Genome duplication increases meiotic recombination frequency

Ou, Fang; Wang, Lin; Zhang, Yuxin; Yang, Jixuan; Tao, Qin; Zhang, Fengjun; Luo, Zewei

DOI:
10.1093/molbev/msaa219

Genome Duplication Increases Meiotic Recombination Frequency: a *Saccharomyces cerevisiae* model

Fang Ou[1], Lin Wang[1], Yuxin Zhang[1], Jixuan Yang[1], Qin Tao[1], Fengjun Zhang[1,3] and Zewei Luo[1,2]

1. Laboratory of Population and Quantitative Genetics, Institute of Biostatistics, Fudan University, Shanghai 200433, China
2. School of Biosciences, University of Birmingham, Birmingham B15 2TT, United Kingdom
3. Qinghai Academy of Agriculture and Forestry Sciences, Xining, Qinghai 810016, China

Running Title: Ploidy-driven increase in meiotic recombination frequency

Key Words: Meiotic recombination frequency, Genome duplication, Tetrasomic linkage analysis, *S. cerevisiae*

* The author to whom all correspondence should be addressed:

Prof. Zewei Luo
School of Biosciences
University of Birmingham
Birmingham B15 2TT, United Kingdom
Email: zwluo@fudan.edu.cn or z.luo@bham.ac.uk
Telephone: +44 121 414 5404
Fax: +44 121 414 5925

Abstract

Genetic recombination characterized by reciprocal exchange of genes on paired homologous chromosomes is the most prominent event in meiosis of almost all sexually reproductive organisms. It contributes to genome stability by ensuring the balanced segregation of paired homologs in meiosis, and it is also the major driving factor in generating genetic variation for natural and artificial selection. Meiotic recombination is subjected to the control of a highly stringent and complex regulating process and meiotic recombination frequency (MRF) may be affected by biological and abiotic factors such as sex, gene density, nucleotide content and chemical/temperature treatments, having motivated tremendous researches for artificially manipulating MRF. Whether genome polyploidization would lead to a significant change in MRF has attracted both historical and recent research interests, however tackling this fundamental question is methodologically challenging due to the lack of appropriate methods for tetrasomic genetic analysis, thus has led to controversial conclusions in the literature. This paper presents a comprehensive and rigorous survey of genome duplication mediated change in MRF using *S. cerevisiae* as a eukaryotic model. It demonstrates that genome duplication can lead to consistently significant increase in MRF and rate of crossovers across all sixteen chromosomes of *S. cerevisiae,* including both cold and hot spots of MRF. This ploidy driven change in MRF is associated with weakened recombination interference, enhanced double-strand break density and loosened chromatin histone occupation. The study illuminates a significant evolutionary feature of genome duplication and opens an opportunity to accelerate response to artificial and natural selection through polyploidization.

INTRODUCTION

In meiosis of all sexually reproductive eukaryotic species, recombination between homologous chromosomes generates reciprocal exchanges of genes on the chromosomes through chromosomal crossing over. Crossovers are essential for ensuring physical connections and balanced segregation between the homologous chromosomes (Jones and Franklin 2006). Reciprocal exchange of parental genetic material allows reshuffling of genes on these chromosomes and thus creates new allelic combinations of them in offspring individuals for natural and/or artificial selection. It has been well documented that meiotic recombination is a strictly programmed and controlled process, characterized by DNA double-strand breaks catalyzed by the topo-isomerase-related enzyme Spo11 and repairing of the breaks involving with a group of evolutionarily highly conserved proteins (Osman *et al* 2011). It is well established that initiation

and distribution of recombination along chromosomes is not random, and meiotic recombination frequency (MRF) may be affected by several biological and genomic factors such as sex (Lenormand and Dutheil 2005), density of genes or transposable elements (Shifman et al. 2006), GC nucleotide content (DeRose-Wilson and Gaut 2007) and genome sequence heterozygosity (Ziolkowski *et al* 2015).

Whether polyploidization of a eukaryotic genome alters the meiotic recombination frequency (MRF) has attracted both historical and recent research interests. Oram (1959) was probably the pioneer in investigating this fundamental question by comparing MRF of two chromosomal regions in diploid and tetraploid maize segregation populations and concluded that MRF in one of the two marker intervals was significantly higher in tetraploids than that in diploids, but the difference is not significant in the other chromosomal interval. Using data collected on three linked dominant/recessive markers from diploid and autotetraploid maize backcross populations, Welch (1962) observed an increased rate of crossovers in diploids in one of the two linked chromosomal intervals but no difference in the crossover rate in the other, and emphasized the weakness of his data analysis for unavailability of an appropriate method for analyzing the autotetraploid data. Sved (1964) presented a theoretical prediction that polyploidization through genome duplication may increase MRF. More recently, Pecinka et al (2011) focused on recombination frequency between a single pair of seed expressing fluorescent markers segregating in large diploid and tetraploid populations of *Arabidopsis* and concluded that recombination frequency between the fluorescent marker loci was markedly higher in the tetraploids than in the diploids. Recognized inappropriate method of statistical analysis with the tetraploid marker data, we re-analyzed their dataset and inferred a significant increase in the recombination frequency in the tetraploids than in the diploids (Wang and Luo 2012). This paper reports a comprehensive study on ploidy driven change in MRF by using *Saccharomyces cerevisiae* as an experimental model and by developing appropriate statistical methods for modeling and analyzing the autotetraploid experimental datasets.

RESULTS

Creation and transformation of fluorescent marker cassettes

Firstly, we modified the lithium acetate method (Lorenz et al. 1995) to transform the two fluorescent cassettes, *GFP* and *RFP* (**Supplementary Figure 1A**), which were driven by the promoter *TEF2* and *TDH1*, onto the same chromosome at pre-designed locations in the haploid

yeast strain S288c. We then developed a feasible but reliable experimental approach to convert two haploid yeast strains, S288c and SK1 respectively, into diploid and then duplicated the diploids into autotetraploid strains as described in **Supplementary Figure 1B** and **Supplementary Method**. Ploidy levels of the created diploid and autotetraploid strains were confirmed in terms of measure of the genomic DNA content by use of the fluorescence-activated cell sorting (**Supplementary Figure 1C**). We repeated transformation of the cassette bearing the two fluorescent markers into each of the sixteen chromosomes at a pre-designed location as illustrated in **Figure 1A**. Precise locations of these fluorescent marker genes are detailed in **Supplementary Table 1** for each of the sixteen yeast chromosomes. Design of the marker locations were mainly arbitrary for a fairly even representation of the marker locations in the sixteen yeast chromosomes except that some of these marker genes have been deliberately designed to locate within previously identified recombination hot (red stars) or cold (blue stars) spot regions (Gerton et al. 2000) as highlighted in **Figure 1A** and **Supplementary Table 1**.

For each of the fluorescent marker cassettes, we created two F_2 segregating populations from crossing the two parental strains, s288cand SK1, in diploid and autotetraploid respectively, and scored a varying number (603 to 2,129) of tetrads from these segregating populations for their phenotype of the two fluorescent markers (**Supplementary Table 2**). Each spore in the 26,281 tetrads scored was phenotyped as either black (**B**), green (**G**), red (**R**) or yellow (**Y**), corresponding to the spores that carry none, only green allele, only red allele or both of the fluorescent marker alleles (**Figure 1B**).

