Sexual antagonism, whereby mutations are favourable in one sex and disadvantageous in the other, is common in natural populations, yet the root causes of sexual antagonism are rarely considered in evolutionary theories of adaptation. Here, we explore the evolutionary consequences of sex-differential selection and genotype-by-sex interactions for adaptation in species with separate sexes. We show that sexual antagonism emerges naturally from sex differences in the direction of selection on phenotypes expressed by both sexes or from sex-by-genotype interactions affecting the expression of such phenotypes. Moreover, modest sex differences in selection or genotype-by-sex effects profoundly influence the long-term evolutionary trajectories of populations with separate sexes, as these conditions trigger the evolution of strong sexual antagonism as a by-product of adaptively driven evolutionary change. The theory demonstrates that sexual antagonism is an inescapable by-product of adaptation in species with separate sexes, whether or not selection favours evolutionary divergence between males and females.

1. Introduction

Species with separate sexes face two important evolutionary challenges that can limit their abilities to adapt to a changing environment. First, discrete sexes are common among complex organisms, in which pleiotropic effects can constrain adaptation [1–3]. Second, patterns of selection and the phenotypic effects of mutations each differ between the sexes [4–6], which can decouple the genetic basis of male versus female fitness. When male and female fitness is decoupled, mutations benefitting one sex will sometimes be deleterious to the other (i.e. their fitness effects are ‘sexually antagonistic’ [7]), which diminishes a population’s ability to respond to selection through fixation of beneficial mutations.

An emerging body of data implies that sexual antagonism is an important feature of animal and plant populations [8–11]. Nevertheless, several fundamental questions regarding sexual antagonism remain unaddressed. For example, how often does sexual antagonism arise owing to sex differences in the direction of selection? How severely does sexual antagonism limit the rate of adaptation in each sex, and how do such constraints evolve over time? How does environmental change mediate opportunities for sexual antagonism, including differences between poorly versus well-adapted populations?

These questions can be addressed with theoretical models that effectively bridge the divide between empirically measurable properties of mutation and selection, and the evolutionary genetic patterns and processes that we seek to understand. Quantitative genetics theory has proved extremely useful for quantifying short-term constraints to sex-specific adaptive evolution [12], yet we currently lack clear theoretical predictions about the underlying population genetics of sexual antagonism, including the fraction of mutations that has sexually antagonistic fitness effects, and the role of sexual antagonism in shaping the long-term evolutionary trajectories of males and females in adaptively evolving populations. With these issues in mind, we developed a two-sex extension of Fisher’s geometric model [13], and used it to characterize the sex-specific distribution of mutant fitness effects, and the population genetic dynamics of adaptation in species with separate sexes.
2. Fisher’s geometric model in species with separate sexes

Fisher’s original model provides a simplified mapping between genotype, phenotype and fitness that captures the basic biological details of adaptation within a complex organism: (i) mutations cause random changes within multi-dimensional trait space; (ii) selection favours those mutations that move the system of traits closer to an optimal phenotypic value that maximizes fitness; and (iii) a population approaches its fitness optimum by successively fixing beneficial mutations \([14,15]\).

Although alternative modes of evolutionary change—e.g. polygenic adaptation involving modest allele frequency shifts at many loci—may also contribute to adaptation over short evolutionary time-scales, such short-term responses may be weak when pleiotropic constraints—including sexual antagonism—are severe \([16–18]\). In such cases, adaptation is likely to be dominated by the fixation of beneficial mutations, provided directional selection is sustained over time \([19,20]\). Fisher’s model also makes simplifying assumptions about the mutational architecture of traits, yet its original assumptions can be relaxed without greatly altering the model’s basic predictions \([21–23]\) and these predictions are adept at explaining a variety of interesting empirical patterns that emerge from the study of real organisms \((e.g. \ [24–28])\).

In extending Fisher’s model to a two-sex system, we characterize the evolution of \(n\) traits that are each expressed, but not necessarily equally, in males and females \(\text{(see the electronic supplementary material)}\). Sets of trait values are depicted using Cartesian coordinates, with each set of coordinates representing an individual’s genetically determined position in \(n\)-dimensional phenotypic space. Following prior work, we assume that fitness depends on the Euclidean distance of each sex to its optimum, and mutational effects are unbiased in direction \(\text{(uniformly oriented in}\ n\text{-dimensional space} \ [13,29,30] )\). These assumptions are not restrictive as long as we interpret \(n\) as the ‘effective complexity’ of a species \(\text{(i.e. the effective number of traits can be less than the actual number} \ [21–23] \text{)}\).

