doi:10.1111/evo.12691

The scale-of-choice effect and how estimates of assortative mating in the wild can be biased due to heterogeneous samples

Emilio Rolán-Alvarez,[1,2,3] Antonio Carvajal-Rodríguez,[1] Alicia de Coo,[1] Beatriz Cortés,[1] Daniel Estévez,[1] Mar Ferreira,[1] Rubén González,[1] and Adriana D. Briscoe[2]

[1]*Departamento de Bioquímica, Xenética e Inmunoloxía, Facultade de Bioloxía, Universidad de Vigo, 36310 Vigo, Spain*

[2]*Department of Ecology and Evolutionary Biology, University of California, Irvine, California 92697*

[3]*E-mail: rolan@uvigo.es*

Received June 7, 2014
Accepted May 13, 2015

The mode in which sexual organisms choose mates is a key evolutionary process, as it can have a profound impact on fitness and speciation. One way to study mate choice in the wild is by measuring trait correlation between mates. Positive assortative mating is inferred when individuals of a mating pair display traits that are more similar than those expected under random mating while negative assortative mating is the opposite. A recent review of 1134 trait correlations found that positive estimates of assortative mating were more frequent and larger in magnitude than negative estimates. Here, we describe the scale-of-choice effect (SCE), which occurs when mate choice exists at a smaller scale than that of the investigator's sampling, while simultaneously the trait is heterogeneously distributed at the true scale-of-choice. We demonstrate the SCE by Monte Carlo simulations and estimate it in two organisms showing positive (*Littorina saxatilis)* and negative (*L. fabalis*) assortative mating. Our results show that both positive and negative estimates are biased by the SCE by different magnitudes, typically toward positive values. Therefore, the low frequency of negative assortative mating observed in the literature may be due to the SCE's impact on correlation estimates, which demands new experimental evaluation.

KEY WORDS: **Correlation bias, mate choice, mating preference, mating pairs, negative assortative mating, positive assortative mating.**

The manner in which a mate is chosen among available members of a population remains a key aspect of reproductive success in most sexual species. This is because mating preferences (mate choice) can affect the probability of transmitting alleles within a species (by mate selection), discrimination of mates among species or both simultaneously (Lewontin et al. 1968; Coyne and Orr 2004; Gavrilets 2004). Therefore, it has been of particular interest to detect and understand the behavioral mechanisms causing deviations from random mating in relation to traits used by the organism for mate choice (Andersson 1994; Gray and McKinnon 2007; Jiang et al. 2013). Deviations from random mating can be measured as a correlation (either positive or negative) between values of a trait across the distribution of observed matings within a particular species (Jiang et al. 2013). Positive assortative mating implies a tendency to mate preferentially among individuals with similar trait values, while negative assortative (disassortative) mating entails the opposite (Merrell 1950). Assortative mating in a context of incipient (ongoing) speciation is called sexual isolation (Coyne and Orr 2004). These different mating patterns (random, positive, or negative) can either be the result of direct selection on mating preferences or just a side effect of temporal, mechanical, or physiological constraints (Gavrilets 2004; Jiang et al. 2013). In cases of direct selection, it has been suggested that positive assortative mating could evolve as a solution to the problem of choosing appropriate intraspecific mates (Kirkpatrick 2000) or in order to avoid interspecific mates (Coyne and Orr 2004), while

negative assortative mating could evolve as a way to avoid inbreeding (Pusey and Wolf 1996). Negative assortative mating can also contribute to the maintenance of polymorphisms by producing patterns of negative frequency-dependent sexual selection in relation to the preferred traits (Takahashi and Hori 2008; Field and Barrett 2012; Holman et al. 2013). Unappreciated complex spatial distributions of the trait among individuals in the species could also potentially bias trait correlations between mating pairs (Langston et al. 1997; see below).

Assortative mating was originally studied in the laboratory by mate choice experiments (reviewed in Knoppien 1985). However, it was quickly noticed that such an approach has limitations, as estimates of assortative mating in the laboratory may not necessarily match estimates obtained in the wild (reviewed in Coyne and Orr 2004; see e.g., in Coyne et al. 2005; Llopart et al. 2005). Assortative mating studies undertaken by direct field observations may thus hold an advantage over laboratory mate choice experiments, because they occur under evolutionarily relevant conditions.

Nonetheless, estimates of assortative mating in the wild can be difficult. First, mating pairs are only found in high enough numbers for sufficient experimental sampling in a few species (Jiang et al. 2013). Second, our knowledge of the scale at which mate decisions are made is generally poor, and may potentially be the source of estimation bias. For example, a typical study of assortative mating in the wild may capture dozens of mating pairs along an extensive geographical area (which could be heterogeneous in trait distribution among mates at a smaller scale), or even pool samples from different localities or time periods to gain statistical power (see example in Cruz et al. 2004). These sampling strategies could potentially cause different spatial or temporal artifacts (Langston et al. 1997). It is a well-known statistical property, for example, that the pooling of heterogeneous data prior to the estimation of correlation coefficients may result in nonsensical and biased correlations (Hassler and Thadewald 2003; Almeida et al. 2013; Cocho et al. 2014). In short, the fundamental issue is that in the wild, it is difficult to know: (1) the exact preference of an individual, and (2) the sampling done by the individual to arrive at a mate-choice decision. Instead what we see is only the final result of the organism's choice and we try to make our inference from that. However, as we will see below such a strategy may potentially cause bias during estimation that leads to incorrect conclusions.

The importance of the scale at which mate choice decisions take place has been previously emphasized (Langston et al. 1997; Gwynne et al. 1998; Cruz et al. 2004; Duraes et al. 2009; Dabire et al. 2013), as it has for the scale of animal resource acquisition in general (Ritchie and Olff 1999; Miller et al. 2009; Ambrosini and Saino 2010; Taborsky et al. 2014). Nonetheless, the main conclusion of Langston et al. (1997) still holds today: "problems of scale probably affect studies of mate choice in avian breeding systems in ways that, to the best of our knowledge, have not previously been investigated by behavioral ecologists." In other words, the biggest obstacle to correctly estimating trait correlations from mating pairs captured in the wild is from unknowingly combining a set of samples that are heterogeneous with respect to the studied trait (see also Hassler and Thadewald 2003). We call this potential source of bias the scale-of-choice effect (SCE), which we define here as the magnitude of assortative mating estimated in the sample minus the assortative mating estimated through the true scale-of-choice. The SCE depends on the relative trait heterogeneity found at the true scale-of-choice. A graphical representation of the SCE for both negative (Fig. 1) and positive (Fig. 2) assortative mating permits visualization of what happens: heterogeneous samples pooled together will typically lead to inferences of positive assortative mating, even when negative assortative mating is occurring.

This SCE could challenge the core conclusion of a recent metaanalysis of assortative mating (see Jiang et al. 2013). Therein a trend toward positive correlations was observed using 1116 trait correlation coefficients from 254 species and was interpreted as evidence that negative assortative mating may be rare in nature. However, this interpretation could be unreliable if a factor (such as the SCE) systematically biased estimates of positive and negative assortative mating in the same direction. In fact, estimates of assortative mating in the wild are typically undertaken assuming that the scale of mate choice (mating preference) is similar to the scale of sampling chosen by investigators (Jiang et al. 2013), which could be incorrect, especially for low mobility organisms.

In the present work, we first describe the consequences of the SCE on assortative mating by means of Monte Carlo simulations. Second, we estimate the SCE in two marine snails that display dissimilar patterns of mate choice: positive assortative mating between ecotypes of *Littorina saxatilis* and negative assortative mating in relation to the shell colors of *L. fabalis*. We obtain the estimates in *L. saxatilis* from previously published statistics and by carrying out a reanalysis of published data. In *L. fabalis*, we provide a reanalysis of previously published data (1990 and 2011 samples) and an analysis of new shell color data (2012 and 2013 samples) (including reflectance spectra in 2013) of mating pairs captured in the wild. To the best of our knowledge, this is the first time that the SCE has been experimentally estimated. Our results provide clear evidence that it biases both positive and negative assortative mating estimates toward positive values. We review the dataset of Jiang et al. (2013) in order to test a basic prediction of this mechanism at a spatial level. Specifically, inferential bias obscuring the detection of negative cases should preferentially affect low mobility organisms.