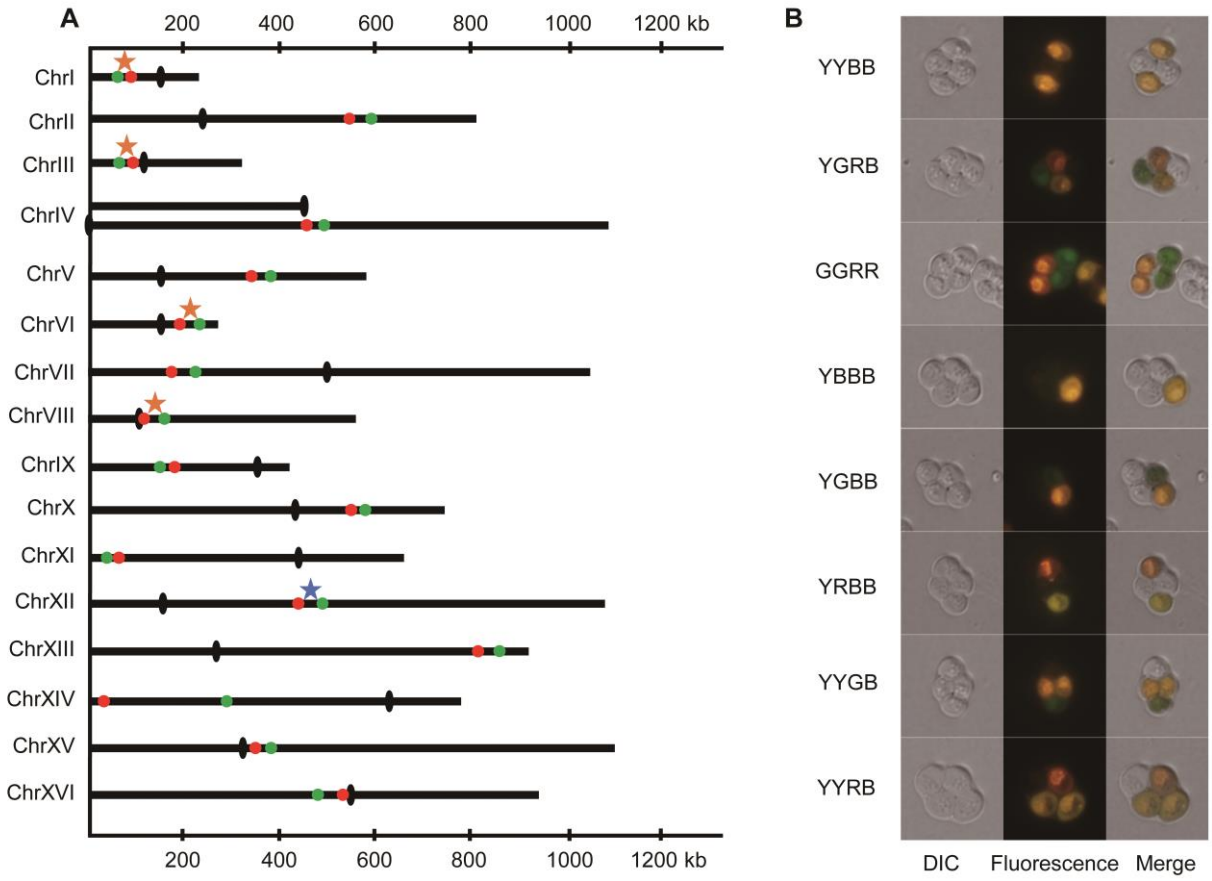

Figure 1 (**A**) Distribution of fluorescent red and green markers on each of sixteen yeast chromosomes with black dots indicating the centromeres, the red and blue stars being the previously defined hot and cold recombination spots. (**B**) Phenotype of spores carrying different alleles at the fluorescent markers. The black olives indicate the centromeres on the yeast chromosomes.

Meiotic recombination frequencies in the diploid and autotetraploid segregating populations

The tetrad data is fully informative in regard to the underlying genotype and recombination events during meiosis of the diploid parents, and thus calculation of recombination frequency between the fluorescent markers is straightforward in the diploid segregating populations. However, the same analysis in the autotetraploid segregating populations is far more complicated, primarily due to the complexities in gene segregation and recombination under tetrasomic inheritance, reflecting in several main aspects. Firstly, in autopolyploids, homologous chromosomes in meiosis may undergo quadrivalent pairing, resulting in the well-known phenomenon of double reduction, i.e. sister chromatids enter into the same gamete (Mather 1936) after the meiosis, leading to systematic allelic segregation distortion in comparison to disomic gene segregation and recombination. Secondly, multiple alleles at individual loci of polyploids cause a substantially wider spectrum of genotypic segregation at the loci. These make tetrasomic linkage analysis a historically

challenging and unsolved question (Bailey 1961) since the pioneer geneticists like RA Fisher (1947), JBS Haldane (Haldane 1930) and K Mather (1936). We have developed a general statistical framework for tetrasomic linkage analysis, which has taken a full account of the key features of tetrasomic inheritance (Luo et al 2004). We modified that general framework to specifically model and analyze the tetrad data of the autotetraploid segregating populations in the present study as detailed in **Statistical Method 1.**

Table 1 lists the maximum likelihood estimates of recombination frequencies ($\hat{r}$) and their sampling standard deviations (s.d.) between the fluorescent markers across the 16 yeast chromosomes in autotetraploid and diploid yeast genomes of *S. serevisae*. It shows that MRF in the autotetraploid genome is consistently highly significantly increased across all the 16 yeast chromosomal regions under investigation when compared to that in the diploid genome. We have converted these estimates of MRF into mapping distances in cM by multiplying the MRF estimates by 100 and illustrated the estimated mapping distances of the marker regions normalized by the corresponding physical distance in (Mbp) in **Figure 2A**. Means of cM/Mbp values are calculated to be 383.5 and 506.7 for the diploid and tetraploid yeasts respectively. It is noted that the significant increase in MRF in the autotetraploids was also observed in both previously characterized hot or cold recombination spots (**Supplementary Table 1**). Listed also in **Table 1** are the maximum likelihood estimates of the coefficient of double reduction at the linked fluorescent marker loci ($\hat{\alpha}$ and $\hat{\beta}$) and the corresponding standard deviations. Moreover, estimates of the coefficient of double reduction are statistically significant for the markers on chromosomes 4, 7, 13, 14 (**Table 1**), revealing a clear signal of the quadrivalent chromosomal pairing in the marked chromosomal regions during meiosis of the autotetraploid yeast cells, suggesting significant double reduction event and tetrasomic inheritance of maker alleles in the autotetraploid yeast strains. Additionally, we tested for significance in allelic deviation from neutral segregation at the marker loci and none of the tests was statistically significant ($P \geq 0.1750$), excluding the possibility of selection on the fluorescent markers and the influence of selection on assessing MRF.

Table 1 Maximum likelihood estimates of recombination frequencies ($\hat{r}$) and the sampling standard deviations (*s.d.*) between the fluorescent markers across 16 yeast chromosomes in autotetraploid and diploid yeast genomes as well as the maximum likelihood estimates of the coefficient of double reduction at the linked fluorescent loci ($\hat{\alpha}$ and $\hat{\beta}$) and the corresponding standard deviation.

Chromosomes	Autotetraploids			Diploids
	$\hat{\alpha} \pm s.d.$	$\hat{\beta} \pm s.d.$	$\hat{r} \pm s.d.$	$\hat{r} \pm s.d.$
1	0.0305 ± 0.0336	0.0361 ± 0.0335	0.2245 ± 0.0086	0.2126 ± 0.0081
2	0.0466 ± 0.0394	0.0435 ± 0.0393	0.1640 ± 0.0088	0.1165 ± 0.0065
3	0.0060 ± 0.0347	0.0229 ± 0.0347	0.1617 ± 0.0076	0.1035 ± 0.0056
4*	0.0659 ± 0.0382	0.0629 ± 0.0382	0.1071 ± 0.0069	0.0797 ± 0.0054
5	0.0431 ± 0.0392	0.0446 ± 0.0392	0.1404 ± 0.0081	0.1004 ± 0.0061
6	0.0147 ± 0.0383	0.0368 ± 0.0383	0.4392 ± 0.0139	0.2776 ± 0.0088
7-1**	0.0618 ± 0.0233	0.0678 ± 0.0234	0.2322 ± 0.0062	0.1599 ± 0.0062
7-2**	0.0629 ± 0.0216	0.0672 ± 0.0217	0.2445 ± 0.0059	0.1755 ± 0.0063
8	0.0000 ± 0.0383	0.0063 ± 0.0398	0.3460 ± 0.0128	0.2571 ± 0.0087
9	0.0252 ± 0.0385	0.0312 ± 0.0385	0.2389 ± 0.0103	0.1918 ± 0.0080
10	0.0338 ± 0.0392	0.0510 ± 0.0392	0.2295 ± 0.0103	0.1765 ± 0.0076
11	0.0404 ± 0.0387	0.0510 ± 0.0387	0.1682 ± 0.0087	0.1645 ± 0.0074
12	0.0495 ± 0.0361	0.0326 ± 0.0361	0.1085 ± 0.0065	0.0642 ± 0.0046
13*	0.0786 ± 0.0377	0.0814 ± 0.0378	0.2110 ± 0.0095	0.1884 ± 0.0078
14**	0.0637 ± 0.0308	0.0751 ± 0.0308	0.6143 ± 0.0133	0.4681 ± 0.0083
15	0.0000 ± 0.0403	0.0098 ± 0.0403	0.0650 ± 0.0056	0.0332 ± 0.0036
16	0.0000 ± 0.0384	0.0217 ± 0.0394	0.2644 ± 0.0110	0.1999 ± 0.0080

* ($P < 0.05$) and ** ($P < 0.01$) indicate the creditability levels of significance of estimates of the coefficients of double reduction at the fluorescent markers.

