Copyright © 2001 by the Genetics Society of America

A General Polyploid Model for Analyzing Gene Segregation in Outcrossing Tetraploid Species

Rongling Wu,* Maria Gallo-Meagher,† Ramon C. Littell* and Zhao-Bang Zeng‡

**Department of Statistics, †Agronomy Department, Institute of Food and Agricultural Sciences, University of Florida, Gainesville, Florida 32611 and ‡Program in Statistical Genetics, Department of Statistics, North Carolina State University, Raleigh, North Carolina 27695*

Manuscript received February 8, 2001
Accepted for publication July 16, 2001

ABSTRACT

Polyploidy has played an important role in higher plant evolution and applied plant breeding. Polyploids are commonly categorized as allopolyploids resulting from the increase of chromosome number through hybridization and subsequent chromosome doubling or autopolyploids due to chromosome doubling of the same genome. Allopolyploids undergo bivalent pairing at meiosis because only homologous chromosomes pair. For autopolyploids, however, all homologous chromosomes can pair at the same time so that multivalents and, therefore, double reductions are formed. In this article, we use a maximum-likelihood method to develop a general polyploid model for estimating gene segregation patterns from molecular markers in a full-sib family derived from an arbitrary polyploid combining meiotic behaviors of both bivalent and multivalent pairings. Two meiotic parameters, one describing the preference of homologous chromosome pairing (expressed as the preferential pairing factor) typical of allopolyploids and the other specifying the degree of double reduction of autopolyploids, are estimated. The type of molecular markers used can be fully informative *vs.* partially informative or dominant *vs.* codominant. Simulation studies show that our polyploid model is well suited to estimate the preferential pairing factor and the frequency of double reduction at meiosis, which should help to characterize gene segregation in the progeny of autopolyploids. The implications of this model for linkage mapping, population genetic studies, and polyploid classification are discussed.

POLYPLOIDY is recognized as an important evolutionary force in flowering plants (Stebbins 1971; Grant 1981; Bever and Felber 1992; Soltis and Soltis 1993, 1999; Ramsey and Schemske 1998). Recent estimates from genomic analyses suggest that as much as 70% of all angiosperms have experienced one or more episodes of polyploidization (Masterson 1994). The frequency of polyploidy in the domesticated plant taxa is also high (75%); alfalfa, banana, canola, coffee, cotton, potato, soybean, strawberry, sugarcane, sweet potato, and wheat represent excellent examples of polyploids of economic importance (Hilu 1993). An upsurge of comparative mapping studies using molecular markers reveals that several crop species traditionally considered as diploids, such as maize and modern species of Brassica, are actually polyploids, whereas for some polyploids like cotton, the level of ploidy is higher than originally recognized (Leitch and Bennett 1997).

Polyploids have been classified as either allopolyploids derived from the chromosome combination of distinct genomes and subsequent chromosome doubling or autopolyploids originated from the chromosome doubling of genetically similar genomes by fusion of unreduced gametes (Stebbins 1950). Allopolyploids are considered to be much more prevalent in nature than are autopolyploids, but, as detected from a growing number of genetic analyses, autopolyploids in nature likely are much more common than typically appreciated (Soltis and Soltis 2000). In allopolyploids, identical or at least fully homologous genomes occur in pairs, but different pairs of a genome have a strong pairing barrier (Sybenga 1996). Because only homologous chromosomes pair, allopolyploids strictly exhibit bivalent formation (two chromosomes pair) at meiosis and undergo disomic inheritance for each locus. For autopolyploids, the chromosomes are all homologous and have equal opportunities to pair at meiosis. Since pairing can start at different chromosomal sites, homologous chromosomes may switch partners, leading to multivalent formation (more than two chromosomes pair) and a type of inheritance called polysomic (Jackson and Jackson 1996; Sybenga 1996; Hauber *et al.* 1999).

Multivalent formation typical of autopolyploids can result in double reduction. The frequency of double reduction, defined as the probability of two sister chromatids occurring in the same gamete, assumes maximum values of 0 (random chromosome segregation), 1/7 (with pure random chromatid segregation), and 1/6 (with complete equational segregation; Muller 1914; Mather 1935). Experiments aimed at estimating the

Corresponding author: Rongling Wu, Department of Statistics, 533 McCarty Hall C, University of Florida, Gainesville, FL 32611. E-mail: rwu@stat.ufl.edu

frequency of double reduction in autotetraploids have yielded values ranging from 0 to almost 0.30 (Fisher 1947, 1949; Welch 1962; Tai 1982a,b; Haynes and Douches 1993). However, double reduction is a position-dependent phenomenon. It may vary depending on which chromosome a locus resides, because chromosomes can vary in their propensity to form multivalents. Also, where a locus resides on a chromosome affects the value of the frequency of double reduction, which will be greater toward the distal-proterminal regions and almost null at loci near the centromeres (Welch 1962). Due to these properties, double reduction can affect the frequency and distribution of homozygosity along the chromosomes and play a role in shaping the evolution of autopolyploid populations (Fisher 1949; Stebbins 1971; Butruille and Boiteux 2000). From a genetic viewpoint, the occurrence and frequency of double reduction is expected to affect the pattern of gene segregation in autopolyploids.

While allopolyploids and autopolyploids are two extremes of polyploids, a number of polyploid taxa actually represent intermediate stages displaying a combination of both allopolyploid and autopolyploid pairing behavior (Jackson and Jackson 1996; Allendorf and Danzmann 1997; Fjellstrom *et al.* 2001). These intermediate polyploids are viewed as a general polyploid model, similar to segmental allopolyploids defined by Stebbins (1950). For a general polyploid model, there is no complete preference of homologous over homeologous pairing, unlike extreme allopolyploids in which the frequency of meiotic pairing is near one between homologous chromosomes as opposed to zero between homeologous chromosomes, resulting from different degrees of evolutionary and taxonomic relatedness of the genomes involved (Sybenga 1988, 1992, 1994, 1999; Jackson and Jackson 1996). The difference between the two types of homologous and homeologous pairings is expressed as the preferential pairing factor, denoted by p (Sybenga 1988). In many proven autotetraploids, such as Tradescantia, Dactylis, Hyoscyamus, and Solanum species, the estimates of the preferential pairing factor significantly greater than zero were obtained using different cytological models (Lenz *et al.* 1983; Matsubayashi 1991; Sybenga 1994), pointing to considerable preferential pairing in some of the sets of four chromosomes.

The occurrence of preferential pairings in a general polyploid model makes the frequency of its multivalent formation lower than expected for extreme autopolyploids possessing fully homologous chromosomes (Sybenga 1996). The reduced frequencies of multivalent formation were observed in a variety of polyploids, as summarized in Soltis and Riesberg (1986) and Sybenga (1996). For a general polyploid model, however, it can be assumed that both preferential pairings and multivalent formation (and therefore double reduction) occur simultaneously at meiotic configuration. Using marker technologies, convincing evidence was found for both tetrasomic inheritance and preferential pairing between parental chromosomes in *Lotus corniculatus*, a perennial forage legume categorized as a segmental allopolyploid (Fjellstrom *et al.* 2001). On the basis of their own results in *Oncorhynchus mykiss* (rainbow trout) and those in fern and treefrog by Hickok (1978), Danzmann and Bogart (1982, 1983), Marsden *et al.* (1987), and Allendorf and Danzmann (1997) suggested that the mosaic of disomy and tetrasomy at various loci might be a general mechanism underlying the inheritance of many tetraploids.

For extreme allopolyploids, segregation ratios at one locus can be described by the preferential pairing factor, whereas for extreme autopolyploids it can be described by the frequency of double reduction. However, for a general polyploid, either the preferential pairing factor or the frequency of double reduction alone is no longer sufficient to specify the frequencies of the different modes of gamete formation. In this article, we develop a generalized statistical method for estimating the preferential pairing factor and the frequency of double reduction, using molecular markers for arbitrary tetraploids. Sybenga (1975, 1988, 1992, 1994, 1995), Haynes *et al.* (1991), and Jackson and Casey (1982) have established a series of mathematical models for estimating preferential pairing and chiasma parameters on the basis of cytogenetic data at diakinesis or metaphase I of meiosis in triploids to autoctoploids. However, theoretical models for these estimates using polymorphic molecular markers are not available. Recent developments in genomic and computational technologies provide a powerful means for examining the behavior of chromosome pairings at the DNA level (Soltis and Soltis 1993; Chen *et al.* 1995; Fjellstrom *et al.* 2001).

Our analysis is based on a full-sib family derived from two outbred tetraploid parents. Thus, many different marker types, fully *vs.* partially informative or codominant *vs.* dominant, can be simultaneously segregating in this family. Unlike a diploid family in which marker genotypes of both parents can be predicted on the basis of a typical segregation pattern, tetraploids may not have a simple one-to-one relationship between parental genotypes and progeny segregation patterns because of a possible multiple-dosage of an allele and double reduction in polysomic inheritance. Luo *et al.* (2000) developed a theoretical model for predicting the marker genotypes of two autotetraploid parents, using their marker phenotypes and the joint segregation information on their progeny's marker phenotypes. Their model provides the foundation on which our statistical analysis is performed to estimate the preferential pairing factor and the frequency of double reduction in polysomic inheritance. Our statistical methods based on a general polyploid model will have great implications for polyploid classification, linkage mapping, and population genetic studies.

A GENERAL TETRAPLOID MODEL

Meiotic pairing configurations: A general polyploid model is viewed as combining the meiotic behaviors of allopolyploids and autopolyploids. As a result of preferential pairing between fully homologous chromosomes over less fully homologous chromosomes, bivalents are formed in the general polyploids. However, preferential pairing is incomplete compared to allopolyploids, and some pairing between the homeologous chromosomes of the parents is possible, where homeologous pairing must compete with homologous pairing. Pairing chromosomes may switch partners but much less frequently than in autopolyploids. In the case of a pairing partner switch, fully homologous partners pair in one segment of the chromosomes and homeologues pair in other segments (Hauber *et al.* 1999). It is not excluded that homeologous chromosomes pair over their entire length and then two homeologous bivalents are formed. If homeologous pairing results in crossing over, quadrivalents are seen at diplotene, diakinesis, and metaphase I, and the resulting recombined chromosomes consist of segments derived from one parent and remaining segments derived from the other parent. This has consequences for their subsequent pairing behavior.

