Supplementary Material

Here, we shall present the results on the consistency and normal limiting distributions of the maximum likelihood estimators of the main parameters of the Y-linked BBP with preference, when the entire family tree up to some generation is observed. First, we shall derive these estimators.

Theorem 1 *The maximum likelihood estimators of α, p^R, and p^r based on the sample $(\mathcal{Z}_N, \mathcal{FM}_N)$ are, respectively,*

$$\hat{\alpha} = \frac{\sum_{n=0}^{N-1} F_{n+1}}{\sum_{n=0}^{N-1}(F_{n+1} + MR_{n+1} + Mr_{n+1})},$$

$$\hat{p}_k^R = \frac{\sum_{n=0}^{N-1} ZR_n(k)}{\sum_{n=0}^{N-1} ZR_n},\ k \in S^R, \qquad \textit{and} \qquad \hat{p}_l^r = \frac{\sum_{n=0}^{N-1} Zr_n(l)}{\sum_{n=0}^{N-1} Zr_n},\ l \in S^r.$$

Proof. It is immediate to verify from Equation (5) in the paper, that the expression for the complete log-likelihood function based on such sample is

$$\begin{aligned} l(p^R, p^r, \alpha | \mathcal{Z}_N, \mathcal{FM}_N) &= C^* + \sum_{n=0}^{N-1}(F_{n+1}\log\alpha + (MR_{n+1} + Mr_{n+1})\log(1-\alpha)) \\ &\quad + \sum_{n=0}^{N-1}\left(\sum_{k\in S^R} ZR_n(k)\log p_k^R + \sum_{l\in S^r} Zr_n(l)\log p_l^r\right), \end{aligned}$$

with C^* some constant.

Given the structure of that function, to maximize this expression subject to the constraints $0 \le \alpha \le 1$, $\sum_{k\in S^R} p_k^R = 1$ and $\sum_{l\in S^r} p_l^r = 1$, with $p_k^R, p_l^r \ge 0$, $k \in S^R$ and $l \in S^r$, it is enough to maximize each corresponding addend. Using the non-negativity of the Kullback-Leibler divergence, it is straightforward to verify that the log-likelihood is maximized by the choice of $\hat{\alpha}$, $\hat{p}_k^R$, and $\hat{p}_l^r$, and therefore they are the MLEs for α, p^R, and p^r. ∎

Corollary 1 *The maximum likelihood estimators of m_R and m_r based on the sample $(\mathcal{Z}_N, \mathcal{FM}_N)$ are, respectively,*

$$\hat{m}_R = \frac{\sum_{n=1}^{N}(FR_n + MR_n)}{\sum_{n=0}^{N-1} ZR_n} \qquad \textit{and} \qquad \hat{m}_r = \frac{\sum_{n=1}^{N}(Fr_n + Mr_n)}{\sum_{n=0}^{N-1} Zr_n}.$$

In the following results some asymptotic properties of the estimators $\hat{\alpha}$, $\hat{p}_k^R$ with $k \in S^R$, $\hat{p}_l^r$ with $l \in S^r$, $\hat{m}_R$, and $\hat{m}_r$ are studied. First, we shall deal with the results about their consistency, establishing previously some properties we shall need in the development of those results.

P1. $\liminf_{n\to\infty} \frac{ZR_{n+1}}{ZR_n} > 1$ a.s. on $A_{\infty,\infty} \cup A_{\infty,0}$

P2. $\liminf_{n\to\infty} \frac{Zr_{n+1}}{Zr_n} > 1$ a.s. on $A_{\infty,\infty} \cup A_{0,\infty}$

P3. $\lim_{n\to\infty} \frac{FR_{n+1}}{ZR_n} = \alpha m_R$ and $\lim_{n\to\infty} \frac{MR_{n+1}}{ZR_n} = (1-\alpha)m_R$ a.s. on $A_{\infty,\infty} \cup A_{\infty,0}$

