

Comparing the estimation properties of different statistics for measuring sexual isolation from mating frequencies

A. PÉREZ-FIGUEROA, A. CABALLERO and E. ROLÁN-ALVAREZ*

Departamento de Bioquímica, Genética e Inmunología, Facultad de Biología, Universidad de Vigo, Campus Universitario, 36310 Vigo, Spain

Received 10 March 2004; accepted for publication 16 August 2004

Sexual isolation is a key component of reproductive isolation, involving mate choice among mature adults. While there are various statistics for estimating sexual isolation from mating frequencies, their ability to produce unbiased estimates varies considerably, depending on the particular situation. We investigated, under different biological scenarios, the estimation properties (statistical bias, efficiency, root mean square, statistical test) of 12 statistics commonly used in the literature for measuring sexual isolation. Yule's Q , V , YA (and related indices) and I_{PSI} are revealed to be the most efficient, with the smallest biases and root mean square deviations. Yule's Q , YA and I_{PSI} show better estimation properties when using infinite sample sizes, while I_{PSI} is preferable using smaller sample sizes. Other statistics investigated should be avoided, at least within the range of conditions considered. Regarding the parametric test of hypothesis, the best alternative is YA . We discuss the advantages and drawbacks of the various estimators, and propose I_{PSI} as the safest for biological sample sizes. © 2005 The Linnean Society of London, *Biological Journal of the Linnean Society*, 2005, 85, 307–318.

ADDITIONAL KEYWORDS: assortative mating – mate choice – mating propensity – multiple choice – pre-mating reproductive isolation.

INTRODUCTION

Sexual isolation is a consequence of the mating strategy of a particular species, where it seeks to avoid mating with individuals from different species (Gilbert & Starmer, 1985). It may occur because natural selection favours any evolutionary strategy that impedes the waste of time and energy of producing unsuccessful (low fitness) progeny ('reinforcement' *sensu* Dobzhansky, 1937). Sexual isolation can originate as the by-product of ecological (adaptive) evolution of traits related to mating behaviour (Schlutter, 2001), or simply as a consequence of the genetic divergence of geographically isolated populations (Turelli, Barton & Coyne, 2001).

The mechanism of sexual isolation results from mate choice, the individual preference for choosing a particular pair type instead of others. This preferen-

tial mating can also occur within species, producing: (1) assortative mating, when there is an excess of similar pair types (i.e. they are more frequent than might be expected by chance), or (2) disassortative mating, when there is an excess of dissimilar pair types (O'Donald, 1980). According to this view, sexual isolation can be responsible for the existence of polymorphisms within species (phenotypes, morphs, ecotypes, or incipient species). There is an implicit assumption that it can finally lead to full prezygotic reproductive isolation.

Sexual isolation can be inferred from mating patterns observed in the wild or in laboratory conditions (multiple choice experiments) using a number of different statistical methods (reviewed by Gilbert & Starmer, 1985; Rolán-Alvarez & Caballero, 2000) based on comparing the relative frequency of similar and dissimilar pair types. There is an obvious interest in quantifying sexual isolation, as it can be fundamental to testing different evolutionary models (Tregenza,

*Corresponding author. E-mail: rolan@uvigo.es

Pritchard & Butlin, 2000; Tregenza, 2002). In two classic studies, Coyne & Orr (1989, 1997) estimated sexual isolation for different pairs of *Drosophila* species. They found that it evolves faster than postzygotic reproductive isolation, as sympatric pairs of species, less genetically divergent, showed stronger sexual isolation estimates than allopatric pairs. This represents a support of the reinforcement hypothesis (Coyne & Orr, 1997; Turelli *et al.*, 2001; but see Endler, 1989).

There have been a few reviews on the estimation properties of statistics of sexual isolation. Marín (1991) pointed out that it is not possible to obtain unbiased estimates of sexual isolation from a single experiment if both mate choice (the cause of sexual isolation) and mating propensity (a tendency for different phenotypes/genotypes to intrinsically mate more frequently than others) are working together.

However, not all statistics are biased to the same extent when trying to estimate sexual isolation under differential mating propensities and type frequencies. Gilbert & Starmer (1985) compared the resampling properties of five estimators (*joint isolation index*, *Yule's V*, *Yule's Q*, *YA* and *Chi-square*). Rolán-Alvarez & Caballero (2000) compared the asymptotic (infinite population) properties of these estimators and a new one, *I_{PSI}*. The results indicate that some estimators, such as the *joint isolation index*, should be avoided (due to extreme bias for differential marginal frequencies and differential mating propensities), whereas others, such as *Yule's V*, *YA* and *I_{PSI}*, behave relatively well.

These reviews, however, did not consider the effect of resampling population frequencies, a key factor when estimates of sexual isolation are obtained with data from the wild. In the present study we extend the comparison of estimation properties to a larger set (12) of sexual isolation statistics, investigating their asymptotic bias, their bias from resampling of mating pairs and population frequencies, as well as other estimation properties.

MODEL AND STATISTICAL PROPERTIES OF THE ESTIMATORS

DESCRIPTION OF THE BASIC MODELS AND THE STATISTICS COMPARED

Previous evaluations of sexual isolation statistics from mating frequency data (Gilbert & Starmer, 1985; Rolán-Alvarez & Caballero, 2000) have assumed a multiple choice framework in which some types or morphs of a particular taxa are placed together to mate for a period of time (see Knoppien, 1985). This framework has been developed to ascertain dimorphic traits in dioecious and polygamous species, but can be easily extended to multiple polymorphism as well to wild populations (see e.g. Rolán-Alvarez *et al.*, 1999;

Cruz *et al.*, 2004). It can also stand for complete (all possible combinations of pairs) no-choice experiments, where single pairs are allowed to mate (see Arnold *et al.*, 1996; Nosil, Crespi & Sandoval, 2002).

The notations used to define the various statistics are listed in Table 1A. Let *A* and *B* be the frequencies of two given morphs of males, and *A'* and *B'* the corresponding frequencies of females, in a particular natural or laboratory population. After a given period of time the observed number of mates for every particular male and female morph is *aa*, *ab*, *ba* and *bb*, respectively, for a total number *t*.

Different sexual isolation estimators used in the literature are listed in Table 2, using the notation of Table 1A. Some of these estimators, such as *YA*, *DI*, *W_i* and *Z_i*, are algebraically related and would be expected to show rather similar statistical behaviour. Most of them, however, fail when there are existing zeros in cells *ab* or *ba* of the mating table, the only exception being the *YA* index, which includes a correction for these cases (see Ringo, Dowse & Lagasse, 1986). The *T_i* index was modified from Tilley, Verrel & Arnold (1990) to be used in the present study, as it was originally developed on the assumption of equal

Table 1. A, notation used for describing a typical multiple choice experiment. *A* and *B*, *A'* and *B'* are the male and female frequencies, respectively, of the two morphs studied. The observed number of copulating pairs is *t*, with *aa*, *ab*, *ba* and *bb*, the observed number of copulating pairs for every male and female combination. B, notation used for simulating the frequencies of mating pairs (see text). *W* and *C* are, respectively, the mating propensity of each morph and sex, and the choice coefficient of each mating type studied.