The meiotic pairing configurations for a general tetraploid model can be modeled mathematically as follows. For each set of four homologous chromosomes, two pairs of chromosomes are homologous and the chromosomes between pairs are homeologous. The pairing affinity between homeologous pairs may be lower than that between the homologues. The pairs cannot be distinguished morphologically, but, for the purpose of the model, the chromosomes are distinguished as 1 and 2 for one pair and 3 and 4 for the other. Each chromosome has two arms (X and Y), and thus the four chromosomes,

$$\left.\begin{matrix}X_1\\Y_1\end{matrix}\right|,\quad \left.\begin{matrix}X_2\\Y_2\end{matrix}\right|,\quad \left.\begin{matrix}X_3\\Y_3\end{matrix}\right|,\quad \left.\begin{matrix}X_4\\Y_4\end{matrix}\right|$$

are distinguished, in which

$$\left.\begin{matrix}X_1\\Y_1\end{matrix}\right| \text{ and } \left.\begin{matrix}X_2\\Y_2\end{matrix}\right|$$

are homologous, as are

$$\left.\begin{matrix}X_3\\Y_3\end{matrix}\right| \text{ and } \left.\begin{matrix}X_4\\Y_4\end{matrix}\right|.$$

The chromosome combinations

$$\left.\begin{matrix}X_1\\Y_1\end{matrix}\right| \text{ with } \left.\begin{matrix}X_3\\Y_3\end{matrix}\right| \text{ or } \left.\begin{matrix}X_4\\Y_4\end{matrix}\right|$$

as well as

$$\left.\begin{matrix}X_2\\Y_2\end{matrix}\right| \text{ with } \left.\begin{matrix}X_3\\Y_3\end{matrix}\right| \text{ or } \left.\begin{matrix}X_4\\Y_4\end{matrix}\right|$$

are homeologous bivalents (Figure 1). Since the two arms are assumed to select a partner independently, each arm may pair with a different chromosome, resulting in a quadrivalent (Figure 1). The homologous combinations are assumed to pair more frequently than the homeologous combinations (Sybenga 1975, 1988, 1992, 1995). The difference is expressed as the preferential pairing factor p (Sybenga 1988, 1994).

With four chromosomes there are three possible combinations for each arm: if 1 pairs with 2, then 3 (if it pairs) must pair with 4; if 1 pairs with 3, then 2 must pair with 4; and if 1 pairs with 4, then 2 must pair with 3. Of the three possibilities, one is homologous and two are homeologous. The homologous combinations have a probability of $\frac{1}{3} + p$ for each arm and the two homeologous combinations each have a probability of $\frac{1}{3} - \frac{1}{2}p$. For extreme allotetraploids in which homeologous chromosomes cannot pair, $p = \frac{2}{3}$. But for extreme autotetraploids having four homologous chromosomes to pair equally, $p = 0$. Therefore, a general polyploid model has the preferential pairing factor bounded on

$$0 < p < \tfrac{2}{3}. \tag{1}$$

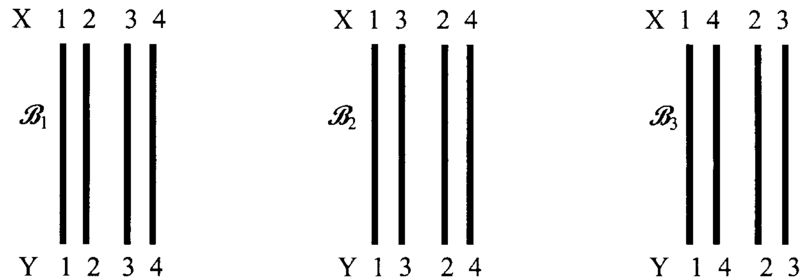
The pairings of two arms produce a total of nine combinations, six of which form a quadrivalent and three of which form a pair of bivalents (Figure 1). One of the three pairs is between homologues ($\mathcal{B}_1$) and the other two are between homeologues ($\mathcal{B}_2$ and $\mathcal{B}_3$). These are complete homeologous pairings. In four of the six quadrivalents, pairing is homeologous in one-half of the arms and homologous in the other half ($\mathcal{Q}_1$–$\mathcal{Q}_4$). In the remaining two quadrivalents, pairing is homeologous in all arms ($\mathcal{Q}_5$ and $\mathcal{Q}_6$). The frequencies of different pairings of four homologous chromosomes are calculated as

$$\begin{aligned}
f(\mathcal{B}_1) &= (\tfrac{1}{3} + p)(\tfrac{1}{3} + p) = \tfrac{1}{9} + \tfrac{2}{3}p + p^2,\\
f(\mathcal{B}_2) &= (\tfrac{1}{3} - \tfrac{1}{2}p)(\tfrac{1}{3} - \tfrac{1}{2}p) = \tfrac{1}{9} - \tfrac{1}{3}p + \tfrac{1}{4}p^2,\\
f(\mathcal{B}_3) &= (\tfrac{1}{3} - \tfrac{1}{2}p)(\tfrac{1}{3} - \tfrac{1}{2}p) = \tfrac{1}{9} - \tfrac{1}{3}p + \tfrac{1}{4}p^2,\\
f(\mathcal{Q}_1) &= (\tfrac{1}{3} + p)(\tfrac{1}{3} - \tfrac{1}{2}p) = \tfrac{1}{9} + \tfrac{1}{6}p - \tfrac{1}{2}p^2,\\
f(\mathcal{Q}_2) &= (\tfrac{1}{3} + p)(\tfrac{1}{3} - \tfrac{1}{2}p) = \tfrac{1}{9} + \tfrac{1}{6}p - \tfrac{1}{2}p^2,\\
f(\mathcal{Q}_3) &= (\tfrac{1}{3} + p)(\tfrac{1}{3} - \tfrac{1}{2}p) = \tfrac{1}{9} + \tfrac{1}{6}p - \tfrac{1}{2}p^2,\\
f(\mathcal{Q}_4) &= (\tfrac{1}{3} + p)(\tfrac{1}{3} - \tfrac{1}{2}p) = \tfrac{1}{9} + \tfrac{1}{6}p - \tfrac{1}{2}p^2,\\
f(\mathcal{Q}_5) &= (\tfrac{1}{3} - \tfrac{1}{2}p)(\tfrac{1}{3} - \tfrac{1}{2}p) = \tfrac{1}{9} - \tfrac{1}{3}p + \tfrac{1}{4}p^2,\\
f(\mathcal{Q}_6) &= (\tfrac{1}{3} - \tfrac{1}{2}p)(\tfrac{1}{3} - \tfrac{1}{2}p) = \tfrac{1}{9} - \tfrac{1}{3}p + \tfrac{1}{4}p^2.
\end{aligned}$$

The frequency of all bivalent pairings equals $f(\mathcal{B}) = \frac{1}{3} + \frac{3}{2}p^2$ and the frequency of all quadrivalent pairings equals $f(\mathcal{Q}) = 1 - f(\mathcal{B}) = \frac{2}{3} - \frac{3}{2}p^2$.

Double reduction: If four homologous chromosomes in autotetraploids pair at meiosis following a quadrivalent pairing mode, two chromatids of a single chromosome can pass to the same gamete, which causes a phenomenon known as double reduction (Darlington 1929; Mather 1936). Double reduction arises from a

A



B

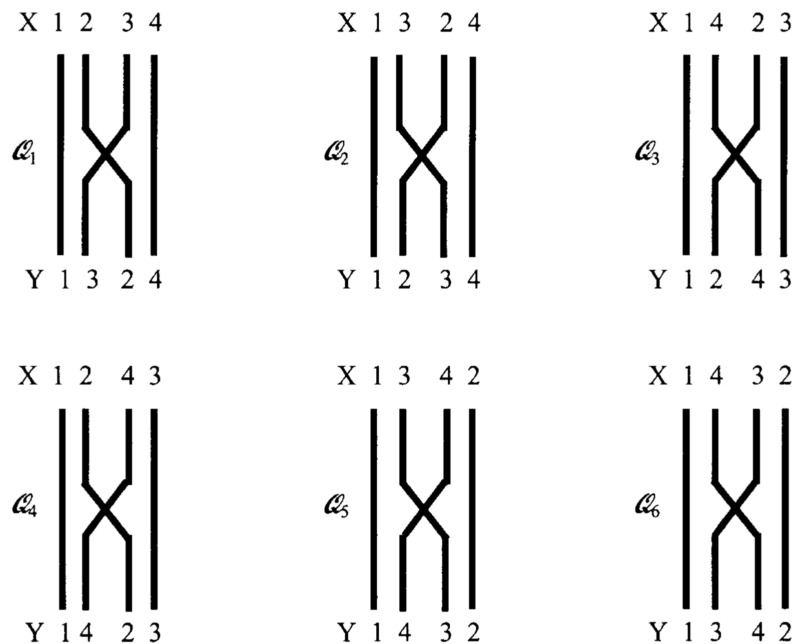


Figure 1.—The nine possibilities of pairing between chromosomes 1–4 with one point of pairing partner exchange in a general tetraploid model; arms X and Y. (A) Three possible bivalent pairings ($\mathscr{B}_1$–$\mathscr{B}_3$). Combinations 1-1 and 2-2 are homologous, as are 3-3 and 4-4 ($\mathscr{B}_1$). The other combinations are homeologous ($\mathscr{B}_2$ and $\mathscr{B}_3$). (B) Six possible quadrivalent pairings ($\mathscr{Q}_1$–$\mathscr{Q}_6$).

combination of three major events during meiosis: crossing over between nonsister chromatids, an appropriated pattern of disjunction, and the subsequent migration of the chromosomal segments carrying a pair of sister alleles to the same gamete. It seems likely that the frequency of double reduction is a constant for any given locus, depending on its distance from the centromere. To clearly describe the process of the formation of double reduction, Figure 2 (modified from Ronfort *et al.* 1998) illustrates possible segregation patterns of a marker locus in an autotetraploid individual following the formation of a quadrivalent. Type I describes the segregation patterns expected when there is no crossover between the centromere and the locus. The first division is then reductional. When a crossover occurs between the centromere and the locus (types II and III), the first division can be either equational (type II) or reductional (type III). Under type III, the second division may then lead to double reduction. In the present case, gametes *aa* and *bb* have undergone double reduction.

Estimation of frequency of double reduction: Consider two outbred autotetraploid parents *P* and *Q* that are crossed to generate a full-sib family of size *N*. Both parents and their progeny are genotyped using dominant and codominant markers. There are up to eight different alleles for a given marker locus, denoted by *a*, *b*, *c*, and *d* for parent *P* and *e*, *f*, *g*, and *h* for parent *Q*. For dominant markers, dominant alleles are indicated by the presence of bands on a gel and recessive alleles (denoted by *o*) are indicated by the absence of bands. For each parent (say *P*), there are a total of 16 possible phenotypes that can be classified into 5 different phenotypes in terms of the number of bands observed: four bands (one genotype, *abcd*), three bands (four genotypes, *abcc*, *abbc*, *aabc*, and *abco*), two bands (six genotypes, *abbb*, *aabb*, *aaab*, *abbo*, *aabo*, and *aboo*), one band (four genotypes, *aaaa*, *aaao*, *aaoo*, and *aooo*), and no band (one genotype, *oooo*). These 16 phenotypes can also be classified into 11 different types on the basis of the number of gamete phenotypes generated and the relative proportions of gamete formations:

