ARTICLE

Check for updates

Meiosis at three loci in autotetraploids: Probabilities of gamete modes and genotypes without and with preferential cross-over formation

Cortland K. Griswold[1] ✉ and Saira Asif[1]

A long-standing goal in the field of polyploid biology has been the derivation of mathematical models of gamete mode formation. These models form the basis of statistical inference and evolutionary theory. Here, we present 3-locus models of gamete mode formation in autotetraploids without and with preferential cross-over formation. The three loci are assumed to occur on one arm of the same chromosome. For preferential cross-over formation, one of the three loci affects the tendency for sets of sister chromatids to pair and therefore affects rates of recombination. The models are derived such that the process of double reduction is a function of rates of synaptic partner switches and recombination, as opposed to being independent of these processes. We assume potentially one synaptic partner switch per meiosis. We also assume the coefficient of coincidence is one, such that cross-over events are independent, given a set of cross-over rates. Illustrative cases are examined demonstrating differences in the gamete mode probabilities without and with preferential cross-over formation. Lastly, we explore the accuracy of maximum likelihood estimates of the probability of synaptic partner switches and preferential cross-over formation when the locus controlling preference is at a proximal, middle, or distal location on the chromosome arm. All Supplementary Information is available at https://github.com/ckgriswold/3-locus-autotetraploid-meiosis.

INTRODUCTION

Autopolyploids are species or populations whose ancestors experienced a whole-genome duplication event. Among present-day species, autopolyploidy is most often studied in plants due to its frequent occurrence (Barker et al., 2016), but also in animals such as salmonids whose genome is a remnant of a whole-genome duplication with some loci still experiencing tetrasomic inheritance (Allendorf and Danzmann, 1997).

Meiosis in autopolyploids, in general, and autotetraploids, in particular, is complex. Whereas diploid meiosis involves the formation of bivalents, autotetraploid meiosis involves the formation of both bivalents and quadrivalents. The formation of quadrivalents allows for previously paired homologous chromosomes to switch partners, thus expanding the possibilities of recombination. Furthermore, if a paired-partner switch is coupled with a recombination event between a locus and the centromere, double reduction may occur at gamete formation, whereby segments of chromosomes on previously sister chromatids segregate together into the same gamete.

The theoretical modeling of meiosis in autotetraploids reflects the complexity of the biology of pairing and recombination. Fisher (1947) advanced the theory of polyploid meiosis by inventing the concept of a "gamete mode". A gamete mode indicates the chromosomal origin of an allele in a gamete. It simplifies and generalizes the potentially large number of gametes generated in an autopolyploid population. Yet, at the same time, Fisher (1947)

showed that there are still 107 modes across three loci and for a tetraploid species. More recently, Chen et al. (2021) derived probabilities of formation of these 107 modes. This work builds on Xu et al. (2013), Lu et al. (2012), and Luo et al. (2004), who decomposed gamete modes into those formed by double reduction, recombination or a combination of both or neither, such that in Chen et al. (2021) the probability of a gamete mode is modeled as a function of the probability of double reduction at the locus proximal to the centromere and recombination rates between loci. Modeling gamete mode probabilities this way is an important advance and leads to a straightforward decomposition of gamete modes. A consideration is that this approach implicitly assumes the probability of double reduction and recombination rates are free parameters, whereas double reduction is a function of probabilities of recombination and paired-partner switches. For example, in a two-locus setting, Rehmsmeier (2013) derived gamete mode probabilities using a model that included the processes of paired-partner switches and recombination, demonstrating that rates of double reduction are a function of recombination rates, as well as rates and locations of paired-partner switches. In statistical contexts such as Chen et al. (2021) and Xu et al. (2013), the dependence of the probability of double reduction on recombination rates is not necessarily problematic because they jointly estimate these variables and potentially allow for their covariance. Allowing the probability of double reduction and rates of recombination to be free variables in other contexts,

---

[1]Department of Integrative Biology, University of Guelph, Guelph, ON N1G 2W1, Canada. ✉email: cgriswol@uoguelph.ca

such as evolutionary genetics, may be problematic. For example, there may be selection on recombination rates (Yant et al., 2013), which would directly affect the rate of double reduction. In addition, the effect of different points of synaptic partner switches on gamete mode probabilities in combination with potentially varying rates of recombination may be of interest.

In this paper, we extend the Rehmsmeier (2013) approach to modeling autotetraploid meiosis to the three-locus case. In addition, we also add the process of preferential cross-over formation. Preferential pairing occurs when homologous chromosomes have different probabilities of pairing (Allendorf et al., 2015, Allendorf and Danzmann, 1997, Jenkins and Chatterjee, 1994). For example, in recent whole-genome duplicates of rye, homologous chromosome segments preferentially pair due to changes in sequence composition (Jenkins and Chatterjee, 1994). In salmonids most genomic segments currently experience preferential disomic inheritance, such that there is complete preferential pairing of previously homologous gene segments. Nevertheless, there are some segments that undergo tetrasomic inheritance and evidence that given tetrasomic inheritance, homologous chromosome segments do not pair with equal probability (Allendorf et al., 2015). Voorrips and Maliepaard (2012) provide a simulation model of tetraploid meiosis for multiple loci and that includes preferential pairing originating near telomeres. Our model differs from theirs in that it is mathematical and includes differential rates of preferential cross-over formation along a chromosome arm, as well as preferential pairing originating near telomeres. Yang et al. (2012) derived a preferential pairing model for allotetraploids in a three-locus setting and with the potential for homeologous pairing. In their model, a parameter affects the tendency for homologous versus homeologous pairing and assumes bivalency following pairing. The focus of the model was to estimate the degree of preferential pairing in allopolyploids and account for it in linkage analysis. Our approach differs in that we include both bivalent and tetravalent pairing (including synaptic partner switches), as well as variation in preferential cross-over formation across a chromosome (see next section). In particular, one of the three loci in our model specifically affects pairing.

Lloyd and Bomblies (2016) noted the link between preferential pairing and preferential cross-over formation. In the context of multiple loci and recombination, preferential pairing will manifest itself in terms of preferential cross-over formation because for a cross-over to occur sets of sister chromatids must be paired and synapsed. Preferential cross-over formation can be directly observed in terms of either differential cross-over events microscopically or indirectly in terms of its effects on recombination rates and gamete formation. Our model examines its effect on gamete formation.

Preferential pairing may have important consequences. Not only does it change the probabilities of gamete modes, it potentially alters the landscape of linkage disequilibrium (LD) and correspondingly may affect inferences associated with LD, such as quantitative trait locus (QTL) mapping. Accordingly, methods have been developed to quantify preferential pairing (Jannoo et al. 2004; Stift et al. 2008; Yang et al. 2012). Conceptually and theoretically, the evolution of preferential pairing may also be of interest. For example, Le Comber et al. (2010) presented a theoretical model of the evolution of preferential pairing, demonstrating that the process of genetic drift and neutral divergence in chromosome composition can result in the evolution of preferential pairing and disomic inheritance, with processes such as neo and subfunctionalization speeding up the transition time to disomy. Although Le Comber et al. (2010) showed that there is a general tendency toward preferential pairing (diploidization), a more general model of gamete mode formation in combination with other selective contexts may be informative about processes that potentially counteract evolution toward preferential pairing and thus maintain the integrity of tetrasomic inheritance and autotetraploidy.

A more general three-locus and mathematical model of meiosis may also be of use from an evolutionary perspective and in the absence of preferential pairing, as well. The central framework to model the evolution of recombination is a three-locus model, with two loci affecting fitness in an epistatic manner and a third locus affecting recombination rate (Nei, 1967, Otto and Feldman, 1997). Accordingly, the absence of a general three-locus meiotic model in autotetraploids is a barrier for mathematical models of the evolution of recombination in autotetraploids. Given that autotetraploids exhibit chromosomal gametic disequilibrium which affects epistatic selection (Griswold and Williamson, 2017), the evolution of recombination in autotetraploids may be different than in diploids. A three-locus model would also be useful in contexts other than the evolution of recombination. For example, the expansion of autotetraploids geographically is associated with genetic load caused by random genetic drift, as well as variation in the rate of selfing (Koshi et al., 2019). This combined with the potential for local adaptation (Griswold, 2021) would require at a minimum a three-locus model, with one affecting genetic load, a second adaptation, and a third selfing rate.

This paper first provides an overview of autotetraploid meiosis. Next, it derives gamete mode probabilities without preferential pairing and for three loci physically linked on one arm of a chromosome. The paper then derives gamete mode probabilities with preferential pairing and again for three loci. Here, it is assumed one locus affects the probability of pairing while the other two loci do not. The loci are physically linked on the same chromosome arm and three locations of the locus affecting the probability of pairing are considered relative to the centromere: distal to the two other loci, in the middle of the two loci, and proximal to the two loci. Besides deriving gamete modes, we also generate transition matrices that give the probability an autotetraploid genotype generates a particular gamete assuming two alleles at each locus. Lastly, we explore whether it is possible to accurately estimate the parameter associated with preferential pairing and whether the location of the locus affecting preferential pairing affects accuracy and/or precision of the estimate. In this context, we consider two parental genotypes that seem informative about estimating the probability of preferential pairing and investigate estimates of this parameter (and its standard deviation) as a function of the number of gametes sampled for the genotypes and for different positions of the locus affecting preferential pairing.

OVERVIEW OF AUTOTETRAPLOID MEIOSIS

An autotetraploid individual consists of four homologous copies of each chromosome. At the onset of meiosis these chromosomes duplicate generating four pairs of sister chromatids per chromosome. Sister chromatids are joined together at centromeres and sets of sister chromatids tend to cohere (Peters and Nishiyama, 2012). A set of sister chromatids then potentially pairs with other sets of homologous sister chromatids. Assuming (a) each set of sister chromatids pairs with at least one other set of sister chromatids, (b) pairing is random and (c) tetravalents can form, the probability sister chromatids pair into two sets of bivalents is 1/3 and the probability sister chromatids pair into one tetravalent is 2/3. This result follows from the observation that given two sets of sister chromatids have paired, one of the remaining sets of sister chromatids can pair with this set or the other unpaired set of sister chromatids. If an unpaired set of sister chromatids can pair with each set of sister chromatids with equal probability, then with probability 1/3 it will pair with the unpaired set of sister chromatids (resulting in two bivalents) and with probability 2/3 with one of the paired sets of sister chromatids (resulting in one tetravalent). In the case of the tetravalent, it is assumed that the remaining set of sister chromatids pairs with the trivalent. For a given autotetraploid population, the probability of bivalent versus

multivalent formation can deviate from these expectations (Parra-Nunez et al., 2019).

Double-stranded breaks occur along chromatids and form the basis for "bridges" between chromatids of paired sets of sister chromatids (Bomblies et al., 2016). Bridges between chromatids can potentially develop into cross-over (recombination) events. In tetravalents a set of paired chromatids can form bridges with chromatids from different sets of sister chromatids. This allows for one or more "synaptic partner switch(es)" (Lloyd and Bomblies, 2016), which is when a set of sister chromatids synapse with more than one other set of sister chromatids. The combination of one or more synaptic partner switch(es) and cross-over events can result in a set of paired sister chromatids recombining with more than one set of sister chromatids. Furthermore, "preferential cross-over formation" (Lloyd and Bomblies, 2016) may occur among sets of sister chromatids. Preferential cross-over formation occurs when sets of sister chromatids experience differential cross-over rates.