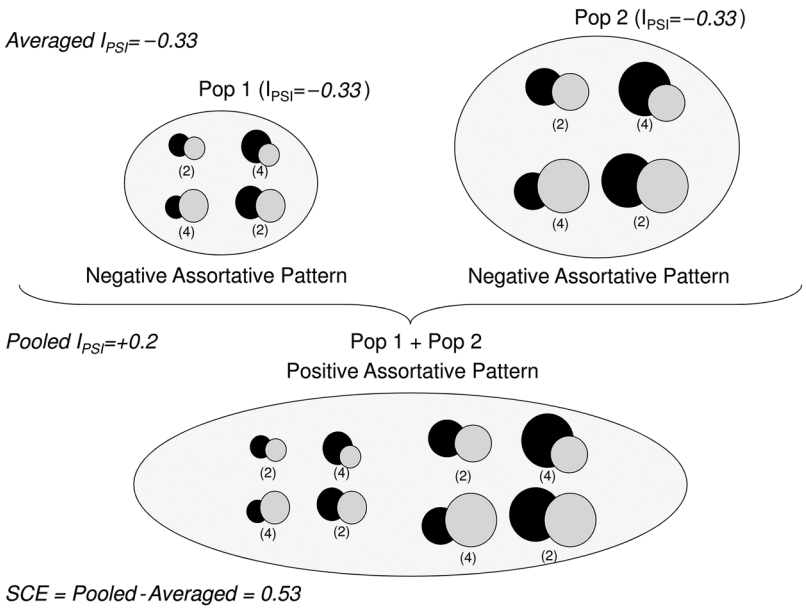

Figure 1. **Description of the SCE for negative assortative mating. Each pair of circles represents a mating pair (black, male; gray, female) type. Different sizes represent different trait values. The upper scheme represents two different populations (Pop 1 and Pop 2), each with a pattern of negative assortative mating. That is, within populations, matings involve more frequent pairings of different sizes (4) than those of similar sizes (2; $I_{PSI} = -0.33$). The lower scheme shows what happens after pooling these two heterogeneous populations (Pop 1 + Pop 2). Although each original subgroup exhibits negative assortative mating, the large differences between them after pooling produces an estimate of assortative mating that is positive ($I_{PSI} = +0.2$). The difference between *Pooled* I_{PSI} minus the *Averaged* I_{PSI} is named the scale-of-choice effect (SCE = 0.2 − (−0.33) = 0.53). All statistics have been calculated by introducing the corresponding population mating (2/4) frequencies in JMATING using 2 × 2 tables (Pop 1 or Pop 2) or 4 × 4 tables (Pop 1 + Pop 2) (Carvajal-Rodríguez and Rolán-Alvarez 2006).**

Material and methods

ASSORTATIVE MATING CONCEPT AND ESTIMATION

Random mating occurs when each male has an equal chance of mating with any other female in the population. This implies that the overall frequency of mating pairs is equal to the product of frequencies of these types in the population (Gavrilets 2004; p. 279). Deviations from random mating can result in different evolutionary phenomena: like mate selection (a component of sexual selection) and/or assortative mating (Gavrilets 2004). Assortative mating (either positive or negative) in itself does not necessarily affect population gene frequencies and so it has been estimated using mating pairs exclusively (ignoring all other information about unmated specimens; reviewed in Jiang et al. 2013; but see our proposal below). Under certain circumstances, however, sexual isolation (an estimate of assortative mating for qualitative traits) can be biased by mating propensity and population trait frequencies (Gilbert and Starmer 1985; Casares et al. 1998; Rolán-Alvarez and Caballero 2000).

Different statistics have been proposed to assess assortative mating. For quantitative traits, Pearson's correlation (r) coefficient (in a range from –1 to 1; with –1 corresponding to a maximum of negative assortative mating, 0 to random mating, and 1 to a maximum of positive assortative mating) is used preferentially (reviewed in Jiang et al. 2013). Whereas for qualitative traits, different statistics (I, Yule's Q, Yule's V, YA, I_{PSI}; all in a range from –1 to 1) have been introduced mainly in a context of sexual isolation between incipient species (Gilbert and Starmer 1985; Rolán-Alvarez and Caballero 2000). All of these statistics (including r) measure the same phenomenon: a tendency to observe an excess of similar (positive) or dissimilar (negative) mating types compared to the random mating pattern. Nonetheless, estimator I_{PSI} has so far been the most suited for the analysis of qualitative (polymorphic) traits, in part due to its usefulness for evaluating assortative mating unbiased by population trait frequencies and mating propensity (Rolán-Alvarez and Caballero 2000). In addition, it has robust estimation properties with respect to both laboratory and wild datasets (Pérez-Figueroa et al. 2005). The I_{PSI} index has been successfully applied to different model organisms (e.g., see Nosil et al. 2002; Cruz et al. 2004; Coyne et al. 2005; Matsubayashi et al. 2013). Having been described in detail elsewhere (Rolán-Alvarez and Caballero 2000), it will only be briefly summarized here:

$$I_{PSI} = \frac{(PSI_{aa} - PSI_{ab} - PSI_{ba} + PSI_{bb})}{(PSI_{aa} + PSI_{ab} + PSI_{ba} + PSI_{bb})}, \qquad (1)$$

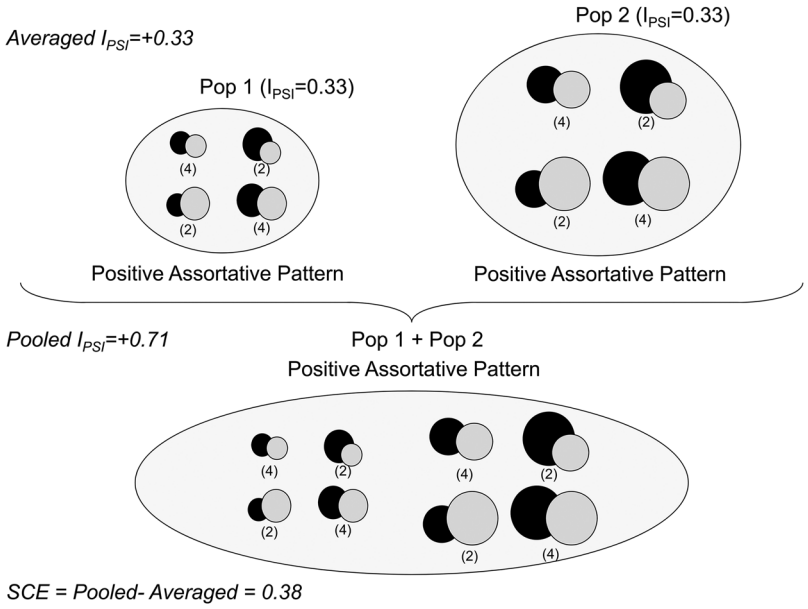

Figure 2. Description of the SCE for positive assortative mating. The upper scheme represents two different populations (Pop 1 and Pop 2), each with a pattern of positive assortative mating. That is, within populations, matings involve less frequent pairings of different sizes (2) than those of similar sizes (4; $I_{PSI} = 0.33$). The lower scheme shows what happens after pooling these two heterogeneous populations (Pop 1 + Pop 2). Both original subgroups and pooled samples exhibit positive assortative mating. The difference between *Pooled* I_{PSI} minus the *Averaged* I_{PSI} is the scale-of-choice effect (SCE = 0.71 – 0.33 = 0.38). Statistics calculated as in Figure 1.

where PSI_{aa}, PSI_{ab}, $PSI_{ba,}$ and PSI_{bb} are the deviations (observed/expected) with respect to random mating (from mated specimens) taking into account the four possible classes of mating types (male–female combinations) tabulated in a mating contingency table of a qualitative trait that comprises two values (a/b, RB/SU, or brown/yellow). In this study, we have estimated two cases of assortative mating based on qualitative traits: two ecotypes in *L. saxatilis* and two classes of shell color in *L. fabalis*. We have provided for convenience both qualitative and quantitative statistics on the traits to show that they ultimately estimate the same phenomenon.

SCALE-OF-CHOICE EFFECT DESCRIPTION

Human investigators concerned with mate choice often have to arbitrarily select a scale at which to conduct their sampling. This scale may or may not match in reality the scale at which the study organisms make their mate choices. This leads to the so called SCE, herein described as the difference between the assortative mating that occurs at the sample scale chosen by a researcher and the actual assortative mating estimated at the true scale of choice. Due to the SCE, estimates of assortative mating will be biased if the scale of choice is smaller than the sampling scale (necessary condition), also if at such a (choice) scale some spatial (or temporal) heterogeneity in the distribution of the trait among individuals in the population being studied is present (sufficient condition; see representative examples in Figs. 1 and 2). Therefore, an alternative to minimize the SCE is to estimate assortative mating in pairs that, before the analysis, were grouped by trait similarity of the mated and unmated specimens in the vicinity. Practically speaking, this avoids the sufficient condition of the SCE and solves the problem of working at the true scale of choice, which is unknown in most species.

