bioRxiv preprint doi: https://doi.org/10.1101/095901; this version posted August 3, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

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.

bioRxiv preprint doi: https://doi.org/10.1101/095901; this version posted August 3, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

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.

bioRxiv preprint doi: https://doi.org/10.1101/095901; this version posted August 3, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

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

bioRxiv preprint doi: https://doi.org/10.1101/095901; this version posted August 3, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

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.

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,

bioRxiv preprint doi: https://doi.org/10.1101/095901; this version posted August 3, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

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 target trait is "magic" sensu Gavrilets (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 Jeffrey 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

bioRxiv preprint doi: https://doi.org/10.1101/095901; this version posted August 3, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

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

bioRxiv preprint doi: https://doi.org/10.1101/095901; this version posted August 3, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

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 … X_k, the frequency of the phenotype X_i in the female population is $p_{1i} = n_{1Xi} / n_1$, 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 $p_{2j} = n_{2Yj} / n_2$.

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

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

Let's represent the relative mutual mating propensity between female phenotype X_i and male phenotype Y_j, for successfully mating in the environment e as

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

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

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

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

bioRxiv preprint doi: https://doi.org/10.1101/095901; this version posted August 3, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

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

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

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 $\Sigma\Delta q_{ij} = 0$ and that $log(M)$ is constant through the summation, then we can measure the mean population change ΔL in relative propensity as

bioRxiv preprint doi: https://doi.org/10.1101/095901; this version posted August 3, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

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

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

The information obtained from J_{PTI} has been computed using the different propensities as classes for classifying the couples i.e. we considered $log(m)$ as the phenotype Z. When the classes are based upon other phenotypes rather than propensities, we are conveying a specific meaning for the change in frequencies, say, the change in mating frequencies due to differential mutual propensities is observed in terms of change in shell color mating frequencies. Therefore, the phenotype can be viewed as other scale on which we can measure this information (Frank, 2013). Of course, different kinds of

bioRxiv preprint doi: https://doi.org/10.1101/095901; this version posted August 3, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

phenotypes can be more or less involved in mate choice and so, different scales are more or less useful for measuring the mate choice information.

2.2 Relative propensity and phenotypes

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

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

$$\Delta_m Z = \sum_{i,j} \Delta q_{ij} Z_{ij} = \sum_{ij} q_{ij}\left(m_{ij} - M\right)\frac{Z_{ij}}{M} = \frac{cov(Z,m)}{M} \qquad (3)$$

Where *cov* is the population covariance in the sense of Price (1972) as highlighted in Frank (2012a).

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.

bioRxiv preprint doi: https://doi.org/10.1101/095901; this version posted August 3, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

The covariance $cov(m,Z)$ is the regression $\beta(Z,m)$ multiplied by the variance $V(m)$. 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

$$\Delta L = cov(m,l) / M = \beta(l,m)\mathrm{V}(m) / M$$

remembering from (2) that $\Delta L = J_{\mathrm{PTI}}$ then $V(m) / M = J_{\mathrm{PTI}} / \beta(l,m)$ so that

$$\Delta_m Z = \beta(Z,m)\frac{V(m)}{M} = \alpha_z J_{PTI}.$$

The term $\alpha_z = \frac{\beta(Z,m)}{\beta(l,m)}$ changes the scale from mutual mating propensity to phenotypes.

From the point of view of the estimation with real data, if we cannot measure directly the values of m then we simply compute J based on trait Z and therefore we are really computing

$$JZ_{\mathrm{PTI}} = \alpha_z J_{\mathrm{PTI}} \text{ so}$$

$$\Delta_m Z = JZ_{PTI}$$

In this case, note that the *PTI* coefficients are no longer the estimate of the mutual mating propensities because the ratio of frequencies q'_{ij}/q_{ij} does not correspond to m_{ij}/M but is a proxy that would be more or less precise depending on the importance of the measured phenotype over the mating choice. For example, is shell size is driving mate choice, the measure of JZ_{PTI} (Z = shell size) would correspond well with J_{PTI}. However, if other phenotype as shell color has nothing to do with mate choice (and is not correlated with shell size) then the measure of JZ_{PTI} (Z = shell color) would be zero. Further details about the distinction between JZ_{PTI} and J_{PTI} are given in the appendix A.

bioRxiv preprint doi: https://doi.org/10.1101/095901; this version posted August 3, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

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.

bioRxiv preprint doi: https://doi.org/10.1101/095901; this version posted August 3, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

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)

