

Article

On Non-Random Mating, Adaptive Evolution, and Information Theory

Antonio Carvajal-Rodríguez 

Centro de Investigación Mariña (CIM), Departamento de Bioquímica, Genética e Inmunología, Universidade de Vigo, 36310 Vigo, Spain; acraaj@uvigo.es

Simple Summary: The evolutionary process can be seen as a process of acquisition, storage, and updating of information by a population about the environment in which it lives. In this paper, I propose a model that starts with the distribution of mating that occurs according to mutual mating fitness and ends with the distribution of viable adult genotypes obtained after this mating. The result of the evolutionary dynamics associated with each stage of the model can be described in terms of information. This informational description facilitates the connection between cause and effect, as well as the development of statistics to test the null model of zero information, i.e., random mating and/or no effect of natural selection. Incorporating the informational perspective into the mathematical formalism of population genetics/genomics contributes to clarifying, expanding, and deepening the mathematical description of evolutionary theory.

Abstract: Population genetics describes evolutionary processes, focusing on the variation within and between species and the forces shaping this diversity. Evolution reflects information accumulated in genomes, enhancing organisms' adaptation to their environment. In this paper, I propose a model that begins with the distribution of mating based on mutual fitness and progresses to viable adult genotype distribution. At each stage, the changes result in different measures of information. The evolutionary dynamics at each stage of the model correspond to certain aspects of interest, such as the type of mating, the distribution of genotypes in regard to mating, and the distribution of genotypes and haplotypes in the next generation. Changes to these distributions are caused by variations in fitness and result in Jeffrey's divergence values other than zero. As an example, a model of hybrid sterility is developed of a biallelic locus, comparing the information indices associated with each stage of the evolutionary process. In conclusion, the informational perspective seems to facilitate the connection between cause and effect and allows the development of statistical tests to perform hypothesis testing against zero-information null models (random mating, no selection, etc.). The informational perspective could contribute to clarify, deepen, and expand the mathematical foundations of evolutionary theory.

Keywords: non-random mating; population genetics model; Kullback–Leibler divergence; information theory



Citation: Carvajal-Rodríguez, A. On Non-Random Mating, Adaptive Evolution, and Information Theory. *Biology* **2024**, *13*, 970. <https://doi.org/10.3390/biology13120970>

Academic Editor: Jacques Demongeot

Received: 31 October 2024

Revised: 18 November 2024

Accepted: 22 November 2024

Published: 25 November 2024



Copyright: © 2024 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The evolutionary process involves the accumulation of information in genomes, which allows living beings to improve their adaptation to the environment. We can trace the mathematical treatment of information back to Shannon [1], who, building on previous work by Nyquist [2] and Hartley [3] on communication and information, published his mathematical theory of communication in 1948, which was soon expanded and became useful across a wide range of fields. In biology, there is an abundance of studies proposing different descriptions and uses of information theory in a biological context, see for example [4–6]. In population genetics, one of the first authors to consider natural selection as a process of obtaining information about the environment during adaptive evolution

was Kimura [7]. Spetner [8] also discussed the transmission of information during adaptive evolution and, later, Maynard Smith [9] and Szathmáry [10] recognized the potential relevance of the concept of information for evolutionary theory. Mac Kay [11] used this formalism to show that recombination allows the acquisition of information due to natural selection at a faster rate than parthenogenesis. In the new century, Adami demonstrated the connection between information theory and the evolution of biological complexity and molecular biology [12–14]. In 2001, Frieden et al. [15] posed evolutionary dynamics as a process that maximizes information and Streliaoff et al. [16] described the mutual information between loci that captures epistasis. However, it seems to be Frank [17–19] who first showed the equivalence between the description in terms of information and the classical dynamic description of population genetics. More recently, other authors have developed and deepened similar ideas [20–22].

Focusing on the selective component of the evolutionary process, the classical dynamic description starts with initial frequencies, based on which the force of natural selection acts, allowing us to predict the final frequencies. However, given the significant difficulty in directly measuring the effect of natural selection, the informational approach becomes relevant, where we consider the fluctuation between the frequencies of the ancestral population and the new one, to obtain the information associated with the natural selection process that has caused that fluctuation [17,20]. Thus, we can obtain information about natural selection from the observed frequencies in two separate population groups, either in space or in time.

In summary, population genetics is the formal theory of evolution, focusing on the concepts of natural selection, genetic drift, mutation, and recombination, etc., and the relationship of these concepts with adaptation to the environment and their effect on genetic diversity. However, in recent years, a reformulation of the evolutionary process in terms of information is gaining ground.

This information-based approach has recently been extended to calculate the information gained by comparing the distribution of phenotypes in regard to mating with their distribution in the population. This has enabled the derivation of sexual selection and assortative mating statistics, as well as the design of a hierarchy of models that have enabled the application of multi-model selection techniques to the study of non-random mating [23–25].

To develop this description in terms of information, the connection between Price's equation, which describes the mean phenotypic change due to the effect of selection, and Jeffrey's divergence was utilized, as established in the work by Frank [18].

In this paper, we propose a model that starts with the distribution of mating that occurs according to mutual mating fitness and ends with the distribution of viable adult genotypes obtained after this mating. The evolutionary dynamics associated with each stage of the model can be described in terms of information.

The structure of the paper is as follows: first, we will briefly review Price's equation and the notions of entropy and information. Second, we will review recent results related to mating fitness and non-random mating to highlight various connections to some measures and properties of information theory. Third, we propose the inclusion of a mating fitness concept within a more general model that will incorporate the genotype–phenotype map and other elements of the life cycle, such as viability. We will show that the mean change in fitness, caused by the shift in genotypic and haplotypic distributions due to sexual and natural selection, can be expressed as Jeffrey's divergence. We will explicitly calculate the information for the simple case of a biallelic locus with non-random mating, according to which we will demonstrate the different information measures obtained for a model of hybrid sterility. Finally, we will briefly comment, in light of the above and recent literature, on whether the concept of information acquisition could serve as an alternative to biological fitness as a design principle and the central axis of evolutionary dynamics.

2. Price's Equation

Price's equation is a mathematical identity that, in its most general form, can be applied to any characteristic that classifies two sets of entities separated in time and/or space [18,19,26–29]. Furthermore, Price's equation has been considered to be a fundamental, unifying, and model-generating theorem in evolutionary theory [30–32]. Indeed, key theorems of evolution, such as the gene selection theorem (average excess), the fundamental theorem of phenotypic selection (Robertson's equation), Breeder's equation, and Fisher's fundamental theorem, can be derived from it [33], although regarding the latter, see the recent work by Ewens [34].

Price's equation can be interpreted in terms of what Kuhn [35] called symbolic generalizations or generalization sketches and, as such, it has been connected with many different types of models, based on which new variants of the equation can be derived to solve different types of problems (reviewed in [30]). Furthermore, Price's equation can also be interpreted in terms of information [18] and, as we have already mentioned, this interpretation has recently been leveraged to generate estimators of sexual selection patterns and assortative mating [23–25].

Price's equation can be formalized in different ways; here, we will follow Frank's notation to express it as the mean change between two sets of entities for a character, z , indexed by different classes, i , as follows:

$$\Delta \bar{z} = \text{cov}(\omega, z) + E(\omega, \Delta z) \quad (1)$$

where $\omega = \frac{w}{\bar{w}}$.

3. Entropy and Information

Information in this regard is concerned with the decrease in entropy or uncertainty in a system. Shannon [1] defined the uncertainty or entropy of a discrete random variable as X , which takes values x_i with probability p_i , as:

$$H(X) = -\sum_{i=1}^N p_i \log(p_i) \quad (2)$$

3.1. Mutual Information

The maximum uncertainty for a variable with N possible states is $H_{\max} = \log(N)$ [36] and we can define the information, I , in a system simply as:

$$I(X) = H_{\max} - H(X).$$

It is also possible to measure the uncertainty of X conditional upon another random variable, Y :

$$H(X|Y) = -\sum_y p(y) \sum_x p(x|y) \log(p(x|y)) \quad (3)$$

and then the mutual information between X and Y is the reduction in uncertainty about X , thanks to the knowledge we have about:

$$I(X;Y) = H(X) - H(X|Y) \quad (4)$$

3.2. Relative Entropy or Kullback–Leibler Divergence

Kullback–Leibler divergence (D_{KL}) or relative entropy allows us to compare two distributions, P and Q , and quantifies the information gained by using P instead of Q .

$$D_{KL}(P||Q) = \sum_x P(x) \log\left(\frac{P(x)}{Q(x)}\right) \quad (5)$$

It can be shown (the information inequality theorem) that $D_{KL} \geq 0$, the equality being fulfilled only when P and Q are equal.

The relationship between D_{KL} and mutual information is:

$$I(X; Y) = D_{KL}(P(X, Y) || P(X)P(Y)) \quad (6)$$

That is, mutual information quantifies the information gained by using the joint distribution of X and Y , instead of the product of the two marginal distributions.

We can also relate D_{KL} to information and maximum entropy, since:

$$I(X) = H_{max} - H(X) = D_{KL}(X || U)$$

where U is the uniform distribution over X [36].

3.3. Jeffrey's Divergency

Jeffrey's divergence is the symmetrical version of the D_{KL} divergence, that is, the sum of the two possible divergences between P and Q and, therefore, it measures the total information that changes when using Q instead of P , plus that of P instead of Q :

$$J(P, Q) = D_{KL}(P || Q) + D_{KL}(Q || P) \quad (7)$$

As a curiosity, J is also known as the Population Stability Index in the world of financial risk modeling and analysis.

Both D_{KL} and J divergences are particular cases of the f divergence, which is defined for the discrete case as:

$$D_f(P || Q) = \sum_x Q(x) f\left(\frac{P(x)}{Q(x)}\right)$$

where f is the generating function, which for D_{KL} is $f(x) = x \log x$ and for J is $f(x) = (x - 1) \log x$.

