Bayesian Tests for Random Mating in Autopolyploids

David Gerard

Department of Mathematics and Statistics, American University, Washington, DC, 20016, USA

Abstract

Hardy-Weinberg equilibrium is often explored to evaluate the assumption of random mating. However, in autopolyploids, organisms with more than two sets of homologous chromosomes, Hardy-Weinberg equilibrium and random mating are different hypotheses that require different statistical testing approaches. Currently, the only available methods to test for random mating in autopolyploids (i) heavily rely on asymptotic approximations and (ii) assume genotypes are known, ignoring genotype uncertainty. Furthermore, these approaches are all frequentist, and so do not carry the benefits of Bayesian analysis, including ease of interpretability, incorporation of prior information, and consistency under the null. Here, we present Bayesian approaches to test for random mating, bringing the benefits of Bayesian analysis to this problem. Our Bayesian methods also (i) do not rely on asymptotic approximations, being appropriate for small sample sizes, and (ii) optionally account for genotype uncertainty via genotype likelihoods. We validate our methods in simulations, and demonstrate on two real datasets how testing for random mating is more useful than testing for equilibrium (in a natural population) and testing for Mendelian segregation (in an experimental S1 population). Our methods are implemented in Version 2.0.0 of the `hwep` R package on GitHub https://github.com/dcgerard/hwep.

1 Introduction

Researchers often test for Hardy-Weinberg equilibrium to evaluate the presence of random mating [Waples, 2015]. Since at autosomal loci in diploids one or two generations of random mating are sufficient to reach equilibrium [Gillespie, 2004], these two hypotheses are effectively equivalent, and so use the same statistical procedures. However, in autopolyploids, organisms with more than two sets of homologous chromosomes, Hardy-Weinberg equilibrium is only approached asymptotically [Haldane, 1930], and so random mating and equilibrium are separate hypotheses. To put notation to these hypotheses, let $\boldsymbol{q} = (q_0, q_1, \ldots, q_K)$ be the genotype frequencies of a K-ploid population at a biallelic locus with allele frequency $r = \frac{1}{K}\sum_{k=0}^{K} kq_k$. That is, q_k is the proportion of the population with k copies of the minor allele. Then the equilibrium hypothesis states that, in the absence of double reduction [Huang et al., 2019], the genotype frequencies follow binomial proportions,

$$q_k = \binom{K}{k} r^k (1-r)^{K-k}. \tag{1}$$

Keywords and phrases: Bayes factors, genotyping errors, Hardy-Weinberg, polyploids, random mating.

While, under the hypothesis of random mating, we have

$$q_k = \sum_{i=\max(0,k-K/2)}^{\min(k,K/2)} p_i p_{k-i}, \tag{2}$$

where $\boldsymbol{p} = (p_0, p_1, \ldots, p_{K/2})$ are the *gamete* frequencies of the population. That is, p_k is the proportion of gametes in the population that have k copies of the minor allele. Equation (2) follows since, under random mating, genotypes are the sums of independent gametes, and so $\boldsymbol{q} = \boldsymbol{p} * \boldsymbol{p}$, where "$*$" denotes discrete linear convolution, which can be written as (2). Hypotheses (1) and (2) are clearly distinct, which indicates the necessity to create separate testing procedures to evaluate them.

Testing for binomial proportions (1) is a fairly straightforward statistical problem, and several researchers have developed permutation approaches to do so effectively, even for small counts [Hardy and Vekemans, 2002, Meirmans and Van Tienderen, 2004]. By contrast, the only approaches that test for random mating (2) heavily rely on asymptotic approximations [Sun et al., 2020, Gerard, 2022]. Such approximations are particularly fraught in autopolyploids where many genotypes are possible, and sample sizes are typically small. Thus, in most applications, there will be genotypes with few/no individuals, limiting the applicability of asymptotic approaches. The current approaches also do not account for genotype uncertainty, a pervasive issue in polyploid data [Gerard et al., 2018, Gerard and Ferrão, 2019], and one that limits the real-world applicability of current methods.

Furthermore, the only approaches that test hypothesis (1) use frequentist procedures. Sun et al. [2020] developed a test for random mating in tetraploids ($K = 4$) that finds the MLEs under model (2) via an EM algorithm, and uses these estimated genotype frequencies to run a chi-squared test. Sun et al. [2020] called this a "A Gamete-Based aHWE [asymptotic Hardy-Weinberg equilibrium] Test", but this is incorrect as it is a test for random mating, not equilibrium. Sun et al. [2020] also uses an incorrect degrees of freedom for their test (4 when it should be 2). Gerard [2022] generalized this EM algorithm to every ploidy ($K \geq 4$) and implemented likelihood ratio tests for random mating for every ploidy. One issue with frequentist procedures is that, for increasing sample size, the p-value threshold should decrease to provide consistency under the null [Wakefield, 2009], but this is rarely done.

Bayesian approaches provide solutions to all of the above issues. Small sample sizes are not an issue for Bayesian methods as all inference is conditional on the observed data, not on hypothetical realizations of data which the current frequentist approaches approximate via asymptotics. Bayesian methods provide a natural foundation to integrate over genotype uncertainty. And, whereas frequentist p-values should have lower significance levels for larger sample sizes so that Type II errors decrease to zero, Bayes factors automatically calibrate themselves with increasing sample size [Shoemaker et al., 1998, Wakefield, 2009]. The appeal of the Bayesian approach can be evidenced by the large number of methods in the diploid literature that implement Bayesian tests for equilibrium [Altham, 1971, Pereira and Rogatko, 1984, Lindley, 1988, Shoemaker et al., 1998, Montoya-Delgado et al., 2001, Consonni et al., 2008, Wakefield, 2010, Consonni et al., 2011, Puig et al., 2017, 2019].

The goal of this paper to produce and implement Bayesian tests for hypotheses

$$H_0 : \boldsymbol{q} = \boldsymbol{p} * \boldsymbol{p} \text{ for some } \boldsymbol{p} \text{ s.t. } \sum_{k=0}^{K/2} p_k = 1, \text{ versus} \tag{3}$$

$$H_1 : \boldsymbol{q} \neq \boldsymbol{p} * \boldsymbol{p}. \tag{4}$$

When genotypes are known, we implement a Bayesian test via an efficient Gibbs sampler that uses a novel data augmentation scheme (Section 2.1). However, genotypes are rarely known in polyploid organisms, and so we account for genotype uncertainty using genotype likelihoods (Section 2.2). Using simulations, we explore small sample behavior for our tests in Section 3.1, and large sample behavior in Section 3.2, before exploring the performance of our procedures under genotype uncertainty in Section 3.3. We demonstrate the utility of our methods on both a natural population of tetraploid white sturgeon (Section 3.4) and an experimental population of hexaploid sweet potatoes (Section 3.5), where we show that competing methods are less useful.