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

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

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

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

$$\Delta_m \bar{X} = \frac{\beta_{xm}}{\beta_{lm}} J_{S1} = JX_{S1}$$

$$\Delta_m \bar{Y} = \frac{\beta_{ym}}{\beta_{lm}} J_{S2} = JY_{S2}$$

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

bioRxiv preprint doi: https://doi.org/10.1101/095901; this version posted August 3, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

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

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

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

In the scale of phenotypes

$$\Delta_{PSS}\bar{Z} = \frac{\beta_{zpss}}{\beta_{lpss}} J_{PSS} = JZ_{PSS}$$

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

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

$$J_{PSS} = J_{S1} + J_{S2}$$

And in the scale of phenotypes

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

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

bioRxiv preprint doi: https://doi.org/10.1101/095901; this version posted August 3, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

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

2.5. Sexual isolation

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

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

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

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

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*.

$$\Delta_{PSI}L = \sum_{i,j} \Delta(p'_{1i}p'_{2j}) log\left(\frac{q'_{ij}}{p'_{1i}p'_{2j}}\right) = J(q', p'_1p'_2) = J_{PSI},$$

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

In the scale of phenotypes

$$\Delta_{PSI}\bar{Z} = \frac{\beta_{z\delta}}{\beta_{l\delta}} J_{PSI} = JZ_{PSI}.$$

bioRxiv preprint doi: https://doi.org/10.1101/095901; this version posted August 3, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

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 propensity of each pair is the product of the marginal types of each partner then δ = 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 (see below).

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}

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

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

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

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

$$log(PTI) = log(PSS) + log(PSI).$$

Therefore

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

bioRxiv preprint doi: https://doi.org/10.1101/095901; this version posted August 3, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

$$J_{PTI} = J_{PSS} + J_{PSI} + E_0$$

where $E_0 = \Sigma_{ij}(\Delta(p_{1i}p_{2j})log(PSI_{ij}) + \Delta(p'_{1i}p'_{2j})log(PSS_{ij}))$. However, note that $\Delta(p'_{1i}p'_{2j})log(PSS_{ij})) = 0$ (see Appendix C) so finally

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

that can also be expressed as

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

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.

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} is not the total information from mating choice. On the other side, when E_0 is negative there is some overlap 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 phenotypic scale.

bioRxiv preprint doi: https://doi.org/10.1101/095901; this version posted August 3, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

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 in the population. 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.

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

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

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-

bioRxiv preprint doi: https://doi.org/10.1101/095901; this version posted August 3, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

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

bioRxiv preprint doi: https://doi.org/10.1101/095901; this version posted August 3, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

were not able of distinguishing between female sexual selection and male preference for *M* females.

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

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

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

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

bioRxiv preprint doi: https://doi.org/10.1101/095901; this version posted August 3, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

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).

bioRxiv preprint doi: https://doi.org/10.1101/095901; this version posted August 3, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

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 *PTI*s 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.

bioRxiv preprint doi: https://doi.org/10.1101/095901; this version posted August 3, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

	Female	**Z**	**M**
Male			
Z		1.42	0.58
M		0.21	1.79

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

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

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

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

bioRxiv preprint doi: https://doi.org/10.1101/095901; this version posted August 3, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

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 were 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

bioRxiv preprint doi: https://doi.org/10.1101/095901; this version posted August 3, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

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

$m_{female,male}$	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_{PSI}/ JZ_{PTI}	100%	0%	75%	
$(JX_{S1}+JY_{S2})$ / JZ_{PTI}	0%	100% (0% + 100%)	49% (0% + 100%)	
ZE_0 / JZ_{PTI}	0%	0%	-24%	

ε: represents an infinitesimal value for which the quantity εlog(ε) will be zero.

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. 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,

bioRxiv preprint doi: https://doi.org/10.1101/095901; this version posted August 3, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

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 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-Leiblerg 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.

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

bioRxiv preprint doi: https://doi.org/10.1101/095901; this version posted August 3, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

$K \times K'$ is the number of different mating categories. The intrasexual selection components correspond to K-1 and K' -1 degrees of freedom for K female and K' male traits respectively. In addition, the sexual isolation component corresponds to (K-1)(K' -1) degrees of freedom. Of course, we may also use randomization tests if we prefer to rely on the empirical distribution approach.

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

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

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

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

bioRxiv preprint doi: https://doi.org/10.1101/095901; this version posted August 3, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

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

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

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

Acknowledgement

I would like to thank Carlos Canchaya 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").

