



A Tracing Model for the Evolutionary Equilibrium of Octoploids

Jing Wang¹, Xuemin Lv¹, Li Feng¹, Ang Dong¹, Dan Liang^{1*} and Rongling Wu^{2*}

¹National Engineering Laboratory for Tree Breeding, Key Laboratory of Genetics and Breeding in Forest Trees and Ornamental Plants, Ministry of Education, The Tree and Ornamental Plant Breeding and Biotechnology Laboratory of National Forestry and Grassland Administration, Center for Computational Biology, College of Biological Sciences and Technology, Beijing Forestry University, Beijing, China, ²Departments of Public Health Sciences and Statistics, Center for Statistical Genetics, The Pennsylvania State University, Hershey, PA, United States

Testing Hardy-Weinberg equilibrium (HWE) is a fundamental approach for inferring population diversity and evolution, but its application to octoploids containing eight chromosome sets has not well been justified. We derive a mathematical model to trace how genotype frequencies transmit from parental to offspring generations in the natural populations of autooctoploids. We find that octoploids, including autooctoploids undergoing double reduction, reach asymptotic HWE (aHWE) after 15 generations of random mating, in a contrast to diploids where one generation can assure exact equilibrium and, also, different from tetraploids that use 5 generations to reach aHWE. We develop a statistical procedure for testing aHWE in octoploids and apply it to analyze a real data set from octoploid switchgrass distributed in two ecologically different regions, demonstrating the usefulness of the test procedure. Our model provides a tool for studying the population genetic diversity of octoploids, inferring their evolutionary history, and identifying the ecological relationship of octoploid-genome structure with environmental adaptation.

Keywords: Hardy-Weinberg equilibrium, polyploid, natural population, EM algorithm, switchgrass

OPEN ACCESS

Edited by:

Simon Charles Heath,
Center for Genomic Regulation (CRG),
Spain

Reviewed by:

Diego Ortega-Del Vecchyo,
National Autonomous University of
Mexico, Mexico
Pankaj Bhardwaj,
Central University of Punjab, India

*Correspondence:

Dan Liang
liangdan@bjfu.edu.cn
Rongling Wu
rww@phs.psu.edu

Specialty section:

This article was submitted to
Statistical Genetics and Methodology,
a section of the journal
Frontiers in Genetics

Received: 14 October 2021

Accepted: 30 December 2021

Published: 28 January 2022

Citation:

Wang J, Lv X, Feng L, Dong A, Liang D
and Wu R (2022) A Tracing Model for
the Evolutionary Equilibrium
of Octoploids.
Front. Genet. 12:794907.
doi: 10.3389/fgene.2021.794907

INTRODUCTION

As an evolutionary force of the organism to buffer against environmental perturbations, the evolutionary mechanisms of polyploidy have been a long-standing subject of population and evolutionary genetic research (Bever and Felber 1992; Ramsey and Schemske 1998; Otto and Whitton 2000; Soltis et al., 2004; Fawcett et al., 2009). While a number of population genetic studies are focused on tetraploids (Haldane 1930; Moody et al., 1993; Butruille and Boiteux 2000; Arnold et al., 2012; Dufresne et al., 2014; Meirmans et al., 2018; Sun et al., 2020), there is little knowledge about population variation in polyploids at a higher ploidy level. Octoploids, whose cells have a chromosome number as many as eight times the basic haploid chromosome number, are an even less-explored group of polyploids (Edger et al., 2019). As compared to tetraploids, octoploids have wide genetic diversity that can better adapt to changing environment (Johnson and Vance-Borland, 2016; Grabowski et al., 2017). Also, octoploids may have a larger body or organ size than tetraploids, implying their more desirable opportunity to be used in breeding programs. For example, basin wildrye have larger leaves, longer culms, and greater crown circumference for octoploids than tetraploids, although the numerical ranges of plant traits and their source climates overlap between ploidy types (Johnson and Vance-Borland, 2016). By creating a ploidy series from tetraploids to octoploid for althea (*Hibiscus syriacus*), Lattier et al. (2019) found an increase in guard cell length and rDNA signals with the level of ploidy.

TABLE 1 | Gamete frequencies generated by different autooctoploid genotypes.

