bioRxiv preprint doi: https://doi.org/10.1101/526970; this version posted March 13, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Adaptation and the Parliament of Genes

Thomas W. Scott[1]* & Stuart A. West[1]

[1]Department of Zoology, University of Oxford, Oxford, United Kingdom

*Corresponding author: E-mail: thomas.scott@zoo.ox.ac.uk

Abstract

Fields such as behavioural and evolutionary ecology are built on the assumption that natural selection leads to organisms that behave as if they are trying to maximise their fitness. However, there is considerable evidence for selfish genetic elements that change the behaviour of individuals to increase their own transmission. How can we reconcile these contradictions? We found theoretically that, when selfish genetic elements have a greater impact at the individual level, they are more likely to be suppressed, and suppression spreads more quickly. Consequently, selfish genetic elements will either have a minor impact at the individual level, or tend to be suppressed. In addition, we found that selection on selfish genetic elements favours higher levels of distortion. Consequently, selfish genetic elements will tend to evolve to make themselves more likely to be suppressed. Overall, our results suggest that even when there is the potential for considerable genetic conflict, this will often have negligible impact at the individual level.

bioRxiv preprint doi: https://doi.org/10.1101/526970; this version posted March 13, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Introduction

There is a clear contradiction between major branches of modern evolutionary biology. On the one hand, fields such as behavioural and evolutionary ecology are based on the assumption that organisms will behave as if they are trying to maximise their fitness[1-4]. Models based on fitness maximisation are used to make predictions about the selective forces (reasons) for adaptation, and these are then tested empirically[5,6]. This approach has been phenomenally successful, explaining many aspects of behaviour, life history and morphology. For example, research on: foraging, resource competition, sexual selection, parental care, sex allocation, signalling and cooperation[7-12].

However, on the other hand, there is considerable evidence for selfish genetic elements, which increase their own contribution to future generations at the expense of other genes in the same organism[13-16]. These 'selfish genes' distort traits away from the values that would maximise individual fitness, to increase their own transmission[14,17-21]. Evidence for such genetic conflict has been found across the tree of life, from simple prokaryotes to complex animals. The contradiction is that selfish genetic elements mess up individual fitness maximisation, and appear to be common. But individual fitness maximisation still appears to occur[22,23]. For example, research on sex allocation has provided phenomenal support for both the individual fitness maximisation approach, and the existence of selfish sex ratio distorters[9,14,24-26].How can we resolve this contradiction?

bioRxiv preprint doi: https://doi.org/10.1101/526970; this version posted March 13, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Leigh[27] provided a potential solution to this contradiction by suggesting that selfish genetic elements would be suppressed by the 'parliament of genes'. Leigh's argument was that, because selfish genetic elements reduce the fitness of the other genes in the organism, the rest of the genome will have a united interest in suppressing selfish genetic elements. Furthermore, because those other genes are far more numerous, they will be likely to win the conflict. Consequently, even when there is considerable potential for conflict within individuals, we would still expect fitness maximisation at the individual level[28-32]. Leigh[27] demonstrated the plausibility of his argument by showing theoretically how a suppressor of a meiotic drive gene could be favoured.

However, there are a number of potential problems with the parliament of genes hypothesis. First, even if a parliament of genes could easily generate suppressors, whether a suppressor spreads can depend upon biological details such as any cost associated with the suppressor, the extent to which a selfish genetic element is distorting a trait, and the population frequency of that selfish genetic element[14,33-38]. Second, even if suppressors can spread, prolonged non-equilibrium trait distortion is possible if the spread of suppressors through populations is slow. Third, selfish genetic elements are themselves also under evolutionary pressure to reach a level of distortion that would maximise their transmission to the next generation – how will this influence the likelihood that they are suppressed?[31] Fourth, segregation at suppressor loci might expose previously suppressed selfish genetic elements[34].

bioRxiv preprint doi: https://doi.org/10.1101/526970; this version posted March 13, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

We investigated the parliament of genes hypothesis theoretically. Our aim was to investigate the extent to which genetic conflict distorts traits away from the value that would maximise individual fitness. We found that: (i) the greater the level of distortion caused by a selfish genetic element, the more likely and the quicker it will be suppressed; (ii) selection on selfish genetic elements leads towards greater distortion, making them more likely to be suppressed. We found the same patterns with an illustrative model, and when examining three specific examples: selfish distortion of the sex ratio by an X chromosome driver; an altruistic helping behaviour encoded by an imprinted gene; and production of a cooperative public good encoded on a horizontally transmitted bacterial plasmid. Furthermore, we found close agreement when analysing scenarios with population genetic analyses and individual based simulations. Our results suggest that even when there is potential for considerable genetic conflict, it will have relatively little impact on traits at the individual level.

Results

A selfish genetic element may be able to gain a propagation advantage through trait distortion ('*distorter*'). Any part of the genome that does not gain the propagation advantage from the trait distortion will be selected to suppress the distorter. This collection of genes will comprise most of the genome, and so will constitute the majority within the parliament of genes[20]. We take into account the large size of this collection of genes by assuming that it is highly likely that a potential suppressor of a distorter can arise by mutation (high mutational accessibility). Consequently, we focus our analyses on when a distorter and its suppressor can spread. Prolonged

bioRxiv preprint doi: https://doi.org/10.1101/526970; this version posted March 13, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

trait distortion may be possible if a suppressor is hard to evolve, but our assumption of suppressor availability allows us to focus on the evolutionary dynamics that succeed the acquisition of a suppressor by mutation[31].

Our overall aim is to assess, given the potential for suppression, the extent to which a distorter can distort the organism trait away from the optimum for individuals (the largest "coreplicon"[21]). In order to elucidate the selective forces in operation, we ask four questions in a step-wise manner, with increasing complexity:

(1) In the absence of a suppressor, when can a distorter invade?

(2) When can a costly suppressor of the distorter invade?

(3) What are the overall consequences of distorter-suppressor dynamics for trait values, at the individual and population level, at evolutionary equilibrium and before equilibrium has been reached?

(4) If the extent to which the distorter manipulates the organism trait can evolve, how will this influence the likelihood that it is suppressed, and hence the individual and population trait values?