2 Methods

2.1 Bayesian test with known genotypes

In this section, we will create a Bayesian test for random mating when genotypes are known. Let x_k be the number of individuals with genotype $k \in \{0, 1, \ldots, K\}$, which we collect into the vector $\boldsymbol{x} = (x_0, x_1, \ldots, x_K)$. Then, given genotype frequencies $\boldsymbol{q}$, we have the following likelihood

$$[\boldsymbol{x}|\boldsymbol{q}] \sim \text{Multinom}(n, \boldsymbol{q}), \tag{5}$$

where $\text{Multinom}(n, \boldsymbol{q})$ denotes the multinomial distribution with size $n = \sum_{k=0}^{K} x_k$ and probability vector $\boldsymbol{q}$. Our goal is to use likelihood (5) to test hypotheses (3)–(4). We will do so by calculating a Bayes factor [Kass and Raftery, 1995], the ratio of the marginal likelihoods under the null and alternative,

$$\text{BF} = \frac{\Pr(\boldsymbol{x}|H_0)}{\Pr(\boldsymbol{x}|H_1)} = \frac{\int \text{Multinom}(\boldsymbol{x}|n, \boldsymbol{p} * \boldsymbol{p})\pi_0(\boldsymbol{p})\mathrm{d}\boldsymbol{p}}{\int \text{Multinom}(\boldsymbol{x}|n, \boldsymbol{q})\pi_1(\boldsymbol{q})\mathrm{d}\boldsymbol{q}}, \tag{6}$$

where $\pi_0(\cdot)$ and $\pi_1(\cdot)$ denote prior densities under the null and alternative, respectively. In this manuscript, we will assume that these prior densities are both Dirichlet (Section 2.3) with concentration hyperparameters $\boldsymbol{\alpha}$ under the null and $\boldsymbol{\beta}$ under the alternative,

$$\boldsymbol{p} \sim \text{Dirichlet}(\boldsymbol{\alpha}), \text{ and} \tag{7}$$

$$\boldsymbol{q} \sim \text{Dirichlet}(\boldsymbol{\beta}). \tag{8}$$

Smaller Bayes factors provide more evidence for H_1, and larger Bayes factors provide more evidence for H_0.

To calculate the Bayes factor (6), we will need the marginal likelihoods under both the null and the alternative. The marginal likelihood under the alternative is straightforward to calculate. Using prior (8) and likelihood (5), we have that $\boldsymbol{x}$ is Dirichlet-Multinomial [Mosimann, 1962], whose probability mass function (PMF) we will denote by

$$Pr(\boldsymbol{x}|H_1) = \text{DM}(\boldsymbol{x}|n, \boldsymbol{\beta}) = \frac{\Gamma\left(\sum_{k=0}^{K} \beta_k\right)\Gamma(n+1)}{\Gamma\left(n + \sum_{k=0}^{K} \beta_k\right)} \prod_{k=0}^{K} \frac{\Gamma(x_k + \beta_k)}{\Gamma(\beta_k)\Gamma(x_k + 1)}, \tag{9}$$

where $\Gamma(\cdot)$ is the gamma function.

To calculate the marginal likelihood under the null, we will need the following data augmentation scheme to represent the multinomial likelihood with probability vector $\boldsymbol{p} * \boldsymbol{p}$. Each individual in our sample corresponds to two parental gametes, and we let $\boldsymbol{A} = (a_{ij})$, with $j \geq i$, denote the number of parental gamete pairs in our sample with genotypes $0 \leq i \leq j \leq K/2$. For example, if an individual's parents provided gametes with genotypes 2 and 0, then this would contribute one count to a_{02}. Let $\boldsymbol{y} = (y_0, y_1, \ldots, y_{K/2})$, where y_k denotes the number of gametes in our sample with dosage k. Then we have the relationships

$$x_k = \sum_{i \text{ s.t. } i \leq k-i} a_{i,k-i}, \text{ and} \tag{10}$$

$$y_k = \sum_{i \text{ s.t. } i < k} a_{ik} + 2a_{kk} + \sum_{j \text{ s.t. } j > k} a_{kj}. \tag{11}$$

See Figure 1 for a visual description of these relationships. We can then represent the likelihood under the null in the following form

$$f(\boldsymbol{x}|n, \boldsymbol{p}) = \sum_{\boldsymbol{A}} \sum_{\boldsymbol{y}} f(\boldsymbol{x}|\boldsymbol{A}) f(\boldsymbol{A}|\boldsymbol{y}) f(\boldsymbol{y}|\boldsymbol{p}), \tag{12}$$

where

$$f(\boldsymbol{x}|\boldsymbol{A}) = 1\left(x_k = \sum_{i \text{ s.t. } i \leq k-i} a_{i,k-i} \right), \tag{13}$$

$$f(\boldsymbol{A}|\boldsymbol{y}) = \frac{n! \prod_{k=0}^{K/2} y_k!}{(2n)! \prod_{j \geq i} a_{ij}!} 2^{\sum_{j>i} a_{ij}}, \text{ and} \tag{14}$$

$$f(\boldsymbol{y}|\boldsymbol{p}) = \text{Multinom}(\boldsymbol{y}|2n, \boldsymbol{p}), \tag{15}$$

where $1(\cdot)$ is the indicator function. To derive (13)–(15), we can consider the following generative process when there is random mating. First, in (15), $2n$ gamete dosages are randomly drawn from the population using the gamete frequencies $\boldsymbol{p}$. Then, these gametes are randomly paired up, with the distribution of the number of each pairing given a fixed number of categories described in Levene [1949] (their equation (5)), and this results in (14). Finally, for (13) we merely calculate $\boldsymbol{x}$ from $\boldsymbol{A}$ using (10).

To calculate the marginal likelihood under the null, let $\mathcal{T}(\boldsymbol{x})$ be the set of non-negative integer matrices $\boldsymbol{A}$ such that (10) holds, and let $\boldsymbol{y}(\boldsymbol{A})$ denote the function (11) that maps $\boldsymbol{A}$ to $\boldsymbol{y}$. Then we have, using prior (7),

$$\Pr(\boldsymbol{x}|H_0) = \int_{\boldsymbol{p}} f(\boldsymbol{x}|n, \boldsymbol{p}) \, \text{Dirichlet}(\boldsymbol{p}|\boldsymbol{\alpha}) \tag{16}$$

$$= \int_{\boldsymbol{p}} \sum_{\boldsymbol{A}} \sum_{\boldsymbol{y}} f(\boldsymbol{x}|\boldsymbol{A}) f(\boldsymbol{A}|\boldsymbol{y}) f(\boldsymbol{y}|\boldsymbol{p}) \, \text{Dirichlet}(\boldsymbol{p}|\boldsymbol{\alpha}) \tag{17}$$

$$= \sum_{\boldsymbol{A}} \sum_{\boldsymbol{y}} f(\boldsymbol{x}|\boldsymbol{A}) f(\boldsymbol{A}|\boldsymbol{y}) \, \text{DM}(\boldsymbol{y}|\boldsymbol{\alpha}) \tag{18}$$

$$= \sum_{\boldsymbol{A} \in \mathcal{T}(\boldsymbol{x})} \sum_{\boldsymbol{y}} f(\boldsymbol{A}|\boldsymbol{y}) \, \mathrm{DM}(\boldsymbol{y}|\boldsymbol{\alpha}) \tag{19}$$

$$= \sum_{\boldsymbol{A} \in \mathcal{T}(\boldsymbol{x})} f(\boldsymbol{A}|\boldsymbol{y}(\boldsymbol{A})) \, \mathrm{DM}(\boldsymbol{y}(\boldsymbol{A})|\boldsymbol{\alpha}). \tag{20}$$

Line (20) follows from the previous line because $f(\boldsymbol{A}|\boldsymbol{y})$ is 0 wherever (11) does not hold.

The summation (20) turns out to be computationally feasible to calculate for tetraploids ($K = 4$) and for small sample sizes of hexaploids ($K = 6$ and $n \lessapprox 100$). We provide details on how to sum over $\mathcal{T}(\boldsymbol{x})$ in Appendix S1. However, for larger ploidies, we resort to simulation approaches to obtain the marginal likelihood under the null. We created a Gibbs sampler [Gelfand and Smith, 1990], via the same data augmentation scheme (13)–(15), to sample from the posterior under the null. We then used the method of Chib [1995] to estimate the marginal likelihood from this Gibbs sampler. The Gibbs sampler is provided in Procedure 2, and its derivation is provided in Appendix S2.

To summarize the results of this section, we obtain the marginal likelihood under the alternative, $P(\boldsymbol{x}|H_1)$, via (9). For small ploidies, we obtain the marginal likelihood under the null, $P(\boldsymbol{x}|H_0)$, via (20). For larger ploidies, we estimate the marginal likelihood under the null via Procedure 2. After obtaining these marginal likelihoods, we obtain the Bayes factor (6).