A		Females		
		<i>A'</i>	<i>B'</i>	
Males	<i>A</i>	<i>aa</i>	<i>ab</i>	<i>aa + ab</i>
	<i>B</i>	<i>ba</i>	<i>bb</i>	<i>ba + bb</i>
		<i>aa + ba</i>	<i>ab + bb</i>	
$t = aa + ab + ba + bb$				
B		Mate propensity	<i>W_{A'}</i>	<i>W_{B'}</i>
		<i>W_A</i>	<i>C_{AA}</i>	<i>C_{AB}</i>
		<i>W_B</i>	<i>C_{BA}</i>	<i>C_{BB}</i>
		Choice coefficients		

Table 2. The different statistics of sexual isolation studied, presented as a function of data from a multiple choice experiment (notation from Table 1A)

Estimator ¹	Formula
Joint isolation index (<i>I</i>)	$I = \frac{aa+bb-ab-ba}{t}$
Merrel's isolation distance (<i>Me</i>)	$Me = \frac{ba+ab}{aa+bb}$
Yule's V (<i>V</i>)	$V = \frac{(aa \times bb) - (ab \times ba)}{\sqrt{(aa+ab)(ba+bb)(aa+ba)(ab+bb)}}$
Yule's Q (<i>Q</i>)	$Q = \frac{(aa \times bb) - (ab \times ba)}{(aa \times bb) + (ab \times ba)}$
Chi-square (χ^2)	$\chi^2 = \frac{4[(aa \times bb) - (ab \times ba)]}{t^2}$
YA index (<i>YA</i>)	$YA = \frac{\sqrt{a'} - 1}{\sqrt{a'} + 1} \quad a' = \frac{aa \times bb}{ab \times ba}$
Coyne & Orr (<i>Co</i>)	$Co = 1 - \frac{ba+ab}{aa+bb}$
Hollocher <i>et al.</i> (<i>DI</i>)	$DI = \ln\left(\frac{ab \times ba}{aa \times bb}\right)$
W_i (<i>W_i</i>)	$W_i = \frac{aa \times bb}{ab \times ba}$
Levene's joint isolation index (<i>Z_i</i>)	$Z_i = \sqrt{\frac{aa \times bb}{ab \times ba}}$
Tilley <i>et al.</i> (<i>Ti</i>)	$Ti = \frac{aa+bb-ab-ba}{aa}$
<i>I_{PSI}</i> index (<i>I_{PSI}</i>)	$I_{PSI} = \frac{PSI_{aa} + PSI_{bb} - PSI_{ab} - PSI_{ba}}{PSI_{aa} + PSI_{ab} + PSI_{ba} + PSI_{bb}}$

¹*I*, *Merrel*: Merrel (1950); *V*, *Q*, χ^2 : see Gilbert & Starmer (1985); *YA*: Ringo *et al.* (1986); *Co*: Coyne & Orr (1989); *DI*: Hollocher *et al.* (1997); *W_i*: Goux & Anxolabehere (1980); *Z_i*: Erhman & Petit (1968); *Ti*: modified from Tilley *et al.* (1990). *I_{PSI}* from Rolán-Alvarez & Caballero (2000).

marginal frequencies (see below) in the mating table. *Co* and *Me* indices are complementary and therefore their properties are identical. Thus, only one (*Co*) will be analysed henceforth. It should be noted that all the statistics analysed range from -1 to 1, except *Co* (-∞ to 1) and *Ti* (-∞ to ∞).

The *PSI* statistic estimates mate choice coefficients for each type of mating pair (Rolán-Alvarez & Caballero, 2000). For example, for the *ab* cell, it is calculated as

Table 3. Parameter sets used to investigate the estimation properties of different sexual isolation indices. The choice coefficients used in the different scenarios are described in the text

	A	A'	B	B'
Morph				
Frequencies				
Equal	0.5	0.5	0.5	0.5
Unequal 1	0.2	0.5	0.8	0.5
Unequal 2	0.2	0.2	0.8	0.8
Unequal 3	0.2	0.8	0.8	0.2
Mate propensity coefficients	W_A	$W_{A'}$	W_B	$W_{B'}$
Neutral	1	1	1	1
Selected 1	1	1	0.1	1
Selected 2	1	1	0.5	0.5
Selected 3	1	0.5	0.5	1

$$PSI_{ab} = \frac{ab \times t}{(aa + ab) \times (ab + bb)}$$

PSI statistics for all mating pairs can be combined in a single statistic using any of the previous indices and, for this purpose, Rolán-Alvarez & Caballero (2000) have suggested the *joint isolation index*, as this is a very intuitive and simple statistic widely used in the literature. Thus, *I_{PSI}* can be obtained as shown in Table 2. We also considered the inclusion of *PSI* coefficients in other statistics (Yule's *Q* and *YA*) but focused on *I_{PSI}*.

ASYMPTOTIC PROPERTIES OF THE STATISTICS

In principle, sexual isolation should be related exclusively to mate choice: the mating preferences of one sex with respect to the different types of the other sex (Gilbert & Starmer, 1985). These preferences are quantified by the mate choice coefficients (*C* in Table 1B) and therefore sexual isolation statistics should estimate the relative importance of the two diagonals of mate choice coefficients in the mating table. In practice, however, different mating propensities (*W* in Table 1B) as well as population frequencies of the mating types (*A*, *A'*, *B* and *B'* in Table 1A) may bias the sexual isolation estimates (Gilbert & Starmer, 1985; Casares *et al.*, 1998; Rolán-Alvarez & Caballero, 2000). Thus, we studied the efficiency of the statistics in estimating the mate choice coefficients under different scenarios, including a variety of mate choice situations, different mating propensities of the morphs and changes in marginal frequencies.

We used an orthogonal combination of four sets of marginal morph frequencies and four sets of mate propensity coefficients (see Table 3), covering a wide

range of possibilities. In addition, 11 combinations of choice coefficients were used, taking $C_{AA} = C_{BB} = 1$ and $C_{AB} = C_{BA}$ with values of 1, 0.8181, 0.6666, 0.5384, 0.4286, 0.3333, 0.25, 0.1765, 0.1111, 0.0526 and 0. The joint isolation index (I) used on these combinations of choice coefficients renders values of a theoretical isolation index $I = 0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9$ and 1, respectively, covering the whole range of scenarios, from no isolation ($I = 0$) to complete isolation ($I = 1$). Thus, a total of 176 different scenarios (combinations of morph frequencies, mate propensities and choice coefficients) were assessed. All these scenarios can potentially bias sexual isolation estimates.