A mathematical description of the potential problems observed while estimating correlation coefficients after pooling heterogeneous samples has been given elsewhere (Hassler and Thadewald 2003). These authors concluded that nonsensical correlations can be obtained irrespective of the magnitude and direction of the true correlations, due to the fact that the bias is dependent on the existing variance between and within the pooled samples. They also confirmed by means of Monte Carlo resampling that the same process occurred at realistic sample sizes ($N = 50$ and 100; Hassler and Thadewald 2003). However, the authors did not describe a systematic bias for both positive and negative correlations. Therefore, we performed computer simulations to study the SCE for mating pairs under different biological scenarios. The programming language was C++ and the design of the simulations comprised the following algorithm:

1. Sample N population means $\mu_1, \mu_2, \ldots, \mu_N$ from a normal $N(1, CV)$.
2. For each of the N populations generate a number of K mating pairs in which each mating consists of a pair of correlated

random normal deviates $N(\mu_i, \sigma_w)$ where the correlation coefficient is the choice C (Press 2007).

3. For the set of mating pairs compute both Pearson's correlation coefficient for each population and the pooled Pearson's correlation coefficient obtained by adding up all of the populations. The SCE is estimated as the r coefficient measured in the pooled set of pairs across subpopulations (r_{pool}) minus assortative mating averaged across subpopulations ($r_{average}$; $SCE = r_{pool} - r_{average}$).
4. Replicate this process 100 times.

We assayed different values for the number of populations $N = \{10, 100\}$, number of mating pairs $K = \{20, 500\}$, coefficient of variation CV (0–1 with increments of 0.1), intrapopulation SD $\sigma_w = \{0.1, 0.3, 0.45\}$ and choice C (–1 to 1 with increments of 0.1), giving a total of 2772 combinations (see Table 1). Simulation results and code have been deposited in the DRYAD Digital Repository (doi:10.5061/dryad.r1h4h; URL: http://dx.doi.org/10.5061/dryad.r1h4h).

SAMPLING DESIGN OF POSITIVE ASSORTATIVE MATING

Details of the *L. saxatilis* study populations are provided in the appendix. The dataset of Rolán-Alvarez et al. (1995a) consisted of 108 mating pairs collected in two areas of the same locality from NW Spain (Table 2). The dataset of Rolán-Alvarez et al. (1999) consisted of 216 mating pairs between ecotypes of *L. saxatilis* from 12 populations (SI194, PED94, CEN94, SI394, AGO94, CET94, BAR94, POR94, CAN94, ESC94, SEN94, SI494; see raw data in Table S1) sampled in June–August of 1994. These populations were separated by 2–5 km of rocky shore, except for the SI samples (SI194, SI394, and SI494), which were only separated by a few dozen meters (Rolán-Alvarez et al. 1999).

SAMPLING DESIGN FOR NEGATIVE ASSORTATIVE MATING

Details of the *L. fabalis* study populations are provided in the appendix. Mating pairs accompanied by unmated specimens were collected on *Fucus vesiculosus* during 4 years in the semiexposed rocky shore of Abelleira located in the Ria of Muros-Noya (NW Spain; 42° 48' 0.30"N and 9° 1' 14.87"W). Two samples (1990 and 2011) were reanalyzed from previous studies (Rolán-Alvarez and Ekendahl 1996 and Rolán-Alvarez et al. 2012), while the remaining two samples are described here for the first time (2012 and 2013; see raw data displayed in Table S1). Overall, data on a total of 918 mated and 1506 unmated specimens was collected. Sampling was always carried out at the same shore level and area (about 150 m^2) and during the same period (1–15th July) throughout all years, a season in which most mating takes place (Rolán-Alvarez and Ekendahl 1996). Specimens of *L. fabalis* can be anatomically distinguished from their sibling species *L. obtusata* (see Reid 1996). Moreover, interspecific mating was confirmed by diagnostic molecular markers in one sample (1990), showing that only 1 out of 250 mating pairs involved crosses between different species (Rolán-Alvarez et al. 1995c). Mating pairs were considered valid when the male had inserted his penis into the female reproductive tract or had just removed it. Mating occurs during an interval of 1–2 h or more in this gastropod species, with males showing mate-guarding behavior (Rolán-Alvarez et al. 1995c). For each mating pair we captured only the four closest unmated specimens within one circular 25 cm diameter microarea. The mated and unmated specimens from every microarea allow estimation of population frequencies at the smallest sampled scale, and are useful for grouping mating pairs *a posteriori* (based on microarea frequency) in analyses to avoid SCE bias in estimates (see below). Further details on methods have been given elsewhere (Rolán-Alvarez and Ekendahl 1996; Rolán-Alvarez et al. 2012).

L. *fabalis* SHELL COLOR MEASUREMENTS

Sex and species determination were based on the reproductive organs after having removed the soft part from the shell. Color detection was exclusively based on shell appearance, having been determined in at least one sample independently by both visual inspection and reflectance spectrometry. All of the specimen's variables were scored under a blind design before undertaking the statistical analysis.

We determined shell color by visual inspection using a printed color model (see Fig. S1). Shell colors were scored according to the eight phenotypic classes present in this locality (Rolán-Alvarez and Ekendahl 1996): brown (Brw), olive (Oli), yellow (Yel), orange (Ora), and corresponding shell band colors (showing two parallel and longitudinal bands over the former colors): Brw2, Oli2, Yel2, and Ora2. Since the banded pattern is produced by a single gene in the sibling species *L. obtusata* (Kozminskii 2011) and the trait does not show any significant effect on mate choice (Rolán-Alvarez et al. 2012), we proceeded to pool similar banded and unbanded colors. We then focused our analysis on the two most common colors: brown (Brw+Brw2) and yellow (Yel+Yel2), which represent more than 90% of the sampled population snails (Rolán-Alvarez et al. 2012). Raw data are shown in Table S1. Use of different color classes (8, 4, or 2) during the analysis, however, did not alter qualitatively the observed pattern with respect to assortative mating or sexual selection estimates (Rolán-Alvarez et al. 2012; verified for the new data but not shown).

QUALITATIVE TRAIT VALIDATION

The qualitative traits assessed in both *L. saxatilis* (RB and SU ecotypes) and *L. fabalis* (brown and yellow shells) could in principle be influenced by the researcher's ability to discriminate

Table 1. Results of pooling heterogeneous samples in a simulation of assortative mating.

Parameters simulated			Results from simulation			
N	K	SD_{within}	SD of averaged $r \pm$ SD	Pooled $r \pm$ SD	Averaged $r \pm$ SD	SCE $\pm$ SD
10	20	0.1	0.048 ± 0.025	0.513 ± 0.211	0.000 ± 0.001	0.513 ± 0.211
		0.3	0.048 ± 0.025	0.344 ± 0.199	0.000 ± 0.001	0.345 ± 0.199
		0.45	0.048 ± 0.024	0.258 ± 0.176	0.000 ± 0.001	0.258 ± 0.175
	500	0.1	0.009 ± 0.005	0.506 ± 0.206	0.000 ± 0.001	0.506 ± 0.206
		0.3	0.009 ± 0.005	0.348 ± 0.201	0.000 ± 0.001	0.348 ± 0.201
		0.45	0.009 ± 0.005	0.259 ± 0.179	0.000 ± 0.001	0.259 ± 0.179
100	20	0.1	0.015 ± 0.008	0.527 ± 0.207	0.000 ± 0.001	0.527 ± 0.207
		0.3	0.015 ± 0.008	0.363 ± 0.206	0.000 ± 0.001	0.363 ± 0.206
		0.45	0.015 ± 0.008	0.282 ± 0.183	0.000 ± 0.001	0.282 ± 0.183
	500	0.1	0.003 ± 0.002	0.527 ± 0.206	0.000 ± 0.001	0.527 ± 0.206
		0.3	0.003 ± 0.002	0.363 ± 0.207	0.000 ± 0.001	0.363 ± 0.207
		0.45	0.003 ± 0.002	0.279 ± 0.185	0.000 ± 0.001	0.279 ± 0.185