Illustrative Model

We assume an arbitrary trait that influences organism fitness. In the absence of distorters, all individuals have the trait value that maximises their individual fitness. The distorter manipulates the trait away from the individual optimum, to increase their own transmission to offspring. We assume a large population of diploid, randomly mating individuals. The aim of this model is to establish key aspects of the population genetics governing distorters and their suppressors, in an abstract

bioRxiv preprint doi: https://doi.org/10.1101/526970; this version posted March 13, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

setting. We will subsequently address analogous issues in three specific biological scenarios.

(1) Spread of a Distorter

We consider a distorter, which we denote by y_1, that is dominant and distorts an organism trait value by some positive amount, denoted by k (k>0). This distortion increases the transmission of the distorter to offspring. Specifically, the distorter (y_1) drives at meiosis, in heterozygotes, against a non-distorter (y_0), being passed into the proportion (1+$t(k)$)/2 of offspring. $t(k)$ denotes the transmission bias (0≤t(k)≤1) and is a monotonically increasing function of trait distortion $\left(\frac{dt}{dk} \geq 0\right)$.

Trait distortion leads to a fitness (viability) cost ($c_{drive}(k)$) at the individual level, reducing an individual's number of offspring from 1 to 1-$c_{drive}(k)$ (0≤$c_{drive}(k)$≤1). Owing to distorter dominance, the fitness cost of trait distortion is borne by heterozygous as well as distorter-homozygous individuals. The fitness cost is a monotonically increasing function of trait distortion $\left(\frac{dc_{drive}}{dk} \geq 0\right)$. We assume that $t(k)$ and $c_{drive}(k)$ do not change with population allele frequencies, but relax this assumption in our specific models.

We first ask what frequency the distorter will reach in the population in the absence of suppression. If we take p and p' as the population frequency of the distorter in two consecutive generations, then the population frequency of the distorter in the latter generation is:

bioRxiv preprint doi: https://doi.org/10.1101/526970; this version posted March 13, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

$$\bar{w}\, p' = (1 - c_{drive}(k))\,(p^2 + (1-p)\,p\,(t(k)+1)), \quad (1)$$

where $\bar{w}$ is the average fitness of individuals in the population in the current generation, and can be written in full as: $\bar{w} = (1 - c_{drive}(k))\big(p^2 + 2p(1-p)\big) + (1-p)^2$. In Appendix 1 we show, with a population genetic analysis of equation (1), that the distorter will spread from rarity and reach fixation when $c_{drive}(k) < t(k)(1 - c_{drive}(k))$. This shows that distortion will evolve when the number of offspring that the distorter gains as a result of distortion ($t(k)(1-c_{drive}(k))$) is greater that the number of offspring bearing the distorter that are lost as a result of reduced individual fitness ($c_{drive}(k)$).

(2) Spread of an autosomal suppressor

We assume that the distorter (y_1) can be suppressed by an unlinked autosomal allele (suppressor), denoted by *sup*. This suppressor (*sup*) is dominant and only expressed in the presence of the distorter (facultative), and its expression may lead to a fitness cost to the individual, c_{sup} ($0 \le c_{sup} \le 1$)[39,40]. This cost can arise for multiple reasons, including energy expenditure, or errors relating to the use of gene silencing machinery, and is likely to be relatively low[41]. Gene silencing generally precedes the translation of the targeted gene, and so we assume that the costs of suppression (c_{sup}) is independent of the amount of distortion caused by the distorter (k).

We can write recursions detailing the generational change in the frequencies of the four possible gametes, y_0/+, y_0/*sup*, y_1/+, y_1/*sup*, with the respective frequencies in

bioRxiv preprint doi: https://doi.org/10.1101/526970; this version posted March 13, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

the current generation denoted by x_1, x_2, x_3 and x_4, and the frequencies in the subsequent generation denoted by an appended dash (’):

$$\bar{w}\, x_1' = x_1^2+x_1x_2+(1-t)(1-c_{drive})x_1x_3+((1-c_{sup})/2)x_1x_4+((1-c_{sup})/2)x_2x_3 \quad (2)$$

$$\bar{w}\, x_2' = x_1x_2+((1-c_{sup})/2)x_1x_4+x_2^2+((1-c_{sup})/2)x_2x_3+(1-c_{sup})x_2x_4$$

$$\bar{w}\, x_3' = (1+t)(1-c_{drive})x_1x_3+((1-c_{sup})/2)x_1x_4+((1-c_{sup})/2)x_2x_3+(1-c_{drive})x_3^2+(1-c_{sup})x_3x_4$$

$$\bar{w}\, x_4' = ((1-c_{sup})/2)x_1x_4+((1-c_{sup})/2)x_2x_3+(1-c_{sup})x_2x_4+(1-c_{sup})x_3x_4+(1-c_{sup})x_4^2.$$

$\bar{w}$ is the average fitness of individuals in the current generation, and equals the sum of the equations’ right-hand sides. In Appendix 2, we show, with a population genetic analysis of our system of equations (2), that a suppressor will spread from rarity if distortion, k, is greater than some threshold value at which the cost of suppression (c_{sup}) is less than the cost of being subjected to trait distortion, $c_{sup}<c_{drive}(k)$. A threshold with respect to the level of distortion (k) arises because the cost of trait distortion ($c_{drive}(k)$) increases with greater distortion, but the cost of suppression (c_{sup}) is constant.