4. Information-Related Interpretation of Price's Equation

The first connection between Price's equation and information metrics was established by Frank [17], connecting Price's equation with Fisher's information. The relationship between Price's equation and relative entropy and Jeffrey's (J) divergence are discussed in later work by the same author [18,19,37,38].

To see the connection with the J divergence, let us return to Equation (1), focusing on the first term, which expresses the average change in the character z caused by selection [18]:

$$\begin{aligned} \Delta \bar{z}_s &= cov(\omega, z) = \sum p_i (z_i - \bar{z})(\omega_i - \bar{\omega}) \\ \Delta \bar{z}_s &= \sum p_i (z_i - \bar{z})(\omega_i - 1) = \sum p_i (z_i - \bar{z}) \left(\frac{p'_i}{p_i} - 1 \right) = \sum \Delta p_i z_i \end{aligned}$$

Note that the relative fitness is $\omega_i = p'_i / p_i$, where p_i is the frequency of class i before selection and p'_i is the frequency after selection. The average relative fitness is equal to 1, $\Delta p_i = p'_i - p_i$ and $\sum \Delta p_i = 0$.

If we take as character z , the logarithm of the relative fitness, we obtain the expression of the first component of Price's equation in terms of information [18]:

$$\Delta \bar{z}_s = \sum \Delta p_i z_i = \sum \Delta p_i \log(\omega_i) = \sum \Delta p_i \log\left(\frac{p'_i}{p_i}\right) = J(P', P) \quad (8)$$

We then see that the mean change in the logarithm of relative fitness is the difference in Jeffrey's divergence between the frequency distribution after selection (P') and before selection (P).

It is interesting to note that:

$$\Delta p = \sum_i p'_i - p_i = \sum_i \omega_i p_i - p_i = \sum_i p_i (\omega_i - 1)$$

and

$$\Delta \bar{z}_s = \sum_i \Delta p_i z_i = \sum_i p_i (\omega_i - 1) \log(\omega_i) = \sum_i p_i f(\omega_i)$$

where f is the f divergence generating function for J , so Equation (8) is an f divergence with generating function $(\omega - 1) \log \omega$, which explains that the evolution of the distribution caused by the fitness parameter in replicator-type models can always be expressed as Jeffrey's divergence.

Recently, Carvajal-Rodríguez used these ideas to reinterpret Price's equation in terms of the comparison between two sets of pairs, one representing random mating (Q) and the other (Q') representing mating according to mutual mating fitness [23–25,39].

5. Information and Non-Random Mating

Sexual selection, defined as any type of selection arising from differential fitness in regard to access to gametes for fertilization [40], can essentially be driven by two biological processes involving non-random mating: competition for mates and mate choice ([25] and the references therein).

In previous work, which has already been mentioned, Carvajal-Rodríguez proposed a simple and abstract framework to describe mate competition and mate choice, along with their consequences in terms of sexual selection and assortative mating. This model is based on the concept of mutual and individual mating fitness. The effects of variation in these fitness measures are described in terms of sexual selection and assortative mating, quantified by the increase in information gained compared to random mating (zero information).

In what follows, we will review how the informational interpretation of Price's equation connects with the information generated by non-random mating.

Let $q'(x)$ be the frequency of females in mating, previously noted as p'_1 in [23], and $q'(y)$ the frequency of males in mating (previously noted as p'_2). Similarly, let $q(x)$ be the population frequency of females (previously p_1) and $q(y)$ the population frequency of males (previously p_2). The joint distribution of random mating is $q(x,y)$, where $q(x,y) = q(x)q(y)$. Then, the joint distribution, $q'(x,y)$, according to mutual mating fitness, will be equal to [23]:

$$q'(x,y) = m_{xy} q(x,y) \quad (9)$$

Model (9) relates the frequency, $q'(x,y)$, of mating guided by the mutual fitness, m_{xy} , for mating between females in class x and males in class y , with the frequency $q(x,y) = q(x)q(y)$ of random mating between those female and males.

For each pair of classes, we can define a feature or property, $z(x,y)$, about those classes, such that Price's equation for the mean change in z caused by the variation in the mutual propensities, m , is:

$$\begin{aligned} \Delta \bar{z}_m &= \text{cov}(m, z) = \sum q(x,y) (z(x,y) - \bar{z}) (m_{xy} - \bar{m}) \\ \Delta \bar{z}_m &= \sum q(x,y) (z(x,y) - \bar{z}) (m_{xy} - 1) \\ &= \sum q(x,y) (z(x,y) - \bar{z}) \left(\frac{q'(x,y)}{q(x,y)} - 1 \right) \\ \Delta \bar{z}_m &= \sum \Delta q_{xy} z(x,y) \end{aligned}$$

where $\Delta q_{xy} = q'(x,y) - q(x,y)$ and $\sum \Delta q_{xy} = 0$.

So, taking $z(x,y) = \log(m_{xy}) = \log(q'(x,y)/q(x,y))$ we get:

$$\Delta \bar{z}_m = \sum \Delta q_{xy} z(x,y) = \sum \Delta q_{xy} \log \left(\frac{q'(x,y)}{q(x,y)} \right) = J(Q', Q) \equiv J_{PTI} \quad (10)$$

as obtained in [23].

Moreover, $q'(x,y)/q(x,y)$, when calculated based on observed frequencies, coincides with the pair total index (PTI) defined in [41]; $J(Q',Q)$, which has previously been called J_{PTI} , and here we will maintain that notation, so $J(Q',Q) \equiv J_{PTI}$.

The J_{PTI} divergence in (10) can be decomposed into the sum of the information associated with sexual selection (J_{PSS}), plus the information on assortative mating (J_{PSI}), plus an interaction that arises when both sexual selection and assortative mating are present [23,25]. However, when the covariance within pairs is zero, J_{PTI} is reduced to the information corresponding to sexual selection, denoted as J_{PSS} (see below). When the covariance within pairs is not zero, but the mean and variance between mating and the population are the same, J_{PTI} is reduced to the information corresponding to assortative mating, referred to as J_{PSI} [25].

5.1. Information and Sexual Selection

As we have said, J_{PSS} information can be obtained from J_{PTI} when the covariance in the pairing is 0. This information is the sexual selection information, J_{PSS} , and it compares the frequencies in the pairing $q'(x)q'(y)$ with the population frequencies $q(x)q(y)$. In [23], it is proven for the discrete case, that J_{PSS} is, in turn, the sum of the sexual selection information within each sex, $J_{PSS} = J_{S1} + J_{S2}$, and, in [25], the same is proven for the continuous case. Here, we will provide another demonstration for both cases, continuous and discrete, based on the chain rule for relative entropy, D_{KL} [36].

We know that:

$$J_{PSS} = J(r(x,y), q(x,y)) = D_{KL}(r(x,y)||q(x,y)) + D_{KL}(q(x,y)||r(x,y)) \quad (11)$$

where $r(x,y) = q'(x) \times q'(y)$ and $q(x,y) = q(x) \times q(y)$.

And considering each sex separately:

$$\begin{aligned} J_{S1} &= D_{KL}(q'(x)||q(x)) + D_{KL}(q(x)||q'(x)) \\ J_{S2} &= D_{KL}(q'(y)||q(y)) + D_{KL}(q(y)||q'(y)) \end{aligned}$$

We want to show that $J_{PSS} = J_{S1} + J_{S2}$. To do this, let us consider the conditional probabilities:

$$\begin{aligned} q(y|x) &= q(x,y)/q(x) = \frac{q(x)q(y)}{q(x)} = q(y) \\ r(y|x) &= r(x,y)/q'(x) = \frac{q'(x)q'(y)}{q'(x)} = q'(y) \end{aligned} \quad (12)$$

According to the chain rule [36], we know that:

$$D_{KL}(r(x,y)||q(x,y)) = D_{KL}(q'(x)||q(x)) + D_{KL}(r(y|x)||q(y|x))$$

which due to (12) is as follows:

$$D_{KL}(r(x,y)||q(x,y)) = D_{KL}(q'(x)||q(x)) + D_{KL}(q'(y)||q(y))$$

Similarly:

$$D_{KL}(q(x,y)||r(x,y)) = D_{KL}(q(x)||q'(x)) + D_{KL}(q(y)||q'(y))$$

Therefore, if we add both divergences and group the x and y components separately, we see that:

$$J_{PSS} = D_{KL}(r(x,y)||q(x,y)) + D_{KL}(q(x,y)||r(x,y)) = J_{S1} + J_{S2}.$$

Sexual Selection in for Normally Distributed Traits

If traits x and y are normally distributed, such that mating frequencies, q' , follow an $N(\mu_1, \Sigma_1)$ distribution, and population frequencies, q , follow an $N(\mu_2, \Sigma_2)$ distribution, where Σ_i is the variance–covariance matrix; then assuming that phenotypic changes are gradual [25], the components J_{S1} and J_{S2} coincide with the opportunity for sexual selection index [42].

On the other hand, if we consider the case where the variance in mating is equal to the population variance, $\sigma_{1xx} = \sigma_{2xx} = \sigma_{xx}$ and $\sigma_{1yy} = \sigma_{2yy} = \sigma_{yy}$, this means that J_{S1} and J_{S2} equal the square of the standardized sexual selection indices for females (x) and males (y), respectively. Furthermore, in this case, if there is no assortative mating, the relative entropy is symmetrical, so that $D_{KL}(q'|q) = D_{KL}(q|q')$, such that $J_{PSS} = 2D_{KL}(q'|q)$. Therefore, when the mean of the trait is the only difference between the mating and the population frequencies, the sum of the squares for both sexes in terms of the standardized sexual selection indices is equal to twice the relative entropy between the mating frequencies and population frequencies.