Appendix A. Relative propensity and phenotypes

We can express the change in the mating phenotypes directly as a function of the mutual mating propensity *m* (or its logarithm, equation 2 in the main text) or as a function of the mating phenotypes for some trait (equation 3).

bioRxiv preprint doi: https://doi.org/10.1101/095901; this version posted August 3, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

If we cannot measure directly the values of *m* then we simply compute *J* based on trait *Z* and therefore we are really computing

$$JZ_{PTI} = \alpha_z J_{PTI} \text{ so}$$

$$\Delta_m Z = JZ_{PTI}$$

The distinction between JZ_{PTI} and J_{PTI} matters because when the information produced by mate choice is computed as JZ_{PTI}, a value of zero could means that *i*) $\alpha_z = 0$ i.e. the trait *Z* do not covariate with the differential propensities so that the mating is random with regard to *Z* or, alternatively *ii*) $J_{PTI} = 0$ meaning that there is no differential mating propensity at all, i.e. the mating is random independently of the trait we focused on. In any other case, when $JZ_{PTI} \neq 0$ it means that the non-random mating is correlated with the trait *Z* i.e. $cov(m,Z) \neq 0$.

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 the term α_z can be separated from J_{PTI}. In this case we could 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* and our proxy for *m*, to compute α_z as the ratio of the corresponding regressions. In this way, we have kept apart the information *J* from the scaling α. The product of information and scale gives $\Delta_m Z$.

Suppose that the estimate of J_{PTI} is different from zero while the scaling is $\alpha_z = 0$, 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

bioRxiv preprint doi: https://doi.org/10.1101/095901; this version posted August 3, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

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

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

$$J_{PSS} = J_{S1} + J_{S2}$$

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

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

and that

$$\Sigma_{ij}\Delta(p_{1i}p_{2j}) = \Sigma_{ij}(p'_{1i}p'_{2j} - p_{1i}p_{2j}) = \Sigma_i p'_{1i}\, \Sigma_j p'_{2j} - \Sigma_i p_{1i}\, \Sigma_j p_{2j} = 0$$

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

bioRxiv preprint doi: https://doi.org/10.1101/095901; this version posted August 3, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

$$J_{PSS} = \Sigma_{ij}\Delta(p_{1i}p_{2j})log((p'_{1i}p'_{2j})/(p_{1i}p_{2j})) = \Sigma_i p'_{1i}log(p'_{1i}/p_{1i}) - \Sigma_i p_{1i}log(p'_{1i}/p_{1i}) + \Sigma_j p'_{2j}log(p'_{2j}/p_{2j}) - \Sigma_j p_{2j}log(p'_{2j}/p_{2j}) = J_{S1} + J_{S2}.$$

Appendix C

Proposition 1

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

then

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

First, recall that

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

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

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

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

bioRxiv preprint doi: https://doi.org/10.1101/095901; this version posted August 3, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

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 $\Sigma_j p'_{2j} = \Sigma_i p'_{1i} = 1$ we note that

$$\Sigma_{ij}(p'_{1i}p'_{2j})\, log(p'_{1i}/p_{1i}) = \Sigma_i(p'_{1i})log(p'_{1i}/p_{1i})\, \Sigma_j p'_{2j} = \Sigma_i\, p'_{1i}log(p'_{1i}/p_{1i})$$

and similarly

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

so we have

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

$$\Sigma_{ij}q'_{ij}log(p'_{1i}/p_{1i}\,) + \Sigma_{ij}q'_{ij}log(p'_{2j}\,/\,p_{2j}\,) - \Sigma_i p'_{1i}log(p'_{1i}/p_{1i}) - \Sigma_j p'_{2j}log(\,(p'_{2i}\,/\,p_{2i}\,)$$

Now note that

$$\Sigma_{ij}q'_{ij}log(p'_{1i}/p_{1i}\,) = \Sigma_i\, log(p'_{1i}/p_{1i}\,)\Sigma_j q'_{ij}$$

and that for each female *i* the sum through males of the observed mating frequencies involving female *i* is, by definition, p'_{1i} i.e. $\Sigma_j q'_{ij} = p'_{1i}$ and similarly for each male *j* we have $\Sigma_i q'_{ij} = p'_{2j}$. Then

$$\Sigma_{ij}q'_{ij}log(p'_{1i}/p_{1i}\,) = \Sigma_i\, log(p'_{1i}/p_{1i}\,)p'_{1i}$$

and

$$\Sigma_{ij}q'_{ij}log(p'_{2j}\,/\,p_{2j}\,) = \Sigma_j log(p'_{2j}\,/\,p_{2j}\,)p'_{2j}$$

therefore

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

$$\Sigma_i\, log(p'_{1i}/p_{1i}\,)p'_{1i} - \Sigma_i p'_{1i}log(p'_{1i}/p_{1i}) + \Sigma_j log(p'_{2j}\,/\,p_{2j}\,)p'_{2j} - \Sigma_j p'_{2j}log(\,(p'_{2i}\,/\,p_{2i}\,) = 0$$

bioRxiv preprint doi: https://doi.org/10.1101/095901; this version posted August 3, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

and so the proposition is true

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

Proposition 2

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

where

$$p' = p'_1 p'_2$$

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

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

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

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

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

bioRxiv preprint doi: https://doi.org/10.1101/095901; this version posted August 3, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

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.