Lloyd and Bomblies (2016) introduced the terms "synaptic partner switch" and "preferential cross-over formation" to overcome confusion regarding the historical use of terms "preferential pairing" and "paired-partner switch". Paired-partner switches occur prior to synapse, but not all of these paired-partner switches are retained at synapse. A synaptic partner switch is a switch in paired-partners at synapse. The introduction of the term "preferential cross-over formation" is useful because it more clearly operationalizes the concept of "preferential pairing". In particular, preferential cross-over formation is directly observed through differential cross-over formation among sets of sister chromatids microscopically or indirectly through the effect of differential cross-over rates on gamete frequencies. Preferential cross-over formation between sets of sister chromatids can vary along a chromosome.

Our three-locus extension of Rehmsmeier (2013) model allows for variable rates of bivalent versus tetravalent formation. The model assumes the three loci occur on the same chromosome arm, with proximal, middle, and distal positions relative to the centromere. Given tetravalent formation, it allows for a single synaptic partner switch either between the centromere and the proximal locus, between the proximal and the middle loci, or between the middle and the distal loci. Cross-over events may occur between loci, as well as between the centromere and the proximal locus. With no preferential pairing, cross-over probabilities are not affected by genotype. With preferential pairing, cross-over formation is affected by genotype.

We study two models of preferential cross-over formation. For both models, either the proximal, middle, or distal locus is biallelic and affects the probability of cross-over formation. In the first model, this locus affects cross-over formation *locally* at regions immediately upstream and downstream of the locus. In the second model, this locus affects cross-over formation at multiple regions *downstream* of the locus toward the centromere. This choice of two models reflects our current state of knowledge. Our understanding has been that homologous chromosome pairing is initialized near the telomeric region and subsequent pairing and cross-over events occur toward the centromere (Sved, 1966, Armstrong et al., 2001, Armstrong and Jones, 2003, Bass et al., 2000, Lopez et al., 2008), with models following these observations (Morgan et al., 2021, Sybenga, 1975, Voorrips and Maliepaard, 2012). It is also possible that although telomeres attach to the nuclear envelope, this does not necessarily initiate pairing. Instead, initial pairing may occur elsewhere (Sybenga, 1999), and cross-over formation is locally controlled. We study both models to determine their consequences on gamete mode probabilities. We do this because our state of knowledge is still in flux and either the local or downstream models, or both, may apply under certain circumstances. If two paired sets of sister chromatids have in common an allele that affects pairing, then they recombine at the baseline rate (Fig. 1a). Otherwise, the rate of recombination is reduced at locations either immediately adjacent
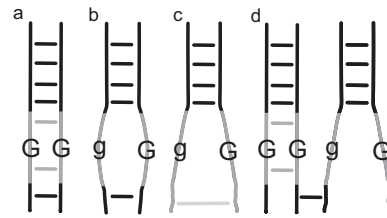

Fig. 1 Local and downstream models of preferential cross-over formation. In parts **a**, **b**, and **c** paired sets of sister chromatids are represented. A set of sister chromatids is represented by a single vertical line, as opposed to two vertical lines for simplicity. Horizontal lines correspond to bridges between paired sets of sister chromatids. In **d** two paired sets of sister chromatids are given. A locus with allelic states G or g affects cross-over formation. For each set of sister chromatids, the state at the $G\|g$ locus is given, where "$G\|g$" indicates a locus with either the G or the g allele. Note, although each set of sister chromatids has two alleles at the $G\|g$ locus, we are assuming they are the same and one letter is used to indicate the state of both alleles of a set of sister chromatids. In part **a**, one set of sister chromatids has G alleles at the $G\|g$ locus and the other set of sister chromatids also has G at the $G\|g$ locus. In contrast in parts **b** and **c**, one set of sister chromatids has G whereas the other set of sister chromatids has g. Different alleles at this locus lead to a loss of affinity between sets of sister chromatids. In part **a**, the paired sets of sister chromatids share the same allele at the $G\|g$ locus and bridges form between chromatids as normal, such that recombination occurs at the baseline rate. In contrast in part **b**, different alleles occur at the $G\|g$ locus and its action is local, which weakens the affinity of the sets of sister chromatids, resulting in fewer bridges and a reduced recombination rate in regions adjacent to the $G\|g$ locus. In part **c**, the action of the $G\|g$ locus is downstream, such that affinity is lost further down the chromosomes. In part **d** a synaptic partner switch occurs and the recombination rate is not affected by the state of the $G\|g$ locus below the synaptic partner switch.

to the locus (Fig. 1b, local model), or all locations downstream toward the centromere (Fig. 1c, downstream model). A synaptic partner switch further upstream or downstream of the locus in the first model, or downstream toward the centromere in the second model decouples the rest of the chromosome from the locus and cross-over rates remain at their baseline levels (Fig. 1d). The locus affecting preferential pairing can be thought of as representing a chromosome segment which affects the tendency for a cross-over event to occur within adjacent segments or segments downstream, up to a synaptic partner switch.

Two models of initial chromosomal pairing are presented. In the first model, it is assumed that sets of paired chromatids pair at random initially. Accordingly, this version of the model focuses on either no preferential pairing or differential preferential pairing along a chromosome. The second version assumes the locus affecting preferential pairing is at the distal location and this locus occurs near the telomere, such that it controls the initialization of pairing at the distal end of a chromosome. This version of the model follows Voorrips and Maliepaard (2012) and has links to allopolyploid meiosis, but with the potential for homeologous pairing.

Before continuing it is worth reviewing terminology in combination with our understanding of meiosis in autotetraploids. A synaptic partner switch is a switch in a paired-partner at the time of synapse. Prior to synapse double-stranded breaks may result in partner switches, but these may not mature to a synaptic partner switch. In our model, we define parameters that give the probabilities of synaptic partner switches because these are the resultant switches that may affect gamete formation. In our model, there are probabilities of synaptic partner switches and separate probabilities for cross-overs. This approach to modeling is consistent with observation from our understanding. For example, a synaptic partner switch does not necessarily

correspond in a one-to-one manner to cross-over. A synaptic partner switch can occur without a cross-over (Morgan et al., 2021). In our model probabilities of synaptic partner switches are realized values. A recent finding is that cross-over interference results in a tendency toward bivalency in established autotetraploids versus neotetraploids (Morgan et al., 2021). Accordingly, the parameter that gives the rate of tetravalency in our model is the realization of this interference process. Similarly, parameters for cross-over rates correspond to the realized rate as a consequence of cross-over interference. This assumption allows for local cross-over interference to reduce the average cross-over rate within a chromosome region, but our model does not directly model cross-over interference among loci, such that we assume coefficients of coincidence (*cf.* Christiansen, 2008) are to equal one among loci. We follow the historical use of the term "preferential". Unfortunately, this suggests "seeks out" as in homologous chromosomes "seek out" other homologous chromosomes. What seems more likely is chromosomes bump into one another more or less at random initially (either near the telomeric end or more centrally), recognize homology, establish bridges, which then build from there. Our approach to modeling follows the "bump around" hypothesis. In particular, we assume homologous chromosomes initially pair at random. The action of the $G \| g$ locus (Fig. 1) mimics whether strong or weak bridges (or numerous versus sparse bridges) are formed depending on allelic state. Strong versus weak bridges result in what appears to be "preferential cross-over formation". In our model a baseline rate of cross-over is reduced by a factor when alleles at the $G \| g$ locus are different, reflecting what would be observed microscopically or in terms of gamete frequencies - an apparent reduction in cross-over rate. There is a lot to discover in terms of what "allelic state" means at a $G \| g$ locus. One possibility is at this locus an allelic state is an insertion of a nonhomologous segment, which would reduce recognition of homology at double-stranded breaks. Others are possible. Lastly, we follow the gamete mode approach of Fisher (1947). In Fisher's approach, an autotetraploid parental genotype is labeled

$$a_1b_1c_1/a_2b_2c_2/a_3b_3c_3/a_4b_4c_4,$$

where letters designate loci and subscript numbers distinguish each of four homologous chromosomes. After chromosome doubling there are four sets of sister chromatids:

$a_1b_1c_1/a_1b_1c_1$, $a_2b_2c_2/a_2b_2c_2$, $a_3b_3c_3/a_3b_3c_3$ and $a_4b_4c_4/a_4b_4c_4$. Subscripts in gamete modes indicates the chromosome of origin of an allele. For example, one possible gamete mode is $a_ib_ic_j/a_kb_kc_k$. This mode indicates that on one of the homologous chromosomes in a diploid gamete, the alleles at the "*a*" and "*b*" loci came from the same chromosome, whereas the allele at the "*c*" locus came from a different chromosome. Furthermore, at the second homologous chromosome, the "*a*" and "*b*" loci also came from the same chromosome, but one that is different from both the "*a*" and "*b*" loci and the "*c*" locus on the first chromosome in the gamete. And, in this case, the allele at the "*c*" locus also came from the same chromosome as the "*a*" and "*b*" loci on this second chromosome. Although there are 107 gamete modes, which is a lot, it is far less than the number of gamete genotypes and in fact gamete genotypes are just special cases of gamete modes. To conclude, it is important to keep in mind that a gamete mode indicates the chromosome origins of alleles in a gamete, and that secondly there are probabilities associated with gamete modes depending on the underlying meiotic process.

GAMETE MODE PROBABILITIES
No preferential cross-over formation nor pairing
Bivalency versus tetravalency. Sets of sister chromatids form a tetravalent with probability τ and two bivalents with probability $(1-\tau)$. Initial pairing occurs with equal probability among sets of
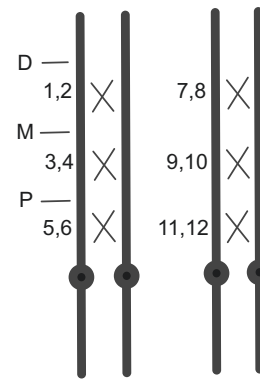

Fig. 2 Two bivalents with potential cross-over locations indicated. A set of sister chromatids is represented by a single line with a circle indicating the location of centromeres. Loci are labeled distal (**D**), middle (**M**), and proximal (**P**). There are potentially twelve cross-over events for the two sets of bivalents. Numbers indicate algorithmically the order events are processed, such that cross-over events occur from the distal locus to the centromere. Two events occur between loci or between the proximal locus and the centromere because each sister chromatid may recombine with one other paired sister chromatid in opposite sets. Note this figure closely parallels Rehmsmeier (2013) Fig. 2, part **E**, except for the addition of the middle locus and corresponding increase in the number of events. In addition, we represent a set of sister chromatids with a single line, as opposed to two adjacent lines, to simplify the figure.

sister chromatids and there are twenty-four possible ordered pairs. For bivalents the twenty-four ordered pairs reduce to three unordered sets of two bivalents. For tetravalents, ordering matters in terms of which sets of sister chromatids initially pair and which potentially undergo a synaptic partner switch (see next section). Once an ordered of set of paired sister chromatids is generated, cross-over and synaptic partner switch (tetravalent) events occur from the distal locus to the centromere.