The efficiency of the different statistics in estimating sexual isolation with large (infinite) sample sizes was investigated by calculating the *asymptotic bias* of the statistics (the difference between the a posteriori and the a priori sexual isolation) in every scenario. The a posteriori sexual isolation estimate in each of the scenarios was calculated using each particular estimator and the expected frequencies of mating pairs (the expected values of aa , ab , ba and bb in Table 1A). For example, the expected frequency of ab is $A \times W_A \times B' \times W_B \times C_{AB}$. The a priori sexual isolation estimate was calculated using each particular estimator and the true mate choice coefficients for every scenario (e.g. C_{AB} ; Table 1B).

SAMPLING PROPERTIES OF THE STATISTICS

Under experimental conditions, the expectations for every mating pair type can vary from one trial to another simply because of sampling error. We incorporated stochastic (sampling) processes by resampling 10^5 times the expected frequencies of mating pairs by Monte Carlo methods (Nooren, 1989). Zero values in a cell of the mating table were substituted for 10^{-6} , as zero values produce errors in the calculation of many indices (e.g. YA and related ones), but similar qualitative results were obtained when 0s were replaced by 1s (not shown). For every simulated sample (20, 40, 60, 80 or 120 mating pairs) we calculated the a posteriori *resampling* estimate of sexual isolation averaged over the 10^5 resamplings per statistic in the 176 different scenarios. Thus, a *total bias* was calculated as the difference between (1) the averaged a posteriori resampling estimate from simulations and (2) the a priori sexual isolation estimate (obtained from the mate choice coefficients, as explained above).

The former approach can be used under two different evolutionary models. First, a *laboratory model*, which can be applied to typical laboratory multiple choice experiments. In this model, the number of individuals used in the experiment is directly controlled by the researcher (morph frequencies are known) and

therefore only mating pairs are resampled in the analysis, as described above.

A second possibility occurs when sexual isolation indices are applied to data obtained from natural populations (see Rolán-Alvarez *et al.*, 1999; Cruz *et al.*, 2004). In this case, the population morph frequencies are unknown and therefore also suffer from stochastic processes. Thus, in order to check the estimation properties of the indices for studies conducted in the wild, we repeated the above stochastic simulations while resampling, in addition to the expected mating pairs (20 in this instance), the frequencies of the mating types (20 males and 20 females). We refer to this situation as the *natural model*, and it was run for identical scenarios to those outlined above for the laboratory model.

For both models we also studied several estimation properties of the different statistics, such as efficiency (inverse of the sampling variance), and root mean square deviation (RMS: square sampling bias plus sampling variance).

STATISTICAL TESTS OF THE ESTIMATORS

Approximations to the sampling variance are known for some statistics (I and I_{PSI} , *Yule's Q*, *Yule's V* and YA ; see Gilbert & Starmer, 1985; Spieth & Ringo, 1983), being used in classical t or chi-square tests of hypotheses (Sokal & Rohlf, 1997). These sampling variances, however, have been developed assuming infinite sample sizes and do not necessarily work properly at biological sample sizes.

Thus, we estimated the type I error (the probability of false rejection of the null hypothesis) of a parametric t -test under the different sampling variances available for those statistics. We calculated the percentage of the resampled sexual isolation estimates under the laboratory model, which was significantly different from the true (a priori) estimate. The resampled estimate minus the a priori estimate divided by the sampling standard deviation of the estimator was compared to a t -value with one degree of freedom (Sokal & Rohlf, 1997).

Another alternative to parametric inference for poorly known statistics is bootstrapping, which works better for unbiased statistics that do not show asymmetry in the resampling distribution (Efron, 1982). Thus, we checked how the different alternative statistics behave under bootstrapping regarding their resampling (simulated) distribution.

RESULTS

The *asymptotic bias* is shown in Figure 1 for the different statistics under the range of scenarios investigated. The figure shows the median of the biases, the

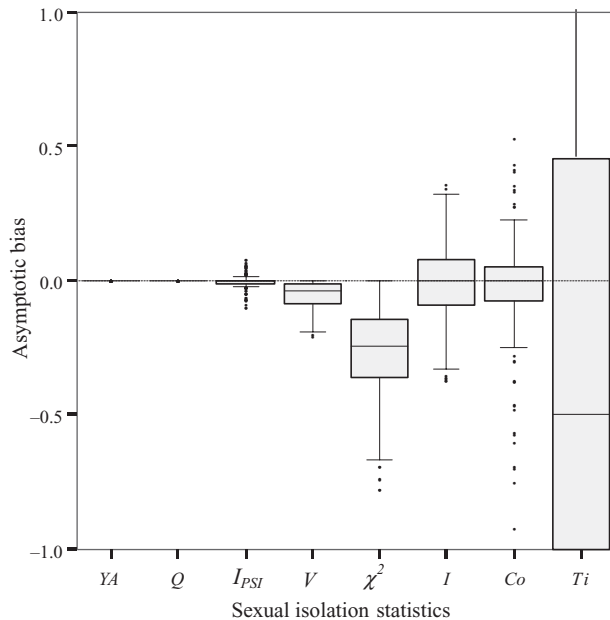


Figure 1. Distribution of the asymptotic bias for the studied sexual isolation estimators over 176 different scenarios. The boxes represent the 25–75th percentiles of the bias distribution across scenarios. The vertical bars include all cases with biases up to 1.5 times the biases limiting the 25–75th percentiles, and the dots represent further outliers. Cases within vertical bars and dots represent 50% of the bias distribution (percentiles 0–25 and 75–100).

25–75 percentiles, as well as the outliers. This allows us to see whether or not the bias distribution is centred around zero as well as the range of biases across scenarios. The DI , W_i and Z_i indices are not shown because their biases were identical to those of YA . These results suggest that YA and Q are the best at infinite sample sizes as they show no bias for any of the scenarios studied.

I_{PSI} and V showed a small bias in most scenarios. I_{PSI} was centred around zero with values ranging between -0.07 and 0.05 for a 90% confidence interval. However, V was biased on average. I and Co should be used with caution. For example, the bias in Co ranged between -0.49 and 0.276 for the 90% confidence interval of the bias distribution, while that of I ranged between -0.26 and 0.25 (see Fig. 1). Furthermore, Co was up to 50% biased under those scenarios with the strongest marginal frequency and mating propensity effects.

Finally, χ^2 and Ti appear to be not very useful estimators, as they were not centred and showed the largest variances of biases across scenarios. The inclusion of PSI coefficients in other statistics (Yule's Q and YA) did not improve their performance, because their bias was zero at infinite sample sizes. Overall, the best choices are YA (and related statistics), Q and I_{PSI} , as they were unbiased on average, and showed zero or

very small variances across scenarios. Good asymptotic estimation properties do not guarantee, however, a similar behaviour at biological (reduced) sample sizes, as will be shown below.

