INTRODUCTION

Mate-choice decisions may be based on innate preferences, individual experience or on the observed choices of others. The last of these, mate-choice copying, is a type of social learning that involves updating mate preferences after observing others' choices (Dugatkin, 1992, 1996; Pruett-Jones, 1992; Wade & Pruett-Jones, 1990). Females may prefer either the successful males themselves (Bowers et al., 2012) or generalize their preferences to males with similar phenotypes ('trait copying' or 'mate-choice copying generalization'; Bowers et al., 2012; Brooks, 1998; Drullion & Dubois, 2008; Godin et al., 2005; Mery et al., 2009; Swaddle et al., 2005; White & Galef Jr., 2000; Witte & Noltemeier, 2002).
Mate-choice copying (synonymous with ‘mate copying’; Danchin et al., 2020) is documented in many taxa (Davies et al., 2020; Jones & DuVal, 2019), e.g., birds (Kniel et al., 2017; Swaddle et al., 2005), fish (Dugatkin & Godin, 1992; Heubel et al., 2008; Schlupp & Ryan, 1997; Witte & Ryan, 1998), mammals (Galef et al., 2008; Kavaliers et al., 2017), insects (Dagaeff et al., 2016; Germain et al., 2016; Mery et al., 2009; Monier et al., 2018) and spiders (Fowler-Finn et al., 2015). Empirical studies typically test whether copying occurs in a species or aim to evaluate conditions that make copying favoured over innate preferences. Theoretical studies have complemented the picture by exploring when copying is expected to invade and spread (Dubois et al., 2012; Losey et al., 1986; Pruett-Jones, 1992; Santos et al., 2017; Servedio & Kirkpatrick, 1996; Stöhr, 1998) and how copying affects the direction and strength of sexual selection (Agrawal, 2001; Kirkpatrick, 1982; Santos et al., 2014).

Mate-choice copying is argued to be adaptive because it allows the female to mate with higher-quality mates (Gibson & Höglund, 1992; Nordell & Valone, 1998; Valone & Templeton, 2002; Danchin et al., 2004; Dugatkin, 2005; Uehara et al., 2005; Wagner & Danchin, 2010; reviews Vakirtzis & Roberts, 2012, Varela et al., 2018). But this yields another question: why is copying needed to achieve a good outcome, i.e., why should any female not already possess preferences for best traits? If temporal changes are fast, i.e., the best mate for a given female varies through time (Ingleby et al., 2010), females might need updated information regarding better adapted (Getty, 2014; Wade, 2014) and/or more popular males (Kokko et al., 2007). Theoretical work that focuses on the ‘informedness’ of individuals has, indeed, considered age effects. Young females are conceivably inexperienced, and although they may possess innate preferences, observing and copying experienced females’ behaviour can be of benefit; still, it is not always clear why informedness should increase with age, i.e., why innate preferences are insufficient. Older and experienced females may also opt for mate-choice copying if, for example, their previous breeding attempt failed (Amlacher & Dugatkin, 2005; Danchin et al., 2008; Hill & Ryan, 2006). In both cases, the probability of copying is argued to increase with the difficulty of the discrimination task (Nordell & Valone, 1998).

Irrespective of age and experience, some individuals may be poorly informed in settings where spatial variation in environmental demands combines with local adaptation and gene flow (Holman & Kokko, 2014). Here, immigrant females may be uninformed of locally best traits. We show that this creates population-level feedback where mate-choice copying may facilitate the evolution of dispersal itself.

The reasoning is the following. Migrants experience an additional cost of dispersal when there is spatial variation in the environment (Berdahl et al., 2015; Blanquart & Gandon, 2014), because their traits are shaped by past selection in their natal environment, with unknown (and likely worse) performance in the environment they disperse to. For females, an additional complication is that their mate preferences—if genetically determined—may be similarly shaped by past selection. If preferences (and not just traits) are locally adapted, immigrant females may mate suboptimally. This problem can be minimized through copying, assuming that dispersal is not so strong that immigrants swamp locally adapted females, which would lead to uninformed immigrants copying each other’s choices. Copying thus allows immigrant-origin lineages gain locally adapted alleles and reduces the costs of dispersal. As shown below, it also ultimately reduces the degree of local adaptation via improved gene flow (reduction of genetic differentiation between different habitats).