Cross-over and synaptic partner switch events. **Bivalents** Only cross-over events may occur (Fig. 2). Six cross-over events may occur per paired set of sister chromatids totaling twelve possible events. The state space per event is *{no cross-over, cross-over}* with associated probabilities $\{1-x, x$ for $x \in \{v, q, r\}$ with *v*, *q*, and *r* corresponding to cross-over probabilities between the centromere and proximal locus, the proximal and middle loci and the middle and distal loci, respectively. Given a state space of size two per event, the total size of the state space across twelve events is $2^{12} = 4096$. Algorithmically, a list of the 4096 cross-over products is generated with associated algebraic expressions for their respective probabilities for each of the three unordered sets of bivalents.

Tetravalent Both cross-over events and a synaptic partner switch event may occur (Fig. 3). Following Rehmsmeier (2013) a cross-over may occur both above and below a synaptic partner switch between two loci. Since *v*, *q*, and *r* equal realized recombination rates between loci, in a synaptic partner switch region, cross-over rates are v', q', and r' for each event immediately above and below the switch, and where $x = 2x'(1-x')$ for $x' \in \{v', q', r'\}$. Synaptic partner switches occur with probabilities p_{cp}, p_{pm}, and p_{md} between the centromere and proximal locus, the proximal and middle loci and the middle and distal loci, respectively, and assuming $p_{cp} + p_{pm} + p_{md} \leq 1$. It is assumed that the "inner" sets of sister chromatids undergo the synaptic partner switch. With probability $1 - p_{cp} - p_{pm} - p_{md}$ no synaptic partner switch occurs. In this case, the tetravalent is treated as a bivalent. With probability $p_{cp} + p_{pm} + p_{md}$ a switch occurs and there is a synaptic tetravalent. The state space per cross-over event remains as *{no cross-over, cross-over}* with
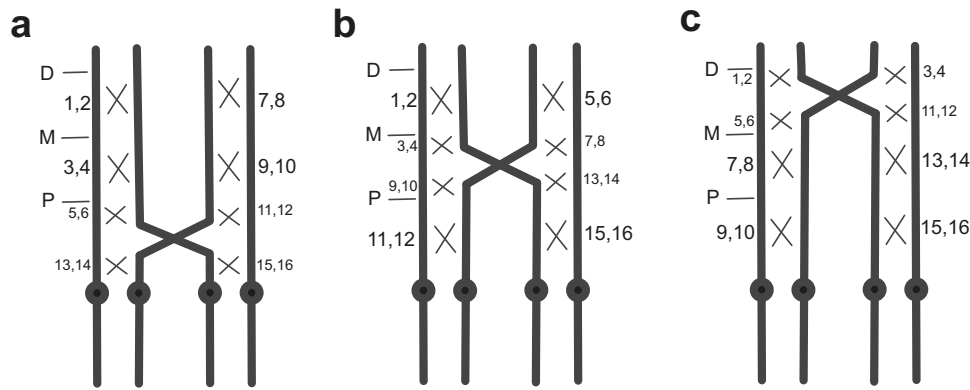
Fig. 3 Three possible tetravalents depending on the location of a synaptic partner switch. Synaptic partner switches may occur between the **a** proximal locus and centromere, **b** middle and proximal loci, and **c** distal and middle loci. A tetravalent without a synaptic partner switch equates to the bivalent case in terms of cross-over events (Fig. 2). Numbers correspond to cross-over events, with the addition of events immediately above and below a synaptic partner switch. This figure closely parallels Rehmsmeier (2013) Fig. 2, parts **a** and **b**, except for the addition of the middle locus, synaptic partner switch location, and cross-over events.

associated probabilities $\{1-x, x\}$ or $\{1-x', x'\}$, as appropriate. Together there are 16 events with a state space of size $2^{16} = 65,536$ combined cross-over and switch products and associated algebraic expressions for their respective probabilities for each of the twenty-four sets of ordered tetravalents and for each switch location. Across all three switch locations, there are $3 \times 65,536 = 196,608$ combined cross-over and switch products.

Meta and anaphases. Paired sets of sister chromatids orient randomly during metaphase 1 and dissociate during anaphase. Similarly, sister chromatids orient randomly during metaphase 2 and dissociate. There are two orientations per paired sets of sister chromatids, such that there are together $2^2 = 4$ combined orientations at metaphase 1. Similarly, there are $2^2 = 4$ orientations of sister chromatids in each divided cell at metaphase 2, such that there are $2^2 \times 2^2 = 16$ gamete products for each of the 4096 cross-over products from bivalents and for each of the 196,608 combined cross-over and switch products from tetravalents.

Mode probabilities. Each of the $16 \times 4096 = 65,536$ gamete products for bivalents and $16 \times 196,608 = 3,145,728$ gamete products for tetravalents are sorted into the 107 gamete modes of Fisher (1947) and for each mode the probabilities of gamete products that form a mode are added to get the gamete mode probability. We use the gamete mode numbering of Lu et al. (2012), Appendix 1. A python file that encodes the meiotic model with no preferential cross-over formation and calculates the probability of a specified mode is provided as Supplementary Information - 1 - No preferential cross-over formation—Python code—Gamete Mode Probability (available at https://github.com/ckgriswold/3-locus-autotetraploid-meiosis, as well as this paper's webpage). Probabilities for the 107 gamete modes are assembled in the python notebook Supplementary Information - 2 - No preferential cross-over formation —Python notebook—Compile Gamete Mode Probabilities and a list of simplified algebraic expressions of gamete mode probabilities is provided in Supplementary Information - 3—No preferential cross-over formation - List of Gamete Mode Probabilities. Note further that a Supplementary Information - 0 - ReadMe file is provided for shared reference information across all files.

Preferential cross-over formation or pairing
One out of the three loci affects preferential cross-over formation. This locus is diallelic with alleles G and g. There is no mutation between allelic states during meiosis. For a given location of the $G\|g$ locus there are $2^4 = 16$ parental genotypes using Fisher (1947) notation. For example, if the $G\|g$ locus occurs in the middle a

generalized Fisher genotype is $a_1G\|gb_1/a_2G\|gb_2/a_3G\|gb_3/a_4G\|gb_4$, where $G\|g$ indicates G or g. Below we assume that sister chromatids with the G allele prefer sister chromatids with a G allele and sister chromatids with the g allele prefer sister chromatids with the g allele. Furthermore, given a match between alleles at the $G\|g$ locus, preferences are equal for G and g alleles. Following this assumption and recognition that the ordering of chromosomes in a parent is exchangeable, there are functionally three distinct parental genotypes in terms of preferential cross-over formation:

$$a_iGb_i/a_jGb_j/a_kGb_k/a_\ell Gb_\ell,$$

$$a_iGb_i/a_jGb_j/a_kGb_k/a_\ell gb_\ell$$

and $a_iGb_i/a_jGb_j/a_kgb_k/a_\ell gb_\ell$. Note that the genotype $a_igb_i/a_jgb_j/a_kgb_k/a_\ell gb_\ell$ is functionally the same in terms of preferential pairing as $a_iGb_i/a_jGb_j/a_kGb_k/a_\ell Gb_\ell$ and $a_igb_i/a_jgb_j/a_kGb_k/a_\ell Gb_\ell$ is the same as $a_iGb_i/a_jGb_j/a_kGb_k/a_\ell gb_\ell$. By functionally the same it is meant that gamete mode probabilities are the same with an exchange of g for G and vice versa.

Besides calculating gamete mode probabilities for parental genotypes

$$a_iGb_i/a_jGb_j/a_kGb_k/a_\ell Gb_\ell,$$

$$a_iGb_i/a_jGb_j/a_kGb_k/a_\ell gb_\ell$$

and $a_iGb_i/a_jGb_j/a_kgb_k/a_\ell gb_\ell$, we also calculate gamete mode probabilities for the parental genotype
$a_iGb_i/a_jGb_j/a_kGb_k/a_\ell G\&\&gb_\ell$, where $G\&\&g$ indicates G and g on duplicates of the ℓth homologous chromosome. We include this because it corresponds the potential for a mutation from the G to the g state.

$G\|g$ *locus is not associated with the telomere and does not affect initial pairing.* **Bivalency versus tetravalency** Initial pairing follows the no preferential pairing case. In the algorithm, there continues to be three unordered sets of two bivalents and twenty-four ordered sets of tetravalents. Some of these sets for both bivalents and tetravalents are equal due to redundancy at the locus affecting cross-over formation. This reduces the efficiency of the algorithm, but not its accuracy. Sets of sister chromatids form a tetravalent with probability τ and two bivalents with probability $(1-\tau)$.

Cross-over and synaptic partner switch events. **Bivalents** There are still three loci, such that six cross-over events may occur per paired
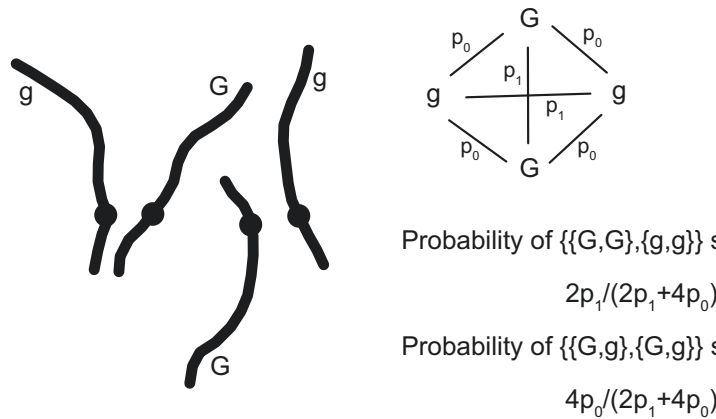
Probability of {{G,G},{g,g}} set of pairings:

$$2p_1/(2p_1+4p_0)$$

Probability of {{G,g},{G,g}} set of pairings:

$$4p_0/(2p_1+4p_0)$$

Fig. 4 Illustrates scramble pairing of sets of sister chromatids and the $G\|g$ locus occurring at the telomeric end of the chromosome. To the left are four duplicated chromosomes, whereas in previous figures sister chromatids are represented together in a single line. The parental genotype was *GGgg* at the $G\|g$ locus. Following the assumption that sister chromatids closely cohere physically, the upper right diagram depicts possible sets of initial pairs of sets of sister chromatids and their corresponding rate of pairing. At the bottom right is the probability a bivalent or tetravalent involves the pairing of two sets of sister chromatids with *G* alleles and two sets with *g* alleles versus two sets of sister chromatids, with each set having different alleles at the $G\|g$ locus.

set of sister chromatids totaling twelve possible events. If paired sets of sister chromatids have different alleles at the $G\|g$ locus, the cross-over rate is reduced by a factor $(1-p_{pref})$ for recombination points adjacent (local) to or downstream of the $G\|g$ locus and toward the centromere. If $p_{pref} = 0$ there is no preferential cross-over formation and if $p_{pref} = 1$ there is complete preferential cross-over formation. For $0 < p_{pref} < 1$ there is intermediate preferential cross-over formation. Accordingly, at recombination points local to or downstream of the $G\|g$ locus, recombination probabilities are $x(1-p_{pref})$ for $x \in \{v, q, r\}$. For the downstream model, recombination probabilities remain equal to x upstream of the $G\|g$ locus, for $x \in \{v, q, r\}$. As with the no preferential cross-over case, a list of the 4096 cross-over products is generated with associated algebraic expressions for their respective probabilities for each of the three unordered sets of bivalents. Note, the factor $(1-p_{pref})$ reflects a reduction in affinity between chromosomes, the formation of bridges, and consequently a reduction in the rate of cross-over formation at the time of synapse that would otherwise occur at rate x for $x \in \{v, q, r\}$.