(3) Consequences for organism trait values

The extent of trait distortion at the individual level shows a discontinuous relationship with the strength of the distorter. When distortion is low, a suppressor will not spread ($c_{sup}>c_{drive}(k)$) and so the level of trait distortion at the individual level will increase with the level of trait distortion induced by the distorter (k). However, once a threshold is reached ($c_{sup}<c_{drive}(k)$), the suppressor spreads. We show in Appendix 3 that the spread of the suppressor causes the distorter (y_1) to lose its selective

bioRxiv preprint doi: https://doi.org/10.1101/526970; this version posted March 13, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

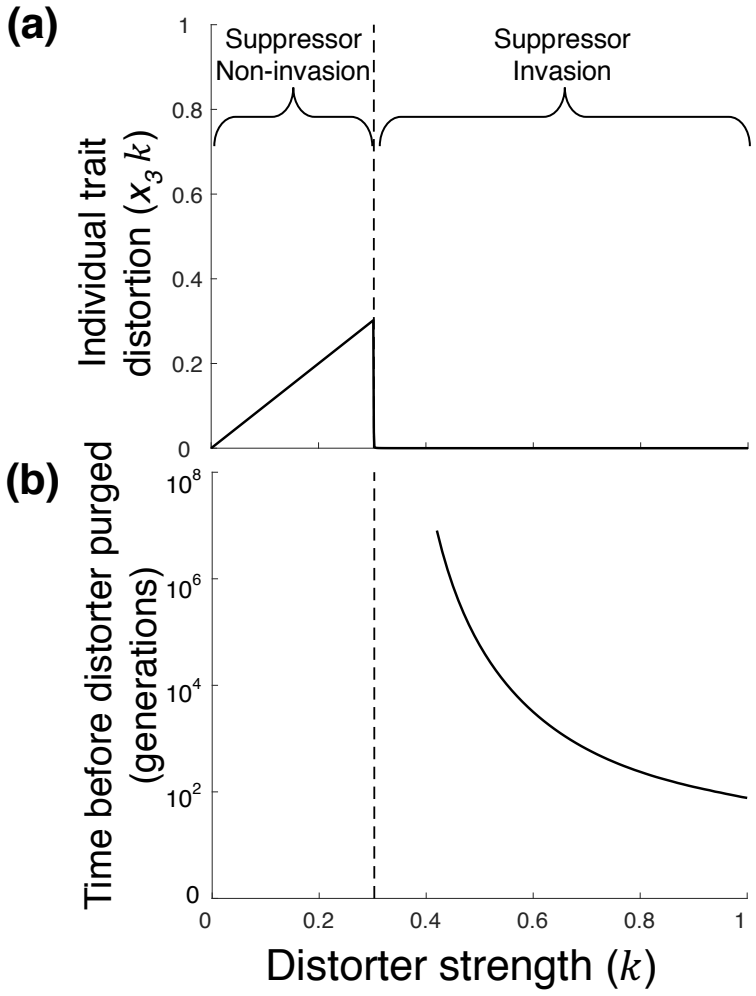

Figure 1. Distorter-suppressor dynamics and consequences for the organism. The trait distorter (y_1) and its suppressor (*sup*) are introduced from rarity. In part (a), the resulting average trait distortion ($x_3\,k$) is plotted at equilibrium, as the solid line, against the extent to which the distorter causes trait values to deviate from the individual optimum (k). Below a certain threshold strength ($c_{sup}>c_{drive}(k)$), to the left of the dashed line, the suppressor does not invade, and so the resulting trait distortion increases with the strength of the distorter (k). However, above this threshold, the suppressor invades, and the distorter is purged, restoring the trait to the individual optimum. In part (b), the number of generations for distorters that are purged at equilibrium (having been suppressed), which lie to the right of the dashed line, is plotted on a $\log_{10}$ scale. Stronger distorters are purged more quickly than weaker distorters. The numerical solutions displayed graphically assume that the cost of suppression, the transmission benefit of distortion, and the individual fitness cost of distortion, are respectively given by: c_{sup}=0.15; t=0.87k and c_{drive}=0.9$k^{1.5}$.

advantage and be eliminated from the population, leading to an absence of distortion at the individual level.

bioRxiv preprint doi: https://doi.org/10.1101/526970; this version posted March 13, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Overall, these results suggest that, given a relatively low cost of suppression (c_{sup}), the level of distortion observed at the individual level will either be low or absent. When a distorter is weak (low k), it will not be suppressed, but it will only have a small influence at the level of the individual. When a distorter is strong (high k), it will be suppressed and so there will be no influence at the level of the individual (Fig. 1a).

We also found that stronger distorters are suppressed more quickly (Fig. 1b). In Appendix 4, we numerically iterated our recursions to determine how many generations it takes for suppressors to reach equilibrium. As long as trait distortion continues to reduce individual fitness non-negligibly after suppression is favoured (such that $\frac{dt}{dk}/\frac{dc_{drive}}{dk}$ is not excessively high after $c_{sup}<c_{drive}(k)$), stronger distorters (higher k) are suppressed and purged more rapidly than weaker distorters, limiting the potential for non-equilibrium trait distortion (Fig. 1b).

(4) Evolution of trait distortion

We then considered the consequence of allowing the level of trait distortion (k) to evolve. We assume a distorter (y_1) that distorts by k, and then introduce a rare mutant (y_2) that distorts by a different amount $\hat{k}$ ($\hat{k}\neq k$). This mutant (y_2) is propagated into the proportion $(1+t(\hat{k})-t(k))/2$ of the offspring of y_2y_1 heterozygotes, and into the proportion $(1+t(\hat{k}))/2$ of the offspring of y_2y_0 heterozygotes. We assume that the stronger of the two distorters is dominant, but found similar results when assuming additivity (Appendix 5). We assume that the similarity in coding sequence and regulatory control means that the original distorter and the mutant are both

bioRxiv preprint doi: https://doi.org/10.1101/526970; this version posted March 13, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