N = number of subgroups considered, K = number of mates within a subgroup, and SD_{within} = trait standard deviation within a subgroup. The standard deviation of Averaged *r* (SD of Averaged *r*) within a scenario is averaged across the 21 correlation coefficients × 11 coefficients of variation (231 scenarios) to show the relatively low sampling error in the study (similar to Pooled SD of *r*; not shown). Averaged *r* = mean (± SD) of the correlation coefficients across pairs for the 11 coefficients of variation. Note that this is a very good estimator of the true correlation, as the expected mean is zero. Pooled *r* = estimated correlation coefficient after pooling pairs for the 11 coefficients of variation. SCE = pooled minus averaged *r* across the 11 coefficients of variation. See Figures 1 and 2 for a graphical definition of Pooled *r*, Averaged *r*, and SCE.

between classes, especially if a nonsymmetrical distribution of specimens with intermediate values exists. With regard to the two ecotypes of *L. saxatilis*, conversely, these have been objectively described based on determined trait values (using simultaneously the presence of ridges and bands; see Rolán-Alvarez 2007), together with molecular markers, which have allowed confirmation of the correct ecotype classification even in the mid shore where they hybridize (Galindo et al. 2013). In addition, to provide an objective measurement of shell color in *L. fabalis*, we have measured their reflectance spectra (Andersson and Prager 2006), following the methodology described by Bybee et al. (2012). The experimental validation of our qualitative shell color classifications by use of objective reflectance spectra is explained in an analysis in the appendix section.

STATISTICAL ANALYSIS

The assortative mating of the qualitative traits was estimated by means of the I_{PSI} index (Rolán-Alvarez and Caballero 2000), with the exception of one of the *L. saxatilis* SCE estimates for which only Yule's V statistic was available (Table 2). The significance of I_{PSI} was calculated by the bootstrapping method using the software package JMATING ver 1.08 that estimates deviations of assortative mating from that of random mating (Carvajal-Rodríguez and Rolán-Alvarez 2006). The statistical bootstrap test of the sampled data works in a similar way to a standard chi-square test of association (between the two trait values in males and females). To verify the former approach, however, we have also calculated Pearson's correlation r coefficient with respect to the two color/ecotype classes recoded into 1 and 0. In addition, we used Kendall's tau nonparametric coefficient to test for their significance (Sokal and Rohlf 1995). While I_{PSI} is the most suited estimator for assortative mating in relation to qualitative traits, the classic correlation coefficient is provided as an independent nonspecific statistic that does not correct for mating propensity or population frequency (see Rolán-Alvarez and Caballero 2000). Mean statistical differences between years or subgroups were estimated by classical *t*-tests. We used discriminant analysis to distinguish yellow and brown shell colors based on the information retrieved from the specimen's raw spectral reflectance data (see Appendix). Classical parametric and nonparametric tests were implemented by either the software SPSS for Windows version 19.0 or Poptools version 3.0 (Hood 2010).

Both I_{PSI} and r statistics implicitly assume that mate choice is accomplished at the same scale that encompasses the whole sample of pooled mating pairs. As it is impossible to undertake analysis of a single microarea with only one mating pair, a different practical approach needs to be taken. In both *L. saxatilis* and *L. fabalis* we have estimated assortative mating from the pooled set of mating pairs as well as within homogeneous subgroups according to trait (ecotype or color) frequency. Mated and unmated specimens within each microarea were used for grouping microareas before embarking on the three class analysis based on the color (or ecotype) frequencies (see also Rolán-Alvarez et al. 1995a): yellow microareas (yellow frequency $\geq$ 66%), brown microareas (yellow frequency $\leq$ 33%), and mixed

Table 2. Summary of estimates of assortative mating and the SCE in *L. saxatilis* and *L. fabalis*.

				Subgroups					
Species	Year	Statistic	Pooled	Yellow/RB	Mixed	Brown/SU	Averaged	CV	$SCE_{Pooled-Averaged}$
L. saxatilis	1991[1]	Yule's V	0.77^{***}	0.85^{**}	0.70^{***}	0.66^{***}	$0.74^{**} \pm 0.100$	–	0.03
L. saxatilis	1999[2]	I_{PSI}	0.83^{***}	–	–	–	$0.86^{***} \pm 0.100$	0.19	–0.03
		r	0.73^{***}				$0.76^{***} \pm 0.150$		–0.03
L. fabalis	1990[3]	I_{PSI}	-0.15^{*}	–	–	–		–	
L. fabalis	2011[4]	I_{PSI}	−0.06	-0.45^{*}	-0.36^{*}	-0.43^{*}	$-0.41^{**} \pm 0.047$	0.16	0.35
		r	−0.06	–0.19	-0.35^{*}	–0.14	$-0.23^{?} \pm 0.110$		0.17
L. fabalis	2012	I_{PSI}	−0.15	$-0.32^{?}$	-0.53^{**}	–	-0.43 ± 0.148	0.13	0.28
		r	−0.14	–0.21	-0.50^{*}	–	-0.35 ± 0.205		0.21
L. fabalis	2013	I_{PSI}	−0.12	-0.52^{*}	–0.19	–0.47	$-0.39^{?} \pm 0.178$	0.13	0.27
		r	−0.12	-0.37^{*}	–0.19	–0.25	$-0.27^{*} \pm 0.092$		0.15
L. fabalis	$Mean_{2011–13}$	I_{PSI}	$-0.12^{*} \pm 0.042$				$-0.41^{*} \pm 0.146$		$0.30^{**} \pm 0.044$
		r	$-0.11^{*} \pm 0.042$				$-0.28^{*} \pm 0.096$		$0.18^{**} \pm 0.031$

In *L. saxatilis* assortative mating is estimated for two sympatric ecotypes (ridged and banded, RB; smooth and unbanded, SU), while in *L. fabalis* it is estimated for two shell colors, yellow and brown. $I_{PSI} \pm$ SD and Pearson's *r* correlation are the estimators of assortative mating while the scale-of-choice effect is measured as SCE = $statistic_{pooled}$ − $statistic_{averaged}$ (see Fig. 1). The analysis is presented for distinct years, both in relation to the whole set of observed experimental data (Pooled) and in the different homogenous subgroups (using microareas categorized according to whether they are mostly yellow/RB, brown/SU, or mixed in frequency; see text). Statistical test of I_{PSI} was based on bootstrapping of the summarized tables (see Table appendix S1 for *L. fabalis* and S2 for *L. saxatilis*), while the average values were tested by means of a classical Student's *t*-test (using 0 as the null hypothesis). CV = coefficient of variation of the yellow and brown shell color frequencies across microareas.

***P* < 0.10, **P* < 0.05, ***P* < 0.01, ****P* < 0.001.

[1]See Table 2 in Rolán-Alvarez et al. (1995a).

[2]Data from Rolán-Alvarez et al. (1999) includes 12 different populations instead of microareas (see Table S2).

[3]Data from Rolán-Alvarez and Ekendahl (1996). This study did not include microarea information.

[4]Data from Rolán-Alvarez et al. (2012).

microareas (intermediate frequencies). The distribution of color frequencies across microareas and years is shown in Figure S2. The SCE was assessed by subtracting the assortative mating averaged across the three groups that differed in frequency from the assortative mating that was estimated from the whole set of pairs (Table 2). For the second dataset of *L. saxatilis*, the SCE was estimated as the assortative mating of the pooled populations minus the assortative mating averaged across populations since there was no available information on microareas within populations (see raw data in Table S2).

THE DATASET OF JIANG ET AL. REANALYZED

Jiang et al. (2013) reviewed 1116 trait correlations linked to 254 species, providing a detailed summary of the data with information concerning the direction of assortative mating (positive; negative; none). The review omitted information about the sampling scale or sampling heterogeneity for each species, yet provided some data which in principle allows testing of a few basic predictions about the SCE mechanism. For example, the SCE mechanism, at least as it relates to the spatial level, is more likely to bias estimates of negative assortative mating (toward positive values) in species with relatively low adult mobility (thereby producing mate choice at a scale smaller than that of the whole sample). Low mobility, however, does not necessarily imply SCE; formal tests for investigating the importance of this mechanism will be required following the strategy outlined in Table 2. Nonetheless, in order to check if the data published by Jiang et al. (2013) suggests that this mechanism is widespread, we counted the number of positive and negative correlations displayed by different taxa that a posteriori were classified as either low or high mobility organisms.