Tetravalents Like bivalents, if paired sets of sister chromatids have different alleles at the $G\|g$ locus the cross-over rate is reduced by a factor $(1-p_{pref})$ for recombination points local to or downstream of the $G\|g$ locus and toward the centromere. Cross-over rates v, q, and r continued to be realized recombination rates and where $x(1 - p_{pref}) = 2x'(1 - x')(1 - p_{pref})$ for $x' \in \{v', q', r'\}$. Synaptic partner switches occur independently of the state of the $G\|g$ locus and with probabilities p_{cp}, p_{pm} and p_{md} for $p_{cp} + p_{pm} + p_{md} \leq 1$. Beyond adjacent sites to the $G\|g$ locus (local model) or downstream of the $G\|g$ locus and a synaptic partner switch, recombination rates are unaffected by the state of the $G\|g$ locus at newly formed pairs, such that recombination rates are equal to x for $x \in \{v, q, r\}$. In principle, paired sets of sister chromatids that share either a G or a g allele may have a lower probability of initiating a switch than a paired set that have different alleles. We leave this possibility for a future extension of this work. The algorithm for generating meiotic products parallels the no preference scenario, such that for a given location of the $G\|g$ locus and for a given parental genotype at the $G\|g$ locus, there are 196,608 combined cross-over and switch products.

Meta and anaphases These phases are the same as the no preferential cross-over scenario.

Mode probabilities Since one of the loci is biallelic, the number of gamete modes reduces to 37. Gamete modes for each of the three parental genotypes and three locations of the $G\|g$ locus are provided in Appendix 2. (Note, chromosome origins are the same for a given

mode across the proximal, middle and distal locations of the $G\|g$ locus, just the location of the $G\|g$ locus changes). For each combination of parental genotype and location of the $G\|g$ locus, gamete products are sorted into modes and the combined probability of generating a mode is calculated. Python files that encode the meiotic model with preferential cross-over formation are provided as Supplementary Information—[4-9]—Preferential cross-over formation - A—Python code—Gamete Mode Probabilities—Q for $A \in \{downstream, adjacent\}$ depending on the action and $Q \in \{proximal, middle, distal\}$ depending on the location of the $G\|g$ locus. Compilation files of gamete modes are provided in Supplementary Information - [10–15] - Preferential Cross-over formation - A - Python notebook—Compile Gamete Mode Probabilities—Q for $A \in \{downstream, adjacent\}$ and $Q \in \{proximal, middle, distal\}$. A list of probabilities for the 37 gamete modes for each location of the $G\|g$ locus are provided in Supplementary Information - [16–21] - Preferential cross-over formation - A - List of Gamete mode probabilities - Q for $A \in \{downstream, adjacent\}$ and $Q \in \{proximal, middle, distal\}$. In these lists the first gamete probability corresponds to a *GGGG* parent, the second a *GGGg* parent, the third a *GGgg* parent, and the fourth a *GGG,G&&g* parent.

Telomere-associated G‖g *locus.* The $G\|g$ locus is distal and in close physical linkage with the telomeric region.

Bivalency versus tetravalency, as well as pairing As with previous scenarios, sister chromatids form a tetravalent with probability τ and two bivalents with probability $(1-\tau)$. Pairing occurs via a "scramble". Sister chromatids with two alleles at the $G\|g$ locus of equal type pair at rate p_1 and alleles of unequal type pair at rate p_0. Consider the meiotic genotype $\{G, G, g, g\}$ immediately after duplication, where a "G" indicates a set of sister chromatids with the allelic state "G" at each chromatid. The probability of a $\{\{G, G\}, \{g, g\}\}$ genotype at synapsis is $2p_1/(4p_0 + 2p_1)$, Fig. 4, where $\{G, G\}$ corresponds to a G pairing with a G. Correspondingly, the probability of a $\{\{G, g\}, \{G, g\}\}$ genotype at synapsis is $4p_0/(4p_0 + 2p_1)$. All other types of genotypes at the G/g locus and for other counts of the G and g alleles occur with equal probability.

Cross-over and synaptic partner switch events, as well as other phases and gamete modes The configuration of paired sets of sister chromatids present the stage for synaptic partner switch events and cross-overs. All recombination events are downstream toward the centromere, such that if alleles at the $G\|g$ locus are not equal for paired sets of sister chromatids, then the recombination rate is reduced by a fraction $(1-p_{pref})$. Synaptic

partner switches follow earlier scenarios. Subsequent meiotic phases are the same as the earlier cases. Gamete modes are the same as the distal case from the local (or downstream) scenario, but their probabilities are different. Under the telomere-associated scenario, gamete modes are a function of rates p_0 and p_1. A python file that encodes the meiotic model with preferential cross-over formation near the telomere is provided as Supplementary Information - 22 - Preferential cross-over formation—Scramble—Python Code—Gamete Mode Probability. The compilation of gamete mode probabilities is provided in Supplementary Information - 23 - Preferential cross-over formation - Scramble—Python Notebook - Compiled Gamete Mode Probabilities. A list of probabilities for the 37 gamete modes is provided in Supplementary Information - 24 - Preferential cross-over formation - Scramble - List of Gamete mode probabilities. As before, for each gamete mode, there are four probabilities associated with the parental genotypes *GGGG*, *GGGg*, *GGgg*, and *GGG,G&&g*.

DIALLELIC GAMETE PROBABILITIES

Across three diallelic loci there are 330 unordered autotetraploid genotypes and 36 gamete genotypes. Numbered lists of these genotypes are provided in Supplementary Information - 25 - Diallelic genotype list and Supplementary Information - 26 - Diallelic gamete list. For each autotetraploid genotype, the set of resultant gamete genotypes can be generated for each gamete mode. From this, a transition matrix can be generated that gives the probability an autotetraploid genotype generates a gamete genotype as a function of gamete mode probabilities. The following python notebooks provide python code that generates transition matrices: Supplementary Information - 27 - No preferential cross-over formation - Python Notebook - Genotype-gamete transition matrix and Supplementary Information - [28–30] - Preferential cross-over formation - Python Notebook - Genotype-gamete transition matrix - Q for $Q \in \{proximal, middle, distal\}$. Note that the proximal, middle, and distal versions are valid for both downstream and adjacent action of the $G\|g$ locus and the non-telomere and telomere-associated cases because they share the same gamete modes and these matrices are presented at the level of modes without specific probabilities substituted in yet. The following files provide transition matrices in terms of gamete modes without substitution of specific probabilities: Supplementary Information - 31 - No preferential cross-over formation - Genotype-gamete transition matrix and Supplementary Information - [32–34] Preferential cross-over formation - Genotype-gamete transition matrix - Q for $Q \in \{proximal, middle, distal\}$. To get transition probabilities in terms of rates of bivalent versus tetravalent formation, cross-over rates, etc. gamete mode probabilities from earlier files can be substituted in for gamete mode probabilities designated by the gamete mode number in the transition matrix files. Python notebooks that do this are provided in Supplementary Information - 35 - No preferential cross-over formation - Substitute algebraic probabilities and Supplementary Information - [36–42] - Preferential cross-over formation - A - Substitute algebraic probabilities - Q for $A \in \{downstream, adjacent\}$ and $Q \in \{proximal, middle, distal, distal-scramble\}$.

ESTIMATING RATES OF SYNAPTIC SWITCHES AND PREFERENTIAL CROSS-OVER FORMATION

Two parental genotypes that seem informative about the occurrences and rates of synaptic switches and preferential cross-over formation are

$A_1B_1G/A_1B_1G/A_1B_1G/A_2B_2g$ and $A_1B_1G/A_1B_1G/A_2B_2g/A_2B_2g$,

where in the examples the $G\|g$ locus is distal, but could also be in the middle or proximal. The reason these genotypes seem informative is that haplotypes across the $A_1\|A_2$ and $B_1\|B_2$ loci

are physically linked to either the G or g allele, such that if a recombination event between a haplotype with a G allele occurs with a haplotype with a g allele, it is detectable. Furthermore, we expect some preferential cross-over formation because of paired sets of sister chromatids with a G and g allele.

Given transition matrices from parental genotype to gamete genotype, it is straightforward to develop a maximum likelihood model for gamete counts as a result of meiosis. Gamete counts are multinomially distributed for a given set of parameters, such that over a sample of n randomly sampled gametes from a genotype the negative log-likelihood of model parameters (M) given the sample (D) follows:

$$-\log L(M|D) \propto -\sum_k n_k \log p_k(M)$$

where $\propto$ indicates proportionality, $p_k(M)$ is the probability of generating the kth type of gamete for model parameters M, and n_k is the number of gametes of type k in the sample. Note the likelihood equation follows from the multinomial distribution which is for K potential gametes in a sample of size n

$$\frac{n!}{n_1!n_2!n_3!\cdots n_K!}p_1^{n_1}p_2^{n2}p_3^{n3}\cdots p_K^{n_K}.$$

The log of this distribution is

$$\log\frac{n!}{n_1!n_2!n_3!\cdots n_K!}+\sum_k n_k \log p_k.$$

In an experiment, n as well as n_k for $k \in \{1, 2, 3,\dots K\}$ are fixed observations, such that $\log n!/(n_1!n_2!n_3!\cdots n_K!)$ is a constant, which corresponds to the use of proportionality in the likelihood equation. The terms p_k are functions of underlying parameters such as the probability of tetravalency, the probabilities of various synaptic partner switches, the probabilities of cross-over, etc.

If all parameters were free to vary, it would be a high dimensional inference, such that it may be difficult to find global maximum likelihoods. Here, we assume that the loci are physically distant, such that recombination rates are assumed to be equal to 1/2. Furthermore, we assume bivalent to tetravalent ratios equal the 1:2 ratio as if these form randomly. Lastly, we assume synaptic partner switch probabilities are equal across locations. Together these assumptions reduce the dimensionality of the inference to two, one parameter being the rate of synaptic switching (p_{switch}) and the second the degree of preferential cross-over formation (p_{pref}).

We examine p_{switch} values equal to 1/100 and 1/4, as well as p_{pref} values equal to 95/100 and 50/100. Power to detect a deviation in p_{pref} from 0 likely arises through cases when a switch occurs and the consequence of upstream or local preferential cross-over is nullified. Accordingly, a limit to detecting a deviation in p_{pref} from 0 may be smaller values of p_{switch}. At the parental genotype level, the genotype $A_1B_1G/A_1B_1G/A_1B_1G/A_2B_2g$ will always incur the expression of preferential cross-over formation for one set of paired sets of sister chromatids, whereas the genotype $A_1B_1G/A_1B_1G/A_2B_2g/A_2B_2g$ will not. This difference may lead to differences in power to detect preferential cross-over formation.

In the inference approach, it is assumed that the state at the $G\|g$ locus is hidden. The location of the $G\|g$ locus may also affect the power to detect preferential cross-over formation. When the action of the $G\|g$ is downstream, a synaptic partner switch can manifest itself across two loci if $G\|g$ is distal, but fewer loci if the $G\|g$ locus is in the middle or proximal. When the action of the $G\|g$ is local, we found little to no effect on gamete mode probabilities if its location is proximal or distal. We therefore only examine the middle case.

