A novel allele of *ASY3* promotes meiotic stability in autotetraploid *Arabidopsis lyrata*

Paul J. Seear[a], Martin G. France[a], Catherine L. Gregory[a], Darren Heavens[e], Roswitha Schmickl[c, d], Levi Yant[b, *], James D. Higgins[a,*].

a. Department of Genetics and Genome Biology, Adrian Building, University Road, University of Leicester, Leicester, LE1 7RH, UK

b. Future Food Beacon of Excellence and the School of Life Sciences, University of Nottingham, Nottingham, UK

c. Department of Botany, Faculty of Science, Charles University, Benátská 2, 128 01 Prague, Czech Republic

d. Institute of Botany, The Czech Academy of Sciences, Zámek 1, 252 43 Průhonice, Czech Republic

e. Earlham Institute, Norwich Research Park Innovation Centre, Colney Lane, Norwich NR4 7UZ

* **Authors for correspondence:** jh555@leicester.ac.uk and levi.yant@nottingham.ac.uk

Keywords: recombination; population genetics; chromosome axis; introgression.

Abstract

Whole genome duplication promotes adaptability but is a dramatic mutation, usually resulting in meiotic catastrophe and genome instability. Here we focus on a rapid case of coordinated stabilization of male meiosis in autotetraploid *Arabidopsis lyrata/Arabidopsis arenosa* populations. We fuse population genomic data with a genotype-phenotype association study, concentrating on the effects of haplotype variation of eight meiosis genes that have undergone selective sweeps in the autotetraploids. Our analysis indicates that a novel allele of the meiotic chromosome axis protein *Asynapsis 3* (the orthologue of budding yeast RED1), is the major determinant of meiotic stability. In addition, gene conversion and widespread allelic chimerism in associated genes *ASY1, PDS5b, PRD3,* and *ZYP1a/b* between alleles originating from both species, as well as between ploidies, indicates that this interacting group of rapidly evolving structurally variable haplotypes provide precise control over meiotic recombination, the very process that gave rise to them.

Introduction

Polyploidy occurs in all eukaryotic kingdoms, and is associated with adaptability, speciation and evolvability (1, 2). At the same time, it is also one of the most dramatic mutations observed, usually resulting in catastrophic problems during meiosis, when ensuring stable chromosome segregation and genome integrity is paramount (3). Because efficient meiosis is required for the formation of euploid gametes during sexual reproduction, selection acts strongly to optimize meiosis gene functions immediately upon the whole genome duplication that gives rise to polyploidy.

In allopolyploids (formed by both genome duplication and interspecies hybridization), loci required for correct chromosome pairing have been identified in wheat (4), oil seed rape (5, 6) and *Arabidopsis suecica* (7). However, in autopolyploids (which form within-species, without hybridization), there has been no functional confirmation of any gene controlling correct chromosome pairing, synapsis and crossing over (CO), although we have detected clear signatures of extreme selection in eight meiosis genes associated with the synaptonemal complex (SC) (*ASY1, ASY3, PDS5b, PRD3, REC8, SMC3, ZYP1a, ZYP1b*) in the young autopolyploid *Arabidopsis arenosa* (8).

The SC is a tripartite protein structure consisting of two lateral elements and a central element, specific to meiotic prophase I that is required for normal levels of COs in the majority of sexually reproducing eukaroytes (9). In Arabidopsis, the chromosome axes (which come to form the SC lateral elements) consist of a scaffold of cohesin proteins (SMC1, SMC3, PDS5, REC8 and SCC3)(10-14) that organise sister chromatids into a loop/base conformation (15). PRD3, the budding yeast MER2 homolog, is required for double-strand break (DSB) formation and is not an SC protein *per se* but may juxtapose the DSB site with the chromosome axis to promote inter-homolog recombination (16, 17). In *Sordaria*, MER2 also transfers and releases recombination complexes to and from the SC central region (16). The meiosis specific proteins ASY1, ASY3 and ASY4 load onto the cohesin scaffold, promoting inter-homolog recombination (18-20). ASY1 and ASY3 are the functional homologs of budding yeast HOP1 and RED1 and, HORMAD1/2 and SCP2 in mammals, that facilitate correct chromosome pairing and synapsis, required for wild-type COs (19, 21, 22). In Arabidopsis, synapsis is

initiated by installation of the transverse filament proteins ZYP1a/b between homologous chromosomes, thus ensuring appropriate levels of COs (23).

The eight meiosis loci displaying highly differentiated alleles in *A. arenosa* were also reported in *A. lyrata* autotetraploids, suggesting extensive bidirectional gene flow between these autotetraploids in an Austrian hybrid zone (24). At meiotic metaphase I in *A. arenosa*, chiasma frequency was reduced in autotetraploids carrying the derived alleles, indicating an ongoing adaptive consequence of their evolution (8). However, thus far the relative importance of each evolved allele to meiotic stabilisation in the hybrid tetraploids is not understood.

In this study we fuse genomic, genetic and cytological approaches to investigate the effects of rapidly evolved adaptive haplotypes in these meiosis genes under strong selection. We measure the consequences of alternative evolved haplotypes at these loci in autotetraploid *A. lyrata, A. arenosa,* and natural introgressants across a hybrid zone. Our analysis reveals functional evidence of novel *ASY3* haplotypes that modulate meiotic recombination in both *A. lyrata* and *A. arenosa* autotetraploids, thus stabilising chromosome bivalent formation and genomic integrity.