Rate of Crossovers in Diploid and Autotetraploid Yeast Genomes

We further characterized the genome-wide distribution of crossovers generated in heterozygous diploid and autotetraploid yeast strains. We first created a heterozygous diploid strain s288c/SK1 (or hs in abbreviation) from the haploid strains, s288c and SK1, which differ at least at 63,000 SNP sites (1 SNP per 190 bp or per 0.06 cM) (Gerton et al. 2000). We designed autotetraploid strains with genome constructs of s288c/SK1/SK1/SK1 (or hsss) and s288c/s288c/s288c/SK1 (or hhhs) (**Material and Methods**). These designed constructs allow crossover detection to be focused on the same single chromosome in both diploid and autotetraploid genomes, and thus enable a direct

comparison of rate of crossovers (CO) involved with the specific single chromosome between diploid and tetraploid genomes. We randomly collected and micro-dissected 5, 3 and 3 tetrads generated from the diploid (hs), autotetraploid (hsss) and autotetraploid (hhhs) strains respectively using a dissection microscope (SINGER MSM400, UK). These tetrads from heterozygous diploid and autotetraploid strains were sequenced using Illumina's Hiseq 2000 sequencer with a design of 2 x 100 bp paired end reads.

From the tetrad sequence datasets, we firstly identified the sequence variant marker sites for the CO analysis, on which the marker alleles show a 2:2 allele configuration in diploid tetrads (hs), a 6:2 or 2:6 configuration in tetraploid tetrads (hhhs) or (hsss) respectively, and thus selected an average of 47,700 markers from the diploid tetrads and an average of 51,969 markers from the tetraploid tetrads for assaying rate of crossovers in these tetrad spores. Use of these selected sequence based markers may effectively avoid influence of sequencing errors, errors from data processing such as nucleotide calling, sequence reads mapping and the compounding meiotic events involving with gene conversion or structural events due to genome instability. A crossover was identified as a reciprocal exchange occurring between chromatids marked by the selected marker sites (**Supplementary Figure 2**). We were able to observe crossovers directly from the diploid tetrad sequence data and from part of tetraploid tetrad sequence data when linkage phase at the linked marker loci can be directly inferred. The COs so derived is referred to as observed COs.

However, it is not feasible to call crossovers (COs) directly from the autotetraploid tetrad sequence data because any tetrad spore is a diploid and linkage phase may be unknown for any spore with a double heterozygote genotype at the flanking sequence markers. We proposed here a statistical method for predicting the number of COs per chromatid of an autotetraploid from the tetrad sequence data as detailed in **Statistical Method 2**. Using the method, we calculated the expected number of COs for each chromosome in the autotetraploid genomes. The COs so derived are referred to as estimated COs. To further remove those observed COs which may be vulnerable to the sequence errors and sequence variants aforementioned, we removed those adjacent observed COs if they were separated by < 10kb (Malkova et al 2004).

Figure 2B shows the mean number of COs observed from diploid yeast tetrads and observed plus estimated COs from tetraploid tetrads. A *t* test shows that the mean number of observed COs from the tetraploid tetrads was significantly higher than that from the diploid tetrads ($P < 0.05$).

The difference would be significant at a much higher statistical confidence if the comparison is made after including the estimated COs of the tetraploid tetrads. It is noted that difference in mean observed number of the observed COs (i.e. hhhs and hsss) was not significant between the two types of tetraploids (*t* test with P = 0.8275). Figure 2C (and also Supplementary Table 6) shows means of the observed COs for each of sixteen chromosomes of the diploid and tetraploid yeast tetrads. A pairwise *t* test shows significant difference in the mean number of observed COs across the sixteen chromosomes between diploid and tetraploid yeasts (P < 0.05), and the corresponding test was not significant when comparison was between the two tetraploid types (P = 0.85).

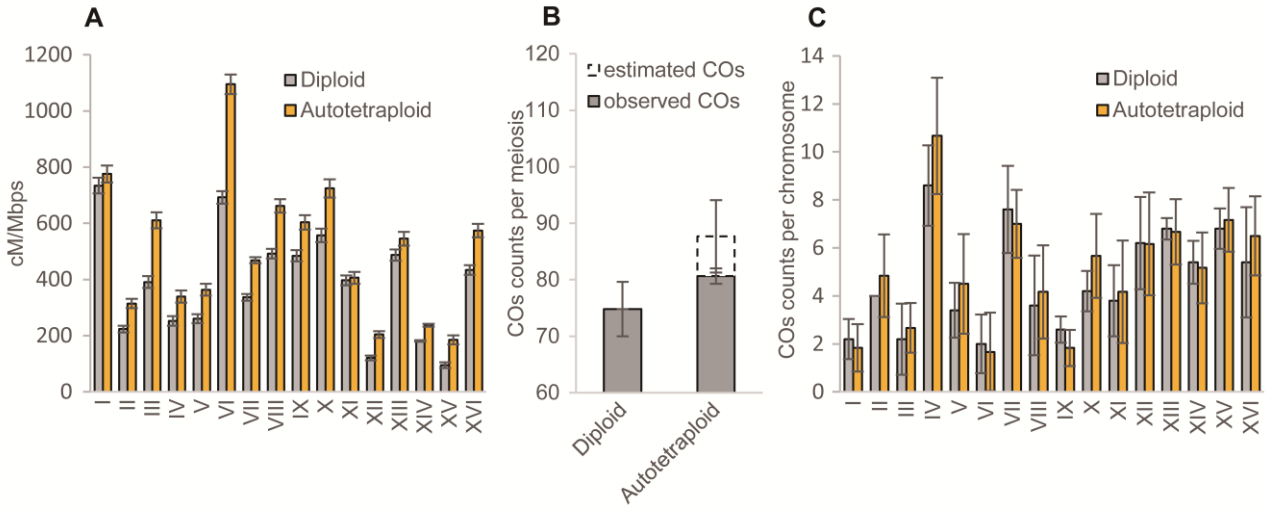

Figure 2. (**A**) The estimated linkage map distance (cM) normalized by the corresponding physical map distance (Mbp) and (**B**) Means of the observed or/and predicted number of crossovers (COs) per meiosis of diploids and autotetraploids. **(C)** The observed number of crossovers (COs) per meiosis across 16 chromosomes (I - XVI) from *n* = 5 diploid tetrads (HS) and *n* = 6 autotetraploid tetrads (hsss or hhhs) yeast *S. cerevisae.*

Comparison in Recombination Interference between Diploid and Autotetraploid Genomes

It has been well established that recombination does not independently occur along chromosome arms, and any recombination at a site may usually prevent others at any nearby sites, the phenomenon is so-called Recombination Interference (RI) (Keeney et al. 1997). RI may be attributed to two types of interference, the chromatid interference, where different pairs between non-sister chromatids are not equally likely to be involved in formation of crossovers, and the position or chiasmata interference, where occurrence of one crossover event at a position along the chromosomal bundle affects chance of an additional crossover to occur in a nearby region (Mieczkowski et al. 2007). We focus here on the latter, and test for a hypothesis that increased

meiotic recombination frequency in the autotetraploid yeast when compared to that in the diploid genome is in association with weakened RI.