2 MATERIALS AND METHODS

We model obligately sexual, haploid populations (which allows us to model one allele per locus, as our focus is not on effects of genetic dominance; see Kokko, 2007), initialized in a locally adapted state with only short-distance dispersal (a negative exponential dispersal kernel with a low mean). We thereafter allow the dispersal kernel to evolve in populations with or without a mate-choice copying locus. This locus, if present, has two alleles c (for innate preference) and C (for copying), only expressed in females. We contrast three scenarios. In the unconditional scenario, females with the C allele copy the mate choice of others at every copying opportunity; in the conditional scenario, they only do so if they have emigrated from their natal patch. In the control scenario, females only express innate preferences (no mate-choice copying evolves), while dispersal still evolves.

The population inhabits a toroid world (a doughnut-shaped continuous surface, each side having length 1). The benefit of using a toroid world is to avoid the need to specify what happens if individuals encounter the edge of the modelled world; under toroid assumptions, they can always continue moving in any direction even if the world is finite. The world is divided into $25^2$ patches with environmental heterogeneity that translates into a spatially varying and positively autocorrelated optimum for individuals’ trait values (note that we include spatial but not temporal variation of the environment). Positive spatial autocorrelation ensures that neighbouring patches do not differ very strongly from each other (details described in step 1 below); in other words, it allows the scale of environmental variation to be broader than a single patch. This is of benefit since broadly similar conditions across several patches, combined with finite dispersal distances, allow local adaptation to proceed and overcome the effects of drift—even if local patches (demes) themselves are kept small for the sake of realistic mate choice (females do not evaluate very many males before mating; Roff & Fairbairn, 2014).

We model genotypes with five different loci (S, T, P, C, D), some of which have sex-limited expression. The S locus determines sex, with alleles 0 (for females) and 1 (for males). The T locus (T stands for trait) has a pleiotropic effect. In both males and females, its value (real number between 0 and 1) relative to an environmental (spatially varying) optimum determines viability; in males, it additionally codes for a phenotype directly observable to females. The P locus likewise takes allelic values between 0 and 1 but is only expressed in
females; it specifies a preference for specific phenotypes of males. Females prefer males whose trait locus matches the female’s preference locus, which allows females to show local adaptation for locally adapted male traits. The C locus, with 2 alleles c and C, is likewise only expressed in females and controls female copying behaviour. Finally, the D locus, expressed in all individuals, determines the mean of the individual’s dispersal kernel, with values between 0 and 1 (where 1 would imply a mean distance equal to the length of the entire world).

Each run of the simulation proceeds as follows: (1) creation of the patch-specific environmental values, (2) population initialization, (3) survival, (4) dispersal, (5) mate choice by females and (6) reproduction. The simulation starts following steps 1 through 6, thereafter repeating steps 3 through 6 for each generation. Note that the order of the events dictates that viability selection is applied at the natal patch. Generations are non-overlapping.

In step 1, to simulate environmental heterogeneity, an environmental value is attributed to each patch using the algorithm described in Holman and Kokko (2014) that creates a matrix of environmental values with an adjustable spatial autocorrelation. The algorithm first generates a random value between 0 and 1 for each patch, then, for 252 × 100 iterations (a large enough number so that each patch, on average, experiences 100 impacts), it updates a randomly chosen patch p by setting its environmental value Ep to μp + r (1 − β), where μp is the mean environmental value of the eight patches surrounding patch p. These successive iterations bring the autocorrelation between neighbouring patches to a level that is controlled by the parameter β (0 ≤ β < 1). β = 0 implies there is no spatial autocorrelation, and when β = 1, neighbouring patches are very similar to each other. After all iterations, the resulting matrices were rescaled to have mean environmental value 0.5 and standard deviation 0.2 (see Figure S1 in Supporting Information for examples).

In step 2, we give 20,000 young individuals random coordinates 0 ≤ x, y < 1. The sex of each individual is randomly chosen (S locus is randomized to be 0 or 1), and we assume initial local adaptation with the following procedure. Values for T and P alleles are drawn randomly from a Gaussian distribution with mean Ep (where p is the patch that the individual’s coordinates imply it resides in) and standard deviation 0.05. Values that fall below 0 or exceed 1 are given a value of 0 or 1 respectively. Mate-choice copying is initially absent, i.e., all individuals are initialized with the c allele at the copying locus; C alleles are introduced later via mutation (see below). For the dispersal-determining D locus, individuals are initialized with random values drawn from a uniform distribution ranging between 0 and 0.005.