Results

NUMERICAL DESCRIPTION OF THE SIMULATED SCE

We simulated the SCE under a set of biologically realistic scenarios (Table 1). The averaged Pearson's *r* coefficient estimated within the *N* (10 or 100) groups, for the *K* (20 or 500) pairs, and the three levels of intragroup variation (0.1, 0.3, or 0.45; the last coincides with the intragroup SD observed in *L. fabalis*) always yielded a value close to zero (Table 1). This means that the averaged *r* estimate did not show any systematic bias during the estimation process. However, the pooled *r* showed an important bias in all cases: always toward positive values (Table 1). It is

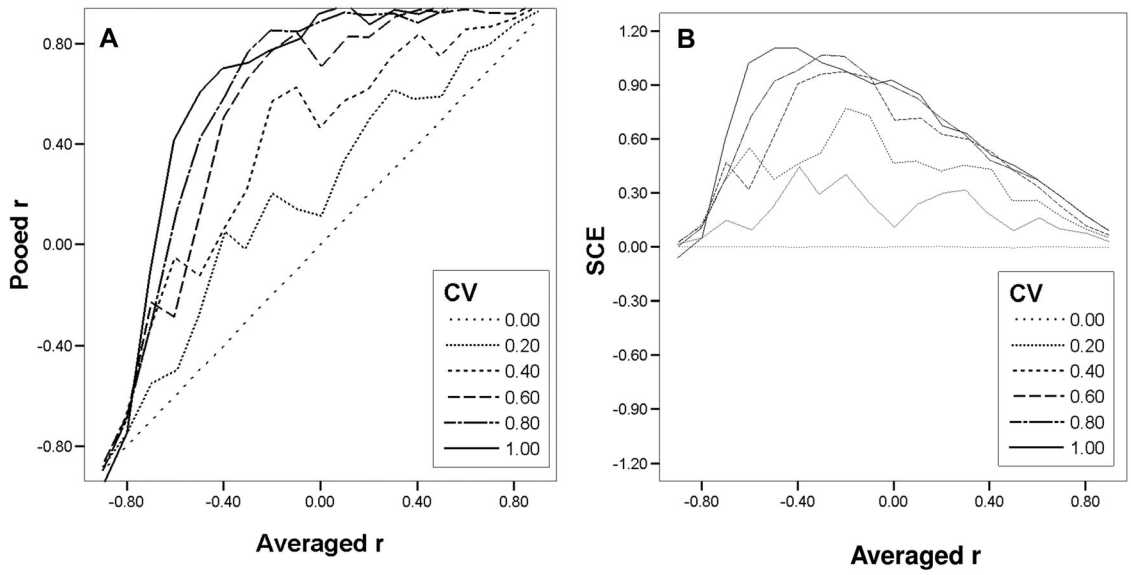

Figure 3. **Description of the SCE from simulations. (A) Relationship between the *Pooled r* and the *Averaged r* coefficients obtained through simulating ($N = 10$, $K = 20$, and $SD_{within} = 0.3$; see Table 1 and text). The different curves represent the relationship between variables for a set of simulated coefficients of variation (CV). The sampling error of each simulation was low (see Table 1), as each point was averaged across 100 simulations (see text). Results falling above the 1:1 line (CV = 0) indicate positive bias, while results falling below the 1:1 line indicate negative bias. (B) Biplot of the SCE (*Pooled r – Averaged r*; see graphic definitions in Fig. 1) and Averaged *r* values estimated on simulated mating pairs. Larger SCE values are produced at intermediate correlations for the simulated set of conditions.**

noteworthy that averaged r (and also pooled r, but not shown) showed relatively low sampling error across the 100 replicated simulations (see column "standard deviation of r"). By contrast, the bias estimated by the SCE ($r_{pool} - r_{averaged}$) was typically large and, while not affected very much by the number of pairs (K) or the number of subgroups (N), was severely affected by the level of intragroup variability (SD_{within}; Table 1). Both the SCE and the pooled r (see corresponding columns) showed relatively high standard deviations across all scenarios, due either to the trait's coefficient of variation across subgroups (CV) or to the simulated level of assortative mating. Pearson's pooled r was plotted against the averaged r for distinct CV (for $N = 10$; $K = 20$; $SD_{within} = 0.3$; Fig. 3A). For any given CV > 0 the pooled estimate of assortative mating was typically positively biased. For the same conditions, the SCE was plotted against the averaged r (Fig. 3B), and usually showed a larger bias under negative assortative mating (SCE $= 0.466 \pm 0.453$) than under positive assortative mating (SCE $= 0.316 \pm 0.280$; see DRYAD simulation data and Fig. 3). While the magnitude of the effect can be influenced by the relative sizes of the intragroup and intergroup variances (SD_{within} and CV from Table 1), the outcome will always lead to an inference of positive assortative mating, even when the opposite is true.

L. saxatilis AND *L. fabalis* SCE ESTIMATES

SCE values in *L. saxatilis* were estimated based on the published ecotype assortative mating estimates, having been computed as the difference between the pooled samples and the mean across the homogenous subgroups using the 1991 dataset (SCE = 0.03; first row in Table 2). In addition, the SCE was also estimated by undertaking a reanalysis of previously published data derived from 12 populations sampled in 1994 (SCE = –0.03; CV = 0.2; second row in Table 2). The simulated SCE for Pearson's r processed under similar conditions ($N = 10$; $K = 20$; CV = 0.2; $SD_{within} = 0.45$; $r = 0.8$) yielded 0.043 (see DRYAD simulation data), displayed the same order of magnitude as the empirical estimate.

The SCE with respect to the pattern of negative (color-based) assortative mating exhibited by *L. fabalis* was estimated based on the analysis of old and new data (Table 2). Negative assortative mating was observed in the same population over a 4-year period with 1990 being the only significant year (2013 was marginally significant; Table 2), although the mean across years was significantly different from zero by a t-test ($P < 0.05$). Remarkably, assortative mating estimates in relation to the homogeneous color frequency subgroups increased on average to –0.41 compared to the averaged pool estimate of –0.12 (see mean subgroups in Table 2). This is the direct consequence of a large scale-of-choice effect for each of the three years (mean 0.30; Table 2). The SCE occurs each year with the corresponding microareas displaying a similar coefficient of variation (ranging from 0.13 to 0.16). The same pattern was revealed using Pearson's correlation coefficient and other nonparametric correlation tests (SCE = 0.18; Table 2), which showed a similar SCE (0.12) value when simulated using

Table 3. Jiang et al. (2013) dataset reanalyzed.

Taxon	*N*	Mobility	% Negative
Bird	134	High	6.7
Insect	101	High	9.9
Fish	25	High	8
Crustacean	101	Low	0
Amphibian	36	Low	5.6
Gastropod	6	Low	0

N **= number of species-trait combinations with assortative mating data.**
Mobility = a qualitative descriptor of adult mobility (averaged) from each taxon.
% Negative = percentage of negative assortative mating.
Low dispersal species should manifest a higher SCE on average, therefore biasing disassortative mating estimates toward positive values.

similar conditions ($N = 10$; $K = 20$; CV = 0.2; $SD_{within} = 0.45$; $r = –0.4$).

JIANG ET AL. 2013 DATA REANALYZED

The Jiang et al. (2013) data were reanalyzed first by separating organisms into high and low mobility classes and then counting the number of taxa displaying negative assortative mating (Table 3). As theory would predict for the SCE mechanism, low mobility organisms exhibited lower frequencies of negative assortative mating, since for these organisms it is more likely that the true scale of choice is smaller than the scale of the entire sample of mating pairs. It is important to note, however, that the possibility that temporal variation has biased these published estimates has not been examined and cannot be ruled out. This result therefore provides indirect evidence for the importance of the SCE in connection with complex sets of assortative mating estimates from different organisms.