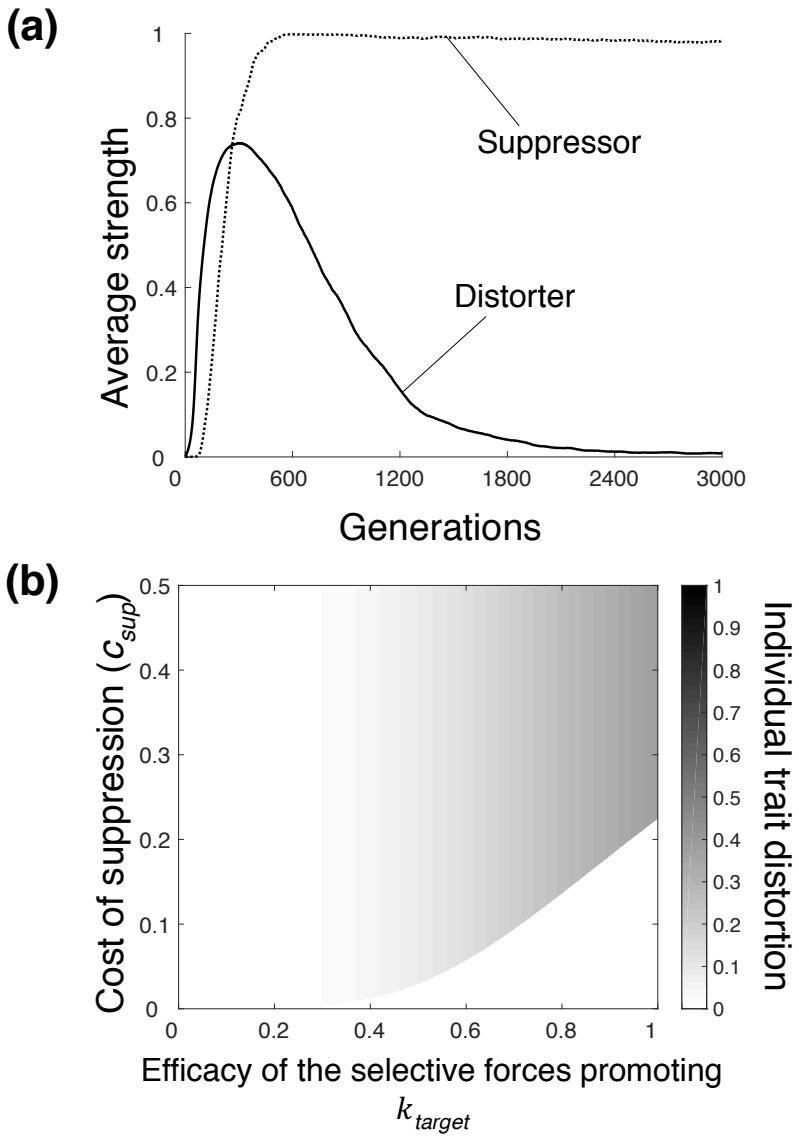

Figure 2. Evolution of trait distortion. In part (a), a distorter and suppressor are introduced in our agent-based simulation model (Appendix 7), and the following parameter regime is chosen: c_{sup}=0.1, t=k and c_{drive}=max(k_a,k_b)/2. The population average distorter and suppressor strengths over 100 simulation runs are plotted for successive generations. Initially, both distorter and suppressor strength increases. Eventually, a threshold is passed, after which, distorters are purged from the population, meaning the trait is undistorted at equilibrium. In part (b), each location on the graph corresponds to a different parameter regime. Along the *x* axis, we vary the rate of increase in the marginal individual cost of trait distortion $\left(\frac{d^2 c_{drive}}{dk^2}\right)$ relative to the rate of decrease in the marginal transmission benefit $\left(-\frac{d^2 t}{dk^2}\right)$, which determines the target level of trait distortion (k_{target}) (Supplementary Information 1). Along the *y* axis, we vary the cost of suppression (c_{sup}). We see that, when there is a low target level of distortion (left hand side of the heat map), suppressors fail to invade, leading to a low level of trait distortion at equilibrium. When there is a high target level of distortion and a relatively low cost of suppression (bottom right hand side), suppressors invade, leading to no trait distortion at equilibrium. When there is a high target level of distortion and a relatively high cost of suppression (top right hand side), suppressors do not invade, leading to appreciable trait distortion at equilibrium.

bioRxiv preprint doi: https://doi.org/10.1101/526970; this version posted March 13, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

suppressed by the same suppressor allele, at the same cost (c_{sup})[49]. In Appendix 5, we write the recursions that detail the generational frequency changes in the different possible gametes (y_0/+, y_0/*sup*, y_1/+, y_1/*sup*, y_2/+, y_2/*sup*).

We found that stronger mutant distorters ($\hat{k}>k$) will invade from rarity when the marginal increase in offspring they are propagated into exceeds the marginal increase in offspring they are lost from as a result of reduced fitness ($\Delta t(1-c_{drive}(\hat{k}))>\Delta c_{drive}$, where Δ denotes marginal change ($\Delta t=t(\hat{k})-t(k)$; $\Delta c_{drive}=c_{drive}(\hat{k})-c_{drive}(k)$)). Consequently, if distortion is initially low, and successive mutant distorters are introduced, each deviating only slightly from the distorters from which they are derived ("δ-weak selection"[42]), invading distorters will approach a 'target' strength, denoted by k_{target}. The target strength is that at which the marginal benefit of transmission is exactly counterbalanced by the marginal individual cost of reduced offspring, which occurs when $\frac{dt}{dk}(1-c_{drive})=\frac{dc_{drive}}{dk}$. The target strength of distortion (k_{target}) will therefore be greater if increased trait distortion (k) leads to a low rate of decrease in marginal transmission benefit $\left(-\frac{d^2t}{dk^2}\right)$ relative to the rate of increase in marginal individual cost $\left(\frac{d^2c_{drive}}{dk^2}\right)$ (Fig. 2b). If mutations are larger (strong selection), invading distorters may overshoot the target strength of distortion ($\hat{k}>k_{target}$). Weaker mutant distorters ($\hat{k}<k$) are recessive so cannot invade from rarity.

As evolution on the distorter increases the level of distortion, it makes it more likely that the distorter reaches the critical level of distortion where suppression will be favoured. When this is the case ($c_{sup}<c_{drive}(k_{target})$), the distorter spreads to high

bioRxiv preprint doi: https://doi.org/10.1101/526970; this version posted March 13, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

frequency, which then causes the suppressor to increase in frequency, reversing the direction of selection on the distorter, towards non-distortion (y_0), resulting in zero trait distortion at equilibrium (k^*=0) (Fig. 2a; Appendix 6). Suppression only fails to spread if the individual fitness cost associated with suppression is greater than the individual fitness cost associated with the target trait distortion ($c_{sup} > c_{drive}(k_{target})$; Fig. 2a). Given that the individual fitness cost of pre-translational suppression at a single locus is likely to be low, then any non-negligible distorter is likely to be suppressed.

Overall, our results suggest that selection on distorters will tend to drive the eventual suppression of those distorters. In Appendix 7 we developed an agent-based simulation, which allowed us to continuously vary the level of both distortion and suppression, and obtained results in close agreement (Fig. 2a; Supplementary Information 3, Fig. S2).