To demonstrate the above, we first obtain the J_{PSS} , assuming that both traits are normally distributed. To do so, we note that the relative entropy between the Q' distribution of the expected pairings, based on mutual fitness, and the Q distribution of expected pairings by chance is [25]:

$$D_{KL}(q'|q) = \frac{1}{2} \left(\log(A) - 2 + \frac{\sigma_{1xx}}{\sigma_{2xx}} + \frac{\sigma_{1yy}}{\sigma_{2yy}} + \frac{(\mu_{2x} - \mu_{1x})^2}{\sigma_{2xx}} + \frac{(\mu_{2y} - \mu_{1y})^2}{\sigma_{2yy}} \right) \quad (13)$$

where:

$$A = \frac{\sigma_{2xx}\sigma_{2yy}}{\sigma_{1xx}\sigma_{1yy} - \sigma_{1xy}^2}$$

where it has been taken into account that the population covariance in the case of random encounters is 0 i.e., $\sigma_{2xy} = 0$.

The other relative entropy is:

$$D_{KL}(q|q') = \frac{1}{2} \left(\log\left(\frac{1}{A}\right) - 2 + B \right)$$

where:

$$B = \frac{\sigma_{1xx}\sigma_{2yy} + \sigma_{1yy}\sigma_{2xx}}{\sigma_{1xx}\sigma_{1yy} - \sigma_{1xy}^2} + \frac{\sigma_{1yy}(\mu_{1x} - \mu_{2x})^2 + \sigma_{1xx}(\mu_{1y} - \mu_{2y})^2 - 2\sigma_{1xy}(\mu_{1x} - \mu_{2x})(\mu_{1y} - \mu_{2y})}{\sigma_{1xx}\sigma_{1yy} - \sigma_{1xy}^2}$$

We have already seen that if there is no assortative mating, $q'(x,y) = q'(x)q'(y) = r(x,y)$, and the covariance within the mating is zero, $\sigma_{1xy} = 0$, then:

$$\begin{aligned} D_{KL}(q'|q) &= D_{KL}(r|q) = \frac{1}{2} \left(\log(A) - 2 + \frac{\sigma_{1xx}}{\sigma_{2xx}} + \frac{\sigma_{1yy}}{\sigma_{2yy}} + \frac{(\mu_{2x} - \mu_{1x})^2}{\sigma_{2xx}} + \frac{(\mu_{2y} - \mu_{1y})^2}{\sigma_{2yy}} \right) \\ D_{KL}(q|q') &= D_{KL}(q|r) = \frac{1}{2} \left(\log\left(\frac{1}{A}\right) - 2 + \frac{\sigma_{2xx}}{\sigma_{1xx}} + \frac{\sigma_{2yy}}{\sigma_{1yy}} + \frac{(\mu_{1x} - \mu_{2x})^2}{\sigma_{1xx}} + \frac{(\mu_{1y} - \mu_{2y})^2}{\sigma_{1yy}} \right) \end{aligned} \quad (14)$$

and:

$$\begin{aligned} J_{PSS} &= D_{KL}(r|q) + D_{KL}(q|r) \\ J_{PSS} &= -1 + \frac{1}{2} \left(\frac{\sigma_{1xx}}{\sigma_{2xx}} + \frac{\sigma_{2xx}}{\sigma_{1xx}} \right) + \frac{1}{2} (\mu_{1x} - \mu_{2x})^2 \left(\frac{1}{\sigma_{1xx}} + \frac{1}{\sigma_{2xx}} \right) + \\ &\quad -1 + \frac{1}{2} \left(\frac{\sigma_{1yy}}{\sigma_{2yy}} + \frac{\sigma_{2yy}}{\sigma_{1yy}} \right) + \frac{1}{2} (\mu_{1y} - \mu_{2y})^2 \left(\frac{1}{\sigma_{1yy}} + \frac{1}{\sigma_{2yy}} \right) \end{aligned} \quad (15)$$

where we have regrouped the terms with x and y , so that they correspond to $J_{S1} + J_{S2}$.

Now, in the case where $\sigma_{1xx} = \sigma_{2xx} = \sigma_{xx}$ and $\sigma_{1yy} = \sigma_{2yy} = \sigma_{yy}$, we see that by substituting Equation (14) we obtain $D_{KL}(q' || q) = D_{KL}(q || q')$ and by substituting Equation (15) we obtain $J_{PSS} = SSI_1^2 + SSI_2^2$, where SSI_1 and SSI_2 are the standardized sexual selection indices for females and males, respectively, and, therefore, we get:

$$J_{PSS} = SSI_1^2 + SSI_2^2 = 2D_{KL}(q' || q)$$

That is, the information on sexual selection, when there are no covariances and the variances do not change in regard to the mating with respect to the population, is the sum of the square of the standardized selection indices, which in turn coincides with twice the relative entropy between the mating frequencies and the population frequencies.

5.2. Information and Assortative Mating

The J_{PTI} divergence, when we ignore differences between the mating and population frequencies and focus only on frequencies within mating, becomes the J_{PSI} divergence, which compares, within mating, the distribution due to mutual mating fitness with random mating:

$$J_{PSI} = J(q'(x, y), r(x, y))$$

In previous work, J_{PSI} has been solved for both the discrete case [23] and the continuous case [25], based on the concept of mutual mating fitness, along with the informational interpretation of Price's equation. Here, for the continuous case and assuming normal distribution applies, we will derive J_{PSI} directly from the sum of the relative entropy values when the mean and variance between the mating sample and the population are equal and we will see that it corresponds to a function of the square of the correlation ρ , which is interesting because ρ is typically used to detect assortative mating, but here we will show that the information corresponding to that pattern is actually captured by $\rho^2/(1 - \rho^2)$.

As before, we assume that x and y are normally distributed, such that q' is $N(\mu_1, \Sigma_1)$ and q is $N(\mu_2, \Sigma_2)$, where Σ_i is the variance–covariance matrix. As we have seen, the J_{PSI} divergence focuses only on the mating and the comparison within mating involves two groups of equal means and variances for the trait, so we assume $\mu_1 = \mu_2$, $\sigma_{1xx} = \sigma_{2xx} = \sigma_{xx}$ and $\sigma_{1yy} = \sigma_{2yy} = \sigma_{yy}$, so what will vary is the covariance within mating, σ_{1xy} .

Then, from Equation (13) we get:

$$D_{KL}(q' || q) = \frac{1}{2} \log(A)$$

and the other divergence:

$$\begin{aligned} D_{KL}(q || q') &= \frac{1}{2} \left(\log\left(\frac{1}{A}\right) - 2 + \frac{2\sigma_{xx}\sigma_{yy}}{\sigma_{xx}\sigma_{yy} - \sigma_{1xy}^2} \right) = \frac{1}{2} \log\left(\frac{1}{A}\right) - 1 + \frac{\sigma_{xx}\sigma_{yy}}{\sigma_{xx}\sigma_{yy} - \sigma_{1xy}^2} \\ J_{PSI} &= D_{KL}(q' || q) + D_{KL}(q || q') \\ J_{PSI} &= -1 + \frac{\sigma_{xx}\sigma_{yy}}{\sigma_{xx}\sigma_{yy} - \sigma_{1xy}^2} = \frac{\sigma_{1xy}^2}{\sigma_{xx}\sigma_{yy} - \sigma_{1xy}^2} = \frac{\rho^2}{1 - \rho^2} \end{aligned} \quad (16)$$

We see that assuming normal distribution in regard to the mating, the assortative mating information is a function of the squared correlation coefficient [43].

6. Information and Evolution of Gene Frequency in Terms of a General Mating and Fitness Scheme

We have seen that the flow of information generated by the deviation from random mating is calculated by comparing the expected distribution, according to mating fitness, with the expected distribution of random mating. This information is obtained in the form of J_{PTI} divergence, regardless of whether the distribution is discrete or continuous, based

on which appropriate statistics are generated for each case [23,25]. However, model (9), according to which we derived J_{PTI} in (10), describes phenotypic frequencies (or classes in an infinitesimal interval in the continuous case) without making any assumptions about the underlying genetic model, or connecting it with the biological fitness of the offspring from those instances of mating.

The aim of this section is precisely to connect the description of the mating distribution with the description of the evolution of gene frequencies subject to natural selection and to quantify all of this in terms of information. That is, to generate a model of gene frequency change for any explicit mating model and multilocus diploid genetic system and describe it in terms of information. This remains a pending task for classical population genetics and speciation models [44]; although, recently, information-based models have been developed that transcend the limitations of previous ones. See, for example, [20,22]. These studies use D_{KL} as a metric that captures the information due to natural selection. However, the distributions these authors use are different; for instance, Smith calculates the information associated with the trajectories by jointly considering the transitions and their states. In any case, the choice of D_{KL} is ad hoc in the sense that it is defined a priori to compare distributions and does not arise naturally from a model of change due to selection, as J divergence does in the case of [18], or in the case of deviations from random mating and sexual selection [23,25].

In what follows, I will propose a model that explicitly incorporates the mating system (the interindividual mating interaction) and that implicitly incorporates the genomic context (mutation, linkage, etc.) into the classical replicator system. From this model, we will see that the mean change in the log fitness caused by the change, at different stages, in the distributions in terms of the mating system, genotypes, and haplotypes, can be captured as information in the form of Jeffrey's divergence.

We will solve the model for the simple case of one locus: (1) with random mating, thereby recovering the classical replicator system; and (2) with non-random mating, but neutral in terms of viability, thus recovering the mating fitness model defined in [23]. In the latter case, we will use as an example a simple one-locus hybrid sterility model to compare the information on the different Jeffrey's divergences obtained.

6.1. General Model

Let us consider a haplotype H (Table 1) of a known frequency in generation t , $f(H)_t$, as well as the frequencies of the corresponding diploid genotypes, HH , Hh , and hh , where h represents any haplotype other than H . Likewise the joint distribution $q'(x,y)$, according to the mutual mating fitness for classes HH , Hh , and hh , is given by (9):

$$q'(x,y) = m_{xy}q(x,y)$$

where x represents any female belonging to any of the three classes and y represents a male. Then, the general equation for the frequency of H in the next generation is:

$$f(H)_{t+1} = \sum_{x,y} q'(x,y) \times \Phi_x \times \left(\omega_{GHH} \pi_{1Hxy} + \sum_h \frac{\pi_{2Hhxy} \omega_{GHh}}{2} \right) \quad (17)$$

Equation (17) connects the haplotype frequency $f(H)$ in the next generation with the distribution of the mating $q'(x,y)$, the fecundity Φ_x of female x , and the G genotypes of the offspring, namely the carriers of the haplotype. Of these genotypes a proportion, π_1 , will be homozygous and, π_2 , heterozygous. The genotypes will survive based on their relative fitness ω (viability) after the action of natural selection.

