1176–1181.
- [71] M. Feldman, B. Liu, G. Segal, S. Abbo, A.A. Levy, J.M. Vega, Rapid elimination of low-copy DNA sequences in polyploid wheat: a possible mechanism for differentiation of homoeologous chromosomes, *Genetics* 147 (1997) 1381–1387.
- [72] H. Ozkan, A.A. Levy, M. Feldman, Allopolyploidy-induced rapid genome evolution in the wheat (*Aegilops–Triticum*) group, *Plant Cell* 13 (2001) 1735–1747.
- [73] H. Shaked, K. Kashkush, H. Ozkan, M. Feldman, A.A. Levy, Sequence elimination and cytosine methylation are rapid and reproducible responses of the

- genome to wide hybridization and allopolyploidy in wheat, *Plant Cell* 13 (2001) 1749–1759.
- [74] K. Kashkush, M. Feldman, A.A. Levy, Gene Loss, Silencing and activation in a newly synthesized wheat allotetraploid, *Genetics* 160 (2002) 1651–1659.
- [75] R.J.A. Buggs, A.N. Doust, J.A. Tate, J. Koh, K. Soltis, F.A. Feltus, A.H. Paterson, P.S. Soltis, D.E. Soltis, Gene loss and silencing in *Tragopogon miscellus* (Asteraceae): comparison of natural and synthetic allotetraploids, *Heredity* 103 (2009) 73–81.
- [76] S. Feng, S.E. Jacobsen, W. Reik, Epigenetic Reprogramming in Plant and Animal Development, *Science* 330 (2010) 622–627.
- [77] J. Bender, DNA methylation and epigenetics, *Annu. Rev. Plant Biol.* 55 (2004) 41–68.
- [78] J.E. Coate, J.J. Doyle, Quantifying whole transcriptome size, a prerequisite for understanding transcriptome evolution across species: an example from a plant allopolyploid, *Genome Biol. Evol.* 2 (2010) 534–546.
- [79] L.E. Flagel, J.F. Wendel, Evolutionary rate variation, genomic dominance and duplicate gene expression evolution during allotetraploid cotton speciation, *New Phytol.* 186 (2010) 184–193.
- [80] H.S. Lee, Z.J. Chen, Protein-coding genes are epigenetically regulated in *Arabidopsis* polyploids, *Proc. Natl. Acad. Sci.* 98 (2001) 6753–6758.
- [81] J. Wang, L. Tian, A. Madlung, H.S. Lee, M. Chen, J.J. Lee, B. Watson, T. Kagochi, L. Comai, Z.J. Chen, Stochastic and epigenetic changes of gene

- expression in *Arabidopsis* polyploids, *Genetics* 167 (2004) 1961–1973.
- [82] O.M. Scheid, K. Afsar, J. Paszkowski, Formation of stable epialleles and their paramutation-like interaction in tetraploid *Arabidopsis thaliana*, *Nat. Genet.* 34 (2003) 450–454.
- [83] C.S. Pikaard, Z.J. Chen, Nucleolar dominance, in: M.R. Paule (Eds.), *RNA polymerase I: transcription of eukaryotic ribosomal RNA*, R.G. Landes, Austin, 1998, pp. 275–293.
- [84] C.S. Pikaard, Nucleolar dominance and silencing of transcription, *Trends Plant Sci.* 4 (1999) 478–483.
- [85] C.S. Pikaard, Nucleolar dominance: uniparental gene silencing on a multi-megabase scale in genetic hybrids, *Plant Mol. Biol.* 43 (2000) 163–177.
- [86] Z.J. Chen, L. Comai, C.S. Pikaard, Gene dosage and stochastic effects determine the severity and direction of uniparental ribosomal RNA gene silencing (nucleolar dominance) in *Arabidopsis* allopolyploids, *Proc. Natl. Acad. Sci.* 95 (1998) 14891–14896.
- [87] Z.J. Chen, C.S. Pikaard, Transcriptional analysis of nucleolar dominance in polyploid plants: biased expression/silencing of progenitor rRNA genes is developmentally regulated in *Brassica*, *Proc. Natl. Acad. Sci.* 94 (1997) 3442–3447.
- [88] M. Lynch, A. Force, The probability of duplicate gene preservation by subfunctionalization, *Genetics*. 154 (2000) 459–473.
- [89] K.L. Adams, R. Cronn, R. Percifield, J.F. Wendel, Genes duplicated by polyploidy