Genotype	Gamete				
	AAAA	AAAa	AAaa	Aaaa	aaaa
AAAAAAA	1	0	0	0	0
AAAAAAa	$9/16 + 1/64\alpha^2 + 3/16\alpha$	$3/8 - 1/16\alpha^2$ $-5/16\alpha$	$1/16 + 3/32\alpha^2 + 1/16\alpha$	$(-1/16)\alpha^2$ $+1/16\alpha$	$1/64\alpha^2$
AAAAAaa	$225/784 + 9/196\alpha^2 + 45/196\alpha$	$45/98 - 9/49\alpha^2$ $-27/98\alpha$	$87/392 + 27/98\alpha^2$ $-6/49\alpha$	$3/98 - 9/49\alpha^2$ $+15/98\alpha$	$1/784 + 9/196\alpha^2 + 3/196\alpha$
AAAAaaa	$25/196 + 225/3136\alpha^2 + 75/392\alpha$	$75/196$ $-225/784\alpha^2$ $-75/784\alpha$	$285/784$ $+675/1568\alpha^2$ $-255/784\alpha$	$45/392$ $-225/784\alpha^2 + 135/784\alpha$	$9/784$ $+225/3136\alpha^2 + 45/784\alpha$
AAAaaaa	$9/196 + 4/49\alpha^2$ $+6/49\alpha$	$12/49 - 16/49\alpha^2 + 4/49\alpha$	$41/98 + 24/49\alpha^2$ $-20/49\alpha$	$12/49 - 16/49\alpha^2$ $+4/49\alpha$	$9/196 + 4/49\alpha^2 + 6/49\alpha$
AAaaaaa	$9/784 + 225/3136\alpha^2 + 45/784\alpha$	$45/392 - 225/784\alpha^2 + 135/784\alpha$	$285/784 +$ $675/1568\alpha^2$ $-255/784\alpha$	$75/196$ $-225/784\alpha^2 - 75/784\alpha$	$25/196$ $+225/3136\alpha^2 + 75/392\alpha$
AAaaaaa	$1/784 + 9/196\alpha^2 + 3/196\alpha$	$3/98 - 9/49\alpha^2 + 15/98\alpha$	$87/392 + 27/98\alpha^2$ $-6/49\alpha$	$45/98 - 9/49\alpha^2$ $-27/98\alpha$	$225/784$ $+9/196\alpha^2$ $+45/196\alpha$
Aaaaaaa	$1/64\alpha^2$	$(-1/16)\alpha^2$ $+1/16\alpha$	$1/16 + 3/32\alpha^2$ $+1/16\alpha$	$3/8 - 1/16\alpha^2$ $-5/16\alpha$	$9/16$ $+1/64\alpha^2 + 3/16\alpha$
aaaaaaa	0	0	0	0	1

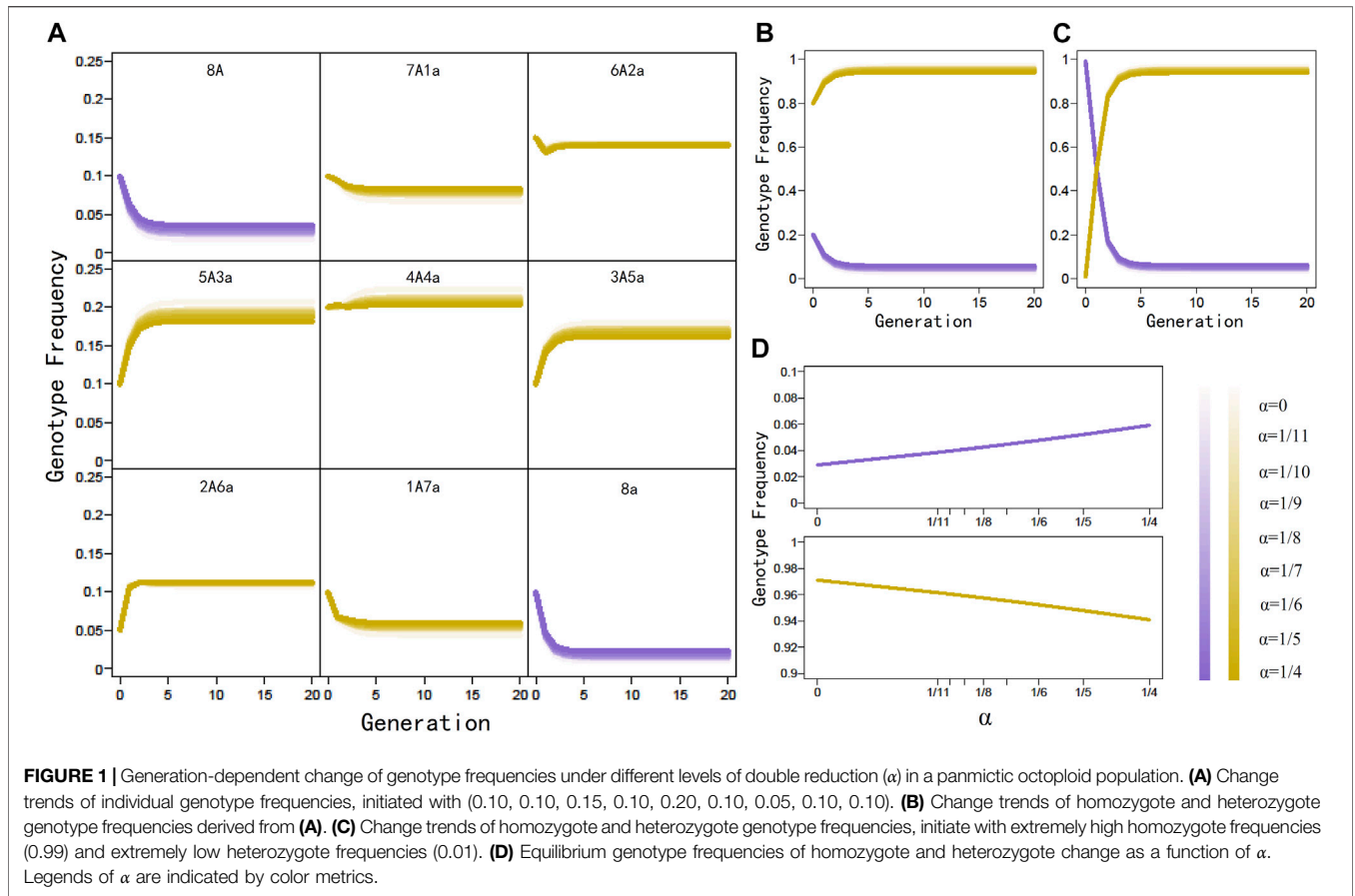
Despite their evolutionary and economic value, it is unclear how octoploids vary and evolve across time and space scales, largely owing to the lack of suitable population genetic analysis tools. Hardy-Weinberg equilibrium (HWE) test has been widely used as an approach for inferring the evolutionary processes of natural populations (Waples 2015), but this approach was established on diploids, leaving its use to polyploids a mystery. Although the diploid-driven HWE test procedure has been modified to accommodate tetraploids (Moody et al., 1993; Meirmans et al., 2018; Sun et al., 2020), the use of such a modification to evaluate population variation in inherently more complicated octoploids is not justified. For example, double reduction has little impact on the asymptotic process of HWE for autotetraploids, but we do not know if this is true for autooctoploids that produce three types of gametes, characterized by two, one, or no double reduction, respectively. Sun et al. (2020) proposed a gamete-based approach for testing autotetraploid aHWE by estimating three equilibrium diploid gamete frequencies. This approach has a power to test the equilibrium of dosage-unknown markers because three gamete frequencies just can be estimated by three observable types of genotypes. However, this approach has no sufficient degrees of freedom to test aHWE for dosage-unknown markers in autooctoploids that produce five tetraploid gametes but still have three observable types of genotypes.