2.2 Bayesian test using genotype likelihoods

Due to the higher number of possible dosages, polyploids are more susceptible than diploids to data variability and data biases, and thus exhibit higher amounts of genotype uncertainty [Gerard et al., 2018, Gerard and Ferrão, 2019, Gerard, 2021a,b]. A common way to account for this is through the use of genotype likelihoods [Li, 2011]. A nice feature of Bayesian methodology is its ability to easily adapt to additional variability by integrating over it. Here, we develop Bayesian testing procedures for random mating that account for genotype uncertainty using genotype likelihoods.

It is possible to modify our Gibbs sampler from Section 2.1 to account for genotype likelihoods via a data augmentation scheme where we introduce latent genotypes that we sample over (Appendix S3). This approach works well for small ploidies and small sample sizes. However, we have found that this approach produces poor mixing when both the ploidy K and the sample size n is large (Figures S6–S8), and thus requires a large number of samples to produce accurate results. We have therefore opted to use the Stan programming language [Stan Development Team, 2022a,b], which uses Hamiltonian Monte Carlo, to more efficiently explore the posterior distribution in fewer samples. Marginal likelihoods and Bayes factors may then be calculated via bridge sampling [Meng and Wong, 1996, Gronau et al., 2020].

Let us discuss how we implemented the model in Stan. Let ℓ_{ik} be the genotype likelihood for individual $i = 1, 2, \ldots, n$ for genotype $k = 0, 1, \ldots, K$. That is, ℓ_{ik} is the probability of the data (sequencing, microarray, or otherwise) for individual i given that the genotype for that individual is k. Genotype likelihoods are provided by many genotyping programs [McKenna et al., 2010, Voorrips et al., 2011, Serang et al., 2012, Gerard et al., 2018, Clark et al., 2019, Gerard and Ferrão, 2019]. Then, given these genotype likelihoods, we have the likelihood for these data is

$$f(\boldsymbol{L}|\boldsymbol{q}) = \prod_{i=1}^{n} \sum_{k=0}^{K} \ell_{ik} q_k. \tag{21}$$

Under the alternative, the prior over $\boldsymbol{q}$ is (8), while under the null, the prior over $\boldsymbol{q} = \boldsymbol{p} * \boldsymbol{p}$ is induced from (7). We run two Markov chains, one using prior (8) and one using prior (7), estimating the marginal likelihoods of each via bridge sampling, and from there calculating the Bayes factor. These implementations in Stan are relatively straightforward so we omit the specific details. The Stan files may be found in the source code for our `hwep` R package (`https://github.com/dcgerard/hwep`).

2.3 Prior elicitation

Here, we describe our default priors. For convenience, as in Sections 2.1 and 2.2, we assume that the gamete frequencies are *a priori* Dirichlet($\boldsymbol{\alpha}$) under the null, and the genotype frequencies are *a priori* Dirichlet($\boldsymbol{\beta}$) under the alternative, and our focus will be on selecting appropriate values of $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$.

To motivate our choice for $\boldsymbol{\alpha}$, we note that at equilibrium, in the absence of double reduction, we have $\boldsymbol{y} \sim \text{Binom}(K/2, r)$, for some allele frequency r [Haldane, 1930]. That is,

$$p_i = \binom{K/2}{i} r^i (1-r)^{K/2-i}. \tag{22}$$

For natural populations, it is reasonable to place a uniform prior over r. Matching moments with the Dirichlet, we set α_i to be proportional to the expected value of (22) when r is uniformly distributed. We have that

$$\alpha_i \propto \int_0^1 \binom{K/2}{i} r^i (1-r)^{K/2-i} \mathrm{d}r = \frac{1}{K+1}, \tag{23}$$

where the calculation in (23) uses the kernel of the Beta density. Thus, each α_k should have the same expected value *a priori*, and this motivates us to take

$$\boldsymbol{\alpha} = \mathbf{1}_{K/2+1}. \tag{24}$$

That is, $\boldsymbol{p}$ is uniformly distributed over the $K/2$ unit simplex. This might appear to be just the naive prior one would choose as "uninformative", but we have also motivated it by placing an uninformative prior on the allele frequency of a randomly mating natural population at equilibrium.

We choose our prior for $\boldsymbol{\beta}$ conditional on our choice of $\boldsymbol{\alpha}$. Our elicitation of $\boldsymbol{\beta}$ makes the following key assumption: *samples of size one are completely uninformative on the status of random mating.* That is, let $\boldsymbol{e}_k$ be a 1-of-$(K+1)$ vector that is 1 at position k and is 0 elsewhere, then our assumption is that

$$\text{BF}(\boldsymbol{\beta}) = \frac{\Pr(\boldsymbol{e}_k | H_0)}{\Pr(\boldsymbol{e}_k | H_1, \boldsymbol{\beta})} = 1 \text{ for all } k = 0, 1, \ldots, K. \tag{25}$$

We can show that, given $\boldsymbol{\alpha} = \mathbf{1}_{K/2+1}$, for any choice of scaling factor $c > 0$, the following choice of $\boldsymbol{\beta}$ will result in property (25),

$$\beta_k = c \left[\min\left(\left\lfloor \frac{k}{2} \right\rfloor, \left\lfloor \frac{K-k}{2} \right\rfloor \right) + 1 \right]. \tag{26}$$

Equation (26) is merely a constant c multiplied by the number of elements in $\boldsymbol{A}$ that correspond to each x_k (Lemma S1). We prove this result in Corollary S1 in Appendix S4.

By default, we choose a value of c such that the total sum of $\boldsymbol{\beta}$ is the same as $\mathbf{1}_{K+1}$. This intuitively means that we place the same overall weight on the prior as a uniform distribution would. When $\boldsymbol{\alpha} = \mathbf{1}_{K/2+1}$, this results in default values of $\boldsymbol{\beta}$,

$$\beta_k = \frac{(K+1)\left(\min[\lfloor k/2 \rfloor, \lfloor (K-k)/2 \rfloor] + 1\right)}{\binom{K/2+2}{K/2}}. \tag{27}$$

The default values of $\boldsymbol{\beta}$ for even ploidies less than or equal to 12 are presented in Table S1.

3 Results

3.1 Small sample behavior

We can gain a sense of how our Bayes procedure works on small samples by applying it to all possible values of $\boldsymbol{x}$ when n is small enough for this to be feasible.

We applied the Bayes procedure from Section 2.1 to all $(K+1)$-tuples ($K \in \{4, 6, 8\}$) that sum to $n \in \{5, 10\}$. This corresponds to $\binom{K+n}{K}$ tuples for a given ploidy and sample size. The histogram of log Bayes factors is presented in Figure S1. We see there that for samples of size 5 the log Bayes factors are almost all between -2 and 2, and for samples of size 10 the log Bayes factors are almost all between -7 and 3. The ploidy of the species does not considerably change the distribution of log Bayes factors. Using the suggested thresholds of Bayes factors from Kass and Raftery [1995], this would indicate that it is possible to obtain "very strong" evidence against the null hypothesis at samples as small as size 10, but not at size 5.

In Table S2, we list the values of $\boldsymbol{x}$ that produce the largest and smallest Bayes factors. We see there that the largest Bayes factors tend to result from very unimodal data, whereas the smallest Bayes factors tend to result from data with a couple prominent modes. This is verified in Figure S2 where we see that monotone $\boldsymbol{x}$ tends to result in positive log Bayes factors, while non-monotone $\boldsymbol{x}$ tends to result in negative log Bayes factors. Most, but not all, of the information in the consistency with a random mating hypothesis appears to be in the form of unimodality.