- show unequal contributions to the transcriptome and organ-specific reciprocal silencing, *Proc. Natl. Acad. Sci.* 100 (2003) 4649–4654.
- [90] E.W. Ganko, B.C. Meyers, T.J. Vision, Divergence in expression between duplicated genes in *Arabidopsis*, *Mol. Biol. Evol.* 24 (2007) 2298–2309.
- [91] M.T. Rutter, K.V. Cross, P.A. Van Woert, Birth, death and subfunctionalization in the *Arabidopsis* genome, *Trends Plant Sci.* 17 (2012) 204–212.
- [92] A. Roulin, P.L. Auer, M. Libault, J. Schlueter, A. Farmer, G. May, G. Stacey, R.W. Doerge, S. a. Jackson, The fate of duplicated genes in a polyploid plant genome, *Plant J.* 73 (2013) 143–153.
- [93] I.J. Leitch, M.D. Bennett, Genome downsizing in polyploid plants, *Biol. J. Linn. Soc.* 82 (2004) 651–663.
- [94] T.C. Osborn, J.C. Pires, J.A. Birchler, D.L. Auger, Z.J. Chen, H.S. Lee, L. Comai, A. Madlung, R.W. Doerge, V. Colot, R.A. Martienssen, Understanding mechanisms of novel gene expression in polyploids, *Trends Genet.* 19 (2003) 141–147.
- [95] N. Senerchia, F. Felber, C. Parisod, Genome reorganization in F1 hybrids uncovers the role of retrotransposons in reproductive isolation, *Philos. Trans. R. Soc. London B Biol. Sci.* 282 (2015) 20142874.
- [96] S. Chen, A.C. Harmon, Advances in plant proteomics, *Proteomics* 6 (2006) 5504–5516.
- [97] W. Albertin, P. Brabant, O. Catrice, F. Eber, E. Jenczewski, A.M. Chèvre, H. Thiellement, Autopolyploidy in cabbage (*Brassica oleracea* L.) does not alter

- significantly the proteomes of green tissues, *Proteomics* 5 (2005) 2131–2139.
- [98] W. Albertin, T. Balliau, P. Brabant, A.M. Chèvre, F. Eber, C. Malosse, H. Thiellement, Numerous and rapid nonstochastic modifications of gene products in newly synthesized *Brassica napus* allotetraploids, *Genetics* 173 (2006) 1101–1113.
- [99] W. Albertin, K. Alix, T. Balliau, P. Brabant, M. Davanture, C. Malosse, B. Valot, H. Thiellement, Differential regulation of gene products in newly synthesized *Brassica napus* allotetraploids is not related to protein function nor subcellular localization, *BMC Genomics* 8 (2007) 56–70.
- [100] S.C. Carpentier, B. Panis, J. Renaut, B. Samyn, A. Vertommen, A.C. Vanhove, R. Swennen, K. Sergeant, The use of 2D-electrophoresis and de novo sequencing to characterize inter- and intra-cultivar protein polymorphisms in an allopolyploid crop, *Phytochemistry* 72 (2011) 1243–1250.
- [101] R.P. Jacoby, A.H. Millar, N.L. Taylor, Investigating the role of respiration in plant salinity tolerance by analyzing mitochondrial proteomes from wheat and a salinity-tolerant amphiploid (wheat $\times$ *lophopyrum elongatum*), *J. Proteome Res.* 12 (2013) 4807–4829.
- [102] K. Kosová, P. Vítámvás, S. Planchon, J. Renaut, R. Vanková, I.T. Prášil, Proteome analysis of cold response in spring and winter wheat (*Triticum aestivum*) crowns reveals similarities in stress adaptation and differences in regulatory processes between the growth habits, *J. Proteome Res.* 12 (2013) 4830–4845.
- [103] J. Koh, S. Chen, N. Zhu, F. Yu, P.S. Soltis, D.E. Soltis, Comparative proteomics