Results

Association of haplotypes with meiotic stability in A. lyrata/A. arenosa autotetraploids

We first performed a thorough examination of allele frequency and meiotic stability in 52 plants sampled from tetraploid populations LIC, MOD, PIL, SCB, KAG, ROK (*A. lyrata* and *A. lyrata*-like hybrids), TBG, SEN and WEK (*A. arenosa* dominant/*A. lyrata* introgressants) covering a range of known genomic backgrounds and demographic histories (24). We obtained individual maternal seed lines from the East Austrian Forealps and the Wachau region (Supplemental Figure 1). Meiotic stability was assessed in each line by cytological analysis of metaphase I (MI) chromosome configurations. We scored MI nuclei as stable when 16 individual bivalents could be observed aligned on the MI plate and unstable if univalents, trivalents, quadrivalents or multivalents were observed (Figure 1a-c,

Supplemental Figure 2). Furthermore, we scored chiasmata (the cytological manifestation of COs) as distal, interstitial or proximal to the centromere (25) (Figure 1D). In unstable nuclei only a proportion of individual bivalents could be scored per nucleus (ranging from 1-11 per cell), so scorable bivalents were randomly grouped into pseudo-nuclei of 16 bivalents for further analysis. Overall, bivalents from diploid MI nuclei contained significantly more chiasmata than the stable and unstable tetraploids (1.52±0.3, n=312; versus 1.12±0.2, n=960 and 1.26±0.3, n=590, respectively, Mann Whitney Test, $P < 0.001$). The frequency of distal chiasmata was not significantly different between the diploid and stable tetraploids, but was reduced in the unstable tetraploids. Bivalents from stable tetraploid nuclei had significantly fewer interstitial and proximal chiasmata compared to bivalents from diploids and those from unstable nuclei, whereas interstitial and proximal chiasmata were not significantly different between bivalents from diploids and unstable tetraploid nuclei. For each maternal line we scored blind the percentage of stable versus unstable nuclei, revealing a range from 20-100% (Figure 1E, Supplementary figure 2).

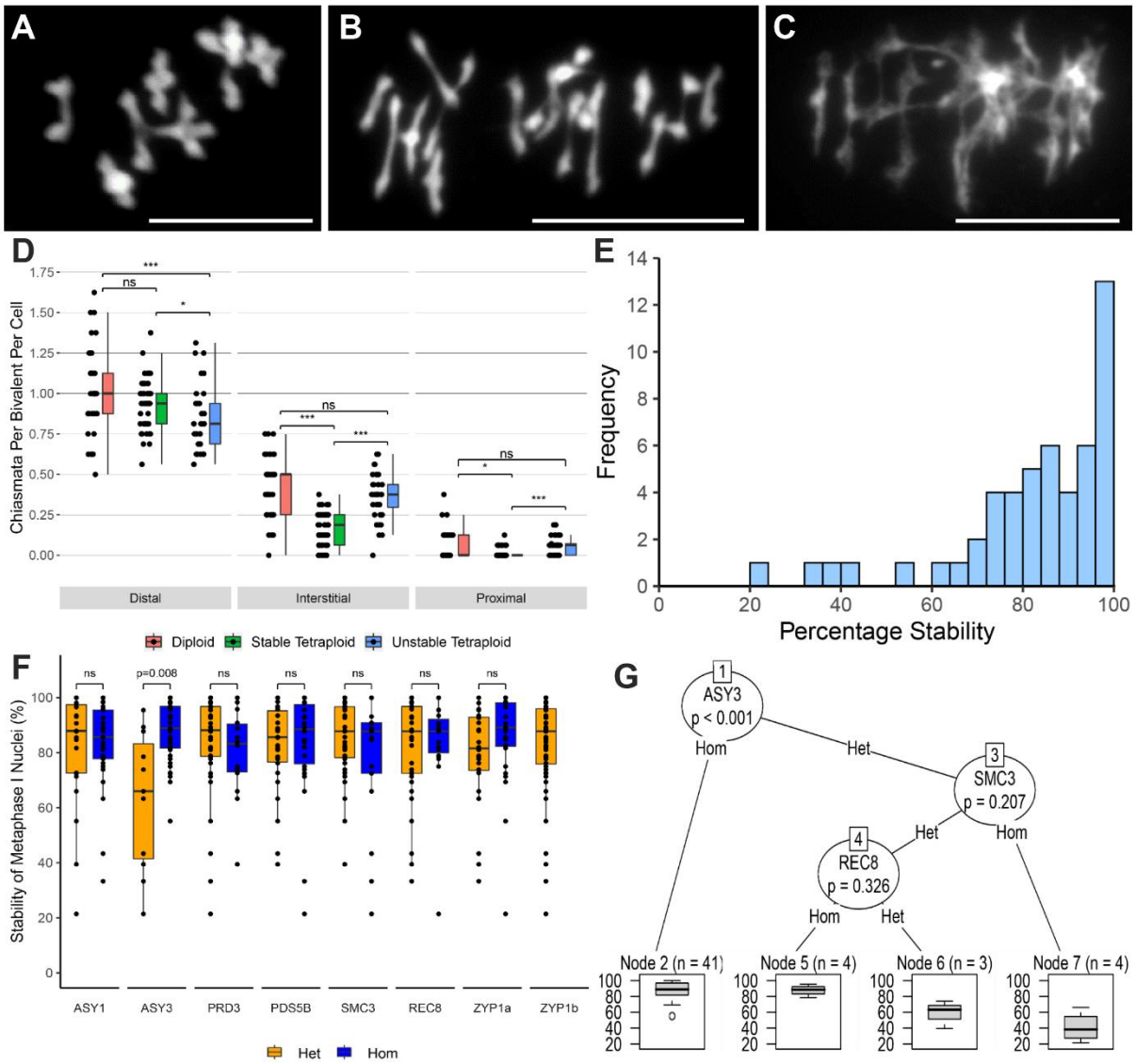

Figure 1. Meiotic adaptation to whole genome duplication in *A. lyrata* autotetraploids. (A) Diploid *A. lyrata* at meiotic metaphase I, exhibiting proximal, interstitial and distal chiasmata. (B) Stable autotetraploid meiotic metaphase I, exhibiting primarily rod bivalents with distal chiasmata. (C) Unstable autotetraploid meiotic metaphase I, exhibiting a multivalent. (D) Comparison of chiasmata position and frequency of diploid (red) and stable (green) or unstable (blue) autotetraploid *A. lyrata*. (E) Frequency of meiotic stability in individual plants from all populations. (F) Meiotic stability of nuclei harbouring either homozygous (blue) or heterozygous (yellow) genotypes for all tested SC genes at metaphase I. (G) Decision tree model calculating probability of *SMC3* and *REC8* states relating to the percentage of meiotic stable nuclei in *ASY3* heterozygotes. Bars for A-C = 10µm.