To test the hypothesis, we created diploid and autotetraploid yeast strains which carried three antibiotic markers (*NAT*, *HGY* and *G418*) in simplex dose on the same arm of the yeast chromosomes III, VI and VIII, and details of these antibiotic markers are listed together with their chromosome locations in **Supplementary Table 2**. These diploid and autotetraploid strains respectively generated the corresponding segregant (gamete) populations. Segregants of the diploid strain were haploids but segregants of the autotetraploid are diploids. The antibiotic markers are indeed dominant 'present' and 'absent' markers, and the number of diploid and autotetraploid segregants is listed in **Supplementary Table 3** together with the phenotype of the markers on each of the three yeast chromosomes.

Statistically, recombination interference (RI) is defined as the coincidence coefficient (CC) in several different forms, which measures degree of independence of recombination events along two adjacent marker intervals flanked by marker loci A, B and C. Let $r_{AB\cap BC}$ be frequency of recombination simultaneously occurring in the two adjacent marker intervals AB and BC. Let $r_{AB|BC}$ (or $r_{AB|\overline{BC}}$) be frequency of recombination between marker loci A and B with (or without) recombination in the interval flanked by markers B and C. The coincidence coefficient can then be defined as $c_{AB-BC} = r_{AB\cap BC} / \left(r_{AB} \times r_{BC}\right)$, $c_{AB/BC} = r_{AB|BC} / r_{AB|\overline{BC}}$ and $c_{BC/AB} = r_{BC|AB} / r_{BC|\overline{AB}}$. In absence of RI, $c_{AB-BC} = c_{AB|BC} = c_{BC|AB} = 1$and significant deviation of the coincidence coefficient from the expected value of 1.0 implies significance of RI. In populations of the diploid segregants that are haploids, recombination events across the antibiotic marker intervals are directly countable, and the coincidence coefficients can thus be calculated directly from the marker data and the estimates are listed in **Table 2**. To test for significance of the CC estimates, we proposed here a simulation based method to generate an empirical significant thresholds for the CC estimates' deviation from the null hypothetical value, and assessed significance of deviation of the CC estimates from their expected value of 1.0 under the null hypothesis. The simulation mimics gametogenesis of any given diploid genotypes at any number of loci assuming independent recombination among different chromosomal regions as detailed elsewhere (Luo et al. 2004), and the simulation parameters were extracted from the real datasets. It can be seen from **Table 2** that recombination interference in the diploid data was detected significant in the marked region on chromosome III, highly significant on chromosome VI but not significant on chromosome VIII. It is

further shown that the CC estimates and their significance were consistent each other, suggesting the equivalence and reliability of the CC estimates from the antibiotic marker data.

Table 2 The maximum likelihood estimates of the coincident coefficients ($\hat{c}_{AB-BC}$, $\hat{c}_{AB/BC}$ and $\hat{c}_{BC/AB}$) from diploid and autotetraploid yeast chromosomes III, VI and VIII and the maximum likelihood estimates of the coefficient of double reduction ($\hat{\alpha}$) and its sampling standard deviation ($s_{\hat{\alpha}}$) at each of antibiotic markers (*NAT*, *HGY* and *G418*) on each of the autotetraploid yeast chromosomes (III, VI and VIII).

Diploids	Chromosome III	Chromosome VI	Chromosome VIII
$\hat{c}_{AB-BC}$	0.8214*	0.5990***	0.9436
$\hat{c}_{AB/BC}$	0.7695*	0.5727***	0.9319
$\hat{c}_{BC/AB}$	0.7978*	0.5322***	0.9241

Autotetraploids	*NAT*	*HGY*	*G418*	*NAT*	*HGY*	*G418*	*NAT*	*HGY*	*G418*
$\hat{\alpha}$	0.0281	0.0506	0.0787	0.0201	0.0718	0.0316	0.0000	0.0315	0.0287
$s_{\hat{\alpha}}$	0.0530	0.0531	0.0536	0.0536	0.0536	0.0539	0.0535	0.0535	0.0535
$\hat{c}_{AB-BC}$	1.0426			1.0376			0.8778		
$\hat{c}_{AB/BC}$	1.0593			1.0442			0.8488		
$\hat{c}_{BC/AB}$	1.0492			1.0557			0.8414		

* $P < 0.05$, *** $P < 0.001$. The P values were calculated from 1,000 permutation simulations.

It is much more sophisticated to model and analyze the marker data for predicting RI from the autotetraploid segregants which are diploi ds because of the possible tetrasomic inheritance at the antibiotic markers. We first tested for significance of double reduction at each of the marker loci on each of the chromosomes under question. For instance, let α be the coefficient of double reduction at a marker locus A. Given an autotetraploid genotype ABBB, the autotetraploid generates two types of gametes AB and BB with probabilities $(2-\alpha)/4$ and $(2+\alpha)/4$ respectively following the principle we previously developed (Luo et al. 2004). In an sample with n_1 of AB and n_2 of BB gametes, the maximum likelihood estimate of α is given by $\hat{\alpha} = 2(n_2 - n_1)/(n_1 + n_2)$, and sampling variance of the estimate can be calculated as $s^2_{\hat{\alpha}} = (4-\alpha^2)^2/[(2+\alpha)^2 n_1 + (2-\alpha)^2 n_2]$. The maximum likelihood estimates of the coefficient of double reduction at the marker loci are listed in lower panel of **Table 2** together with the corresponding standard deviations. It shows that double reduction is not significant on every marker loci, suggesting bivalent pairing of homologous chromosomes at the loci under question during meiosis.

Based on the quadrivalent chromosomal pairing and simplex marker alleles linked on the same chromosome, we calculated the coincidence coefficient (CC) from the tetraploid segregant marker data. Again, we implemented a simulation based approach to evaluate significance of the CC estimates' deviation from their expected value of 1.0 under the null hypothesis. The simulation mimics gametogenesis of any autotetraploid genotype at any number of loci under either bivalent pairing or quadrivalent pairing of homologous chromosomes as previously described (Luo et al. 2004). Specifically, the simulated parental strain had a genotype ABC/abc/abc/abc where the capital letters correspond to the three antibiotic marker alleles, and other simulation parameters such as the number of segregants and recombination frequencies between the marker loci were directly extracted from the real datasets. The simulation was repeated 1,000 times and recombination was simulated independently among the two marker intervals. **Table 2** shows that none of the marked chromosomal regions in the tetraploid yeast was detected to show significant recombination interference. Comparison of the RI estimates for the same marked chromosomal regions in the same genome but at different ploidy levels indicates that recombination interference in the tetraploid yeast genome was significantly weakened when compared to that in the corresponding diploid genome.

Double Strand Breaks and Histone Occupation in Diploid and Autotetraploid Genomes

Crossing over between paired homologous chromosomes in meiosis is initiated by developmentally programmed DNA double-strand breaks (DSB), which are catalysed by the topoisomerase-like protein Spo11 (Keeney et al. 1997). Change in DSB frequency will thereby affect frequency of crossovers and, in turn, recombination frequency (Mieczkowski et al. 2007). In the yeast strains with RAD50S mutated, the epitope-tagged Spo11 remains bound to the sheared DNA even after completion of DSB. This allows the DNA fragments surrounding the DSB sites to be enriched through immuno-precipitation of the Spo11-DNA complex (Gerton et al. 2000; Alani et al. 1990; Prieler et al. 2005). On basis of the principle, we compared the density and frequency of DSB in the diploid and autotetraploid yeast genomes. Immuno-precipitated Spo11-oligos from diploids and autotetraploids were deeply sequenced with two biological replicates for each of the two yeast strains. We obtained more than 1.5 million sequence reads of 50bp length for per sample. More than 80% of the sequence reads of the sequenced samples (2 biological replicates for diploid and tetraploid cells) were uniquely mapped to the genome of the yeast strain SK1, suggesting a good quality of the sequencing data. In parallel, genomic DNA from the strain SK1 was also sequenced as input control with the same sequence depth in order to achieve a

comparable number of uniquely mapped sequence reads to that of the corresponding ChIP-seq experiment. We split the yeast genome into bins of 1kb and calculated RPKM (reads per kilobases per million mapped reads) from the ChIP-seq and input control sequence data of the diploid and autotetraploid samples.