Viability selection (step 3) is applied in a density-dependent manner. At most 16 individuals survive in each patch. The value is chosen to allow each female to observe a low number of males and females (we expect patches to contain close to 8 females and 8 males; in nature, examples of more than 10 males sampled before mating appear rare; Roff & Fairbairn, 2014), and also to produce a global population of maximally 10,000 adults. No viability selection occurs in patches containing up to 16 individuals. In patches with more than 16 individuals, we first compute an adaptedness value for each individual i in patch p:

$$V_{Ai} = e^{-S_{n_i}(a_{ni} - a_{p})^2},$$

where $S_{n_i} ≥ 0$ scales the strength of natural selection and $a_{ni}$ is the allele value of individual i at the T locus. This expression is close to 1 when the difference between $a_{ni}$ and $E_{p}$ is small, indicating little mismatch between the phenotype and the environment, and declines at a rate controlled by $S_{n_i}$ as the mismatch increases. Actual survival probability depends on $V_{Ai}$ relative to competitors, and to generate stochasticity that allows mild differences in individual $V_{Ai}$ to translate to different rank orders when competing for the 16 survival ‘slots’, we compute the relative success $V_{Ai}^*$ for individual i by drawing from an exponential distribution with mean $V_{Ai}$. At each generation, the 16 individuals with the highest $V_{Ai}^*$ of each patch are retained, while the others die.

In step 4, dispersal distances are drawn for each individual from a negative exponential distribution with a mean equal to the allelic value at the D locus. The direction of dispersal is random. The toroidal arrangement of the patches ensures that the dispersal kernel can be applied even if dispersal distances exceed 1, the width of the world. More importantly (given that very long-distance dispersal is unlikely), short dispersal distances may mean not leaving the natal patch; coordinates are still updated, but the environmental value that the individual experiences does not change. We apply dispersal mortality of 0.05 to each individual whose new coordinates bring it outside its natal patch.

In step 5 (mate choice), females only perceive males who reside in the same patch as potential mates. Females only mate once; males may mate multiply. While we impose no restrictions on male mating capacity, their realized success is limited by the fact that males can only be chosen by females residing in the same patch. Within each patch, non-copier females mate first, after which copier females choose mates based on observations of all non-copier females’ matings. There are three categories of non-copiers: (a) all females with the c allele, (b) females with a C allele who have not left their natal patch in the conditional scenario and (c) females who attempt to copy but their patch offers no non-copier female whose behaviour they could observe. Non-copier females (of any category) observe all the males in their patch and preferentially mate with a male whose value at the T locus is close to the female’s value at her P locus. Specifically, each non-copier female f assigns a preference value $V_{Pmf}$ for each male m according to the equation:

$$V_{Pmf} = e^{-S_{j}(a_{j} - a_{m})^2},$$

where $S_{j} ≥ 0$ scales the strength of sexual selection that female choice can impose on males, $a_{pf}$ is the female’s allelic value at the P locus and $a_{tm}$ is the male’s allelic value at the T locus. The expression reaches its highest possible value, 1, when the male trait perfectly matches the female’s preference, and declines towards zero for increasing levels of mismatch. The probability that female f chooses male m* in the presence of other competitors is

$$\text{Prob}(f, m^*) = \frac{V_{Pmf}}{\sum_{m=1}^{M_p} V_{Pmf}},$$

where $M_p$ is the number of males in patch p.
A higher value of $S$ means that females are choosier, i.e. realized matings more closely match their innate preferences. Values at the $P$ locus do not modulate the strength of preferences, but instead indicate which phenotypes of males are preferred by each female. If $S = 0$, mating is random, i.e. values at the $P$ locus do not impact realized mate choice.

While non-copier females choose, copier females observe. We assume trait-based mate-choice copying (Danchin et al., 2018), i.e., copying females update their preferences for certain phenotypes rather than specific males (thus a male with zero success so far can become favoured should he be phenotypically similar to a male observed to mate). If choices made by non-copier females yield a single winner among the males, copier females replace, phenotypically, their innate preferences with a value that equals the $T$ of this winner: $a_{Pj}^* = a_{Tm}$, where $m^*$ is the identity of the winner. In case of a tie, one of the males is randomly chosen to be the winner, independently so for each of the copier females.

A copier female $f$ then assigns the value $v_{Pm}$ to each male $m$ analogously to the non-copier case above,

$$v_{Pm} = e^{-S_i(a_{Tm} - a_{Pm})^2}, \quad (4)$$

and the actual choice of a mate follows Equation 3 above.