Specific Models

We then tested the robustness of our above conclusions by developing models for three different biological scenarios: a sex ratio distorter on an X chromosome (X driver); an imprinted gene that is only expressed when maternally inherited; and a gene for the production of a public good by bacteria, which is encoded on a mobile genetic element. We examined these cases because they are different types of distortion, involving different selection pressures, in very different organisms. We obtained qualitatively similar results in all three cases.

bioRxiv preprint doi: https://doi.org/10.1101/526970; this version posted March 13, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

In all of our specific models, we assume that the suppressor: is dominant; is only expressed in the presence of the distorter (facultative); completely suppresses the distorter; and may incur a fitness (viability) cost to the individual when it is expressed, independent of distorter strength, denoted by c_{sup} ($0 \leq c_{sup} \leq 1$)[43,44]. These assumptions fit well to a molecularly characterised suppressor ("*nmy*") of a sex ratio distorter ("*Dox*")[39,40]; and more generally to suppressors that act pre-translationally[45,46]. We also relax a simplifying assumption of our illustrative model, by allowing the transmission benefit and individual fitness cost of trait distortion to vary with the population frequency of the distorter.

Sex Ratio Distortion

We examined sex ratio evolution in a diploid species, in a large outbreeding (panmictic) population, with non-overlapping generations, and where males and females are equally costly to produce. Fisher[1] and many others have shown that, in this scenario, individuals would be selected to invest equally in male and female offspring[9,24]. We assumed genetic sex determination, with males as XY and females as XX, and that females mate with λ mates per generation. The distorter (y_1) that we considered is an X driving chromosome, which acts in males, killing Y-bearing sperm, and causing the male's mating partners to produce a higher proportion of female (XX) offspring. The proportion is given by $(1+k)/2$, where k denotes the proportion of Y-bearing sperm that are killed ($0<k\leq1$). We assumed that the sex ratio distorter can be suppressed by a costly autosomal suppressor (*sup*). This biology corresponds to sex ratio distortion in flies[25].

bioRxiv preprint doi: https://doi.org/10.1101/526970; this version posted March 13, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

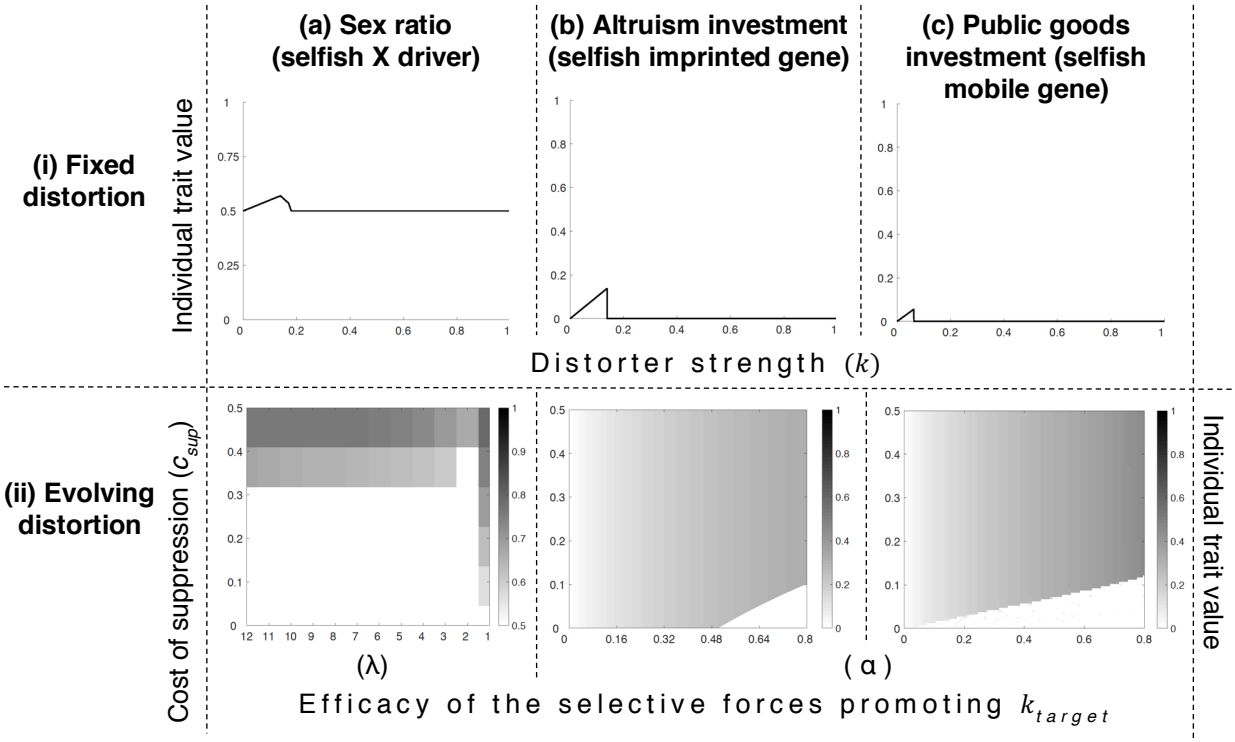

Figure 3. Specific Biological Scenarios. We consider three biological scenarios: (a) sex ratio distortion by an X driver; (b) cooperative investment by an imprinted gene, in which the cost of cooperation is assumed to be $c=k$ and the benefit is $b=k^{\alpha}$; (c) cooperative public goods investment by a mobile gene, in which the cost of cooperation is assumed to be $c=k$ and the benefit is $b=8k^{\alpha}$. Part (i) plots equilibrium trait values after short term coevolution between distorters of different strengths (k) and their suppressors. In part (i), the cost of suppression is taken to be $c_{sup}=0.05$; double female mating ($\lambda=2$) is assumed in (ai); and slightly decelerating returns on cooperation ($\alpha=0.9$) is assumed in (bi) and (ci). Part (i) shows that trait distortion will be relatively low or absent. Part (ii) shows the level of distortion that will be evolved to, when distorters can evolve. Along the x axis, we vary a model parameter that affects the target level of trait distortion (k_{target}). In (aii), we vary the female mating rate (λ), and in (bii) and (cii), we vary α, which mediates the rate of decrease in marginal cooperative benefits $\left(-\frac{d^2b}{dk^2}\right)$ relative to the rate of increase in marginal cooperative costs $\left(\frac{d^2c}{dk^2}\right)$. Part (ii) demonstrates that trait distortion will evolve to be zero or relatively low, except when there is a high cost of suppression (c_{sup}) alongside a high target level of distortion (low λ / high α), corresponding to the top right regions of the heat maps.