Table 1. A table of main symbols used in Section 6.

Symbol	Meaning
q	Mating frequency if mating is random.
m_{xy}	Relative mutual mating fitness between female with phenotype x and male with phenotype y , i.e., the propensity of mating with fertilization between a given pair of phenotypic classes.
q'	Mating frequency after the action of mutual mating fitness.
Φ_G	Fecundity, which is included within mutual mating fitness.
M	Average mutual mating fitness.
m_G	The relative (divided by the average M) marginal mating fitness of genotype G .
ω_G	The relative (divided by the average fitness W) genotype viability, i.e., the potential for survival to adulthood of a given genotype.
Ω_{xyG}	The product of m_{xy} Φ_G ω_G .
H	A given haplotype, i.e., a sequence that carries a particular set of genes that tend to be inherited together.
Π_{1Hxy}	A fraction of the offspring from the mating (x, y) that are homozygous for the H haplotype.
Π_{2Hxy}	A fraction of the offspring from the mating (x, y) that are heterozygous for the H haplotype.
p_1, p_2, p_3	For one locus, frequencies of the genotypes AA , Aa , and aa , respectively. These symbols are used to simplify the notation of mating frequencies in regard to single-locus examples.

The equation is general in the sense that the haplotype component, H , incorporates any multilocus genetic system and the q' component can incorporate any mating scheme. However, it is true that the values of π will depend on any mutations or linkage disequilibrium, which in turn depends on the recombination frequency. Therefore, the model will not be dynamically sufficient unless we specify functions for the proportions π , which will depend on the case being studied. In each case, we will obtain new statistics for detecting sexual selection and assortative mating, as well as new insights into the underlying evolutionary dynamics.

If there are only two types of haplotypes, the above formula can be simplified as:

$$f(H)_{t+1} = f(HH)_{t+1} + \frac{f(Hh)_{t+1}}{2} = \sum_{x,y} q'(x,y) \times \Phi_x \times \left(\omega_{GHH} \pi_{1Hxy} + \omega_{GHh} \frac{\pi_{2Hxy}}{2} \right)$$

We will now apply (17) to the simple case of a biallelic locus without mutation.

6.2. One Locus Model

Let us consider a biallelic locus that codes for a discrete trait with three phenotypes, which are directly associated with their corresponding genotypes and follow a mating scheme described by the mutual mating fitnesses, m_{xy} . Within these mating fitnesses, we also include the fecundity component Φ . We wish to determine the frequency $f(A)$ of the allele A after mating, and after fecundity and viability selection. The frequency $f(A)$ after selection will be the sum of the frequency of the new AA genotypes, plus half of the Aa genotypes. The structure of the model is as follows:

$$f(A)_{t+1} = p' = f(AA)_{t+1} + \frac{f(Aa)_{t+1}}{2} = \sum_{xy} q'_{xy} \times \left(\omega_{AA} \pi_{1Axy} + \omega_{Aa} \frac{\pi_{2Axy}}{2} \right) \quad (18)$$

For convenience, we have denoted $q'(x,y)$ as q'_{xy} , which therefore corresponds to Equation (9). Moreover, π_{1Axy} is the fraction of offspring from the mating $x \times y$ that are homozygous for the allele A , π_{2Axy} is the fraction of offspring from the mating $x \times y$ that are heterozygous for the allele A , ω_{AA} is the relative viability (divided by the average fitness) associated with the homozygote AA , and ω_{Aa} is the relative viability associated with the heterozygote.

To simplify the notation concerning pair parameters, we will use the subscript 1 for the AA genotype, 2 for the heterozygote, and 3 for aa . Thus, the genotype frequencies in the population before mating are p_1 (AA frequency), p_2 (Aa frequency), and p_3 (aa frequency). If mating is random and we assume that there is no selection in regard to fecundity, we take values $m_{xy} = 1$ for all x, y ; the proportion of AA and Aa genotypes in the offspring before selection and without mutations are shown in Table 2.

Table 2. Proportion $p_x p_y \pi$ of offspring genotypes AA and Aa as a function of mating frequencies, when $m_{xy} = 1$ for every x, y .

Mating	Offspring	
	π_{1A} (AA)	π_{2A} (Aa)
$p_1 \times p_1 \times m'_{11}/M$	p_1^2	0
$p_1 \times p_2 \times m'_{12}/M$	$(p_1 p_2)/2$	$(p_1 p_2)/2$
$p_1 \times p_3 \times m'_{13}/M$	0	$p_1 p_3$
$p_2 \times p_1 \times m'_{21}/M$	$(p_1 p_2)/2$	$(p_1 p_2)/2$
$p_2 \times p_2 \times m'_{22}/M$	$(p_2^2)/4$	$(p_2^2)/2$
$p_2 \times p_3 \times m'_{23}/M$	0	$(p_2 p_3)/2$
$p_3 \times p_1 \times m'_{31}/M$	0	$p_1 p_3$
$p_3 \times p_2 \times m'_{32}/M$	0	$(p_2 p_3)/2$
$p_3 \times p_3 \times m'_{33}/M$	0	0
$M = \sum_{xy} p_x p_y m'_{xy}$		

Note that $m_{xy} = m'_{xy}/M$.

If we assume that random mating is taking place, according to (9) we know that the probability of mating between x and y is $q'_{xy} = q(x,y) = p_x p_y$ and, noting the allele frequencies of the previous generation denoted as p and q (without subscripts), we get $p_1 = p^2$, $p_2 = 2pq$, and $p_3 = q^2$. Therefore, substituting the values from Table 2 into Equation (18) gets:

$$p' = [(p^4 + 2p^3q + p^2q^2)(1 - \mu_{1-}) + \mu_{1-}f_1(p, q)]\omega_{AA} + \frac{1}{2}[(2p^3q + 4p^2q^2 + 2pq^3)(1 - \mu_{2-}) + \mu_{2-}f_2(p, q)]\omega_{Aa}$$

where μ_{1-} and μ_{2-} are any mutations producing genotype AA or Aa , respectively, μ_{1-} is any mutation from AA to any other genotype and μ_{2-} from Aa to any other genotype, and f_1 and f_2 are functions of allele frequencies.

Assuming there is no mutation, $\mu_{-} = 0$, and after simplifying, we obtain:

$$p' = p^2\omega_{AA} + pq\omega_{Aa} = p(p\omega_{AA} + q\omega_{Aa}) = pW_A$$

where $W_A = p\omega_{AA} + q\omega_{Aa}$ is the marginal fitness of allele A [45].

We see that, for a biallelic locus and with random mating, Equations (17) and (18) correspond to the basic model of selection for a biallelic locus [45,46].

Genotypic Mating Fitness

Alternatively, if we assume a neutral model, such that $\omega_{AA} = \omega_{Aa} = \omega_{aa} = 1$, Equation (18) becomes

$$p' = p^2[m_{11}p^2 + m_{12}2pq + m_{13}q^2] + pq[m_{12}p^2 + m_{22}2pq + m_{23}q^2]$$

where we have assumed that the mutual mating fitnesses between classes x and y do not depend on the sex of the genotype, that is, $m_{xy} = m_{yx}$, and that the genotypic frequencies are equal in males and females.

In these conditions, the individual mating fitness (Formula (4) in [24]) for type *AA* females is:

$$m_{FemAA} = m_{11}p_1 + m_{12}p_2 + m_{13}p_3$$

and for males:

$$m_{MaleAA} = m_{11}p_1 + m_{21}p_2 + m_{31}p_3$$

And, since $m_{xy} = m_{yx}$, we get $m_{FemAA} = m_{MaleAA} = m_{AA}$, thus:

$$m_{AA} = m_{11}p_1 + m_{12}p_2 + m_{13}p_3$$

In this context, where each phenotypic class corresponds to a genotype, the individual mating fitness m_{AA} corresponds to the genotypic mating fitness for the *AA* genotype. Then, the frequency of the *AA* genotype in regard to mating is:

$$p'_{f1} = p'_{m1} = p_1 m_{AA} = p^2 m_{AA} \quad (19)$$

Similarly, for heterozygotes, the genotypic mating fitness m_{Aa} is:

$$m_{Aa} = m_{21}p_1 + m_{22}p_2 + m_{23}p_3$$

being the frequency of *Aa* in regard to mating:

$$p'_{f2} = p'_{m2} = p^2 m_{Aa} = 2pqm_{Aa}$$

Therefore, using a non-random mating selectively neutral model, the equation for the new frequency of allele *A* in the next generation can be rewritten as:

$$p' = p'_1 + \frac{p'_2}{2} = p^2 m_{AA} + pqm_{Aa} = pM_A \quad (20)$$

where M_A is the marginal mating fitness of allele *A*.

The frequency of allele *A* in the new generation is the sum of the frequencies of genotype *AA*, plus half of the frequency of the heterozygous genotype *Aa* after mating selection. Note that if the genotypic mating fitnesses are 1, meaning that mating is random, then $p' = p$, as expected.

Similarly, for the genotype *aa*, we get:

$$m_{aa} = m_{31}p_1 + m_{32}p_2 + m_{33}p_3$$

Then, $p'_3 = q^2 m_{aa}$, and the frequency of allele *a* in the next generation is:

$$q' = q^2 m_{aa} + pqm_{Aa} = qM_a \quad (21)$$

where M_a is the marginal mating fitness of allele *a*.