This paper derived gamete probabilities for the $A_1B_1G/A_1B_1G/A_1B_1G/A_2B_2g$ and $A_1B_1G/A_1B_1G/A_2B_2g/A_2B_2g$ genotypes (see

previous sections). For a set of underlying parameters, we then randomly generated samples of gametes for each genotype using the *random.multinomial* function in the *numpy* python library (vers. 1.13.1). Based on these samples, we then numerically determined maximum likelihood estimates using the Scipy function *minimize* (vers. 0.19.0).

RESULTS AND DISCUSSION

The main results of this paper are the gamete mode probabilities and genotype-gamete transition matrices (Supplementary Information files). Below we compare gamete mode probabilities from the two-locus case of Rehmsmeier ([2013](#)) and without preferential cross-over formation versus the three-locus case with preferential cross-over formation for two informative cases. In addition, for the same cases, we compare gamete mode probabilities for the distal location of the $G\|g$ locus when that locus is not associated versus associated with the telomeric region. Lastly, we examine the accuracy of estimates of the rates of synaptic partner switches and preferential cross-over formation.

Comparisons of gamete mode probabilities

With and without preferential cross-over formation, as well as among locations of the G∥g locus. We begin by assuming the $G\|g$ locus is distal, which leads to a more direct comparison to Rehmsmeier ([2013](#)) two-locus result. The two modes are a_ib_i/a_jb_j and a_ib_j/a_kb_ℓ, and where we have suppressed the state of the $G\|g$ locus for the three-locus case because we sum across all states at the $G\|g$ locus. In addition, we assume the parental genotype at the $G\|g$ locus is *GGgg*. The mode probabilities for the a_ib_i/a_jb_j mode are

Rehmsmeier ([2013](#)):

$$1 - 2q + q^2 + p_{cp}\left(-\frac{v}{4} + \frac{q}{4} + \frac{vq}{2} - \frac{q^2}{4} - \frac{vq^2}{4}\right)\tau + p_{pm}\left(\frac{5q}{8} - \frac{13q^2}{16} - q' + 7\frac{qq'}{8}\right)\tau$$

3-locus model, with local affects on cross-over formation: equal to Rehmsmeier ([2013](#)) without preferential pairing

3-locus model, with downstream affects on cross-over formation:

$$1 - 2\frac{q}{3} + \frac{q^2}{3} - 4\frac{(1-p_{pref})q}{3} + 2\frac{(1-p_{pref})^2q^2}{3}$$
$$+ p_{cp}\left(-\frac{v'}{6} + \frac{v'^2}{6} + \frac{q}{12} + \frac{qv'}{3} - \frac{qv'^2}{3} - \frac{q^2}{12} - \frac{v'q^2}{6} + \dots \{8\}\dots + \frac{(1-p_{pref})^4v'^2q^2}{3}\right)\tau$$
$$+ p_{pm}\left(-31\frac{q'}{12} + 29\frac{q'^2}{12} - 13\frac{q'^3}{12} + \frac{q'^4}{4} + \dots \{9\}\dots - 2\frac{(1-p_{pref})^2q^2}{3}\right)\tau$$
$$+ p_{md}\left(-4\frac{q}{3} + 2\frac{q^2}{3} + 4\frac{(1-p_{pref})q}{3} - 2\frac{(1-p_{pref})^2q^2}{3}\right)\tau$$

(1)

where parameters in Rehmsmeier are recast to correspond to this paper's use of parameters. We use the notation …{X}… to indicate *X* additional terms in the expression for the three-locus model. We do not include these terms to make the equation readable in the main paper. The full expression is provided in the Supplementary Information. A point to draw out from comparisons of mode probabilities is that when the $G\|g$ locus is distal and has local effects on pairing, this does not affect gamete mode probabilities, such that they are the same as the no preferential case. A second point is the addition of the $(1-p_{pref})$ term when the $G\|g$ locus has downstream effects on pairing, which captures preferential cross-over formation, and the addition of a synaptic partner switch between the middle and distal loci.

The mode probabilities for the a_ib_j/a_kb_ℓ mode are Rehmsmeier ([2013](#)):

$$q^2 - p_{cp}\frac{q^2\tau}{2} + p_{pm}\left(-\frac{q}{4} - 5\frac{q^2}{8} + \frac{q'}{2} - \frac{qq'}{4}\right)\tau$$

3-locus model, with local affects on cross-over formation: equal to Rehmsmeier ([2013](#)) without preferential pairing

3-locus model, with downstream affects on cross-over formation:

$$\frac{q^2}{3} + 2\frac{(1-p_{pref})^2q^2}{3} + p_{cp}\left(-\frac{q^2}{6} - \frac{(1-p_{pref})^2q^2}{3}\right)\tau$$
$$+ p_{pm}\left(7\frac{q'^2}{6} - 5\frac{q'^3}{6} + \frac{q'^4}{2} + \dots \{5\}\dots - 2\frac{(1-p_{pref})^2q^2}{3}\right)\tau$$
$$+ p_{md}\left(2\frac{q^2}{3} - 2\frac{(1-p_{pref})^2q^2}{3}\right)\tau$$

(2)

such that there is again the addition of the $(1-p_{pref})$ term and a synaptic partner switch between the middle and distal locus with downstream affects of the $G\|g$ locus on cross-over formation. Numerically, with downstream affects of the $G\|g$ locus on cross-over formation, an increase in preferential cross-over (p_{pref}) leads to an increase in the probability of the a_ib_i/a_jb_j mode and a decrease in probability of the a_ib_j/a_kb_ℓ mode (Fig. [5](#)).

Next we use the probabilities for the a_ib_i/a_jb_j and a_ib_j/a_kb_ℓ modes for the 3-locus preferential cross-over formation model and a distal $G\|g$ locus as baselines for comparisons with corresponding probabilities when the $G\|g$ locus is proximal or in the middle:

a_ib_i/a_jb_j, proximal $G\|g$ locus with local affects on cross-over formation:

$$(1 - (p_{cp} + p_{pm} + p_{md})\tau)(1-r)^2$$
$$+ p_{cp}(1-r)\left(1 - \frac{v'}{6} + \frac{v'^2}{6} - \frac{3r}{4} + \dots \{29\}\dots - \frac{4(1-p_{pref})^4q^2rv'^2}{3}\right)\tau$$
$$+ p_{pm}(1-r)\left(1 - \frac{3r}{4} - \frac{q'}{2} + \frac{q'r}{2} + \frac{q^2}{2} - \frac{q'^2r}{2}\right)\tau + p_{md}\frac{(1-r')^3(4-3r')}{4}\tau$$

(3)

a_ib_i/a_jb_j, proximal $G\|g$ locus with downstream affects on cross-over formation:

$$(1 - (p_{cp} + p_{pm} + p_{md})\tau)(1-r)^2$$
$$+ p_{cp}(1-r)\left(1 - \frac{v'}{3} + \frac{v'^2}{3} - \frac{3r}{4} + \dots \{25\}\dots - \frac{2(1-p_{pref})^2q^2rv'^2}{3}\right)\tau$$
$$+ p_{pm}(1-r)\left(1 - \frac{3r}{4} - \frac{q'}{2} + \frac{q'r}{2} + \frac{q^2}{2} - \frac{q'^2r}{2}\right)\tau + p_{md}\frac{(1-r')^3(4-3r')}{4}\tau$$

(4)

a_ib_j/a_kb_ℓ, proximal $G\|g$ locus with local affects on cross-over formation: equal to proximal $G\|g$ locus with downstream affects on cross-over formation (see next)

a_ib_j/a_kb_ℓ, proximal $G\|g$ locus with downstream affects on cross-

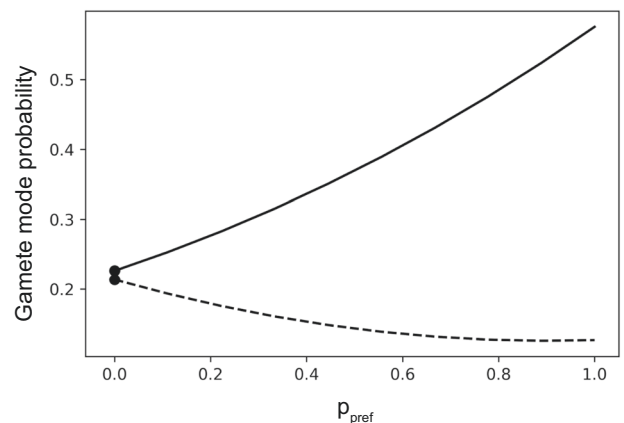**Fig. 5 The probability of gamete modes a_ib_i/a_jb_j (solid line) and a_ib_j/a_kb_ℓ (dashed line) without and with preferential cross-over formation and when the $G\|g$ locus is distal and acts downstream.** Points correspond to Rehmsmeier ([2013](#)) predictions without preferential cross-over formation. Note, for the distal location of the $G\|g$ locus and local effects, gamete mode probabilities are the same as without preferential pairing. Lines are preferential cross-over formation as a function of p_{pref}. Other parameters are $v = q = r = 1/2$, $p_{cp} = p_{pm} = p_{md} = 1/4$, and $\tau = 2/3$.

over formation:

$$(1 - (p_{cp} + p_{pm} + p_{md})\tau)r^2$$
$$+ p_{cp}\frac{r^2}{2}\tau + p_{pm}\frac{r^2}{2}\tau + p_{md}\frac{r'^2(3r'^2 - 5r' + 3)}{2}\tau \quad (5)$$

a_ib_i/a_jb_j, middle $G\|g$ locus with local affects on cross-over formation:

$$\begin{aligned}
&(1 - (p_{cp} + p_{pm} + p_{md})\tau)(1 - \tfrac{2r}{3} + \tfrac{r^2}{3} - \tfrac{2q}{3} + 2qr\ldots\{11\}\ldots + \tfrac{8(1-p_{pref})^4 q^2 r^2}{3}) \\
&+ p_{cp}(1 - \tfrac{v'}{2} + \tfrac{v'^2}{2} - \tfrac{7r}{12} + \tfrac{rv'}{4} + \ldots \{53\}\ldots + \tfrac{4(1-p_{pref})^4 q^2 r^2 v'^2}{3})\tau \\
&+ p_{pm}(1 - \tfrac{7r}{12} + \tfrac{r^2}{4} - \tfrac{31q'}{12} + \tfrac{17q'r}{6} + \ldots \{33\}\ldots + 2(1-p_{pref})^4 q'^4 r^2)\tau \\
&+ p_{md}(1 - \tfrac{29r'}{12} + \tfrac{9r'^2}{4} - \tfrac{13r'^3}{12} + \tfrac{r'^4}{4} + \ldots \{32\}\ldots 2(1-p_{pref})^4 q^2 r'^4)\tau
\end{aligned}$$
$$(6)$$

a_ib_i/a_jb_j, middle $G\|g$ locus with downstream affects on cross-over formation:

$$\begin{aligned}
&(1 - (p_{cp} + p_{pm} + p_{md})\tau)(1 - 2r + r^2 - \tfrac{4q}{3} + 4qr + \ldots \{9\}\ldots + \tfrac{8(1-p_{pref})q^2 r^2}{3}) \\
&+ p_{cp}(1 - \tfrac{v'}{3} + \tfrac{v'^2}{3} - \tfrac{7r}{4} + \tfrac{2rv'}{3} + \ldots \{48\}\ldots + \tfrac{2(1-p_{pref})^4 q^2 r^2 v'^2}{3})\tau \\
&+ p_{pm}(1 - \tfrac{7r}{4} + \tfrac{3r^2}{4} - \tfrac{19q'}{6} + \tfrac{41q'r}{6} + \ldots \{30\}\ldots + \tfrac{(1-p_{pref})^2 q'^4 r^2}{3})\tau \\
&+ p_{md}(1 - \tfrac{15r'}{4} + \tfrac{21r'^2}{4} - \tfrac{13r'^3}{4} + \tfrac{3r'^4}{4} + \ldots \{19\}\ldots + 2(1-p_{pref})q^2 r'^4)\tau
\end{aligned}$$
$$(7)$$

