Non-random mating and information theory

A. Carvajal-Rodríguez

Departamento de Bioquímica, Genética e Inmunología. Universidad de Vigo, 36310
Vigo, Spain.

*: A. Carvajal-Rodríguez. Departamento de Bioquímica, Genética e Inmunología.

Universidad de Vigo, 36310 Vigo, Spain. Phone: +34 986813828

email: acraaj@uvigo.es

Keywords: Mate choice, sexual selection, sexual isolation, Price equation, Kullback-Leibler divergence, population genetics.

Abstract

In this work, mate choice is modeled by means of the abstract concept of mutual mating propensity. This only assumes that different type of couples can have different mating success. The model is adequate for any population where mating occurs among distinct types. There is no extra assumption about particular mating scheme or preference model. The concept of mutual mating propensity permits to express the observed change in the mating phenotypes as the gain in information with respect to random mating. The obtained expression is a form of the Price equation in which the mapping between ancestral and descendant population is substituted by a mapping between random mating and non random mating population.

At the same time, this framework provides the connection between mate choice and the exact mathematical partition of the choice effects, namely sexual isolation, sexual selection and a mixed effect. The sexual selection component is the sum of the intra-sexual male and female selection.

The proposed framework helps to unveil previously hidden invariants. For example, if the mutual preference between partner types is multiplicative there is no sexual isolation (inter-sexual selection) effect on the frequencies, i.e. the only possible effect of mate choice is intra-sexual selection. On the contrary, whatever the contribution of each partner to the mutual preference, if it comes as a non-multiplicative factor, there is at least an inter-sexual selection detectable effect.

This new view over the mate choice problem, permits to develop general mating propensity models and to make predictions of the mate choice effects that may emerge from such models. This possibility opens up the way for setting a general theory of model fitting and multimodel inference for mate choice.

24 Thus, it is suggested that the proposed framework, by describing mate choice as the
25 flow of information due to non-random mating, provides a new important setting for
26 exploring different mating models and their consequences.

27

1. Introduction

Mate choice is arguably one of the most active areas of evolutionary research. There has been a lot controversy regarding the concept of mate choice. The debate around mate choice was due in part to its importance for fields so diverse as population genetics, evolutionary-ecology, animal behavior, sociology, or psychology. In addition, there has been an excess of verbal models and imprecise terminology regarding different aspects of mate choice (Edward, 2015). Mate choice can be broadly described as the effect of some expressed traits leading to non-random mating. Under this broad definition there are various aspects that can be considered. Yet Darwin (1871) distinguishes between intrasexual selection and intersexual selection. The first arises directly from competition among individuals of the same sex while the second arises from choice of mates by the other sex (Kuijper et al., 2012). Alternatively, from a population genetics point of view, mate choice is defined as the observed mating frequency deviation with respect to random mating, considering population gene or phenotype frequencies. So defined, mate choice can be partitioned into (intra)sexual selection, defined as the observed change in gene or phenotype frequencies in mated individuals with respect to population frequencies, and sexual isolation (behavioral isolation or intersexual selection), defined as the deviation from random mating in mated individuals (Rolán-Alvarez and Caballero, 2000). In this work I followed these definitions of mate choice, intrasexual and intersexual selection.

For an alternative description of these concepts and a discussion about some of the most widely used descriptions of evolutionary change within the context of sexual selection, I refer the reader to (Kuijper et al., 2012; Rosenthal, 2017).

The many aspects and complexity of mate choice justifies the extensive research that has been made in the last decades producing several theoretical models and empirical tests. Related to modeling and detection of mate choice, there is the question about the correct null hypothesis for testing the evolution of mate choice. The Lande-Kirpatrick (L-K) model has been proposed as a null model (Kirkpatrick, 1982; Lande, 1981; Prum, 2010; Roff and Fairbairn, 2014). This model assumes neutral genetic variation for the mating preference trait while the target trait can be under natural selection. However, the L-K role as a null model is not clear when the preference is set by similarity (preference and target trait coincide) and the trait is under divergent selection (Servedio et al., 2011), i.e. the trait is "magic" sensu Gavrillets (2004), because in this case the preference trait is already under selection (Hughes, 2015).

Therefore, there is still a need for both, null models and a general framework, where the key essential facts of the mate choice can be adequately described. Here, I argue that the formalism provided by the information theory in the form of the Jeffreys' divergence is the right tool to do so.

The information theory has already been elegantly applied for describing evolutionary change (Frank, 2009; Frank, 2012b; Frank, 2013). The present work takes advantage of that mathematical structure and applies it for modeling the change in mating frequencies due to mate choice. As far as I know there is no previous attempt of describing mate choice from the viewpoint of the information theory. Nevertheless, the potential of the informational view for evolutionary ecology has been already suggested (Dall et al., 2005).

First, I defined a general model that only requires an abstract functional relationship connecting the observed mating frequencies with the expected by random mating from

the population gene or phenotype frequencies. This suffices for developing a general information equation for mate choice that can be adequately partitioned into intrasexual and intersexual information components, plus a mixed term provoked by the confounding effect of the marginal frequencies when the mating propensity effects are asymmetric. Interestingly, the three terms can be separately estimated from the observed frequencies and so, the researcher can study how different models and parameters translate into the different mate choice components. Also, it is proposed that this setting provides the baseline for solving the mate choice null hypothesis problem, since the null model emerges naturally from the idea of zero information. Thus, the correct null should not rely on neutral preference or trait genes but on zero information.

The utility of this framework is shown by analyzing a real data example. I will show how the view obtained from the unveiled relationships can be utilized to classify different general models from its consequences which facilitates the multimodel inference of the mate choice. However, a deeper study on the outcomes of different forms of the mating preference functions is out of the scope of the present article and is part of a different paper.

2. Model of Mate Choice

As mentioned above, the following model is as a particular specification of the information theory interpretation for evolutionary models, proposed in (Frank, 2012b; Frank, 2013). The general framework developed by this author fits perfectly for the purpose of describing the occurrence of non-random mating and the flow of information that it provokes. Remarkably, once the basic equation for the gain in information due to non-random mating is formalized, the relationship between mate choice and its different

evolutionary outcomes emerges naturally, providing a clear and useful picture of the intrasexual and intersexual selection effects.

2.1 General model

Let consider a population with a number of n_1 females and n_2 males. For a given female phenotype X (e.g. shell color) with K different classes having values $X_1, X_2 \dots X_k$, the frequency of the phenotype X_i in the female population is $p_{1i} = n_{1X_i} / n_1$, i.e. the number of females with that phenotypic value divided by the total number of females. Similarly, for the male phenotype Y (could be the same as X) with K' classes, the frequency of Y_j in the male population is $p_{2j} = n_{2Y_j} / n_2$.

In this way, by using the frequency of the phenotype for each sex, the expected mating frequencies if mating is at random is

$$q_{ij} = p_{1i} \times p_{2j}.$$

Now, given a female phenotype X_i and a male phenotype Y_j , let's define the mutual mating propensity $m_{ij}(x, y, e)$ as the number of matings of X_i with Y_j after their encounter in the environment e . The normalized mating propensity is

$$m_{ij}(x, y, e) / M$$

where $M = \sum_{i,j} q_{ij} m_{ij}(x, y, e)$.

Then, the observed mating frequencies in a given environment e can be expressed as

$$q'_{ij} = q_{ij} \frac{m_{ij}(x, y, e)}{M} \quad (1)$$

Therefore, the observed mating frequencies are the result of the functions $m_{ij}(x,y,e)$ (hereafter noted as m_{ij}), that can be any kind of composition of the preference of female X_i for male Y_j , and vice versa, in the environment e .

Note that random mating is a particular case of the model in (1) when the propensities are equal for every mating pair. The mutual mating propensity functions can represent empirical or analytical functions, as for example the Gaussian-like preference functions (reviewed in Carvajal-Rodriguez and Rolán-Alvarez, 2014). Moreover, each m_{ij} can be composed of female and male preferences, so mutual mate choice models (Bergstrom and Real, 2000) are also available under this setting. The standardized m_{ij} values could also be estimated a posteriori from the data. In this case they coincide with the pair total index i.e. the ratio of the frequency of the observed types divided by the expected pair types calculated from the total frequencies ($PTI_{ij} = q'_{ij}/q_{ij}$, Rolán-Alvarez and Caballero, 2000) which becomes an observation of the mutual mating propensity from the mating phenotypes (see below).

Once we have the mating frequencies as defined in (1), the change with respect to random mating is

$$\Delta q_{ij} = q'_{ij} - q_{ij} = q_{ij} \left(\frac{m_{ij}}{M} - 1 \right)$$

The mean population change for a combined phenotype $Z = X * Y$ is

$$\Delta Z = \sum_{i,j} \Delta q_{ij} Z_{ij}$$

Because the relationship in (1) is defined by ratios is more natural to express the quantities in the logarithmic scale and so we can express m_{ij} as

$$m_{ij} = M \frac{q'_{ij}}{q_{ij}}$$

which in the logarithmic scale becomes

$$l_{ij} = \log(m_{ij}) = \log(M) + \log\left(\frac{q'_{ij}}{q_{ij}}\right)$$

Thus, if we take the logarithm of the propensity as the combined phenotype Z and by noting that $\sum \Delta q_{ij} = 0$ and that $\log(M)$ is constant through the summation, then we can measure the mean population change ΔL in relative propensity as

$$\Delta L = \sum_{i,j} \Delta q_{ij} l_{ij} = \sum_{i,j} \Delta q_{ij} \log\left(\frac{q'_{ij}}{q_{ij}}\right) = J(q', q) \equiv J_{PTI} \quad (2)$$

which is the Kullback-Leibler symmetrized divergence (noted as Jeffreys in Frank, 2012b), that measures the gain in information when the differential mating propensity moves the population from mating frequencies q to q' or vice versa. Note that if the propensity is equal for every pair i.e. $M = m_{ij} \forall i, j$ then $q' = q$ so that $J = 0$ which is the minimum information value since J cannot be negative.