3.2 Large sample behavior

We can verify that the Bayes factors increase under the null, and decrease under the alternative, as the sample size increases. To do this, we simulated genotype counts for ploidies $K \in \{4, 6, 8\}$ with $n \in \{10, 100, 1000\}$ individuals under two scenarios. The first scenario was where the null was satisfied with $\boldsymbol{p} = \mathbf{1}_{K/2+1}/(K/2+1)$ and $\boldsymbol{q} = \boldsymbol{p} * \boldsymbol{p}$, and the second scenario was where the alternative was satisfied with $\boldsymbol{q} = \mathbf{1}_{K+1}/(K+1)$. For each scenario, we generated genotype counts via $\boldsymbol{x} \sim \text{Multinom}(n, \boldsymbol{q})$. Each replication, we obtained the Bayes factor from Section 2.1 and the p-value from Gerard [2022]. For each combination of parameters, we replicated these simulations 1000 times.

The results are in Figure 2. There, we see that the Bayes factor is consistent under both the alternative and null scenarios. We also plot the computation time in Figure S3 under these scenarios. We see that, when using simulation, the Bayesian approach has about the same computation time for any ploidy and sample size. The Bayesian test is even slightly faster than the likelihood ratio test from Gerard [2022], but both are still fast enough for genome-wide applications. Perhaps the biggest disadvantage of the likelihood ratio test is that the p-values, under the null, are not uniformly

distributed for small sample sizes (Figure 3) because of its reliance on asymptotics. The Bayesian approach does not depend on asymptotics and so may be used at any sample size.

3.3 Performance under genotype uncertainty

To explore the performance of our approaches under genotype uncertainty, we incorporated genotype likelihoods into the simulation study from Section 3.2. For each replication, using the same simulation settings as in Section 3.2, we simulated read counts using `rflexdog()` from the `updog` R package [Gerard et al., 2018, Gerard and Ferrão, 2019] at a read depth of 10, a sequencing error rate of 0.01, an overdispersion parameter of 0.01, and no allelic bias. From these read counts, we then used `updog` to generate genotype likelihoods and genotype posteriors. We used these genotype likelihoods to run the Bayesian test from Section 2.2. To compare this approach to that of Section 2.1, we need an estimate of $\boldsymbol{x}$. We did this by summing up the posterior probabilities for each genotype across individuals and then rounding these sums to the nearest integers.

The results are presented in Figure S4. We see there that only the genotype likelihood approach from Section 2.2 is consistent, with the Bayes factor increasing with the sample size when the null is satisfied. Assuming known genotypes, the method of Section 2.1 appears to be anti-conservative in the presence of genotype uncertainty, generally providing Bayes factors that are too small. This indicates the need to account for genotype uncertainty in this testing situation to provide accurately calibrated test statistics.

We also evaluated the Gibbs sampler from Appendix S3 in these simulations. The number of iterations in the Gibbs sampler was chosen to have the same run-time as the Stan approach when $K = 8$ and $n = 1000$ (Figure S5). The Bayes factors calculated using the Gibbs sampler and those calculated using Stan are mostly the same except when both the sample size and ploidy are large (Figures S6–S7). The differences are due to poor mixing in the Gibbs sampler, as this approach is not consistent for $K = 8$ (Figure S8), at least given the number of sampling iterations used. This demonstrates the benefits of using Stan, at least for large sample sizes and large ploidies, to improve mixing.

3.4 A natural population

In this section, we analyze the tetraploid white sturgeon (*Acipenser transmontanus*) data from Delomas et al. [2021]. These data were collected from a central California caviar farm with broodstock from the Sacramento river, and so likely exhibit random mating. We genotyped the $n = 19$ individuals at 325 loci using `updog` [Gerard et al., 2018, Gerard and Ferrão, 2019]. These data were sequenced to a very high average read depth of 5342, and so it is reasonable to take genotypes at most loci as known. We ran our Bayesian test (Section 2.1) and the likelihood ratio test from Gerard [2022] on these loci.

Figure 4 contains a histogram of the log Bayes factors. We see there that the mode of the log Bayes factors is above 0, with most log Bayes factors near or slightly above zero. This provides strong evidence that the population is undergoing random mating. There are just three loci that appear to deviate from this trend (Table S3).

The test for random mating immediately reveals its use as, looking deeper at the three loci with the smallest Bayes factors, it appears that all three are a result of genotyping errors. Let us take, as an example, the SNP with the lowest Bayes factor, Atr_20529-52. When we fit `updog` on these data, there are two modes of the log-likelihood (Figure 5). The maximum likelihood estimate results in

$\boldsymbol{x} = (0, 6, 2, 1, 10)$ (log-likelihood of -90) while a smaller local mode results in $\boldsymbol{x} = (0, 0, 6, 3, 10)$ (log-likelihood of -105). This local mode estimates the allelic bias (the ratio of mapping probabilities of the minor and major alleles) to be 2.2, which is large but not too out of the ordinary for these data (Figure S12). Interestingly, when we run our Bayesian test for random mating on this local mode with $\boldsymbol{x} = (0, 0, 6, 3, 10)$, we get a log Bayes factor of 1.1, which is in alignment with the rest of these data (Figure 4). Thus, it is likely that this SNP was a genotyping error. When we do the same analysis with the SNPs with the next two smallest Bayes factors, Atr_33123-54 and Atr_26310-63, we also see that the global modes are likely genotyping errors and local modes likely provide the correct genotype estimates (Figures S9–S10). This demonstrates the practical usefulness of our approach.

The likelihood ratio test from Gerard [2022] produces log p-values that are linearly associated with the log Bayes factors, particularly in the tail (Figure S13). However, standard frequentist approaches would indicate that no loci exhibit strong deviations from random mating. For example, the smallest three q-values [Storey, 2003] are 0.06, 0.17, and 0.17, which many researchers would not raise an eyebrow to since they are all above 0.05.

Tests for equilibrium are less useful for these data. The p-values, using the U-statistic based test from Gerard [2022], are highly non-uniform (Figure S11). Indeed, using these p-values, the `qvalue` package [Storey, 2003] estimates that 53% of loci do not satisfy the equilibrium hypothesis, with 14 loci being rejected at the 0.05 FDR level. This indicates that, though these data exhibit random mating, there are some severe deviations from equilibrium, likely caused by unobserved relatedness between individuals. This limits the usefulness of this approach for detecting genotyping errors.

3.5 An S1 population

Here, we analyze the S1 population of 142 autohexaploid sweet potatoes (*Ipomoea batatas*) from Shirasawa et al. [2017]. An S1 population consists of a single generation of selfing from the same parent. It might seem strange, but random mating is fulfilled in S1 populations because (i) all offspring gametes are drawn *iid* from the distribution of gametes produced by the parent, and (ii) these gametes are randomly paired during fertilization.

Data were prepared and filtered as in Gerard [2022]. That is, SNPs were selected that had allele frequencies between 0.1 and 0.9, average read depths of at least 100, and no more than 50% missing data. This resulted in 5074 SNPs. Genotype likelihoods were obtained by `updog` [Gerard et al., 2018, Gerard and Ferrão, 2019]. We used these genotype likelihoods to fit the Bayesian test from Section 2.2. The numbers of individuals with each dosage were estimated by summing posterior posterior genotype probabilities across individuals and then rounding these sums. We used these estimated counts to fit the Bayesian test from Section 2.1 on these loci.

The Bayes factors calculated using either genotype likelihoods (Section 2.2) or assuming known genotypes (Section 2.1) are only moderately correlated (Figure S14), indicating that there is enough genotype uncertainty that we should only use the genotype likelihood approach for these data (Section 3.3). The log Bayes factors for these loci are almost all positive (Figure S15) indicating strong evidence of random mating, as it should for these data since random mating is indeed fulfilled.