Individual plants were then genotyped for the proportion of each meiosis gene haplotype by high-throughput sequencing. Initially, population genomic data (26) was used to design degenerate primers to amplify coding regions of meiosis genes from floral bud cDNA that were cloned and sequenced from diploid *A. arenosa* SNO and *A. lyrata* PER populations, as well as tetraploid *A. arenosa*,

SEN, TBG and WEK and *A. lyrata* KAG and MAU. This approach provided accurate sequence polymorphism data for 320 contiguous coding regions from the 8 meiosis genes, as well as validating primers for full gene amplification (exons and introns) for genotyping haplotypes. Full gene amplicons of the 8 meiosis genes from these 52 cytologically analysed plants were used to construct Nextera LITE libraries. Libraries were barcoded per plant and sequenced by MiSeq, generating an average sequence depth across all loci of >2000x, from which we determined the proportion of each allele per plant by SNP frequency. Genotype data was integrated with the cytological analysis (Supplemental Table 1), revealing that only the meiotic axis gene *ASY3* had an effect on meiotic stability (Figure 1F). Plants that were heterozygous for the *ASY3* ancestral diploid allele and derived tetraploid allele were significantly more meiotically unstable, compared to the homozygous derived allele (4n Hom $\tilde{x}$ = 88.9, IQR = 15.1, n= 41, $\tilde{x}$ = 66, IQR =41.7, n=11, Bonferroni adjusted Wilcox Test p=0.008). Given that no effect was observed for the other genes alone and there was a large range of meiotic stability within the *ASY3* heterozygotes, we performed a conditional inference tree analysis to recursively partition the data using Bonferroni corrected permutation tests to determine when to split the data. The results again showed that the derived *ASY3* allele was predictive of meiotic stability. Furthermore, no other splits could be made at a cut-off of p=0.05, but *ASY3* heterozygotes could be further divided depending on their alleles, first for *SMC3* and then *REC8* (Figure 1G).

Meiosis gene flow between A. lyrata/A. arenosa tetraploid populations

We next split the tetraploid populations into *A. lyrata* dominant (LIC, MOD, SCB, KAG, PIL and ROK) or *A. arenosa* dominant (WEK, SEN and TBG), based on a detailed demographic analysis (24). The 320 coding sequences acquired above provided high quality reference data to determine allelic origin and infer direction of gene flow. The *ASY3* allele associated with meiotic stability clusters with the *A. lyrata* diploid sequence (Supplemental Figure 3a, b), as does *PDS5b*. Conversely, *ASY1, PRD3, REC8, SMC3* and *ZYP1a/b* had highest homology with diploid *A. arenosa*. The adaptive *A. lyrata ASY3* allele is

present at 99% frequency in tetraploid *A. arenosa* populations tested and 95% in tetraploid *A. lyrata* populations (Figure 2A). The residual percentages were due to the presence of diploid *arenosa* or *lyrata* alleles in these populations that were observed for all genes except *ZYP1a* in *A. arenosa* that was completely homozygous. In contrast, the *ASY1* allele derived from diploid *A. arenosa* had a frequency of 94% in *A. arenosa* tetraploid and 93% in *A. lyrata* tetraploid, indicating bidirectional gene flow (Figure 2A). We also see evidence of bidirectional gene flow of diploid alleles, particularly for *SMC3* and *REC8* (Figure 2A).