In step 6, reproduction occurs in all patches with at least one male and one female. These patches produce 32 offspring each. Each offspring has a mother and a father. The mother is chosen randomly (as we assume no fecundity differences among females), and the sire is the mother’s chosen mate as determined above. Breeding is density dependent with these assumptions: a female breeding in a patch with a total of $F_p$ females present will produce, on average, $32/F_p$ offspring. The expected number of offspring produced by a focal male is $32/F_p$ times the number of females who chose this male. Density dependence is relevant as it causes selection for dispersal to avoid competing with kin (Hamilton & May, 1977; Li & Kokko, 2019). Offspring are initially placed at the same coordinates as their mothers.

For each locus, offspring inherit one allele, randomly chosen from either parent, without any linkage between loci. The $T$, $P$ and $D$ loci have a mutation probability of 0.01. Mutation is implemented by adding, to the original allelic value, a random value drawn from a Gaussian distribution with mean 0 and standard deviation 0.05, 0.05 or 0.005 for $T$, $P$ and $D$ loci respectively. Post-mutation allelic values below 0 or above 1 are assigned the value of 0 or 1 respectively. For the biallelic mate-choice copying allele, a mutation changes the value of the allele from $c$ to $C$ and vice versa, and the mutation probability was time-dependent: it was first set to zero for a burn-in period of 7,500 generations (such that $C$ individuals remain absent), and to 0.001 thereafter, except for the control scenario (which never permits mate-choice copying). The ‘burning in’ phase of 7,500 generations allows for all other loci to evolve to equilibrium values, and for dispersal specifically to show a balance between avoidance of kin competition on the one hand and avoidance of dispersal costs (direct and indirect) on the other.

After breeding, all adults die, and the offspring experience viability selection as described above. Each simulation was run for 15,000 generations, 20 times for each parameter value of $S$ (taking the values of 0, 25, 50, 75 and 100) and for each scenario (control, unconditional and conditional). From generation 7,500 onwards, we recorded the genotype and coordinates of all individuals every 2,500 generations. This census was performed after the dispersal step and its associated mortality.

To understand the impact of mate-choice copying on local adaptation, we created a local adaptation score ($L$) for each population. This measure quantifies the degree to which trait values, $a_{Tm}$, align with the environmental value in the $E_p$ in the breeding patch:

$$L = -\frac{\sum_p (a_{Tm} - E_p)^2}{N}, \quad (5)$$

where $N$ is the global number of surviving adults after dispersal. The negative sign in Equation 5 implies that high $L$ corresponds to better local adaptation.

We compared mean allelic $D$ values, $C$ allele proportions, the local adaptation score $L$ and trait-preference correlation coefficients across scenarios using one-way ANOVA tests (among all three scenarios) or Welch two samples $t$ tests (between the conditional and unconditional scenarios). Pairwise $t$ tests were used for comparisons of two different time points within populations. Significant differences were followed by a post hoc Tukey’s honest significance test. To account for multiple testing, we applied Bonferroni corrections so that the corrected $p$-value is $\min (n \times p, 1)$, where $p$ is the original $p$-value and $n$ is the number of hypothesis being tested. This may lead to a higher number of false negatives than other methods (McDonald, 2014), thus our conservative approach should provide a strong argument for all significant cases that we find and discuss.

The Shapiro–Wilk test of normality revealed that some of our data deviate significantly from normality. Although the tests we employed are robust against deviations from normality, as a precaution, we repeated all the analyses for the significantly non-normal data using nonparametric statistical tests. Since the results were similar, we only report parametric results.

The code for the mathematical model (programmed in C, using the GNU Scientific Library, version 2.3; Galassi et al., 2009) have been deposited in the Dryad Digital Repository (Sapage et al., 2020). The statistical analyses were done with R, version 3.6.1 (R Core Team, 2019).

3 | RESULTS

3.1 | Presence of mate-choice copying increases dispersal

Scenarios did not differ with respect to dispersal at generation 7,500 (all corrected $p > 0.05$ for differences in mean $D$ allele values; Table S1), an expected outcome since copying was not yet present.
in any of the scenarios. The control scenario that continued to lack mate-choice copying in subsequent generations did not show any significant difference in dispersal tendency between generations 7,500 and 15,000 (all corrected \( p > 0.05 \); Table S2). In scenarios where the C allele was introduced by allowing mutations to occur from generation 7,500 onwards, dispersal experienced a new evolutionary boost followed by reaching a new equilibrium (Figure 1), as evidenced by no significant difference when contrasting mean D allele values between generations 12,500 and 15,000 (all corrected \( p > 0.05 \); Table S2).