Finally, the change in gene frequency due to differential mating fitness is:

$$\Delta p = p' - p = p^2(m_{AA} - 1) + pq(m_{Aa} - 1) = p(M_A - 1) \quad (22)$$

$$\Delta q = q' - q = q^2(m_{aa} - 1) + pq(m_{Aa} - 1) = q(M_a - 1)$$

The sum of frequencies $p' + q' = p'_1 + p'_2 + p'_3$ must equal 1. To check this, remember that the mutual mating fitnesses $m_{xy} = m'_{xy}/M$ are normalized (Table 2) and, therefore, the genotypic fitnesses are also $m_{AA} = m'_{AA}/M$, $m_{Aa} = m'_{Aa}/M$, and $m_{aa} = m'_{aa}/M$, where the tilde is used to indicate the unnormalized fitness. The mean fitness, $M = \sum_{xy} p_x p_y m'_{xy}$, can be expressed in terms of the unnormalized genotypic fitnesses as:

$$M = p_1 m'_{AA} + p_2 m'_{Aa} + p_3 m'_{aa} = p^2 m'_{AA} + 2pqm'_{Aa} + q^2 m'_{aa}$$

Therefore using (19)–(21), the sum $p' + q' = (p^2m'_{AA} + 2pqm'_{Aa} + q^2m'_{aa})/M = M/M = 1$.

6.3. Information and Fitness

Equation (17) can be expressed in terms of diploid genotype frequencies and combined fitness by substituting it with Equation (9), $q'_{xy} = m_{xy}q_{xy}$, where we obtain:

$$f(H)_{t+1} = \sum_{x,y} q(x,y)m_{xy}\Phi_x \left(\omega_{GHH}\pi_{1Hxy} + \sum_h \frac{\pi_{2Hhxy}\omega_{GHh}}{2} \right)$$

Therefore, since the haplotype frequency is simply the sum of the frequencies of the genotypes containing that haplotype, we can also express the frequency of any genotype G as:

$$f(G)_{t+1} = \sum_{x,y} q(x,y)m_{xy}\Phi_x\pi_{Gxy}\omega_G = \sum_{x,y} q(x,y)\pi_{Gxy}\Omega_{xyG} \quad (23)$$

where $\Omega_{xyG} = m_{xy}\Phi_x\omega_G$.

The genotype frequencies in the previous generation can be expressed as:

$$f(G)_t = \sum_{x,y} q(x,y)\pi_{Gxy}$$

For example, if $G = AA$, then $f(AA)_t$ is the sum of the expected frequency according to the chance of crosses $AA \times AA$, $Aa \times Aa$, $AA \times Aa$, and $Aa \times AA$, and the corresponding proportion of the AA genotype in the offspring of each cross, i.e.:

$$f(AA)_t = p^4 + \frac{4p^2q^2}{4} + \frac{2p^3q}{2} + \frac{2p^3q}{2} = p^2(p^2 + q^2 + 2pq) = p^2$$

Note that, in general, the overall fitness component of G , Ω_G , is different from the frequency ratio before and after selection:

$$\Omega_G \neq \frac{f(G)_{t+1}}{f(G)_t}$$

because for each mating, we have $\Omega_{xyG} = m_{xy}\Phi_x\omega_G$ that can differ for different matings x, y .

The only fitness element common to genotype G will be viability ω_G , so if we assume that there are no differences in the mating fitness or fertility, we get:

$$f(G)_{t+1} = \omega_G \sum_{x,y} q(x,y)\pi_{Gxy}$$

so:

$$\omega_G = \frac{f(G)_{t+1}}{f(G)_t}$$

And, in this latter case, the mean change in the log fitness caused by selection, due to differential viability, $\log(\omega_G)$, is:

$$J_{\omega G} = \sum_{x,y,G} (f(G)_{t+1} - f(G)_t) \log(\omega_G) \quad (24)$$

6.3.1. Information on Mating by Genotype

However, we are interested in the information due to changes in genotypic frequencies when the different fitnesses causing that change may vary. The problem is that the final frequency of each genotype, which is due to the different mating fitness of its parents, the fecundity of the mother, and ultimately the viability of the genotype itself, cannot be expressed as the result of a single measure of fitness.

One option is to obtain the information associated with each genotype separately. That is, the average change in the logarithm of the overall genotypic fitness, Ω_{xyG} , after mating (x, y) , associated with the change in frequency of that genotype from one generation to the next, is:

$$\begin{aligned} J_G &= \sum_{x,y} \left(f(G)_{(t+1)xy} - f(G)_{txy} \right) \log(\Omega_{xyG}) \\ &= \sum_{x,y} \left(q(x,y) \pi_{Gxy} \Omega_{xyG} - q(x,y) \pi_{Gxy} \right) \log(\Omega_{xyG}) \end{aligned}$$

where it is noted that:

$$f(G)_{txy} = q(x,y) \pi_{Gxy}$$

Therefore, for the entire set of possible genotypes G , whose frequency changes from one generation to the next due to the effect of selection, we get:

$$\Delta G_\Omega = \sum_G J_G \quad (25)$$

That is, if the combined effect of mating fitness, fecundity, and viability produces a shift in genotype frequencies, this can be captured as the sum of the applicable Jeffrey's divergences.

6.3.2. Genotype Distribution Information

Alternatively, and from Equation (23), we can directly consider the frequency of a genotype as:

$$f(G)_{t+1} = \sum_{x,y} q(x,y) \pi_{Gxy} \Omega_{xyG} = f(G)_t \times \sum_{x,y} f_{xy}(\gamma) \pi_{Gxy} \Omega_{xyG} = f(G)_t \times \Omega_G$$

where:

$$f_{xy}(\gamma) \Pi_{Gxy}$$

is a function of the genotype frequencies after the (x, y) mating and Ω_G is the overall marginal fitness of genotype G .

$$\Omega_G = \sum_{x,y} f_{xy}(\gamma) \pi_{Gxy} \Omega_{xyG} = \frac{f(G)_{t+1}}{f(G)_t}$$

So, the mean change in the overall log fitness caused by the variation in genotype frequencies because of selection takes the form of Jeffrey's divergence, as follows:

$$I_{\Omega G} = \sum_{\gamma} (f(\gamma)_{t+1} - f(\gamma)_t) \log(\Omega_\gamma) \quad (26)$$

6.3.3. Haplotype Distribution Information

From Equation (17), we can directly consider the frequency of haplotype H as:

$$\begin{aligned} f(H)_{t+1} &= \sum_{x,y} q(x,y) \left(\Omega_{xyGHH} \pi_{1Hxy} + \sum_{\eta \neq H} \frac{\Omega_{xyGH\eta} \pi_{2H\eta xy}}{2} \right) = \\ &f(H)_t \left(g_1(f(H)) \Omega_{xyGHH} \pi_{1Hxy} + \sum_{\eta \neq H} g_2(f(\eta)) \frac{\Omega_{xyGH\eta} \pi_{2H\eta xy}}{2} \right) \end{aligned}$$

where η is any allele other than H , and g_1 and g_2 are functions of the frequency in the previous generation of H and η , respectively.

Then:

$$f(H)_{t+1} = f(H)_t \Omega_H$$

where:

$$\Omega_H = g_1(f(H))\Omega_{xyGHH}\pi_{1Hxy} + \sum_{\eta \neq H} g_2(f(\eta)) \frac{\Omega_{xyGH\eta}\pi_{2H\eta xy}}{2}$$

is the marginal overall fitness of haplotype H .

So, the total mean change in the log fitness caused by the change in haplotype frequencies is:

$$J_{\Omega\eta} = \sum_{\eta} (f(\eta)_{t+1} - f(\eta)_t) \log(\Omega_{\eta}) \quad (27)$$

Therefore, from model (17), if we measure the mean change in the log fitness, we can obtain different measures of information associated with the change in genotype frequency or the change in haplotype frequency. These types of information are generally different, since there can be genotypic change without a change in gene frequency. As for the change in genotypes, we calculate the information in two ways, one linked to the variation in each genotype obtained from different instances of mating and another linked to the variation in the genotype frequencies as a whole in the new population.

Below are some simple examples, in this regard, for a biallelic locus.

6.3.4. Sexual Selection Information for a Biallelic Locus

We will assume, like before, that the genotypic frequencies in males and females are equal; in the case of females, we already saw that, $p'_{f1} = p_{f1}m_{AA}$, $p_{f1} = p^2$, $p'_{f2} = p_{f2}m_{Aa}$, $p_{f2} = 2pq$, etc. The difference between the population frequencies, p_f , and the frequencies within the mating, p'_f , is $\Delta p_{f1} = p'_{f1} - p_{f1} = (p_1)m_{AA} - (p_1)$, etc.

Information due to sexual selection in females in terms of genotypic mating fitness can be described as:

$$J_{S1} = \sum \Delta p_{fx} \log \frac{p'_{fx}}{p_{fx}}$$

$$J_{S1} = [(p_1m_{AA} - p_1)\log m_{AA} + (p_2m_{Aa} - p_2)\log m_{Aa} + (p_3m_{aa} - p_3)\log m_{aa}]$$

and since the genotypic frequencies in males and females are equal, $p_{fx} = p_{my}$, and we have assumed that the mutual fitnesses are symmetrical, $m_{xy} = m_{yx}$, then, sexual selection in males is equal to that of females, $J_{S2} = J_{S1}$, and, since $J_{PSS} = J_{S2} + J_{S1}$, then:

$$J_{PSS} = 2[p^2(m_{AA} - 1)\log(m_{AA}) + 2pq(m_{Aa} - 1)\log(m_{Aa}) + q^2(m_{aa} - 1)\log(m_{aa})] \quad (28)$$

where we can see that Jeffreys' divergence for sexual selection captures the information generated by the change in genotype frequencies, weighted by the logarithm of the genotypic mating fitness.

We can now compare, for this same scenario, what the genotypic information (Equations (25) and (26) and the haplotypic information (Equation (27)) would be.