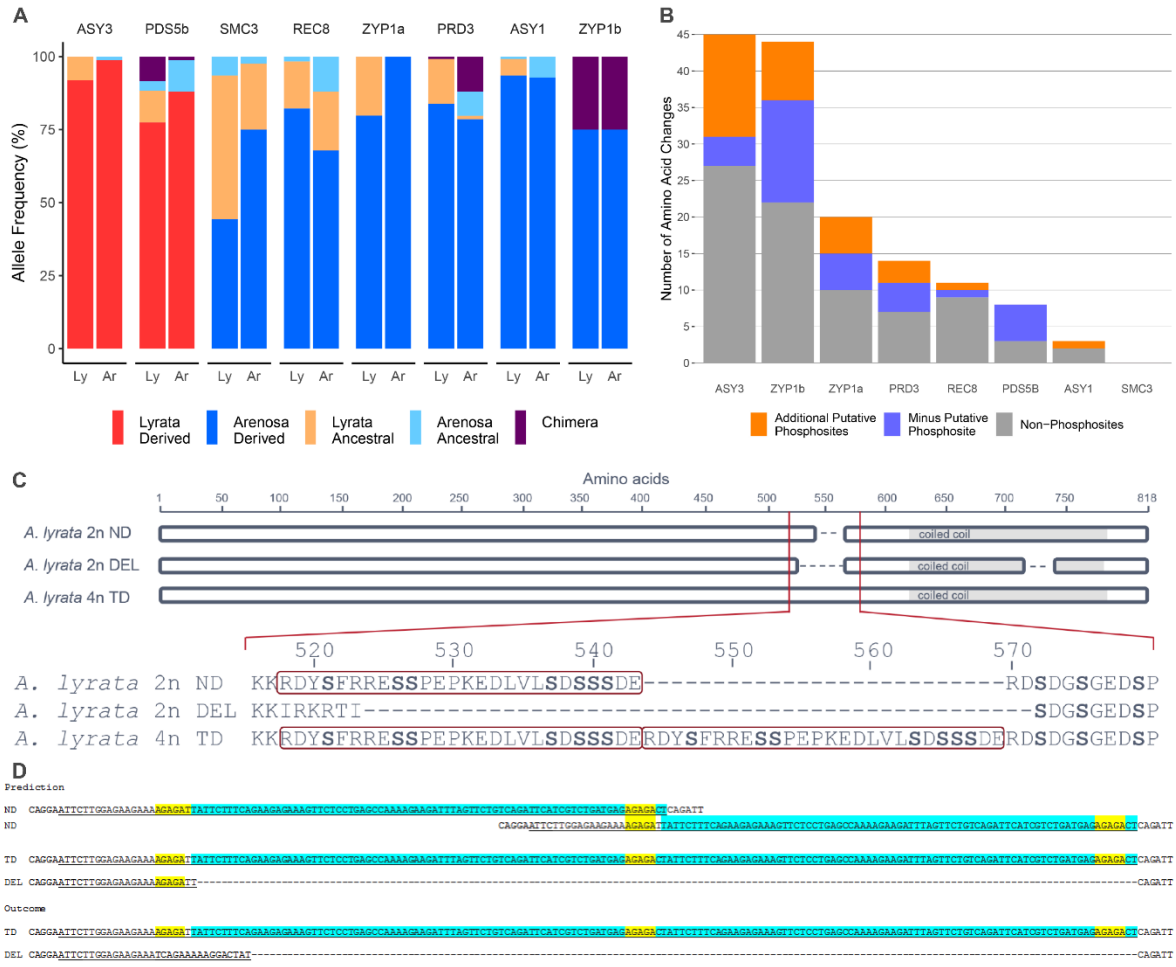

Figure 2. Contrasting origins and polymorphisms of adaptive meiosis gene haplotypes in tetraploid introgressants of *A. lyrata*/*A. arenosa*. (A) Haplotype frequencies of adaptive meiosis alleles from *A. lyrata* (*Ly*) and *A. arenosa* (*Ar*) populations. (B) Conserved amino acid polymorphisms in autotetraploids compared to ancestral diploid alleles including gains and losses of predicted phosphosites. (C) Structural variants of the three *ASY3* haplotypes with the non-duplicated (*ND*) ancestral allele showing the serine-rich region (red box) with serines highlighted in bold, the deletion (*DEL*) allele lacking the serine-rich region and the derived tandem duplicated (*TD*) allele possessing two serine-rich regions (red boxes) and putative SUMO sites at K517, K531 and K556. (D) Major structural variation at the ASY3 locus suggesting that microhomology (highlighted yellow) between the ancestral alleles may have led to the formation of the *TD* allele (underline=exon, blue=region of duplication).

Adaptive polymorphisms in meiosis genes

All previous studies aligned short read sequencing to the *A. lyrata* reference to infer polymorphic amino acid changes between diploids and tetraploids in *A. arenosa* and *A. lyrata,* and could not infer contiguous autotetraploid haplotypes (8, 24, 27). To overcome this and resolve individual haplotypes, we cloned the coding regions of all 8 meiosis genes from representative populations and performed Sanger sequencing, identifying 320 alleles. This allowed us to detect both structural variation and more divergent SNP variation (Figure 2B, 2C, and Supplemental Tables 2-4). The encoded proteins are largely structural, forming the physically interacting synaptonemal complex and associated proteins, so it is difficult to infer functional or non-functional amino acid polymorphisms with perfect confidence, but overall 45% were either loss or gain of putative phosphosites and 55% were non-phosphosites. ASY3 and ZYP1b exhibited the greatest quantity of residue changes conserved amongst populations (45 and 44, respectively) and SMC3 none (although there were population specific SMC3 polymorphisms). Of the 45 polymorphic residues for ASY3, 27 were due to a tandem duplication (*TD* allele) in a serine-rich region of the protein upstream of the coiled-coil domain, possessing putative ATM/CKII phosphosites and a predicted SUMO site (K556, GPS-SUMO), compared to the ancestral *A. lyrata* non-duplicated allele (*ND*) (Figure 2C). As the *TD* allele had highest sequence similarity to the diploid *A. lyrata ND* allele, to better understand its origin, we screened the diploid *A. lyrata* population (PER) geographically adjacent to the *A. lyrata* autotetraploids LIC and MOD (Supplemental Figure 1) for the *TD* allele. One hundred and twenty-eight plants from PER were screened indicating the current absence (or vanishingly low frequency) of the *TD* allele, but did reveal the presence of a deletion (DEL) allele at 7% frequency, where the entire serine-rich region is absent. Genomic DNA from *ASY3 ND*, *TD* and *DEL* was cloned, sequenced, aligned, and compared to the coding regions. A 78bp region of exon 2 is duplicated in-frame in the TD allele that is missing in the DEL allele between two AGAGA sites (Figure 2D).

Chimeric novel meiosis alleles

Our analysis identified a novel chimeric allele of *ZYP1b* in all tetraploid populations; *PRD3* in SCB, SEN and WEK; and *PDS5b* in MOD, ROK and WEK. At the 3' end of all *ZYP1b* alleles we detect a 474bp gene conversion (GC) to *ZYP1a* (Supplemental Figure 4). We also find evidence of GC in *PRD3* between *A. arenosa* and *A. lyrata* ancestral diploid alleles. In one of the diploid *PRD3* alleles in the WEK population, the first 740bp is more similar to the diploid *A. arenosa* than to *A. lyrata* (7 vs 21 SNPs), while the remaining 625bp of coding sequence has a higher similarity to diploid *A. lyrata* than to *A. arenosa* (5 vs 24 SNPs) (Supplemental Figure 4). In *PDS5b,* the 3' sequence of some diploid alleles are converted to the tetraploid-favoured allele providing evidence of GC (or CO) between ploidy levels. In addition, analysis of whole genome resequencing data (24) revealed a 3' GC from a diploid *ASY1 A. lyrata* allele to a tetraploid *A. arenosa* allele in the KAG population (Supplemental Figure 4). The widespread presence of such evidence of gene conversion products in these loci exhibiting the most dramatic signatures of selection suggests a mechanism by which the peaks of differentiation generally found in this system are so narrow (3, 8, 24, 26, 28), despite a recent origin of the tetraploids (24).

Discussion

Here we aimed to determine the impact of strongly selected meiosis haplotypes that underwent recent selective sweeps on the rapid evolution of autotetraploid meiotic stability and to trace their evolutionary origin. By associating genotypic and cytological phenotypic data we provide evidence that *ASY3* is the major locus currently stabilising autotetraploid male meiosis in these populations. We identified structural variation of meiosis allele haplotypes including a novel derived, tandemly duplicated (TD) *ASY3* allele that underlies the stable chromosome meiotic phenotype in the tetraploids, as well as novel *ASY1, PDS5b, PRD3* and *ZYP1b* chimeric alleles between diploids and tetraploids and *A. arenosa* and *A. lyrata*.

A cytological MI analysis revealed that chiasmata in the autotetraploids were significantly reduced in both stable and unstable nuclei compared to diploid *A. lyrata*. Moreover, chiasmata frequencies in meiotically stable nuclei were significantly reduced in regions proximal and interstitial to the centromere. A shift in chiasma distribution may reflect a fundamental mechanism for meiotic adaptation to autopolyploidy (29). The meiotically unstable nuclei occur due to unregulated meiotic recombination between multiple chromosomes, either homologous or non-homologous, as previously observed in MI rDNA FISH experiments (24, 30).