In this article, we propose an approach for HWE testing in octoploids using any type of molecular markers. We derive a system of recursive equations that transmit individual octoploid genotype frequencies from the parental to offspring generation under random mating. We find that, as opposed to diploids in which exact equilibrium can be reached after one generation of random mating, both allooctoploids and autooctoploids can only gradually approach aHWE after 15 generations, different from tetraploids that use 5

generations to reach aHWE. We propose specific statistical procedures for testing aHWE using dosage-known and dosage-unknown markers in octoploids. We investigate the power of octoploid HWE detection through computer simulation. By analyzing population genetic data of allooctoploid switchgrass (Grabowski et al., 2017), we validate the usefulness of our testing procedure.

Mathematical Model

Our analysis is based on the segregation of biallelic single nucleotide polymorphisms (SNPs). At such a SNP, there are a total of nine genotypes, AAAAAAAAA (8A), AAAAAAAa (7A1a), AAAAAAaa (6A2a), AAAAAaaa (5A3a), AAAAaaaa (4A4a), AAaaaaaa (3A5a), Aaaaaaaa (2A6a), Aaaaaaaaa (1A7a), and aaaaaaaaa (8a) in an octoploid population. These genotypes will produce five types of tetraploid gametes during meiosis, AAAA, AAAa Aaaa, Aaaa, and aaaa, with different frequencies determined by Mendel's first law and the rate of double reduction (α) defined as the probability of two sister chromatids occurring in the same gamete (Darlington 1929; Mather 1935; Haynes and Douches 1993). For 8A or 8a, the same gamete type is identified although its formation results from either double reduction or non-double reduction. **Table 1** tabulates the frequencies of tetraploid gametes produced by each octoploid genotype. We use AAAAAAAAA (written as $A_1A_2A_3A_4A_5A_6A_7a$ for the identification of allele A) as an example to demonstrate how we derive these gamete frequencies. This genotype produces 8 diploids $A_1A_1, A_2A_2, A_3A_3, A_4A_4, A_5A_5, A_6A_6, A_7A_7, aa$, with the total frequency of α , through double reduction, and 28 diploids $A_1A_2, A_1A_3, A_1A_4, A_1A_5, A_1A_6, A_1A_7, A_1a, A_2A_3, A_2A_4, A_2A_5, A_2A_6, A_2A_7, A_2a, A_3A_4, A_3A_5, A_3A_6, A_3A_7, A_3a, A_4A_5, A_4A_6, A_4A_7, A_4a, A_5A_6, A_5A_7, A_5a, A_6A_7, A_6a, A_7a$, with the total frequency



of $1-\alpha$, without double reduction. Together, we have the frequencies of AA as $(1/8)\alpha \times 7 + (1/28)(1-\alpha) \times 21 = 3/4 + 1/8\alpha$, the frequencies of Aa as $(1/8)\alpha \times 0 + (1/28)(1-\alpha) \times 7 = 1/4 - 1/4\alpha$, and the frequencies of aa as $(1/8)\alpha \times 1 + (1/28)(1-\alpha) \times 0 = 1/8\alpha$. Tetraploid gametes produced by genotype $AAAAAa$ include $AAAA$, with frequency calculated as $(3/4 + 1/8\alpha)^2 = 9/16 + 1/64\alpha^2 + 3/16\alpha$, $AAAa$, with frequency as $2(3/4 + 1/8\alpha)(1/4 - 1/4\alpha) = 3/8 - 1/16\alpha^2 - 5/16\alpha$, $AAaa$, with frequency as $2(3/4 + 1/8\alpha)(1/8\alpha) + (1/4 - 1/4\alpha)^2 = 1/16 + 3/32\alpha^2 + 1/16\alpha$, and $Aaaa$, with frequencies as $2(1/4 - 1/4\alpha)(1/8\alpha) = (-1/16)\alpha^2 + 1/16\alpha$, and $aaaa$, with frequency as $(1/8\alpha)^2 = 1/64\alpha^2$. All other gamete frequencies can be derived in a similar way.