P4. $\lim_{n\to\infty} \frac{Fr_{n+1}}{Zr_n} = \alpha m_r$ and $\lim_{n\to\infty} \frac{Mr_{n+1}}{Zr_n} = (1-\alpha)m_r$ a.s. on $A_{\infty,\infty} \cup A_{0,\infty}$

where $A_{\infty,0} = \{ZR_n \to \infty, Zr_n \to 0\}$, $A_{0,\infty} = \{ZR_n \to 0, Zr_n \to \infty\}$ and $A_{\infty,\infty} = \{ZR_n \to \infty, Zr_n \to \infty\}$.

Intuitively, $A_{\infty,0}$ (resp. $A_{0,\infty}$) means the fixation of the R allele (resp. r allele) and $A_{\infty,\infty}$ the survival or coexistence of both genotypes. Moreover, notice that $A_{\infty,0} \cup A_{\infty,\infty} = \{ZR_n \to \infty\}$, which corresponds to the survival of the R allele independently of the behaviour of the r allele, and that $A_{0,\infty} \cup A_{\infty,\infty} = \{Zr_n \to \infty\}$ with analogous meaning.

Remark 1 *Sufficient conditions for the sets $A_{\infty,0}$, $A_{0,\infty}$, and $A_{\infty,\infty}$ to have positive probability are given in González et al. (2006) and González et al. (2008), and conditions which guarantee **P1**-**P2** have been studied in González et al. (2008). Notice that, from **P1**-**P2** and using the conditioned Borel-Cantelli lemma, one can obtain **P3**-**P4**.*

Theorem 2 *The maximum likelihood estimators of p^R, p^r, and α based on $(\mathcal{Z}_N, \mathcal{FM}_N)$ verify:*

*i) If **P1** holds, then for each $k \in S^R$, $\hat{p}_k^R$ is strongly consistent for p_k^R on $A_{\infty,\infty} \cup A_{\infty,0}$.*

*ii) If **P2** holds, then for each $l \in S^r$, $\hat{p}_l^r$ is strongly consistent for p_l^r on $A_{\infty,\infty} \cup A_{0,\infty}$.*

*iii) If **P3** and **P4** hold and $\lim_{n\to\infty} \frac{ZR_n}{Zr_n}$ exists a.s. on $A_{\infty,\infty}$ (it could be ∞), then $\hat{\alpha}$ is strongly consistent for α on $A_{\infty,0} \cup A_{0,\infty} \cup A_{\infty,\infty}$.*

Proof. We start by proving *i*). The proof of *ii*) is analogous using the property **P2**. Firstly, we define the filtration $\mathcal{F}_n = \sigma(ZR_0, Zr_0, F_k, MR_k, Mr_k, k = 1, 2, ..., n)$, $n \geq 1$. Let $\varepsilon > 0$, $k \in S^R$ and define $A_n = \{|ZR_n(k) - p_k^R ZR_n| \geq \varepsilon ZR_n\}$, $n \geq 0$. Taking into account that the conditional distribution of $(ZR_n(k), k \in S^R)$ given ZR_n is a multinomial distribution with size ZR_n and probability p^R, then $E[ZR_n(k)|ZR_n] = ZR_n p_k^R$ a.s. and $Var[ZR_n(k)|ZR_n] = ZR_n p_k^R(1-p_k^R)$ a.s. Applying Chebyshev's inequality, from **P1** one obtains

$$\sum_{n=1}^{\infty} P(A_n|\mathcal{F}_n) \leq \sum_{n=1}^{\infty} \frac{Var[ZR_n(k)|ZR_n]}{\varepsilon^2 ZR_n^2} = \frac{p_k^R(1-p_k^R)}{\varepsilon^2} \sum_{n=1}^{\infty} \frac{1}{ZR_n} < \infty$$

a.s. on $\{ZR_n \to \infty\}$.