Equation (25) computes the information generated by the change in the distribution of genotypes as a sum of the information:

$$\Delta G_{\Omega} = \sum_G J_G$$

where:

$$J_G = \sum_{x,y} (f(G)_{(t+1)xy} - f(G)_{txy}) \log(\Omega_{xyG})$$

In our case, this implies that $m_{xy} = m_{yx}$ and that $\Omega_{xyG} = m_{xyG}$, then:

$$\begin{aligned} J_{AA} = & (p^{4m_{11}} - p^4) \log(m_{11}) + 2(p^3 q m_{12} - p^3 q) \log(m_{12}) \\ & + (p^2 q^{2m_{22}} - p^2 q^2) \log(m_{22}) \\ & p^2(p^{2m_{11}} - p^2) \log(m_{11}) + p^2(2pqm_{12} - 2pq) \log(m_{12}) \\ & + p^2(q^{2m_{22}} - q^2) \log(m_{22}) \\ & p^2[p^2(m_{11} - 1) \log(m_{11}) + 2pq(m_{12} - 1) \log(m_{12}) \\ & + q^2(m_{22} - 1) \log(m_{22})] \end{aligned}$$

Similarly:

$$\begin{aligned} J_{aa} = & q^2[q^2(m_{33} - 1) \log(m_{33}) + 2pq(m_{32} - 1) \log(m_{32}) \\ & + p^2(m_{22} - 1) \log(m_{22})] \end{aligned}$$

and, finally:

$$\begin{aligned} J_{Aa} = & 2pq[p^2(m_{12} - 1) \log(m_{12}) + pq(m_{22} - 1) \log(m_{22}) \\ & + q^2(m_{23} - 1) \log(m_{23}) + pq(m_{13} - 1) \log(m_{13})] \end{aligned}$$

Now recall that:

$$\begin{aligned} m_{AA} &= p^{2m_{11}} + 2pqm_{12} + q^{2m_{13}} \\ m_{Aa} &= p^{2m_{21}} + 2pqm_{22} + q^{2m_{23}} \\ m_{aa} &= p^2m_{31} + 2pqm_{32} + q^2m_{33} \end{aligned}$$

We can see that the difference between J_{PSS} and J_G is where we focus the unit of information. That is, for the sexual selection information, J_{PSS} , the focus is on each genotype that is involved in various instances of mating. So, for m_{AA} , the focus is on the combination of mating that involves one or two AA genotypes as parents. However, for J_G information, the focus is within a specific genotype produced by the mating.

As for the other measure of information that we have defined for the genotypes (Equation (26)):

$$\begin{aligned} J_{\Omega\eta} &= \sum_{\eta} (f(\eta)_{t+1} - f(\eta)_t) \log(\Omega_{\eta}) \\ f(G)_{t+1} &= \sum_{x,y} q(x,y) \pi_{Gxy} \Omega_{xyG} = f(G)_t \times \sum_{x,y} f_{xy}(\gamma) \pi_{Gxy} \Omega_{xyG} = f(G)_t \times \Omega_G \end{aligned}$$

which in regard to our sexual selection scenario implies:

$$\begin{aligned} f(AA)_{t+1} &= p^{2(p^{2m_{11}} + 2pqm_{12} + q^{2m_{13}})} \\ f(AA)_t &= p^{2(p^2 + 2pq + q^2)} = p^2 \\ \Omega_{AA} &= p^2m_{11} + 2pqm_{12} + q^2m_{22} \\ f(Aa)_{t+1} &= 2pq[p^{2m_{12}} + pq(m_{22} + m_{13}) + q^{2m_{23}}] \\ f(Aa)_t &= 2pq \\ \Omega_{Aa} &= p^{2m_{12}} + pq(m_{22} + m_{13}) + q^{2m_{23}} \\ f(aa)_{t+1} &= q^{2[q^{2m_{33}} + 2pqm_{32} + p^{2m_{22}}]} \\ f(aa)_t &= q^2 \\ \Omega_{aa} &= q^2m_{33} + 2pqm_{32} + p^2m_{22} \end{aligned}$$

so that:

$$\begin{aligned} J_{\Omega\eta} &= p^2(\Omega_{AA} - 1) \log(\Omega_{AA}) + 2pq(\Omega_{Aa} - 1) \log(\Omega_{Aa}) \\ &+ q^2(\Omega_{aa} - 1) \log(\Omega_{aa}) \end{aligned}$$

We can see that the difference between J_{PSS} in Equation (28) and $J_{\Omega\eta}$ is due to the different fitness weights of each genotype and, since $m_{xy} = m_{yx}$, when comparing these aspects, we can see that:

$$\begin{aligned}\Omega_{AA} - m_{AA} &= q^2(m_{22} - m_{13}) \\ \Omega_{Aa} - m_{Aa} &= pq(m_{13} - m_{22}) \\ \Omega_{aa} - m_{aa} &= p^2(m_{22} - m_{13})\end{aligned}$$

The difference lies in the crosses, $Aa \times Aa$ (m_{22}) and $AA \times aa$ (m_{13} or m_{31}). This is because the J_{PSS} calculation of sexual selection focuses on the variation generated by the mating on the marginal frequency of each genotype. For example, the marginal frequency of AA comes from the mating of AA with other genotypes, which includes the cross $AA \times aa$ (mating fitness m_{13}), but not $Aa \times Aa$ (mating fitness m_{22}). The same applies to the marginal frequency of the aa genotype; it includes the cross $aa \times AA$, but not the cross between heterozygotes.

On the other hand, the calculation of the genotypic information, $J_{\Omega\eta}$, directly studies the variation in the frequency of each genotype resulting from a set of mating. In this case, the pattern is reversed, namely for the genotypes AA and aa , the cross between heterozygotes is included, but not the $AA \times aa$. The opposite holds true for the Aa genotype.

In this specific case involving a biallelic locus with no differences in terms of viability and fertility, but with mating fitness under the constraint $m_{xy} = m_{yx}$, if the mating fitness among heterozygotes coincides with the mating fitness among different homozygotes, the information obtained from the genotypes aligns with the information from sexual selection. If this condition is not met, the information is similar, but differs by a factor associated with the difference between $m_{22} - m_{13}$.

Finally, for the haplotype information measure (Equation (27)), the information generated by the change in the distribution of haplotypes is:

$$J_{\Omega\eta} = \sum_{\eta} (f(\eta)_{t+1} - f(\eta)_t) \log(\Omega_{\eta})$$

where:

$$\begin{aligned}f(\eta)_{t+1} &= f(\eta)_t \Omega_{\eta} \\ \Omega_{\eta} &= g_1(f(\eta)) \Omega_{xyG\eta\eta} \pi_{1\eta xy} + \sum_{h \neq \eta} g_2(f(h)) \frac{\Omega_{xyG\eta h} \pi_{2\eta xy}}{2}\end{aligned}$$

which in our case implies:

$$\begin{aligned}f(A)_{t+1} &= p[(p^3m_{11} + 2p^2qm_{12} + pq^2m_{22}) \\ &\quad + (p^2qm_{12} + pq^2(m_{22} + m_{13}) + q^3m_{23})]\end{aligned}$$

so:

$$\Omega_A = p^3m_{11} + 3p^2qm_{12} + pq^2(2m_{22} + m_{13}) + q^3m_{23}$$

and:

$$\begin{aligned}f(a)_{t+1} &= q[(q^3m_{33} + 2pq^2m_{32} + pq^2m_{22}) \\ &\quad + (p^2qm_{12} + pq^2(m_{22} + m_{13}) + q^3m_{23})]\end{aligned}$$

so:

$$\Omega_a = q^3m_{33} + 3pq^2m_{32} + p^2q(2m_{22} + m_{13}) + p^3m_{12}$$

Information about the change in allele frequency caused by selection is defined as:

$$J_{\Omega\eta} = p(\Omega_A - 1) \log(\Omega_A) + q(\Omega_a - 1) \log(\Omega_a)$$

Note that if the mating fitnesses, m_{ij} , are equal to one, then $\Omega_A = \Omega_a = 1$ and, therefore, the information is zero.

6.3.5. Toy Example: Information for Hybrid Sterility

Let us compare the different information indices we have defined using a simple hybrid sterility model. We continue with the one-locus model without mutations and assume that heterozygotes have lower fertility, so that their frequency in terms of successful mating (with fertilization) is lower. That is, $m'_{12} = m'_{21} = m'_{22} = m'_{23} = m'_{32} = \varepsilon < 1$. The rest of the mating fitnesses equal one, $m'_{11} = m'_{13} = m'_{31} = m'_{33} = 1$. The average mating fitness is:

$$M = (p^2 + q^2)m'_\varepsilon + 2pq\varepsilon$$

where:

$$m'_\varepsilon = p^2 + 2pq\varepsilon + q^2$$

Since the frequencies of males and females are equal and $m'_{ij} = m'_{ji}$, then the marginal fitnesses of females and males are equal:

$$\begin{aligned} m'_{AA} &= m'_{aa} = m'_\varepsilon \\ m'_{Aa} &= \varepsilon \end{aligned}$$

The normalized frequencies will be:

$$\begin{aligned} m_{11} &= m_{13} = m_{31} = m_{33} = \frac{1}{M} = m \\ m_{12} &= m_{21} = m_{22} = m_{23} = m_{32} = \frac{\varepsilon}{M} = E \\ m_{AA} &= \frac{m'_{AA}}{M} = m_{aa} = m'_{aa} = m_\varepsilon = \frac{m'_\varepsilon}{M} \\ m_{Aa} &= E \end{aligned}$$

Sexual Selection Information

Substituting (28) we obtain:

$$J_{PSS} = 2[(p^2 + q^2)(m_\varepsilon - 1)\log(m_\varepsilon) + 2pq(E - 1)\log(E)]$$

J_{PSS} represents the information obtained from the change in the phenotypic frequencies of each sex (in this case, equal to the genotypic frequencies) within the mating, with respect to the population frequencies.

Note that in the particular case of equal allelic frequency, $p = q$, we get:

$$\begin{aligned} m'_\varepsilon &= \frac{1+\varepsilon}{2} \\ M &= \frac{1+3\varepsilon}{4} \\ m_\varepsilon &= \frac{2(1+\varepsilon)}{1+3\varepsilon} \\ E &= \frac{4\varepsilon}{1+3\varepsilon} \end{aligned}$$

and the value of J_{PSS} is not 0:

$$J_{PSS} = \frac{(1-\varepsilon)}{1+3\varepsilon} \log\left(\frac{(1+\varepsilon)}{2\varepsilon}\right)$$

We can see that if, for example, we take $\varepsilon = 1/3$, we obtain $J_{PSS} = (1/3)\log(2) \neq 0$. This is interesting because this model with equal allele frequencies does not affect the distribution of genotypes in the following generation and, therefore, the information obtained from the distributions of genotypes and haplotypes (alleles) will be 0 (see below).