We explored the 5 SNPs with log Bayes factors less than -5. Genotype plots for the `updog` fits of these SNPs are provided in Figure S16. `Updog` has its own measure of genotyping performance, the estimated proportion of individuals mis-genotyped, and yet all but one of these SNPs have reasonable values of this measure (Table S4), so would likely pass filters using it. The numbers of individuals with each genotype deviate strongly from that expected under Mendelian segregation

(Table S4), indicating that these are, indeed, problematic SNPs.

Since this is an S1 population, it is possible to use this knowledge to specifically test for genotype frequencies that would result from Mendelian segregation. We could then compare this approach with our more general test for random mating. We created a Bayesian test for S1 proportions in Appendix S5 and applied this test to the sweet potato data. There was significant agreement on which SNPs satisfy the random mating assumption (Table S5). For SNPs that support the random mating hypothesis, the S1 test tends to produce larger Bayes factors, indicating greater discriminative ability, as expected (Figure S17). However, the S1 test provides many more SNPs that support the alternative hypothesis. Genotype plots of a random sample of these SNPs, where the S1 test indicates violations in the null and the random mating test indicates no violations, are presented in Figure S18. These SNPs appear to be relatively well behaved, except for two or three individuals in each SNP, where Mendelian segregation would predict those genotypes to be impossible. Indeed, those problematic individuals are the reasons why the S1 test is so keen on the alternative hypothesis, and if we remove them then the S1 test produces positive log Bayes factors (Table S6). If the goal of a test is to find SNPs where *any* individual might be mis-genotyped, then the S1 test might be appropriate. But it is very sensitive to minor violations in the genotype frequencies predicted by an S1 population, and so might be too finicky for applied work. By contrast, the random mating test only detects broad genotyping issues. We note, though, that these problematic individuals could possibly be a result of double reduction [Stift et al., 2010], in which case a more sophisticated method to test for S1 genotype frequencies in the presence of double reduction might produce better results.

4 Discussion

In this manuscript, we developed Bayesian tests for random mating using either known genotypes or genotype likelihoods. When there is non-ignorable genotype uncertainty, the genotype likelihood approach must be used to be consistent in the sample size. We demonstrated our methods in both simulations and on real data, where we were able to detect problematic SNPs with genotyping errors missed by other methods.

In the diploid literature, it is common to test for Hardy-Weinberg equilibrium to check for genotyping errors [Hosking et al., 2004]. We would argue that in autopolyploids, tests for random mating are more useful for this purpose than tests for equilibrium. One reason is that equilibrium is only approached asymptotically [Haldane, 1930] and is therefore a much more stringent criterion than random mating, and so less-likely to be satisfied in the real world. This can be seen in both of our real data examples (Sections 3.4 and 3.5). In both cases, equilibrium was not fulfilled, and so tests for equilibrium are useless for detecting genotyping issues. But random mating was fulfilled, and we successfully used our tests for random mating to detect loci with gross genotyping errors.

Most researchers are not used to the Bayesian approach to hypothesis testing, and so a common question is how to choose the threshold of the Bayes factor before one concludes violations in the random mating hypothesis? The canonical solution [Section 4.4.3 of Berger, 2013, e.g.] is for a researcher to elicit prior probabilities π_0 and $(1 - \pi_0)$ for the null and alternative hypotheses, respectively, and losses L_I and L_{II} for type I and type II errors, respectively. One would then reject H_0 if the Bayes factor is less than $(1 - \pi_0)L_{II}/(\pi_0 L_I)$. Eliciting these quantities can be difficult for most researchers. So our recommendations would be to (i) plot the Bayes factors and see if any loci are anomalous, (ii) in settings with few loci (~ 100) use the thresholds from Kass and Raftery

[1995] as a guide, and (iii) in GWAS settings with many loci use the threshold from Wakefield [2010] as a guide. For (ii) this would correspond to a log Bayes factor being less than -3 indicating strong evidence, and a log Bayes factor less than -5 indicating very strong evidence for violations of random mating. For (iii) this would correspond to a log Bayes factor less than -16 indicating very strong evidence for violations of random mating.

Acknowledgments

This material is based upon work supported by the National Science Foundation under Grant No. 2132247.

Most analyses were performed using the R statistical language [R Core Team, 2022].

Data availability

All methods discussed in this manuscript are implemented Version 2.0.0 of the `hwep` R package on GitHub:
https://github.com/dcgerard/hwep.
Scripts to reproduce all of the results of this manuscript are available on GitHub:
https://github.com/dcgerard/rmbayes_sims.
All datasets used in this manuscript are open and available at:
https://www.doi.org/10.5061/dryad.crjdfn33r, and
http://sweetpotato-garden.kazusa.or.jp/

Supplementary material

Additional details, figures, and tables are available in the Supplementary Material online.

5 Tables, figures, and procedures

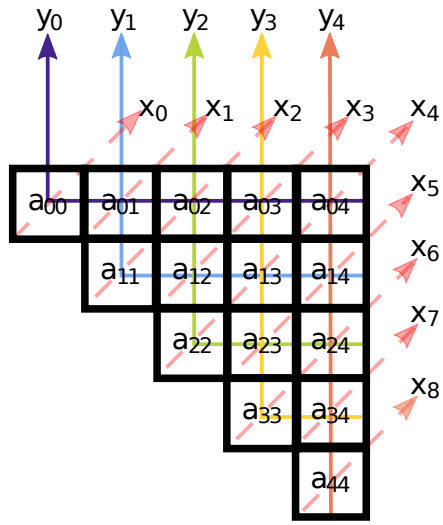

Figure 1: Visualization of data augmentation from (10) and (11) for an octoploid species ($K = 8$). The elements of $\boldsymbol{A}$ are in each cell. The x_k's are the sums of the anti-diagonal elements, indicated by the dashed red lines. The y_k's are the sums of the color-coded solid lines. We sample $[\boldsymbol{A}|\boldsymbol{x}, \boldsymbol{p}]$ by sampling each anti-diagonal by a multinomial with size x_k and probabilities proportional to the cell probabilities of the a_{ij}'s (S13). The sampled $\boldsymbol{y}$ is then calculated from this $\boldsymbol{A}$.

Procedure 1 Sampling from the full conditional of $[\boldsymbol{y}|\boldsymbol{x}, \boldsymbol{p}]$.

Input: Genotype counts $\boldsymbol{x} = (x_0, x_1, \ldots, x_K)$ for a K-ploid population, and gamete frequencies $\boldsymbol{p} = (p_0, p_1, \ldots, p_{K/2})$.

Calculate r_{ij} for $0 \leq i \leq j \leq K/2$:

$$r_{ij} = \begin{cases} p_i^2 & \text{if } i = j, \text{ and} \\ 2p_i p_j & \text{if } i \neq j. \end{cases} \tag{28}$$

for $k = 0, 1, \ldots, K$ **do**

  $\boldsymbol{c}_k = (r_{ij})_{0 \leq i \leq j \leq K/2 \text{ s.t. } i+j=k}$

  $\boldsymbol{d}_{k\ell} = c_{k\ell} / \sum_\ell c_{k\ell}$

  Sample $\boldsymbol{a}_k = (a_{ij})_{0 \leq i \leq j \leq K/2 \text{ s.t. } i+j=k}$ from Multinom$(x_k, \boldsymbol{d}_k)$

end for

$\boldsymbol{A}$ is now populated.

for $k = 0, 1, \ldots, K/2$ **do**

  $y_k = \sum_{i \text{ s.t. } i<k} a_{ik} + 2a_{kk} + \sum_{j \text{ s.t. } j>k} a_{kj}$.

end for

Return: Sampled gamete counts $\boldsymbol{y}$.

Procedure 2 Gibbs sampler and marginal likelihood estimation under the null of random mating when genotypes are known.