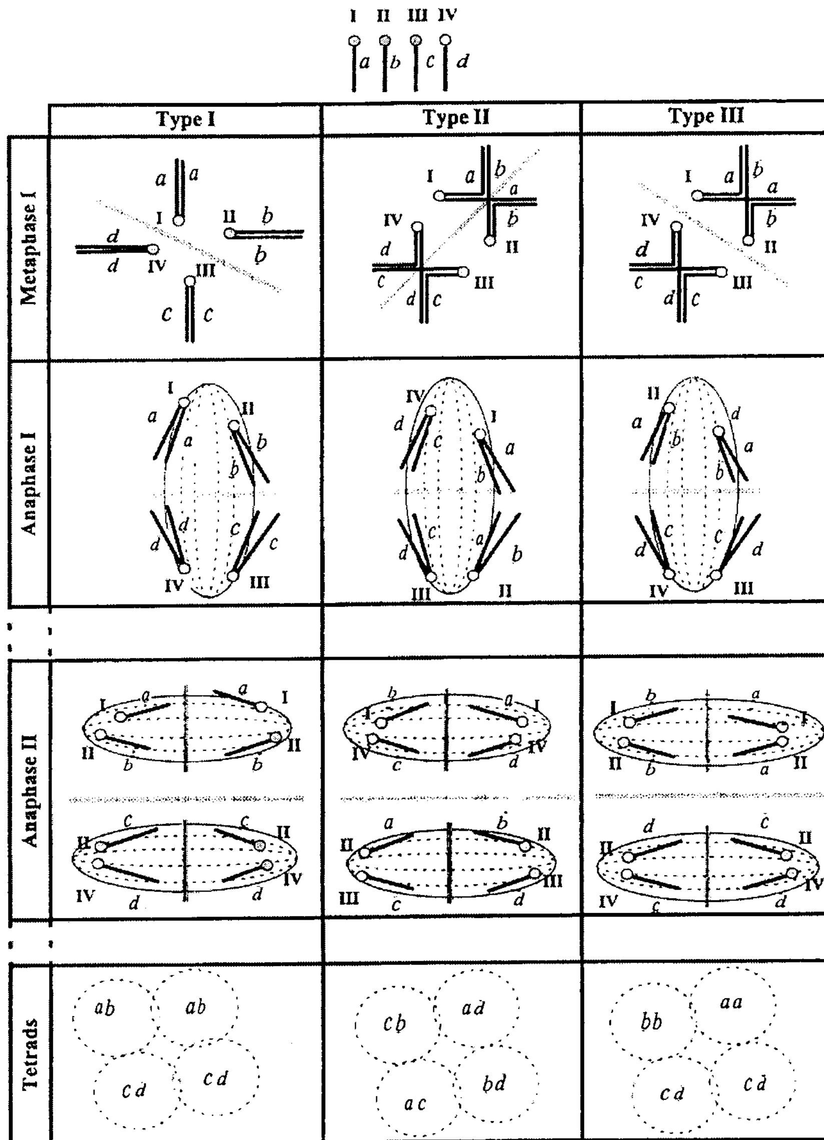


Figure 2.—Possible segregation patterns of a locus in an autotetraploid individual following the formation of a quadrivalent.

A_1, 10 gametes with formation proportion 1(*aa*):1(*bb*):1(*cc*):1(*dd*):1(*ab*):1(*ac*):1(*ad*):1(*bc*):1(*bd*):1(*cd*) for genotype *abcd*;

A_2, 7 gametes with formation proportion 1(*oo*):1(*ab*):1(*ac*):1(*bc*): 2(*a*_):2(*b*_):2(*c*_) for genotype *abco*;

A_3, 6 gametes with formation proportion 1(*aa*):1(*bb*):1(*ab*):2(*ac*):2(*bc*):3(*cc*) for genotype *abcc*, 1(*aa*):1(*cc*):1(*ac*):2(*ab*):2(*bc*):3(*bb*) for genotype *abbc*; or 1(*bb*):1(*cc*):1(*bc*):2(*ab*):2(*ac*):3(*aa*) for genotype *aabc*;

A_4, 4 gametes with formation proportion 1(*oo*):2(*a*_):2(*ab*):5(*b*_) for genotype *abbo* or 1(*oo*):2(*ab*): 2(*b*_):5(*a*_) for genotype *aabo*;

A_5, 4 gametes with formation proportion 1(*ab*):3(*oo*):3(*a*_):3(*b*_) for genotype *aboo*;

A_6, 3 gametes with formation proportion 1(*aa*):3(*ab*):6(*bb*) for genotype *abbb* or 1(*bb*):3(*ab*):6(*aa*) for genotype *aaab*;

A_7, 3 gametes with formation proportion 3(*aa*):3(*bb*):4(*ab*) for genotype *aabb*;

A_8, 2 gametes with formation proportion 1(*oo*):9(*a*_) for genotype *aaao*;

A_9, 2 gametes with formation proportion 3(*oo*):7(*a*_) for genotype *aaoo*;

A_{10}, 2 gametes with formation proportion 4(*a*_):6(*oo*) for genotype *aooo*;

A_{11}, 1 gamete *aa* for genotype *aaaa* and *oo* for genotype *oooo*.

It should be noted that the gamete types derived from the process of double reduction (Darlington 1929; Mather 1936) are considered in the above classifications. For example, for type A_1, genotype *abcd* produces double reduction-type gametes *aa, bb, cc,* and *dd*. Similar classifications can also be made for the second parent *Q* with alleles denoted by *e, f, g, h,* and *o*. For one parent, the markers of type A_1 are fully informative because all of the 10 gamete types can be phenotypically distinguished on the basis of their genotypes, whereas the

markers of types A_2 to A_{10} are partially informative because some of the gamete types have identical phenotypes. The markers from A_{11} are noninformative given its single gamete phenotype. For a real data set, all or some of the four alleles at a given marker for parent Q may be identical to those for parent P. All possible marker cross types between the two parents can be sorted into two groups, A and B. Group A includes all marker cross types in which the number of tetraploid progeny phenotypes is the product of the number of the diploid gametes from parent P and the number of the diploid gametes from parent Q. Group B comprises all marker cross types in which the number of tetraploid progeny phenotypes is less than the product of the number of the diploid gametes from parent P and the number of the diploid gametes from parent Q. Thus, whereas the phenotype of each progeny in group A is uniquely dependent on the phenotypes of two gametes derived from each parent, the phenotype of some progeny in group B can be generated by different combinations of the gametes from each parent. Marker group B results if one of the two following events is true: (1) there are at least two alleles common to the two parents; and (2) there is a common allele to the two parents, one of which has one or more nulls. All possible, if any, marker cross types that belong to group B are listed in Table 1.

Consider a marker with four alleles each assigned to one of the four chromosomes. The four alleles are labeled P_1, P_2, P_3, and P_4 for parent P and Q_1, Q_2, Q_3, and Q_4 for parent Q. Consider first parent P. For bivalent pairings, this parent generates six gametes, P_1P_2, P_1P_3, P_1P_4, P_2P_3, P_2P_4, and P_3P_4, whose frequencies are $\frac{1}{2}(\frac{1}{9} - \frac{1}{3}p + \frac{1}{4}p^2)$, $\frac{1}{4}(\frac{2}{9} + \frac{1}{3}p + \frac{5}{4}p^2)$, $\frac{1}{4}(\frac{2}{9} + \frac{1}{3}p + \frac{5}{4}p^2)$, $\frac{1}{4}(\frac{2}{9} + \frac{1}{3}p + \frac{5}{4}p^2)$, $\frac{1}{4}(\frac{2}{9} + \frac{1}{3}p + \frac{5}{4}p^2)$, and $\frac{1}{2}(\frac{1}{9} - \frac{1}{3}p + \frac{1}{4}p^2)$, respectively. For quadrivalent pairings, two types of diploid gametes are generated: (1) double reductions in which a gamete is derived from two sister chromatids of a single chromosome, *i.e.*, P_1P_1, P_2P_2, P_3P_3, and P_4P_4; and (2) random pairings in which a gamete results from two sister chromatids, each from one of two different chromosomes, *i.e.*, P_1P_2, P_1P_3, P_1P_4, P_2P_3, P_2P_4, and P_3P_4. If the frequency of double reduction during quadrivalent pairings is denoted by α for parent P, then the frequencies of the first-type gametes are each $\frac{1}{4}\alpha f(\mathcal{Q}) = \frac{1}{4}\alpha(\frac{2}{3} - \frac{3}{2}p^2)$ and the frequencies of the second-type gametes are each $\frac{1}{6}(1-\alpha)f(\mathcal{Q}) = \frac{1}{6}(1-\alpha)(\frac{2}{3} - \frac{3}{2}p^2)$. The second-type gametes resulting from quadrivalent pairings are mixed with the gametes from bivalent pairings. Thus, all the gametes from both bivalent and quadrivalent pairings can be arrayed in order by

$$_P\mathbf{G} = (P_1P_1\ P_2P_2\ P_3P_3\ P_4P_4\ P_1P_2\ P_1P_3\ P_1P_4\ P_2P_3\ P_2P_4\ P_3P_4)^T,$$

assuming a particular assignment of the four alleles among homologous chromosomes $P_1|P_2|P_3|P_4|$, where T denotes the transpose of the vector. The frequencies of the gametes are arrayed by

$$_P\mathbf{F} = [f(P_1P_1)\ f(P_2P_2)\ f(P_3P_3)\ f(P_4P_4)\ f(P_1P_2)\ f(P_1P_3)\ f(P_1P_4)\ f(P_2P_3)\ f(P_2P_4)\ f(P_3P_4)]^T,$$

where $f(P_1P_1) = f(P_2P_2) = f(P_3P_3) = f(P_4P_4) = \frac{1}{4}\alpha(\frac{2}{3} - \frac{3}{2}p^2)$, $f(P_1P_2) = f(P_3P_4) = \frac{1}{2}(\frac{1}{9} - \frac{1}{3}p + \frac{1}{4}p^2) + \frac{1}{6}(1 - \alpha)(\frac{2}{3} - \frac{3}{2}p^2)$, $f(P_1P_3) = f(P_1P_4) = f(P_2P_3) = f(P_2P_4) = \frac{1}{4}(\frac{2}{9} + \frac{1}{3}p + \frac{5}{4}p^2) + \frac{1}{6}(1 - \alpha)(\frac{2}{3} - \frac{3}{2}p^2)$. The gamete frequency vector $_P\mathbf{F}$ is partitioned into two components due to bivalent and quadrivalent pairings,

$$_P\mathbf{F} = {_P\mathbf{F}^{\mathcal{B}}} + {_P\mathbf{F}^{\mathcal{Q}}} = {_P\mathbf{F}^{\mathcal{B}}} + \mathbf{P}\ {_P\mathbf{A}},$$

where

$$_P\mathbf{F}^{\mathcal{B}} = [0\ 0\ 0\ 0\ \tfrac{1}{2}(\tfrac{1}{9} - \tfrac{1}{3}p + \tfrac{1}{4}p^2)\ \tfrac{1}{4}(\tfrac{2}{9} + \tfrac{1}{3}p + \tfrac{5}{4}p^2)\ \tfrac{1}{4}(\tfrac{2}{9} + \tfrac{1}{3}p + \tfrac{5}{4}p^2)\ \tfrac{1}{4}(\tfrac{2}{9} + \tfrac{1}{3}p + \tfrac{5}{4}p^2)\ \tfrac{1}{4}(\tfrac{2}{9} + \tfrac{1}{3}p + \tfrac{5}{4}p^2)\ \tfrac{1}{2}(\tfrac{1}{9} - \tfrac{1}{3}p + \tfrac{1}{4}p^2)]^T,$$

$$\mathbf{P} = [(\tfrac{2}{3} - \tfrac{3}{2}p^2)\ (\tfrac{2}{3} - \tfrac{3}{2}p^2)\ (\tfrac{2}{3} - \tfrac{3}{2}p^2)\ (\tfrac{2}{3} - \tfrac{3}{2}p^2)\ (\tfrac{2}{3} - \tfrac{3}{2}p^2)\ (\tfrac{2}{3} - \tfrac{3}{2}p^2)\ (\tfrac{2}{3} - \tfrac{3}{2}p^2)\ (\tfrac{2}{3} - \tfrac{3}{2}p^2)\ (\tfrac{2}{3} - \tfrac{3}{2}p^2)\ (\tfrac{2}{3} - \tfrac{3}{2}p^2)]^T,$$

and

$$_P\mathbf{A} = [\tfrac{\alpha}{4}\ \tfrac{\alpha}{4}\ \tfrac{\alpha}{4}\ \tfrac{\alpha}{4}\ \tfrac{(1-\alpha)}{6}\ \tfrac{(1-\alpha)}{6}\ \tfrac{(1-\alpha)}{6}\ \tfrac{(1-\alpha)}{6}\ \tfrac{(1-\alpha)}{6}\ \tfrac{(1-\alpha)}{6}]^T.$$

$_P\mathbf{A}$ is the vector for the frequencies of the gametes generated through quadrivalent pairings.