Recall from equation (1), that each m_{ij}/M is the ratio of the frequency of the observed types divided by the expected pair types from the total frequencies. This is, by definition, the pair total index PTI (Rolán-Alvarez and Caballero, 2000) and so, the logarithmic term in ΔL is the logarithm of the PTI values. Therefore, $J(q', q)$ measures the gain in information as captured by the PTI coefficients, confronting the hypothesis of mate choice against random mating. Hereafter, we note this J as J_{PTI} .

Interestingly enough, the Jeffreys' divergence computed as J_{PTI} (by taking the natural logarithm and multiplying (2) by the total number of matings) is well approximated by a

159 chi-square for the null hypothesis of random mating with $KK'-1$ degrees of freedom
160 (Evren and Tuna, 2012).

161 The information obtained from J_{PTI} has been computed using the different propensities
162 as classes for classifying the couples i.e. we considered $\log(m)$ as the phenotype Z .
163 When the classes are based upon other phenotypes rather than propensities, we are
164 conveying a specific meaning for the change in frequencies, say, the change in mating
165 frequencies due to differential mutual propensities is observed in terms of change in
166 shell color mating frequencies. Therefore, the phenotype can be viewed as other scale
167 on which we can measure this information (Frank, 2013). Of course, different kinds of
168 phenotypes can be more or less involved in mate choice and so, different scales are
169 more or less useful for measuring the mate choice information.

170

171 *2.2 Relative propensity and phenotypes*

172 When we observe any mating pair (i,j) , we need to identify the mating by a given
173 characteristic (e.g. shell color) since we cannot directly classify it by the value of the
174 propensity function m_{ij} . In general, we ignore the specific form of the mutual mating
175 propensity function m and so, we may assume that some phenotype matches it perfectly,
176 as we did above (each phenotypic pair was perfectly differentiated by specific m_{ij}
177 mating propensity).

178 Thus, if T is the trait that is the target of the choice, we call J_{PTI} to the change in the
179 numbers of matings when these matings were classified by T .

180

Therefore, we may think on different traits Z that classify the mating pairs; Z can be a composition of female trait X , e.g. preference, and male target Y , or can be any kind of different traits or alternatively the same trait in both sexes as size, age or color. In any case, we measure the mean change in Z caused by differences in m , as

$$\Delta_m Z = \sum_{\alpha, \beta} \Delta q_{\alpha\beta} Z_{\alpha\beta} = \sum_{\alpha, \beta} q_{\alpha\beta} (m_{\alpha\beta} - M_z) \frac{Z_{\alpha\beta}}{M_z} = \frac{\text{cov}(Z, m_z)}{M_z} \quad (3)$$

Where cov is the population covariance in the sense of Price (1972) as highlighted in Frank (2012a). The subscripts α, β emphasize that we are looking at pairs with observed phenotypes that not necessarily are the phenotypes exactly connected with the choice. Therefore, the propensities for the matings classified under these phenotypes can be different to the propensities for the trait T , then we note m_z and M_z to distinguish from the propensities (m) and mean propensity (M) measured directly from the real choice trait.

Equation (3) is in fact, a form of the Price equation with a different mapping for the populations involved. While the Price equation (Frank, 2012a; Price, 1972) describes the change in phenotype between two connected ancestor and descendant populations; in our equation (3), the mapping is between the random mating population and the one obtained under a given mutual mating propensity scheme.

The variable Z can be any desired trait including, as we assumed above, the logarithm of the propensities. So, if we take Z equal to the logarithm of m , then by substituting in (3) we obtain the mean population change ΔL_z as

$$\Delta L_z = \text{cov}(m_z, l_z) / M_z$$

where the subscript z indicates that the propensities are now indexed by the trait Z .

203 Recalling the relationship in (2), we now define

$$204 \quad JZ_{PTI} \equiv \Delta L_z = cov(m_z, l_z) / M_z = \alpha_z J_{PTI}$$

205 where if $J_{PTI} > 0$

$$206 \quad \alpha_z = (\beta(m_z, l_z) \times D_z) / (\beta(m, l) \times D)$$

207 with $l_z = \log(m_z)$, $D_z = V(m_z)/M_z$, $l = \log(m)$ and $D = V(m) / M$

208 or $\alpha_z = 1$ if $J_{PTI} = 0$.

209 Note that D and D_z are the indexes of dispersion over the choice and Z traits

210 respectively, so α_z is the quotient of the regressions multiplied by the index of

211 dispersion at each phenotypic scale.

212 From the point of view of the estimation with real data, if we cannot measure directly

213 the values of m then we simply compute J based on trait Z and therefore we are really

214 computing JZ_{PTI} .

215 In this case, note that the PTI coefficients are no longer the exact estimate of the mutual

216 mating propensities because the ratio of frequencies $q'_{\alpha\beta}/q_{\alpha\beta}$ does not correspond to

217 m_{ij}/M but to $m_{\alpha\beta}/M_z$ which is a proxy that would be more or less precise depending on

218 the importance of the measured phenotype over the mating choice. For example, if shell

219 size is driving mate choice, the measure of JZ_{PTI} (Z = shell size) would correspond well

220 with J_{PTI} ($J_{PTI} > 0$; $\alpha_z \approx 1$). However, if other phenotype as shell color has nothing to do

221 with mate choice (and is not correlated with shell size) then the measure of JZ_{PTI} (Z =

222 shell color) would be zero ($J_{PTI} > 0$; $\alpha_z = 0$). Further details about the distinction

223 between JZ_{PTI} and J_{PTI} are given in appendix A.

The mate choice mediated by the differences in mutual mating propensity would produce a deviation from random mating. At the same time, this may cause two different effects, namely, intrasexual selection and intersexual selection, hereafter noted as sexual selection and sexual isolation, respectively.

2.3. Sexual selection

Sexual selection is defined as the observed change in gene or phenotype frequencies in mated individuals with respect to total population frequencies (Rolán-Alvarez and Caballero, 2000). This change can be studied using the frequencies within each sex, or considering jointly both sexes, by using the pair sexual selection coefficient (*PSS*, Rolán-Alvarez and Caballero, 2000). I will show that, when the *PSS* coefficients are considered a priori as the marginal propensities for the mating types, the information gained due to sexual selection is the sum of the information from each sex. When the focus is on the phenotypes instead on the propensities, the partition continue to be true, provided that the same phenotypic scale is applied when computing the *PSS* coefficients and the intrasexual components.

From the general model, the population frequency of the female phenotype X_i is p_{1i} . The observed frequency of X_i in mated individuals, p'_{1i} , is computed as the sum of the mating frequencies involving a female X_i

$$p'_{1i} = \sum_j^{males} q'_{ij} = p_{1i} \sum_j p_{2j} \frac{m_{ij}}{M} = p_{1i} \frac{m'_{fi}}{M}$$

where m'_{fi} is the marginal mating propensity for the female type i .

Similarly for males, the frequency of phenotype Y_j is p_{2j} and the frequency for the male type j in mated individuals is

$$p'_{2j} = p_{2j} \frac{m'_{mj}}{M}$$

where m'_{mj} is the marginal mating propensity for the male type j .

The mean change in information due to sexual selection within each sex is, in terms of the female marginal propensity (female intrasexual selection)

$$\Delta_m L_x = \sum_i^{females} \Delta p_{1i} X_i = \sum_{ij} p_i (m'_{fi} - M) \frac{X_i}{M} = J(p'_1, p_1) \equiv J_{S1}$$

and, in terms of male marginal propensity (male intrasexual selection)

$$\Delta_m L_y = \sum_j^{males} \Delta p_{2j} Y_j = J(p'_2, p_2) \equiv J_{S2}$$

The term J_S has been obtained in a similar way as for the general case, i.e. by expressing each marginal m'_{fi} and m'_{mj} in function of their respective ratio of frequencies multiplied by the mean propensity M and substituting the phenotype X or Y , by the logarithm of the corresponding (female or male) marginal m' .

The change to the scale of phenotypes produces

$$\Delta_m \bar{X} = \alpha_x J_{S1} \equiv JX_{S1}$$

with

$$\alpha_x = [\beta(m'_{xf}, l_{xf}) \times D_{xf}] / [\beta(m'_f, l_f) \times D_f] \text{ or } 1 \text{ if } J_{PS1} = 0$$

where $l_{xf} = \log(m'_{xf})$, $D_{xf} = V(m'_{xf})/M_z$, $l_f = \log(m'_f)$, $D_f = V(m'_f)/M$.