**FIGURE 1** Mate-choice copying selects for higher dispersal. Line plots: evolution of D for each sexual selection strength \( (S_S) \) as indicated in the plot and mate-choice copying scenario ('control', 'unconditional' and 'conditional' as indicated by colour). Lines depict the means (across 20 simulation runs) of (population-wide) D and the standard deviations of the means (shading); the vertical line denotes the generation at which mutation towards copying is introduced, and the coevolution with dispersal begins. Box plot: distribution of D at generation 15,000. Thick lines depict the medians of the distribution of each population's mean D; box, the 25% and 75% interquartile range; vertical dashed lines, the most extreme values within 1.5 of the interquartile range; opened circles, extreme values outside of this range. Stars indicate pairwise differences between populations within each case (Tukey HSD test for cases where the ANOVA test with the Bonferroni correction was significant, see Section 2). *\( p < 0.05 \), **\( p < 0.01 \), ***\( p < 0.001 \). Other parameters: \( \beta = 0.99 \), \( S_N = 50 \)
We only expect differences in dispersal across scenarios if \( S_S > 0 \): non-random mating is required for copying to mitigate the costs of dispersal that we envisage. Indeed, \( S_S = 0 \) yielded no differences across scenarios (corrected \( p > 0.05 \); Table S1) for allelic values of \( D \) at generation 15,000, while differences emerged when \( S_S > 0 \): most populations with copying evolved significantly higher mean values of \( D \) than control populations (Figure 1). The exception, which we attribute to a type II error, was the comparison between the conditional and the control scenarios at \( S_S = 25 \), where the \( p \)-value remained marginally above significance (Tukey HSD \( p = 0.05006 \)).

3.2 | If sexual selection is strong, conditional copying increases dispersal more strongly than unconditional copying

Significant differences in the mean \( D \) allelic value at generation 15,000 arose between the conditional and the unconditional copying scenarios, when \( S_S \geq 75 \) (Figure 1). Here, conditional scenarios consistently produced more dispersal.

3.3 | Associations of the \( C \) allele and high dispersal within a population remain weak

Are the above results driven by (a) \( C \) individuals (potential copiers) themselves dispersing at higher rate than individuals with the allele \( c \) or (b) the presence of \( C \) in a population elevating dispersal for everyone? At generation 15,000, there was a tendency for \( C \) individuals to disperse more than \( c \) individuals (in both scenarios where \( C \) alleles existed), but the difference was usually too small to be significant (Figure 2). Thus, any evidence for a statistical association between \( C \) and the allelic value for \( D \) remains too weak to favour the first interpretation.

3.4 | Unconditional copying only evolves if sexual selection is suitably weak; conditional copying does so across a range of strengths of sexual selection

The frequency of the \( C \) allele remained constant (did not change significantly) between generations 12,500 and 15,000 (all corrected \( p > 0.05 \); Table S3), thus the proportion of copiers had reached an equilibrium by generation 12,500. In conditional as well as unconditional scenarios, random mating \( (S_S = 0) \) should make the \( C \) allele neutral, and with sufficient time its frequency should reach 0.5. This prediction was supported: the frequency of the \( C \) allele not differ significantly from 0.5 at generation 15,000 (Figure 3). Under non-random mating \( (S_S > 0) \), both scenarios deviated significantly from the neutral expectation, but the effect depended on the strength of sexual selection. At a low value of \( S_S \) \( (S_S = 25) \), the frequency of the copier allele \( C \) was elevated more in the unconditional than in the conditional scenario and exceeded 0.5 in both cases. This conclusion reversed for \( S_S \geq 50 \), where unconditional copying led to a significantly lower frequency of the \( C \) allele than the conditional scenario.
3.5 | Mate-choice copying weakens local adaptation and trait-preference correlations

The local adaptation score $L$ showed no significant differences between scenarios in generation 7,500 (all corrected $p > 0.05$, Table S4). As expected, control scenarios without copying also yielded no significant differences in $L$ between generations 7,500 and 15,000 (all corrected $p > 0.05$; Table S5), and $L$ also did not differ between scenarios in generation 15,000 under random mating ($S = 0$, $p > 0.05$).