a_ib_j/a_kb_ℓ, middle $G\|g$ locus with local affects on cross-over formation:

$$\begin{aligned}
&(1 - (p_{cp} + p_{pm} + p_{md})\tau)(\tfrac{r^2}{3} + \tfrac{2qr}{3} - \tfrac{4qr^2}{3} + \tfrac{q^2}{3} + \ldots \{7\}\ldots + \tfrac{8(1-p_{pref})^4 q^2 r^2}{3}) \\
&+ p_{cp}(\tfrac{r^2}{6} + \tfrac{qr}{6} - \tfrac{2qr^2}{3} + \tfrac{q^2}{3} + \ldots \{7\}\ldots + \tfrac{4(1-p_{pref})^4 q^2 r^2}{3})\tau \\
&+ p_{pm}(\tfrac{r^2}{6} + \tfrac{q'r}{3} - q'r^2 + \tfrac{7q'^2}{6} + \ldots \{27\}\ldots + 4(1-p_{pref})^4 q'^4 r^2)\tau \\
&+ p_{md}(\tfrac{5r'^2}{6} - \tfrac{5r'^3}{6} + \tfrac{r'^4}{2} + \tfrac{2qr'}{3} + \ldots \{27\}\ldots + 4(1-p_{pref})^4 q^2 r'^4)\tau
\end{aligned}$$
$$(8)$$

a_ib_j/a_kb_ℓ, middle $G\|g$ locus with downstream affects on cross-over formation:

$$\begin{aligned}
&(1 - (p_{cp} + p_{pm} + p_{md})\tau)(r^2 + \tfrac{4qr}{3} - \tfrac{8qr^2}{3} + \tfrac{q^2}{3} + \ldots \{6\}\ldots + \tfrac{8(1-p_{pref})q^2 r^2}{3}) \\
&+ p_{cp}(\tfrac{r^2}{2} + \tfrac{2qr}{3} - \tfrac{4qr^2}{3} + \tfrac{q^2}{3} + \ldots \{6\}\ldots + \tfrac{4(1-p_{pref})q^2 r^2}{3})\tau \\
&+ p_{pm}(\tfrac{r^2}{2} + \tfrac{2q'r}{3} - \tfrac{7q'r^2}{3} + \tfrac{7q'^2}{3} + \ldots \{21\}\ldots + \tfrac{2(1-p_{pref})^2 q'^4 r^2}{3})\tau \\
&+ p_{md}(\tfrac{3r'^2}{2} - \tfrac{5r'^3}{2} + \tfrac{3r'^4}{2} + \tfrac{4qr'}{3} + \ldots \{16\}\ldots 4(1-p_{pref})q^2 r'^4)\tau
\end{aligned}$$
$$(9)$$

These results demonstrate that when the $G\|g$ locus is either in the middle or proximal to the focal loci, it can have different effects on gamete mode probabilities when it acts either locally or downstream toward the centromere. There are noticeable increases in complexity of the middle versus proximal and distal cases in the number of terms, which likely corresponds to an increase in events that can bring about a gamete mode. When the $G\|g$ locus is in the middle, it has an effect on gamete mode probabilities, whether its action is local or downstream.

Although gamete modes in the proximal case involve terms with p_{pref}, variation in p_{pref} has little to no effect for these mode probabilities (Fig. 6) for both local (solid lines) and downstream (dash-dotted lines) effects at the $G\|g$ locus. In the middle case and the downstream $G\|g$ model, p_{pref} only has a weak effect (Fig. 6, long-dash, short-dash lines). In the middle case and local $G\|g$ effects, p_{pref} has strong effects (Fig. 6, dashed lines). There was no effect when the $G\|g$ locus was distal for the two modes that were investigated when it acts locally. Overall, preferential cross-over formation had a greater effect in the distal position when the $G\|g$ locus acts downstream (Fig. 5), likely because it manifests its effect across more potential cross-over points. A lack of an effect when it acts locally arises because chromosome pairing is initially random and a distal location of the $G\|g$ locus does not affect the probability of double reduction,
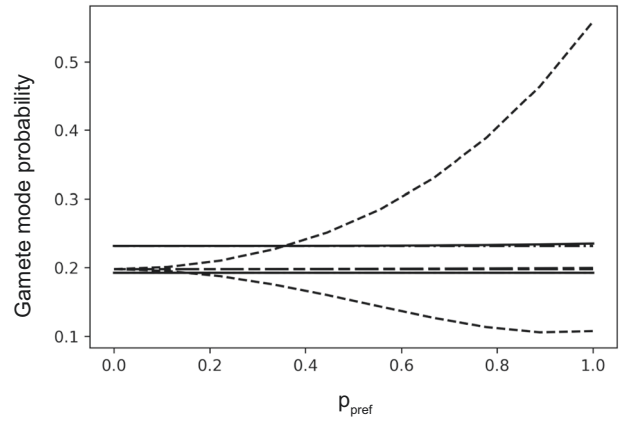

Fig. 6 The probability of gamete modes a_ib_i/a_jb_j and a_ib_j/a_kb_ℓ as a function of p_{pref}. The $G\|g$ locus is proximal (solid lines) or in the middle (dashed lines) and the $G\|g$ locus acts locally, or the $G\|g$ locus is proximal (dash-dot lines) or in the middle (long-dash, short-dash lines) and the $G\|g$ locus acts downstream. For both the proximal and middle cases, upper lines correspond to the mode a_ib_i/a_jb_j and lower lines the mode a_ib_j/a_kb_ℓ. Other parameters are the same as in Fig. 5.

nor recombination between focal loci in the gamete mode.

For readers interested in probabilities of double reduction, these are provided in Supplementary Information - 43 - Double Reduction - Python Notebook. With substitutions of $x = 2x'(1 - x')$ for $x' \in \{v', q', r'\}$ it can be shown that probabilities of double reduction at the proximal and middle loci are equal to Rehmsmeier (2013) cases for two loci (their equations 1 and 2), and without preferential cross-over formation. Double reduction at the distal locus for three loci and without preferential cross-over formation is a function of the three synaptic partner switch locations and the recombination events between the middle and distal loci, as well as recombination events in locations toward the centromere. In addition, we give probabilities of double reduction across combinations of loci, such as at both the proximal and middle locations, proximal and distal locations, as well as all three locations simultaneously. Lastly, Supplementary Information - 43 provides probabilities of double reduction with preferential cross-over formation and at the locus nearest the centromere when the $G\|g$ locus is proximal, in the middle, and distal. Cases when both the $G\|g$ locus acts locally and downstream toward the centromere are explored. It is worthwhile to note that the probability of double reduction depends on both the location of and parental genotype at the $G\|g$ locus. This may be a consideration in approaches such as Chen et al. (2021) which define gamete mode probabilities in terms of the probability of double reduction at a locus nearest the centromere and in situations when there may be preferential cross-over formation.

A telomeric location of the $G\|g$ locus. Probabilities for the gamete modes a_ib_i/a_jb_j and a_ib_j/a_kb_ℓ are
   a_ib_i/a_jb_j:

$$\begin{aligned}
\Big(&(1 - (p_{cp} + p_{pm} + p_{md})\tau)(p_1 - 2p_1 q + \ldots \{3\}\ldots + p_0(1-p_{pref})^2 q^2) \\
&+ p_{cp}(p_1 - 2p_1 q + \ldots \{27\}\ldots + 4p_0(1-p_{pref})^6 q^2 r^2 v'^2)/2\tau \\
&+ p_{pm}(1 - q')^2(p_1 - 2p_1 q' + \ldots \{9\}\ldots + 2p_0(1-p_{pref})^4 q^2 r^2)/2\tau \\
&+ p_{md}(1 - q)^2(p_1 + 3p_0 - 2p_0(1-p_{pref})r' + 2p_0(1-p_{pref})^2 r'^2)/2\tau\Big)/(2p_0 + p_1) \\
&+ \Big((1 - (p_{cp} + p_{pm} + p_{md})\tau)(8p_0(1 - (1-p_{pref})q)^2) \\
&+ p_{pc}(2p_1 - 2p_1 v' + \ldots \{27\}\ldots - 8p_0(1-p_{pref})^6 q^2 r^2 v'^2)\tau \\
&+ p_{pm}2(1 - q')^2(2p_1 - 3p_1 q' + \ldots \{9\}\ldots - 4p_0(1-p_{pref})^4 q^2 r^2)\tau \\
&+ p_{md}4(1 - q)^2(p_1 + p_0 + 2p_0(1-p_{pref})r' - 2p_0(1-p_{pref})^2 r'^2)\tau\Big)/(16p_0 + 8p_1)
\end{aligned}$$
$$(10)$$

$a_i b_j / a_k b_\ell$:

$$
\begin{aligned}
&\Big(\big((1-(p_{cp}+p_{pm}+p_{md})\tau)(4p_0(1-p_{pref})^2 q^2) \\
&+p_{cp}(2p_0(1-p_{pref})^2 q^2)\tau \\
&+p_{pm}q'^2(2p_1-2p_1 q'+\dots\{9\}\dots-8p_0(1-p_{pref})^4 q'^2 r^2)\tau \\
&+p_{md}q^2(2p_1+2p_0+4p_0(1-p_{pref})r'-4p_0(1-p_{pref})^2 r'^2)\tau\big)/(8p_0+4p_1) \\
&+\Big((1-(p_{cp}+p_{pm}+p_{md})\tau)(p_1 q^2+p_0(1-p_{pref})^2 q^2) \\
&+p_{cp}(p_1 q^2+p_0(1-p_{pref})^2 q^2)/2\tau \\
&+p_{pm}q'^2(2p_1-4p_1 q'+\dots\{11\}\dots+4p_0(1-p_{pref})^4 q'^2 r^2)/2\tau \\
&+p_{md}q^2(p_1+3p_0-2p_0(1-p_{pref})r'+2p_0(1-p_{pref})^2 r'^2)\tau\Big)/(2p_0+p_1)
\end{aligned}
$$

$$(11)$$

Figure 7 plots these mode probabilities as a function of p_0 and assuming $p_1 = 1$. As p_0 approaches p_1 these gamete mode probabilities converge on the gamete mode probabilities for the non-telomere-associated and a distal location of the $G \| g$ locus. This makes sense because if $p_0 = p_1$, then initial pairing is random, like the non-telomere-associated scenario.

Note, Python Notebooks used to generate Figs. 5–7 are provided as Supplementary Information (files 44–48), and with corresponding identifiers.