And

$$\Delta_m \bar{Y} = \alpha_y J_{S2} \equiv JY_{S2}$$

$$\alpha_y = [\beta(m'_{ym}, l_{ym}) \times D_{ym}] / [\beta(m'_m, l_m) \times D_m] \text{ or } 1 \text{ if } J_{PS2} = 0$$

$$\text{where } l_{ym} = \log(m'_{ym}), D_{ym} = V(m'_{ym})/M_z.$$

Note that the subscripts x (females) or y (males) refer to the matings classified by phenotype instead of the true choice trait, also note that the mean of both female and male marginals is the same and equal to the mean propensity (M_z or M depending on the scale).

JX_{S1} and JY_{S2} are the Jeffrey's divergence that expresses the gain of information due to intrasexual selection measured on the combined phenotypic scale Z .

2.4. Pair sexual selection

In addition to the computation within each sex, we can compare the expected pair types under random mating calculated in mated individuals, with the expected pair types from total numbers (PSS , see above). Thus, $PSS_{ij} = (p'_{1i}p'_{2j}) / (p_{1i}p_{2j}) = m'_{fi}m'_{mj} / M^2$. The latter term can be viewed as an a priori expression of the PSS coefficients. Again, the difference between the observed and the expected distribution can be expressed as

$$\Delta_{PSS}L = \sum_{i,j} \Delta(p_{1i}p_{2j}) \log \left(\frac{p'_{1i}p'_{2j}}{p_{1i}p_{2j}} \right) = J(p'_1p'_2, p_1p_2) \equiv J_{PSS},$$

$$\text{where } \Delta(p_{1i}p_{2j}) = p'_{1i}p'_{2j} - p_{1i}p_{2j}.$$

In the scale of phenotypes

$$\Delta_{PSS}L_z = \alpha_{pss}J_{PSS} \equiv JZ_{PSS}$$

with

$$\alpha_{pss} = (\beta(m'_z, l'_z) \times D'_z) / (\beta(m', l') \times D') \text{ or } 1 \text{ if } J_{PSS} = 0$$

285 where $m'_z = m'_{xf} \times m'_{ym}$, $m' = m'_f \times m'_m$, $D'_z = V(m'_z)/M_z^2$, $l'_z = \log(m'_z)$, $l' = \log(m')$, D'
 286 $= V(m')/M_z^2$.

287 The change in the phenotype due to sexual selection is driven by the aprioristic version
 288 of *PSS*, and is expressed in term of the information accumulated and rescaled from the
 289 marginal propensities to *Z*.

290 The relationship between sexual selection measured within sex and the pair sexual
 291 selection measured by *PSS* is (details in Appendix B)

$$292 \quad J_{PSS} = J_{S1} + J_{S2}$$

293 And in the scale of phenotypes

$$294 \quad JZ_{PSS} = JX_{S1} + JY_{S2} \quad (4)$$

295 provided that the same phenotypic scale applies in the pair sexual selection statistic and
 296 in the intrasexual components (i.e. the criteria utilized for classifying the different
 297 couples is the same).

298 The information captured in the *PSS* coefficients is the sum of the sexual selection
 299 within each sex.

300

301 2.5. Sexual isolation

302 Sexual isolation is defined as the deviation from random mating in mated individuals
 303 (Rolán-Alvarez and Caballero, 2000). The pair sexual isolation statistic (*PSI*) is the
 304 number of observed pair types divided by the expected pair types from mates. In terms
 305 of our model this is the ratio of frequencies

$$PSI_{ij} = q'_{ij} / (p'_{1i} p'_{2j}) = (m_{ij} / M) / (m'_{fi} m'_{mj} / M^2) \equiv \delta_{ij} \quad (5)$$

The term δ refers to an aprioristic (depends on the m 's from the model) definition of the PSI s. The joint isolation index for PSI can be expressed as

$$I_{PSI} = \frac{(k-1) \sum_i \delta_{ii} - \sum_{i,j} \delta_{ij}}{(k-1) \sum_i \delta_{ii} + \sum_{i,j} \delta_{ij}}$$

where k is the number of phenotypic classes involved in the classification of the matings (Carvajal-Rodriguez and Rolan-Alvarez, 2006).

As with the previous pairwise statistics, we may obtain the equations of change between observed and expected pair types in terms of J .

$$\Delta_{PSI} L = \sum_{i,j} \Delta(p'_{1i} p'_{2j}) \log \left(\frac{q'_{ij}}{p'_{1i} p'_{2j}} \right) = J(q', p'_1 p'_2) \equiv J_{PSI},$$

where $\Delta(p'_{1i} p'_{2j}) = q'_{ij} - p'_{1i} p'_{2j}$.

In the scale of phenotypes

$$\Delta_{PSI} L_z = \alpha_\delta J_{PSI} \equiv JZ_{PSI}.$$

The scaling factor α_δ is not always easy to compute. Provided that there is no sexual selection ($J_{PSS} = 0$) then $\alpha_\delta = \alpha_z$ and so $JZ_{PTI} = JZ_{PSI} = \alpha_z J_{PSI}$. Otherwise we need to rescale the factor E_0 (see below) to finally get the transformation between JZ_{PSI} and J_{PSI} (see Appendix C).

The JZ_{PSI} index provides the correct metric to express the part of change in mating information that translates into sexual isolation. Presenting the PSI 's under this formalism allow us to appreciate some facts that are not obvious from the a posteriori definition of coefficients estimated from data. We must realize (see equation 5) that if the normalized propensity of each pair (m_{ij} / M) is the product of the normalized

marginal types of each partner then $\delta = 1$ and so, both, the values of I_{PSI} and J_{PSI} are zero indicating no sexual isolation at all. Thus, in any model in which the mutual mating propensity is multiplicative, the only possible outcome from mate choice is intrasexual selection.

We can illustrate the multiplicative effect by means of a simple model based on a real species scenario. The bird sage grouse (*Centrocercus urophasianus*) has elaborate courtship rituals. In the spring season, males congregate in leks that are visited by the females that actively choose one of the males for mating. The number of females visiting a male seems to be related with the male long-range acoustic broadcasts whereas the probability of mating once visited is related to the visual display (Gibson, 1996).

It has been suggested that both traits, acoustic broadcast and display rate, yield a multiplicative preference for males with specific acoustic conditions and high display rates (Gibson, 1996; Rosenthal, 2017).

Thus, we can define a model of the multiplicative effect of the aforementioned traits (see details in Appendix D). Obviously, the real mating scenario is by far more complex, but the example suffices to illustrate the point.

The females are the choosers and so our model assumes a single female phenotypic class (X) and two male traits with two phenotypic classes each, B/b for acoustic broadcast, and D/d for display rate, where in both cases the upper case refers to the higher value of the trait. We define a multiplicative preference effect for acoustic broadcast and display rate, so that the female propensity for males BD can be expressed as the product of the female propensities for Bd and dB , i.e. $m_{XBD} = m_{XBd} \times m_{XbD}$ (supplementary Table S1).

Under this model, the mean propensity M coincides with the female marginal, $m'_{fX} = M$.

The four male marginal propensities (m'_{mBd} , m'_{mBD} , m'_{mbD} , m'_{mbd}) have the same values as their corresponding mutual propensities ($m_{XBd} = \alpha$, $m_{XBD} = \alpha\beta$, $m_{XbD} = \beta$, $m_{Xbd} = 1$; see Appendix D).

The model is multiplicative since each normalized mutual propensity is equal to the product of the normalized marginals, e.g. $m_{XBd}/M = (m'_{fX}/M) \times (m'_{mBd}/M)$.

By computing the aprioristic expressions in (5), we see that $\delta_{Bd} = \delta_{BD} = \delta_{bD} = \delta_{bd} = 1$.

Thus, provided that the mating reflects the propensities, the result is that independently of the phenotypes, the number of observed pair types would be equal to the expected pair types from mates, which means that there is no sexual isolation.

On the other hand, the model predicts male sexual selection whenever m_{XBd} and/or $m_{XbD} \neq 1$.

3. Relationship between Mate Choice, Sexual Selection and Sexual Isolation

The information as captured by the PTI coefficients can be partitioned in terms of PSS and PSI . Recall the expression (2) for J_{PTI}

$$\Delta L = \sum_{i,j} \Delta q_{ij} \log \left(\frac{q'_{ij}}{q_{ij}} \right) = J(q', q) \equiv J_{PTI}$$

The term Δq_{ij} can be expressed as the sum of the frequency changes for sexual selection and isolation

$$\Delta q_{ij} = \Delta(p_{1i}p_{2j}) + \Delta(p'_{1i}p'_{2j})$$

371 The logarithmic term $\log(q'_{ij}/q_{ij})$ which we have also noted as $\log(PTI)$ is also
 372 partitioned in the sexual selection and isolation components

$$373 \quad \log(PTI) = \log(PSS) + \log(PSI).$$

374 Therefore

$$375 \quad J_{PTI} = \sum_{ij} ([\Delta(p_{1i}p_{2j}) + \Delta(p'_{1i}p'_{2j})] \times [\log(PSS_{ij}) + \log(PSI_{ij})]) \text{ i.e.}$$

$$376 \quad J_{PTI} = J_{PSS} + J_{PSI} + E_0$$

377 where $E_0 = \sum_{ij} (\Delta(p_{1i}p_{2j})\log(PSI_{ij}) + \Delta(p'_{1i}p'_{2j})\log(PSS_{ij}))$. However, note that

378 $\Delta(p'_{1i}p'_{2j})\log(PSS_{ij}) = 0$ (see Appendix E) so finally

$$379 \quad E_0 = \sum_{ij} (\Delta(p_{1i}p_{2j})\log(PSI_{ij}) = \text{cov}(m'_f m'_m, \log(PSI)) / M^2$$

380 thus E_0 is proportional to the covariance between the marginal propensities and the
 381 logarithm of the $PSIs$.