Input: Genotype counts $\boldsymbol{x} = (x_0, x_1, \ldots, x_K)$ for a K-ploid population, hyperparameter $\boldsymbol{\alpha} = (\alpha_0, \alpha_1, \ldots, \alpha_{K/2})$, B sampling iterations, and T burn-in iterations.
Initialize $\boldsymbol{p} \sim \text{Dirichlet}(\boldsymbol{\alpha})$
Initialize $\tilde{\boldsymbol{p}} \leftarrow \boldsymbol{p}$
Initialize $\tilde{L} \leftarrow \text{Multinom}(\boldsymbol{x}|n, \boldsymbol{p} * \boldsymbol{p})\,\text{Dirichlet}(\boldsymbol{p}|\boldsymbol{\alpha})$
Initialize $\hat{\pi} \leftarrow 0$
for $j = 1, 2, \ldots, T + B$ **do**
    Sample $[\boldsymbol{y}|\boldsymbol{x}, \boldsymbol{p}]$ using Procedure 1
    Sample $[\boldsymbol{p}|\boldsymbol{y}] \sim \text{Dirichlet}(\boldsymbol{y} + \boldsymbol{\alpha})$
    if $(\text{Multinom}(\boldsymbol{x}|n, \boldsymbol{p} * \boldsymbol{p})\,\text{Dirichlet}(\boldsymbol{p}|\boldsymbol{\alpha}) > \tilde{L})$ **and** $(j \leq T)$ **then**
        $\tilde{\boldsymbol{p}} \leftarrow \boldsymbol{p}$
        $\tilde{L} \leftarrow \text{Multinom}(\boldsymbol{x}|n, \boldsymbol{p} * \boldsymbol{p})\,\text{Dirichlet}(\boldsymbol{p}|\boldsymbol{\alpha})$
    else if $j > T$ **then**
        $\hat{\pi} \leftarrow \hat{\pi} + \frac{1}{B}\,\text{Dirichlet}(\tilde{\boldsymbol{p}}|\boldsymbol{y} + \boldsymbol{\alpha})$
    end if
end for
Estimate marginal likelihood

$$\widehat{\Pr}(\boldsymbol{x}|H_0) = \frac{\text{Multinom}(\boldsymbol{x}|n, \tilde{\boldsymbol{p}} * \tilde{\boldsymbol{p}})\,\text{Dirichlet}(\tilde{\boldsymbol{p}}|\boldsymbol{\alpha})}{\hat{\pi}} \tag{29}$$

Return: Estimate of marginal likelihood $\widehat{\Pr}(\boldsymbol{x}|H_0)$.

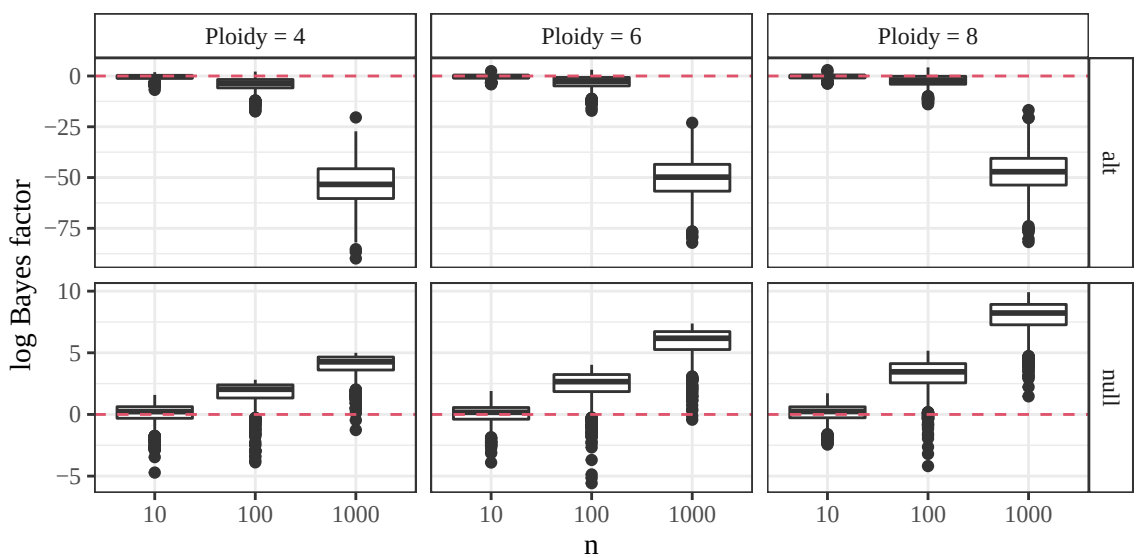

Figure 2: Log Bayes factors (y-axis) at different sample sizes (x-axis) at different ploidies (column facets) when either the null is true ("null") or the alternative is true ("alt"). The Bayes factor increases gradually with sample size when the null is true and decreases rapidly with sample size when the alternative is true.