We studied the sampling properties (using 20 mating pairs) of the sexual isolation indices under the laboratory and natural models in the same scenarios considered above. The total bias distribution across scenarios for the different statistics is presented in Figure 2 for both models, and the particular biases for each of the simulation scenarios under the natural model are shown in Table 4.

χ^2 and Ti did not work properly at reduced sample sizes for any of the models and therefore their use should be avoided. For example, the bias of χ^2 ranged between -0.67 and 0 for a confidence interval of 90% of the bias distribution. Co and I revealed high biases under the natural and laboratory models, respectively.

Note that the biases shown by the statistics under the different scenarios do not show very clear patterns (Table 4). For example, Co showed a low bias for equal morph frequencies and two mate propensity sets (cases selected 1 and 2), but showed a large bias in case 3. However, for the unequal set of morph frequencies it performed poorly in case selected 2. The bias for different mating propensity effects but equal marginal frequencies is particularly relevant, as this cannot be experimentally avoided in multiple choice experiments.

Four indices (YA , Q , I_{PSI} and V) showed the best behaviour at reduced sample sizes under both models and can therefore be considered as the most useful statistics for estimating sexual isolation at both reduced and infinite sample sizes. However, the indices related to YA (DI , Z_i and W_i) are not recommended, as they suffer from important resampling biases. They do not correct for zeros in cells (the total bias was always larger than two in these cases). The smallest bias under both models was achieved by I_{PSI} , which performed better in most scenarios.

The inclusion of PSI coefficients in other statistics (Yule's Q and YA) did not improve the performance of I_{PSI} . In fact, the former showed somewhat larger average bias (0.034 for YA_{PSI} ; -0.012 for Q_{PSI}) and larger standard deviation of the bias (0.054 for YA_{PSI} ; 0.038 for Q_{PSI}) than the latter (average $I_{PSI} = -0.009$; $SD = 0.027$). Henceforth, the discussion relates only to the four statistics with the best performance: YA , Q , I_{PSI} and V .

Figure 3 shows the total bias of YA , Q , I_{PSI} and V at different levels of sexual isolation under the laboratory model. I_{PSI} showed low general bias, whereas Q and V showed substantial bias at intermediate values. YA showed a nearly linear increase of bias except at extreme isolation values, and could not be calculated for $YA = 1$ because it is not possible to obtain the a priori estimate when $a' = 0$ (see Table 2).

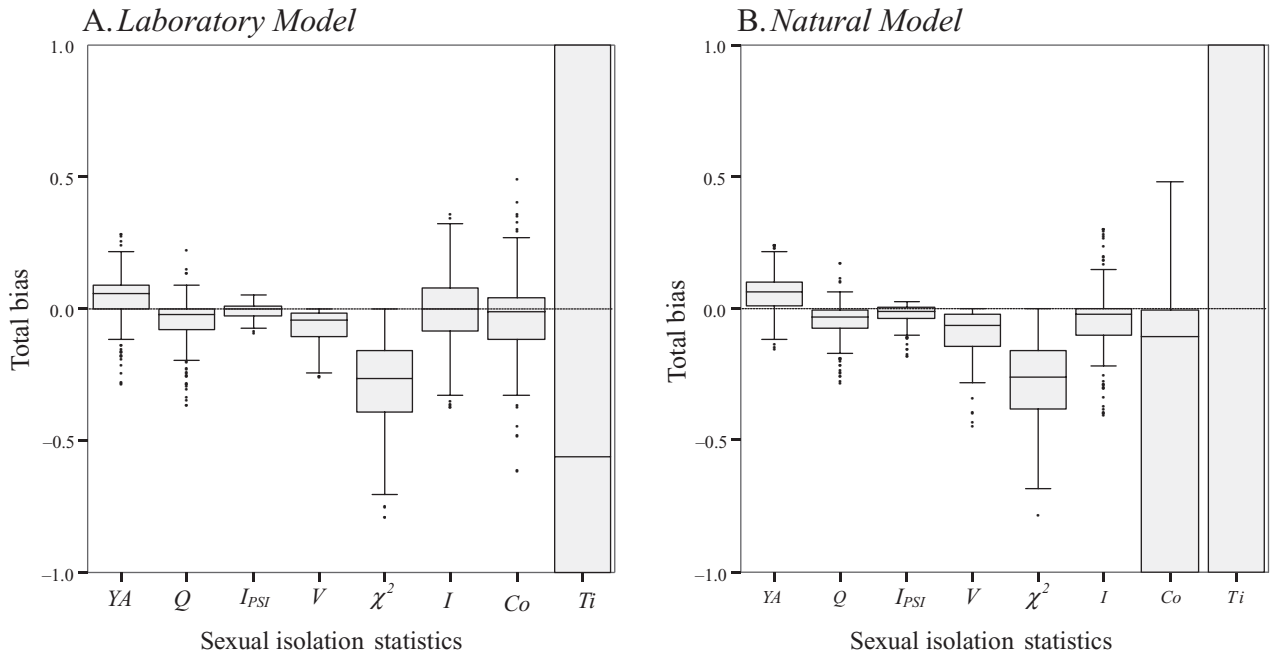


Figure 2. Distribution of the total bias (asymptotic and resampling biases) for the studied sexual isolation estimators over 176 different scenarios under (A) the laboratory model (resampling only mating pairs), and (B) the natural model (resampling both mated and unmated individuals). Boxes, vertical bars and dots as in Fig. 1.

Table 4. Mean total bias across all isolation levels in every simulated scenario under the natural model. See Table 3 for case legends. Range for YA, Q, I_{PSI}, V, X² and I is [-1, 1]. Range for Co is [-∞, 1] and for Ti is [-∞, +∞]

Morph frequencies	Mate propensities	Sexual isolation statistics							
		YA	Q	I _{PSI}	V	χ ²	I	Co	Ti
Equal	Neutral	0.04	-0.02	0.01	0.00	-0.13	0.00	-0.04	< -100
Unequal 1		0.06	-0.04	0.00	-0.06	-0.30	0.00	-0.04	-0.77
Unequal 2	Neutral	-0.16	-0.23	-0.03	-0.14	-0.48	0.21	0.20	> 100
Unequal 3		0.21	0.04	-0.03	-0.10	-0.24	-0.27	< -100	> 100
Equal	Selected 1	0.02	-0.10	-0.03	-0.15	-0.44	0.00	-0.04	-0.77
	Selected 2	0.02	-0.05	0.01	-0.03	-0.28	0.07	0.05	-0.37
	Selected 3	0.08	-0.01	0.00	-0.03	-0.16	-0.08	-27.42	26.62
Unequal 1	Selected 1	0.05	-0.02	0.01	-0.03	-0.21	0.00	-0.04	-0.77
	Selected 2	0.08	-0.01	0.00	-0.03	-0.19	-0.08	-27.42	26.62
	Selected 3	0.02	-0.05	0.01	-0.03	-0.28	0.07	0.05	-0.86
Unequal 2	Selected 1	0.16	0.02	-0.02	-0.07	-0.23	-0.18	< -100	> 100
	Selected 2	0.02	-0.05	0.00	-0.03	-0.28	0.07	0.05	> 100
	Selected 3	-0.11	-0.19	-0.03	-0.16	-0.48	0.16	0.15	-0.95
Unequal 3	Selected 1	-0.09	-0.16	-0.02	-0.10	-0.42	0.15	0.14	-0.95
	Selected 2	-0.01	0.00	-0.03	-0.12	-0.35	-0.19	-45.82	45.01
	Selected 3	0.08	-0.01	0.00	-0.03	-0.16	-0.08	-27.42	26.62
Mean		0.04	-0.06	-0.01	-0.07	-0.29	-0.01	-9.11	-10.86
SD		(0.10)	(0.08)	(0.02)	(0.05)	(0.11)	(0.13)	(15.67)	(16.91)