The newly-produced gametes combine randomly to generate offspring genotypes. We use $P_j(t)$ ($j = 1$ for $8A$, 2 for $7A1a$, 3 for $6A2a$, 4 for $5A3a$, 5 for $4A4a$, 6 for $3A5a$, 7 for $2A6a$, 8 for $1A7a$, and 9 for $8a$) to denote the frequencies of nine octoploid genotypes in the t th (parental) generation. Random mating of these parental genotypes produces 45 possible combination types, each of which forms different offspring genotypes, with frequencies depending on **Table 1**'s gamete frequencies and the frequencies of 45 parental combinations (**Supplementary Table S1**). We derive the mathematical expression of each genotype frequency in the $(t+1)$ th (offspring) generation derived from

those of the t th (parental) generation after random mating. Let $P_{j_1 j_2}^j(t+1)$ denote the frequency of offspring genotype j derived from parental mating type $j_1 \times j_2$ ($j_1 \leq j_2 = 1, \dots, 9$). Then, the total frequency of offspring genotype j is the sum of these offspring genotype frequencies, weighted by mating frequencies, expressed as

$$P_j(t+1) = \sum_{j_1 \leq j_2=1}^9 (P_{j_1}(t) \times P_{j_2}(t)) P_{j_1 j_2}^j(t+1) \quad (1)$$

This explicit expression of **Eq. 1** is a group of recursive equations, indicating that the frequency of genotype j in the $(t+1)$ th generation is jointly determined by the mating frequencies of relevant genotypes in the t th generation, Mendelian segregation, and double reduction rate. As can be seen from **Table S1**, $P_j(t+1)$ has a complicated but explicit relationship with $P_j(t)$.

To numerically explore how $P_j(t)$ transmits to $P_j(t+1)$, we randomly sample nine parental genotype frequencies from their space, substitute these sampled values under different double reduction rates into the expression of offspring genotype frequencies given in **Eq. 1**, use these estimated offspring genotype frequencies to calculate the subsequent offspring genotype frequencies, and repeat this process until 20 generations have passed. By plotting genotype frequencies

against generation (**Figure 1A**), we find that genotype frequencies will not reach absolutely stable values even after many generations of random mating. However, genotype frequencies will quickly approach the stationarity after 15 generations. We repeat the above procedure by sampling 1,000 sets of initial genotype frequencies, from which a similar conclusion is reached, i.e., octoploid genotype frequencies do not attach exact HWE but aHWE over generation, a similar phenomenon detected in tetraploids. Generation 15 can be regarded as one after which the octoploid population is at aHWE.

We plot the overall frequencies of genotypes 8A and 8a and those of genotypes 7A1a, 6A2a, 5A3a, 4A4a, 3A5a, 2A6a, and 1A7a over generation (**Figure 1B**), from which we find that homozygotes and heterozygotes change their frequencies differently during the approximation of aHWE. While homozygote frequency tends to decrease with generation, heterozygote frequency displays a generation-dependent increase. In an extreme example, we start the homozygote frequency of near one and the heterozygote frequency of near zero as initial values, in which the same trends are observed (**Figure 1C**). This phenomenon suggests that by approaching aHWE, octoploids are equipped to increase genetic diversity in natural populations.

We also investigate how double reduction influences the approximation of aHWE in octoploids. Under a range of double reduction rates, we find that octoploids are always close to stabilize their genotype frequencies in 15 generations, suggesting that double reduction is neutral for aHWE (**Figure 1A**). The case of no double reduction in the autopolyploid model reduces to allopolyploids if no preferential pairing is assumed. Thus, it is postulated that both autopolyploids and allopolyploids follow the same rule of aHWE. In autooctoploids, double reduction does influence the distribution of equilibrium genotype frequencies to some extent; the frequency of double reduction is positively associated with equilibrium frequencies of homozygotes but negatively associated with the equilibrium frequencies of heterozygotes (**Figure 1D**). Thus, double reduction may be regarded as a determinant of genetic diversity in autopolyploids.