Estimating synaptic switch rate and probability of preferential cross-over formation

For both the parental genotype *GGGg* and *GGgg* and across locations of the $G \| g$ locus, accurate estimates of the probability of synaptic partner switches are possible for both low and high switch rates and for sample sizes of 5000 and 50,000 (Table 1). If the $G \| g$ locus acts downstream and is at the distal position (but is not associated with the telomere), accurate estimates of p_{pref} are possible for both the parental genotype *GGGg* and *GGgg*, low versus high synaptic partner switch rates and sample sizes of 5000 and 50,000. In contrast, for a middle or proximal location of the $G \| g$ locus estimates of p_{pref} are consistent and reasonably accurate only when synaptic partner switch rates are higher and the sample size is 50,000. If the $G \| g$ locus acts locally and is at the middle position, accurate estimates of p_{pref} occur for both low and high synaptic partner switch rates, low and high values of p_{pref}, and lower and higher sample sizes - except the combination of a higher synaptic switch rate (0.25) and a lower value of p_{pref} (0.50).

Overall, the estimation results thus far correspond to our earlier results for a subselection of gamete modes. Changes in p_{pref} had a greater effect on gamete mode probabilities for the distal location of the $G \| g$ locus and when its action was downstream compared to other locations, and this combination of distal location and downstream action is where estimates of p_{pref} are consistently accurate. Additionally, a higher rate of synaptic partner switches aids estimates of p_{pref}, particularly when the $G \| g$ locus is proximal or in the middle and acts downstream. This makes sense because a higher rate of synaptic partner switches generates more contrasts in gametic genotypes to reveal the effect of the $G \| g$ locus on cross-over formation. When the $G \| g$ locus is in the middle and acts locally, estimates of p_{pref} are generally more accurate than when the $G \| g$ locus is in the middle and acts downstream, which is consistent with Fig. 6 where changes in p_{pref} have a strong effect on gamete mode probabilities.

Files that form the basis for Table 1 are provided as Supplementary Information (files 49, 50), with corresponding identifiers.

Relating the theoretical model to empirical observations of meiosis

No preferential pairing. Over the last decade *A. arenosa* has become a model species to understand non-preferential meiosis in autotetraploids (Bomblies et al., 2016, Morgan et al., 2021, Yant
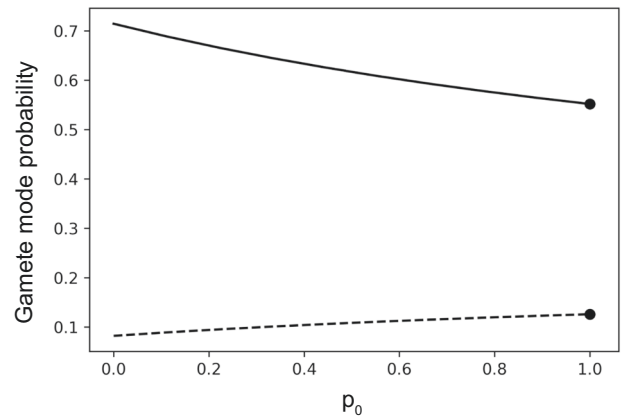

Fig. 7 The probability of gamete modes $a_i b_i / a_j b_j$ **(solid line) and** $a_i b_j / a_k b_\ell$ **(dashed line) as a function of** p_0 **when the** $G \| g$ **locus is closely associated with the telomere of a chromosome.** The point is the case when the $G \| g$ locus is not associated with the telomere. Other parameters are $v = q = r = 1/2$, $p_{cp} = p_{pm} = p_{md} = 1/4$, $p_{pref} = 95/100$, and $\tau = 2/3$ (when the $G \| g$ locus is not associated with the telomere).

et al., 2013). A key finding is that established versus neo-autotetraploids have a reduced cross-over rate and an increased rate of bivalency versus tetravalency. The mechanism appears to be related to interference, whereby adjacent cross-overs occur further apart in established versus neo-autotetraploids (Morgan et al., 2021). Our no preferential pairing model can capture the observations of *A. arenosa* and the shift to bivalency by reducing τ (the probability of tetravalency) and/or by reducing probabilities of partner switches given a tetravalent has formed. Morgan et al. (2021) framed cross-overs in terms of the probability that two cross-overs share the same partner. This frame of reference is captured in our model by recognizing that a synaptic partner switch results in two cross-overs sharing different partners. It is also possible for a switch to occur, but no cross-over (e.g., the cis configuration, Fig. 5M, Morgan et al. 2021). This can also occur in our model; a proportion of switches will not experience cross-overs in the direction of maturation of pairing and chiasmata.

Preferential pairing. A major process that may be shaping the evolution of autotetraploids is a reduction in cross-over rates and the rate of tetravalency (Yant et al., 2013). Nevertheless, it is important to note that in *A. arenosa* there is as yet no evidence of preferential pairing. In contrast, in rye, there is evidence for preferential pairing due to changes in sequence composition along homologous chromosomes (Jenkins and Chatterjee, 1994). In salmon, there is also evidence for preferential pairing. Most genomic segments currently experience preferential disomic inheritance, while other segments undergo tetrasomic inheritance, with some homologous chromosome segments not pairing with equal probability (Allendorf et al., 2015). Results from salmon indicate there could be variation in cross-over probabilities along homologous chromosomes.

This paper presents three models of preferential pairing. All three models assume there is a locus that affects whether two homologous chromosomes that come into proximity will form bridges and pair. In the "local" and "downstream" models it is assumed that initial pairing is random and preferential cross-over formation may occur along a chromosome. In the "scramble" model, it is assumed that initial pairing is preferential and there is subsequently the possibility for synaptic partner switches that break-up this initial preferential pairing. The "scramble" model can mimic allopolyploid preferential pairing at the whole-chromosome level by setting the synaptic partner switch rates equal to zero.

Table 1. Mean estimates of the probability of a synaptic partner switch (upper) and p_{pref} (lower), as well as standard deviations (in parentheses) across 100 replicates.

Sample size	Location of G‖g locus							
	Distal	Middle	Middle (local)	Proximal	Distal	Middle	Middle (local)	Proximal
	GGGg parent: $p_{cp} = p_{pm} = p_{md} = 0.01$, $p_{pref} = 0.95$				*GGGg* parent: $p_{cp} = p_{pm} = p_{md} = 0.01$, $p_{pref} = 0.50$			
5000	0.01 (0.004)	0.01 (0.002)	0.01 (0.003)	0.01 (0.003)	0.01 (0.003)	0.01 (0.003)	0.01 (0.002)	0.01 (0.002)
	0.95 (0.05)	0.75 (0.36)	0.95 (0.007)	0.68 (0.41)	0.50 (0.02)	0.42 (0.044)	0.50 (0.024)	0.53 (0.46)
50000	0.01 (0.001)	0.01 (0.0008)	0.01 (0.001)	0.01 (0.004)	0.01 (0.001)	0.01 (0.001)	0.01 (0.001)	0.01 (0.01)
	0.95 (0.003)	0.90 (0.122)	0.95 (0.002)	0.79 (0.32)	0.50 (0.007)	0.54 (0.34)	0.50 (0.007)	0.56 (0.40)
	GGGg parent: $p_{cp} = p_{pm} = p_{md} = 0.25$, $p_{pref} = 0.95$				*GGGg* parent: $p_{cp} = p_{pm} = p_{md} = 0.25$, $p_{pref} = 0.50$			
5000	0.25 (0.015)	0.25 (0.01)	0.25 (0.02)	0.25 (0.01)	0.25 (0.02)	0.25 (0.01)	0.25 (0.01)	0.25 (0.01)
	0.95 (0.14)	0.92 (0.08)	0.95 (0.02)	0.91 (0.11)	0.50 (0.04)	0.57 (0.27)	0.50 (0.03)	0.55 (0.29)
50000	0.25 (0.004)	0.25 (0.004)	0.25 (0.005)	0.25 (0.003)	0.25 (0.005)	0.25 (0.004)	0.25 (0.004)	0.25 (0.004)
	0.95 (0.004)	0.95 (0.02)	0.95 (0.006)	0.95 (0.04)	0.50 (0.01)	0.51 (0.09)	0.50 (0.01)	0.53 (0.13)
	GGgg parent: $p_{cp} = p_{pm} = p_{md} = 0.01$, $p_{pref} = 0.95$				*GGgg* parent: $p_{cp} = p_{pm} = p_{md} = 0.01$, $p_{pref} = 0.50$			
5000	0.01 (0.004)	0.01 (0.01)	0.01 (0.005)	0.01 (0.01)	0.01 (0.01)	0.01 (0.01)	0.01 (0.01)	0.01 (0.01)
	0.95 (0.009)	0.82 (0.40)	0.95 (0.005)	0.23 (0.35)	0.50 (0.02)	0.31 (0.43)	0.50 (0.02)	0.17 (0.34)
50000	0.01 (0.002)	0.01 (0.005)	0.01 (0.001)	0.01 (0.005)	0.01 (0.006)	0.01 (0.004)	0.01 (0.005)	0.01 (0.004)
	0.95 (0.003)	0.75 (0.40)	0.95 (0.002)	0.91 (0.25)	0.50 (0.008)	0.22 (0.40)	0.50 (0.008)	0.01 (0.10)
	GGgg parent: $p_{cp} = p_{pm} = p_{md} = 0.25$, $p_{pref} = 0.95$				*GGgg* parent: $p_{cp} = p_{pm} = p_{md} = 0.25$, $p_{pref} = 0.50$			
5000	0.25 (0.02)	0.25 (0.02)	0.25 (0.02)	0.25 (0.01)	0.25 (0.03)	0.25 (0.01)	0.24 (0.02)	0.25 (0.05)
	0.95 (0.02)	0.91 (0.15)	0.95 (0.01)	0.75 (0.34)	0.50 (0.04)	0.64 (0.35)	0.58 (0.19)	0.56 (0.42)
50000	0.25 (0.006)	0.25 (0.004)	0.25 (0.005)	0.25 (0.006)	0.25 (0.01)	0.25(0.004)	0.25 (0.008)	0.25 (0.05)
	0.95 (0.007)	0.94 (0.04)	0.95 (0.003)	0.91 (0.11)	0.50 (0.01)	0.54 (0.19)	0.51 (0.07)	0.53 (0.24)

Estimates of parameters are based on parental genotypes *GGGg* or *GGgg* and for proximal, middle and distal locations of the G‖g locus. Unless noted, the action of the G‖g locus is downstream toward the centromere, except for an examination of performance when the G‖g locus is in the middle and acts locally. For a given parental genotype either 5000 or 50,000 gametes are produced. True values for the probability of a synaptic partner switch were 0.01 or 0.25 and for p_{pref} was 0.95 or 0.50.

Model considerations and areas for advancement

Our model assumed potentially one synaptic partner switch per meiotic event, where this switch could occur between each pair of loci or the proximal locus and centromere. Sets of sister chromatids can have multiple synaptic partners, but it is also the case that typically only one or a few physical cross-over events occur per meiotic event (Bomblies et al., 2016), such that consequences of more than one switch may be limited by a lack of crossovers. Rehmsmeier (2013) included a two-partner switch version of their two-locus model of meiosis. In principle, a three-locus model like ours could be extended to allow for potentially three partner switches during a single meiotic event. This would increase the number of potential crossover events to 24 per pairing configuration. In addition, for each partner switch there are two possible rearrangements after the first switch, such that if three switches are possible, there would be $2 \times 2 \times 2^{24} = 67, 108, 864$ combinations of events if three partner switches occurred. The scale of this analysis is computationally possible and left for consideration. Presently, our model allows for three synaptic partner switch locations, but the realization of switches across these locations in a single lineage would happen during separate meiotic events.