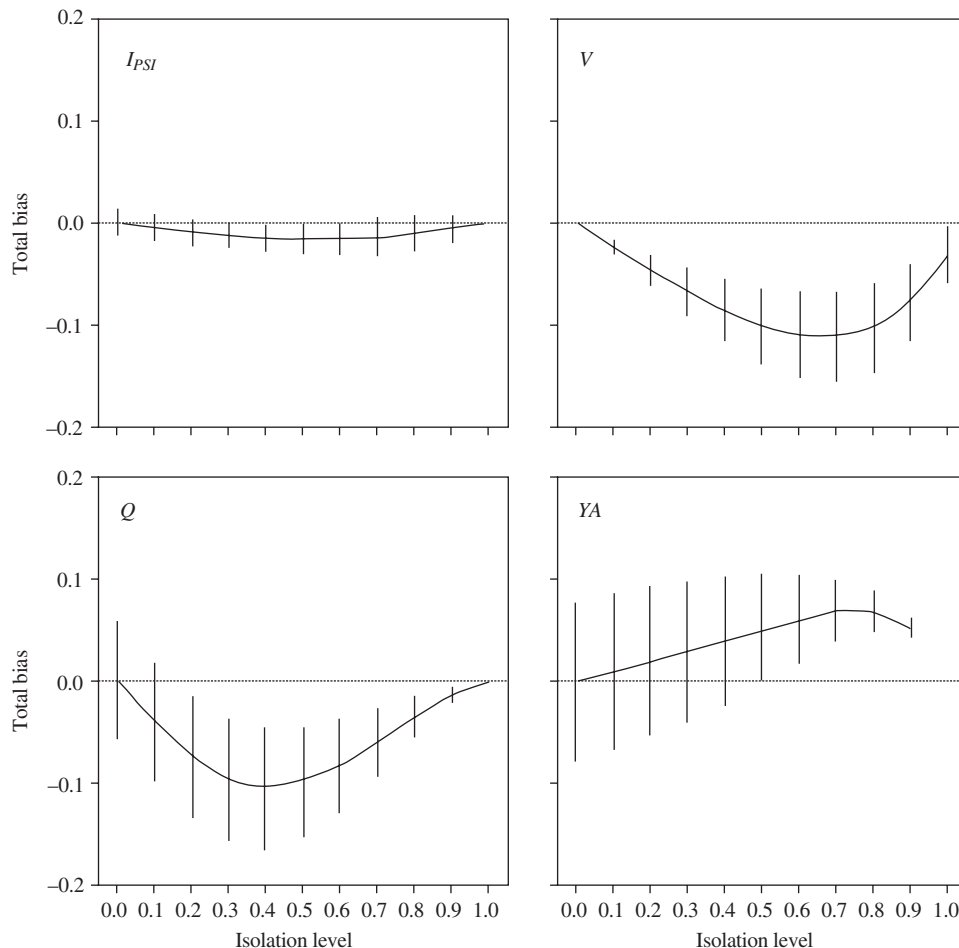


Figure 3. Distribution of total bias (asymptotic and resampling biases) for the studied sexual isolation estimators over 176 different scenarios under the laboratory model.

Overall, the results suggest that I_{PSI} is the best choice for small sample sizes (Figs 2, 3), whereas YA and Q (and, to a somewhat lesser extent, I_{PSI}) are the best choices for infinite sample sizes (Fig. 1). In order to see how the bias evolves at different sample sizes, we repeated the analyses for the laboratory model using 20, 40, 60, 80 or 120 mating pairs. The results are presented in Figure 4 and show that I_{PSI} and V do not change their total bias because of changes in sample size, although the latter showed a 3.5-fold larger bias than the former. Biases for YA and Q were reduced with larger sample sizes. However, at biological sample sizes (say less than 100 mating pairs) I_{PSI} remains the safest choice.

We also investigated other sampling properties of the statistics, namely the efficiency (inverse of the estimator variance), and the root mean squared deviation (RMS), which combines the bias and the sampling variance of the indices. For most indices, efficiency showed a trend qualitatively similar to that for RMS, so only results for the latter are shown. In the

laboratory model (Fig. 5), the smallest RMS deviations were found for V at low and intermediate values of sexual isolation, followed by I_{PSI} . YA and Q showed RMS deviations up to six times larger than those of V . χ^2 , I , Co and Ti always showed larger RMS deviations than the former indices (not shown). The results for RMS regarding the natural model were basically identical to those shown in Figure 5, except that Q and YA were more sensitive to mating propensity effects (not shown). In summary, the best indices regarding these sampling properties in general were I_{PSI} and V .

Finally, we compared the statistical test associated with the best four estimators (V , Q , YA and I_{PSI}) at two different sample sizes. Figure 6 shows the true type I error for an α of 0.05 with the parametric test for different degrees of sexual isolation across different scenarios under the laboratory model. In all cases the true type I error was large (> 0.1) and this happened at both reduced and relatively large sample sizes. This points towards a relatively poor performance of all these parametric tests, because a type I error higher