Our genotype-phenotype association study revealed that the allele state of the structurally variable meiotic chromosome axis protein ASY3 was the major factor governing whether nuclei were stable. We hypothesise that the ancestral diploid *ASY3 ND* allele promotes high levels of proximal and interstitial chiasmata, but in the tetraploid it acts dominantly over the evolved *TD* allele, promoting interstitial and proximal chiasmata as well as multivalents. Such multivalents have been observed in diploid *A. thaliana ZYP1*^{RNAi} lines where the authors postulated that chiasmata may have formed between extensive duplications on non-homologous chromosomes (23). The tandemly duplicated serine-rich region in the TD allele may function in a manner similar to the budding yeast N-terminus MSH4 degron in destabilizing the protein (31), thus creating a hypomorph. The analysed brassica SC phosphoproteome (32) did not recover peptides for ASY3 in the serine-rich region, although similar (serine-aspartic acid) residues recovered from ASY1 were phosphorylated. The *TD* allele also contains 19 derived residues outside the serine-rich region, of which 10 are predicted phosphosite gains or losses, that cannot be ruled out as functionally important along with unknown *trans* effects. Chiasmata are distalized in axis mutants *asy1, asy3* and *asy4*, presumably due to telomere proximity enabling sufficient inter-homolog pairing, whereas a complete meiotic axis is required to promote high levels of recombination between spatially separated regions in nuclei along the arms of the chromosomes (19, 20, 33). Our data suggest a model wherein the *TD* allele is hypomorphic, thus distalizing chiasmata, whereas when heterozygous with the ancestral *ND* allele, rates of inter-homolog and non-homologous recombination increase. We hypothesise that axis components that favour interstitial

and proximal recombination in diploids promote associations with non-homologous chromosomes in the tetraploids, especially in regions with high sequence homology. However, for stable bivalents, once a distal CO site is designated, interference may prevent further COs forming. In budding yeast, the TopoII interference pathway requires SUMOylation of TopoII and Red1, the ASY3 orthologue (22, 34). The serine-rich duplicated region in ASY3 possesses a predicted SUMO site, which could play a role in protein function, although this hypothesis needs functional testing.

In *ASY3 ND/TD* heterozygotes, we observed a trend for alleles of two cohesins to influence meiotic stability; these are the *structural maintenance of chromosomes 3* (*SMC3*) and the α-kleisin *REC8*. Our sample size was limited by the very low frequency of *ASY3* heterozygotes in the wild populations, so the effect of the cohesins was not statistically significant, although suggestive. SMC3 and REC8 loading onto meiotic chromosomes is upstream of ASY3 and required for normal patterning of ASY3 on the meiotic axis (15). REC8 and SMC3 could alter loading of ASY3 onto the chromosome axes, thus promoting a similar outcome to the TD homozygote. Alternatively, SMC3 and REC8 may differentially organise the chromosome loop axis arrangement, thus influencing inter-homolog recombination (15).

Phylogenetic analysis revealed that the *ASY3 TD* allele most likely originated from diploid *A. lyrata*, consistent with (24). A screen of the diploid *A. lyrata* population geographically adjacent to the tetraploids for the *ASY3 TD* did not detect the allele, but instead identified one with the same region deleted (*DEL*). The *DEL* allele coding region is in-frame and contains a second deletion in the coiled-coil. It may be a coincidence that the same region is lost and gained in the *ASY3* alleles, or that this sequence is more susceptible to DNA replicative errors. The region of interest is flanked by DNA microhomology (AGAGA) that is positioned at the putative exchange points and may have led to DNA polymerase slippage or aberrant replication fork repair. It is unlikely to have occurred during meiotic recombination due to the low level of sequence homology. However, we observe numerous examples of gene conversion that appear to have arisen during meiotic recombination in *ASY1, PRD3, PDS5b* and *ZYP1a/b* between both species and ploidies. The bidirectional gene flow between *A. arenosa* and

A. lyrata tetraploids has enlarged the gene pool for beneficial alleles to be borrowed and selected upon and the novel GC chimeric alleles may precisely coalesce advantageous sequences from differing origins for adaptation, although this would require further testing with an even larger sample size.

The *A. lyrata* autotetraploid populations contain plants with variable levels of meiotic stability as well as relatively high frequencies of diploid-like alleles. The diploid alleles may persist due to: a) continuous gene flow via unreduced gametes (35); b) being beneficial in certain environmental conditions e.g. high altitudes (36); c) being beneficial during female meiosis (COs are reduced in *A. thaliana* female compared to male (37); d) in the case of SMC3, diploid alleles appear beneficial for male meiosis; e) inability to purge genetic load in autotetraploids (38); f) limited effect on overall male pollen fecundity, due to an excess of grains transmitted, despite variable quality.

Taken together, our data indicate multiple mechanisms for rapid meiotic evolution in autotetraploid *A. lyrata*. They reveal a dominant association of a duplication of the serine-rich region in the *ASY3 TD* allele with MI stability. Furthermore, tetraploid *A. lyrata* has introgressed *SMC3* and *REC8* alleles from *A. arenosa* by gene flow (as well as *ZYP1a/b*, *PRD3* and *ASY1*). Finally, novel chimeric genes of *ZYP1a/b, PRD3, PDS5b* and *ASY1* have arisen evidently through gene conversion, suggesting highly dynamic mechanisms to generate variation that may be selected upon by evolution to ensure meiotic success in these populations.

Methods

Cloning and sequencing of meiosis gene transcripts

Plants were grown from seed to obtain fresh flower buds from diploid *A. arenosa* SNO and *A. lyrata* PER populations, and tetraploid *A. arenosa*, SEN, TBG and WEK and *A. lyrata* KAG and MAU populations (24). These buds were collected, flash frozen in liquid nitrogen and stored at -80C until RNA extraction. Total RNA was extracted using a Bioline ISOLATE II RNA Plant Kit (Bioline Ltd, London, UK), following manufacturer's instructions, eluting into a final volume of 60 µl nuclease free water.