The calculated RPKM per bin from the diploid ChIP-seq data was highly significantly positively correlated with that from the autotetraploid ChIP-seq data (r = 0.92, P=0), suggesting a high degree of consistency in distribution of DSB across the genome between the yeast strains at two different ploidy levels (**Figure 3A**). Among the enriched regions with an enrichment fold change of ≥ 3 observed in the diploid and autotetraploid samples, there were 443 shared by the two samples (**Figure 3B**), which include previously identified DSB hotspots such as HIS4, HIS2, ARG4, CYS3 etc (Gerton et al. 2000).

By sequencing Spo11-bound oligos, Pan *et al* identified 3,600 DSB hotspots in diploid budding yeast (Pan et al. 2011), which included all the 443 enriched bins we detected in the present study. We compared the Spo11-oligos between the diploid and tetraploid at the 443 bins and all the 3,600 hotspots through a paired *t* test. The test showed the autotetraploid fold enrichment was consistently significantly higher than that of the diploid at the 3600 DSB sites (P=3e-22, **Figure 3C**). In contrast, no difference was observed in the fold enrichment at the control sites (1kb downstream of the DSB sites, **Figure 3D**) between the diploid and autotetraploid samples (P = 1.0, **Figure 3C**).

To further validate the ChIP-seq analysis, we compared the DSB density of diploid and autotetraploid yeasts at several previously identified sites, one at YCR047C and two within 3' region of YFR25C (Petes 2001), through the standard Southern blotting assay in which the target DNA was detected by the probes labelled with digoxin-11-dUTP. The Southern blotting assay shows that DSB density at these sites was markedly higher in the tetraploid than in the diploid (**Figure 3E**), agreeing with that revealed by the ChIP-seq data. Moreover, the higher DSB frequency at 3' region of YFR025C was also confirmed by an additional return-to-growth assay **(Figure 3F**). Additionally, we conducted the return-to-growth assay using genetically modified diploid and autotetraploid strains. The diploid strain carried two nonsense mutants, his2-A (blue box) and his2-X (purple box) of HIS2, whilst the corresponding autotetraploid carried two combinations of the mutant alleles on a chromosome (**Figure 3F**). The mutant carriers were auxotrophy but would recover back to be normal when a functionally normal HIS2+ was created from DSB and the

following gene conversion between the two mutant genes on different chromosomes. Thus, proportion of the yeast cells carrying the mutant alleles, which could grow on the medium lacking of histidine (i.e. HIS2+ carrying cells), reflects the density of DSB surrounding the gene conversion. The right panel of **Figure 3F** shows the percentage of the HIS2+ carrying cells with either a diploid or autotetraploid genome and reveals that the autotetraploid yeast had significantly higher density of DSB than the diploid cells. This provides further evidence supporting the autotetraploid genome has a denser DSB than the diploid and agrees well with the above ChIP-seq and Southern blotting assays.

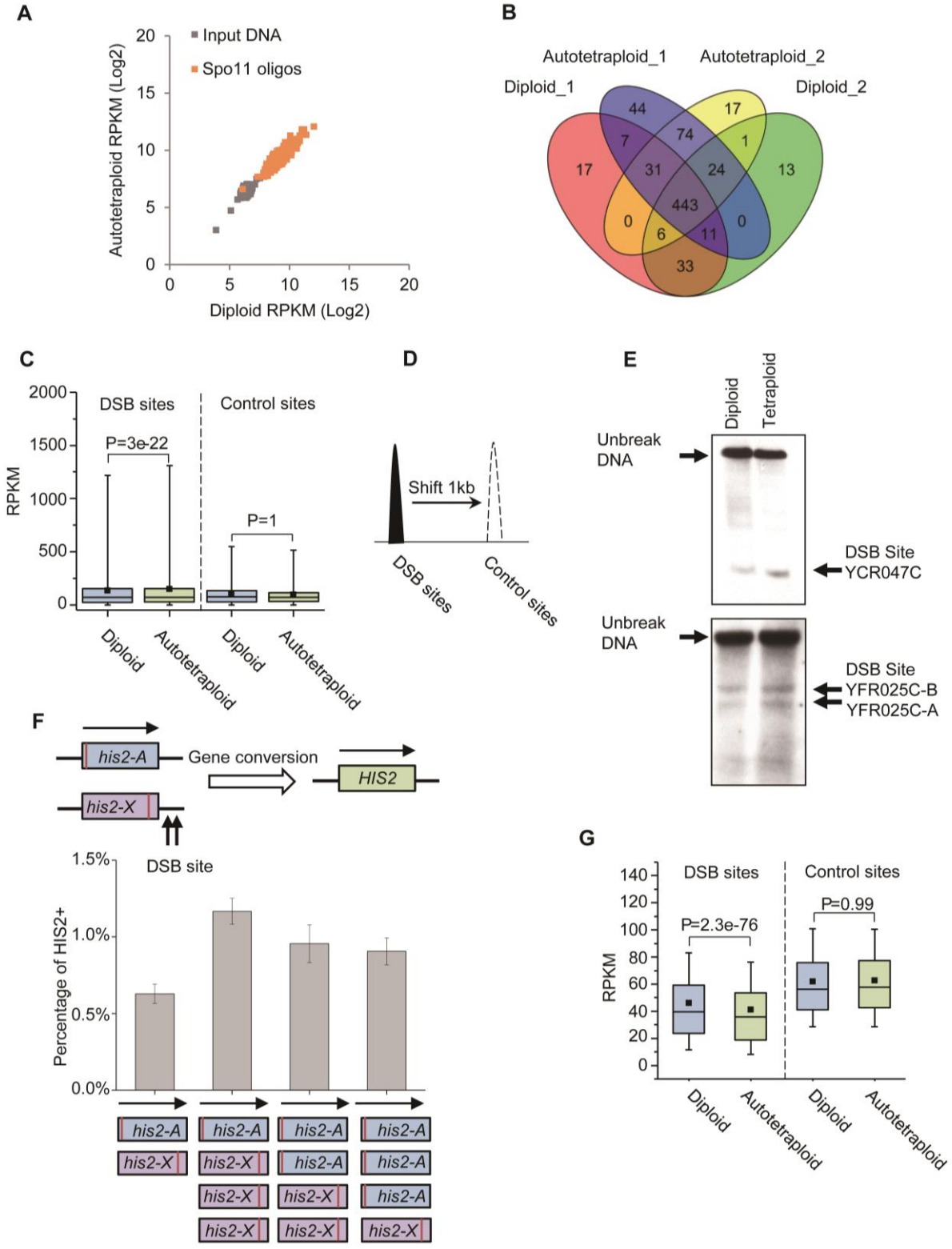

Figure 3 ChIP-seq assays of double strand breaks (DSB) and histone occupation in diploid and autotetraploid yeast strains. **(A)** Scatterplot of Log_2 transformed RPKM values of input DNA (grey) and spo11-oligos at the 443 DSB sites common to both diploid and autotetraploid cells. **(B)** The number of DSB sites identified from 2 biological replicates of independently cultured diploid and autotetraploid cells. **(C)** Illustration of ChIP-seq assay at the 3600 DSB sites and control sites of 1kb downstream of the corresponding DSB sites. **(D)** Boxplots of RPKM from ChiP-seq assay at the 3600 DSB sites and the control sites in diploid and tetraploid cells. **(E)** Southern blotting assay of DBS density at three identified sites in genomes of diploid and tetraploid yeasts. **(F)** Rerun to

growth assay of DSB at site *HIS2* (or YFR025C). **(G)** Boxplots of RPKM of micrococcal nuclease (MNase)-resistant mononucleosome DNA at the 3600 DSB sites and the control sites in diploid and tetraploid cells.