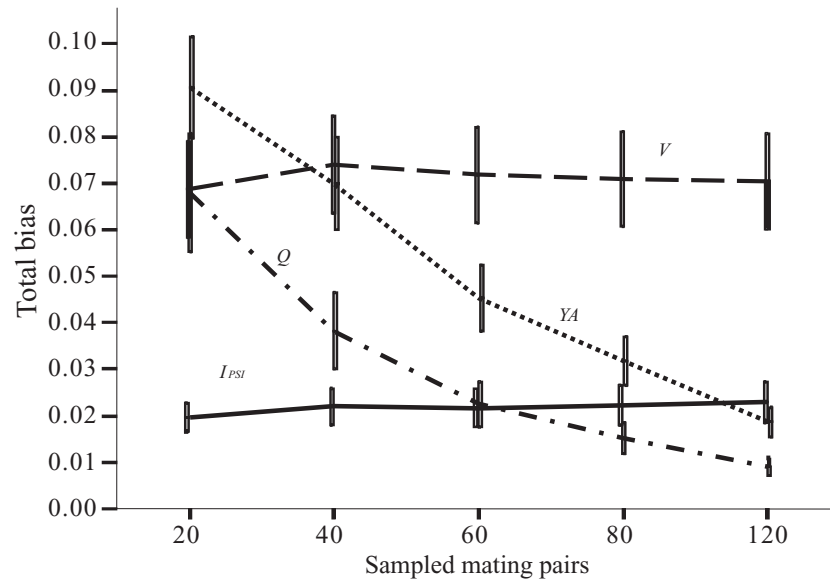


Figure 4. Evolution of total bias (asymptotic and resampling biases) for different sample sizes for some sexual isolation estimators over 176 different scenarios under the laboratory model.

than that determined a priori would produce a common erroneous rejection of the null hypothesis when it is in fact true. The most extreme case is seen with Q , for which a parametric approach should be definitively avoided.

I_{PSI} and V showed rather symmetrical bias distributions as well as the best sampling properties (see Figs 1, 2). Therefore, they can be considered the most appropriate for statistical inference based on bootstrapping. We were able to estimate the true type I error ($\alpha = 0.05$) of these statistics when using a bootstrap test of hypothesis ($I_{PSI} = 0.007$; $V = 0.001$), which suggests that this approach is more appropriate (although more conservative) than the parametric tests available.

DISCUSSION

A traditional method of studying mating behaviour under laboratory conditions involves the employment of multiple choice (Merrel, 1950). It is considered to provide an experimental set-up close to that occurring in nature. Males and females of at least two forms (genotypes, ecotypes or incipient species) are placed together and mating monitored for a certain period (Knoppin, 1985). The pattern of mating can be used to measure both sexual isolation and sexual selection effects with a degree of statistical independence (Merrel, 1950; Rolán-Alvarez & Caballero, 2000).

Sexual selection (also known in this context as mating propensity) can be considered as contributing to overall fitness (Andersson, 1994), and is estimated

from multiple choice experiments by comparing the frequency of mated and unmated individuals (Merrel, 1950; Hartl & Clark, 1989). Sexual isolation can be defined as the deviation from random mating in mated individuals, which leads to assortative (disassortative) mating when similar (opposite) phenotypes mate more often than expected at random (Merrel, 1950; Lewontin, Kirk & Crow, 1968; Spieth & Ringo, 1983).

A classical strategy for the statistical partition of these two effects is based on a chi-square or G -test partition of mating components (Merrel, 1950) and, more recently, by the use of PSS (pair sexual selection) and PSI (pair sexual isolation) estimators for each pair combination (Rolán-Alvarez & Caballero, 2000). This partitioning has an evolutionary justification, because these two components have different evolutionary consequences: sexual selection changes gene frequencies producing microevolution, whereas sexual isolation is directly involved in speciation processes (Lewontin *et al.*, 1968).

Sexual selection and sexual isolation (effects) can be originated, although not necessarily, by different mechanisms. In laboratory multiple choice experiments, the former can appear due to intrasexual competition (mate propensity) or intersexual choice (mate choice), whereas the latter is mainly caused by mate choice. However, in the wild other mechanisms may also contribute to this. Therefore, the causes of sexual isolation cannot be safely inferred directly from any single experiment on mating behaviour (Marín, 1991; Casares *et al.*, 1998).

Here we have reviewed, so far as we know, all the statistics of sexual isolation used in the literature under a wide range of different biological scenarios, in which the estimation of the mate choice coefficient

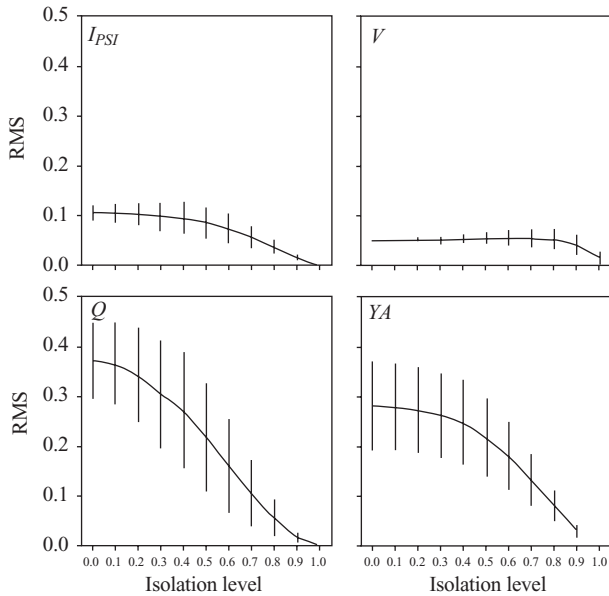


Figure 5. Distribution of the root mean square deviation (RMS: squared bias plus resampling variance) for the studied sexual isolation estimators over 176 different scenarios under the laboratory model.

can be confounded by two different processes: asymmetrical marginal frequencies and differential mating propensities. The first can be corrected in the laboratory by using a balanced experimental design, but it may provide a fundamental estimation problem for descriptive or experimental work conducted in the wild. This can be of particular importance, for example, when the two taxa being studied vary considerably in their relative population densities. The second process is very difficult to rule out, unless a very well-known experimental model species is being used.

A summary of the biases incurred by the different statistics under the scenarios investigated is given in Table 5. Some of the estimators (χ^2 and T_i) should be straightforwardly avoided, as they are significantly biased by asymmetrical marginal frequencies and differential mating propensities at any sample size (see also Gilbert & Starmer, 1985; Rolán-Alvarez & Caballero, 2000). They are so sensitive to this source of bias that any study based on them should be approached with caution or repeated using a more appropriate estimator of sexual isolation (see Figs 1, 2).