382 The covariance expression is useful for defining a scaling factor (see Appendix C) i.e.

$$383 \quad \alpha_E = \beta(m'_z, l_{zpsi}) \times D'_z / \beta(m', l_{psi}) \times D' \text{ or } 1 \text{ if } E_0 = 0$$

384 where $l_{zpsi} = \log(PSI_z)$, $l_{psi} = \log(PSI)$ and m'_z , m' , D'_z and D' are the same as defined for

385 α_{pss} . The subscript in PSI_z indicates that this is the value obtained under trait Z contrary

386 to PSI which is obtained directly under the choice trait.

387 Then, α_E permits to interchange between the scale of phenotypes and choice so, $ZE_0 =$

$$388 \quad \alpha_E E_0.$$

389 Alternatively, we can also express E_0 (see Appendix E for details) as

$$390 \quad E_0 = D'_{KL}(w, q' || p')$$

which is a Kullback–Leibler-like divergence with weights $w_{ij} = (PSS_{ij} - 1) / PTI_{ij}$ in the observations q' . Note that contrary to the standard K-L divergence, E_0 can be negative depending on the weights.

The total information is separated into the sexual selection (J_{PSS}) and isolation (J_{PSI}) components plus the mixed term E_0 . Note that E_0 appears only when both sexual selection and sexual isolation effects occur so that the above given covariance is not null.

If $E_0 = 0$ this means that J_{PSS} and/or J_{PSI} capture the complete information from mate choice. When E_0 is positive it indicates that the information gathered from J_{PSS} and J_{PSI} separately is not the total information from mating choice. On the other side, when E_0 is negative there is some inverse relationship between sexual selection and sexual isolation information.

In the scale of phenotypes the partition still holds provided that the same phenotypic classification is applied when computing the different measures

$$JZ_{PTI} = JZ_{PSS} + JZ_{PSI} + ZE_0 \quad (6)$$

where ZE_0 is the value of E_0 in the given phenotypic scale.

For any given logarithmic base, the amount of the total information, JZ_{PTI} , depends on the magnitude of the differences among the mutual mating propensity values linked to the population phenotypes under study. The higher the differences encountered the higher the value of JZ_{PTI} . Without loss of generality, from herein we consider the natural logarithm because this facilitates testing against the null hypothesis of no information by means of the chi-square distribution.

We have given formulae for the change in the phenotypic scale for every term in (6) except JZ_{PSI} . In this case, we have to predict the change in the scale by computing the remaining factors and solving for JZ_{PSI} (Appendix C).

If, as expected, the observations used to compute the information statistics come from the same sample, the sum in (6) is exact so it recovers the whole information gathered from mate choice. On the contrary, if the computations has been performed using different samples, it could be a remaining part of mate choice information that is non-explained by the above statistics but that can be recovered by the error term

$$E_{PTI} = JZ_{PTI} - (JZ_{PSS} + JZ_{PSI} + ZE_0) \quad (7)$$

that reflects how much information may be lost due to differences in the measurement of the involved phenotypes when computing the different information components from separate samples.

4. Real Data Application

The theoretical framework I have presented so far has been defined in a general way, for any number of male and female phenotypic classes, and for any kind of mutual mating propensities. The application of the J statistics to a data sample of dimorphic traits (two classes), is immediate. For clarity, I will use the same example that appears in the pairwise statistics (PTI , PSI and PSS) original article (Rolán-Alvarez and Caballero, 2000). The correspondence between the pairwise statistics notation used by (Rolán-Alvarez and Caballero, 2000) and ours is as follows. The two phenotypic types are noted as A and B , the total number of observed matings is t and the number of A type females (A' in Rolán-Alvarez and Caballero, 2000) becomes, under our notation, $p_{1A}n_1$,

and so B' is $p_{1B}n_1$; the number of A males becomes $p_{2A}n_2$ and B males are $p_{2B}n_2$. The observed absolute number for each pair (i, j) would be $q'_{ij}t$ with $i, j \in \{A, B\}$ (see Table 1). The total number of expected mating pairs from population frequencies is n_1n_2 corresponding to the quantity S in (Rolán-Alvarez and Caballero, 2000).

TABLE 1. The mating model for two phenotypic classes identified as types A and B . The number of observed mating pairs is t .

Females				
	$p_{1A}n_1$	$p_{1B}n_1$		
Males	$p_{2A}n_2$	$q'_{AA}t$	$q'_{BA}t$	$p'_{2A}t$
	$p_{2B}n_2$	$q'_{AB}t$	$q'_{BB}t$	$p'_{2B}t$
	$p'_{1A}t$	$p'_{1B}t$		

p_{1i} : observed relative frequency of type $i \in \{A, B\}$ in population females ; n_1 : number of females in the population; p_{2i} : observed relative frequency of type $i \in \{A, B\}$ in population males ; n_2 : number of males in the population; p'_{1i} : observed relative frequency of type i in mating females; p'_{2i} : observed relative frequency of type i in mating males; q'_{ij} : observed relative frequency of mating pair i, j .

The data correspond to a multiple-choice experiment involving two different lines of *Drosophila melanogaster* so called M -like and Z -like (Hollocher et al., 1997). Rolán-Alvarez & Caballero applied the pairwise statistics to this data and confirmed the previous results from Hollocher *et al* indicating stronger sexual isolation than sexual selection. They also suggested a fitness advantage of females versus males but they were not able of distinguishing between female sexual selection and male preference for M females.

455 To perform the analysis, we expressed the observed data from that experiment in terms
456 of the information model as presented in Table 1. In doing so, and noting that the
457 observed number of mating pairs was $t = 1704$, we obtained the necessary quantities in
458 terms of our model (Table 2).

459 The total mate choice information obtained in JZ_{PTI} is partitioned in 89% of sexual
460 isolation ($JZ_{PSI} / JZ_{PTI} = 0.468 / 0.526 = 0.89$; $I_{PSI} = 0.63$), 6% of sexual selection and
461 5% of mixed effects which explains the 100% of JZ_{PTI} . The information coming from
462 sexual isolation is 14 times that from sexual selection, result that matches pretty well the
463 outcome in (Rolán-Alvarez and Caballero, 2000).

464 The value of JZ_{PTI} multiplied by the number of matings can be approximated by a chi-
465 square variable with 3 degrees of freedom under the expectation of $JZ_{PTI} = 0$, the p -
466 value obtained was below 0.00001 which indicates non-random mating. The test against
467 $JZ_{PSI} = 0$ with 1 degree of freedom, also had a p -value below 0.00001. The test against
468 $JZ_{PSS} = 0$ was also below 0.0001. However, testing separately the female and male
469 sexual selection cases (with one degree of freedom each) produced a p -value below
470 0.0001 for the female case but a p -value of 0.77 for males.

471 Thus, we detected significant sexual isolation and selection effects as previously
472 reported by (Rolán-Alvarez and Caballero, 2000). The sexual selection component is
473 caused by a significant intrasexual effect in females. The mixed term E_0 is positive thus
474 indicating that not all the information is recovered by the PSS and PSI coefficients. This
475 is due to the confounding effect which explains as far as the 5% from the total
476 information.

477

TABLE 2. Analysis using the mate choice information model (Table 1 and equations 7) on *D. melanogaster* mating data from (Hollocher et al., 1997). The number of observed copulating pairs is $t = 1704$.

		Females Z	Females M	
		0.5×1440	0.5×1440	
Males Z	0.5×1440	0.3585×1704	0.145×1704	0.5035×1704
Males M	0.5×1440	0.051×1704	0.4455×1704	0.4965×1704
		0.4095×1704	0.5905×1704	
JZ_{PTI}	0.526			
$JZ_{PSI} (I_{PSI})$	0.468 (0.63)			
$JX_{S1} + JY_{S2}$	$0.033 + 0.00005$			
ZE_0	0.024			
E_{PTI}	0			

Number of females in the population: $n_1 = 1440$; number of males in the population: $n_2 = 1440$; From Table 1, p'_{1i} : observed relative frequency of type i in mating females; p'_{2i} : observed relative frequency of type i in mating males; q'_{ij} : observed relative frequency of copulating pair i, j .

4.1 Exploring models

In the analysis performed above we used the information partition for testing if the observations can be explained by random mating, in a similar way as we do when using the I_{PSI} statistic for testing sexual isolation (Carvajal-Rodriguez and Rolan-Alvarez, 2006; Rolán-Alvarez and Caballero, 2000).

However, the proposed theoretical framework permits going further than just testing random mating. We can rely on the described properties of mutual propensities under

sexual selection and isolation, for defining different effects models. If we can define models from which we can predict the effects, then we can try to fit and infer significant parameters from the available data.

As an example, I have used the software InfoMating (Carvajal-Rodriguez, 2017) to estimate the mutual mating propensity parameters associated to the data in Table 2. The software uses the J information framework to a priori construct (before data) different effects models, and then compare the fitting of random mating, sexual selection and sexual isolation models to the data. There are models having sexual selection only in females, only in males or in both. The models with sexual isolation will have or not sexual selection depending on the frequencies (the conditions on marginal propensities for sexual selection are frequency dependent). The most complex model is also considered. Under this model the mutual mating propensities are estimated by the $PTIs$ that are indeed the maximum likelihood estimates.