- of the recently and recurrently formed natural allopolyploid *Tragopogon mirus* (Asteraceae) and its parents, *New Phytol.* 196 (2012) 292–305.
- [104] G. Hu, J. Koh, M.J. Yoo, K. Grupp, S. Chen, J.F. Wendel, Proteomic profiling of developing cotton fibers from wild and domesticated *Gossypium barbadense*, *New Phytol.* 200 (2013) 570–582.
- [105] G. Hu, J. Koh, M.J. Yoo, D. Pathak, S. Chen, J. Wendel, Proteomics profiling of fiber development and domestication in upland cotton (*Gossypium hirsutum* L.), *Planta*, 240 (2014) 1237–1251.
- [106] G. Hu, J. Koh, M.J. Yoo, S. Chen, J.F. Wendel, Gene-expression novelty in allopolyploid Cotton: a proteomic perspective, *Genetics* 200 (2015) 91–104.
- [107] Y. Shen, Y. Zhang, J. Zou, J. Meng, J. Wang, Comparative proteomic study on *Brassica* hexaploid and its parents provides new insights into the effects of polyploidization, *J. Proteomics* 112 (2015) 274–284.
- [108] S. Alvarez, S. Roy Choudhury, K. Sivagnanam, L.M. Hicks, S. Pandey, Quantitative proteomics analysis of *Camelina sativa* seeds overexpressing the AGG3 gene to identify the proteomic basis of increased yield and stress tolerance, *J. Proteome Res.* 14 (2015) 2606–2616.
- [109] J. Hu, C. Rampitsch, N.V. Bykova, Advances in plant proteomics toward improvement of crop productivity and stress resistance, *Front. Plant Sci.* 6 (2015) 1–15.
- [110] J. Wang, L. Tian, H.S. Lee, N.E. Wei, H. Jiang, B. Watson, A. Madlung, T.C. Osborn, R.W. Doerge, L. Comai, Z.J. Chen, Genomewide nonadditive gene

- regulation in *Arabidopsis* Allotetraploids, *Genetics* 172 (2006) 507–517.
- [111] G. Hu, N.L. Houston, D. Pathak, L. Schmidt, J.J. Thelen, J.F. Wendel, Genomically biased accumulation of seed storage proteins in allopolyploid cotton, *Genetics* 189 (2011) 1103–1115.
- [112] B. Yaakov, K. Kashkush, Methylation, transcription, and rearrangements of transposable elements in synthetic allopolyploids, *Int. J. Plant Genomics* 2011 (2011) 569826.
- [113] K. Song, P. Lu, K. Tang, T.C. Osborn, Rapid genome change in synthetic polyploids of *Brassica* and its implications for polyploid evolution, *Proc. Natl. Acad. Sci.* 92 (1995) 7719–7723.
- [114] K.L. Adams, R. Percifield, J.F. Wendel, Organ-specific silencing of duplicated genes in a newly synthesized cotton allotetraploid, *Genetics* 168 (2004) 2217–2226.
- [115] K. Skalická, K.Y. Lim, R. Matyasek, M. Matzke, A.R. Leitch, A. Kovarik, Preferential elimination of repeated DNA sequences from the paternal, *Nicotiana tomentosiformis* genome donor of a synthetic, allotetraploid tobacco, *New Phytol.* 166 (2005) 291–303.
- [116] J. Wang, L. Tian, H.S. Lee, N.E. Wei, H. Jiang, B. Watson, A. Madlung, T.C. Osborn, R.W. Doerge, L. Comai, Z.J. Chen, Genomewide nonadditive gene regulation in *Arabidopsis* allotetraploids, *Genetics* 172 (2006) 507–517.
- [117] N. Bahrman, H. Thiellement, Parental genome expression in synthetic wheats (*Triticum turgidum* sp. $\times$ *T. tauschii* sp.) revealed by two-dimensional