STATISTICAL TESTING OF AHWE

Recursive Test

This test was proposed by Sun et al. (2020) to test aHWE in tetraploids. It is straightforward to extend it to test octoploid aHWE. As shown from recursive **Eq. 1**, an octoploid population reaches asymptotic equilibrium at approximately generation 15 (**Figure 1A**). Thus, the genotype frequencies at generation 15 are regarded as a proxy of equilibrium frequencies, denoted as \hat{P}_j for a SNP. Let N_j (totaling to N) and P_j denote the observations and frequencies of nine genotypes in the current octoploid natural population. We formulate the likelihood of genotypic observations as

$$L(N) = c + \sum_{j=1}^9 N_j \log(P_j) \quad (2)$$

where c is the constant and the maximum likelihood estimates (MLEs) of P_j can be solved as $\hat{P}_j = N_j/N$. We test whether equilibrium genotype frequencies are not different from observed genotype frequencies by using Pearson's chi-square testing approach, expressed as

$$\chi^2 = N \sum_{j=1}^9 \frac{(\hat{P}_j - \bar{P}_j)^2}{\bar{P}_j} \quad (3)$$

which is thought to follow a chi-square distribution with eight degrees of freedom. If test statistics χ^2 is larger than the critical value $\chi_{95\%,8}^2$, then we claim that the population deviates from aHWE at the SNP considered. Otherwise, the population is at equilibrium for this locus. Multiple comparisons will be needed if a set of SNPs are tested in a population genetic study.

Gamete-Based Model

Under HWE, genotype frequencies are expressed as the frequency products of the gametes that form the genotypes. Let P_{AAAA} , P_{AAAa} , P_{AAaa} , P_{Aaaa} , and P_{aaaa} denote the frequencies of five tetraploid gametes produced by octoploid genotypes in a natural population. Equilibrium genotype frequencies (Q_j) are expressed as

$$\begin{aligned} Q_{8A} &= P_{AAAA}^2 \\ Q_{7A1a} &= 2P_{AAAA}P_{AAAa} \\ Q_{6A2a} &= 2P_{AAAA}P_{AAaa} + P_{AAAa}^2 \\ Q_{5A3a} &= 2P_{AAAA}P_{Aaaa} + 2P_{AAAa}P_{AAaa} \\ Q_{4A4a} &= 2P_{AAAA}P_{aaaa} + 2P_{AAAa}P_{Aaaa} + P_{AAaa}^2 \\ Q_{3A5a} &= 2P_{AAAa}P_{aaaa} + 2P_{AAaa}P_{Aaaa} \\ Q_{2A6a} &= 2P_{AAaa}P_{aaaa} + P_{Aaaa}^2 \\ Q_{1A7a} &= 2P_{Aaaa}P_{aaaa} \\ Q_{8a} &= P_{aaaa}^2 \end{aligned} \quad (4)$$

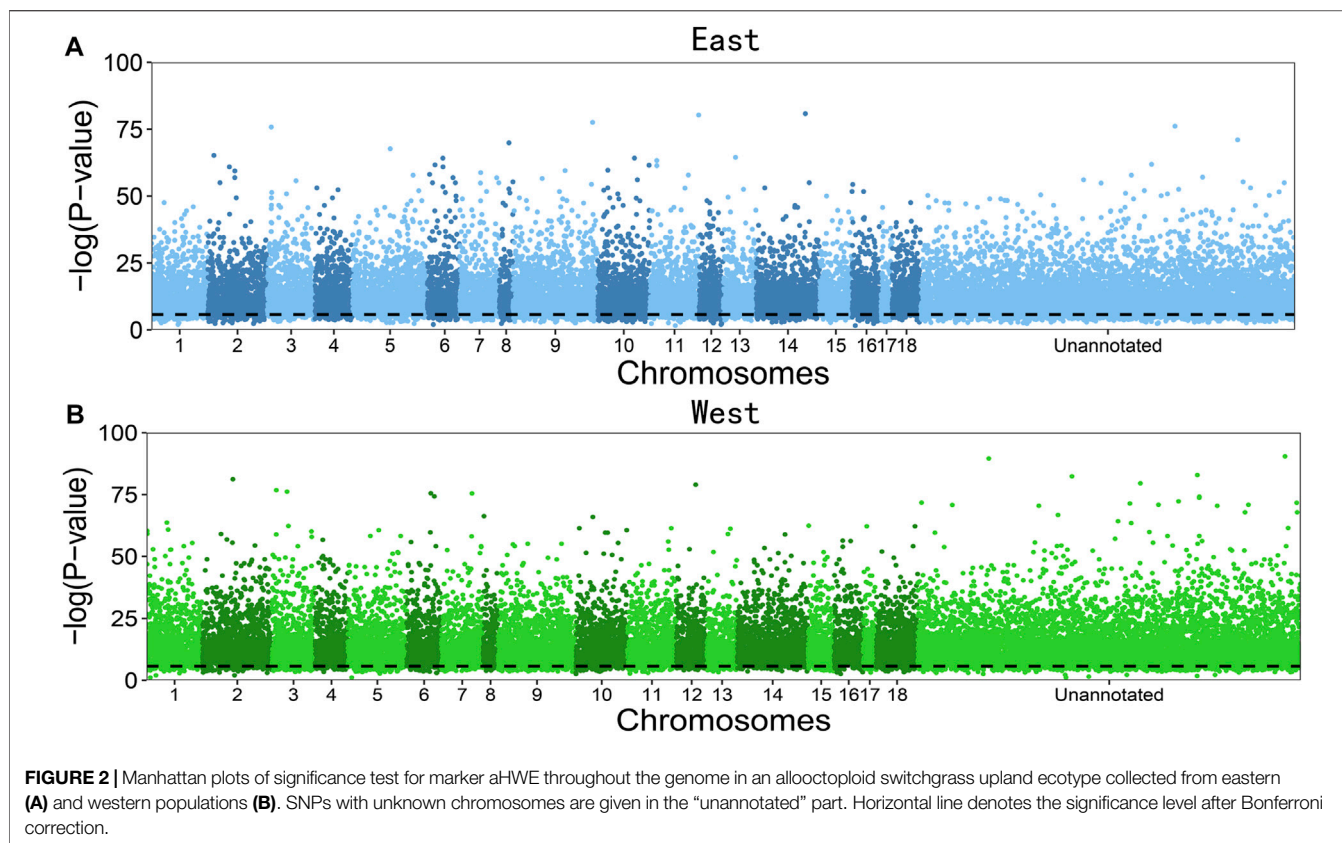
We formulate a likelihood under HWE, i.e.,