Co and Me show typically moderate, but occasionally very large, biases due to marginal frequencies and differential mating propensities. In the natural model they present extreme downward bias and thus should not be used for studies in the wild (Fig. 2B). In addition, for half of the scenarios in the laboratory they

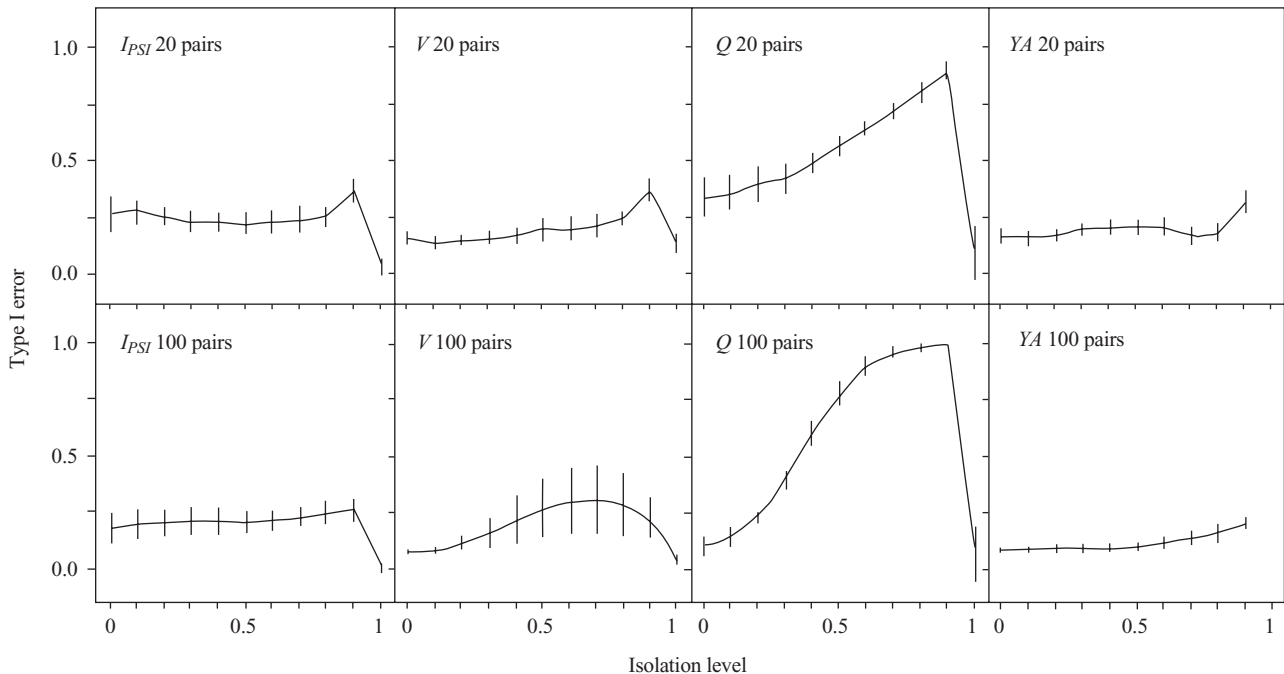


Figure 6. Type I error for some sexual isolation indices under different degrees of isolation. The bars represent the 95% confidence intervals for the distribution over the 16 different combinations of scenarios from Table 3 under a laboratory model.

Table 5. Summary of the performance of sexual isolation statistics under different scenarios and sample sizes. Low (< 0.05), Moderate (0.05–0.25), High (0.25–1), Very High (> 1)

Sexual isolation statistics	Total bias			
	Scenarios		Sample size	
	Unequal morph frequencies	Mate propensity	Small	Large
YA	moderate	moderate	moderate	low
Q	moderate	moderate	moderate	low
I_{PSI}	low	low	low	low
V	moderate	moderate	moderate	moderate
χ^2	high	high	high	high
I	high	moderate	high	high
Co	very high	very high	very high	very high
Ti	very high	very high	very high	very high

show moderate to large biases at any sample size (Figs 1, 2A).

Co has been used in well known reviews of mating behaviour in *Drosophila* (Coyne & Orr, 1989, 1997). These authors found different patterns regarding reproductive isolation in different species; one of the most relevant trends detected was the observation that prezygotic reproductive isolation was larger among sympatric than allopatric species. This has been considered to be one of the clearest supports for the reinforcement hypothesis (Noor, 1999; Turelli *et al.*, 2001; but see Endler, 1989). The analyses used balanced marginal frequencies in laboratory experiments. However, the existence of mate propensity effects cannot be ruled out systematically in all the species tested, and even low to moderate biases (due to confounding mate propensity in the experiments) could affect the significance of the observed trends.

The magnitude of any possible bias can be checked by reanalysing Coyne & Orr's (1989, 1997) data with a more reliable statistic, such as I_{PSI} . This would serve to verify that the conclusions of the analysis are not flawed by biases in the estimates. Such reanalyses for the works of Hollocher *et al.* (1997) and Arnold *et al.* (1996), which used sexual isolation estimators possibly biased by mate propensity, have been reported by Rolán-Alvarez & Caballero (2000) and Rolán-Alvarez (2004), respectively.

The joint isolation index (I) is one of the most commonly used estimators in the literature, even if it should not be used with asymmetrical marginal frequencies (see Merrel, 1950; Gilbert & Starmer, 1985; Marín, 1991) and/or large mating propensities (Gilbert & Starmer, 1985). The index does not function correctly for experiments conducted in the wild (Fig. 2B), and even employing a multiple choice labo-

ratory design shows moderate to large biases for 50% of scenarios (Fig. 2A). It is, however, intuitive and simple, and this may be the reason for its prevalence in the literature. For safe estimation it should be used in the PSI coefficients rather than on the raw data.

Yule's V shows quite good sampling properties (Fig. 5), but fails somewhat regarding its asymptotic properties (Figs 1–4). The estimates are slightly biased downwards on average, but it can be still considered an useful alternative (Gilbert & Starmer, 1985). It has been used in descriptive studies conducted in the wild (Rolán-Alvarez *et al.*, 1999), although in such circumstances the estimate is expected to be slightly biased downwards (Fig. 2B).

The YA (and related DI , W_i and Z_i) and Q statistics do not show asymptotic bias. These indices are the best alternatives if the sample size is larger than 100 mating pairs (Fig. 4) (Gilbert & Starmer, 1985). However, they suffer from upward or downward biases, both for the laboratory and natural models at biological sample sizes (Figs 2, 4) (Goux & Anxolabehere, 1980), or show estimation problems (Knoppin, 1985). DI , W_i and Z_i , however, should be avoided because they do not correct for zero values within cells and therefore show extreme total biases at biological sample sizes. YA also shows some bias at small sample sizes (Fig. 4). The sampling properties of Q are slightly better than those of YA and related indices, but the parametric test available for Q described by Gilbert & Starmer (1985) should be avoided due to its large type I error.

Finally, I_{PSI} has a very small asymptotic bias (a maximum of 9%; Fig. 1) and it is the least biased alternative at reduced sample sizes under both laboratory and natural models (Figs 2, 3). These useful properties are maintained at any sample size, although YA and Q work slightly better for more than 100 pairs (Fig. 4).

In general, I_{PSI} can be considered as the safest alternative for estimating sexual isolation caused by differential choice coefficients. However, the parametric test available for I (and consequently for I_{PSI}) shows relatively high type I errors, and thus bootstrapping should be preferable to parametric inference.

PSI coefficients were also included in the other statistics (*Yule's Q* and *YA*), but they performed slightly less well (in terms of average and variance of bias) than I_{PSI} with reduced sample sizes. In addition, *Yule's Q* and, analogously, T_i , χ^2 and Co , do not show a direct relationship with the degree of isolation. For example, when the choice coefficients of C_{AB} and C_{BA} are half the value of C_{AA} and C_{BB} in Table 1 and therefore the asymptotic sexual isolation estimate for I_{PSI} is exactly 0.5 for equal marginal frequencies and no mating propensity effects, *Yule's Q* has a value of 0.8. The joint isolation index on PSI is a more intuitive statistic and should be preferred over *Yule's Q*, *YA* and related indices.