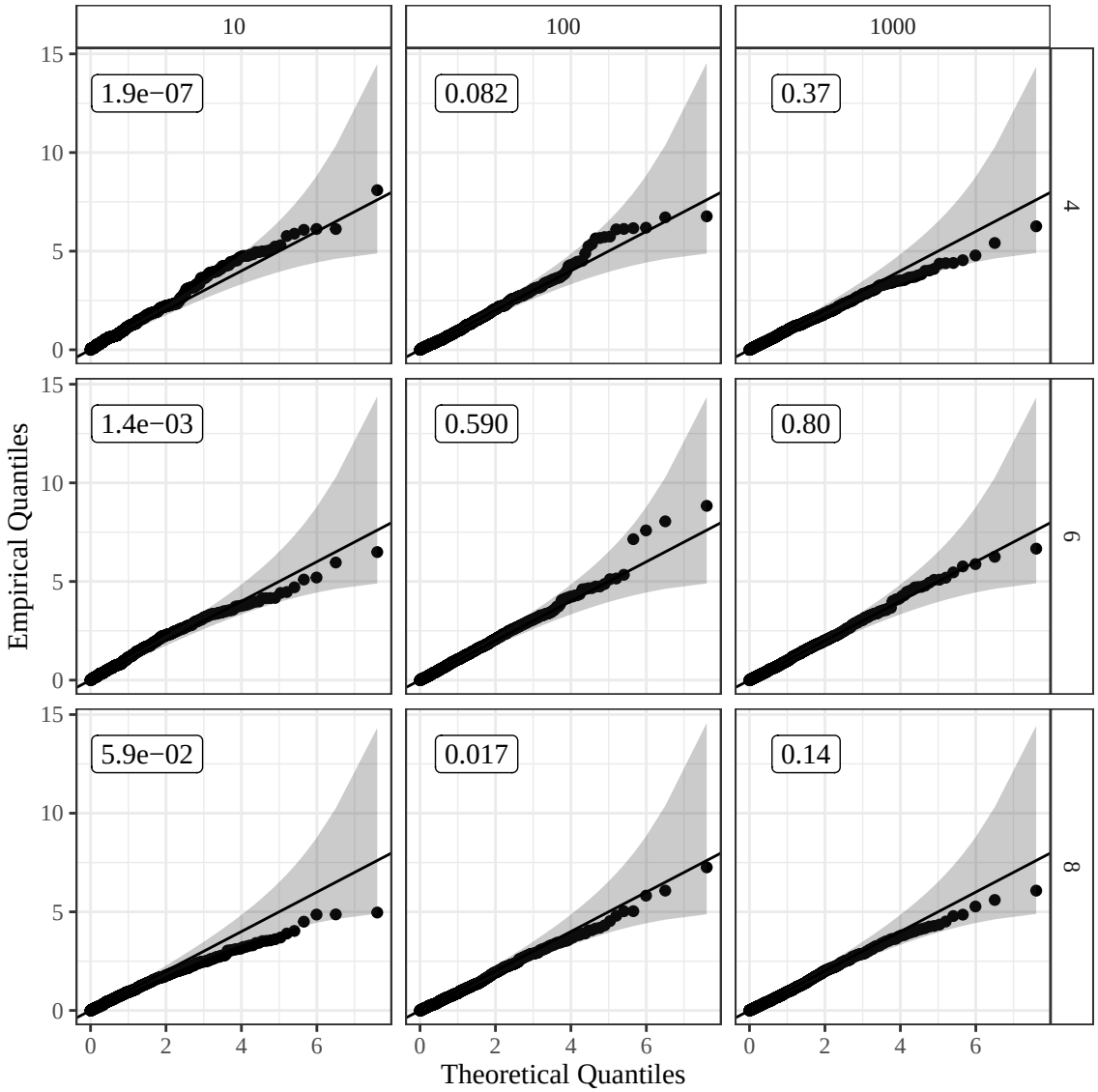

Figure 3: QQ-plots of p-values against the uniform distribution, on the $-\log$ scale, with the 95% simultaneous confidence bands from Aldor-Noiman et al. [2013], stratified by ploidy (row facets) and sample size (column facets). These p-values come from the simulations of Section 3.2 when the null is satisfied, and so should follow a uniform distribution. The numbers in the top-left of each plot are the p-values from a Kolmogorov-Smirnov test against the null of uniformity. Only for larger sample sizes are the p-values appropriately calibrated.

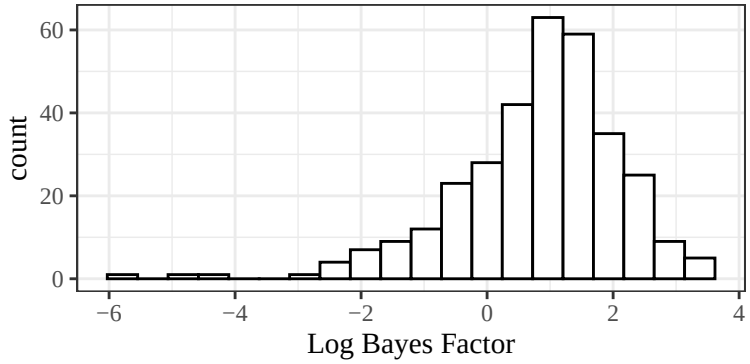

Figure 4: Histogram of log Bayes factors applied to all 325 loci from the white sturgeon data from Section 3.4. We see that most log Bayes factors are near or slightly above zero, indicating strong evidence that the population is undergoing random mating.

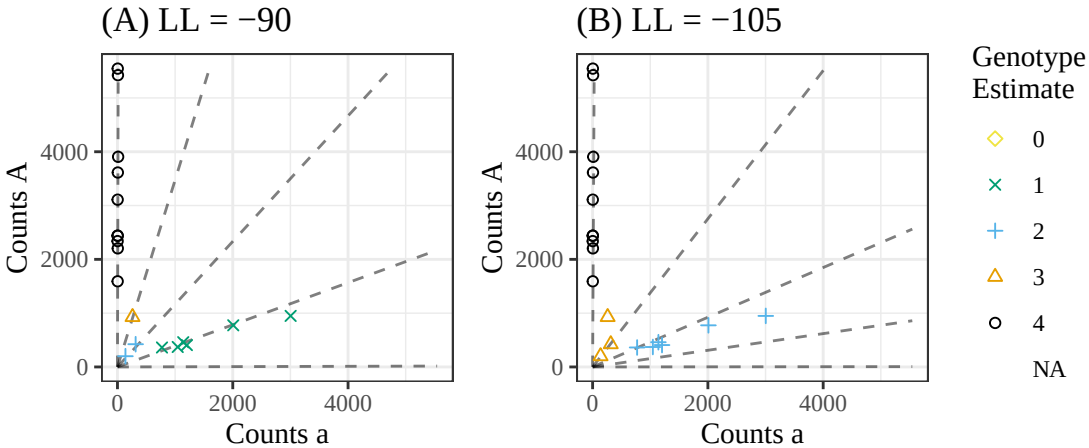

Figure 5: Genotype plots [Gerard et al., 2018] for two fits on SNP Atr_20529-52 from the white sturgeon data of Section 3.4. Each point is an individual, the color and shape of which indicate the estimated genotype under the two fits. We have the number of alternative reads on the x-axis and the number of reference reads on the y-axis. The lines radiating from the origin indicate the expected number of reference/alternative reads of a given genotype under that given fit. The left plot **(A)** is the global mode with a log-likelihood of -90, and the right plot **(B)** is a local mode with a log-likelihood of -105. The local mode exhibits high allelic bias, which can be seen by the lines clustering towards the x-axis. The test for random mating using the global mode of $\boldsymbol{x} = (0, 6, 2, 1, 10)$ yields a log Bayes factor of -5.9, while the test for random mating using the local mode of $\boldsymbol{x} = (0, 0, 6, 3, 10)$ yields a log Bayes factor of 1.1. The global mode is thus likely a genotyping error.

References

S. Aldor-Noiman, L. D. Brown, A. Buja, W. Rolke, and R. A. Stine. The power to see: A new graphical test of normality. *The American Statistician*, 67(4):249–260, 2013. doi: 10.1080/00031305.2013.847865.

P. M. E. Altham. Exact Bayesian analysis of an intraclass 2×2 table. *Biometrika*, 58(3):679–680, 1971. doi: 10.1093/biomet/58.3.679.

J. Berger. *Statistical Decision Theory: Foundations, Concepts, and Methods*. Springer Series in Statistics. Springer New York, 2013. ISBN 9781475717273. URL `https://books.google.com/books?id=TT_jBwAAQBAJ`.

S. Chib. Marginal likelihood from the Gibbs output. *Journal of the American Statistical Association*, 90(432):1313–1321, 1995. doi: 10.1080/01621459.1995.10476635.

L. V. Clark, A. E. Lipka, and E. J. Sacks. polyRAD: Genotype calling with uncertainty from sequencing data in polyploids and diploids. *G3: Genes |Genomes |Genetics*, 9(3):663–673, 2019. doi: 10.1534/g3.118.200913.

G. Consonni, E. Gutiérrez-Peña, and P. Veronese. Compatible priors for Bayesian model comparison with an application to the Hardy-Weinberg equilibrium model. *Test*, 17(3):585–605, 2008. doi: 10.1007/s11749-007-0057-7.

G. Consonni, E. Moreno, and S. Venturini. Testing Hardy–Weinberg equilibrium: An objective Bayesian analysis. *Statistics in Medicine*, 30(1):62–74, 2011. doi: 10.1002/sim.4084.

T. A. Delomas, S. C. Willis, B. L. Parker, D. Miller, P. Anders, A. Schreier, and S. Narum. Genotyping single nucleotide polymorphisms and inferring ploidy by amplicon sequencing for polyploid, ploidy-variable organisms. *Molecular Ecology Resources*, 21(7):2288–2298, 2021. doi: 10.1111/1755-0998.13431.

A. E. Gelfand and A. F. M. Smith. Sampling-based approaches to calculating marginal densities. *Journal of the American Statistical Association*, 85(410):398–409, 1990. doi: 10.1080/01621459.1990.10476213.

D. Gerard. Pairwise linkage disequilibrium estimation for polyploids. *Molecular Ecology Resources*, 21(4):1230–1242, 2021a. doi: 10.1111/1755-0998.13349.

D. Gerard. Scalable bias-corrected linkage disequilibrium estimation under genotype uncertainty. *Heredity*, 127(4):357–362, 2021b. doi: 10.1038/s41437-021-00462-5.

D. Gerard. Double reduction estimation and equilibrium tests in natural autopolyploid populations. *Biometrics*, 2022. doi: 10.1111/biom.13722. (in press).