Then, using the conditioned Borel-Cantelli lemma,

$$\{ZR_n \to \infty\} \subseteq \left\{ \sum_{n=1}^{\infty} P(A_n|\mathcal{F}_n) < \infty \right\} = \liminf_{n\to\infty} A_n^c \quad \text{a.s.}$$

So, taking into account that A_n is equal to $\{|ZR_n(k)ZR_n^{-1} - p_k^R| \geq \varepsilon\}$ on $\{ZR_n \to \infty\}$, one has that $\lim_{n\to\infty} ZR_n(k)ZR_n^{-1} = p_k^R$ a.s. on $\{ZR_n \to \infty\}$. The proof is completed by applying the Toeplitz lemma.

To finish, we prove *iii*). This will be done by proving that

$$\lim_{n\to\infty} \frac{F_{n+1}}{F_{n+1} + MR_{n+1} + Mr_{n+1}} = \alpha \quad \text{a.s.} \tag{1}$$

on each of the sets $A_{\infty,0}$, $A_{0,\infty}$, and $A_{\infty,\infty}$. Again, the Toeplitz lemma is used to conclude the proof.

We shall prove (1) on $A_{\infty,0}$. The proof on $A_{0,\infty}$ is analagous. Taking into account **P3**, and from one generation onwards (which depends on the realization of the process), the offspring given by r couples is null on $A_{\infty,0}$. Then, recalling that $F_n = FR_n + Fr_n$, for $n = 1, 2, \ldots$,

$$\begin{aligned}
\lim_{n\to\infty} \frac{F_{n+1}}{F_{n+1} + MR_{n+1} + Mr_{n+1}} &= \lim_{n\to\infty} \frac{\frac{FR_{n+1}}{ZR_n}}{\frac{FR_{n+1}}{ZR_n} + \frac{MR_{n+1}}{ZR_n}} \\
&= \frac{\alpha m_R}{\alpha m_R + (1-\alpha) m_R} \\
&= \alpha \text{ a.s. on } A_{\infty,0}.
\end{aligned}$$

To prove the result on $A_{\infty,\infty}$, the relation between ZR_n and Zr_n must be taken into account because, a.s. on $A_{\infty,\infty}$,

$$\lim_{n\to\infty} \frac{F_{n+1}}{F_{n+1} + MR_{n+1} + Mr_{n+1}} = \lim_{n\to\infty} \frac{\frac{FR_{n+1}}{ZR_n}\frac{ZR_n}{Zr_n} + \frac{Fr_{n+1}}{Zr_n}}{\frac{FR_{n+1}}{ZR_n}\frac{ZR_n}{Zr_n} + \frac{Fr_{n+1}}{Zr_n} + \frac{MR_{n+1}}{ZR_n}\frac{ZR_n}{Zr_n} + \frac{Mr_{n+1}}{Zr_n}}. \tag{2}$$

Then, as by hypothesis there exists $\lim_{n\to\infty} ZR_n Zr_n^{-1}$ a.s. on $A_{\infty,\infty}$ (it could be ∞), one has:

a) If $\lim_{n\to\infty} ZR_n Zr_n^{-1} = 0$ a.s. on $A_{\infty,\infty}$, i.e., if $\{Zr_n\}_{n\geq 0}$ has a faster growth than $\{ZR_n\}_{n\geq 0}$, taking into account **P3** and **P4**, the right-hand side of (2) is a.s. on $A_{\infty,\infty}$ equal to
$$\frac{\alpha m_r}{\alpha m_r + (1-\alpha) m_r} = \alpha.$$

b) If $\lim_{n\to\infty} Zr_n ZR_n^{-1} = 0$ a.s. on $A_{\infty,\infty}$, i.e., if $\{ZR_n\}_{n\geq 0}$ has a faster growth than $\{Zr_n\}_{n\geq 0}$, from **P3** and **P4** one obtains an analogous result to a).