$$L_g(N) = c + \sum_{j=1}^9 N_j \log(Q_j) \quad (5)$$

which includes multiple mixture terms with component proportions determined by gamete frequencies. We implement the derivative-free EM algorithm to obtain the MLEs of gamete frequencies (\hat{Q}_j) under HWE. Plugging the \hat{Q}_j values into $L_g(N)$ (4) obtains the likelihood value under the null hypothesis that the population is at HWE. The likelihood under the alternative hypothesis, i.e., the population deviates from HWE, is calculated by **Eq. 2**. The log-likelihood ratio (LR) is then calculated as

$$LR_g = -2(L_g(N) - L(N)) \quad (6)$$

which follows a chi-square distribution with four degrees of freedom. By comparing the LR test statistics with critical value



$\chi^2_{95\%,4}$, we determine whether the SNP considered is at HWE in the octoploid population.

Allele-Based Model

Because of technical and economic reasons, many studies may genotype octoploids at a limited resolution, in which case seven heterozygotes 7A1a, 6A2a, 5A3a, 4A4a, 5A3a, 6A2a, and 7A1a cannot be distinguished from each other. Thus, for such dosage-unknown markers, there are only three distinguishable genotypes, i.e., two homozygotes each for an alternative allele and one mixed heterozygote. The frequencies and sizes of two homozygotes and the heterozygote at a dosage-unknown SNP are denoted as P_{8A} , P_{8a} , and P_- and N_{8A} , N_{8a} , and N_- , respectively. It is impossible that five gamete frequencies are estimated from three distinguishable genotypes. To make HWE testable for dosage-unknown markers, we make an assumption; i.e., alleles randomly unite to generate gametes during meiosis. Let p and q denote allele frequencies of A and a , respectively. Under HWE, we express genotype frequencies of homozygotes and heterozygote in terms of allele frequencies, i.e.,

$$\begin{aligned} Q_{8A} &= p^8 \\ Q_{--} &= 8p^7q + 28p^6q^2 + 56p^5q^3 + 70p^4q^4 + 56p^3q^5 + 28p^2q^6 + 8pq^7 \\ Q_{8a} &= q^8 \end{aligned} \quad (7)$$

Under the null hypothesis of HWE, the likelihood is formulated as

$$L_e(N) = c + N_{8A} \log(Q_{8A}) + N_- \log(Q_-) + N_{8a} \log(Q_{8a}) \quad (8)$$

in a comparison with the likelihood under the alternative hypothesis of no HWE,

$$L(N) = c + N_{8A} \log(P_{8A}) + N_- \log(P_-) + N_{8a} \log(P_{8a}) \quad (9)$$

Likelihood Eq. 8 contains a mixture term, whose solution can be made by implementing the derivative-free EM algorithm. We calculate the LR from likelihoods Eqs 8, 9 as a test statistic to test whether the dosage-unknown SNP considered is at HWE. This test statistic is thought of as being following a chi-square distribution with one degree of freedom.

RESULTS

Example

In a genome-wide association study of allopolyploid switchgrass, Grabowski et al. (2017) collected samples from the southern-adapted upland ecotype and northern-adapted upland ecotype of this species in the United States. The sampled plants include allotetraploids from both ecotypes and allooctoploids from the upland ecotype. Here, we test whether upland octoploids, distributed in the east and west

TABLE 2 | Power analysis and false positive rate of aHWE detection under different sample size.

Degree of deviation	50	100	200	400
0	0.03	0.02	0.01	0.01
20	0.37	0.58	0.94	1.00

Note: Degree of deviation describes how much genotypes frequencies deviate from equilibrium frequencies.

regions, deviate from aHWE. East and west populations include 66 and 101 samples, for which 24,859 and 23,795 quality SNPs are, respectively, available for our analysis. All markers are dosage-unknown, at each of which there are three distinguishable genotypes. Thus, the allele-based model is used to test whether they are at aHWE in the populations.

Figure 2 illustrates the significance test of deviation from aHWE for dosage-unknown SNPs distributed throughout the switchgrass genome in the east and west populations of the upland ecotype. The majority of SNPs are detected to deviate from aHWE, with the western population having a slightly larger proportion (93.3%) than the eastern population (90.7%). This result suggests that the segregating genes of allootoploid switchgrass plants, especially those from the western population, throughout the entire genome are on their way toward equilibrium. In a similar aHWE test for allotetraploids of the same species, only a small proportion of SNPs was detected to deviate from aHWE (Sun et al., 2020). This comparison suggests that while lower-ploidy switchgrass tends to be evolutionarily stable, higher-ploidy switchgrass is still experiencing a strong evolutionary change.