D. Gerard and L. F. V. Ferrão. Priors for genotyping polyploids. *Bioinformatics*, 36(6):1795–1800, 2019. doi: 10.1093/bioinformatics/btz852.

D. Gerard, L. F. V. Ferrão, A. A. F. Garcia, and M. Stephens. Genotyping polyploids from messy sequencing data. *Genetics*, 210(3):789–807, 2018. doi: 10.1534/genetics.118.301468.

J. Gillespie. *Population Genetics: A Concise Guide*. Population Genetics. Johns Hopkins University Press, 2004. ISBN 9781421401706. URL `https://books.google.com/books?id=KAcAfiyHpcoC`.

Q. F. Gronau, H. Singmann, and E.-J. Wagenmakers. bridgesampling: An R package for estimating normalizing constants. *Journal of Statistical Software*, 92(10):1–29, 2020. doi: 10.18637/jss.v092.i10.

J. Haldane. Theoretical genetics of autopolyploids. *Journal of Genetics*, 22(3):359–372, 1930. doi: 10.1007/BF02984197.

O. J. Hardy and X. Vekemans. SPAGeDi: a versatile computer program to analyse spatial genetic structure at the individual or population levels. *Molecular Ecology Notes*, 2(4):618–620, 2002. doi: 10.1046/j.1471-8286.2002.00305.x.

L. Hosking, S. Lumsden, K. Lewis, A. Yeo, L. McCarthy, A. Bansal, J. Riley, I. Purvis, and C.-F. Xu. Detection of genotyping errors by Hardy-Weinberg equilibrium testing. *European Journal of Human Genetics*, 12(5):395–399, 2004. doi: 10.1038/sj.ejhg.5201164.

K. Huang, T. Wang, D. W. Dunn, P. Zhang, X. Cao, R. Liu, and B. Li. Genotypic frequencies at equilibrium for polysomic inheritance under double-reduction. *G3: Genes |Genomes |Genetics*, 9(5):1693–1706, 2019. doi: 10.1534/g3.119.400132.

R. E. Kass and A. E. Raftery. Bayes factors. *Journal of the American Statistical Association*, 90 (430):773–795, 1995. doi: 10.1080/01621459.1995.10476572.

H. Levene. On a matching problem arising in genetics. *The Annals of Mathematical Statistics*, 20 (1):91–94, 1949. doi: 10.1214/aoms/1177730093.

H. Li. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics*, 27(21):2987, 2011. doi: 10.1093/bioinformatics/btr509.

D. Lindley. Statistical inference concerning Hardy-Weinberg equilibrium. In J. Bernardo, M. DeGroot, D. Lindley, and A. Smith, editors, *Bayesian Statistics 3: Proceedings of the Third Valencia International Meeting, June 1-5, 1987*, pages 307–326. Oxford University Press, 1988. ISBN 9780198522201. URL https://books.google.com/books?id=2s4x-ZeMTkIC.

A. McKenna, M. Hanna, E. Banks, A. Sivachenko, K. Cibulskis, A. Kernytsky, K. Garimella, D. Altshuler, S. Gabriel, M. Daly, and M. DePristo. The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome research*, 20(9):1297–1303, 2010. doi: 10.1101/gr.107524.110.

P. G. Meirmans and P. H. Van Tienderen. GENOTYPE and GENODIVE: two programs for the analysis of genetic diversity of asexual organisms. *Molecular Ecology Notes*, 4(4):792–794, 2004. doi: 10.1111/j.1471-8286.2004.00770.x.

X.-L. Meng and W. H. Wong. Simulating ratios of normalizing constants via a simple identity: a theoretical exploration. *Statistica Sinica*, 6(4):831–860, 1996. doi: 10.2307/24306045.

L. E. Montoya-Delgado, T. Z. Irony, C. A. de B. Pereira, and M. R. Whittle. An unconditional exact test for the Hardy-Weinberg equilibrium law: Sample-space ordering using the Bayes factor. *Genetics*, 158(2):875–883, 2001. doi: 10.1093/genetics/158.2.875.

J. E. Mosimann. On the compound multinomial distribution, the multivariate β-distribution, and correlations among proportions. *Biometrika*, 49(1/2):65–82, 1962. doi: 10.2307/2333468.

C. Pereira and A. Rogatko. The Hardy-Weinberg equilibrium under a Bayesian perspective. *Revista Brasileira de Genética*, 7(4):689–707, 1984.

X. Puig, J. Ginebra, and J. Graffelman. A Bayesian test for Hardy-Weinberg equilibrium of biallelic X-chromosomal markers. *Heredity*, 119(4):226–236, 2017. doi: 10.1038/hdy.2017.30.

X. Puig, J. Ginebra, and J. Graffelman. Bayesian model selection for the study of Hardy-Weinberg proportions and homogeneity of gender allele frequencies. *Heredity*, 123(5):549–564, 2019. doi: 10.1038/s41437-019-0232-0.

R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2022. URL https://www.R-project.org/.

O. Serang, M. Mollinari, and A. A. F. Garcia. Efficient exact maximum a posteriori computation for Bayesian SNP genotyping in polyploids. *PLOS ONE*, 7(2):1–13, 2012. doi: 10.1371/journal.pone.0030906.

K. Shirasawa, M. Tanaka, Y. Takahata, D. Ma, Q. Cao, Q. Liu, H. Zhai, S.-S. Kwak, J. C. Jeong, U.-H. Yoon, H.-U. Lee, H. Hirakawa, and S. Isobe. A high-density SNP genetic map consisting of a complete set of homologous groups in autohexaploid sweetpotato (*Ipomoea batatas*). *Scientific Reports*, 7, 2017. doi: 10.1038/srep44207.

J. Shoemaker, I. Painter, and B. S. Weir. A Bayesian characterization of Hardy-Weinberg disequilibrium. *Genetics*, 149(4):2079–2088, 1998. doi: 10.1093/genetics/149.4.2079.

Stan Development Team. *RStan: the R interface to Stan*, 2022a. URL `https://mc-stan.org/`. R package version 2.21.5.

Stan Development Team. *Stan Modeling Language Users Guide and Reference Manual*, 2022b. URL `https://mc-stan.org`. Version 2.29.

M. Stift, R. Reeve, and P. H. Van Tienderen. Inheritance in tetraploid yeast revisited: segregation patterns and statistical power under different inheritance models. *Journal of Evolutionary Biology*, 23(7):1570–1578, 2010. doi: 10.1111/j.1420-9101.2010.02012.x.

J. D. Storey. The positive false discovery rate: a Bayesian interpretation and the q-value. *The Annals of Statistics*, 31(6):2013–2035, 2003. doi: 10.1214/aos/1074290335.

L. Sun, J. Gan, L. Jiang, and R. Wu. Recursive test of Hardy-Weinberg equilibrium in tetraploids. *Trends in Genetics*, 2020. doi: 10.1016/j.tig.2020.11.006.

R. E. Voorrips, G. Gort, and B. Vosman. Genotype calling in tetraploid species from bi-allelic marker data using mixture models. *BMC Bioinformatics*, 12(1):172, 2011. doi: 10.1186/1471-2105-12-172.

J. Wakefield. Bayes factors for genome-wide association studies: comparison with p-values. *Genetic Epidemiology*, 33(1):79–86, 2009. doi: 10.1002/gepi.20359.

J. Wakefield. Bayesian methods for examining Hardy–Weinberg equilibrium. *Biometrics*, 66(1): 257–265, 2010. doi: 10.1111/j.1541-0420.2009.01267.x.

R. S. Waples. Testing for Hardy-Weinberg proportions: have we lost the plot? *Journal of Heredity*, 106(1):1–19, 2015. doi: 10.1093/jhered/esu062.