Discussion

In many organisms, studies of assortative mating in the wild may require sampling at a much smaller scale than most investigators typically use. Otherwise, misleading conclusions might result if, at the smaller choice scale, unappreciated heterogeneity exists. Pearson's correlation coefficients were previously known to produce nonsensical estimates when heterogeneous samples are pooled (Hassler and Thadewald 2003)—a process influenced by the relative sample sizes and variances across and within the pooled subgroups. We have further extended this finding by establishing that such a bias may also affect other statistics used to measure correlation such as I_{PSI}. Our analysis reveals that under biologically realistic scenarios of mate choice, a systematic (and occasionally huge) bias toward positive values is always produced irrespective of the true degree of positive or negative assortative mating (Figs. 1 and 2). Interestingly, the SCE may bias assortative mating estimates toward positive values even under circumstances of extremely low sample sizes (20 mating pairs; Fig. 3). A similar statistical bias was noticed when pooling correlation coefficients in the context of heterogeneous microarray data (Almeida-de-Macedo et al. 2013). More recently, the same phenomenon has been shown to bias different statistics estimating intragenomic correlation (Cocho et al. 2014). While our results are quite robust under the simulated conditions, it is useful to consider our empirical examples of the SCE as it affects the two dissimilar scenarios of assortative mating in turn.

Strong (positive) assortative mating has been confirmed in *L. saxatilis* between ecotypes that live in sympatry, contributing to their partial reproductive isolation (reviewed in Rolán-Alvarez 2007). Published estimates of ecotype assortative mating in the wild have allowed us to estimate the scale-of-choice effect ($SCE_{saxatilis} = \pm 0.03$; Table 2). Positive assortative mating between ecotypes is produced by positive size assortative mating linked to an adaptive size divergence between the two ecotypes present at the mid shore, where they meet in sympatry. In *L. saxatilis*, mate choice has been attributed to males following female mucus trails in a size-dependent manner (Davies and Beckwith 1999; Johannesson et al. 2008). This phenomenon may result in sexual isolation where the two ecotypes, differing in mean size, meet in sympatry (Conde-Padín et al. 2008; Johannesson et al. 2008). Similar positive assortative mating mechanisms have also been observed to contribute to the partial reproductive isolation of stick-insects, cichlid fish and palms (Nosil et al. 2002; Barluenga et al. 2006; Savolainen et al. 2006; Martin 2013) or to operate in organisms where strong divergent selection affects the chosen trait (Ariyomo and Watt 2013).

By contrast, negative assortative mating for shell color has been thoroughly described in *L. fabalis* by evaluating the same population over a period of different years (Rolán-Alvarez et al. 2012). Snails did show negative assortative mating during the four years of the study. However, if mate choice is in fact produced within the range of the microareas studied, the data pooling that yielded the former estimates could have biased them (as expected from the scale-of-choice effect). In an attempt to reduce such a potential bias we proceeded to reanalyze the data by first grouping homogeneous frequency classes within each microarea. These results yielded even stronger negative assortative mating estimates, due to a large scale-of-choice effect (0.30 on average; Table 2). This finding supports the view that even strong negative assortative mating at the microarea scale results in positively-biased values when microareas that are heterogeneous for color frequency are pooled. This fact also implies that mate choice is made at a smaller scale than the hundreds of square meters used for sampling.

The reflectance spectra analysis further confirmed our prior classifications of shell color from the 2013 sample, as the SCE

estimates were nearly identical using either reflectance spectra variables or qualitative numerical values (see Appendix). While we have specifically selected shell color for study, it is possible that mate choice in *L. fabalis* is in fact influenced by unknown chemicals present in the mucus or in the skin, which are tightly correlated with shell pigment composition. Under this scenario, mate choice would be caused by a correlative effect linked to the color rather than due to the color itself. In other organisms, color-based choice during mating may be the result of visual perception directly (Handcox et al. 2010; Bybee et al. 2012). Regardless of the mechanism, negative assortative mating based on shell color remains a powerful force for maintaining color polymorphism in the population (see Takahashi and Hori 2008; Holman et al. 2013). A similar contribution of negative assortative mating to the maintenance of polymorphisms has been observed in cichlid fish (Takahashi and Hori 2008) and maple tree populations (Field and Barrett 2012).

We have described a new process (the scale-of-choice effect) that could potentially bias estimates of assortative mating in the wild. The scale-of-choice effect biases estimates only when (1) the choice happens at a smaller scale than that of the pair sampling, and (2) when some heterogeneity for the trait persists at such a scale. Usually the former is unknown in most species, yet due to the statistical properties of the SCE we are in a position to provide a method that yields unbiased (or minimally biased) assortative mating estimates. The same method can then be used indirectly to estimate whether or not the SCE exists in a particular species. By having a sample of unmated specimens surrounding every pair, it is possible to subdivide a posteriori the mating pairs into homogeneous subgroups based on the trait studied (in this case we have used three frequency classes). This procedure allows comparison of assortative mating estimates within homogeneous groups, as has been done in Table 2. The resulting SCE estimate, if significant, represents an indirect way of confirming that the organism's true scale-of-choice is in fact smaller than the sampling scale.

A related matter is whether or not negative assortative mating is more frequent in nature than has previously been estimated by meta-analysis of experimental cases (Jiang et al. 2013). Definitive corroboration of this possibility will either require new experiments among the different species surveyed, or by undertaking a more sophisticated meta-analysis that considers the possibility of incorporating the SCE. Nonetheless, our brief reanalysis of the frequency of negative assortative mating from the former review suggests that low dispersal species display the lowest negative assortative mating frequencies, in agreement with theoretical expectation under the SCE. In order to imagine the potential impact that the SCE could have on the meta-analysis provided by Jiang and coworkers (2013), one could simplistically assume that either all of the estimates were moderately biased (using the averaged SCE obtained from all of the simulated scenarios with $CV = 0.1$; $SCE = 0.128$) or else just one out of five estimates were randomly affected by a strong SCE (0.7). Correcting the correlation coefficients reviewed by these authors accordingly (by subtracting the latter estimated SCE from the observed estimates) would generate a mean r value of 0.097 and 0.080 under the first and second scenarios, respectively. These estimates are substantially different from the estimate of 0.28 in Jiang et al. (2013). It would also increase the number cases of negative assortative mating in Jiang et al. (2013) from 19 to 38% or 37%, respectively. Of course, we do not claim that these new "estimates" are correct; we just want to emphasize the impressive impact that either frequent and moderate or rare and high levels of SCE could have on such a meta-analysis. Another prediction under the SCE is that a higher percentage of negative assortative mating cases should be detectable in the laboratory compared to the wild, as estimate bias (by pooling heterogeneous samples) is not possible in the laboratory. Although these observations remain far from being conclusive, they do suggest that the problem deserves further study. It would be extremely unfortunate to miss an important evolutionary process in the wild such as negative assortative mating only because of an overlooked and subtle methodological error.

ACKNOWLEDGMENTS

We thank N. Santamaría for administrative contributions and Ian Dworkin and Spencer Hall together with one anonymous referee for key comments and suggestions for revising the manuscript. This work was supported by the Xunta de Galicia (Grupo con Potencial de Crecimiento, GPC2013-011), the Ministerio de Economía y Competitividad (BFU2013-44635-P) and also by Fondos FEDER ("*Unha maneira de facer Europa*"). All authors declare to have no conflict of interest. E-R-A. enjoyed a grant from the Ministerio de Educación, Cultura y Deporte for a sabbatical period at U.C.I. (PRX14/00005).

DATA ARCHIVING

Doi: 10.1111/evo.12691

LITERATURE CITED

Ariyomo, T. O., and P. J. Watt. 2013. Disassortative mating for boldness decreases reproductive success in the guppy. Behav. Ecol. 24:1320–1326.

Almeida-de-Macedo, M. M., N. Ransom, Y. Feng, J. Hurst, and E. S. Wurtele. 2013. Comprehesive analysis of correlation coefficients estimated from pooling heterogeneous microarray data. BMC Bioinformatics 14:214.

Ambrosini, R., and N. Saino. 2010. Environmental effects at two nested spatial scales on habitat choice and breeding performance of barn swallow. Evol. Ecol. 24:491–508.

Andersson, M. 1994. Sexual selection. Princeton Univ. Press, Princ eton, NJ.

Andersson, S. 1999. Morphology of UV reflectance in a whistling-thrush: implications for the study of structural color signaling in birds. J. Avian. Biol. 30:193–204.

Andersson, S., and M. Prager. 2006. Quantifying colors. *In* G.E. Hill and K.J. McGraw, eds. Bird coloration volume I. Mechanisms and measurements. Harvard Univ. Press, London.

Barluenga, M., K. N. Stölting, W. Salzburger, M. Muschick, and A. Meyer. 2006. Sympatric speciation in Nicaraguan crater lake cichlid fish. Nature 439:719–723.