Our sex ratio model showed very similar results to our illustrative model. In Supplementary Information 4 we show with population genetic analyses that when distortion is weak (low k), suppressors are not favoured. In contrast, if distortion is strong (high k), then suppression is favoured, and so there is no net influence on the individual trait value. Consequently, the extent to which the sex ratio deviates from a

bioRxiv preprint doi: https://doi.org/10.1101/526970; this version posted March 13, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

50:50 investment will be small or zero[47] (Fig. 3ai). In addition to being more likely to be suppressed, the stronger distortion is, the quicker suppressors spread (Fig. S5). Finally, when we allowed the X chromosome driver to evolve, it evolved to high levels of sex ratio distortion (high k_{target}), increasing the likelihood of suppression.

We only obtained appreciable and detectable levels of sex ratio distortion (>60% females) if the cost of suppression exceeded a 15-35% viability reduction (Fig. S8). This is a much greater cost than what we would expect from natural gene suppression pathways[48]. A suppressor will only fail to spread when the individual cost of sex ratio distortion is less than the cost of suppressing the distorter (Equation S3). Given that the cost of suppression is likely to be low, we would only expect distorters that have relatively little impact at the individual level to evade suppression. We tested the robustness of our population genetic analysis with an agent-based simulation, and a game theory model, and found close agreement (Fig. S6).

Our predictions are consistent with the available data on X drivers in *Drosophila*. As predicted by our model: (1) Across natural populations of *D. simulans*, there is a positive correlation between the extent of sex ratio distortion and the extent of suppression[49]. (2) In both *D. mediopunctata* and *D. simulans* the presence of an X linked driver led to the experimental evolution of suppression[50,51]. In addition, consistent with our model: (3) In natural populations of *D. simulans* the prevalence of an X driver has been shown to sometimes decrease under complete suppression[52]. (4) Crossing different species of *Drosophila* has been shown to lead to appreciable

bioRxiv preprint doi: https://doi.org/10.1101/526970; this version posted March 13, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

sex ratio deviation, by unlinking distorters from their suppressors, and hence revealing previously hidden distorters[53]. Work on other sex ratio distorters has also shown that suppressors can spread extremely quickly from rarity, reaching fixation in as little as ~5 generations[54].

Genomic Imprinting and Altruism

Genomic imprinting occurs at a minority of genes in mammals and flowering plants. An imprinted allele has different epigenetic marks, and corresponding expression levels, when maternally and paternally inherited[55]. We examined the evolution of an altruistic helping behaviour in a population capable of genomic imprinting. A behaviour is altruistic if it incurs a cost (c) to perform, by the actor, and provides a benefit (b) to another individual, the recipient. Altruism is favoured if the genetic relatedness (R) between the actor and recipient is sufficiently high, such that $Rb>c$[3].

An individual may be more closely related to their social partners via their maternal or paternal genes[14,15,18]. For example, if a female mates two males, then on average her offspring would be related by R_m=1/2 at maternal genes and R_p=1/4 at paternal genes. If genes can 'gain information' about where they came from, by imprinting, then they could be selected to adjust traits accordingly. Assume that relatedness to social partners is R_p *and* R_m at paternal and maternal genes respectively. In this case, altruistic helping would be favoured at: maternally imprinted genes when $R_mb>c$; paternally imprinted genes when $R_pb>c$; and unimprinted genes when $((R_p+R_m)/2)b>c$[18,56,57]. Consequently, if $R_mb>c>((R_p+R_m)/2)b$, then altruistic helping is

bioRxiv preprint doi: https://doi.org/10.1101/526970; this version posted March 13, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

favoured at maternally imprinted genes, when it is disfavoured at unimprinted genes (selfish trait distortion).

We modelled the evolution of altruism in a large population of diploid, sexually reproducing individuals. The distorter (y_1) increases altruistic investment by some amount (k), at a fitness cost to the individual ($0 \leq c_{GI}(k) \leq 1$) and benefit to the social partner ($b_{GI}(k) > c_{GI}(k)$) that are both monotonically increasing functions of investment $\left(\frac{d\{b_{GI}, c_{GI}\}}{dk} \geq 0\right)$. The distorter ($y_1$) is only expressed when maternally inherited, whereas its potential suppressor (*sup*) is unimprinted[58]. Every generation, individuals associate in pairs with kin that they are maximally related to via their maternally inherited genes (R_m=1) but minimally related to via their paternally inherited genes (R_p=0). Individuals then have the opportunity to be altruistic to their partner, before mating at random in proportion to their fitness (fecundity), reproducing, then dying (non-overlapping generations).

In Supplementary Information 5, we showed with a population genetic analysis that our imprinting model produces very similar results to our illustrative model. When distortion is weak (low k), such that the cost of trait distortion is less than the cost of suppression, suppressors are not favoured (Equation S8). Given that the distorter suppression cost is likely to be low, these distorters that evade suppression will have relatively little impact at the individual level. When distortion is strong (high k), then suppression is favoured, and so there is no influence on the individual trait value[58]. Consequently, the extent to which altruistic investment deviates from the individual optimum of zero investment will be small or zero (Fig. 3bi). Finally, when we allowed

bioRxiv preprint doi: https://doi.org/10.1101/526970; this version posted March 13, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

404 the imprinted gene to evolve, we found that cooperative distortion increased until it
405 reached the value (k_{target}) at which the marginal benefit of cooperation is cancelled
406 out by the marginal cost of cooperation ($\frac{\partial b_{GI}}{\partial k} = \frac{\partial c_{GI}}{\partial k}$; Equation S10), increasing the
407 likelihood of suppression (Fig. 3bii).
408
409 Although there have been no direct tests, our predictions are consistent with data on
410 imprinted genes. There is no evidence that traits influenced by imprinted genes
411 deviate significantly from individual level optima under normal development[14].
412 Significant deviation is only observed when imprinted genes are deleted, implying
413 that imprinted trait distorters are either suppressed, or counterbalanced by
414 oppositely imprinted genes pulling the trait in the opposite direction[32,58]. Furthermore,
415 although many different parties (coreplicons) have vested interests in genomic
416 imprinting, our analysis suggests why the unimprinted majority could win control[59].
417 This could help explain both why imprinting appears to be relatively rare within the
418 genome[18,55,60], and why imprints are removed and re-added every generation in
419 mice, handing control of genomic patterns of imprinting to unimprinted genes[18,59,61].
420