To simplify the presentation, we model mutation and evolution in a haploid population; our results also apply to diploids, with the caveat that mutations contributing to adaptation in diploids will sometimes involve a transient, balanced polymorphic state \([30]\). Finally, although we focus on evolution in dioecious \(\text{(gonochoristic)}\) species with distinct sexes, sexual antagonism may also manifest within hermaphrodites, by way of allocation trade-offs between male and female reproductive structures in simultaneous hermaphrodites, or antagonistic pleiotropy between male and female stages in sequential hermaphrodites \([31]\). Our results may imperfectly characterize adaptation in hermaphrodites, yet our underlying mathematical framework is sufficiently flexible to permit future theoretical extensions to mating systems that we do not specifically consider here.

Patterns of selection depend on the distance and orientation of each sex to its phenotypic optimum. Let \(z_m\) and \(z_f\) represent the Euclidean distance between the optimal phenotype and the actual phenotype expressed by each sex \(\text{(subscripts,} \ m\text{ and} \ f\text{, hereafter refer to} \text{‘male’ and} \text{‘female’)}\). Fitness functions for each sex are Gaussian, with \(w_m = \exp(-\omega_m z_m^2)\) and \(w_f = \exp(-\omega_f z_f^2)\), and \(\omega_m\) and \(\omega_f\) are positive constants specifying the strength of the fitness decline with increased distance from the optimum. The relative orientation of each sex to its optimum—\(\text{the direction of selection—can be described by a pair of vectors that extend from the current locations of each sex within phenotypic space to the location of its optimum (figure 1). The strength of selection is defined as} \ \beta_m = \frac{\partial \ln(w_m)}{\partial z_m} = 2\omega_m z_m\) and \(\beta_f = \frac{\partial \ln(w_f)}{\partial z_f} = 2\omega_f z_f\), which parallel the definition of the selection gradient from quantitative genetics \(\text{(see} \ [32, \ pp. 122–123])\). \(\theta_{\text{mut}}\) represents the angle between male and female selection vectors, with \(\rho_{\text{mut}} = \cos(\theta_{\text{mut}})\) representing the correlation between male and female orientations of directional selection \((-1 < \cos(\theta_{\text{mut}}) < 1)\), with \(\rho_{\text{mut}} = 1\) representing identical directions of selection; \(\text{see} \ [33]\).

Phenotypic effects of mutations are similarly described using paired vectors, with each mutation having a unique magnitude \((r_m\) and \(r_f)\) and orientation within each sex \(\text{(figure 1)}\). For each mutation, \(\theta_{\text{mut}}\) is a random variable, which represents the angle between male and female orientations of phenotypic change caused by the mutation. For mutations with specified magnitude \(r_m\) and \(r_f\), the phenotypic correlation between the sexes for each trait axis is given by \(\rho_{\text{mut}}\). In the two-sex extension of Fisher’s model, there is a geometric relationship between \(\rho_{\text{mut}}\) and the angle between mutation vectors: \(\rho_{\text{mut}} = \cos(\theta_{\text{mut}})\), \text{where the angle brackets denote the mean among random mutations (see the electronic supplementary material)}\).

Unique mutations arise randomly within a population that is initially fixed at each locus for a resident \(\text{(wild-type)}\) allele. For a given mutation, its sex-specific fitness effect