I have considered BIC (Schwarz, 1978) and AIC (Akaike, 1973) selection criteria. Both gave similar results. The best fit model was a two parameter model with sexual isolation and female sexual selection effects. The model uses two parameters a and b to define the four mating propensity values as $m_{ZZ} = a$, $m_{ZM} = 1-b$, $m_{MZ} = 1$, $m_{MM} = a+b$.

The obtained estimates under the BIC criterion were $a = 2.47$ and $b = 0.64$ which after normalization, provide the mutual mating propensity estimates as they appear in Table 3.

TABLE 3. Mutual-propensity estimates from multimodel inference.

	Female	Z	M
Male			

Z	1.42	0.58
M	0.21	1.79

513

514 The obtained estimates are almost identical to the corresponding *PTI* values but we have
 515 only needed two parameters instead of three for defining the model. Therefore, the two
 516 parameter model may provide some insight into the biology of the mating relationships.

517 The obtained estimates imply positive assortative mating because the homotype mutual
 518 propensities (main diagonal in Table 3) are higher than the heterotype ones
 519 (antidiagonal, m_{ZM} and m_{MZ}). If we compare the mean homotype versus the mean
 520 heterotype mating propensities, the difference is $a + b - 1$. The value 1 is the value under
 521 random mating so, the increase of homotype mating with respect to random mating is a
 522 $+ b$.

523 Moreover, the chosen model has no male sexual selection effect. This is clear when we
 524 measure the mean effect of changing the male type in the matings. We see that the
 525 effect is 0 i.e. $(m_{MM} - m_{MZ} + m_{ZM} - m_{ZZ}) / 2 = 0$. On the contrary, the mean effect of
 526 changing female Z by M is b .

527 Thus, the deviation from random mating in the data from Table 2 is composed of a
 528 sexual isolation effect captured by the parameter a plus an effect b of sexual selection
 529 focused on the M females which may imply that this females are more receptive to
 530 mating in general.

531

532 **5. Female preference and male display models**

The example we have considered involves the same trait in female and male. However, there are several situations where the female preference is for a male display trait (Pomiankowski and Iwasa, 1998). In this case, the female trait is the exerted preference and the male trait is the target phenotype. In the preference-display context, the traits involved are different between sexes so that the crosses cannot be classified in homotypic versus heterotypic, which prevents the calculation of I_{PSI} and other similar indices that are only applicable to mating models in which the female and male phenotype is the same (similarity/dissimilarity models).

The mutual mating propensity framework can easily capture the preference-display scenario to express the components of mate choice in terms of information.

In Table 4 we appreciate three examples of such preference-display models. There are two types of females which have preference for males displaying phenotypic values A , B or C . The frequencies for the different phenotypes are equal. The mating propensities have been defined with only one parameter and three possible values, namely a , $a/2$ or virtually 0 (ϵ). In the first column the female preference generates a situation of complete isolation; in the second column the resultant effect of the female preference is of full intrasexual selection in males and the third column corresponds to a mixed scenario where both sexual selection and isolation occur with a mixed effect of -24% than indicates an strong overlap between both effects.

TABLE 4. Mating propensity models of female preference for male display traits. Two types of females ‘0’ or ‘1’ might have different preferences for males presenting distinct values for some secondary trait ($a = 1$, $a/2$ or ϵ). Females are the

556 **choosy sex so that the generic model implies only the female acceptance (or**
557 **preference) function f_1 .**

$m_{\text{female} \cdot \text{male}}$	Isolation	Sexual selection	Mixed	Generic
m_{0A}	a	a	a	$f_1(0,A)$
m_{0B}	$a/2$	$a/2$	ϵ	$f_1(0,B)$
m_{0C}	ϵ	ϵ	ϵ	$f_1(0,C)$
m_{1A}	ϵ	a	ϵ	$f_1(1,A)$
m_{1B}	$a/2$	$a/2$	ϵ	$f_1(1,B)$
m_{1C}	a	ϵ	a	$f_1(1,C)$
JZ_{PTI}	13	13	26	
$JZ_{\text{PSI}} / JZ_{\text{PTI}}$	100%	0%	75%	
$(JX_{S1} + JY_{S2}) / JZ_{\text{PTI}}$	0%	100% (0% + 100%)	49% (0% + 100%)	
ZE_0 / JZ_{PTI}	0%	0%	-24%	

558 ϵ : represents an infinitesimal value for which the quantity $\epsilon \log(\epsilon)$ will be zero.

559

560 However, upon inspecting the propensities in the Table 4, the effects (isolation,
561 selection, and mixed) of the preference-display scenarios are not so intuitive, which
562 stresses the usefulness of the information partition. For example, the column “Mixed”
563 can be represented in a two-way table (Table 5).

564

565 **Table 5. Mating preferences involved in the mixed model from Table 4.**

Males	A	B	C
-------	---	---	---

Females			
0	a	ε	ε
1	ε	ε	a

566

567 The pattern in Table 5 is a clear case of isolation that splits females 0 and males A from
568 females 1 and males C. Recall that a mixed model implies isolation+ sexual selection.

569 The model is mixed because there is strong sexual selection against B males that
570 virtually do not mate. At this stage, we do not care if this is because A and C have more
571 vigour than B in the searching for mates, or because females in general do not like B
572 males. The result is male sexual selection (against B males), so the model is mixed
573 because the preferences in the model produce both sexual isolation and selection.

574 We can perform a similar exercise with the other models in Table 4 and see, for
575 example, that the isolation model provokes isolation because females 0 prefer A while
576 skip C, and vice versa, females 1 skip A and prefer C (both female phenotypes have the
577 same preference for B).

578 Under uniform frequencies in both sexes, this isolation model does not generate sexual
579 selection. The marginal propensity of females 0 and 1 is the same, $m'_{f0} = m'_{f1} = (a + a/2$
580 $+ \varepsilon) / 3$ so there is no sexual selection in females. In males, the marginal propensity of A
581 and C is $(a + \varepsilon)/2$ while is $a/2$ for B males; they are equal except for the addition of the
582 factor $\varepsilon/2$ which is virtually 0 and therefore there is no detectable effect of sexual
583 selection.

584

585 *Discussion*

The mate choice model defined in (1) is valid for phenotypes and genotypes, and it only requires the abstract representation of any kind of relative mutual mating propensity.

The model in (1) is similar to the model for the mating pattern predicted from encounter-mating (EM) scenarios when the availability of individuals is not affected by the matings that have already occurred (equation 19 in Gimelfarb, 1988). The latter happens in polygamous species, or even with monogamous, when only a small fraction of individuals of both sexes successfully mate (i.e. the process of the encounter corresponds to sampling with replacement).

On the contrary, when the species are monogamous and the population size is small, the mating pattern will depend on the kind of pair formation process (Gimelfarb, 1988). In the latter case, the information framework should still be valid but the equations must be updated after each mating round. Therefore, the pair formation process without replacement, would introduce some noise in the obtained mating patterns. The application of the proposed methodology in such situations is left for future work.

At the same time, (1) is analogous to the Wright's selection equation for the change in gene frequencies so, from the viewpoint of that analogy, the relative propensity would play the role of fitness referred to each mating couple. By defining the relationship between observed and expected mating frequencies as a function of relative mating propensity, the choice is expressed as a potentiality which is also a key characteristic of fitness (Wagner, 2010).

As with the fitness concept, the mate propensity faces two main aspects, namely the measurement of differences between couples, and the intrinsic causes that provokes the propensity values. By expressing the equation of change in terms of the choice information and its components, this work focused in the first aspect.

I have connected the cause of mating choice, which is modeled by the abstract concept of mutual mating propensity, with the different possible outcomes. Notably, the connection between mate choice and its consequences appears in terms of information. The general equation (J_{PTI}) represents the information gained by mate choice with respect to random mating. This general information is the sum of the information due to sexual isolation and sexual selection, plus a mixed effect term that can be computed separately from the others. The mixed term measures the adjustment of the partition components with respect to the total mate choice information. In addition, the information from sexual selection is the sum of the male and female intrasexual selection information.

Although the model has been constructed assuming discrete phenotypes, it is possible to estimate the Kullback-Leibler divergence for the continuous case (Pérez-Cruz, 2008) in order to apply a similar mate choice information partition for quantitative traits.

The information framework also provides a baseline for defining adequate null hypotheses for the distinct aspects of the mate choice problem. In fact, the information terms are mean log-likelihood ratios, so we can apply them for contrasting the different null hypothesis about random mating, sexual selection, and isolation.

Therefore, the statistical test defined as nJ (total number of matings n , multiplied by the Jeffreys' divergence) is similar to a G -test. In fact, if we note G for the G -test with nq expected counts, and G' for the G -test with nq' expected counts, then $nJ = (G + G') / 2$.

Indeed, it has been shown that the G -test G , is highly correlated (0.99) with the Jeffreys' statistic (Evren and Tuna, 2012).