Differences emerged when mating was non-random, and they became significant once sexual selection was strong: when $S > 75$, scenarios with copying showed significantly less local adaptation than the control scenario (Figure 4a). Whether mate-choice copying operated conditionally or unconditionally did not significantly impact the local adaptation score.

To understand the impact of mate-choice copying on the correlation between male traits and female preferences, we calculated Pearson’s correlation coefficients for the corresponding alleles. They were not significantly different between scenarios in generation 7,500 (all corrected $p > 0.05$; Table S6), control scenarios did not show any significant difference in the correlation coefficient between generations 7,500 and 15,000 (all corrected $p > 0.05$; Table S7), and the correlation coefficient did not differ between scenarios in generation 15,000 when mating was random ($S = 0$, corrected $p > 0.05$). Introducing sexual selection ($S > 0$) increased these correlations, but the magnitude of this response was scenario-dependent: at generation 15,000, control populations showed a significantly higher correlation coefficient than either scenario with mate-choice copying. This was true across all positive values of $S$ (Figure 4b). The conditional scenario showed a significantly higher correlation coefficient than the unconditional one when sexual selection was weak ($S = 25$), but this result was reversed as $S$ increased, and the reversed finding became significant for $S = 75$ or higher (Figure 4b).

3.6 | Sexual selection increases the effect of mate-choice copying on dispersal

To investigate whether stronger sexual selection (high $S$) can amplify the effect of mate-choice copying on dispersal, we created a linear model where the response variable was the mean $D$ value at generation 15,000 and the predictors were $S$, the scenario, and their interaction. The control scenario was used as the baseline. We also tested a linear model with the explanatory variable $S$ transformed to log($S + 1$), because of the nonlinearity in Equation 4: the model with the transformation fitted the data better (AIC calculated by the `extractAIC` function from the R statistical software [Venables & Ripley, 2002], without transformation: $-4312$; with transformation: $-4393$). Diagnostic plots did not show any major deviation from model assumptions. According to this model, when log($S + 1$) = 0 (i.e. $S = 0$), mean dispersal did not differ between control populations and unconditionally copying populations (allelic values for $D$: difference estimate: $-2.40 \times 10^{-7}$; $t = -1.12$; $p = 0.233$), but both were significantly higher than dispersal in conditionally copying populations (difference estimate: $-4.24 \times 10^{-6}$; $t = -2.11$; $p = 0.035$). Given that all populations should evolve identically under random mating, and the previous results showed no significant difference between populations in this case (Figure 1), we regarded this particular result as an artefact of a model fit that is predominantly impacted by the evolutionary responses to higher values of $S$.

---

**FIGURE 4** Mate-choice copying (a) reduces local adaptation when sexual selection is high and (b) weakens trait-preference correlations. The box plots (design as in Figures 1–3) are complemented with stars that indicate statistical significance (*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$) of pairwise differences between scenarios at generation 15,000 over 20 simulations for each set of parameters (Tukey HSD test, calculated for cases where the ANOVA test with the Bonferroni correction was significant, see details in Section 2). Parameters as in Figure 1.
The linear model shows $S_S$ to increase dispersal in all scenarios. An increase of one log($S_S + 1$) unit creates a significant increase of $1.96 \times 10^{-4}$ in the mean allelic value of D in control populations lacking copying ($t = 5.02; p < 0.001$), with an additional significant increase of $3.14 \times 10^{-4}$ in the unconditional copying scenario ($t = 5.67; p < 0.001$) and an additional significant increase of $5.38 \times 10^{-5}$ in the conditional copying scenario ($t = 9.76; p < 0.001$). Overall, the dispersal-enhancing effect of sexual selection is significantly strengthened by mate-choice copying, and it reaches its maximum effect if copying is conditional on dispersal.

4 | DISCUSSION

Results show that mate-choice copying increases dispersal, especially when females are choosier (sexual selection is stronger). Dispersal, in turn, has knock-on effects on gene flow and local adaptation. Gene flow becomes stronger, and local adaptation becomes weaker, when immigrant females can compensate for lack of local knowledge by copying the choices of others. Mate-choice copying also decreased the correlation between the male trait and the female preference alleles. These phenomena, taken together, counteract the depletion of male trait genetic variation by persistent innate female preferences.