\textbf{Figure 1.} Fisher’s geometric model in two dimensions and with two sexes. Trait values are represented by the \(x\)- and \(y\)-axes, with each of the traits expressed by both sexes. \(A_f\) and \(A_m\) depict the trait values expressed by females and males in the population, and \(O_f\) and \(O_m\) are the female and male trait optimum. Black arrows represent the sex-specific selection vectors that reflect the distance of each sex to its optimum \(\text{(vector lengths,} \ z_f\text{ and} \ z_m)\). Sex differences in directional selection are described by the angle between selection vectors \(\theta_{\text{mut}}\). The fitness effects of each mutation depend on its phenotypic effect in each sex: its size \(\text{(vector length,} \ r_m\text{ and} \ r_f)\) and orientation in phenotypic space. Sex differences in a mutation’s orientation are described by the angle between mutation vectors \(\theta_{\text{mut}}\). Mutations that bring a given sex closer to its optimum \(\text{(i.e. within the relevant circle)}\) are beneficial to that sex, whereas those causing movement away from the optimum will be deleterious \(\text{(i.e. outside the relevant circle)}\). Two mutations are shown: one that is beneficial to both sexes \(\text{(blue)},\) and the other that is sexually antagonistic \(\text{(red)}\).
in the jth sex is $s_j = (w_j(z_j)/w_j(z_{j+})) - 1$, where $z_{j+}$ and $z_j$ represent distances to the optimum for wild-type and mutant individuals, respectively. Letting $t_j = \ln(1 + s_j)$ represent the logarithm of relative fitness, $f(t_{m}, t_{f}, r_{m}, r_{f})$ is the bivariate distribution of $t_{m}$ and $t_{f}$ among mutations with arbitrary magnitude, $r_{m}$ and $r_{f}$ ($r_{m} \geq r_{f} > 0$). In many dimensions ($n > 1$), this distribution converges to a bivariate normal distribution, with marginal mean and variance ($t_j = -aw_j^2$ and $\text{var}(t_j) = 2\omega w_j^2/n$, respectively, and between-sex correlation coefficient of $\rho_{w} = \cos(\theta_{w})/\sin(\theta_{w}) = \rho_{w}\rho_{b}$ (see the electronic supplementary material), noting that $s_{w}$ will also be bivariate normal in the limit: $s_{w} \rightarrow 0$. This simple identity for $\rho_{w}$ implies that genotype-by-sex interactions and sex-differential selection contribute multiplicatively to the fitness effect correlation between the sexes. Moreover, either factor is sufficient to generate an imperfect genetic correlation between male and female fitness.

Adaptive evolutionary change involves the sequential fixation of mutations with sex-averaged beneficial fitness effects ($s_{\text{avg}} = s_{m} + s_{f}/2 > 0$). Similar to previous work in Fisher’s genetic model (e.g. [29]), we characterize evolutionary change in a large population where beneficial mutations are rare, selection coefficients of beneficial alleles are small ($s_{\text{m}}, s_{f} \ll 1$; see [34]), and therefore each mutation’s probability of fixation depends only on its net fitness effect. Provided the sex-averaged fitness effect is positive ($s_{\text{avg}} > 0$), a mutation’s probability of fixation is $1 - \exp(-2s_{\text{avg}})$, and otherwise its fixation probability is zero. Mutations that are successfully fixed trigger a change in the location of males and females within phenotypic space. Following each substitution, new distances and relative orientations of males and females to their optima may be calculated (i.e. new values of $s_{\text{m}}, z_{m}, \theta_{m}$), and the next step of adaptation can proceed from this new population state (see the electronic supplementary material).

Because substitutions in Fisher’s model are discrete (see above and [29]), each time point during evolution is associated with specific values for $s_{\text{m}}, z_{m}$ and $\theta_{m}$, and these terms may therefore be treated as population parameters appropriate for a given time point. We assume throughout that the phenotypic distributions of mutations remain constant during evolution (distributions of $t_{m}, t_{f}$, and $\theta_{out}$ remain constant). Therefore, $\rho_{nout}$ can be treated as a population parameter, albeit one that can evolve over the course of an adaptive walk. Accordingly, we first consider the sex-specific fitness effects of new mutations and the rate of adaptation, for an arbitrary population with parameters $s_{\text{m}}, z_{m}$ and $\theta_{m}$. We then characterize the long-term evolution of such populations as they adapt, with each sex approaching its fitness optimum.

3. Results and discussion

Adaptation requires beneficial genetic variation, which serves as the fuel for evolutionary change. In populations with separate sexes, random mutations can be deleterious to both sexes, beneficial to both, or beneficial to one sex and deleterious to the other (‘sexually antagonistic’). Among mutations that are beneficial to at least one sex ($s_{\text{m}} > 0$ and/or $s_{f} > 0$), the fraction that is sexually antagonistic is:

$$f_{SA} = \frac{\text{Pr}(s_{m}/s_{f} < 0)}{1 - \text{Pr}(s_{m} < 0; s_{f} < 0)}.$$  

which, in the limit of small mutation size, reduces to a simple function of the fitness effect correlation ($\rho_{nout}$):

$$\lim_{n_{\text{out}} \rightarrow 0} f_{SA} = \frac{2\pi - 4\sin^{-1} (\rho_{nout})}{3\pi - 2\sin^{-1} (\rho_{nout})}.$$  