A second consideration is our assumption that the probability of a synaptic partner switch is independent of the state of the G‖g locus. Presently, the G‖g locus affects either cross-over formation given a pairing, telomere-associated pairing or both. It seems reasonable that if a locus like the G‖g locus occurs and it leads to weaker bridges (or no bridges) between paired sets of sister chromatids, then it may also be associated with a greater tendency for a synaptic partner switch to occur because chromosomes would tend to dissociate and potentially bump into another homologous chromosome and establish bridges. We

leave this for future research and the current model as a base-point for comparison.

Our analysis of estimating rates of synaptic partner switches and preferential pairing should be viewed as illustrative, but not definitive. In particular, we assume baseline cross-over rates equal to 1/2. This is motivated by more of a "null" perspective in that we assume recombination is otherwise very high. In reality, particularly species like *A. arenosa* baseline rates of recombination may be less than 0.5. Empirically, one may not be free to assume a particular baseline rate of cross-over, requiring an increase in complexity of statistical inference.

Lastly, a question is the extent to which our model captures cross-over interference. Cross-over interference is not unique to polyploids, but inherent in diploid meiosis. There seems to be perhaps two points of consideration in the context of cross-over interference. One point is the extent to which it reduces realized rates of recombination. The second point is the extent to which it causes the coefficient of coincidence to deviate below a value of one. The coefficient of coincidence is the ratio of observed to expected number of cross-overs (assuming independence). A value below one indicates interference and non-independent cross-over, given a set of realized or observed cross-over rates (Christiansen, 2008). The results of Yant et al. (2013) and Morgan et al. (2021) indicate cross-over rates are reduced in established autotetraploid *A. arenosa*. This can be captured in our model by reducing cross-over/recombination rates, as well as by having lower rates for loci that are close together compared to further apart. Nevertheless, implicit in our model is the assumption that the probability of one recombination event is independent of the probability of another recombination event when recombination (or cross-over) rates are modeled as realized rates. The assumption of independence may contradict a strict application of

interference. For example, there are three regions a cross-over can occur in our model, such that there are potentially three cross-over events. If it is assumed probabilities of cross-over are 1/2, then 12.5% of meiotic products involve three simultaneous cross-overs. Three physical cross-overs are vanishingly rare in established autotetraploids of *A. arenosa* and occur 3% of the time in neo-autotetraploids (Morgan et al., 2021). Nevertheless, it is important to note that a map distance corresponding to a 1/2 cross-over probability from the centromere to a proximal locus and twice more between the other loci would take up several linkage groups in diploid *A. arenosa* (Dukic and Bomblies, 2022), such that lower cross-over rates are more realistic and the difference between predicted and observed simultaneous cross-over would be less. For example, the longest chromosome arm in diploid *A. arenosa* is about 70 cM (Dukic and Bomblies, 2022). Dividing this arm in three, results in a cross-over probability of about 0.23 (Haldane, 1919) per segment, such that the probabilities of double and triple cross-overs along the same chromosome arm are 0.05 and 0.01, respectively. A 1% triple cross-over rate is just below 3% and in line with neo-autotetraploid *A. arenosa*. Established autotetraploid *A. arenosa* have lower observed recombination rates and the extent to which coefficients of coincidence deviate from one is not known from our understanding. Our model can be extended to account for cross-over interference and coefficients of coincidence that deviate from one by, for example, adding a process whereby cross-over initiation between two loci has a probability of interfering with cross-over initiation elsewhere. This is left for future work because of the complexity of extending meiosis to include three loci and preferential cross-over formation, as presented here.

CONCLUSIONS

This paper presented three-locus models of gamete mode probabilities with and without preferential cross-over formation. Our initial analysis of the effect of preferential cross-over formation shows that it has an effect on gamete mode probabilities and this effect appears to be greater at the distal versus middle and proximal positions of the $G\|g$ locus when the action of the $G\|g$ locus is downstream toward the centromere, whereas when its action is local, a middle position has the strongest effect. The greater effect at the distal position (when action is downstream) or the middle position (when action is local) leads to apparently more accurate estimates of the degree of preferential cross-over formation compared to the other positions. In both animals and plants (Allendorf et al., 2015, Jenkins and Chatterjee, 1994) preferential cross-over formation may be variable across a chromosome; the consequence of this for microevolutionary processes and the landscape of linkage disequilibrium are open questions, as well as the extent to which selective processes may overcome a natural tendency for the diploidization of autotetraploid genomes.

DATA AVAILABILITY
Supplementary files are divided between Data Dryad (https://doi.org/10.5061/dryad.0k6djhb48) and Zenodo (https://doi.org/10.5281/zenodo.7569103), depending on whether the files include analyses and results (Data Dryad) or pure code (Zenodo). In addition, analyses and result in files to generate Table 1 are provided in the Data Dryad archive.

REFERENCES

Allendorf FW, Basham S, Cresko WA, Limborg MT, Seeb LW, Seeb JE (2015) Effects of crossovers between homeologs on inheritance and population genomics in polyploid-derived salmonid fishes. J Hered 106:217–227

Allendorf FW, Danzmann RG (1997) Secondary tetrasomic segregation of MDH-B and preferential pairing of homeologues in rainbow trout. Genetics 145:1083–1092

Armstrong AJ, Franklin FCH, Jones GH (2001) Nucleolus-associated telomere clustering and pairing precede meiotic chromosome synpasis in Arabidopsis thaliana. J Cell Sci 114:4207–4217

Armstrong SJ, Jones GH (2003) Meiotic cytology and chromosome behaviour in wild-type Arabidopsis thaliana. J Exp Bot 54:1–10

Barker MS, Arrigo N, Baniaga AE, Li Z, Levin DA (2016) On the relative abundance of autopolyploids and allopolyploids. New Phytol 210:391–398

Bass HW, Riera-Lizarazu O, Ananiev EV, Bordoll SJ et al. (2000) Evidence for the coincident initiation of homolog pairing and synpasis during the telomere-clustering (bouquet) stage of meiotic prophse. J Cell Sci 113:1033–1042

Bomblies K, Jones G, Franklin C, Zickler D, Kleckner N (2016) The challenge of evolving stable polyploidy: could an increase in "crossover interference distance" play a central role? Chromosoma 125:287–300

Chen J, Leach L, Yang J, Zhang F, Tao Q, Dang Z, Chen Y, Luo Z (2021) A tetrasomic inheritance model and likelihood-based method for mapping quantitative trait loci in autotetraploid species. New Phytol 230:387–398

Christiansen FB (2008) Theories of population variation in genes and genomes. Princeton University, Princeton.

Dukic M, Bomblies K (2022) Male and female recombination landscapes of diploid Arabidopsis arenosa. Genetics 220:1–16

Fisher RA (1947) The theory of linkage in polysomic inheritance. Phil Trans R Soc Lond B 233:55–87

Griswold CK (2021) The effects of migration load, selfing, inbreeding depression, and the genetics of adaptation on autotetraploid versus diploid establishment in peripheral habitats. Evolution 75:39–55

Griswold CK, Williamson MW (2017) A two-locus model of selection in auto-tetraploids: Chromosomal gametic disequilibrium and selection for an adaptive epistatic gene combination. Heredity 119:314–327

Haldane JBS (1919) The combination of linkage values and the calculation of distances between the loci of linked factors. J Genet 8:299–309

Jannoo N, Grivet L, David J, D'Hont A, Glaszmann J-C (2004) Differential chromosome pairing affinities at meiosis in polyploid sugarcane revealed by molecular markers. Heredity 93:460–467

Jenkins G, Chatterjee R (1994) Chromosome structure and pairing preferences in au totetraploid rye (Secale cereale). Genome 37:784–793

Koshi MH, Layman NC, Busch JW, Galloway LF (2019) Selfing ability and drift load evolve with range expansion. Evol Lett 3:500–512

Le Comber SC, Ainouche ML, Kovarik A, Leitch AR (2010) Making a functional diploid: from polysomic to disomic inheritance. New Phytol 186:113–122

Lloyd A, Bomblies K (2016) Meiosis in autopolyploid and allopolyploid Arabidopsis. Curr Opin Plant Biol 30:116–122

Lopez E, Pradillo M, Romero C, Santos JL, Cuñado N (2008) Pairing and synapsis in wild type Arabidopsis thaliana. Chromosome Res 16:701–708

Lu Y, Yang X, Tong C, Li X, Feng S, Wang Z, Pang X, Wang Y, Wang N, Tobias CM, Wu R (2012) A multivalent three-point linkage analysis model of autotetraploids. Brief Bioinform 14:460–468

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

Morgan C, White MA, Franklin FC, Zickler D, Kleckner N, Bomblies K (2021) Evolution of crossover interference enables stable autopolyploidy by ensuring pairwise partner connections in Arabidopsis arenosa. Curr Biol 31:4713–4726

Nei M (1967) Modification of linkage intensity by natural selection. Genetics 57:625–641

Otto SP, Feldman MW (1997) Deleterious mutations, variable epistatic interactions, and the evolution of recombination. Theor Popul Biol 51:134–147

Parra-Nunez P, Pradillo M, Santos JL (2019) Competition for Chiasma formation between identical and homologous (but not identical) chromosomes in synthetic autotetraploids of Arabidopsis thaliana. Fronti Plant Sci 9:1924

Peters J-M, Nishiyama T (2012) Sister Chromatid Cohesion. Cold Spring Harb Perspect Biol 4:a011130

Rehmsmeier M (2013) A computational approach to developing mathematical models of polyploid meiosis. Genetics 193:1083–1094

Stift M, Berenos C, Kuperus P, van Tienderen PH (2008) Segregation models for disomic, tetrasomic and intermediate inheritance in tetraploids: a general procedure applied to Rorippa (Yellow Cress) Microsatellite Data. Genetics 179:2113–2123

Sved JA (1965) Telomere attachment of chromsomes: Some genetical and cytological consequences. Genetics 53:747–756

Sybenga J (1975) The quantitative analysis of chromosome pairing and chiasma formation based on the relative frequencies of M I configurations VII. Auto-tetraploids. Chromosoma 50:211–222

Sybenga J (1999) What makes homologous chromosomes find each other in meiosis? A review and an hypothesis. Chromosoma 108:209–219

Voorrips RE, Maliepaard CA (2012) The simulation of meiosis in diploid and tetraploid organisms using various genetic models. BMC Bioinformatics 13:248

Xu F, Lyu Y, Tong C, Wu W, Zhu X, Yin D, Yan Q, Zhang J, Pang X, Tobias CM, Wu R (2013) A statistical model for QTL mapping in polysomic autotetraploids underlying double reduction. Brief Bioinform 15:1044–1056

Yang X, Lv Y, Pang X, Tong C, Wang Z, Li X, Feng S, Tobias CM, Wu R (2012) A unifying framework for bivalent multilocus linkage analysis of allotetraploids. Brief Bioinform 14:96–108

Yant L, Hollister JD, Wright KM, Arnold BJ, Higgins JD, Franklin FC et al. (2013) Meiotic adaptation to genome duplication in Arabidopsis arenosa. Curr Biol 23:2151–2156

ACKNOWLEDGEMENTS

AUTHOR CONTRIBUTIONS
CKG and SA derived and coded the models, as well as generated gamete mode probabilities. CKG performed the likelihood analysis. CKG wrote the paper, with input from SA.

COMPETING INTERESTS
The authors declare no competing interests.

ADDITIONAL INFORMATION
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41437-023-00595-9.

Correspondence and requests for materials should be addressed to Cortland K. Griswold.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.