Majority of DSB sites or crossover hotspots are nuclease-hypersensitive and share a common open chromatin structure without nucleosome occupation, which are recognized to be necessary for Spo11 to access the DNA substrate and break DNA double strands (Petes 2001). The 3,600 DSB hotspots aforementioned were also observed to be markedly less occupied by nucleosomes (Prieler et al. 2005). We examined the histone occupation landscape in both diploid and autotetraploid genomes by sequencing DNA at micrococcal nuclease (MNase)-resistant mononucleosomes. The RPKM was calculated for per chromosome from the DNA sequence data and compared between the diploids and autotetraploids. The analysis shows that the RPKM of the autotetraploids is significantly decreased when compared to that of the diploids at all the 3,600 DSB sites (P=2.3e-76, **Figure 3G**), indicating that the autotetraploid chromosomes were less wrapped by nucleosomes than the diploid chromosomes. This agrees well with the above ChIP based DSB assay and further supports the ploidy driven increase in meiotic recombination frequency in budding yeast.

RPKM was proposed to detect differential gene expression between two samples from RNA sequence data when total amount of expression is comparable between samples under comparison (Evans et al 2018). We implemented the normalization method to compare level of Spo11 bound and MNase-resistant oligos at the DSB sites between diploid and tetraploid cells. In the Spo11 pulling down ChIP-seq experiment, sequence reads from Spo11 bound or MNase resistant oligos account for only 19.13 and 21.96% at the DSB regions or 7.1 and 6.7% of total sequence reads at the control regions of the diploid and tetraploid yeast respectively. In the MNase-seq experiments, these figures are 3.9 and 3.4% at the DSB regions or 5.0 and 5.1% at the control regions (Supplementary Table 7). Additionally, a markedly high level of positive correlation in ChIP-seq or MNase-seq data was observed between the diploid and tetraploid yeasts. These suggest that the RPKM normalization would be recognized appropriate for comparing the sequence data between diploid and tetraploid yeasts in the present study

DISCUSSION

This study reports the first comprehensive survey of the ploidy driven change in meiotic recombination frequency (MRF) in *Saccharomyces cerevisiae*. It demonstrates that a significant

increase in MRF occurred after genome duplication in all examined regions covering both recombination hot and cold spots on all 16 yeast chromosomes. Given that formation of crossovers is a prerequisite for recombination, we assayed profile of crossovers on the all 16 chromosomes of the diploid and tetraploid yeast genomes using genome sequencing and found that rate of crossovers is, on single chromosome basis, highly significantly higher in the autotetraploid genome than in the corresponding diploid genome. It should be acknowledged that crossing over and gene conversion are associated in meiosis. Gene conversion and crossing over may be regulated independently, and crossing over may involve interference but gene conversion may not (Engebrecht et al 1990). The present tetrad sequence data should allow to work out the rate of gene conversion in addition to the rate of crossovers. However, tetraploid tetrads are diploids and how to estimate gene conversion from the tetraploid tetrad sequence data remains an open question for further exploitation.

To explain the ploidy driven increase in MRF, we compared level of recombination interference (RI) between diploid and tetraploid genomes and observed that RI is a clearly weakened in the genomic regions investigated. Additionally, a Spo11 pulled down DNA sequencing experiment indicates a significantly higher frequency of single chromosome based double-strand breaks (DSB) in the tetraploid genome than in the corresponding diploid counterpart at the 443 commonly shared and previously identified sites in meiosis of both the diploid and tetraploid yeasts. The observation from the ChIP-seq data is further verified independently by the Southern blotting assay and Return to growth experiment. DNA sequence data collected at micrococcal nuclease (MNase)-resistant mononucleosomes shows that the tetraploid chromosomes are comparatively less wrapped by nucleosomes than the diploid chromosomes, having provided the structural basis for the difference in DSB frequency between the diploid and tetraploid genomes. Although there has yet been mechanistic explanation for association of genome duplication with nucleosome occupancy to our best knowledge, it has been well established that genome duplication is in parallel with the dynamic alteration in genome structure and functionality such as genome wide gene expression and phenotypic change (Song et al 1995, Comai et al 2005). On the other hand, nucleosome occupancy is partially encoded by intrinsic antinucleosomal DNA sequences as well as by binding sites for trans-acting factors that can evict nucleosomes in yeast (Jiang and Pugh 2009). Moreover, binding site gain or loss events at nucleosome depleted regions in yeast genome may cause more expression differences than those in nucleosome occupied regions (Swamy et al 2011).

Thus, the present study could stimulate further study to elucidate the molecular mechanism underpinning genome duplication associated change in nucleosome occupancy.

Although it has been theoretically predicted that meiotic recombination would be more frequent in tetraploids than in diploids (Sved 1964), there are only a few experimental studies of limited scale by far performed to test this theoretical prediction (Oram 1959; Welch 1962). More recently, Pecinka *et al* compared frequency of meiotic recombination of two fluorescent makers between diploid and artificially synthesized tetraploid in *Arabidopsis* (Pecinka et al. 2011), and concluded a higher recombination frequency at the marked chromosomal interval in tetraploid than in diploid. In fact, appropriate statistical methods for linkage analysis under a tetrasomic inheritance model are essential for statistically appropriate evaluation of genetic recombination frequency ((Wang and Luo 2012). One of other important distinctions of the present study from its rivals in the literature is the development of statistically appropriate methods for modeling and analyzing the experimental data under a tetrasomic inheritance basis. This not only enables the data specific analysis in the present study but also provides useful analytical tools for the tetrasomic linkage and genetic analyses with other tetraploid species.

Meiotic recombination is the major mechanism for genetic variation blocked within individual genomes to be released for natural and artificial selection, and thus is recognized as one of major driving factors for the evolution and/or speciation (Otto and Barton 1997) as well as for breaking limits of artificial selection (Hill and Robertson 2007). On the other hand, polyploidy has played an important role in the evolution of eukaryotes, particularly flowering plants, with 30-80% of angiosperms being currently polyploid, and the rest existing as paleopolyploids, having evolved from and/or reverted to a diploid state over evolutionary time (Otto and Whitton 2000). It has been well established that a genome in polyploidy has three distinct advantages. Firstly, a high level of heterozygosity maintained enables polyploids to be more vigorous than their diploid progenitors. Secondly, a larger number of segregating alleles shield the polyploid genome from deleterious mutation. Thirdly, if not finally, most polyploids may reproduce asexually and thus can be propagated much efficiently into large populations (Comai 2005). In addition to these, the present study contributes a significant feature to polyploidization through genome duplication, i.e. the significantly increased meiotic recombination frequency. Thus, the present study, on one hand, fills a gap between plolyploidization and evolution of species via the significant effect of genome duplication on the genome's recombination frequency, and on the other, opens an opportunity for

artificial manipulation of meiotic recombination through polyploidization breeding, as has been seeking for by plant scientists and breeders for many years (Henderson 2012).

MATERIAL AND METHODS

All strains used in this study were isogenic to s288c and SK1. Autotetraploids were constructed through fusion between the mating-type switched a/a and α/α diploids, which were induced and screened from the normal a/α diploids transformed with pTetra plasmid, as described in our previous study (Fang et al. 2018). Fluorescent and antibiotic markers were generated by using of PCR, transformed into strains through the PEG-LiAC method and inserted into predesigned loci of chromosomes. Unless specified, both wild type and genetically modified strains were grown in the standard rich medium YPD (1% yeast extract, 2% polypeptone, 2% glucose, plus 2% agar if necessary). Meiosis was induced in the SPM medium (1% potassium acetate, plus 2% agar if necessary). Mating type switch was stimulated in YPGal (1% yeast extract, 2% polypeptone, 2% galactose, 2% ranffinose). Test strains were synchronized at the same meiosis stage in the YPG plates (3% glycerol, 2% polypeptone, 1% yeast extract, 2% agar) and SPS medium (1% potassium acetate, 1% w/v polypeptone, 0.5% yeast extract, 0.17% yeast nitrogen base with ammonium sulfate and without amino acids, 0.5% ammonium sulfate, 0.05M potassium biphtalate, 2 drops per liter antifoam, pH to 5.5 with 10 N KOH). The ChIP-seq and MNase-seq assay was implemented according to the classic protocol, but with minor modification (Prieler et al. 2005; Kaplan et al. 2009). Detailed description of all experimental materials and methods used can be found in **Supplementary Information**.