Equation (3.1b) approximates the minimum proportion of beneficial mutations that are sexually antagonistic (see the electronic supplementary material). Numerical evaluation of equation (3.1), along with exact computer simulations across a distribution of $r_{m}$ and $r_{f}$ values, shows that $f_{SA}$ increases with dimensionality and mutation size—factors that also increase the ‘scaled size’ of mutations in Fisher’s geometric model (i.e. the scaled mutation size in sex $j$ is defined as $s_{j} = r_{j}/(2\sigma_{j})$; see [13,29]). Overall, sexual antagonism is pervasive, and is a dominant feature of beneficial mutations, across the entire range of $\rho_{nout}$ (figure 2).

Adaptation represents a compromise between fitness benefits and costs to males and females, with selection favouring mutations with positive, sex-averaged fitness effects (i.e. $s_{\text{avg}} = s_{m} + s_{f}/2 > 0$). Rates of adaptation in each sex can be approximated by taking into account both the mutation rate to positively selected alleles and the fixation probabilities of individual mutations [1]. To quantify the effect of an imperfect fitness correlation on the rate of adaptation, constraint is expressed as the ratio of the expected rate of adaptation in a population with imperfectly correlated male and female fitness effects ($\rho_{nout} < 1$) to the rate of adaptation in an ideal population with perfectly correlated fitness effects ($\rho_{nout} = 1$). We formally define this constraint as $C_{j} = (dz_{j}/dt | \rho_{nout} < 1)/(dz_{j}/dt | \rho_{nout} = 1)$, where $dz_{j}/dt$ is the rate of evolutionary change of sex $j$ towards its optimum (see the electronic supplementary material; $C_{j} = 1$ in the absence of sexually antagonistic constraints, and otherwise, $C_{j} < 1$). In the limit of small mutation size ($r_{m}, r_{f} \rightarrow 0$), these ratios in males and females are approximately:

$$C_{m} = 1 - (1 - \rho_{nout})\frac{r_{f}\beta_{f}}{r_{f}\beta_{f} + r_{m}\beta_{m}}$$  

and  

$$C_{f} = 1 - (1 - \rho_{nout})\frac{r_{m}\beta_{m}}{r_{m}\beta_{m} + r_{f}\beta_{f}},$$  

which compares well with exact computer simulations of small to moderately sized mutations, and it underestimates the true magnitude of constraint, particularly when scaled mutation sizes are large (figure 2, where $s_{f} = r_{f}/(2\sigma_{f})$ is the scaled size). An imperfect fitness correlation between the sexes significantly decreases the rate of adaptation ($C_{m}$ and $C_{f}$ can be substantially lower than one), with the relative magnitude of constraint dependent upon mutation and selection asymmetries between the sexes. With equally strong directional selection in each sex and similar-sized mutations ($r_{m}\beta_{m} = r_{f}\beta_{f}$), males and females experience equal degrees of constraint: $C_{m} = C_{f} = 1 - (1 - \rho_{nout})/2$. Asymmetries between the sexes ($r_{m}\beta_{m} \neq r_{f}\beta_{f}$) shift the burden of constraint to the sex that experiences weaker selection or mutation (figure 2).

The results above describe features of selection and adaptation at single time points during evolution, and consequently they treat fitness effect distributions and sex-specific rates of adaptation and constraint as static rather than evolutionarily labile population properties. However, each mutation that is fixed during adaptation will carry the population to a new location in phenotypic space and will alter the distance and orientation of each sex to its optimum. Fitness effect distributions and subsequent opportunities
for adaptation will therefore change following each genetic substitution. To characterize such changes during the course of adaptive evolution, we analysed adaptive walks of populations towards stationary male and female fitness optima (see the electronic supplementary material; figure 3). We focused on the strength and orientation of sex-specific directional selection ($b_m$, $b_f$) and sex-specific mutation and selection parametrizations, two factors, as reflected in the identity: $r_m/r_f = o_m/o_f$. Additional details of the simulation procedure can be found in the electronic supplementary material.