- electrophoresis of seedling proteins, *Theor. Appl. Genet.* 74 (1987) 218–223.
- [118] N. Islam, H. Tsujimoto, H. Hirano, Proteome analysis of diploid, tetraploid and hexaploid wheat: towards understanding genome interaction in protein expression, *Proteomics* 3 (2003) 549–557.
- [119] D.W.K. Ng, C. Zhang, M. Miller, Z. Shen, S.P. Briggs, Z.J. Chen, Proteomic divergence in *Arabidopsis* autopolyploids and allopolyploids and their progenitors, *Heredity* 108 (2012) 419–430.
- [120] B. McClintock, The significance of responses of the genome to challenge, *Science* 226 (1984) 792–801.
- [121] M.J. Hegarty, G.L. Barker, I.D. Wilson, R.J. Abbott, K.J. Edwards, S.J. Hiscock, Transcriptome shock after interspecific hybridization in *Senecio* is ameliorated by genome duplication, *Curr. Biol.* 16 (2006) 1652–1659.
- [122] R.T. Gaeta, J.C. Pires, F. Iniguez-Luy, E. Leon, T.C. Osborn, Genomic changes in resynthesized *Brassica napus* and their effect on gene expression and phenotype, *Plant Cell* 19 (2007) 3403–3417.
- [123] A. Marmagne, P. Brabant, H. Thiellement, K. Alix, Analysis of gene expression in resynthesized *Brassica napus* allotetraploids: transcriptional changes do not explain differential protein regulation, *New Phytol.* 186 (2010) 216–227.
- [124] J. Koh, G. Chen, M.J. Yoo, N. Zhu, D. Dufresne, J.E. Erickson, H. Shao, S. Chen, Comparative proteomic analysis of *Brassica napus* in response to drought stress, *J. Proteome Res.* 14 (2015) 3068–3081.
- [125] H. Jia, M. Shao, Y. He, R. Guan, P. Chu, H. Jiang, Proteome dynamics and

- physiological responses to short-term salt stress in *Brassica napus* leaves, PLoS One, 10 (2015) e0144808.
- [126] F. Kong, S. Mao, J. Jiang, J. Wang, X. Fang, Y. Wang, Proteomic changes in newly synthesized *Brassica napus* allotetraploids and their early generations, Plant Mol. Biol. Report 29 (2011) 927–935.
- [127] X.C. Wang, Q. Li, X. Jin, G.H. Xiao, G.J. Liu, N.J. Liu, Y.M. Qin, Quantitative proteomics and transcriptomics reveal key metabolic processes associated with cotton fiber initiation, J. Proteomics 114 (2015) 16–27.
- [128] A.L. Gancel, J. Grimplet, F.X. Sauvage, P. Ollitrault, E.M. Brillouet, Predominant expression of diploid mandarin leaf proteome in two citrus mandarin-derived somatic allotetraploid hybrids, J. Agric. Food Chem. 54 (2006) 6212–6218.
- [129] S. Carpentier, Exploring the potential of genetic diversity via proteomics: past, present, and future perspectives for banana, in: R.M. Ahuja, M.S. Jain (Eds.), Genetic diversity and erosion in plants: indicators and prevention, Springer International Publishing, Switzerland, 2015, pp. 311–323.
- [130] A.C. Vanhove, W. Vermaelen, R. Swennen, S.C. Carpentier, A look behind the screens: characterization of the HSP70 family during osmotic stress in a non-model crop, J. Proteomics 119 (2015) 10–20.
- [131] E. Bancel, T. Bonnot, M. Davanture, G. Branlard, M. Zivy, P. Martre, Proteomic approach to identify nuclear proteins in wheat grain, J. Proteome Res. 14 (2015) 4432–4439.
- [132] P.O. Havanapan, A. Bourchookarn, A.J. Kettermann, C. Krittanai, Comparative