c) If $\lim_{n\to\infty} ZR_n Zr_n^{-1} = X$ a.s. on $A_{\infty,\infty}$ with X a random variable, $0 < X < \infty$, i.e., if both have a similar growth, from **P3** and **P4**, the right hand of (2) is a.s. on $A_{\infty,\infty}$ equal to
$$\frac{\alpha m_R X + \alpha m_r}{\alpha m_R X + \alpha m_r + (1-\alpha) m_R X + (1-\alpha) m_r} = \frac{\alpha(m_R X + m_r)}{m_R X + m_r} = \alpha. \quad \blacksquare$$

Corollary 2 *The maximum likelihood estimators of m_R and m_r based on $(\mathcal{Z}_N, \mathcal{FM}_N)$ verify:*

*i) If **P3** holds, then $\hat{m}_R$ is strongly consistent for m_R on $A_{\infty,\infty} \cup A_{\infty,0}$.*

*ii) If **P4** holds, then $\hat{m}_r$ is strongly consistent for m_r on $A_{\infty,\infty} \cup A_{0,\infty}$.*

Finally, we shall obtain some results on the asymptotic distribution of the derived maximum likelihood estimators. Previously, we shall need to assume some working hypotheses in order to develop these results.

H1. $P(A_{\infty,0}) > 0$, and there exist $\rho_R > 1$ and a random variable W_R such that $\{\rho_R^{-n} ZR_n\}_{n\geq 0}$ converges to W_R a.s. on $A_{\infty,0}$ and $A_{\infty,0} \subseteq \{0 < W_R < \infty\}$ a.s.

H2. $P(A_{\infty,\infty}) > 0$, and there exist $\rho_R^* > 1$ and a random variable W_R^* such that $\{{\rho_R^*}^{-n} ZR_n\}_{n\geq 0}$ converges to W_R^* a.s. on $A_{\infty,\infty}$ and $A_{\infty,\infty} \subseteq \{0 < W_R^* < \infty\}$ a.s.

Remark 2 *Conditions which guarantee **H1** and **H2** have been studied in González et al. (2008).*

We shall denote $P_{\mathcal{B}}(\cdot) = P(\cdot|\mathcal{B})$ for any set $\mathcal{B}$, and write $[x]$ to indicate the greatest integer number less than or equal to x.

The maximum likelihood estimator of p^R based on $(\mathcal{Z}_N, \mathcal{FM}_N)$ verifies the following asymptotic properties.

Theorem 3 *If P' is an absolutely continuous probability with respect to $P_{\mathcal{D}}$ ($P' \ll P_{\mathcal{D}}$) then, for any $x \in \mathbb{R}$, the maximum likelihood estimator of p_k^R, with $k \in S^R$, verifies that*

$$\lim_{N\to\infty} P'\left((p_k^R(1-p_k^R))^{-1/2} \left(\sum_{n=1}^{N} ZR_{n-1}\right)^{1/2} (\hat{p}_k^R - p_k^R) \leq x \right) = \phi(x),$$

with $\phi(x)$ being the standard normal distribution function, and where

i) if **H1** *holds, $\mathcal{D} = A_{\infty,0}$;*

ii) if **H2** *holds $\mathcal{D} = A_{\infty,\infty}$.*

Proof. Defining $TR_{01} = FR_{01} + MR_{01}$, the following equality is verified in distribution

$$\hat{p}_k^R = \frac{\sum_{n=1}^{N} ZR_{n-1}(k)}{\sum_{n=1}^{N} ZR_{n-1}} \stackrel{d}{=} \frac{\sum_{i=1}^{\sum_{n=1}^{N} ZR_{n-1}} I_{\{TR_{0i}=k\}}}{\sum_{n=1}^{N} ZR_{n-1}},$$