Bybee, S. M., F. Yuan, M. D. Ramstetter, J. Llorente-Bousquets, R. D. Reed, D. Osorio, and A. D. Briscoe. 2012. UV Photoreceptors and UV-Yellow wing pigments in *Heliconius* butterflies allow a color signal to serve both mimicry and intraspecific communication. Am. Nat. 179:38–51.

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

Casares, P., M. C. Carracedo, B. D. Rio, R. Piñeiro, L. García-Flórez, and A. R. Barros. 1998. Disentangling the effects of mating propensity and mating choice in *Drosophila*. Evolution 52:126–133.

Conde-Padín, P., R. Cruz, J. Höllander, and E. Rolán-Alvarez. 2008. Revealing the mechanisms of sexual isolation in a case of sympatric and parallel ecological divergence. Biol. J. Linn. Soc. 94:513–526.

Cocho, G., P. Miramontes, R. Mansilla, and W. Li. 2014. Bacterial genomes lacking long-range correlations may not be modeled by low-order Markov chains: the role of mixing statistics and frame shift of neighboring genes. Comput. Biol. Chem. 53:15–25.

Cruz, R., M. Carballo, P. Conde-Padín, and E. Rolán-Alvarez. 2004. Testing alternative models for sexual isolation in natural populations of *Littorina saxatilis*: indirect support for by-product ecological speciation? J. Evol. Biol. 17:288–293.

Coyne, J. A., and H. A. Orr. 2004. Speciation. Sinauer Associates, Inc. Chicago, IL.

Coyne, J. A., S. Elwyn, and E. Rolán-Alvarez. 2005. Impact of experimental design on *Drosophila* sexual isolation studies: direct effects and comparison to field hybridization data. Evolution 59:2588–2601.

Dabire, K. R., S. Sawadodgo, A. Diabate, K. H. Toe, P. Kengne, A. Ouari, C. Constantini, C. Gouagna, F. Simard, T. Baldet et al. 2013. Assortative mating in mixed swarms of the mosquito *Anopheles gambiae* s.s. M and S molecular forms, in Burkina Faso, West Africa. Med. Vet. Entomol. 27:298–312.

Davies, M. S., and P. Beckwith. 1999. Role of mucus trails and trail-following in the behaviour and nutrition of the periwinkle *Littorina littorea*. Mar. Ecol. Prog. Ser. 179:247–257.

Duraes, R., B. A. Loiselle, P. G. Parker, and J. G. Blake. 2009. Female mate choice across spatial scales: influence of lek and male attributes on mating success of blue-crowned manakins. Proc. R. Soc. B. 276:1875–1881.

Field, D. L., and S. C. H. Barrett. 2012. Disassortative mating and the maintenance of sexual polymorphism in painted maple. Mol. Ecol. 21:3640–3643.

Galindo, J., M. Martínez-Fernández, S. Rodríguez-Ramilo, and E. Rolán-Alvarez. 2013 The role of local ecology during hybridisation at the initial stages of ecological speciation in a marine snail. J. Evol. Biol. 26:1472–1487.

Gavrilets, S. 2004. Fitness landscapes and the origin of species. Princeton Univ. Press, Princeton, NJ.

Gilbert, D. G., and W. T. Starmer. 1985. Statistics of sexual isolation. Evolution 39:1380–1383.

Gray, S. M., and J. S. McKinnon. 2007. Linking color polymorphism maintenance and speciation. Proc. R. Soc. B. 22:71–79.

Gwynne, D. T., W. J. Bailey, and A. Annells. 1998. The sex in short supply for matings varies over small spatial scales in a katydid (*Kawanaphila nartee*, Orthoptera: Tettigoniidae). Behav. Ecol. Sociobiol. 42:157–162.

Handcox, D., C. J. Hoskin, and R. S. Wilson. 2010. Evening up the score: sexual selection favours both alternatives in the colour-polymorphic ornate rainbowfish. Anim. Behav. 80:845–851.

Hárosi, F. I. 1996. Visual pigment types and quantum-catch ratios: implications from three marine teleosts. Biol. Bull. 190:203–212.

Hassler, U., and T. Thadewald. 2003. Nonsensical and biased correlations due to pooling heterogeneous samples. Statistician 52:367–379.

Holman, S., L. J. van Zweden, T. A. Linksvayer, and P. d'Ettorre. 2013. Crozier's paradox revisited: maintenance of genetic recognition systems by disassortative mating. BMC Evol. Biol. 13:211.

Hood, G. M. 2010. PopTools version 3.2.5. Available at www.poptools.org. Accessed April 9, 2015.

Jiang, Y., D. I. Bolnick, and M. Kirkpatrick. 2013. Assortative mating in animals. Am. Nat. 181:E125–E138.

Johannesson K., J. N. Havenhand, P. R. Jonsson, M. Lindegarth, A. Sundin, and J. Hollander. 2008. Male discrimination of female mucous trails permits assortative mating in a marine snail species. Evolution 62:3178–3184.

Kirkpatrick, M. 2000. Reinforcement and divergence under assortative mating. Proc. R. Soc. B. 267:1649–1655.

Knoppien, P. 1985. Rare male mating advantage: a review. Biol. Rev. 60:81–117.

Kozminskii, E. V. 2011. Inheritance of longitudinal shell bands in the snail *Littorina obtusata* and *L. saxatilis* (Gastropdoa: Prosobranchia). Russ. J. Genet. 47:1112–1119.

Langston, N., S. Rohwer, and D. Gori. 1997. Experimental analyses of intra- and intersexual competition in red-winged blackbirds. Behav. Ecol. 8:524–533.

Lewontin, R., D. Kirk, and J. Crow. 1968. Selective mating, assortative mating, and inbreeding: definitions and implications. Eugen. Q. 15:141–143.

Llopart, A., D. Lachaise, and J. A. Coyne. 2005. An anomalous hybrid zone in *Drosophila*. Evolution 59:2602–2607.

Martin, C. H. 2013. Strong assortative mating by diet, color, size, and morphology but limited progress toward sympatric speciation in a classic example: Cameroon crater lake cichlids. Evolution 67:2114–2123.

Matsubayashi, K. W., S. Kahono, and H. Katakura. 2013. Divergent host plant preference causes assortative mating between sympatric host races of the ladybird beetle, *Henosepilachna diekei*. Biol. J. Linn. Soc. 110:606–614.

Montgomerie, R. 2006. Analyzing colors. *In* G. E. Hill and K. J. McGraw, eds. Bird coloration volume I. Mechanisms and measurements. Harvard Univ. Press, London.

Merrell, D. J. 1950. Measurement of sexual isolation and selective mating. Evolution 4:326–331.

Miller, A. M., C. McArthur, and P. J. Smethurst. 2009. Spatial scale and opportunities for choice influence browsing and associational refuges of focal plants. J. Anim. Ecol. 78:1134–1142.

Nosil, P., B. J. Crespi, and C. P. Sandoval. 2002. Host-plant adaptation drives the parallel evolution of reproductive isolation. Nature 417:440–443.

Pérez-Figueroa, A., A. Caballero, and E. Rolán-Alvarez. 2005. Comparing the estimation properties of different statistics for measuring sexual isolation from mating frequencies. Biol. J. Linn. Soc. 85:307–318.

Press, W. H. 2007. Numerical recipes: the art of scientific computing. Cambridge Univ. Press, Cambridge.

Pusey, A., and M. Wolf. 1996. Inbreeding avoidance in animals. Trends Ecol. Evol. 11:201–206.

Reid, D. G. 1996. Systematics and evolution of Littorina. Dorset Press, Dorchester.

Reimchen, T. E. 1989. Shell colour ontogeny and tubeworm mimicry in a marine gastropod *Littorina mariae*. Biol. J. Linn. Soc. 36:97–109.

Ritchie, M. E., and H. Olff. 1999. Spatial scaling laws yield a synthetic theory of biodiversity. Nature 400:557–560.

Rolán-Alvarez, E. 2007. Sympatric speciation as a by-product of ecological adaptation in the Galician *Littorina saxatilis* hybrid zone. J. Mollus. Stud. 73:1–10.

Rolán-Alvarez, E., and A. Caballero. 2000. Estimating sexual selection and sexual isolation effects in polymorphic traits. Evolution 54:30–36.

Rolán-Alvarez, E., and A. Ekendahl. 1996. Sexual selection and non-random mating for shell colour in a natural population of the snail *Littorina mariae* (Gastropoda: Prosobranchia). Genetica 97:39–46.