We can perform the test against random mating by considering a chi-square distribution with $KK'-1$ degrees of freedom (Evren and Tuna, 2012; Sokal and Rohlf, 1981), where

634 $K \times K'$ is the number of different mating categories. The intrasexual selection
635 components correspond to $K-1$ and $K'-1$ degrees of freedom for K female and K' male
636 traits respectively. In addition, the sexual isolation component corresponds to $(K-1)(K'-$
637 $1)$ degrees of freedom.

638 Of course, we may also use randomization tests if we prefer to rely on the empirical
639 distribution approach.

640 Therefore, if we want to contrast mate choice for a given trait Z , we test deviations from
641 zero information in JZ_{PTI} and its components. However, if we want to contrast mate
642 choice in general, we must test deviations from zero information in J_{PTI} which should be
643 the same that testing a flat preference function across all trait values (Edward, 2015).

644 In addition to contrasting the null hypothesis of random mating, we may take advantage
645 of the informational partition of mate choice effects to develop different kind of general
646 models defined by their effects. This is possible because the developed relationships
647 expose and clarify useful general properties, such as the requirement of non-
648 multiplicative mutual propensity functions for obtaining sexual isolation effects and the
649 connection of the marginal propensities at each sex with sexual selection.

650 As an example of the possible insight that can be gained relying in the informational
651 framework, I reanalyzed the well-known example of *D. melanogaster* mating data from
652 (Hollocher et al., 1997) and besides confirming previous results on the components of
653 sexual isolation and selection effects, I have been able to fit a simple two-parameter
654 model that explains the data by means of a component of sexual isolation plus a sexual
655 selection component favoring the mating of the M-type females.

656 In addition to the similarity models in which the same phenotype is involved in both
657 sexes, the preference-display models are also easily interpreted in terms of information

658 and we have been able of inspecting models of full isolation, full intrasexual selection,
659 and mixed effect models.

660 We have also seen an example with multiplicative mutual propensity by means of a
661 simple preference-display model based on the bird sage grouse (*Centrocercus*
662 *urophasianus*) in which the traits ‘acoustic broadcast’ and ‘display rate’, act
663 multiplicatively over the preference.

664 To conclude, it is worth mentioning that the concept of mate choice is important in the
665 evolutionary theory and other disciplines. It has been approached from a diversity of
666 fields and inference methodologies, which has provoked that the terminology has not
667 always been very precise. This may have contributed to some confusion in terms of
668 causes and effects jointly with plenty discussion (Ah-King and Gowaty, 2016; Edward,
669 2015; Janicke et al., 2016; Roughgarden et al., 2015).

670 Here, I have shown that the mean change in the mating phenotypes can be expressed as
671 the information gained due to mate choice. Overall, the obtained results lead to the
672 suggestion that the information interpretation of mate choice is an interesting avenue
673 that may help to improve the study of the causes as well as the effects of this important
674 evolutionary phenomenon.

Acknowledgement

I would like to thank Carlos Canchaya and two anonymous reviewers for their valuable comments on the manuscript. This work was supported by Xunta de Galicia (Grupo de Referencia Competitiva, ED431C2016-037), Ministerio de Economía y Competitividad (BFU2013-44635-P and CGL2016-75482-P) and by Fondos FEDER (“Unha maneira de facer Europa”).

675

676 **Appendix A. Relative propensity and phenotypes**

677 Let T the trait that is the target of the choice while Z is any other trait that can be more
678 or less related to the choice.

679 We call J_{PTI} to the change in the number of matings when these matings were classified
680 by T (equation 2 in the main text). On the other hand, the change in the number of
681 matings when they were classified by Z is JZ_{PTI} .

682 If we cannot measure directly the values of m (trait T) then we simply compute J based
683 on trait Z and therefore we are really computing

$$684 \quad JZ_{PTI} = \alpha_z J_{PTI}$$

685 The scaling factor α_z is

$$686 \quad \alpha_z = (\beta(m_z, l_z) \times D_z) / (\beta(m, l) \times D) \text{ if } J_{PTI} > 0; \text{ or } 1 \text{ otherwise.}$$

687 The values $D_m = V(m) / M$ and $D_z = V(m_z) / M_z$ are the indexes of dispersion over T and Z
688 respectively; the subscript z indicates that the matings were classified by phenotype Z
689 instead of by T (see Appendix C for more details in the scaling formulae).

690 The distinction between JZ_{PTI} and J_{PTI} matters because when the information produced
691 by mate choice is computed as JZ_{PTI} , a value of zero could mean that $i) \text{cov}(m_z, l_z) = 0$
692 so $\alpha_z = 0$ i.e. the trait Z do not covariate with the differential propensities (the mating is
693 random with regard to Z) or, alternatively $ii) J_{PTI} = 0$ meaning that there is no
694 differential mating propensity at all, i.e. the mating is random independently of the trait
695 we focused on.

Let's see an example of the first situation i.e. there is mate choice but the trait Z is not involved in the mate choice process. Thus, assume that some unknown trait X that is involved in an assortative mating process exists. There are two phenotypic classes '1' and '2' so that $m_{11} = 2, m_{12} = 1, m_{21} = 1, m_{22} = 2$; the phenotype frequencies are uniform in males and females, $p = 0.5$, and mean propensity $M = 1.5$. This results in $J_{PTI} = J_{PSI} = 0.1155$.

However, when counting the matings, we evaluated a phenotype Z with classes A/B that are independent of the mating choice process. If the trait responsible of the mate choice is uniformly distributed over the phenotypes A/B (i.e. half of A individuals have trait value '1' and the other half have value '2' and the same is true for B individuals) then the expected preference for the phenotype pairs are

$$m_{AA} = m_{AB} = m_{BA} = m_{BB} = (m_{11} + m_{12} + m_{21} + m_{22}) / 4 = 1.5 = M = M_z.$$

Thus it is clear that the normalized preferences are 1 and $JZ_{PTI} = 0$.

Consider now a different case in which $JZ_{PTI} \neq 0$, this means that the non-random mating is correlated to some extent with the trait Z . For example, consider the same mate choice scenario as above with $J_{PTI} = J_{PSI} = 0.1155$ but now the phenotype under study is partially linked to the mate choice so $m_{AA} = 1.7, m_{AB} = 1.2, m_{BA} = 1.2, m_{BB} = 1.7$.

Recall that the frequencies are uniform. If we compute directly the information index over the phenotypes Z , we get $JZ_{PTI} = 0.03$. The mean propensity for these phenotypes is $M_z = 1.45$ and $cov(m_z, l_z) = 0.0435$. However, $M = 1.5$ and $cov(m, l) = 0.1733$ for the real mate choice trait (T). The scaling is $\alpha_z = [cov(m_z, l_z) / M_z] / [cov(m, l) / M] = 0.2597$ so $J_{PTI} = JZ_{PTI} / \alpha_z = 0.03 / 0.2597 = 0.1155$, as expected.

If we have an estimate or a computable proxy for the propensity function m , as for example, a measure of distance between female and male traits $|D|$, or a model with Gaussian functions (Carvajal-Rodriguez and Rolán-Alvarez, 2014), then JZ_{PTI} and J_{PTI} can be estimated separately. We obtain J_{PTI} by means of $J(q', q)$ using the estimated mating propensities to classify the frequencies, and we still can use the phenotypes Z to compute JZ_{PTI} . The relationship between both measures may give an idea about the linkage between the phenotypes Z and the mate choice.

Suppose that the estimate of J_{PTI} is different from zero while $JZ_{PTI} = 0$, then mate choice do exist but it is not linked with the phenotype Z . An interested researcher could compare different traits looking for the ones having the best scaling for the information J_{PTI} , i.e. the one that is more involved in the mate choice. It seems that if we are able of having good proxies for mating propensity, this could pave the way for testing the impact of different traits on mate choice.

Additionally, we still can compute directly $\Delta Z = Z' - Z$, i.e. the difference in phenotype frequencies between observed and expected by random mating. Therefore, we have two values, $\Delta_m Z$ and ΔZ , for the change in Z , the discrepancy between them gives an estimate of the change in Z caused by other factors than mating propensity (e.g. predators) so $e_z = \Delta Z - \Delta_m Z$.

Thus the total change in mean Z is

$$\Delta Z = \alpha_z J_{PTI} + e_z$$

Appendix B

739 The relationship between sexual selection measured within sex and the pair sexual
740 selection measured by PSS is

$$741 \quad J_{PSS} = J_{S1} + J_{S2}$$

742 To see this, recall that J_{PSS} is the sum of products $\Delta(p_{1i}p_{2j}) \times \log((p'_{1i}p'_{2j})/(p_{1i}p_{2j}))$.

743 Then note that

$$744 \quad \log((p'_{1i}p'_{2j})/(p_{1i}p_{2j})) = \log(p'_{1i}/p_{1i}) + \log(p'_{2j}/p_{2j})$$

745 and that

$$746 \quad \sum_{ij} \Delta(p_{1i}p_{2j}) = \sum_{ij} (p'_{1i}p'_{2j} - p_{1i}p_{2j}) = \sum_i p'_{1i} \sum_j p'_{2j} - \sum_i p_{1i} \sum_j p_{2j} = 0$$

747 because each summation is 1. Then, after some algebraic rearrangement we obtain

$$748 \quad J_{PSS} = \sum_{ij} \Delta(p_{1i}p_{2j}) \log((p'_{1i}p'_{2j})/(p_{1i}p_{2j})) = \sum_i p'_{1i} \log(p'_{1i}/p_{1i}) - \sum_i p_{1i} \log(p'_{1i}/p_{1i}) +$$

$$749 \quad \sum_j p'_{2j} \log(p'_{2j}/p_{2j}) - \sum_j p_{2j} \log(p'_{2j}/p_{2j}) = J_{S1} + J_{S2}.$$

Appendix C. Scaling factors

750 We can compute the scaling factors that translate the information between different
751 phenotypic scales.

752 We have used the notation J_{PTI} , J_{PSS} , J_{PSI} and E_0 for indicating the information when
753 measured from phenotypes that are the choice targets i.e. the phenotypes that the mates
754 care about in choosing each other. On the other hand, we note JZ_{PTI} , JZ_{PSS} , JZ_{PSI} and
755 ZE_0 when the phenotypes may or may not be related with the choice.