Information on Mating by Genotype

We can calculate the sum of the J_G indices, which for each genotype compute the frequency of the genotype after the mating with respect to the expected frequency, if the

instances of mating are random. This mating information on each genotype corresponds to Equation (25) and, in our example, is as follows:

$$\begin{aligned} J_{AA} &= p^2[p^2(m-1)\log(m) + (1-p^2)(E-1)\log(E)] \\ J_{Aa} &= 2pq[pq(m-1)\log(m) + (1-pq)(E-1)\log(E)] \\ J_{aa} &= q^2[q^2(m-1)\log(m) + (1-q^2)(E-1)\log(E)] \end{aligned}$$

The total change for the three genotypes will be the result of the following sum:

$$\Delta G_{\Omega} = J_{AA} + J_{Aa} + J_{aa}$$

Note that in the case $p = q$, $J_{AA} = J_{aa} = J$, $J_{Aa} = 2J$. Then, $\Delta G_{\Omega} = 4J$.

$$\begin{aligned} J &= \frac{-3}{16} \frac{(1-\varepsilon)}{(1+3\varepsilon)} \log(\varepsilon) \\ \Delta G_{\Omega} &= \frac{-3}{4} \frac{(1-\varepsilon)}{(1+3\varepsilon)} \log(\varepsilon) \end{aligned}$$

taking $\varepsilon = 1/3$, $\Delta G_{\Omega} = (1/4)\log(3) \neq 0$.

We see that the mating information by genotype is not 0 either, even though the allele frequencies are equal. This is because the distribution of genotypes according to mating changes, even though the total distribution remains constant.

Genotype Distribution Information

We can use Equation (26) to calculate the information obtained when the distribution of genotypes changes. In this case, the fitness associated with each genotype is:

$$\begin{aligned} \Omega_{AA} &= p^2m + (1-p^2)E \\ \Omega_{Aa} &= pqm + (1-pq)E \\ \Omega_{aa} &= q^2m + (1-q^2)E \end{aligned}$$

And the information is:

$$\begin{aligned} J_{G\Omega} &= p^2(\Omega_{AA} - 1)\log(\Omega_{AA}) + 2pq(\Omega_{Aa} - 1)\log(\Omega_{Aa}) \\ &\quad + q^2(\Omega_{aa} - 1)\log(\Omega_{aa}) \end{aligned}$$

Note that if $p = q = 0.5$, the three values of Ω_G are equal and $m = 4/(1+3\varepsilon)$ and $E = 4\varepsilon/(1+3\varepsilon)$, so that

$$\Omega_G = \left(\frac{1}{4}\right)m + \left(\frac{3}{4}\right)E = \frac{1}{(1+3\varepsilon)} + \frac{3\varepsilon}{1+3\varepsilon} = 1$$

and the information is $J_{\Omega G} = 0$.

This is expected, since selection acts on heterozygotes and the gain in frequency in terms of homozygotes is distributed equally (in the case of $p = q$), generating heterozygotes again and, therefore, the genotypic distribution is not altered when the frequencies of both alleles are equal.

Haplotype Distribution Information

We calculate this information using Equation (27) and, as such, we obtain:

$$\begin{aligned} \Omega_A &= p\Omega_{AA} + q\Omega_{Aa} \\ \Omega_a &= q\Omega_{Aa} + q\Omega_{aa} \end{aligned}$$

and the corresponding information is:

$$J_{\eta} = p(\Omega_A - 1)\log(\Omega_A) + q(\Omega_a - 1)\log(\Omega_a)$$

which, just like in the previous case, will be 0 if $p = q$.

7. Discussion

In this work, we have proposed a population genetics model that begins with a mating system guided by mutual mating fitness between individuals, continues with the distribution of genotypes resulting from each mating, and follows with the subsequent survival of these genotypes, based on their viability fitness until they reach adulthood. At each stage of the model, variation in the distribution of the genotypes is captured in terms of information.

Population genetics models are fundamental both for testing proposed hypotheses, whether general or associated with a specific species and evolutionary scenario, and for deepening our understanding of the history of species and biodiversity on Earth. They also help reveal invariances and logical connections that, in turn, may suggest new questions and pose new hypotheses [47,48].

As mentioned in the introduction, information theory seems to be increasingly relevant in terms of providing new descriptions of the evolutionary process, allowing us to explore certain aspects more deeply or facilitate the development of models and estimators for parameters of interest. Recent studies have shown how selection accumulates information in the genome [20] and also how the flow of information throughout the process of evolutionary change can capture different degrees of adaptation and can help develop new statistics [22].

It should be noted that when we consider typical equations in terms of classical population genetics, or the replicator equation as $p' = \Omega p$, where Ω is some kind of relative fitness, and we express the mean change in the logarithm of the fitness of the character, what we obtain is an equation of the type:

$$\Delta \bar{z}_s = \sum_i p_i f(\Omega_i)$$

where f is the generating function $(\Omega - 1)\log(\Omega)$ of an f divergence corresponding to Jeffrey's divergence.

Therefore, every component of the evolutionary process that is guided by some kind of selection, be it sexual selection, mate choice, or viability, etc., such that we can express the expected change in terms of $(\Omega - 1)\log(\Omega)$, can be described in terms of Jeffrey's divergence. That is, the dynamics of evolutionary change itself, at least as it can be expressed in the given form $p' = p\Omega$, can be measured in terms of Jeffrey's divergence, because that is what it is mathematically, in this sense it is not an ad hoc statistic.

In the model presented here, the average change in fitness at different stages of the life cycle can be calculated as Jeffrey's information. This has allowed us to calculate the information associated with various stages of the process, namely the information obtained from the distribution of mating, the information associated with mating for each genotype, as well as the information associated with the distribution of genotypes and haplotypes in the new generation. We have found that these indices contain distinct information and, therefore, could allow us to draw inferences about associated phenomena.

Population genetics models typically assume random mating [20,47] or, at most, consider any mating system implicitly, as in [21,22], where different mating systems may be reflected in the topology of a graph. In our case, by explicitly modeling the mating system in terms of information, on the one hand, we unify different aspects related to the detection of sexual selection and assortative mating and facilitate the development of statistical measures [25]. And, on the other hand, by incorporating mating fitness into a more general model, we reveal connections between different mating processes and other parts of the gene frequency evolution process, described in terms of information.

An interesting aspect of the proposed model is that the average changes in terms of different types of fitness, due to variations between the corresponding distributions, can be easily translated into statistical tests, since Jeffrey's divergence approximates well

a chi-square distribution, when the hypothesis of equality between the distributions is met [36,49].

All this has allowed the development of a multi-model selection methodology applied to the estimation of non-random mating [24] that can be used in empirical studies, see for example [50–52].

Limitations of the Model

The proposed model connects the mating in the parental generation with the distribution of adult genotypes in the next generation. The model explicitly accounts for the mating system defined by mutual mating fitness, female fecundity, and the viability associated with genomic genotypes. The effects of mutation and recombination are not explicitly modeled, as they are implicitly included in the calculation of genotypic proportions following mating. However, in its current form, the model does not consider migration, genetic drift, epigenetic inheritance, or other potential factors related to niche constructions or habitat choice [53].

However, these aspects are not central to the main message of the work, which emphasizes the connection between different aspects of evolutionary change and the measurement of information through shifts in distributions, a point also highlighted in recent studies [20,22].

8. Conclusions

In conclusion, it may be worth considering the relationship between biological fitness and information, whether the increase in information is favored by selection [54], or even whether we can think of fitness in terms of multilevel information alignment between populations and their environment [55,56]. Additionally, we might reflect on the advantages of shifting from a fitness-driven approach to evolution toward one guided by information and whether this could help clarify, broaden, and deepen the mathematical description of evolutionary theory.

Funding: This study was funded by Xunta de Galicia (ED431C 2024/22), the Ministerio de Ciencia e Innovación (PID2022-137935NB-I00), the Centro singular de investigación de Galicia accreditation 2024-2027 (ED431G 2023/07), and “ERDF A way of making Europe”.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data are contained within the article.

Conflicts of Interest: The authors declare that there are no conflicts of interest.

References

- Shannon, C.E. A Mathematical Theory of Communication. *Bell Syst. Tech. J.* **1948**, *27*, 623–656. [\[CrossRef\]](#)
- Nyquist, H. Certain Factors Affecting Telegraph Speed. *Bell Syst. Tech. J.* **1924**, *3*, 324–346. [\[CrossRef\]](#)
- Hartley, R.V.L. Transmission of Information. *Bell Syst. Tech. J.* **1928**, *7*, 535–563. [\[CrossRef\]](#)
- Wagner, A. Information Theory, Evolutionary Innovations and Evolvability. *Philos. Trans. R. Soc. B Biol. Sci.* **2017**, *372*, 20160416. [\[CrossRef\]](#)
- Sherwin, W.B. Entropy, or Information, Unifies Ecology and Evolution and Beyond. *Entropy* **2018**, *20*, 727. [\[CrossRef\]](#)
- Cartwright, J.H.E.; Čejková, J.; Fimmel, E.; Giannerini, S.; Gonzalez, D.L.; Goracci, G.; Grácio, C.; Houwing-Duistermaat, J.; Matić, D.; Mišić, N.; et al. Information, Coding, and Biological Function: The Dynamics of Life. *Artif. Life* **2024**, *30*, 16–27. [\[CrossRef\]](#)
- Kimura, M. Natural Selection as the Process of Accumulating Genetic Information in Adaptive Evolution. *Genet. Res.* **1961**, *2*, 127–140. [\[CrossRef\]](#)
- Spetner, L. Information Transmission in Evolution. *IEEE Trans. Inf. Theory* **1968**, *14*, 3–6. [\[CrossRef\]](#)
- Smith, J.M. The Idea of Information in Biology. *Q. Rev. Biol.* **1999**, *74*. [\[CrossRef\]](#)
- Szathmáry, E. Biological Information, Kin Selection, and Evolutionary Transitions. *Theor. Popul. Biol.* **2001**, *59*, 11–14. [\[CrossRef\]](#)
- MacKay, D.J. Rate of Information Acquisition by a Species Subjected to Natural Selection. 1999. Available online: www.inference.phy.cam.ac.uk/mackay/gene.pdf (accessed on 21 November 2024).