Statistical Method 1 *Statistical modelling and analysis of the fluorescent maker data*

To model and analyze the fluorescent marker data collected from the diploid and autotetraploid segregating populations, we proposed here the probabilistic models for modelling the marker data and statistical methods for analyzing the datasets. The model and statistical methods have properly accounted for the key features of disomic and tetrasomic inheritance of gene segregation and recombination and solved properly the challenges in the statistical analysis. Our formulation here focuses on the autotetraploids. We considered a general tetraploid genotype at two loci, $A_1A_2A_3A_4/B_1B_2B_3B_4$, and assumed that the alleles at the two loci are linked with a recombination frequency *r*, and the coefficient of double reduction at locus A is *a*. We have worked out the probability distribution of 136 possible diploid gamete genotypes generated by the autotetraploid individual genotype in term of *r* and *a* (Table 1 in Luo et al 2004). In the present context, we worked out the probability distribution of 10 possible diploid gamete genotypes generated by the

parental genotype, GBBB/RCCC by setting A_1 = G, B_1 = R, A_2 = A_3 = A_4 = B and B_2 = B_3 = B_4 = C and summing up the corresponding genotype frequencies for the same gamete genotype. Table 1 lists the probability distribution for the 10 gamete genotype $g_i(\alpha, r)$ (*i* = 1, 2, …, 10), which describes the probability distribution of the genotypes in the yeast tetrad spore population created in the study. However, genotypes of the diploid spores cannot be directly observable but the spores can be grouped according to four possible fluorescent phenotype classes (yellow, green, red and black). Probabilities of the phenotype groups are given as

$$f_y(\alpha,r) = g_1(\alpha,r)+g_2(\alpha,r)+g_3(\alpha,r)+g_4(\alpha,r)+g_5(\alpha,r) = \frac{1}{12}\left[2(3-3r+r^2)-\alpha(3-6r+5r^2)\right] \quad (1)$$

$$f_g(\alpha,r) = g_6(\alpha,r)+g_7(\alpha,r) = \frac{r}{12}\left[6-2r-\alpha(6-5r)\right] \quad (2)$$

$$f_r(\alpha,r) = g_8(\alpha,r)+g_9(\alpha,r) = \frac{1}{36}(2+\alpha)r(6-r) \quad (3)$$

$$f_b(\alpha,r) = g_{10}(\alpha,r) = \frac{1}{36}(2+\alpha)(3-r)^2 \quad (4)$$

Let n_i denote by the number of diploid spores with the *i*th phenotype (*i* = 1, 2, 3, 4 corresponding to *y*, *g*, *r*, *b* respectively) and $n = n_1+n_2+n_3+n_4$. The log-likelihood of the model parameters, α and *r*, given the observed n_i'*s*, is given by

$$L(a,r|\, n_i) = \sum_{i=1}^{4} n_i \log[\, f_i(a,r)] \quad (5)$$

Because *a* indicates the coefficient of double reduction at the locus nearer to the centromere, information about segregation of alleles at the locus is sufficient to estimate the parameter [21]. To work out the double reduction parameter, we set *r* = 0 in the likelihood function (5), solved the equation

$$(n_1+n_2)\partial\{\log[\, f_1(\alpha,0)+f_2(\alpha,0)]\}/\partial\alpha+(n_3+n_4)\partial\{\log[\, f_3(\alpha,0)+f_4(\alpha,0)]\}/\partial\alpha = 0 \quad (6)$$

for α, and obtained the maximum likelihood (MLE)

$$\hat{\alpha} = 2(n_3+n_4-n_1-n_2)/n \quad (7)$$

The asymptotic sampling variance of the MLE can be calculated according to the Fisher's information metric from

$$-\left[(n_1+n_2)\partial^2\{\log[f_1(\alpha,0)+f_2(\alpha,0)]\}/\partial\alpha^2+(n_3+n_4)\partial^2\{\log[f_3(\alpha,0)+f_4(\alpha,0)]\}/\partial\alpha^2\right]^{-1}_{\alpha=\hat{\alpha}} \quad (8)$$
$$=16\times(n_1+n_2)(n_3+n_4)/n$$

The formulation can be modified by exchanging between $f_1(\alpha,0)$ and $f_2(\alpha,0)$ and also exchanging between n_2 and n_3 to calculate $\hat{\alpha}$, the coefficient of double reduction at the red fluorescent locus, which is distal to the centromere in the model.

We calculated the MLE of recombination frequency *r* directly from solving $\partial L(\hat{a},r|\,n_i)/\partial r=0$, which is equivalent to a polynomial equation of grade 5 and has no a simple and close form for the solution. The equation can be numerically solved and the root within in the range 0.0 and 0.75 was taken as the MLE $\hat{r}$. The asymptotic sampling variance for $\hat{r}$ can be calculated from $-1/\left[\partial^2 L(\hat{a},r|\,n_i)/\partial r^2\right]_{r=\hat{r}}$.

Distribution of phenotype at the single marker locus is given by $f_{R|G}(\alpha)=(4-\alpha)/12$ and $f_B(\alpha)=(8+\alpha)/12$ for the individuals with and without carrying the fluorescent marker respectively. If the number of the two groups of individuals is denoted by n_1 and n_2 respectively,

Statistical Method 2 *Predicting the average number of crossovers from tetraploid gamete data*
Consider a marker interval and let α be the coefficient of double reduction at the flanking marker locus, which is nearer to the centromere, and *p* represents the probability of 1 crossover in the marker interval. We focused here gametogenesis of an autotetraploid individual with the genotype, *AB/ab/ab/ab*, with A and B corresponding to s288c (SK1) alleles, and *a* and *b* to SK1 (or s288c) alleles in the autotetraploid strain s288c/SK1/SK1/SK1 (or SK1/s288c/s288c/s288c). We considered the crossover occurring between all possible non-sister chromatids and all possible configurations of diploid gamete generation under a tetrasomic model, and worked out distribution of phenotype of five possible tetrads at the two marker loci in term of α and *p*, which was listed as **Supplementary Table 4**. In the distribution, a tetrad phenotype was presented as two sequential integers representing two chromosomes. A non-zero integer in the sequence represented the number of A or B alleles and the four integers referred to the four spores.

For a sample of *n* tetrads, let n_i (*i*=1,2,...,5) be the number of tetrads with the i^{th} marker phenotype. The log-likelihood of the model parameters, α and *p*, given the observed n_i is given by

$$L(\alpha, p \mid n_i) \propto \sum_{i=1}^{5} n_i \log(f_i) = n_1 \log\{\alpha(1-p/2)\} + n_2 \log\{(1-\alpha)(1-5p/12)\} + n_3 \log\{\alpha p/2\} + n_4 \log\{(1-\alpha)p/12\} + n_5 \log\{(1-\alpha)p/3\} \quad (9)$$

Differentiating the above and solving the normal equation led to the maximum likelihood estimate of *p* as

$$\hat{p} = \left(11n - 5n_1 - 6n_2 \pm \sqrt{\Delta}\right) / 5n \quad (10)$$

with $\Delta = n^2 + 2n(5n_1 - 6n_2) + (5n_1 + 6n_2)^2$. Choice of the alternatives is feasible because a meaningful estimate of *p* must fall in [0, 1]. Sum of estimates of *p* over all marker intervals on a chromosome gives an estimate of expected number of crossovers of that chromosome.

DATA AVAILABILITY STATEMENT

The Data underlying the article are available in its online supplementary information and all sequence data are available in SRA (Sequence Read Archive) with the access number PRJNA650187 at website https://dataview.ncbi.nlm.nih.gov/object/PRJNA650187?reviewer=45j1d8c1hcm53evd0jkbch35q6).