756 Therefore, it is interesting to shown how the information changes between one measures
757 or others. So that

$$JZ_{PTI} = \alpha_z J_{PTI}; JZ_{PSS} = \alpha_{pss} J_{PSS}; JX_{S1} = \alpha_x J_{S1}; JY_{S2} = \alpha_y J_{S2}; ZE_0 = \alpha_E E_0 \text{ and } JZ_{PSI} = \alpha_z J_{PTI} - (\alpha_{pss} J_{PSS} + \alpha_E E_0)$$

The scalings are as follows.

$$\alpha_z = [\text{cov}(m_z, l_z) / M_z] / [\text{cov}(m, l) / M] \text{ with } l_z = \log(m_z) \text{ and } l = \log(m) \text{ if } J_{PTI} > 0$$

or $\alpha_z = 1$ otherwise ($J_{PTI} = 0$).

However, we can also express $\text{cov}(m_z, l_z) / M_z$ as $\beta(m_z, l_z) \times D_z$ which is the regression of the propensity under the trait Z over its logarithm multiplied by the index of dispersion. Then if $D_m = V(m) / M$ and $D_z = V(m_z) / M_z$ are the indexes of dispersion over the choice and Z traits respectively, we obtain

$$(\text{cov}(m_z, l_z) / M_z) / (\text{cov}(m, l) / M) = (\beta(m_z, l_z) \times D_z) / (\beta(m, l) \times D)$$

Then if $J_{PTI} > 0$ define

$$\alpha_z = (\beta(m_z, l_z) \times D_z) / (\beta(m, l) \times D)$$

where $l_z = \log(m_z)$ and $l = \log(m)$

or $\alpha_z = 1$ if $J_{PTI} = 0$

Similarly,

$$\alpha_x = [\text{cov}(m'_{zf}, l_{zf}) / M_z] / [\text{cov}(m'_f, l_f) / M] = (\beta(m'_{zf}, l_{zf}) \times D_{zf}) / (\beta(m'_f, l) \times D_f)$$

or 1 if $J_{PS1} = 0$

where $l_{zf} = \log(m'_{zf})$, $D_{zf} = V(m'_{zf}) / M_z$, $l_f = \log(m'_f)$, $D_f = V(m'_f) / M$

$$\alpha_y = [\text{cov}(m'_{zm}, l_{zm}) / M_z] / [\text{cov}(m'_m, l_m) / M] = (\beta(m'_{zm}, l_{zm}) \times D_{zm}) / (\beta(m'_m, l) \times D_m) \text{ or } 1 \text{ if }$$

$J_{PS2} = 0$

778 where $l_{zm} = \log(m'_{zm})$, $D_{zm} = V(m'_{zm})/M_z$.

779 $\alpha_{pss} = [\text{cov}(m'_{zf} m'_{zm}, l'_z) / M_z^2] / [\text{cov}(m'_f m'_m, l') / M^2] = (\beta(m'_z, l'_z) \times D'_z) / (\beta(m', l') \times D')$

780 or 1 if $J_{PSS} = 0$

781 where $m'_z = m'_{zf} \times m'_{zm}$, $m' = m'_f m'_m$, $D'_z = V(m'_z) / M_z^2$, $l'_z = \log(m'_z)$, $l' = \log(m')$, $D' =$

782 $V(m') / M^2$.

783 $\alpha_E = [\text{cov}(m'_{zf} m'_{zm}, \log(PSI_z)) / M_z^2] / [\text{cov}(m'_f m'_m, \log(PSI)) / M^2] = (\beta(m'_z, l_{zpsi}) \times D'_z) /$

784 $(\beta(m', l_{psi}) \times D')$ or 1 if $E_0 = 0$

785 where $l_{zpsi} = \log(PSI_z)$, $l_{psi} = \log(PSI)$.

786 Finally

787 if $J_{PSI} > 0$

788 $\alpha_\delta = \alpha_z$ when $J_{PSS} = 0$ or in general

$$\alpha_\delta = \frac{\alpha_T J Z_{PSI}}{\alpha_{pss} \alpha_E J Z_{PTI} - \alpha_z (\alpha_E J Z_{PSS} + \alpha_{pss} Z E_0)}$$

789 with $\alpha_T = \alpha_z \times \alpha_{pss} \times \alpha_E$; so that $J Z_{PSI} = \alpha_\delta J_{PSI}$.

Appendix D. Multiplicative preference example

790 Consider two male traits, B/b for acoustic broadcast and D/d for display rate, where in

791 both cases the upper case refers to the higher value of the trait; consider also one female

792 trait X with a single phenotypic class. The females are the choosers and so we may

793 define multiplicative preferences as appear in Table S1.

794

Table S1. Mating preferences for acoustic broadcast and display rate.

	Males	Bd	BD	bD	bd
Females					
X	α	$\alpha\beta$	β	1	

The model assumes a single female phenotypic class and so the frequency is 1. Let the frequencies of the different male classes be p_{Bd} , p_{BD} , p_{bD} and p_{bd} . Then the mean preference is $M = \alpha p_{Bd} + \alpha\beta p_{BD} + \beta p_{bD} + p_{bd}$. There is only one female marginal preference, that in fact coincides with the mean preference, $m'_{fX} = M$. There are four male marginal preferences (m'_{mBd} , m'_{mBD} , m'_{mbD} , m'_{mbd}) that match the corresponding male propensity with the single female phenotype i.e. $m'_{mBd} = \alpha$, $m'_{mBD} = \alpha\beta$, $m'_{mbD} = \beta$ and $m'_{mbd} = 1$.

In (5) we have given an aprioristic expression (δ) of the pair sexual isolation (PSI) coefficients in terms of the preferences so that $\delta_{ij} = (m_{ij} / M) / (m'_{fi} m'_{mj} / M^2)$. Therefore, $\delta_{Bd} = (m_{Bd} / M) / (m'_{fX} m'_{mBd} / M^2) = (\alpha / M) / (M \alpha / M^2) = 1$ and similarly for the other phenotypes so finally $\delta_{Bd} = \delta_{BD} = \delta_{bD} = \delta_{bd} = 1$, which implies no sexual isolation.

On the other hand, the model predicts male sexual selection whenever α and/or $\beta \neq 1$.

Appendix E

Proposition 1

$$\sum_{ij} \Delta(p'_{1i} p'_{2j}) \log(PSS_{ij}) = 0$$

then

$$E_0 = \sum_{ij} (\Delta(p_{1i}p_{2j})\log(PSI_{ij}) + \Delta(p'_{1i}p'_{2j})\log(PSS_{ij})) = \sum_{ij} \Delta(p_{1i}p_{2j})\log(PSI_{ij}).$$

First, recall that

$$\sum_{ij} \Delta(p'_{1i}p'_{2j})\log(PSS_{ij}) = \sum_{ij} (q'_{ij} - p'_{1i}p'_{2j})\log(PSS_{ij})$$

and also that by definition of PSS

$$\log(PSS_{ij}) = \log((p'_{1i}p'_{2j}) / (p_{1i}p_{2j}))$$

that can be expressed as

$$\log(PSS_{ij}) = \log(p'_{1i} / p_{1i}) + \log(p'_{2j} / p_{2j})$$

then by simple substitution and rearranging the terms

$$\sum_{ij} \Delta(p'_{1i}p'_{2j})\log(PSS_{ij}) =$$

$$\sum_{ij} (q'_{ij} - p'_{1i}p'_{2j})\log(PSS_{ij}) = \sum_{ij} q'_{ij}\log(p'_{1i}/p_{1i}) + \sum_{ij} q'_{ij}\log(p'_{2j} / p_{2j}) -$$

$$\sum_{ij} (p'_{1i}p'_{2j})\log(p'_{1i} / p_{1i}) - \sum_{ij} (p'_{1i}p'_{2j})\log(p'_{2j} / p_{2j})$$

Now recall that the i subscript refers to females and subscript j to males, then the double

summatory is the sum through females and males, thus by reminding that $\sum_i p'_{2j} = \sum_i p'_{1i}$

$= 1$ we note that

$$\sum_{ij} (p'_{1i}p'_{2j})\log(p'_{1i}/p_{1i}) = \sum_i (p'_{1i})\log(p'_{1i}/p_{1i}) \sum_j p'_{2j} = \sum_i p'_{1i}\log(p'_{1i}/p_{1i})$$

and similarly

$$\sum_{ij} (p'_{1i}p'_{2j})\log(p'_{2j}/p_{2j}) = \sum_j p'_{2j}\log(p'_{2j}/p_{2j})$$

so we have

$$\sum_{ij} \Delta(p'_{1i}p'_{2j})\log(PSS_{ij}) =$$

$$\sum_{ij} q'_{ij}\log(p'_{1i}/p_{1i}) + \sum_{ij} q'_{ij}\log(p'_{2j} / p_{2j}) - \sum_i p'_{1i}\log(p'_{1i}/p_{1i}) - \sum_j p'_{2j}\log(p'_{2j} / p_{2j})$$