12. Adami, C.; Ofria, C.; Collier, T.C. Evolution of Biological Complexity. *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 4463–4468. [[CrossRef](#)] [[PubMed](#)]
13. Adami, C. Information Theory in Molecular Biology. *Phys. Life Rev.* **2004**, *1*, 3–22. [[CrossRef](#)]
14. Adami, C. The Use of Information Theory in Evolutionary Biology. *Ann. N. Y. Acad. Sci.* **2012**, *1256*, 49–65. [[CrossRef](#)] [[PubMed](#)]
15. Frieden, B.R.; Plastino, A.; Soffer, B.H. Population Genetics from an Information Perspective. *J. Theor. Biol.* **2001**, *208*, 49–64. [[CrossRef](#)]
16. Strelioff, C.C.; Lenski, R.E.; Ofria, C. Evolutionary Dynamics, Epistatic Interactions, and Biological Information. *J. Theor. Biol.* **2010**, *266*, 584–594. [[CrossRef](#)]
17. Frank, S.A. Natural Selection Maximizes Fisher Information. *J. Evol. Biol.* **2009**, *22*, 231–244. [[CrossRef](#)]
18. Frank, S.A. Natural Selection. V. How to Read the Fundamental Equations of Evolutionary Change in Terms of Information Theory. *J. Evol. Biol.* **2012**, *25*, 2377–2396. [[CrossRef](#)]
19. Frank, S.A. Universal Expressions of Population Change by the Price Equation: Natural Selection, Information, and Maximum Entropy Production. *Ecol. Evol.* **2017**, *7*, 3381–3396. [[CrossRef](#)]
20. Hledík, M.; Barton, N.; Tkačik, G. Accumulation and Maintenance of Information in Evolution. *Proc. Natl. Acad. Sci. USA* **2022**, *119*, e2123152119. [[CrossRef](#)]
21. Smith, E. Beyond Fitness: The Nature of Selection Acting through the Constructive Steps of Lifecycles. *Evolution* **2023**, *77*, 1967–1986. [[CrossRef](#)]
22. Smith, E. Beyond Fitness: The Information Imparted in Population States by Selection throughout Lifecycles. *Theor. Popul. Biol.* **2024**, *157*, 86–117. [[CrossRef](#)] [[PubMed](#)]
23. Carvajal-Rodríguez, A. Non-Random Mating and Information Theory. *Theor. Popul. Biol.* **2018**, *120*, 103–113. [[CrossRef](#)] [[PubMed](#)]
24. Carvajal-Rodríguez, A. Multi-Model Inference of Non-Random Mating from an Information Theoretic Approach. *Theor. Popul. Biol.* **2020**, *131*, 38–53. [[CrossRef](#)] [[PubMed](#)]
25. Carvajal-Rodríguez, A. Unifying Quantification Methods for Sexual Selection and Assortative Mating Using Information Theory. *Theor. Popul. Biol.* **2024**, *158*, 206–215. [[CrossRef](#)] [[PubMed](#)]
26. Price, G.R. Extension of Covariance Selection Mathematics. *Ann. Hum. Genet.* **1972**, *35*, 485–490. [[CrossRef](#)]
27. Price, G.R. The Nature of Selection. *J. Theor. Biol.* **1995**, *175*, 389–396. [[CrossRef](#)]
28. Frank, S.A. George Price's Contributions to Evolutionary Genetics. *J. Theor. Biol.* **1995**, *175*, 373–388. [[CrossRef](#)]
29. Gardner, A. Price's Equation Made Clear. *Philos. Trans. R. Soc. B Biol. Sci.* **2020**, *375*, 20190361. [[CrossRef](#)]
30. Luque, V.J. One Equation to Rule Them All: A Philosophical Analysis of the Price Equation. *Biol. Philos.* **2017**, *32*, 97–125. [[CrossRef](#)]
31. Frank, S. The Price Equation Program: Simple Invariances Unify Population Dynamics, Thermodynamics, Probability, Information and Inference. *Entropy* **2018**, *20*, 978. [[CrossRef](#)]
32. Luque, V.J.; Baravalle, L. The Mirror of Physics: On How the Price Equation Can Unify Evolutionary Biology. *Synthese* **2021**, *199*, 12439–12462. [[CrossRef](#)]
33. Queller, D.C. Fundamental Theorems of Evolution. *Am. Nat.* **2017**, *189*, 345–353. [[CrossRef](#)] [[PubMed](#)]
34. Ewens, W.J. The Fundamental Theorem of Natural Selection: The End of a Story. *Evolution* **2024**, *78*, 803–808. [[CrossRef](#)] [[PubMed](#)]
35. Kuhn, T.S. Second Thoughts on Paradigms. *Struct. Sci. Theor.* **1974**, *2*, 459–482.
36. Cover, T.M.; Thomas, J.A. *Elements of Information Theory*; John Wiley & Sons: Hoboken, NJ, USA, 2012; ISBN 978-1-118-58577-1.
37. Frank, S.A. Natural Selection. VI. Partitioning the Information in Fitness and Characters by Path Analysis. *J. Evol. Biol.* **2013**, *26*, 457–471. [[CrossRef](#)]
38. Frank, S.A. D'Alembert's Direct and Inertial Forces Acting on Populations: The Price Equation and the Fundamental Theorem of Natural Selection. *Entropy* **2015**, *17*, 7087–7100. [[CrossRef](#)]
39. Carvajal-Rodríguez, A. A Generalization of the Informational View of Non-Random Mating: Models with Variable Population Frequencies. *Theor. Popul. Biol.* **2019**, *125*, 67–74. [[CrossRef](#)]
40. Shuker, D.M.; Kvarnemo, C. The Definition of Sexual Selection: A Response to Comments on Shuker and Kvarnemo. *Behav. Ecol.* **2021**, *32*, 801–802. [[CrossRef](#)]
41. Rolán-Alvarez, E.; Caballero, A. Estimating Sexual Selection and Sexual Isolation Effects from Mating Frequencies. *Evolution* **2000**, *54*, 30–36.
42. Wade, M.J.; Arnold, S.J. The Intensity of Sexual Selection in Relation to Male Sexual Behaviour, Female Choice, and Sperm Precedence. *Anim. Behav.* **1980**, *28*, 446–461. [[CrossRef](#)]
43. Carvajal-Rodríguez, A. The Information Theory Formalism Unifies the Detection of the Patterns of Sexual Selection and Assortative Mating for Both Discrete and Quantitative Traits. *bioRxiv* **2023**. [[CrossRef](#)]
44. Bürger, R. Multilocus Population-Genetic Theory. *Theor. Popul. Biol.* **2020**, *133*, 40–48. [[CrossRef](#)] [[PubMed](#)]
45. Bürger, R. *The Mathematical Theory of Selection, Recombination, and Mutation*; Wiley series in mathematical and computational biology; John Wiley: Chichester, UK, 2000; ISBN 0-471-98653-4.
46. Hedrick, P.W. *Genetics of Populations*, 3rd ed.; Jones & Bartlett Publishers: Boston, MA, USA, 2005; ISBN 0-7637-4772-6.
47. Hofrichter, J.; Jost, J.; Tran, T.D. *Information Geometry and Population Genetics: The Mathematical Structure of the Wright-Fisher Model; Understanding Complex Systems*; Springer International Publishing: Cham, Switzerland, 2017; ISBN 978-3-319-52044-5.

48. Otsuka, J. *The Role of Mathematics in Evolutionary Theory*; Elements in the Philosophy of Biology; Cambridge University Press: Cambridge, UK, 2019. [[CrossRef](#)]
49. Kullback, S. *Information Theory and Statistics, New Edition*; Dover Publications: Mineola, NY, USA, 1997; ISBN 978-0-486-69684-3.
50. Gefaell, J.; Vigo, R.; Galindo, J.; Rolán-Alvarez, E. Experimental Evidence of Mate Choice as the Driving Mechanism behind Negative Assortative Mating for Shell Colour in a Marine Snail. *Biol. J. Linn. Soc.* **2024**, *142*, 441–451. [[CrossRef](#)]
51. Roggero, A.; Alù, D.; Laini, A.; Rolando, A.; Palestini, C. Color Polymorphism and Mating Trends in a Population of the Alpine Leaf Beetle *Oreina Gloriosa*. *PLoS ONE* **2024**, *19*, e0298330. [[CrossRef](#)]
52. Lau, S.L.Y.; Williams, G.A.; Carvajal-Rodríguez, A.; Rolán-Alvarez, E. An Integrated Approach to Infer the Mechanisms of Mate Choice for Size. *Anim. Behav.* **2021**, *175*, 33–43. [[CrossRef](#)]
53. Edelaar, P.; Otsuka, J.; Luque, V.J. A Generalised Approach to the Study and Understanding of Adaptive Evolution. *Biol. Rev.* **2023**, *98*, 352–375. [[CrossRef](#)]
54. Mirmomeni, M.; Punch, W.F.; Adami, C. Is Information a Selectable Trait? *arXiv* **2014**, arXiv:1408.3651.
55. Hilbert, M. *Fitness as Informational Fit: The Communication Channel Between the Evolving Population and Its Environment*; SSRN: Atlanta, GA, USA, 2015.
56. Hilbert, M. *Fitness as Informational Fit: An Information Theoretic Approach to Multilevel Requisite Variety*; SSRN: Atlanta, GA, USA, 2015.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.