Similarly, for parent Q, we have

$$\mathbf{G}_Q = (Q_1Q_1\ Q_2Q_2\ Q_3Q_3\ Q_4Q_4\ Q_1Q_2\ Q_1Q_3\ Q_1Q_4\ Q_2Q_3\ Q_2Q_4\ Q_3Q_4)^T,$$

$$\mathbf{F}_Q = [f(Q_1Q_1)\ f(Q_2Q_2)\ f(Q_3Q_3)\ f(Q_4Q_4)\ f(Q_1Q_2)\ f(Q_1Q_3)\ f(Q_1Q_4)\ f(Q_2Q_3)\ f(Q_2Q_4)\ f(Q_3Q_4)]^T,$$

$$\mathbf{F}_Q = \mathbf{F}_Q^{\mathcal{B}} + \mathbf{F}_Q^{\mathcal{Q}} = \mathbf{F}_Q^{\mathcal{B}} + \mathbf{Q}\mathbf{A}_Q,$$

where $f(Q_1Q_1) = f(Q_2Q_2) = f(Q_3Q_3) = f(Q_4Q_4) = \frac{1}{4}\beta(\frac{2}{3} - \frac{3}{2}q^2)$, $f(Q_1Q_2) = f(Q_3Q_4) = \frac{1}{2}(\frac{1}{9} - \frac{1}{3}q + \frac{1}{4}q^2) + \frac{1}{6}(1 - \beta)(\frac{2}{3} - \frac{3}{2}q^2)$, $f(Q_1Q_3) = f(Q_1Q_4) = f(Q_2Q_3) = f(Q_2Q_4) = \frac{1}{4}(\frac{2}{9} + \frac{1}{3}q + \frac{5}{4}q^2) + \frac{1}{6}(1 - \beta)(\frac{2}{3} - \frac{3}{2}q^2)$, q is the preferential pairing factor for parent Q, β is the frequency of double reduction for this parent,

$$\mathbf{F}_Q^{\mathcal{B}} = [0\ 0\ 0\ 0\ \tfrac{1}{2}(\tfrac{1}{9} - \tfrac{1}{3}q + \tfrac{1}{4}q^2)\ \tfrac{1}{4}(\tfrac{2}{9} + \tfrac{1}{3}q + \tfrac{5}{4}q^2)\ \tfrac{1}{4}(\tfrac{2}{9} + \tfrac{1}{3}q + \tfrac{5}{4}q^2)\ \tfrac{1}{4}(\tfrac{2}{9} + \tfrac{1}{3}q + \tfrac{5}{4}q^2)\ \tfrac{1}{4}(\tfrac{2}{9} + \tfrac{1}{3}q + \tfrac{5}{4}q^2)\ \tfrac{1}{2}(\tfrac{1}{9} - \tfrac{1}{3}q + \tfrac{1}{4}q^2)]^T,$$

$$\mathbf{Q} = [(\tfrac{2}{3} - \tfrac{3}{2}q^2)\ (\tfrac{2}{3} - \tfrac{3}{2}q^2)\ (\tfrac{2}{3} - \tfrac{3}{2}q^2)\ (\tfrac{2}{3} - \tfrac{3}{2}q^2)\ (\tfrac{2}{3} - \tfrac{3}{2}q^2)\ (\tfrac{2}{3} - \tfrac{3}{2}q^2)\ (\tfrac{2}{3} - \tfrac{3}{2}q^2)\ (\tfrac{2}{3} - \tfrac{3}{2}q^2)\ (\tfrac{2}{3} - \tfrac{3}{2}q^2)\ (\tfrac{2}{3} - \tfrac{3}{2}q^2)]^T,$$

and

$$\mathbf{A}_Q = [\tfrac{\beta}{4}\ \tfrac{\beta}{4}\ \tfrac{\beta}{4}\ \tfrac{\beta}{4}\ \tfrac{(1-\beta)}{6}\ \tfrac{(1-\beta)}{6}\ \tfrac{(1-\beta)}{6}\ \tfrac{(1-\beta)}{6}\ \tfrac{(1-\beta)}{6}\ \tfrac{(1-\beta)}{6}]^T.$$

Marker group A: For a fully informative marker that

TABLE 1

Marker cross types of group B segregating in a full-sib family derived from two tetraploid parents *P* and *Q*

Type	Cross ($P \times Q$)	Parent *P*		Parent *Q*		No. zygote phenotypes
		No. gametes	Formation proportion	No. gametes	Formation proportion	
B_1	*abcd* × *abef*	10	1:1:1:1:1:1:1:1:1:1	10	1:1:1:1:1:1:1:1:1:1	50
B_2	*abcd* × *abce*	10	1:1:1:1:1:1:1:1:1:1	10	1:1:1:1:1:1:1:1:1:1	28
B_3	*abcd* × *abcd*	10	1:1:1:1:1:1:1:1:1:1	10	1:1:1:1:1:1:1:1:1:1	15
B_4	*abcd* × *aefo*	10	1:1:1:1:1:1:1:1:1:1	7	1:1:1:1:2:2:2	48
B_5	*abcd* × *abeo*	10	1:1:1:1:1:1:1:1:1:1	7	1:1:1:1:2:2:2	29
B_6	*abcd* × *abco*	10	1:1:1:1:1:1:1:1:1:1	7	1:1:1:1:2:2:2	15
B_7	*abcd* × *abee*	10	1:1:1:1:1:1:1:1:1:1	6	1:1:1:2:2:3	22
B_8	*abcd* × *aabc*	10	1:1:1:1:1:1:1:1:1:1	6	1:1:1:2:2:3	11
B_9	*abcd* × *aabo*	10	1:1:1:1:1:1:1:1:1:1	4	1:2:2:5	15
B_{10}	*abcd* × *aeoo*	10	1:1:1:1:1:1:1:1:1:1	4	1:3:3:3	23
B_{11}	*abcd* × *aboo*	10	1:1:1:1:1:1:1:1:1:1	4	1:3:3:3	15
B_{12}	*abcd* × *aabb*	10	1:1:1:1:1:1:1:1:1:1	3	3:3:4	12
B_{13}	*abcd* × *aaab*	10	1:1:1:1:1:1:1:1:1:1	3	1:3:6	12
B_{14}	*abcd* × *aaao*	10	1:1:1:1:1:1:1:1:1:1	2	1:9	13
B_{15}	*abcd* × *aaoo*	10	1:1:1:1:1:1:1:1:1:1	2	3:7	13
B_{16}	*abcd* × *aooo*	10	1:1:1:1:1:1:1:1:1:1	2	4:6	13
B_{17}	*abco* × *adeo*	7	1:1:1:1:2:2:2	7	1:1:1:1:2:2:2	27
B_{18}	*abco* × *abdo*	7	1:1:1:1:2:2:2	7	1:1:1:1:2:2:2	16
B_{19}	*abco* × *abco*	7	1:1:1:1:2:2:2	7	1:1:1:1:2:2:2	8
B_{20}	*abco* × *adde*	7	1:1:1:1:2:2:2	6	1:1:1:2:2:3	25
B_{21}	*abco* × *abdd*	7	1:1:1:1:2:2:2	6	1:1:1:2:2:3	14
B_{22}	*abco* × *aabc*	7	1:1:1:1:2:2:2	6	1:1:1:2:2:3	7
B_{23}	*abco* × *aabo*	7	1:1:1:1:2:2:2	4	1:2:2:5	8
B_{24}	*abco* × *adoo*	7	1:1:1:1:2:2:2	4	1:3:3:3	14
B_{25}	*abco* × *aboo*	7	1:1:1:1:2:2:2	4	1:3:3:3	8
B_{26}	*abco* × *aadd*	7	1:1:1:1:2:2:2	3	3:3:4	11
B_{27}	*abco* × *aabb*	7	1:1:1:1:2:2:2	3	3:3:4	7
B_{28}	*abco* × *addo*	7	1:1:1:1:2:2:2	4	1:2:2:5	16
B_{29}	*abco* × *aado*	7	1:1:1:1:2:2:2	4	1:2:2:5	16
B_{30}	*abco* × *aabo*	7	1:1:1:1:2:2:2	4	1:2:2:5	7
B_{31}	*abco* × *aaab*	7	1:1:1:1:2:2:2	3	1:3:6	6
B_{32}	*abco* × *aaao*	7	1:1:1:1:2:2:2	2	1:9	5
B_{33}	*abco* × *aaoo*	7	1:1:1:1:2:2:2	2	3:7	5
B_{34}	*abco* × *aooo*	7	1:1:1:1:2:2:2	2	4:6	5
B_{35}	*aabc* × *aabd*	6	1:1:1:2:2:3	6	1:1:1:2:2:3	13
B_{36}	*aabc* × *abdd*	6	1:1:1:2:2:3	6	1:1:1:2:2:3	13
B_{37}	*aabc* × *bcdd*	6	1:1:1:2:2:3	6	1:1:1:2:2:3	13
B_{38}	*aabc* × *aabc*	6	1:1:1:2:2:3	6	1:1:1:2:2:3	7
B_{39}	*aabc* × *abbc*	6	1:1:1:2:2:3	6	1:1:1:2:2:3	7
B_{40}	*aabc* × *aado*	6	1:1:1:2:2:3	4	1:2:2:5	14
B_{41}	*aabc* × *abbo*	6	1:1:1:2:2:3	4	1:2:2:5	7
B_{42}	*aabc* × *aabo*	6	1:1:1:2:2:3	4	1:2:2:5	7
B_{43}	*aabc* × *adoo*	6	1:1:1:2:2:3	4	1:3:3:3	14
B_{44}	*aabc* × *bdoo*	6	1:1:1:2:2:3	4	1:3:3:3	14
B_{45}	*aabc* × *aboo*	6	1:1:1:2:2:3	4	1:3:3:3	7
B_{46}	*aabc* × *bcoo*	6	1:1:1:2:2:3	4	1:3:3:3	6
B_{47}	*aabc* × *aabb*	6	1:1:1:2:2:3	3	3:3:4	6
B_{48}	*aabc* × *bbcc*	6	1:1:1:2:2:3	3	3:3:4	7
B_{49}	*aabc* × *aaab*	6	1:1:1:2:2:3	3	1:3:6	6
B_{50}	*aabc* × *abbb*	6	1:1:1:2:2:3	3	1:3:6	6
B_{51}	*aabc* × *aaao*	6	1:1:1:2:2:3	2	1:9	7
B_{52}	*aabc* × *aaoo*	6	1:1:1:2:2:3	2	3:7	7