Concentration and purity were determined using a NanoDrop spectrophotometer (LabTech International, Lewes, UK) and one microgram of total RNA was electrophoresed on a non-denaturing 1% (w/v) agarose gel to check for degradation. First strand cDNA was reverse transcribed from 0.5 µg of total RNA using a Quantitect Reverse Transcription Kit (Qiagen, Hilden, Germany) that incorporates genomic DNA removal prior to reverse transcription. The coding regions of 8 meiosis genes (*ASY1*, *ASY3*, *PRD3*, *PDS5b*, *REC8*, *SMC3*, *ZYP1a* and *ZYP1b*) were amplified by PCR using 0.2 µM primers (Supplemental Table 5) designed using HiSeq data (26) and Platinum™ *Taq* DNA Polymerase High Fidelity (ThermoFisher Scientific, MA, USA). PCR conditions were as follows: 94 °C for 2 min, followed by 35 cycles of 94 °C for 15 s, 60-65 °C for 30 s and 68 °C for 2-5 min (see Supplemental Table 5), with a final extension of 68 °C for 5-10 min. PCR products were electrophoresed on a 2% (w/v) agarose gel, and single bands of the expected size were excised and purified with a Monarch® DNA Gel Extraction kit (New England Biolabs, MA, USA). Purified PCR products were cloned into pCR-XL-TOPO™ vector using a TOPO™ XL PCR Cloning Kit following the manufacturer's instructions. For each gene a total of 8 clones from each plant were isolated from overnight LB cultures using an ISOLATE II Plasmid Mini Kit (Bioline) prior to sequencing with universal M13F and M13R primers by GATC Biotech (Konstanz, Germany). Nucleotide sequences of the cDNAs were processed in Geneious 11.1.2 (https://www.geneious.com) to remove vector and low-quality sequence before using BLASTN to search the June 2010 (v.1.0/INSDC) assembly of the North American *A. lyrata* reference genome (39) and NCBI nonredundant (nr) database for confirmation that the obtained cDNAs were the expected gene transcripts. Primer walking was then used to sequence the entire length of the transcript. For each meiosis gene, cDNAs from each population were aligned with the respective Ensembl gene predictions from the *A. lyrata* reference genome (Supplemental Table 6), and to act as outgroups, *A. thaliana* transcripts obtained from The Arabidopsis Information Resource (TAIR) using the MUSCLE 3.8.425 plugin with default settings (40). Phylogenetic trees were produced using the Geneious Tree Builder with Jukes-Cantor genetic distance model and Neighbor-Joining tree build method.

Meiotic haplotype genotyping

Genomic DNA was extracted from leaf material of each of the 52 plants in the study using a DNeasy Plant Mini Kit (Qiagen) and eluting into 100 µl nuclease free water. Full length coding regions of each of the 8 meiosis genes (including introns) were amplified from this genomic DNA by PCR using Platinum™ SuperFi™ Green PCR Master Mix (ThermoFisher Scientific) and 0.5 µM primers (Supplemental Table 5) designed against the Sanger sequenced cDNA of the 8 meiosis genes (see above). PCR conditions were as follows: 98 °C for 30 s, followed by 35 cycles of 98 °C for 30 s, 60-63 °C for 10 s and 72 °C for 2.5-10 min (see Supplemental Table 5), with a final extension of 72 °C for 5-10 min. PCR products were electrophoresed on a 1% (w/v) agarose gel, and single bands of the expected size were excised and purified with a Monarch DNA Gel Extraction kit (New England Biolabs). Libraries were constructed using 1ng of input DNA in a Low Input, Transposase Enabled (LITE) pipeline developed at the Earlham Institute (Norwich, UK) and based on the Illumina Nextera kits (Illumina, San Diego, CA, USA) (41). Each library was constructed using unique 9 bp dual index combinations allowing samples to be multiplexed. Pooled libraries were size selected between 600 and 750bp on a BluePippin (Sage Science, Beverly, MA, USA) 1.5% Cassette and then sequenced with a 2x250bp read metric on an Illumina MiSeq sequencer.

MiSeq fastq files were imported into Geneious 11.1.2 (https://www.geneious.com) and R1 and R2 reads paired. Trimming of reads was performed with the BBDuk Adaptor/Quality Trimming v.37.64 plugin with default settings. For each of the 52 individuals, trimmed reads were mapped to each of the 8 meiosis genes (Supplemental Table 6) from the North American *A. lyrata* reference genome (39). SNPs relative to the reference genome genes were then called and used to identify different 2n and 4n alleles for each gene using a set of allele specific SNPs (Supplemental Table 7).

Cloning and sequencing of ASY3 alleles

Genomic DNA was extracted from diploid *A. lyrata* PER plants as above and a short section of *ASY3* was PCR amplified using 0.5 µM primers (Supplemental Table 5) and MyTaq™ Red Mix (Bioline). PCR conditions were as follows: 95 °C for 2 min, followed by 35 cycles of 95 °C for 30 s, 69 °C for 30 s and 72 °C for 30 s, with a final extension of 72 °C for 5 mins. PCR products were electrophoresed on a 2% (w/v) agarose gel and the lower band (~125 bp) corresponding to a partial *ASY3 DEL* allele was excised and purified with a Monarch® DNA Gel Extraction kit (New England Biolabs). Purified PCR products were cloned into pCR™4-TOPO® vector using a TOPO™ TA Cloning™ for Sequencing Kit (ThermoFisher Scientific) following manufacturer's instructions. A total of 12 clones were isolated, purified and Sanger sequenced as above. Nucleotide sequences were processed in Geneious 11.1.2 (https://www.geneious.com) to remove vector and low-quality sequence before aligning with *ASY3 ND*/*TD* transcript alleles as described above. Primers designed against the partial *DEL* sequence were used to obtain the 3' end of the transcript from 3 µg *ASY3 ND/DEL* heterozygous 2n *A. lyrata* (PER) floral bud total RNA using a GeneRacer™ Kit (ThermoFisher Scientific) following manufacturer's instructions. Purified PCR products were cloned into pCR™4-TOPO® vector, sequenced and processed as above.