831 Now note that

$$832 \quad \sum_{ij} q'_{ij} \log(p'_{1i}/p_{1i}) = \sum_i \log(p'_{1i}/p_{1i}) \sum_j q'_{ij}$$

833 and that for each female i the sum through males of the observed mating frequencies

834 involving female i is, by definition, p'_{1i} i.e. $\sum_j q'_{ij} = p'_{1i}$ and similarly for each male j we

835 have $\sum_i q'_{ij} = p'_{2j}$. Then

$$836 \quad \sum_{ij} q'_{ij} \log(p'_{1i}/p_{1i}) = \sum_i \log(p'_{1i}/p_{1i}) p'_{1i}$$

837 and

$$838 \quad \sum_{ij} q'_{ij} \log(p'_{2j}/p_{2j}) = \sum_j \log(p'_{2j}/p_{2j}) p'_{2j}$$

839 therefore

$$840 \quad \sum_{ij} \Delta(p'_{1i} p'_{2j}) \log(PSS_{ij}) =$$

$$841 \quad \sum_i \log(p'_{1i}/p_{1i}) p'_{1i} - \sum_i p'_{1i} \log(p'_{1i}/p_{1i}) + \sum_j \log(p'_{2j}/p_{2j}) p'_{2j} - \sum_j p'_{2j} \log(p'_{2j}/p_{2j}) = 0$$

842 and so the proposition is true

$$843 \quad E_0 = \sum_{ij} (\Delta(p_{1i} p_{2j}) \log(PSI_{ij}))$$

844

845 *Proposition 2*

$$846 \quad E_0 = D'_{KL}(w, q' || p')$$

847 where

$$848 \quad p' = p'_1 p'_2$$

$$849 \quad w_{ij} = (PSS_{ij} - 1) / PTI_{ij}$$

$$D'_{KL}(w, q' || p') = \sum_{ij} w_{ij} q'_{ij} \log(q'_{ij} / (p'_{1i} p'_{2j}))$$

From the model (1) and the partitions (4) and (5) in the main text we know that

$$(p_{1i} p_{2j}) = (p'_{1i} p'_{2j}) / PSS_{ij}$$

$$(p'_{1i} p'_{2j}) = q'_{ij} / PSI_{ij}$$

therefore

$$\Delta(p_{1i} p_{2j}) = (p'_{1i} p'_{2j}) - (p_{1i} p_{2j}) = (q'_{ij} / PSI_{ij}) - (q'_{ij} / PTI_{ij}) = q'_{ij} [(1 / PSI_{ij}) - (1 / PTI_{ij})]$$

and since $PTI_{ij} = PSI_{ij} \times PSS_{ij}$ we obtain

$$\Delta(p_{1i} p_{2j}) = q'_{ij} (PSS_{ij} - 1) / PTI_{ij}$$

and so

$$E_0 = \sum_{ij} (\Delta(p_{1i} p_{2j}) \log(PST_{ij})) = \sum_{ij} w_{ij} q'_{ij} \log(PST_{ij}) = D'_{KL}(w, q' || p')$$

which is Kullback–Leibler-like divergence with weights w_{ij} in the observations q' .

References

- Ah-King, M., Gowaty, P. A., 2016. A conceptual review of mate choice: stochastic demography, within-sex phenotypic plasticity, and individual flexibility. *Ecology and Evolution*. 6, 4607-4642.
- Akaike, H., Information theory and an extension of the maximum likelihood principle. In: B. N. Petrov, F. Csaki, Eds.), Second International Symposium on Information Theory, Budapest: Akademiai Kiado, 1973, pp. 267-281.
- Bergstrom, C. T., Real, L. A., 2000. Toward a theory of mutual mate choice: Lessons from two-sided matching. *Evolutionary Ecology Research*. 493-508.
- Carvajal-Rodriguez, A., InfoMating. 2017, pp. Model selection and multimodel inference from mating frequency data to study sexual selection and sexual isolation effects.

872 Carvajal-Rodriguez, A., Rolan-Alvarez, E., 2006. JMATING: a software for the analysis of sexual
873 selection and sexual isolation effects from mating frequency data. *BMC Evol Biol.* 6, 40.
874 Carvajal-Rodriguez, A., Rolán-Alvarez, E., 2014. A comparative study of Gaussian mating
875 preference functions: a key element of sympatric speciation models. *Biological Journal*
876 *of the Linnean Society.* 113, 642-657.
877 Dall, S. R. X., et al., 2005. Information and its use by animals in evolutionary ecology. *Trends in*
878 *Ecology & Evolution.* 20, 187-193.
879 Darwin, C., 1871. *The descent of man, and selection in relation to sex.* Murray.
880 Edward, D. A., 2015. The description of mate choice. *Behavioral Ecology.* 26, 301-310.
881 Evren, A., Tuna, E., 2012. On some properties of goodness of fit measures based on statistical
882 entropy. *International Journal of Research and Reviews in Applied Sciences.* 13, 192-
883 205.
884 Frank, S. A., 2009. Natural selection maximizes Fisher information. *J Evol Biol.* 22, 231-44.
885 Frank, S. A., 2012a. Natural selection. IV. The Price equation. *J Evol Biol.* 25, 1002-19.
886 Frank, S. A., 2012b. Natural selection. V. How to read the fundamental equations of
887 evolutionary change in terms of information theory. *J Evol Biol.* 25, 2377-96.
888 Frank, S. A., 2013. Natural selection. VI. Partitioning the information in fitness and characters
889 by path analysis. *Journal of Evolutionary Biology.* 26, 457-471.
890 Gavrillets, S., 2004. *Fitness landscapes and the origin of species.* Princeton University Press,
891 Princeton, N.J.
892 Gibson, R. M., 1996. Female choice in sage grouse: the roles of attraction and active
893 comparison. *Behavioral Ecology and Sociobiology.* 39, 55-59.
894 Gimelfarb, A., 1988. Processes of Pair Formation Leading to Assortative Mating in Biological
895 Populations: Encounter-Mating Model. *The American Naturalist.* 131, 865-884.
896 Hollocher, H., et al., 1997. Incipient Speciation by Sexual Isolation in *Drosophila melanogaster*:
897 Variation in Mating Preference and Correlation Between Sexes. *Evolution.* 51, 1175-
898 1181.
899 Hughes, A. L., 2015. Sexual Selection and Mate Choice: Insights from Neutralist Perspectives.
900 *Evolutionary Biology.* 42, 366-378.
901 Janicke, T., et al., 2016. Darwinian sex roles confirmed across the animal kingdom. *Science*
902 *Advances.* 2.
903 Kirkpatrick, M., 1982. Sexual selection and the evolution of female choice. *Evolution.* 1-12.
904 Kuijper, B., et al., 2012. A guide to sexual selection theory. *Annual Review of Ecology,*
905 *Evolution, and Systematics.* 43, 287-311.
906 Lande, R., 1981. Models of speciation by sexual selection on polygenic traits. *Proceedings*
907 *National Academy of Sciences, USA.* 78, 3721-3725.
908 Pérez-Cruz, F., Kullback-Leibler divergence estimation of continuous distributions. 2008 IEEE
909 international symposium on information theory. IEEE, 2008, pp. 1666-1670.
910 Pomiankowski, A., Iwasa, Y., 1998. Runaway ornament diversity caused by Fisherian sexual
911 selection. *Proceedings of the National Academy of Sciences of the United States of*
912 *America.* 95, 5106-5111.
913 Price, G. R., 1972. Extension of covariance selection mathematics. *Annals of human genetics.*
914 35, 485-490.
915 Prum, R. O., 2010. The Lande-Kirkpatrick mechanism is the null model of evolution by
916 intersexual selection: implications for meaning, honesty, and design in intersexual
917 signals. *Evolution.* 64, 3085-3100.
918 Roff, D. A., Fairbairn, D. J., 2014. The evolution of phenotypes and genetic parameters under
919 preferential mating. *Ecology and Evolution.* 4, 2759-2776.
920 Rolán-Alvarez, E., Caballero, A., 2000. Estimating sexual selection and sexual isolation effects
921 from mating frequencies. *Evolution.* 54, 30-6.
922 Rosenthal, G. G., 2017. *Mate choice: the evolution of sexual decision making from microbes to*
923 *humans.* Princeton University Press.

- 924 Roughgarden, J., et al., 2015. Sexual selection studies: a NESCent catalyst meeting. PeerJ
- 925 PrePrints 3:e680v3.
- 926 Schwarz, G., 1978. Estimating the dimension of a model. The Annals of Statistics. 6, 461-464.
- 927 Servedio, M. R., et al., 2011. Magic traits in speciation: 'magic' but not rare? Trends in Ecology
- 928 & Evolution. 26, 389-397.
- 929 Sokal, R. R., Rohlf, F. J., 1981. Biometry. W. H. Freeman and Co., New York.
- 930 Wagner, G. P., 2010. The measurement theory of fitness. Evolution. 64, 1358-1376.