Rolán-Alvarez, E., J. Erlandsson, K. Johannesson, and R. Cruz. 1999. Mechanisms of incomplete prezygotic reproductive isolation in an intertidal snail: testing behavioral models in wild populations. J. Evol. Biol. 12:879–890.

Rolán-Alvarez, E., K. Johannesson, and A. Ekendahl. 1995a. Frequency- and density-dependent sexual selection in natural populations of Galician *Littorina saxatilis* (Olivi). Hydrobiologia 309:167–172.

Rolán-Alvarez E., M. Saura, A. P. Diz, M. J. Rivas, M. Alvarez, B. Cortés, A. deCoo, D. Estévez, and L. Iglesias. 2012. Can sexual selection and disassortative mating contribute to maintain shell color polymorphism in the intertidal snail *Littorina fabalis*? Curr. Zool. 58:460–471.

Rolán-Alvarez, E., C. Zapata, and G. Alvarez. 1995b. Distinct genetic subdivision in sympatric populations of two sibling species of the genus *Littorina* (Gastropoda: Littorinidae). Heredity 74:1–9.

Rolán-Alvarez, E., C. Zapata, and G. Alvarez. 1995c. Multilocus heterozygosity and sexual selection in a natural population of the marine snail *Littorina mariae* (Gastropoda: Prosobranchia). Heredity 74:17–25.

Sokal, R. R., and F. J. Rohlf. 1995. Biometry. W.H. Freeman and Co., New York.

Saks, L., K. McGraw, and P. Hörack. 2003. How feather colour reflects its carotenoid content. Funct. Ecol. 17:555–561.

Savolainen, V., M.-C. Anstett, C. Lexer, I. Hutton, J. J. Clarkson, M. V. Norup, M. P. Powell, D. Springate, N. Salamin, and W. J. Baker. 2006. Sympatric speciation in palms on an oceanic island. Nature 441:210–213.

Taborsky, B., L. Guyer, and P. Demus. 2014. 'Prudent habitat choice': a novel mechanism of size-assortative mating. J. Evol. Biol. 27:1217–1228.

Takahashi, T., and M. Hori. 2008. Evidence of disassortative mating in a T*anganyikan cichlid* fish and its role in the maintenance of intrapopulation dimorphism. Biol. Lett. 4:497–499.

Associate Editor: I. Dworkin
Handling Editor: J. Conner

Appendix

BIOLOGICAL INFORMATION OF *LITTORINA SAXATILIS*

Littorina saxatilis (Olivi 1972) is a North Atlantic gastropod with a low dispersal ability that grazes directly on rocky shores (Reid 1996). The species is dioecious, ovoviviparous with internal reproduction, and mating can be directly observed in the wild (Rolán-Alvarez et al. 1995a; Cruz et al. 2004). A striking adaptive polymorphism, ridged, and banded (RB) and smooth and unbanded (SU) ecotypes, has been described in the Galician populations associated with different shore-level microhabitats, with the two sympatric ecotypes having evolved partial reproductive isolation ($I_{PSI} = 0.7$, range –1 to 1; reviewed in Rolán-Alvarez 2007). We used previously published estimates of assortative mating in order to calculate the scale-of-choice effect in this species.

BIOLOGICAL INFORMATION OF *L. FABALIS*

L. fabalis (Turton 1825; former *L. mariae*) is a North Atlantic gastropod with low dispersal ability and that grazes on microalgae and diatoms growing on different macroalgae species (Reid 1996). In NW Spain, *L. fabalis* typically lives on the mid-shore tidal zone by foraging on *Fucus vesiculosus* (Rolán-Alvarez et al. 1995b). Negative assortative mating for shell color was detected in one particularly stable and dense population back in 1990 (Rolán-Alvarez and Ekendahl 1996), although the same sample also showed random mating with respect to genotypes of nine allozyme loci (Rolán-Alvarez et al. 1995c). In addition, negative assortative mating for shell color was corroborated 21 years later (Rolán-Alvarez et al. 2012). These studies used qualitative descriptors of shell colors.

QUALITATIVE TRAIT VALIDATION

Description of reflectance spectra methods

The reflectance spectra of shells were estimated at a point centered at the last whorl with the snail resting on the aperture. The probe holder (Ocean Optics RPH-1) was placed over the shell in such a way that the axis of the illuminating and detecting fiber (Ocean Optics R400-7-UV/VIS) was placed at an elevation of approximately 45˚ to the plane of the shell surface. Illumination was performed by a DH-2000 deuterium-halogen lamp, and reflectance spectra were measured with an Ocean Optics USB2000 spectrometer. In total, we examined 495 specimens obtaining a spectral reflectance curve from 300 to 700 nm wavelengths (with three readings taken between every nm), which was further processed with MATLAB. To establish the technical error involved among the 400 reflectance values, several specimens of different colors were measured twice (a correlation greater than 0.99 between replicated measurements was always observed).

Several strategies have been proposed to translate data of reflectance spectra into biologically useful variables. One possibility is to convert the reflectance spectra into brightness (B), saturation (S), and hue color (H) variables, depending on the particular color perception scenario assumed (Andersson and Prager 2006). We estimated B according to Andersson (1999). However, different statistics for S and H have been developed (see Table 3.2, page 108 in Montgomerie 2006), depending on the number of photoreceptor spectral sensitivity peaks assumed in the species. The number of spectral sensitivity peaks used for color vision is unknown in littorinids, but they are typically eaten by several blennid fish species (Reimchen 1989), which as teleosts may have between 3 and 4 spectrally distinct classes of

photoreceptor (see Hárosi 1996). Therefore, S and H were estimated in line with Saks et al. (2003), which assume four spectral classes of photoreceptor. Briefly,

$$B = \frac{\sum_{\lambda_{300}}^{\lambda_{700}} R_i}{400}. \quad \text{(A1)}$$

with *Ri* being the reflectance intensity and λ the corresponding wavelength (from 300 to 700 nm).

$$S = \sqrt{(B_r - B_g)^2 + (B_y - B_g)^2}. \quad \text{(A2)}$$

$$H = \text{Arc tan gent}\left(\frac{\frac{(B_y - B_b)}{B_1}}{\frac{(B_r - B_g)}{B1}}\right). \quad \text{(A3)}$$

with B_b, B_g, B_y, and B_r being the sum of the reflectance for the blue (400–474 nm), green (475–549 nm), yellow (550–624 nm), and red (625–700 nm) range, while B_1 comprises the sum of the reflectance spectra curve from 400–700 nm.

Results of the qualitative shell color validation

Each of the basic colors presented visually distinct reflectance spectra (Fig. S1b). Discriminant analysis permitted differentiation among the color classes by making use of the entire 400 spectra obtained ($P < 0.0001$). The cross-validation comparing the a priori visual and a posteriori spectra-defined color classes revealed a 99% overlap, thus validating our visual color classification.

In addition we repeated our assortative mating and SCE estimates using reflectance spectra. Table S3 provides the analysis of the individual visual measurements B, S, and H, as well as for the average of the three (BSH) and a discriminate score. Results are compared with the available qualitative estimates (using Pearson's *r* coefficient). Although the individual variables showed some heterogeneity, the overall means under all approaches were rather similar. Even the SCE estimates were nearly identical in value and typically significant. Ultimately, this analysis has demonstrated that our qualitative measurements are a biologically valid approximation of a more objective shell color measurement.

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's website:

Supplementary Figure S1. (a) Printed color model used to determine the qualitative color classes in *L. fabalis* specimens. (b) Representative reflectance spectra of the four main colors observed (brown, yellow, olive and orange).

Supplementary Figure S2. Frequency distribution of microarea colors in *L. fabalis* for different years: (a) 2011, (b) 2012, and (c) 2013. Each year the distribution of microareas did not deviate from a Normal distribution by a Kolmogorov–Smirnov test (P2011 = 0.247, P2012 = 0.187, P2013 = 0.143). The arrows divide the distribution into three equal sections that were used to define the homogeneous subgroups based on yellow frequencies (as in Table 2).

Supplementary Table S1. Data used to estimate the parameters of Table 1.

Supplementary Table S2. *L. saxatilis* mating type distributions among the two sympatric ecotypes (RB and SU) of 12 populations presented in Rolán-Alvarez et al. (1999).

Supplementary Table S3. The first row (*r* statistics) is the average across years (2011–2013) of color assortative mating using Pearson's correlation coefficient (see Table 1).