ASY3 ND, *TD* and *DEL* alleles were PCR amplified from genomic DNA extracted from diploid and tetraploid *A. lyrata* respectively using 0.2 µM primers designed against *ASY3* cDNA sequences obtained above (Supplemental Table 5) and Q5® High-Fidelity DNA Polymerase (New England Biolabs). PCR conditions were as follows: 98 °C for 2 min, followed by 35 cycles of 98 °C for 10 s, 63°C for 30 s and 72 °C for 4 min, with a final extension of 72 °C for 10 min. PCR products were electrophoresed on a 1% (w/v) agarose gel, and single bands of the expected size were excised and purified as above. Purified PCR products were cloned into pDrive (Qiagen) and sequenced by Eurofins Genomics (Ebersberg, Germany).

Cytology

Chromosome spreads were performed (42) (24) on all populations used in this study. Nikon Eclipse Ci and Ni-E microscopes installed with NIS Elements software were used to capture images of chromosomes.

Protein predictions

Protein post-translational predictions were provided by KinasePhos2.0 (43) and NetPhos3.1 (44) and SUMO sites were predicted by GPS-SUMO (45).

Statistical analyses

Statistical analysis and data figures were prepared using R. The conditional inference tree was fitted using the ctree function from the party Package (46), with Bonferroni adjusted p values calculated.

Data availability

All sequences in this study including cDNA transcripts and genomic DNA sequences have been deposited in the DDBJ/EMBL/GenBank databases under accession numbers MN512718 - MN513026 and MN520243 - MN520257. MiSeq amplicon reads have been deposited in the NCBI Sequence Read Archive (SRA; https://www.ncbi.nlm.nih.gov/sra) database under BioProject ID PRJNA575228.

Acknowledgments

We would like to thank Neelam Dave and Maria Spezzati for providing technical support and Chris Franklin, Holger Puchta and Kirsten Bomblies for helpful discussions during the project. DEEPSeq at the University of Nottingham provided MiSeq data. This work was funded by BBSRC New Investigator grant BB/M01973X/1. This work was supported by the European Research Council (ERC) under the

European Union's Horizon 2020 research and innovation programme [grant number ERC-StG 679056 HOTSPOT], via a grant to L.Y.; and the Biotechnology and Biological Sciences Research Council [grant number BB/P013511/1], via a grant to the John Innes Centre (L.Y.).

Author contributions

PJS, CLG, DH, JDH performed lab work. PJS, MGF and JDH analysed data. JDH and LY conceived the project. PJS, MGF, DH, RS, LY and JDH wrote the manuscript.

References

1. K. Alix, P. R. Gerard, T. Schwarzacher, J. S. P. Heslop-Harrison, Polyploidy and interspecific hybridization: partners for adaptation, speciation and evolution in plants. *Ann Bot* **120**, 183-194 (2017).
2. A. M. Selmecki *et al.*, Polyploidy can drive rapid adaptation in yeast. *Nature* **519**, 349-352 (2015).
3. P. Baduel, S. Bray, M. Vallejo-Marin, F. Kolar, L. Yant, The "Polyploid Hop": Shifting Challenges and Opportunities Over the Evolutionary Lifespan of Genome Duplications. *Front Ecol Evol* **6** (2018).
4. M. D. Rey *et al.*, Exploiting the ZIP4 homologue within the wheat Ph1 locus has identified two lines exhibiting homoeologous crossover in wheat-wild relative hybrids. *Mol Breed* **37**, 95 (2017).
5. A. Gonzalo *et al.*, Reducing MSH4 copy number prevents meiotic crossovers between non-homologous chromosomes in Brassica napus. *Nat Commun* **10**, 2354 (2019).
6. E. Jenczewski *et al.*, PrBn, a major gene controlling homeologous pairing in oilseed rape (Brassica napus) haploids. *Genetics* **164**, 645-653 (2003).
7. I. M. Henry *et al.*, The BOY NAMED SUE quantitative trait locus confers increased meiotic stability to an adapted natural allopolyploid of Arabidopsis. *Plant Cell* **26**, 181-194 (2014).
8. L. Yant *et al.*, Meiotic adaptation to genome duplication in Arabidopsis arenosa. *Curr Biol* **23**, 2151-2156 (2013).
9. S. L. Page, R. S. Hawley, The genetics and molecular biology of the synaptonemal complex. *Annu Rev Cell Dev Biol* **20**, 525-558 (2004).
10. W. S. Lam, X. Yang, C. A. Makaroff, Characterization of Arabidopsis thaliana SMC1 and SMC3: evidence that AtSMC3 may function beyond chromosome cohesion. *J Cell Sci* **118**, 3037-3048 (2005).
11. M. Pradillo *et al.*, Involvement of the Cohesin Cofactor PDS5 (SPO76) During Meiosis and DNA Repair in Arabidopsis thaliana. *Front Plant Sci* **6**, 1034 (2015).
12. A. M. Bhatt *et al.*, The DIF1 gene of Arabidopsis is required for meiotic chromosome segregation and belongs to the REC8/RAD21 cohesin gene family. *Plant J* **19**, 463-472 (1999).