(*continued*)

TABLE 1

(Continued)

Type	Cross ($P \times Q$)	Parent P No. gametes	Parent P Formation proportion	Parent Q No. gametes	Parent Q Formation proportion	No. zygote phenotypes
B_{53}	*aabc* × *aooo*	6	1:1:1:2:2:3	2	4:6	7
B_{54}	*aabo* × *aaco*	4	1:2:2:5	4	1:2:2:5	8
B_{55}	*aabo* × *acco*	4	1:2:2:5	4	1:2:2:5	8
B_{56}	*aabo* × *aabo*	4	1:2:2:5	4	1:2:2:5	4
B_{57}	*aabo* × *abbo*	4	1:2:2:5	4	1:2:2:5	4
B_{58}	*aabo* × *bcoo*	4	1:2:2:5	4	1:3:3:3	3
B_{59}	*aabo* × *acoo*	4	1:2:2:5	4	1:3:3:3	8
B_{60}	*aabo* × *aacc*	4	1:2:2:5	3	3:3:4	6
B_{61}	*aabo* × *bbcc*	4	1:2:2:5	3	3:3:4	6
B_{62}	*aabo* × *aabb*	4	1:2:2:5	3	3:3:4	3
B_{63}	*aabo* × *aaac*	4	1:2:2:5	3	1:3:6	6
B_{64}	*aabo* × *bbbc*	4	1:2:2:5	3	1:3:6	6
B_{65}	*aabo* × *aaab*	4	1:2:2:5	3	1:3:6	3
B_{66}	*aabo* × *abbb*	4	1:2:2:5	3	3:3:4	3
B_{67}	*aabo* × *aaao*	4	1:2:2:5	2	1:9	3
B_{68}	*aabo* × *aaoo*	4	1:2:2:5	2	3:7	3
B_{69}	*aabo* × *bbbo*	4	1:2:2:5	2	1:9	4
B_{70}	*aabo* × *bboo*	4	1:2:2:5	2	3:7	4
B_{71}	*aabo* × *aooo*	4	1:2:2:5	2	4:6	3
B_{72}	*aabo* × *booo*	4	1:2:2:5	2	4:6	4
B_{73}	*aboo* × *acoo*	4	1:3:3:3	4	1:3:3:3	8
B_{74}	*aboo* × *aboo*	4	1:3:3:3	4	1:3:3:3	4
B_{75}	*aboo* × *aacc*	4	1:3:3:3	3	3:3:4	5
B_{76}	*aboo* × *aabb*	4	1:3:3:3	3	3:3:4	3
B_{77}	*aboo* × *aaac*	4	1:3:3:3	3	1:3:6	5
B_{78}	*aboo* × *accc*	4	1:3:3:3	3	1:3:6	5
B_{79}	*aboo* × *aaab*	4	1:3:3:3	3	1:3:6	3
B_{80}	*aboo* × *aaao*	4	1:3:3:3	2	1:9	4
B_{81}	*aboo* × *aaoo*	4	1:3:3:3	2	3:7	4
B_{82}	*aboo* × *aooo*	4	1:3:3:3	2	4:6	4
B_{83}	*aabb* × *aabb*	3	3:3:4	3	3:3:4	3
B_{84}	*aabb* × *aaab*	3	3:3:4	3	1:3:6	3
B_{85}	*aabb* × *aaao*	3	3:3:4	2	1:9	3
B_{86}	*aabb* × *aaoo*	3	3:3:4	2	3:7	3
B_{87}	*aabb* × *aooo*	3	3:3:4	2	4:6	3
B_{88}	*aaab* × *aaab*	3	1:3:6	3	1:3:6	3
B_{89}	*aaab* × *abbb*	3	1:3:6	3	1:3:6	3
B_{90}	*aaab* × *aaao*	3	1:3:6	2	1:9	3
B_{91}	*aaab* × *bbbo*	3	1:3:6	2	1:9	3
B_{92}	*aaab* × *aaoo*	3	1:3:6	2	3:7	3
B_{93}	*aaab* × *bboo*	3	1:3:6	2	3:7	3
B_{94}	*aaab* × *aooo*	3	1:3:6	2	4:6	3
B_{95}	*aaab* × *booo*	3	1:3:6	2	4:6	3
B_{96}	*aaao* × *aaao*	2	1:9	2	1:9	2
B_{97}	*aaao* × *aaoo*	2	1:9	2	3:7	2
B_{98}	*aaao* × *aooo*	2	1:9	2	4:6	2
B_{99}	*aaoo* × *aaoo*	2	3:7	2	3:7	2
B_{100}	*aaoo* × *aooo*	2	3:7	2	4:6	2
B_{101}	*aooo* × *aooo*	2	4:6	2	4:6	2

As defined, the number of the zygote phenotype in group B is less than the product of the number of the gamete phenotype from parent P and the number of the gamete phenotype from parent Q. The marker cross types within the same boxes can be distinguished in terms of the arrangements of different zygotes in the matrix $\ddot{\mathbf{F}}$.

generates $10 \times 10 = 100$ different zygotes, the progeny's phenotypes are exactly consistent with their genotypes. On the basis of the observations of each phenotype or genotype in the full-sib family, the maximum-likelihood estimate of the frequencies of double reduction can be obtained by using an explicit expression. When the two parents are crossed, the zygote genotypes in the full-sib family can be expressed as

$$\dot{\mathbf{G}} = {}_P\mathbf{G}\mathbf{G}_Q^{\mathrm{T}},$$

where $\dot{\mathbf{G}}$ is the (10×10) matrix in which each element ${}_{r_1 r_2}G_{u_1 u_2}$ represents a zygote genotype $P_{r_1}P_{r_2}Q_{u_1}Q_{u_2}$ at the marker considered (r_1, $r_2 = 1, 2, 3, 4$ are the two marker alleles contributed by parent P and u_1, $u_2 = 1, 2, 3, 4$ are the two alleles contributed by parent Q). The corresponding (10×10) matrix for the frequencies of the zygotes in the full-sib family is denoted by

$$\dot{\mathbf{F}} = {}_P\mathbf{F}\mathbf{F}_Q^{\mathrm{T}},$$

assuming that the formation of gametes is independent between the two parents. The occurrence of double reduction in each progeny genotype can also be expressed in a (10×10) matrix form. But this matrix differs depending on which parent contributes to double reduction at meiosis, expressed as

$$
{}_P\dot{\mathbf{D}} = \begin{array}{c} \\ P_1P_1 \\ P_2P_2 \\ P_3P_3 \\ P_4P_4 \\ P_1P_2 \\ P_1P_3 \\ P_1P_4 \\ P_2P_3 \\ P_2P_4 \\ P_3P_4 \end{array}
\begin{array}{c} \begin{array}{cccccccccc} Q_1Q_1 & Q_2Q_2 & Q_3Q_3 & Q_4Q_4 & Q_1Q_2 & Q_1Q_3 & Q_1Q_4 & Q_2Q_3 & Q_2Q_4 & Q_3Q_4 \end{array} \\
\left[\begin{array}{cccccccccc}
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{array}\right] \end{array},
$$

if double reductions are contributed by parent P, and

$$
\dot{\mathbf{D}}_Q = \begin{array}{c} \\ P_1P_1 \\ P_2P_2 \\ P_3P_3 \\ P_4P_4 \\ P_1P_2 \\ P_1P_3 \\ P_1P_4 \\ P_2P_3 \\ P_2P_4 \\ P_3P_4 \end{array}
\begin{array}{c} \begin{array}{cccccccccc} Q_1Q_1 & Q_2Q_2 & Q_3Q_3 & Q_4Q_4 & Q_1Q_2 & Q_1Q_3 & Q_1Q_4 & Q_2Q_3 & Q_2Q_4 & Q_3Q_4 \end{array} \\
\left[\begin{array}{cccccccccc}
1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0
\end{array}\right] \end{array},
$$

if double reductions are contributed by parent Q.

For marker group A, distinct zygote phenotypes can be predicted on the basis of the genotypes of two gametes each from a parent; thus the vector ($\boldsymbol{\pi}$) of unknown parameters, including the preferential pairing factors p and q and the frequencies of double reduction α and β, can be estimated by formulating the likelihood function of marker data from gametes given the unknown parameter vector $\boldsymbol{\pi}$,

$$P({}_P\mathbf{m}, \mathbf{m}_Q|\boldsymbol{\pi}) = \prod_{j=1}^{N} P[{}_P\mathbf{m}_{(j)}, \mathbf{m}_{Q(j)}|\boldsymbol{\pi}], \tag{2}$$

where ${}_P\mathbf{m}$ and $\mathbf{m}_Q$ are a $(N \times \xi)$ and $(N \times \zeta)$ 1/0 matrix describing ξ and ζ gamete phenotypes (ξ, $\zeta = 10$ for A_1, 7 for A_2, 6 for A_3, 4 for A_4 and A_5, 3 for A_6 and A_7, 2 for A_8, A_9 and A_{10}, and 1 for A_{11}) generated by parent P or Q, respectively. For a fully informative marker type, the maximum-likelihood estimate (MLE) of each unknown has an explicit expression. However, for many other partially informative marker types, such explicit expressions cannot be written. In this case, the expectation-maximization (EM) algorithm can be implemented to estimate these parameters (Dempster *et al.* 1977).

In step E, the expected number of double reductions contained in each zygote phenotype is calculated for parent P,

$${}_{r_1r_2}\Xi_{u_1u_2j}^{\{\tau+1\}} = \frac{{}_{r_1r_2}\mathbf{m}_j^{\mathrm{T}}[{}_P\mathbf{I}^{\mathrm{T}}({}_P\dot{\mathbf{D}} \circ \dot{\mathbf{F}}^{\{\tau\}})\mathbf{I}_Q]\mathbf{m}_{u_1u_2j}}{{}_{r_1r_2}\mathbf{m}_j^{\mathrm{T}}[{}_P\mathbf{I}^{\mathrm{T}}\dot{\mathbf{F}}^{\{\tau\}}\mathbf{I}_Q]\mathbf{m}_{u_1u_2j}}, \tag{3a}$$

and for parent Q,

$${}_{r_1r_2}\mathrm{Z}_{u_1u_2j}^{\{\tau+1\}} = \frac{{}_{r_1r_2}\mathbf{m}_j^{\mathrm{T}}[{}_P\mathbf{I}^{\mathrm{T}}(\dot{\mathbf{D}}_Q \circ \dot{\mathbf{F}}^{\{\tau\}})\mathbf{I}_Q]\mathbf{m}_{u_1u_2j}}{{}_{r_1r_2}\mathbf{m}_j^{\mathrm{T}}[{}_P\mathbf{I}^{\mathrm{T}}\dot{\mathbf{F}}^{\{\tau\}}\mathbf{I}_Q]\mathbf{m}_{u_1u_2j}}, \tag{3b}$$

where ${}_{r_1r_2}\mathbf{m}_j$ is the jth row of ${}_P\mathbf{m}$ representing the gamete phenotype $P_{r_1}P_{r_2}$ from parent P, which the jth individual has received; and ${}_P\mathbf{I}$ is the $(\xi \times 10)$ design matrices relating the gamete genotypes to the gamete phenotypes for parent P. Similarly, $\mathbf{m}_{u_1u_2j}$ and $\mathbf{I}_Q$ can be defined for parent Q.