Figure 2. The fitness effect correlation between sexes ($p_{mut}$) predicts the fraction of sexually antagonistic mutations, and the severity of constraints to sex-specific adaptation. The minimum fraction of beneficial mutations that are sexually antagonistic ($f_{SA}$) is based on the approximation in equation (3.1b) (solid curve), and from exact simulation results, using three different mutation size distributions (triangles). Shown are results using a bivariate exponential distribution for $r_m$ and $r_f$ with $corr(r_m, r_f) = 0.9$ and equal marginal distributions for $r_m$ and $r_f$ [35]. The mean scaled mutation size is $\langle x_i \rangle = \langle x_i \rangle/\langle n_i \rangle/\langle 2z_i \rangle$, with $n = 25$ dimensions, and $z_i = 1/2; j = \{m, f\}$. Each simulation result is based on 10 000 randomly generated beneficial mutations (e.g. beneficial to at least one sex). The relative reduction in the rate of adaptation in females ($C_f$; (b)) is based on the small mutation approximation in equation (3.2) (solid lines). Exact simulation results are depicted by triangles, based on the parameters: $r_m = r_f = 0.05$, $z_m = z_f = 1/2$, $n = 25$ ($k_f = 0.25$) and $\beta_m/\beta_f = o_m/o_f$. Additional details of the simulation procedure can be found in the electronic supplementary material.

Our model yields two general insights into the evolution of species with separate sexes. First, conditions for sexual antagonism among beneficial mutations are extremely permissive and readily emerge when directional selection or the phenotypic effects of mutations differ between the sexes. Indeed, the correlation of mutant fitness effects is equally sensitive to both factors, as reflected in the identity: $P_{mut} = P_{mut} \cos(\theta_{male}) \cos(\theta_{female})$. Second, decoupling the male and female fitness effects of random mutations significantly reduces the rate of adaptation of both sexes, with sexual antagonism and the magnitude of adaptive constraints increasing during the process of adaptation. Evolution of sexual antagonism occurs whether or not selection favours phenotypic divergence between the sexes (figure 3 and electronic supplementary material, figure S2). Opposing selection between the sexes—for homologous traits expressed by both sexes and mutations affecting these traits—is therefore an inescapable by-product of adaptation in populations with separate sexes.

Given that $p_{mut}$ eventually evolves to be negative, sexual antagonism is expected to ultimately reduce the subsequent rate of adaptation in each sex by at least twofold ($C_m, C_f < 1/2$), with a strong majority of beneficial mutations having sexually antagonistic fitness effects ($f_{SA} > 2/3$).
between the sexes. Estimates of selection on single traits suggest that opposing directional selection is common, and often of similar magnitude between the sexes [5,38]. Of the few studies estimating the angle between male and female directional selection on multiple traits, two support the general prediction that opposing selection readily evolves (ρ_{sel}/C^2 = 0.61 and 0.73 in a moth and fly population, respectively [39,40]), whereas a third reveals a positive correlation between male and female directional selection, albeit in a modern human population (ρ_{sel}/C^2 = 0.22 [41]). To systematically test whether opposing selection between the sexes evolves during adaptation, future studies could contrast patterns of sex-specific directional selection in poorly versus well-adapted populations, with the latter expected to exhibit more strongly opposing selection between males and females (for a similar prediction, see [42]).

We have identified general conditions yielding sexually antagonistic selection, and although such antagonism does not require that selection favour evolutionary divergence between the sexes, the pervasiveness and intensity of...
antagonism is expected to increase when selection does favour sexual dimorphism. Discordance between the fitness landscapes of males and females—the root cause of selection for dimorphism—emerges from the unique interactions between each sex and its environment [8–11], which can lead to sex-differential opportunities for niche partitioning, reproductive investment, competition for mates or their susceptibilities to predation and parasitism. We emphasize the population genetic consequences of distinct male and female fitness landscapes, while placing no special emphasis on the specific ecological causes for divergent male and female optima. The ubiquity of sexually dimorphic phenotypes [43], including conspicuous differences between closely related species [44], attests to the biological fact of fitness landscape dimorphism and implies a central role for antagonistic selection during the process of adaptation within lineages and speciation between them.

Acknowledgements. We thank Deborah Charlesworth, Angela Early, Ted Morrow, Chris Stieha, Akane Uesugi, Rob Unckless, Cris van Hout, and members of the Clark laboratory for valuable discussion and comments on an early version of the manuscript, Allen Orr for valuable correspondence during the initial stages of the project, and two anonymous referees for thoughtful comments and suggestions that substantially improved the paper.

Funding statement. This work was funded by NIH grant GM64590 to A.G.C. and A. B. Carvalho.

References