Power Analysis

We perform simulation studies to examine the power of HWE detection. Using given gamete frequencies for dosage-known markers and allele frequencies for dosage-unknown markers, we simulate genotype data by allowing simulated genotype frequencies to deviate from equilibrium genotype frequencies by 20%. Under this deviation, genotype data are simulated for different sample sizes ($N = 50, 100, 200, 400$). We use the gamete- and allele-based models to test HWE for dosage-known and dosage-unknown markers, respectively. This simulation and estimation process is repeated 1,000 times to empirically calculate HWE detection power, defined as the proportion of the simulations for which LR tests are significant.

We find that to detect aHWE for octoploid populations, a sample size of $n = 200$ is required, under which the power of nonequilibrium detection reaches 0.94 (**Table 2**). A sample size of $n = 100$ can only have about 0.60 power, whereas $n = 50$ fails to detect deviation from aHWE in most cases. We also analyze the false positive rate of the model by simulating the data with genotype frequencies deviating from equilibrium genotype frequencies by zero. In all cases, the model has a low type I error (**Table 2**).

DISCUSSION

It has been recognized that polyploids gradually reach HWE through random mating, but it is not very clear how many

generations they mate to approach equilibrium (Geiringer 1949; Bever and Felber 1992). Through a mathematical derivation, we showed that tetraploids never attain absolute HWE but aHWE after 5 generations of random mating (Sun et al., 2020). This phenomenon, in contrast to diploids in which exact HWE can be achieved only after one generation of random mating, suggests that polyploids are a “perpetual machine” of evolution; i.e., they can evolve even without actions of evolutionary forces, such as natural selection, mutation, genetic drift, admixture and so on.

In this article, we show that octoploids require more generations (i.e., 15) to approach aHWE. As compared to low ploidy-level polyploids, high ploidy-level polyploids have more allelic combinations and, thus, larger genetic diversity (Johnson and Vance-Borland, 2016; Grabowski et al., 2017). More generations required to reach aHWE suggest that high ploidy-level polyploids can maintain a longer-standing time of evolution than low ploidy-level polyploids. This theoretical postulation is well in agreement with empirical observations for many plants whose high-ploidy polyploids have greater diversity than low-ploidy relatives in the same habitat (Johnson and Vance-Borland, 2016). Double reduction is an important phenomenon in autopolyploids. It has been thought to affect evolutionary processes due to genetic drift (Moody et al., 1993) and gametophytic selection (Butruille and Boiteux 2000). Double reduction can change population structure and diversity by increasing the frequencies of homozygotes, but it does not strikingly accelerate the evolution of autopolyploid populations because of its subtle impact on deviation from HWE.

Based on mathematical derivations of frequency transmission, we propose three approaches for testing aHWE in octoploids using dosage-known and dosage-unknown markers. By analyzing dosage-unknown marker data of octoploid switchgrass collected from its natural distribution (Grabowski et al., 2017), we validate the usefulness of the gamete-based equilibrium-detecting model. Computer simulation is performed to determine the sample size required to detect aHWE, as guidance for designing octoploid population and evolutionary genetic studies. An increasing number of studies have begun to investigate the ecological relationship between HWE deviation and the organisms’ adaptation to heterogeneous environments (Tyukmaeva et al., 2011; Tarazona et al., 2019; Carvalho et al., 2021). Our aHWE testing model provide a generic tool to infer the genetic variation and evolutionary status of octoploid populations.

DATA AVAILABILITY STATEMENT

The data and code used in this article can be freely downloaded at <https://github.com/CCBBeijing/OctoploidDeer>.

AUTHOR CONTRIBUTIONS

JW derived the model, analyzed the data, and developed the code. XL, LF, and AD participated in model development. DL and RW supervised the study and drafted the manuscript with inputs from all other authors.

ACKNOWLEDGMENTS

We thank Paul P. Grabowski for providing his switchgrass data to validate our model and Libo Jiang for his contribution to this work. This work is supported by Beijing Forestry University Fund.