In the M step, the frequencies of double reduction are calculated using the equations

$$\alpha^{\{\tau+1\}} = \frac{1}{N}\sum_{r_1=1}^{4}\sum_{r_2=1}^{4}\sum_{u_1=1}^{4}\sum_{u_2=1}^{4}[{}_{r_1r_2}\Xi_{u_1u_2j}^{\{\tau+1\}}], \tag{4a}$$

or parent P, and

$$\beta^{\{\tau+1\}} = \frac{1}{N}\sum_{r_1=1}^{4}\sum_{r_2=1}^{4}\sum_{u_1=1}^{4}\sum_{u_2=1}^{4}[{}_{r_1r_2}\mathrm{Z}_{u_1u_2j}^{\{\tau+1\}}], \tag{4b}$$

for parent Q. Also, the preferential pairing factors p and q are calculated by solving the log-likelihood equations

$$\sum_{r_1=1}^{4}\sum_{r_2=1}^{4}\sum_{j=1}^{N}\left\{{}_{r_1r_2}X_j^{\{\tau+1\}}\frac{\partial}{\partial p}[{}_{r_1r_2}m_j\,{}_P\mathbf{I}\;\mathbf{P}^{\{\tau\}}] + {}_{r_1r_2}U_j^{\{\tau+1\}}\frac{\partial}{\partial p}[{}_{r_1r_2}m_j\,{}_P\mathbf{I}\;{}_P\mathbf{F}^{\mathcal{B}\{\tau\}}]\right\} = 0, \tag{4c}$$

$$\sum_{u_1=1}^{4}\sum_{u_2=1}^{4}\sum_{j=1}^{N}\left\{Y_{u_1u_2j}^{\{\tau+1\}}\frac{\partial}{\partial q}[m_{u_1u_2j}\mathbf{I}_Q\mathbf{Q}^{\{\tau\}}] + V_{u_1u_2j}^{\{\tau+1\}}\frac{\partial}{\partial q}[m_{u_1u_2j}\,\mathbf{I}_Q\mathbf{F}_Q^{\mathcal{B}\{\tau\}}]\right\} = 0, \tag{4d}$$

where

$${}_{r_1r_2}X_j^{(\tau+1)} = \frac{{}_{r_1r_2}\mathbf{m}_{jP}\mathbf{I}\ {}_P\mathbf{A}^{(\tau)}}{{}_{r_1r_2}\mathbf{m}_{jP}\mathbf{I}\ {}_P\mathbf{F}^{(\tau)}}, \qquad {}_{r_1r_2}U_j^{(\tau+1)} = \frac{1}{{}_{r_1r_2}\mathbf{m}_{jP}\mathbf{I}\ {}_P\mathbf{F}^{(\tau)}},$$

$$Y_{u_1u_2j}^{(\tau+1)} = \frac{\mathbf{m}_{u_1u_2j}\ \mathbf{I}_Q\mathbf{A}_Q^{(\tau)}}{\mathbf{m}_{u_1u_2j}\ \mathbf{I}_Q\mathbf{F}_Q^{(\tau)}}, \qquad V_{u_1u_2j}^{(\tau+1)} = \frac{1}{\mathbf{m}_{u_1u_2j}\ \mathbf{I}_Q\mathbf{F}_Q^{(\tau)}},$$

The E and M steps are repeated until the estimates converge to stable values. The estimates at convergence are the MLEs of the unknowns.

Marker group B: For the markers from group B, zygote phenotypes can be determined only after two gametes are fused. Thus, marker analysis for group B should be based on zygotic phenotypes. In this case, 100-dimension vectors for zygotic phenotypes and their frequencies are expressed as

$$\ddot{\mathbf{G}} = {}_P\mathbf{G} \otimes \mathbf{G}_Q,$$

$$\ddot{\mathbf{F}} = {}_P\mathbf{F} \otimes \mathbf{F}_Q.$$

Correspondingly, 100-dimension vectors for the occurrence of double reductions with parents P and Q are expressed as

$${}_P\ddot{\mathbf{D}} = \Big(\overbrace{11\ldots1}^{40}\quad\overbrace{00\ldots0}^{60}\Big)^T,$$

$$\ddot{\mathbf{D}}_Q = \Big(\underbrace{1111000000}_{1}\ \underbrace{1111000000}_{2}\ \ldots\ \underbrace{1111000000}_{10}\Big)^T.$$

As for marker group A, in step E, calculate

$${}_{r_1r_2}\Xi_{u_1u_2j}^{(\tau+1)} = \frac{{}_{r_1r_2}\mathbf{M}^{\mathrm{T}}_{u_1u_2j}[\mathbf{I}^{\mathrm{T}}({}_P\ddot{\mathbf{D}}\circ\ddot{\mathbf{F}}^{(\tau)})]}{{}_{r_1r_2}\mathbf{M}^{\mathrm{T}}_{u_1u_2j}(\mathbf{I}^{\mathrm{T}}\ddot{\mathbf{F}}^{(\tau)})},$$

$${}_{r_1r_2}Z_{u_1u_2j}^{(\tau+1)} = \frac{{}_{r_1r_2}\mathbf{M}^{\mathrm{T}}_{u_1u_2j}[\mathbf{I}^{\mathrm{T}}(\ddot{\mathbf{D}}_Q\circ\ddot{\mathbf{F}}^{(\tau)})]}{{}_{r_1r_2}\mathbf{M}^{\mathrm{T}}_{u_1u_2j}(\mathbf{I}^{\mathrm{T}}\ddot{\mathbf{F}}^{(\tau)})},$$

where ${}_{r_1r_2}\mathbf{M}_{u_1u_2j}$ is the jth row of the $(N \times \phi)$ matrix $\mathbf{M}$ for marker genotypes with ϕ being the number of distinguishable zygotic genotypes in the full-sib family, whose element is 1 if the jth individual has the genotype $P_{r_1}P_{r_2}Q_{u_1}Q_{u_2}$ and is 0 otherwise, and $\mathbf{I}$ is the $(\phi \times 100)$ incidence matrices relating the zygotic genotypes to zygotic phenotypes. The form and structure of $\mathbf{I}$ depend on marker cross types in group B (Table 1). In the M step, the frequencies of double reduction are estimated using Equations 4a and 4b. The preferential pairing factors p and q can be estimated by solving the corresponding log-likelihood equations as shown in Equations 4c and 4d.

Tests for the preferential pairing factor and frequency of double reduction: The existence and magnitude of preferential pairings and double reduction have particular evolutionary significance and implications for genetic and breeding research. If p or $q = 0$, this means full homology among four single chromosomes, typical of allopolyploids derived from the combination of different genomes. If p or $q = \frac{2}{3}$, this means that homeologous pairings characterized by autopolyploids do not exist. Similarly, α or β can take any value from 0 (with pure random chromosome segregation) to $\frac{1}{7}$ (with pure random chromatid segregation) to $\frac{1}{6}$ (with complete equational segregation; Mather 1936; Fisher and Mather 1943). It is possible to test whether the estimated p, q, α, or β is statistically different than a particular value. This can be done by formulating the alternative hypotheses and calculating the likelihood-ratio (LRT) test statistic

$$\mathrm{LRT_A} = -2\log\left[\frac{P({}_P\mathbf{m}, \mathbf{m}_Q|\alpha = \beta = 0)}{P({}_P\mathbf{m}, \mathbf{m}_Q|\hat{\pi})}\right],$$

for marker group A, and

$$\mathrm{LRT_B} = -2\log\left[\frac{P(\mathbf{M}|\alpha = \beta = 0)}{P(\mathbf{M}|\hat{\pi})}\right],$$

for marker group B, if one intends to test whether double reduction occurs in both parents. Each of the two LRTs follows approximately a chi-square distribution with 2 d.f. Other LRTs also can be formulated in a similar way.

SIMULATION

Simulation experiments are performed to demonstrate the statistical properties of the MLEs of the preferential pairing factors and the frequencies of double reduction at meiosis in autopolyploids. The experiments are designed to consider the effects of different marker types, different degrees of preferential pairs, and different frequencies of double reduction on the parameter estimation. Because it is difficult and also unnecessary to consider all possible marker cross types (Table 1), only six representative types, three from each marker group, were chosen to reflect different informativeness of markers (Table 2). The experiments allow for changes of the preferential pairing factors (0 and $\frac{1}{3}$) and the frequencies of double reduction (0, 0.08, and 0.15) within their respective boundaries. For simplicity, the preferential pairing factors are assumed equal between the two parents ($p = q$), and so are the frequencies of double reduction ($\alpha = \beta$). As shown in Table 2, our simulation experiments are also created to examine the interaction effects of these factors on parameter estimation. Given the hypothesized marker cross types and hypothesized parameter values, meioses for two parents P and Q are simulated and the phenotypes of the progeny at a given marker are generated. Luo *et al.* (2000) compared the power to detect double reduction under different sample sizes and suggested that a sample of size 100 would be adequate for providing a reasonable estimate for double reduction. In this study, our simulation is based on a sample size of 100.

TABLE 2

MLEs and standard errors (in parentheses) of the preferential pairing factors and the frequencies of double reduction for different marker cross types

		$\alpha = \beta$								
		0			0.08			0.15		
$p = q$	Cross type	$\hat{p} = \hat{q}$	$\hat{\alpha} = \hat{\beta}$	Power[a]	$\hat{p} = \hat{q}$	$\hat{\alpha} = \hat{\beta}$	Power[a]	$\hat{p} = \hat{q}$	$\hat{\alpha} = \hat{\beta}$	Power[a]
0	*abcd* × *efgh*	0.000 (0.000)	0.000 (0.000)	0	0.000 (0.000)	0.082 (0.014)	100	0.000 (0.000)	0.150 (0.016)	100
	abcd × *efgg*	0.004 (0.003)	0.001 (0.000)	0	0.004 (0.004)	0.076 (0.019)	98	0.005 (0.004)	0.153 (0.029)	98
	abcc × *efgg*	0.006 (0.004)	0.004 (0.002)	0	0.006 (0.005)	0.086 (0.024)	87	0.006 (0.005)	0.152 (0.036)	92
	abcd × *abcd*	0.010 (0.009)	0.007 (0.003)	1	0.011 (0.010)	0.073 (0.034)	64	0.015 (0.013)	0.143 (0.047)	75
	abcd × *abcc*	0.014 (0.012)	0.009 (0.006)	1	0.015 (0.014)	0.087 (0.041)	40	0.017 (0.014)	0.141 (0.053)	51
	abco × *aaoo*	0.018 (0.017)	0.010 (0.012)	2	0.019 (0.017)	0.089 (0.050)	32	0.022 (0.019)	0.142 (0.061)	44
1/3	*abcd* × *efgh*	0.332 (0.018)	0.000 (0.000)	0	0.331 (0.019)	0.083 (0.017)	98	0.330 (0.019)	0.152 (0.017)	98
	abcd × *efgg*	0.321 (0.021)	0.001 (0.000)	0	0.320 (0.023)	0.084 (0.025)	90	0.320 (0.025)	0.154 (0.033)	94
	abcc × *efgg*	0.322 (0.024)	0.007 (0.006)	4	0.319 (0.026)	0.072 (0.032)	78	0.309 (0.031)	0.146 (0.045)	86
	abcd × *abcd*	0.310 (0.035)	0.009 (0.008)	5	0.338 (0.039)	0.086 (0.045)	50	0.345 (0.046)	0.157 (0.054)	62
	abcd × *abcc*	0.343 (0.049)	0.012 (0.015)	7	0.352 (0.054)	0.088 (0.055)	34	0.324 (0.057)	0.158 (0.062)	38
	abco × *aaoo*	0.355 (0.060)	0.023 (0.021)	10	0.359 (0.067)	0.072 (0.063)	21	0.362 (0.067)	0.142 (0.078)	30

[a] Power is expressed as the percentage of the number of simulation trials in which significant double reduction is detected.