Carvajal-Rodriguez, A., Rolan-Alvarez, E., 2006. JMATING: a software for the analysis of sexual selection and sexual isolation effects from mating frequency data. BMC Evol Biol. 6, 40.

Carvajal-Rodriguez, A., Rolán-Alvarez, E., 2014. A comparative study of Gaussian mating preference functions: a key element of sympatric speciation models. Biological Journal of the Linnean Society. 113, 642-657.

Dall, S. R. X., et al., 2005. Information and its use by animals in evolutionary ecology. Trends in Ecology & Evolution. 20, 187-193.

Darwin, C., 1871. The descent of man, and selection in relation to sex. Murray.

Edward, D. A., 2015. The description of mate choice. Behavioral Ecology. 26, 301-310.

Evren, A., Tuna, E., 2012. On some properties of goodness of fit measures based on statistical entropy. International Journal of Research and Reviews in Applied Sciences. 13, 192-205.

Frank, S. A., 2009. Natural selection maximizes Fisher information. J Evol Biol. 22, 231-44.

Frank, S. A., 2012a. Natural selection. IV. The Price equation. J Evol Biol. 25, 1002-19.

Frank, S. A., 2012b. Natural selection. V. How to read the fundamental equations of evolutionary change in terms of information theory. J Evol Biol. 25, 2377-96.

Frank, S. A., 2013. Natural selection. VI. Partitioning the information in fitness and characters by path analysis. Journal of Evolutionary Biology. 26, 457-471.

Gavrilets, S., 2004. Fitness landscapes and the origin of species. Princeton University Press, Princeton, N.J.

Hollocher, H., et al., 1997. Incipient Speciation by Sexual Isolation in Drosophila melanogaster: Variation in Mating Preference and Correlation Between Sexes. Evolution. 51, 1175-1181.

Hughes, A. L., 2015. Sexual Selection and Mate Choice: Insights from Neutralist Perspectives. Evolutionary Biology. 42, 366-378.

Janicke, T., et al., 2016. Darwinian sex roles confirmed across the animal kingdom. Science Advances. 2.

Kirkpatrick, M., 1982. Sexual selection and the evolution of female choice. Evolution. 1-12.

Kuijper, B., et al., 2012. A guide to sexual selection theory. Annual Review of Ecology, Evolution, and Systematics. 43, 287-311.

Lande, R., 1981. Models of speciation by sexual selection on polygenic traits. Proceedings National Academy of Sciences, USA. 78, 3721-3725.

bioRxiv preprint doi: https://doi.org/10.1101/095901; this version posted August 3, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Pérez-Cruz, F., Kullback-Leibler divergence estimation of continuous distributions. 2008 IEEE international symposium on information theory. IEEE, 2008, pp. 1666-1670.

Pomiankowski, A., Iwasa, Y., 1998. Runaway ornament diversity caused by Fisherian sexual selection. Proceedings of the National Academy of Sciences of the United States of America. 95**,** 5106-5111.

Price, G. R., 1972. Extension of covariance selection mathematics. Annals of human genetics. 35**,** 485-490.

Prum, R. O., 2010. The Lande-Kirkpatrick mechanism is the null model of evolution by intersexual selection: implications for meaning, honesty, and design in intersexual signals. Evolution. 64**,** 3085-3100.

Roff, D. A., Fairbairn, D. J., 2014. The evolution of phenotypes and genetic parameters under preferential mating. Ecology and Evolution. 4**,** 2759-2776.

Rolán-Alvarez, E., Caballero, A., 2000. Estimating sexual selection and sexual isolation effects from mating frequencies. Evolution. 54**,** 30-6.

Roughgarden, J., et al., 2015. Sexual selection studies: a NESCent catalyst meeting. PeerJ PrePrints 3:e680v3.

Schwarz, G., 1978. Estimating the dimension of a model. The Annals of Statistics. 6**,** 461-464.

Sokal, R. R., Rohlf, F. J., 1981. Biometry. W. H. Freeman and Co., New York.

Wagner, G. P., 2010. The measurement theory of fitness. Evolution. 64**,** 1358-1376.