The evolution of mate-choice copying can be hampered when innate female preferences coevolve with the male trait (Santos et al., 2017; Servedio & Kirkpatrick, 1996), raising the question of how sufficient variation in preferences can be maintained for a female to have any reason to switch from her innate to an observed preference (in the absence of variation, a switch should never be worthwhile). In their population-genetic models, Servedio and Kirkpatrick (1996) maintained the polymorphism in preferences by mutation, while Santos et al. (2017) did so by randomly distributing the preference locus each generation. In our case, environmental heterogeneity was the responsible factor. Our finding that mate-choice copying can spread, but that it can also make local populations less well adapted, is a novel twist on the idea that local adaptation of both traits and preferences, combined with dispersal, can maintain a persistent supply of locally acting selection gradients (thus helping to make the so-called lek paradox disappear; Holman & Kokko, 2014).

Sexual selection can be switched off in our model by assuming random mating. Dispersal evolves to be low in such a setting. Complete philopatry still does not evolve as it would make individuals compete for breeding resources (females) or matings (males) with close kin (Li & Kokko, 2019), but dispersal mortality, as well as the risk of being maladapted to faraway parts of the landscape, keeps dispersal rates low. Sexual selection can make this latter cost smaller, but this only works if dispersing females subsequently mate with locally adapted males (and not, e.g., with males who are themselves maladapted immigrants). Dispersal readily responds to this cost reduction, and evolves to be higher, when the option of mate-choice copying is available. Copying allows choosy females to express a different phenotype than her P locus would dictate. An immigrant female’s offspring will therefore, all else being equal, be better locally adapted than in the absence of mate-choice copying. However, we also show that all else is not equal: when the evolutionary process with mate-choice copying involves higher dispersal, this also improves gene flow and reduces local adaptation in the global population as a whole.

Dispersal was particularly enhanced if mate-choice copying was conditionally expressed by females who had dispersed, confirming the intuitive prediction that copying is best performed in unfamiliar surroundings. For philopatric females, copying may lead to counterproductive updating of preferences to those of immigrants, whose preferences have been shaped by selection elsewhere; expressing one’s own innate preferences may then be favoured. Conditional copying, then, appears the superior way to alleviate the local adaptation cost of dispersal; it accordingly produces a more robust feedback between copying and dispersal across different strengths of sexual selection, compared with unconditional copying. This result is in line with earlier findings that indiscriminate copying can be maladaptive (Dubois et al., 2012; see also Kendal et al., 2018). We are unaware of empirical studies documenting differences in copying behaviour between immigrants and philopatric females; such plasticity appears worth looking for.

Our results do not rely on assortative mating or other processes (e.g. physical linkage) creating a statistical association between the C allele (copying) and the D allele (dispersal). While, in principle, only high-D individuals ‘need’ the C allele (assuming philopatric individuals should not copy), such associations remain weak at the genetic level. This does not strictly exclude any possibility for linkage disequilibria: an immigrant non-copier would presumably mate non-ideally and thus high-D combined with c would be selected against at the stage when her offspring viabilities are tested. But as her choices are blind with respect to the dispersal status or (unexpressed) C locus of the sire, any effects here remain weak. Instead, it appears that a high frequency of C in a population facilitates high D mainly via (a) an overall expectation that preference updating is frequently an option for individuals in a lineage (all female offspring who have inherited C from the mother or from the father can do so) and (b) a reduction of the overall importance of local adaptation. The latter is a general finding in the literature on local adaptation and dispersal: once dispersal is relatively frequent, the costs of settling in a new location are reduced, because recurrent gene flow prevents strong local adaptation in the first place (Berdahl et al., 2015; Blanquart & Gandon, 2011). Our contribution is to show that mate-choice copying can amplify this effect quite substantially.

Note that the C allele in our model only experienced indirect selection, via its effects on the identity of a female’s chosen mate. Earlier population genetic work has investigated both direct and indirect selection on copying (Santos et al., 2017; Servedio & Kirkpatrick, 1996). Direct selection, which we ignore, could either favour or disfavour copying: copying might help reduce the time and energy costs of assessing male quality, or there may be mild costs inherent in the copying process itself (e.g. costs of learning) or
pleiotropic effects (Servedio & Kirkpatrick, 1996). While our chosen focus on the coevolution of copying and dispersal (via local adaptation) made us leave direct costs outside our study, it is intuitively clear that copying may reach higher frequencies than reported by us if it also helps to avoid direct costs.