REFERENCES

- Arnold, B., Bomblies, K., and Wakeley, J. (2012). Extending Coalescent Theory to Autotetraploids. *Genetics* 192, 195–204. doi:10.1534/genetics.112.140582
- Bever, J., and Felber, F. (1992). The Theoretical Population Genetics of Autopolyploids. *Oxford Surv. Evol. Biol.* 8, 185–217.
- Butruille, D. V., and Boiteux, L. S. (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.* 97, 6608–6613. doi:10.1073/pnas.100101097
- Carvalho, C. S., Forester, B. R., Mitre, S. K., Alves, R., Imperatriz-Fonseca, V. L., Ramos, S. J., et al. (2021). Combining Genotype, Phenotype, and Environmental Data to Delineate Site-adjusted Provenance Strategies for Ecological Restoration. *Mol. Ecol. Resour.* 21 (1), 44–58. doi:10.1111/1755-0998.13191
- Darlington, C. D. (1929). Chromosome Behaviour and Structural Hybridity in the Tradescantia. *Journ. Gen.* 21, 207–286. doi:10.1007/bf02984208
- Dufresne, F., Stiff, M., Vergilino, R., and Mable, B. K. (2014). Recent Progress and Challenges in Population Genetics of Polyploid Organisms: an Overview of Current State-Of-The-Art Molecular and Statistical Tools. *Mol. Ecol.* 23, 40–69. doi:10.1111/mec.12581
- Edger, P. P., Poorten, T. J., VanBuren, R., Hardigan, M. A., Colle, M., McKain, M. R., et al. (2019). Origin and Evolution of the Octoploid Strawberry Genome. *Nat. Genet.* 51, 541–547. doi:10.1038/s41588-019-0356-4
- Fawcett, J. A., Maere, S., and Van de Peer, Y. (2009). Plants with Double Genomes Might Have Had a Better Chance to Survive the Cretaceous-Tertiary Extinction Event. *Proc. Natl. Acad. Sci.* 106, 5737–5742. doi:10.1073/pnas.0900906106
- Geiringer, H. (1949). Chromatid Segregation of Tetraploids and Hexaploids. *Genetics* 34, 665–684. doi:10.1093/genetics/34.6.665
- Grabowski, P. P., Evans, J., Daum, C., Deshpande, S., Barry, K. W., Kennedy, M., et al. (2017). Genome-wide Associations with Flowering Time in Switchgrass Using Exome-capture Sequencing Data. *New Phytol.* 213 (1), 154–169. doi:10.1111/nph.14101
- Haldane, J. B. S. (1930). Theoretical Genetics of Autopolyploids. *Journ. Gen.* 22, 359–372. doi:10.1007/bf02984197
- Haynes, K. G., and Douches, D. S. (1993). Estimation of the Coefficient of Double Reduction in the Cultivated Tetraploid Potato. *Theoret. Appl. Genet.* 85, 857–862. doi:10.1007/bf00225029
- Johnson, R. C., and Vance-Borland, K. (2016). Linking Genetic Variation in Adaptive Plant Traits to Climate in Tetraploid and Octoploid Basin Wildrye [*Leymus Cinereus* (Scribn. & Merr.) A. Love] in the Western U.S. *PLoS One* 11 (2), e0148982. doi:10.1371/journal.pone.0148982
- Lattier, J. D., Chen, H., and Contreras, R. N. (2019). Variation in Genome Size, Ploidy, Stomata, and rDNA Signals in *Althea*. *J. Amer. Soc. Hort. Sci.* 144 (2), 130–140. doi:10.21273/jashs04618-18
- Mather, K. (1935). Reductional and Equational Separation of the Chromosomes in Bivalents and Multivalents. *Journ. Genet.* 30, 53–78. doi:10.1007/bf02982205
- Meirns, P. G., Liu, S., and van Tienderen, P. H. (2018). The Analysis of Polyploid Genetic Data. *J. Hered.* 109, 283–296. doi:10.1093/jhered/esy006
- Moody, M. E., Mueller, L. D., and Soltis, D. E. (1993). Genetic Variation and Random Drift in Autotetraploid Populations. *Genetics* 134, 649–657. doi:10.1093/genetics/134.2.649
- Otto, S. P., and Whitton, J. (2000). Polyploid Incidence and Evolution. *Annu. Rev. Genet.* 34, 401–437. doi:10.1146/annurev.genet.34.1.401
- Ramsey, J., and Schemske, D. W. (1998). Pathways, Mechanisms, and Rates of Polyploid Formation in Flowering Plants. *Annu. Rev. Ecol. Syst.* 29, 467–501. doi:10.1146/annurev.ecolsys.29.1.467
- Soltis, D. E., Soltis, P. S., and Tate, J. A. (2004). Advances in the Study of Polyploidy since Plant Speciation. *New Phytol.* 161, 173–191. doi:10.1046/j.1469-8137.2003.00948.x
- Sun, L., Gan, J., Jiang, L., and Wu, R. L. (2020). Recursive Test of Hard-Weinberg Equilibrium in Tetraploids. *Trends Genet.* 37, 504. doi:10.1016/j.tig.2020.11.006
- Tarazona, E., Hahn, C., Franch-Gras, L., García-Roger, E. M., Carmona, M. J., and Gómez, A. (2019). Ecological Genomics of Adaptation to Unpredictability in Experimental Rotifer Populations. *Sci. Rep.* 9, 19646. doi:10.1038/s41598-019-56100-y
- Tyukmaeva, V. I., Salminen, T. S., Kankare, M., Knott, K. E., and Hoikkala, A. (2011). Adaptation to a Seasonally Varying Environment: a strong Latitudinal Cline in Reproductive Diapause Combined with High Gene Flow in *Drosophila montana*. *Ecol. Evol.* 1 (2), 160–168. doi:10.1002/ece3.14
- Waples, R. S. (2015). Testing for Hardy-Weinberg Proportions: Have We Lost the Plot? *J. Hered.* 106, 1–19. doi:10.1093/jhered/esu062

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2021.794907/full#supplementary-material>

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Wang, Lv, Feng, Dong, Liang and Wu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.