(recall that I_A is the indicator function of a set A). From this, one has, for all $x \in \mathbb{R}$, that

$$\begin{aligned} & P'\left((p_k^R(1-p_k^R))^{-1/2} \left(\sum_{n=1}^{N} ZR_{n-1}\right)^{1/2} (\hat{p}_k^R - p_k^R) \leq x \right) \\ = \quad & P'\left((p_k^R(1-p_k^R))^{-1/2} \left(\sum_{n=1}^{N} ZR_{n-1}\right)^{-1/2} \sum_{i=1}^{\sum_{n=1}^{N} ZR_{n-1}} (I_{\{TR_{0i}=k\}} - p_k^R) \leq x \right). \end{aligned}$$

First we shall deal with the proof of the result in the case *i*). Taking into account that **H1** holds and Cesaro's lemma, one has that, as $N \to \infty$,

$$(\rho_R)^{-N} \sum_{n=1}^{N} ZR_{n-1} \to (\rho_R - 1)^{-1} W_R \quad \text{a.s. on } A_{\infty,0}.$$

Thus to conclude it is sufficient to apply Theorem I in Dion (1974), with

$$a_N = \rho_R^N, \quad \nu_N = \sum_{n=1}^{N} ZR_{n-1}, \quad \Theta = (\rho_R - 1)^{-1} W_R$$

and, for $0 \leq t \leq 1$,

$$Y_N(t,\omega) = \left(p_k^R(1-p_k^R) \sum_{n=1}^{N} ZR_{n-1}(\omega) \right)^{-1/2} \sum_{i=1}^{[\sum_{n=1}^{N} ZR_{n-1}(\omega)t]} (I_{\{TR_{0i}=k\}}(\omega) - p_k^R).$$

The proof in case *ii*) is analogous. ∎

Corollary 3 *If* **H1** *and* **H2** *hold, then, for any $x \in \mathbb{R}$, the maximum likelihood estimator of p_k^R, with $k \in S^R$, verifies that*

$$\lim_{N\to\infty} P_{\{ZR_n \to \infty\}}\left((p_k^R(1-p_k^R))^{-1/2} \left(\sum_{n=1}^{N} ZR_{n-1}\right)^{1/2} (\hat{p}_k^R - p_k^R) \leq x \right) = \phi(x),$$

with $\phi(x)$ being the standard normal distribution function.

Remark 3 *By Lemma 2.3 in Guttorp (1991), the probability $P_{\{ZR_n \to \infty\}}$ in Corollary 3 can be replaced by $P_{\{ZR_{N-1}>0\}}$. Hence, taking into account i) in Proposition 2 and applying the Slutsky theorem, one obtains that if $ZR_{N-1} > 0$ then the $(1-\gamma)$-level asymptotic confidence interval for p_k^R is*

$$\hat{p}_k^R \pm z_\gamma \sqrt{\hat{p}_k^R (1 - \hat{p}_k^R) \left(\sum_{n=1}^{N} ZR_{n-1} \right)^{-1}},$$

where z_γ satisfies $\phi(z_\gamma) = 1 - \gamma/2$ with $\gamma \in (0,1)$, and $\phi(x)$ is the standard normal distribution function.

Remark 4 *Analogous asymptotic distribution results to those related to $\hat{p}_k^R$, with $k \in S^R$, can be obtained for $\hat{p}_l^r$, with $l \in S^r$, using similar working hypotheses to* **H1** *and* **H2**. *Moreover, the asymptotic normality of the (suitably normalized) estimators $\hat{m}_R$ and $\hat{m}_r$ can be established by following a similar reasoning to that given in González et al. (2007). Also, asymptotic normality can be derived for $\hat{\alpha}$.*

Supplementary Figures

Here, we shall present the figures corresponding to the simulated example given in Section 4.