- proteome analysis of rubber latex serum from pathogenic fungi tolerant and susceptible rubber tree (*Hevea brasiliensis*), J. Proteomics 131 (2016) 82–92.
- [133] L. Dai, G. Kang, Z. Nie, Y. Li, R. Zeng, Comparative proteomic analysis of latex from *Hevea brasiliensis* treated with Ethrel and methyl jasmonate using iTRAQ-coupled two-dimensional LC-MS/MS, J. Proteomics 132 (2016) 167–75.
- [134] A.Y.A. Rahman, A.O. Usharraj, B.B. Misra, G.P. Thottathil, K. Jayasekaran, Y. Feng, S. Hou, S.Y. Ong, F.L. Ng, L.S. Lee, H.S. Tan, M.K.L.M. Sakaff, B.S. Teh, B.F. Khoo, S.S. Badai, N.A. Aziz, A. Yuryev, B. Knudsen, A. Dionne-Laporte, N.P. Mchunu, et al., Draft genome sequence of the rubber tree *Hevea brasiliensis*, BMC Genomics 14 (2013) 1–15.
- [135] P.F. Wong, S. Abubakar, Post-germination changes in *Hevea brasiliensis* seeds proteome, Plant Sci. 169 (2005) 303–311.
- [136] L. Dai, G. Kang, Y. Li, Z. Nie, C. Duan, R. Zeng, In-depth proteome analysis of the rubber particle of *Hevea brasiliensis* (para rubber tree), Plant Mol. Biol. 82 (2013) 155–168.
- [137] Q. Xiang, K. Xia, L. Dai, G. Kang, Y. Li, Z. Nie, C. Duan, R. Zeng, Proteome analysis of the large and the small rubber particles of *Hevea brasiliensis* using 2D-DIGE, Plant Physiol. Biochem. 60 (2012) 207–213.
- [138] X. Wang, M. Shi, X. Lu, R. Ma, C. Wu, A. Guo, M. Peng, W. Tian, A method for protein extraction from different subcellular fractions of laticifer latex in *Hevea brasiliensis* compatible with 2-DE and MS, Proteome Sci. 8 (2010) 1–10.
- [139] F. An, J. Fan, J. Li, Q.X. Li, K. Li, W. Zhu, F. Wen, L.J.C.B. Carvalho, S. Chen,

- Comparison of leaf proteomes of cassava (*Manihot esculenta* Crantz) cultivar NZ199 diploid and autotetraploid genotypes, PLoS One, 9 (2014) e85991.
- [140] A. Podda, G. Checcucci, W. Mouhaya, D. Centeno, V. Rofidal, R. Del Carratore, F. Luro, R. Morillon, P. Ollitrault, B.E. Maserti, Salt-stress induced changes in the leaf proteome of diploid and tetraploid mandarins with contrasting Na⁺ and Cl⁻ accumulation behaviour, J. Plant Physiol. 170 (2013) 1101–1112.
- [141] Z. Wang, M. Wang, L. Liu, F. Meng, Physiological and proteomic responses of diploid and tetraploid black locust (*Robinia pseudoacacia* L.) subjected to salt stress, Int. J. Mol. Sci. 14 (2013) 20299–20325.
- [142] C.M. Pacheco, M.C. Pestana-Calsa, F.C. Gozzo, R.J. Mansur Custodio Nogueira, M. Menossi, T. Calsa, Differentially delayed root proteome responses to salt stress in sugar cane varieties, J. Proteome Res. 12 (2013) 5681–5695.
- [143] A.S. Heringer, T. Barroso, A.F. Macedo, C. Santa-Catarina, G.H.M.F. Souza, E.I.S. Floh, G.A. de Souza-Filho, V. Silveira, Label-free quantitative proteomics of embryogenic and non-embryogenic callus during sugarcane somatic embryogenesis, PLoS One 10 (2015) e0127803.
- [144] J. Koh, P.S. Soltis, D.E. Soltis, Homeolog loss and expression changes in natural populations of the recently and repeatedly formed allotetraploid *Tragopogon mirus* (Asteraceae), BMC Genomics 11 (2010) 97.
- [145] J. Wang, L. Tian, H.S. Lee, Z.J. Chen, Nonadditive regulation of *FRI* and *FLC* loci mediates flowering-time variation in *Arabidopsis* allopolyploids, Genetics 173 (2006) 965–974.