421 Horizontal Gene Transfer and Public Goods

422 Bacteria produce and excrete many extracellular factors that provide a benefit to the
423 local population of cells and so can be thought of as public goods[62]. We modelled
424 the evolution of investment in a public good in a large, clonally reproducing
425 population. We assume a public good that costs *c* to produce, and provides a benefit
426 *b* to the group. We assume a well-mixed population, meaning genetic relatedness at

bioRxiv preprint doi: https://doi.org/10.1101/526970; this version posted March 13, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

vertically inherited genes is zero ($R_{vertical}$=0), and so public good production is disfavoured at the individual level ($R_{vertical}b$=0<c)[3,63].

We consider a distorter (y_1) of public goods production that is on a mobile gene, such as a plasmid. Mobile genes can spread within groups, increasing genetic relatedness at the mobile locus ($R_{horizontal}$>0), potentially favouring public goods production[64,65]. We assume that the distorter increases public goods investment by some amount (k), at a fitness cost to the individual ($0 \leq c_{HGT}(k) \leq 1$) and benefit shared within the group ($b_{HGT}(k) > c_{HGT}(k)$), that are both monotonically increasing functions of investment $\left(\frac{d\{b_{HGT},c_{HGT}\}}{dk} \geq 0\right)$. We also assume a potential suppressor (*sup*) that is immobile[46,66,67]. Each generation, individuals randomly aggregate into groups, and one allele at the mobile locus (y_0,y_1,y_2) spreads horizontally within each group, each with equal likelihood, increasing relatedness at the mobile locus. Public goods may then be produced and shared within groups, before individuals reproduce and die (non-overlapping generations).

In Supplementary Information 6, we show with a population genetic analysis that our plasmid model produces similar results to our illustrative model. When distortion is weak (low k), suppressors are not favoured, but the distorter has relatively little impact at the individual level (Fig. 3ci). When distortion is strong (high k), then suppression is favoured, and so there is no influence on the individual trait value[67]. Consequently, the extent to which public goods investment deviates from the individual optimum of zero investment will be relatively small or zero (Fig. 3ci). Finally, when we allowed the mobile distorter to evolve, we found that higher levels

bioRxiv preprint doi: https://doi.org/10.1101/526970; this version posted March 13, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

of public goods investment (high k_{target}) are favoured, leading to selection for suppression (Fig. 3cii). We lack empirical data that would allow us to test our model of mobile public goods genes.

Discussion

We have found that the individual level consequences of selfish genetic elements ('*distorters*') will be either small or non-existent. If distorters lead to only small distortions of traits, then they will not be suppressed, but they will only have small effects on traits (Figs. 1a & 3i). Specifically, distorters will only remain unsuppressed if trait distortion compromises individual fitness less than suppression of the trait distorter does ($c_{drive}(k)<c_{sup}$). Given that the individual fitness cost of pre-translational suppression at a single locus is likely to be low, we can say that trait distortion conferred by unsuppressed distorters is likely to be relatively negligible.

However, if distorters lead to large distortions of traits then this selects for their suppression, and so they will have no net effect on traits at the individual level (Figs. 1a & 3i). Stronger distorters will also be suppressed more quickly (Fig. 1b). Furthermore, selection on distorters favours higher levels of distortion, which will render them more likely to be ultimately suppressed. Consequently, the evolution of distorters will often drive their own demise (Figs. 2, 3ii & 4). These results suggest that even if there is substantial potential for genetic conflict, distorters will have relatively little influence at the individual level, in support of Leigh's[27] parliament of genes hypothesis.

bioRxiv preprint doi: https://doi.org/10.1101/526970; this version posted March 13, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

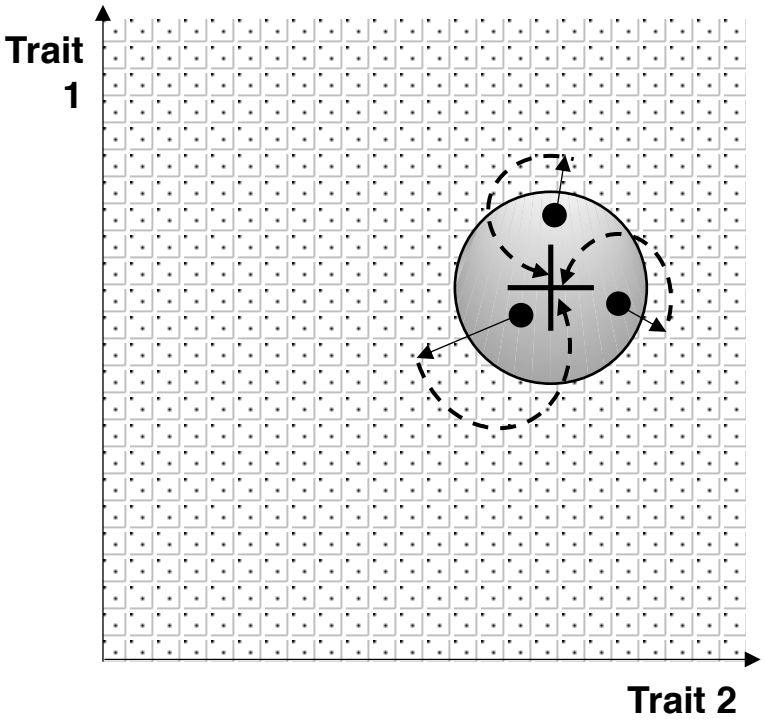

Figure 4: Selfish genetic elements evolve to be suppressed by the parliament of genes. The cross represents the position in phenotype space, here defined with respect to two traits, *1* and *2*, that maximises the fitness of an individual. The circle surrounding the cross represents the phenotype space where suppression of selfish genetic elements, that have distorted traits 1 or 2, would not be selected for. The surrounding area represents the phenotype space in which the parliament of genes is selected to suppress selfish genetic elements. The three dots represent three possible individuals, which, owing to weakly selfish genetic elements, are not maximising individual fitness (the dots do not lie exactly on the cross), although they are approximately (e.g. within the bounds of experimental error in measurement). Because these deviations from individual fitness maximisation are only slight, costly suppression of the weakly selfish genetic elements does not evolve. However, the selfish genetic elements will evolve to become more distorting (solid arrows), bringing individuals into the non-tolerated area of phenotype space, where they will be suppressed and individual fitness maximisation (the black cross) is regained (dashed arrows).