13. X. Cai, F. Dong, R. E. Edelmann, C. A. Makaroff, The Arabidopsis SYN1 cohesin protein is required for sister chromatid arm cohesion and homologous chromosome pairing. *J Cell Sci* **116**, 2999-3007 (2003).
14. L. Chelysheva *et al.*, AtREC8 and AtSCC3 are essential to the monopolar orientation of the kinetochores during meiosis. *J Cell Sci* **118**, 4621-4632 (2005).
15. C. Lambing, Tock, A.J., Choi, K., Topp, S.D., Kuo, P.C., Blackwell, A.R., Zhao, X., Osman, K., Higgins, J.D., Franklin, F.C.H., Henderson, I.R, **REC8-cohesin, chromatin and transcription orchestrate meiotic recombination in the Arabidopsis genome**. https://doi.org/10.1101/512400.
16. S. Tesse *et al.*, Asy2/Mer2: an evolutionarily conserved mediator of meiotic recombination, pairing, and global chromosome compaction. *Genes Dev* **31**, 1880-1893 (2017).
17. A. De Muyt *et al.*, A high throughput genetic screen identifies new early meiotic recombination functions in Arabidopsis thaliana. *Plos Genet* **5**, e1000654 (2009).
18. S. J. Armstrong, A. P. Caryl, G. H. Jones, F. C. Franklin, Asy1, a protein required for meiotic chromosome synapsis, localizes to axis-associated chromatin in Arabidopsis and Brassica. *J Cell Sci* **115**, 3645-3655 (2002).
19. M. Ferdous *et al.*, Inter-homolog crossing-over and synapsis in Arabidopsis meiosis are dependent on the chromosome axis protein AtASY3. *Plos Genet* **8**, e1002507 (2012).
20. A. Chambon *et al.*, Identification of ASYNAPTIC4, a Component of the Meiotic Chromosome Axis. *Plant Physiol* **178**, 233-246 (2018).
21. A. P. Caryl, S. J. Armstrong, G. H. Jones, F. C. Franklin, A homologue of the yeast HOP1 gene is inactivated in the Arabidopsis meiotic mutant asy1. *Chromosoma* **109**, 62-71 (2000).
22. A. M. V. West *et al.*, A conserved filamentous assembly underlies the structure of the meiotic chromosome axis. *Elife* **8** (2019).
23. J. D. Higgins, E. Sanchez-Moran, S. J. Armstrong, G. H. Jones, F. C. Franklin, The Arabidopsis synaptonemal complex protein ZYP1 is required for chromosome synapsis and normal fidelity of crossing over. *Genes Dev* **19**, 2488-2500 (2005).
24. S. Marburger, Monnahan, P., Seear, P.J., Martin, S.H., Koch, J., Paajanen, P., Bohutinska, M., Higgins, J.D., Schmickl, R., Yant, L., Interspecific introgression mediates adaptation to whole genome duplication. https://doi.org/10.1101/636019.
25. E. Sanchez Moran, S. J. Armstrong, J. L. Santos, F. C. Franklin, G. H. Jones, Chiasma formation in Arabidopsis thaliana accession Wassileskija and in two meiotic mutants. *Chromosome Res* **9**, 121-128 (2001).
26. P. Monnahan *et al.*, Pervasive population genomic consequences of genome duplication in Arabidopsis arenosa. *Nat Ecol Evol* **3**, 457-468 (2019).
27. K. M. Wright *et al.*, Selection on meiosis genes in diploid and tetraploid Arabidopsis arenosa. *Mol Biol Evol* **32**, 944-955 (2015).
28. B. Arnold, S. T. Kim, K. Bomblies, Single Geographic Origin of a Widespread Autotetraploid Arabidopsis arenosa Lineage Followed by Interploidy Admixture. *Mol Biol Evol* **32**, 1382-1395 (2015).
29. M. H. Hazarika, H. Rees, Genotypic control of chromosome behaviour in rye X. Chromosome pairing and fertility in autotetraploids. *Heredity* **22**, 317-332 (1967).
30. J. D. Higgins, K. M. Wright, K. Bomblies, F. C. H. Franklin, Cytological techniques to analyze meiosis in Arabidopsis arenosa for investigating adaptation to polyploidy. *Front Plant Sci* **4** (2013).
31. W. He, Rao, H.B.D.P., Tang, S., Bhagwat, N., Kulkarni, D.S., Chang, M.A.W., Hall, C., Singh, L., Chen, X., Hollingsworth, N.M., Cejka, P., Hunter, N., The crossover function of MutSγ is activated via Cdc7-dependent stabilization of Msh4. https://doi.org/10.1101/386458.
32. K. Osman *et al.*, Affinity proteomics reveals extensive phosphorylation of the Brassica chromosome axis protein ASY1 and a network of associated proteins at prophase I of meiosis. *Plant J* **93**, 17-33 (2018).

33. E. Sanchez-Moran, J. L. Santos, G. H. Jones, F. C. Franklin, ASY1 mediates AtDMC1-dependent interhomolog recombination during meiosis in Arabidopsis. *Genes Dev* **21**, 2220-2233 (2007).
34. L. Zhang *et al.*, Topoisomerase II mediates meiotic crossover interference. *Nature* **511**, 551-556 (2014).
35. M. H. Jorgensen, D. Ehrich, R. Schmickl, M. A. Koch, A. K. Brysting, Interspecific and interploidal gene flow in Central European Arabidopsis (Brassicaceae). *BMC Evol Biol* **11**, 346 (2011).
36. K. Bomblies, J. D. Higgins, L. Yant, Meiosis evolves: adaptation to external and internal environments. *The New phytologist* **208**, 306-323 (2015).
37. L. Giraut *et al.*, Genome-Wide Crossover Distribution in Arabidopsis thaliana Meiosis Reveals Sex-Specific Patterns along Chromosomes. *Plos Genet* **7** (2011).
38. S. P. Otto, The evolutionary consequences of polyploidy. *Cell* **131**, 452-462 (2007).
39. T. T. Hu *et al.*, The Arabidopsis lyrata genome sequence and the basis of rapid genome size change. *Nat Genet* **43**, 476-+ (2011).
40. R. C. Edgar, MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* **32**, 1792-1797 (2004).
41. M. Mascher *et al.*, A chromosome conformation capture ordered sequence of the barley genome. *Nature* **544**, 427-433 (2017).
42. J. D. Higgins, K. M. Wright, K. Bomblies, F. C. Franklin, Cytological techniques to analyze meiosis in Arabidopsis arenosa for investigating adaptation to polyploidy. *Front Plant Sci* **4**, 546 (2014).
43. Y. H. Wong *et al.*, KinasePhos 2.0: a web server for identifying protein kinase-specific phosphorylation sites based on sequences and coupling patterns. *Nucleic Acids Research* **35**, W588-W594 (2007).
44. N. Blom, T. Sicheritz-Ponten, R. Gupta, S. Gammeltoft, S. Brunak, Prediction of post-translational glycosylation and phosphorylation of proteins from the amino acid sequence. *Proteomics* **4**, 1633-1649 (2004).
45. Q. Zhao *et al.*, GPS-SUMO: a tool for the prediction of sumoylation sites and SUMO-interaction motifs. *Nucleic Acids Research* **42**, W325-W330 (2014).
46. T. Hothorn, K. Hornik, A. Zeileis, Unbiased recursive partitioning: A conditional inference framework. *J Comput Graph Stat* **15**, 651-674 (2006).