Numerical analyses of the simulated data are carried out according to the procedures presented above. Each simulation trial was run 100 times over which the means and standard errors of the MLEs and the power to detect double reduction were calculated.

The precision of the estimate of unknown parameters is strongly affected by marker cross types, regardless of the values for the preferential pairing factor and the frequency of double reduction (Table 2). The most precise estimates are obtained for the most informative marker type of eight different alleles, *abcd* × *efgh*, with the precision being reduced when there are identical alleles in one parent (*e.g.*, *abcd* × *efgg*) and further reduced when there are identical alleles in both parents (*e.g.*, *abcc* × *efgg*). The precision of parameter estimation is also reduced if common alleles are shared between the two parents (*e.g.*, *abcd* × *abcd* or *abcd* × *abcc*) or if dominant alleles occur in both parents and affect the phenotypes of the progeny (*e.g.*, *abco* × *aaoo*). Similar trends are observed for the power to detect significant double reduction using our method. For example, when double reduction is moderately large (0.08) or extremely large (0.15), the power of detection drops from 100% for the most informative marker to 20–40% for dominant marker type *abco* × *aaoo*. For these less informative dominant markers, it is possible to generate type I error, in which significant double reduction is occasionally detected even though no double reduction is assumed.

The effects of different degrees of preferential pairings and double reduction on parameter estimation are also examined (Table 2). Given a fixed preferential pairing factor, the precision of the estimate of the preferential pairing factor is slightly affected by changes in the frequency of double reduction, but a change in the preferential pairing factor significantly affects the precision of the estimate of the frequency of double reduction. The estimate of the frequency of double reduction is subjected to larger deviations when there is no preference than when there is a preference in chromosome pairings. For example, the standard error of the MLE of the frequency of double reduction for a marker cross type at a given frequency 0.08 is 0.063 when preferential pairings are assumed, compared to 0.050 when no preferential pairings are assumed. A similar trend is held for the power to detect double reduction.

Marker cross type, preferential pairing factor, and

the frequency of double reduction can display strong interaction effects on the precision and power of parameter estimates (Table 2). For example, at a given frequency of double reduction, the power of detecting double reduction is reduced from a more informative marker type to a less informative type, but the extent of reduction is much larger when there are preferential pairings than when there are no preferential pairings.

DISCUSSION

The major distinction between true allopolyploids and true autopolyploids is in the origin of their genomes. The genomes of the former are well differentiated, whereas all genomes of the latter are identical or very closely related (Stebbins 1950). In allopolyploids, homologous chromosomes pair at meiosis, but there is a strong pairing barrier between homeologous genomes. Therefore, the preferential pairing factor was suggested in order to describe the bivalent formation of allopolyploids (Sybenga 1988, 1994, 1995). On the other hand, autopolyploids have only homologous chromosomes that pair with equal opportunity and, therefore, multivalent formation and a resulting double reduction may occur. Between these two extremes of polyploids there exist many intermediate types, defined as a general polyploid model in this study or called segmental allopolyploids by Stebbins (1950), which combine both bivalent and multivalent pairing behaviors. Recent cytological and molecular data suggest that many traditionally recognized autopolyploids can be indeed treated as a general polyploid model (Sybenga 1996; Allendorf and Danzmann 1997; Fjellstrom *et al.* 2001).

In this article, we presented a maximum-likelihood-based statistical method for simultaneously estimating the preferential pairing factor and the frequency of double reduction using molecular markers to examine gene segregation patterns in a full-sib polyploid family. In spite of the importance of the preferential pairing factor and double reduction in describing the behavior of chromosome pairing and chromosome recombination (Darlington 1929; Mather 1936), the estimates of these two phenomena are inadequate due to the lack of a suitable analytical method. Our method proposed here can make use of all possible marker types segregating in a family, as opposed to simple dominant marker systems currently used to construct genetic maps in polyploids (Wu *et al.* 1992). In practical molecular experiments, a mixed set of marker types, including dominant (*e.g.*, random amplified polymorphic DNA or amplified fragment length polymorphism) and codominant markers (*e.g.*, restriction fragment length polymorphism or microsatellite), is often used to characterize the entire genome of outbred polyploids. Simulation studies were performed to examine the statistical properties of our method when different types of markers are used. It was suggested that there was an advantage in estimating both the preferential pairing factor and the frequency of double reduction with more informative markers than less informative markers. The estimate of the frequency of double reduction is also affected by sample sizes (Luo *et al.* 2000). However, as shown in Luo *et al.* (2000) and more comprehensively demonstrated in this study, a sample of size 100 in a tetraploid family can provide reasonable estimates for the frequency of double reduction over different types of markers.

The proposed method has three major implications. First, our method can provide more accurate information about the classification of polyploids. According to Stebbins (1950), polyploids are classified into allopolyploids and autopolyploids. But such a distinction is blurred in some cases because no precise, quantitative criterion for classification is available. Observing no quadrivalents in the natural tetraploid *Festuca mairei*, Chen *et al.* (1995) concluded that it was an allopolyploid. However, a further hybridization experiment suggested that this species might be an autotetraploid. Such a dilemma can be solved if actual estimates of the preferential pairing factor and the frequency of double reduction are available. If this species has a mixed behavior of allopolyploids and autopolyploids, the estimated preferential pairing factor should be significantly greater than zero but less than two-thirds (Equation 1). The estimated frequency of double reduction provides additional information about polyploid classification. Yet, no double reduction should not be seen as sole evidence for allopolyploid behavior because double reduction may not occur in autopolyploids when homologous chromosomes pair randomly (Bever and Felber 1992). In other species, such as *L. corniculatus* (Fjellstrom *et al.* 2001) and rainbow trout (Allendorf and Danzmann 1997), the estimates of these two parameters can eliminate the ambiguity of their classification and position them correctly.

Second, results obtained from our method can help to design an efficient linkage mapping experiment. A number of genome projects are now under way to develop molecular linkage maps of the polyploid plant genomes (Wu *et al.* 1992; Yu and Pauls 1993; da Silva *et al.* 1995; Grivet *et al.* 1996; Hackett *et al.* 1998; Meyer *et al.* 1998; Ming *et al.* 1998; Brouwer and Osborn 1999; Ripol *et al.* 1999). These maps constructed from polymorphic markers are essential for understanding the genome structure and organization of polyploids and identifying quantitative trait loci responsible for complex autopolyploid traits of economic importance. However, these maps are based on a limited number of marker types (mostly single-dose dominant markers) and their construction is conditioned on the simplified assumption of random bivalent chromosome pairings. In contrast to diploids, estimates of gene segregation and linkage in polyploids are expected to depend upon how single chromosomes pair to generate gametes at meiosis. Empirical results from cytogenetic data suggest

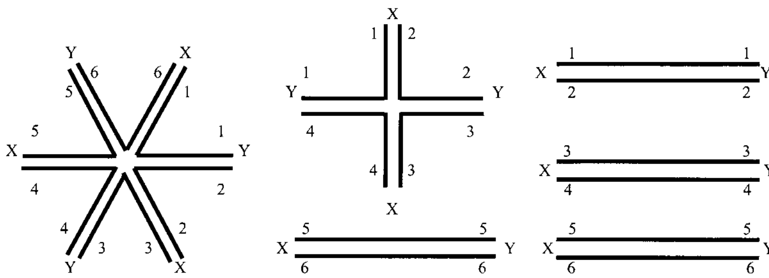


FIGURE 3.—One example of hexavalent pairing (left), quadrivalent + bivalent pairing (middle), and bivalent pairing (right) in general hexaploids; chromosomes numbered 1–6; chromosome arms X and Y. Hexavalent pairing: one pairing partner switch at the middle of the chromosomes; pairing between X1 and X6, Y1 and Y2, X2 and X3, Y3 and Y4, X4 and X5, and Y5 and Y6. Quadrivalent + bivalent pairing: the quadrivalent has one partner switch. Of the 225 possible combinations, 120 are hexavalents, 90 are quadrivalent + bivalent, and 15 are bivalents.

that the modes of chromosome pairings are not random even in those polyploids proven to be autopolyploids (Khawaja *et al.* 1995), which is thus inconsistent with the theoretical prediction based on random pairings (Jackson and Casey 1982). Chromosome pairings in general polyploids can be suggested to be a function of the homology between the genomes involved, with a propensity in pairing between homologous over homeologous chromosomes, which is defined as the preferential pairing factor (Sybenga 1988, 1994, 1995). Also, the frequency of double reduction typically occurring in multivalents affects the distribution and frequency of genotypes in autopolyploid populations (Bever and Felber 1992; Butruille and Boiteux 2000). For these reasons, the construction of genetic maps may be inaccurate without considering the effects of the preferential pairing factor and the frequency of double reduction.

Third, the estimates of the preferential pairing factor and the frequency of double reduction when extended to include multiple families are of interest to population and evolutionary genetic studies of polyploids because both the parameters affect the allele frequencies and genotype frequencies of a gene in a population (Bever and Felber 1992; Ronfort *et al.* 1998). The preferential pairing factor can provide information about the degree of homology between different genomes. Estimates of evolutionary relatedness based on meiotic pairing can shed light on the possibility of interspecific gene exchange. Meanwhile, knowledge about the occurrence and frequency of double reduction, depending on the frequency at which a locus recombines with its centromere and on the frequency of multivalent formation, provides additional insights into the population genetic structure and biological conservation of polyploids. Because genetic loci near the centromere are more protected against inbreeding, the preservation of proximal heterozygosity would be more critical than the preservation of heterozygosity at central and distal loci in polysomic polyploid species (Butruille and Boiteux 2000).