Population geneticists have traditionally broken down mating behaviour into two statistically independent mating components: sexual selection (mating propensity) and sexual isolation (Merrel, 1950). However, because these two effects have different evolutionary consequences and can also be caused (although not necessarily) by different evolutionary mechanisms, it is convenient to estimate them separately in practice. A possible strategy would be to estimate the effects first and then infer their causes later (Rolán-Alvarez & Caballero, 2000; Cruz *et al.*, 2004). An alternative method might involve the development of precise hypotheses of a series of factors interacting to produce the observed levels of sexual isolation, and then testing them by chi-square or likelihood methods (see O'Donald, 1980; Tregenza *et al.*, 2000; Tregenza, 2002).

ACKNOWLEDGEMENTS

We are grateful to an anonymous referee for comments. We thank the European Commission (EUMAR project EVK3-CT-2001-00048), Ministerio de Educación y Ciencia y Fondos FEDER (BMC 2003-03022 and CGL2004-03920/BOS), Xunta de Galicia and Universidade de Vigo for financial support. A. P.-F. was supported by a Fellowship from Ministerio de Ciencia y Tecnología.

REFERENCES

- Andersson M.** 1994. *Sexual selection*. Princeton, NJ: Princeton University Press.
- Arnold SJ, Verrell PA, Tilley SG.** 1996. The evolution of asymmetry in sexual isolation: a model and a test case. *Evolution* **50**: 1024–1033.

- Casares P, Carracedo MC, Del Ría B, Piñeiro R, García-Florez L, Barros AR.** 1998. Disentangling the effects of mating propensity and mating choice in *Drosophila*. *Evolution* **52**: 126–133.
- Coyne JA, Orr HA.** 1989. Patterns of speciation in *Drosophila*. *Evolution* **43**: 362–381.
- Coyne JA, Orr HA.** 1997. Patterns of speciation in *Drosophila* revisited. *Evolution* **51**: 295–303.
- Cruz R, Carballo M, Conde-Padín P, Rolán-Alvarez E.** 2004. Testing alternative models for sexual isolation in natural populations of *Littorina saxatilis*: indirect support for by-product ecological speciation? *Journal of Evolutionary Biology* **17**: 288–293.
- Dobzhansky TH.** 1937. *Genetics and the origin of species*. New York: Columbia University Press.
- Efron B.** 1982. *The jackknife, the bootstrap and other resampling plans*. Philadelphia, PA: Society for Industrial and Applied Mathematics.
- Ehrman L, Pettit C.** 1968. Genotype frequency and mating success in the Willistoni species group of *Drosophila*. *Evolution* **22**: 649–658.
- Endler JA.** 1989. Conceptual and other problems in speciation. In: Otte D, Endler JA, eds. *Speciation and its consequences*. Sunderland, MA: Sinauer, 625–648.
- Gilbert DG, Starmer WT.** 1985. Statistics of sexual isolation. *Evolution* **39**: 1380–1383.
- Goux JM, Anxolabehere D.** 1980. The measurement of sexual isolation and selection: a critique. *Heredity* **45**: 255–262.
- Hartl DL, Clark AG.** 1989. *Principles of population genetics*, 2nd edn. Sunderland, MA: Sinauer.
- Hollocher H, Ting C-T, Pollack F, Wu C-I.** 1997. Incipient speciation by sexual isolation in *Drosophila melanogaster*: variation in mating preference and correlation between sexes. *Evolution* **51**: 1175–1181.
- Knoppien P.** 1985. Rare male mating advantage: a review. *Biological Reviews* **60**: 81–117.
- Lewontin R, Kirk D, Crow J.** 1968. Selective mating, assortative mating, and inbreeding: definitions and implications. *Eugenics Quarterly* **15**: 141–143.
- Marín I.** 1991. Sexual isolation in *Drosophila* I. Theoretical Models for multiple-choice experiments. *Journal of Theoretical Biology* **152**: 271–284.
- Merrel DJ.** 1950. Measurement of sexual isolation and selective mating. *Evolution* **4**: 326–331.
- Noor MAF.** 1999. Reinforcement and other consequences of sympatry. *Heredity* **83**: 503–508.
- Nooren EW.** 1989. *Computer intensive methods for testing hypothesis: an introduction*. New York: Wiley.
- Nosil P, Crespi BJ, Sandoval CP.** 2002. Host-plant adaptation drives the parallel evolution of reproductive isolation. *Nature* **417**: 440–443.
- O'Donald P.** 1980. *Genetic models of sexual selection*. Cambridge: Cambridge University Press.
- Ringo JM, Dowse HB, Lagasse S.** 1986. Symmetry versus asymmetry in sexual isolation experiments. *Evolution* **40**: 1071–1083.
- Rolán-Alvarez E.** 2004. Evolution of asymmetry in sexual

- isolation: a criticism of a test case. *Evolution Ecological Research* **6**: 1099–1106.
- Rolán-Alvarez E, Caballero A. 2000.** Estimating sexual selection and sexual isolation effects from mating frequencies. *Evolution* **54**: 30–36.
- Rolán-Alvarez E, Erlandson J, Johannesson K, Cruz R. 1999.** Mechanisms of incomplete prezygotic reproductive isolation in an intertidal snail: testing behavioural models in wild populations. *Journal of Evolutionary Biology* **12**: 879–890.
- Schlutter D. 2001.** Ecology and the origin of species. *Trends in Ecology and Evolution* **16**: 372–380.
- Sokal RR, Rohlf FJ. 1997.** *Biometry*. New York: WH Freeman.
- Spieth HT, Ringo JM. 1983.** Mating behavior and sexual isolation in *Drosophila*. In: Ashburner M, Carson HL, Thompson JN, eds. *The genetics and biology of Drosophila*. New York: Academic Press, 223–284.
- Tilley SG, Verrel PA, Arnold SJ. 1990.** Correspondence between sexual isolation and allozyme differentiation: a test in salamander *Desmognathus ochrophaeus*. *Proceedings of the National Academy of Sciences, USA* **87**: 2715–2719.
- Tregenza T. 2002.** Divergence and reproductive isolation in the early stages of speciation. *Genetica* **116**: 291–300.
- Tregenza T, Pritchard V, Butlin RK. 2000.** The origins of premating reproductive isolation: testing hypotheses in the grasshopper *Chorthippus parallelus*. *Evolution* **54**: 1687–1698.
- Turelli M, Barton NH, Coyne JA. 2001.** Theory and speciation. *Trends in Ecology and Evolution* **16**: 330–343.