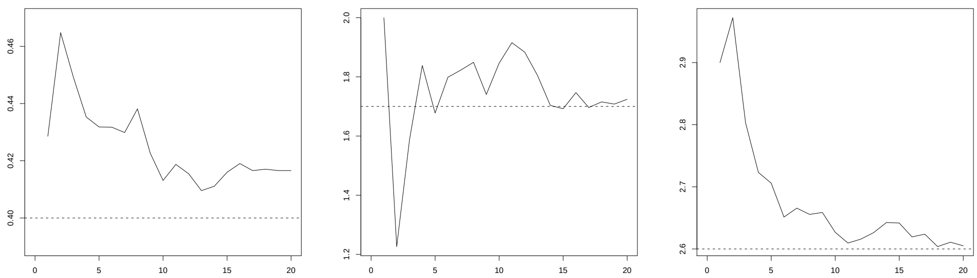

Figure 1: Evolution of $\hat{\alpha}_{EM}$ (left), $\hat{m}_R^{EM}$ (centre), and $\hat{m}_r^{EM}$ (right) over the course of the generations, together with the true value of each parameter (dashed line).

References

Dion, J. P. (1974). Estimation of the mean and the initial probabilities of a branching process. *J. Appl. Probab.* **11**, 687–694.

González, M., Hull, D. M., Martínez, R., & Mota, M. (2006). Bisexual branching processes in a genetic context: The extinction problem for Y-linked genes. *Math. Biosci.* **202**, 227–247.

González, M., Martínez, R., & Mota, M. (2008). Bisexual branching processes in a genetic context: Rates of growth for Y-linked genes. *Math. Biosci.* **215**, 167–176.

González, M., Mota, M., & Ramos, A. (2007). Moment estimation in the class of bisexual branching processes with population-size dependent mating. *Aust. N. Z. J. Stat.* **49(1)**, 37–50.

Guttorp, P. (1991). *Statistical Inference for Branching Processes.* John Wiley and Sons, Inc.

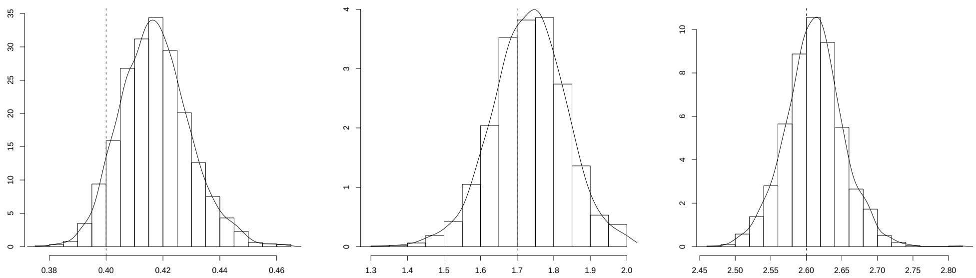

Figure 2: Bootstrap approximation to the sampling distribution of $\hat{\alpha}_{EM}$ (left), $\hat{m}_R^{EM}$ (centre), and $\hat{m}_r^{EM}$ (right), at generation 20, together with the true value of each parameter (dashed line) and kernel density estimates (solid line).

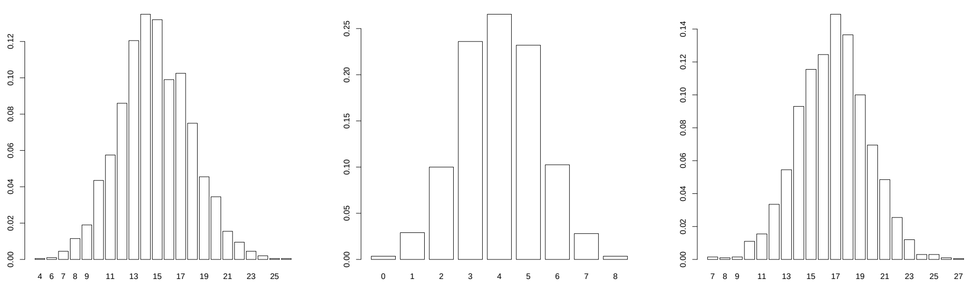

Figure 3: Histogram of the estimated predictive distribution of F_{21} (left), MR_{21} (centre), and Mr_{21} (right), when fm_{20} is observed.