Our contrasts between different scenarios were run in parallel, without the conditional strategy directly competing with the unconditional one. Even so, it is still useful to reflect whether the differences in the evolved frequencies of the C allele are a measure of ‘success’ of a certain kind of copying strategy. Usually, high frequencies of C associate with conditional expression, but an exception is found at weak sexual selection ($S_2 = 25$), potentially explicable as a result of the quite variable realized matings at low value of $S_2$ (Box 1). Matings are in this case impacted by stochasticity, irrespective of whether the choices follow innate preferences or updated (copied) ones; the net effect is that non-copier females may lose to copier females (in terms of mate quality). Given that females with a C allele are much more likely to be copiers in the unconditional scenario than in the conditional one, the situation described by Box 1 arises more often in the unconditional scenario. This provides new insight into the debate (Giraldeau et al., 2002; Kendal et al., 2018; Nordell & Valone, 1998; Vakirtzis, 2011; Varela et al., 2018; Witte et al., 2015) regarding the conditions under which mate-choice copying can be adaptive.

Note that although the model has many parameters, it simultaneously presents a simplified view of dispersal and mate-choice copying in a heterogeneous environment. For example, we modelled dispersal as a simple exponential kernel, ignoring e.g. sex-biased dispersal (Li & Kokko, 2019), dispersal kernels with flexibilities offered by more than one variable (Bonte et al., 2010; Chapman et al., 2007; Nathan et al., 2012; Poethke et al., 2011; Tung et al., 2017) or any decision-making during dispersal (which can have a strong impact on the kernel in settings with two sexes; Shaw & Kokko, 2014). Also, to avoid having to specify effects of genetic dominance on several traits, we also chose to investigate haploidy. Dominance relationships between two alleles at a diploid locus can take very many forms when fitness effects also depend on spatial location; we ignored these real-life complications to focus on a minimal genetic set-up that permits all the intended feedbacks between mate-choice copying, dispersal and local adaptation to occur. This obviously leaves avenues for further study.

More generally, mate-choice copying is just one example of social learning, which happens to be expressed in a sexual selection context (Kendal et al., 2018; Verzijden et al., 2012). If there is a need to locally adapt to new conditions after dispersal, and social learning offers a way to acquire a new set of behaviours (Varela et al., 2020), then our model’s conclusions should generalize to situations where behaviours other than mate choice are learned. The prediction that social learning enhances dispersal appears quite generally applicable—as does its flipside, i.e. local adaptation may become less strong in the long term due to homogenization of traits over the spatial range of a species when learning and dispersal together promote strong gene flow (Varela et al., 2018). This finding, should it generalize, could be of importance in cultural evolution as a whole.

ACKNOWLEDGEMENTS

We thank Margarida Matos and Mauro Santos for helpful insights. This work was financed by Portuguese National Funds through Fundação para a Ciência e a Tecnologia (FCT), Portugal, within the cE3c Unit funding UIDB/00329/2020, MS PhD grant (ref. PD/BD/128349/2017) and SAMV research grant (ref. PTDC/BIA-COM/31887/2017). H.K. was financed by the Swiss National Science Foundation. The authors have no conflict of interest to declare. We also thank the associate editor and two anonymous reviewers for their very helpful comments.

AUTHORS’ CONTRIBUTIONS

All authors helped to conceive the problem and contributed to the design of the study. M.S. wrote the code to the mathematical model, ran the simulations, analysed the data and wrote the first draft of the manuscript. All authors contributed substantially to the writing.
DATA AVAILABILITY STATEMENT
Data deposited in the Dryad Digital Repository https://doi.org/10.5061/dryad.dncjsxkz1 (Sapage et al., 2020).

ORCID
Manuel Sapage https://orcid.org/0000-0002-3007-6482
Susana A. M. Varela https://orcid.org/0000-0002-9947-4024

REFERENCES
Davies, A. D., Lewis, Z., & Douherty, L. R. (2020). A meta-analysis of factors influencing the strength of mate-choice copying in animals.


**SUPPORTING INFORMATION**

Additional supporting information may be found online in the Supporting Information section.

**How to cite this article:** Sapage M, Varela SAM, Kokko H. Social learning by mate-choice copying increases dispersal and reduces local adaptation. Funct Ecol. 2021;35:705–716. [https://doi.org/10.1111/1365-2435.13735](https://doi.org/10.1111/1365-2435.13735)