Although our method is developed for tetraploids, its extension to hexaploid, octoploid, and dexaploid species is not difficult in principle but can be much more tedious technically. For a hexaploid plant, triploid gametes are generated at meiosis. Theoretical hexaploid models based on random pairings propose three different modes for the formation of triploid gametes in autohexaploids: (1) hexavalent pairing, (2) quadrivalent + bivalent pairing, and (3) bivalent pairing, with the respective frequencies 8/15, 6/15, and 1/15 (Jackson and Casey 1982; Figure 3). But empirical data did not support such a prediction for the frequencies of chromosome pairings (Khawaja *et al.* 1995). As in the tetraploid model, each of the three modes is affected by preferential pairings, and also the first two modes undergo double reduction because of multivalent pairings at meiosis. The occurrence of preferential pairings results in a lower multivalent pairing than predicted on the basis of a random pairing (Sybenga 1995), which should be considered in model derivations. In addition, a sex-specific difference in chromosome pairing may exist in some species. For example, only disomic segregation was detected in the females of salmonid fishes, but segregation ratios in the males were best explained by a mixture of disomic and tetrasomic inheritance (Allendorf and Danzmann 1997). It is our hope that statistical methods proposed for tetraploids can stimulate further research into higher ploidy plants and more realistic situations to ultimately unravel the genetic mechanisms underlying the evolution and domestication of polyploidy.

We thank Dr. George Casella for his support on this and other studies, Dr. George Casella and Dr. Mark Yang for stimulating discussions regarding this study, and two anonymous reviewers for their constructive comments on an earlier version of this manuscript. This manuscript was approved for publication as journal series no. R-08029 by the Florida Agricultural Experiment Station.

LITERATURE CITED

Allendorf, F. W., and R. G. Danzmann, 1997 Secondary tetrasomic segregation of *MDH-B* and preferential pairing of homeologues in rainbow trout. Genetics **145:** 1083–1092.

Bever, J. D., and F. Felber, 1992 The theoretical population genetics of autopolyploidy. Oxf. Surv. Evol. Biol. **8:** 185–217.

Brouwer, D. J., and T. C. Osborn, 1999 A molecular marker linkage map of tetraploid alfalfa (*Medicago sativa* L.). Theor. Appl. Genet. **99:** 1194–1200.

Butruille, D. V., and L. S. Boiteux, 2000 Selection-mutation balance in polysomic tetraploids: impact of double reduction and gametophytic selection on the frequency and subchromosomal

localization of deleterious mutations. Proc. Natl. Acad. Sci. USA **97:** 6608–6613.

Chen, C., D. A. Sleper and C. P. West, 1995 RFLP and cytogenetic analysis of hybrids between *Festuca mairei* and *Lolium perenne*. Crop Sci. **35:** 720–725.

Danzmann, R. G., and J. P. Bogart, 1982 Evidence for a polymorphism in gametic segregation in the tetraploid treefrog *Hyla versicolor* using a glutamate oxaloacetic transaminase locus. Genetics **103:** 753–769.

Danzmann, R. G., and J. P. Bogart, 1983 Further evidence for a polymorphism in gametic segregation using a malate dehydrogenase locus in the tetraploid treefrog *Hyla versicolor*. Genetics **100:** 287–306.

Darlington, C. D., 1929 Chromosome behaviour and structural hybridity in the *Tradescantiae*. J. Genet. **21:** 207–286.

da Silva, J., M. E. Sorrells, W. L. Burnquist and S. D. Tanksley, 1995 RFLP linkage map and genome analysis of *Saccharum spontaneum*. Genome **36:** 782–791.

Dempster, A. P., N. M. Laird and D. B. Rubin, 1977 Maximum likelihood from incomplete data via EM algorithm. J. R. Stat. Soc. Ser. B **39:** 1–38.

Fisher, R. A., 1947 The theory of linkage in polysomic inheritance. Philos. Trans. R. Soc. B **233:** 55–87.

Fisher, R. A., 1949 *The Theory of Inbreeding*. Hafner, New York.

Fisher, R. A., and K. Mather, 1943 The inheritance of style length in *Lythrum salicaria*. Ann. Eugen. **7:** 265–280.

Fjellstrom, R. G., P. R. Beuselinck and J. J. Steiner, 2001 RFLP marker analysis supports tetrasomic inheritance in *Lotus corniculatus* L. Theor. Appl. Genet. **102:** 718–725.

Grant, V., 1981 *Plant Speciation*, Ed. 2. Columbia University Press, New York.

Grivet, L., A. D'Hont, D. Roques, P. Feldmann, C. Lanaud *et al.*, 1996 RFLP mapping in cultivated sugarcane (*Saccharum* spp): genome organization in a highly polyploid and aneuploid interspecific hybrid. Genetics **142:** 987–1000.

Hackett, C. A., J. E. Bradshaw, R. C. Meyer, J. W. McNicol, D. Milbourne *et al.*, 1998 Linkage analysis in tetraploid species: a simulation study. Genet. Res. **71:** 143–154.

Hauber, D. P., A. Reeves and S. M. Stack, 1999 Synapsis in a natural autotetraploid. Genome **42:** 936–949.

Haynes, K. G., and D. S. Douches, 1993 Estimation of the coefficient of double reduction in the cultivated tetraploid potato. Theor. Appl. Genet. **85:** 857–862.

Haynes, K. G., W. E. Potts and M. J. Camp, 1991 Estimation of preferential pairing in tetraploid × diploid hybridizations. Theor. Appl. Genet. **81:** 504–508.

Hickok, L. G., 1978 Homeologous chromosome pairing: frequency difference in inbred and intraspecific hybrid polyploid ferns. Science **202:** 982–984.

Hilu, K. W., 1993 Polyploidy and the evolution of domesticated plants. Am. J. Bot. **80:** 1491–1499.

Jackson, R. C., and J. Casey, 1982 Cytogenetic analysis of autopolyploids: models and methods for triploids to octoploids. Am. J. Bot. **69:** 487–501.

Jackson, R. C., and J. W. Jackson, 1996 Gene segregation in autotetraploids: prediction from meiotic configurations. Am. J. Bot. **83:** 673–678.

Khawaja, H. I. T., J. R. Ellis and J. Sybenga, 1995 Cytogenetics of natural autohexaploid *Lathyrus palustris*. Genome **38:** 827–831.

Leitch, I. J., and M. D. Bennett, 1997 Polyploidy in angiosperms. Trends Plant Sci. **2:** 470–476.

Lenz, E. M., C. F. Grane, D. A. Sleper and W. Q. Loegering, 1983 An assessment of preferential chromosome pairing at meiosis in *Dactylis glomerata*. Can. J. Genet. Cytol. **25:** 222–232.

Luo, Z. W., C. A. Hackett, J. E. Bradshaw, J. W. McNicol and D. Milbourne, 2000 Predicting parental genotypes and gene segregation for tetrasomic inheritance. Theor. Appl. Genet. **100:** 1067–1073.

Marsden, J. E., S. J. Schwager and B. May, 1987 Single locus inheritance in the tetraploid treefrog Hyla versicolor with an analysis of expected progeny ratios in tetraploid organisms. Genetics **116:** 299–311.

Masterson, J., 1994 Stomatal size in fossil plants—evidence for polyploidy in majority of angiosperms. Science **264:** 421–424.

Mather, K., 1935 Reductional and equational separation of the chromosomes in bivalents and multivalents. J. Genet. **30:** 53–78.

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

Matsubayashi, M., 1991 Phylogenetic relationships in the potato and related species, pp. 93–118 in *Chromosome Engineering in Plants: Genetics, Breeding, Evolution*, Part B, edited by T. Tsuchiya and P. K. Gupta. Elsevier, Amsterdam.

Meyer, R. C., D. Milbourne, C. A. Hackett, J. E. Bradshaw, J. W. McNichol *et al.*, 1998 Linkage analysis in tetraploid potato and association of markers with quantitative resistance to late blight (*Phytophthora infestans*). Mol. Gen. Genet. **259:** 150–160.

Ming, R., S. C. Liu, Y. R. Lin, J. da Silva, W. Wilson *et al.*, 1998 Detailed alignment of *Saccharum* and *Sorghum* chromosomes: comparative organization of closely related diploid and polyploid genomes. Genetics **150:** 1663–1682.

Muller, H. J., 1914 A new mode of segregation in Gregory's tetraploid primulas. Am. Nat. **48:** 508–512.

Ramsey, J., and D. W. Schemske, 1998 Pathways, mechanisms, and rates of polyploid formation in flowering plants. Annu. Rev. Ecol. Syst. **29:** 467–501.

Ripol, M. I., G. A. Churchill, J. A. G. da Silva and M. Sorrells, 1999 Statistical aspects of genetic mapping in autopolyploids. Gene **235:** 31–41.

Ronfort, J., E. Jenczewski, T. Bataillon and F. Rousset, 1998 Analysis of population structure in autotetraploid species. Genetics **150:** 921–930.

Soltis, D. E., and L. H. Riesberg, 1986 Autopolyploidy in *Tolmeiea menziesii* (Saxifragaceae): genetic insight from enzyme electrophoresis. Am. J. Bot. **73:** 310–318.

Soltis, D. E., and P. S. Soltis, 1993 Molecular data and the dynamic nature of polyploidy. Crit. Rev. Plant Sci. **12:** 243–273.

Soltis, D. E., and P. S. Soltis, 1999 Polyploidy: recurrent formation and genome evolution. Trends Ecol. Evol. **14:** 348–352.

Soltis, P. S., and D. E. Soltis, 2000 The role of genetic and genomic attributes in the success of polyploids. Proc. Natl. Acad. Sci. USA **97:** 7051–7057.

Stebbins, G. L., 1950 *Variation and Evolution in Plants*. Columbia University Press, New York.

Stebbins, G. L., 1971 *Chromosomal Evolution in Higher Plants*. Addison-Wesley, Reading, MA.

Sybenga, A., 1975 The quantitative analysis of chromosome pairing and chiasma formation based on the relative frequencies of MI configurations. VII. Autotetraploids. Chromosoma **50:** 211–222.

Sybenga, A., 1988 Mathematical models for estimating preferential pairing and recombination in triploid hybrids. Genome **30:** 745–757.

Sybenga, A., 1992 *Cytogenetics in Plant Breeding*. Springer-Verlag, Berlin/Heidelberg/New York.

Sybenga, A., 1994 Preferential pairing estimates from multivalent frequencies in tetraploids. Genome **37:** 1045–1055.

Sybenga, A., 1995 Meiotic pairing in autohexaploid Lathyrus: a mathematical model. Heredity **75:** 343–350.

Sybenga, A., 1996 Chromosome pairing affinity and quadrivalent formation in polyploids: Do segmental allopolyploids exist? Genome **39:** 1176–1184.

Sybenga, A., 1999 What makes homologous chromosomes find each other in meiosis? A review and a hypothesis. Chromosoma **108:** 209–219.

Tai, G. C. C., 1982a Estimation of double reduction and genetic parameters of autotetraploids. Heredity **49:** 63–70.

Tai, G. C. C., 1982b Estimation of double reduction and genetic parameters in autotetraploids based on 4x-2x and 4x-4x matings. Heredity **49:** 331–335.

Welch, J. E., 1962 Linkage in autotetraploid maize. Genetics **47:** 367–396.

Wu, K. K., W. Burnquist, M. E. Sorrells, T. L. Tew, P. H. Moore *et al.*, 1992 The detection and estimation of linkage in polyploids using single-dose restriction fragments. Theor. Appl. Genet. **83L:** 294–300.

Yu, K. F., and K. P. Pauls, 1993 Segregation of random amplified polymorphic DNA markers and strategies for molecular mapping in tetraploid alfafa. Genome **36:** 844–851.

Communicating editor: M. A. Asmussen