ACKNOWLEDGMENTS

We gratitude to two anonymous reviewers and the handling editors for their constructively critical comments and suggestions that have greatly helped clarify several ambiguities and errors, and improve presentation of the paper. The study was supported by research grants from the National Nature Science Foundation (grant numbers 31671328 and 31871240) and the State Key project (2016YFC1305501) of China. ZWL is also supported by BBSRC (grant number BB/N008952/1) and the Leverhulme Trust in in the United Kingdom.

AUTHOR CONTRIBUTION

ZL conceived the study, developed statistical methods and conducted the data analysis with input from OF, LW and QT. OF, YZ and FZ conducted the experiment. ZL and OF wrote the paper.

DISCLOSURE DECLARATION

The authors declare no competing financial interests.

Reference cited

Alani E, Padmore R, Kleckner N. 1990. Analysis of wild-type and rad50 mutants of yeast suggests an intimate relationship between meiotic chromosome synapsis and recombination. Cell 61:419–436.

Comai L. 2005. The advantages and disadvantages of being polyploid. Nat. Rev. Genet. 6:836–846.

DeRose-Wilson LJ, Gaut BS. 2007. Transcription-related mutations and GC content drive variation in nucleotide substitution rates across the genomes of Arabidopsis thaliana and Arabidopsis lyrata. BMC Evol. Biol. 7:66.

Engebrecht J, Jeanne Hirsch J, Roeder GS 1990 Meiotic gene conversion and crossing over: Their relationship to each other and to chromosome synapsis and segregation. Cell 6: 927-37.

Evans C, Hardin J, and Stoebel DM (2018) Selecting between-sample RNA-Seq normalization methods from the perspective of their assumptions *Briefings in Bioinformatics* 9(5): 776–792.

Fang O, Hu X, Wang L, Jiang N, Yang J, Li B, Luo Z. 2018. Amn1 governs post-mitotic cell separation in Saccharomyces cerevisiae.Barsh GS, editor. PLOS Genet. 14:e1007691.

Gerton JL, DeRisi J, Shroff R, Lichten M, Brown PO, Petes TD. 2000. Global mapping of meiotic recombination hotspots and coldspots in the yeast Saccharomyces cerevisiae. Proc. Natl. Acad. Sci. 97:11383–11390.

Glaser RL, Leach TJ, Ostrowski SE. 1997. The structure of heterochromatic DNA is altered in polyploid cells of Drosophila melanogaster. Mol. Cell. Biol. 17:1254–1263.

Henderson IR. 2012. Control of meiotic recombination frequency in plant genomes. Curr. Opin. Plant Biol. 15:556–561.

Hill WG, Robertson A. 2007. The effect of linkage on limits to artificial selection. Genet. Res. (Camb). 89:311–336.

Haldane JBS. 1930. Theoretical genetics of autotetraploids. J. Genet 22:359–372.

Jiang C, Pugh BF. 2009. Nucleosome positioning and gene regulation: advances through genomics. Nat Rev Genet 10: 161–172.

Jones GH, Franklin FCH. 2006. Meiotic Crossing-over: Obligation and Interference. Cell 126:246–248.

Kaplan N, Moore IK, Fondufe-Mittendorf Y, Gossett AJ, Tillo D, Field Y, LeProust EM, Hughes TR, Lieb JD, Widom J, et al. 2009. The DNA-encoded nucleosome organization of a eukaryotic genome. Nature 458:362–366.

Keeney S, Giroux CN, Kleckner N. 1997. Meiosis-Specific DNA Double-Strand Breaks Are Catalyzed by Spo11, a Member of a Widely Conserved Protein Family. Cell 88:375–384.

Lenormand T, Dutheil J. 2005. Recombination Difference between Sexes: A Role for Haploid Selection.Hurst LD, editor. PLoS Biol. 3:e63.

Lewis WH. 1980. Polyploidy: biological relevance. Springer US.

Lorenz MC, Muir RS, Lim E, McElver J, Weber SC, Heitman J. 1995. Gene disruption with PCR products in Saccharomyces cerevisiae. Gene 158:113–117.

Luo ZW, Zhang RM, Kearsey MJ. 2004. Theoretical basis for genetic linkage analysis in autotetraploid species. Proc. Natl. Acad. Sci. 101:7040–7045.

Malkova A, Swanson J, German M et al 2004. Gene conversion and crossing over along 405-kb left arm of *Saccharomyces cerevisiae* Chromosome VII. Genetics 168: 49–63.

Mather K. 1936. Segregation and linkage in autotetraploids. J. Genet. 32:287–314.

Mieczkowski PA, Dominska M, Buck MJ, Lieb JD, Petes TD. 2007. Loss of a histone deacetylase dramatically alters the genomic distribution of Spo11p-catalyzed DNA breaks in Saccharomyces cerevisiae. Proc. Natl. Acad. Sci. 104:3955–3960.

Bailey NTJ. 1961. Introduction to the mathematical theory of genetic linkage. Clarendon Press,Oxford

Oram KN. 1959. The Genetics of autotetraploid maize. PHD Univ. Adelaide.

Osman K, Higgins JD, Sanchez-Moran E, Armstrong SJ, Franklin FCH. 2011. Pathways to meiotic recombination in Arabidopsis thaliana. New Phytol. 190:523–544.

Otto SP, Barton NH. 1997. The evolution of recombination: Removing the limits to natural selection. Genetics 147:879–906.

Otto SP, Whitton J. 2000. Polyploid Incidence and Evolution. Annu. Rev. Genet. 34:401–437.

Pan J, Sasaki M, Kniewel R, Murakami H, Blitzblau HG, Tischfield SE, Zhu X, Neale MJ, Jasin M, Socci ND, et al. 2011. A Hierarchical Combination of Factors Shapes the Genome-wide Topography of Yeast Meiotic Recombination Initiation. Cell 144:719–731.

Pecinka A, Fang W, Rehmsmeier M, Levy AA, Mittelsten Scheid O. 2011. Polyploidization increases meiotic recombination frequency in Arabidopsis. BMC Biol. 9:24.

Petes TD. 2001. Meiotic recombination hot spots and cold spots. Nat. Rev. Genet. 2:360–369.

Prieler S, A P, V B, F K. 2005. The control of Spo11's interaction with meiotic recombination hotspots. Genes Dev. 19:255–269.

Fisher RA. 1947. The theory of linkage in polysomic inheritance. Philos. Trans. R. Soc. Lond. B. Biol. Sci. 233:55–87.

Shifman S, Bell JT, Copley RR, Taylor MS, Williams RW, Mott R, Flint J. 2006. A High-Resolution Single Nucleotide Polymorphism Genetic Map of the Mouse Genome.Gibbs R, editor. PLoS Biol. 4:e395.

Sved JA. 1964. The relationship between diploid and tetraploid recombination frequencies. Heredity (Edinb). 19:585–596.

Song K et al 1995 Rapid genome change in synthetic polyploids of Brassica and its implications for polyploid evolution *PNAS* **92**: 7719-23.

Swamy KBS, Chu WY, Wang CY, Tsai HK and Wang D. 2011. Evidence of association between nucleosome occupancy and the evolution of transcription factor binding sites in yeast BMC Evol Biol 11:150.

Wang L, Luo Z. 2012. Polyploidization increases meiotic recombination frequency in Arabidopsis: a close look at statistical modeling and data analysis. BMC Biol. 10:30.

Welch JE. 1962. Linkage in Autotetraploid Maize. Genetics 47:367–396.

Zhang H, Zheng R, Wang Y, Zhang Y, Hong P, Fang Yaping, Li G, Fang Yuda. 2019. The effects of Arabidopsis genome duplication on the chromatin organization and transcriptional regulation. Nucleic Acids Res. 47:7857–7869.

Ziolkowski PA, Berchowitz LE, Lambing C, Yelina NE, Zhao X, Kelly KA, Choi K, Ziolkowska L, June V, Sanchez-Moran E, et al. 2015. Juxtaposition of heterozygous and homozygous regions causes reciprocal crossover remodelling via interference during Arabidopsis meiosis. Elife 4.