- [146] Z. Ni, E.D. Kim, M. Ha, E. Lackey, J. Liu, Y. Zhang, Q. Sun, Z.J. Chen, Altered circadian rhythms regulate growth vigor in hybrids and allopolyploids, *Nature* 457 (2009) 327–331.
- [147] A. Braszewska-Zalewska, A. Dziurlikowska, J. Maluszynska, Histone H3 methylation patterns in *Brassica nigra*, *Brassica juncea*, and *Brassica carinata* species, *Genome* 55 (2011) 68–74.
- [148] C. Vogel, E.M. Marcotte, Insights into the regulation of protein abundance from proteomic and transcriptomic analyses, *Nat. Rev. Genet.* 13 (2012) 227–232.
- [149] L. Anderson, J. Seilhamer, A comparison of selected mRNA and protein abundances in human liver, *Electrophoresis* 18 (1997) 533–537.
- [150] S.P. Gygi, B. Rist, S.A. Gerber, F. Turecek, M.H. Gelb, R. Aebersold, Quantitative analysis of complex protein mixtures using isotope-coded affinity tags, *Nat. Biotechnol.* 17 (1999) 994–999.
- [151] T. Maier, M. Güell, L. Serrano, Correlation of mRNA and protein in complex biological samples, *FEBS Lett.* 583 (2009) 3966–73.
- [152] B. Schwanhauser, D. Busse, N. Li, G. Dittmar, J. Schuchhardt, J. Wolf, W. Chen, M. Selbach, Global quantification of mammalian gene expression control, *Nature* 473 (2011) 337–342.
- [153] M. Hegarty, S. Hiscock, Polyploidy: doubling up for evolutionary success, *Curr. Biol.* 17 (2007) R927– R929.
- [154] L. Li, Y. Huang, X. Xia, Z. Sun, Preferential duplication in the sparse part of yeast protein interaction network, *Mol. Biol. Evol.* 23 (2006) 2467–2473.

- [155] R.A. Veitia, S. Bottani, J.A. Birchler, Cellular reactions to gene dosage imbalance: genomic, transcriptomic and proteomic effects, *Trends Genet.* 24 (2008) 390–397.
- [156] S.E. Stevenson, Y. Chu, P. Ozias-Akins, J.J. Thelen, Validation of gel-free, label-free quantitative proteomics approaches: applications for seed allergen profiling, *J. Proteomics* 72 (2009) 555–566.
- [157] B. Chalhoub, F. Denoeud, S. Liu, I.A.P. Parkin, H. Tang, X. Wang, J. Chiquet, H. Belcram, C. Tong, B. Samans, M. Corra, C. Da Silva, J. Just, C. Falentin, C.S. Koh, I. Le Clainche, M. Bernard, P. Bento, B. Noel, K. Labadie, et al., Early allopolyploid evolution in the post-Neolithic *Brassica napus* oilseed genome, *Science* 345 (2014) 950–953.
- [158] J.E. Coate, H. Bar, J.J. Doyle, Extensive translational regulation of gene expression in an Allopolyploid (*Glycine dolichocarpa*), *Plant Cell* 26 (2014) 136–150.
- [159] W.B. Barbazuk, Y. Fu, K.M. McGinnis, Genome-wide analyses of alternative splicing in plants: opportunities and challenges, *Genome Res.* 18 (2008) 1381–1392.
- [160] N.H. Syed, M. Kalyna, Y. Marquez, A. Barta, J.W.S. Brown, Alternative splicing in plants – coming of age, *Trends Plant Sci.* 17 (2012) 616–623.
- [161] A.S.N. Reddy, Y. Marquez, M. Kalyna, A. Barta, Complexity of the alternative splicing landscape in plants, *Plant Cell* 25 (2013) 3657–83.
- [162] Y. Marquez, J.W.S. Brown, C. Simpson, A. Barta, M. Kalyna, Transcriptome survey reveals increased complexity of the alternative splicing landscape in

- Arabidopsis*, Genome Res. 22 (2012) 1184–1195.
- [163] M.L. Roose, L.D. Gottlieb, Genetic and biochemical consequences of polyploidy in *Tragopogon*, Evolution 30 (1976) 818–830.
- [164] M.L. Roose, L.D. Gottlieb, Biochemical properties and level of expression of alcohol dehydrogenases in the allotetraploid plant *Tragopogon miscellus* and its diploid progenitors, Biochem. Genet. 18 (1980) 1065–1085.
- [165] L. Bianco, L. Lopez, A.G. Scalone, M. Di Carli, A. Desiderio, E. Benvenuto, G. Perrotta, Strawberry proteome characterization and its regulation during fruit ripening and in different genotypes, J. Proteomics 72 (2009) 586–607.
- [166] N. Lu, Z. Xu, B. Meng, Y. Sun, J. Zhang, S. Wang, Y. Li, Proteomic analysis of etiolated juvenile tetraploid *Robinia pseudoacacia* branches during different cutting periods, Int. J. Mol. Sci. 15 (2014) 6674–6688.

Figure 1

Figure 2