We have assumed that there will be a much greater number of genes where suppression is favoured, and so it is relatively easy for a suppressor to be reached by mutation[31]. We have therefore examined, given the potential for suppression, what direction would we expect natural selection to take on average. We are not claiming that appreciable trait distortion will never evolve, and there are clearly cases where it has[14]. A number of biological details will matter for different systems, including: (i) whether a cost of suppression is incurred when the distorter is not present; (ii) whether distorters can outstrip suppressors in coevolutionary arms

bioRxiv preprint doi: https://doi.org/10.1101/526970; this version posted March 13, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

497 races, leading to distorters going unsuppressed for extended time periods[31,54,68]; (iii)
498 the rate at which trait-distorting selfish genetic elements arise by mutation, relative to
499 the rate at which their suppressors arise by mutation. Our results are supported by
500 cases where appreciable distortion is only revealed in hybrid crosses, implying that
501 they are generally suppressed[53,69].
502
503 We emphasise that when the assumption of individual fitness maximisation is made
504 in behavioural and evolutionary ecology, it is not being assumed that natural
505 selection produces perfect fitness maximisers[5]. Many factors could constrain
506 adaptation, such as genetic architecture, mutation and phylogenetic constraints[70,71].
507 Instead, the assumption of fitness maximisation is used as a basis to investigate the
508 selective forces that have favoured particular traits (adaptations). The aim is not to
509 test if organisms maximise fitness, or behave 'optimally', but rather to try to
510 understand the selective forces favouring particular traits or behaviours[2].
511
512 To conclude, debate over the validity of assuming individual level fitness
513 maximisation has revolved around whether selfish genetic elements are common or
514 rare[4,19,20,23,72]. We have shown that that even if selfish genetic elements are
515 common, they will tend to be either weak and negligible, or suppressed. This
516 suggests that even if there is the potential for appreciable genetic conflict, individual
517 level fitness maximisation will still often be a reasonable assumption. This allows us
518 to explain why certain traits, especially the sex ratio, have been able to provide such
519 clear support for both individual level fitness maximisation and genetic conflict.
520

bioRxiv preprint doi: https://doi.org/10.1101/526970; this version posted March 13, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Appendix

Appendix 1: Distorter population frequency

We ask when a rare distorter (y_1) can invade a population fixed for the non-distorter (y_0). We take Equation (1), set $p'=p=p^*$, and solve to find two possible equilibria: $p^*=0$ (non-distorter fixation) and $p^*=1$ (distorter fixation). The distorter (y_1) can invade from rarity when the $p^*=0$ equilibrium is unstable, which occurs when the differential of p' with respect to p, at $p^*=0$, is greater than one. The distorter invasion criterion is therefore $c_{drive}(k)<t(k)(1-c_{drive}(k))$.

We now ask what frequency the distorter (y_1) will reach after invasion. The distorter (y_1) can spread to fixation if the $p^*=1$ equilibrium is stable, which requires that the differential of p' with respect to p, at $p^*=1$, is less than one. This requirement always holds true, demonstrating that there is no negative frequency dependence on the distorter, and that it will always spread to fixation after its initial invasion.

Appendix 2: Suppressor invasion condition

We ask when the suppressor (*sup*) can spread from rarity in a population in which the distorter (y_1) and non-suppressor (+) are fixed at equilibrium. We derive the Jacobian stability matrix for this equilibrium, which is a matrix of each genotype frequency (x_1', x_2', x_3', x_4') differentiated by each genotype frequency in the prior generation (x_1, x_2, x_3, x_4), at the equilibrium position given by $x_1^*=0$, $x_2^*=0$, $x_3^*=1$, $x_4^*=0$:

bioRxiv preprint doi: https://doi.org/10.1101/526970; this version posted March 13, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

$$J = \begin{pmatrix} 1-t & \frac{1-c_{sup}}{2(1-c_{drive})} & 0 & 0 \\ 0 & \frac{1-c_{sup}}{2(1-c_{drive})} & 0 & 0 \\ t-1 & \frac{-3(1-c_{sup})}{2(1-c_{drive})} & 0 & \frac{-(1-c_{sup})}{1-c_{drive}} \\ 0 & \frac{1-c_{sup}}{2(1-c_{drive})} & 0 & \frac{1-c_{sup}}{1-c_{drive}} \end{pmatrix}. \quad \text{(A1)}$$

The suppressor can invade when the equilibrium is unstable, which occurs when the leading eigenvalue is greater than one. The leading eigenvalue is $(1-c_{sup})/(1-c_{drive})$, meaning the suppressor invasion criterion is $c_{drive}>c_{sup}$.

Appendix 3: Equilibrium distorter and suppressor frequencies

We ask what frequency the distorter (y_1) and suppressor (*sup*) will reach after initial suppressor (*sup*) invasion. We assume that the suppressor is introduced from rarity when the distorter has reached the population frequency given by f ($x_1 \to f$, $x_3 \to 1-f$, $\{x_2, x_4\} \to 0$). We numerically iterate Equations (2), over successive generations, until equilibrium has been reached. At equilibrium, for all parameter combinations $(f, t, c_{sup}, c_{drive})$, the suppressor reaches an internal equilibrium and the distorter is lost from the population ($x_1^* + x_2^* = 1$, $x_3^* = 0$, $x_4^* = 0$). This equilibrium arises because distorter-presence gives the suppressor (*sup*) a selective advantage, leading to high suppressor frequency, which in turn reverses the selective advantage of the distorter (y_1), leading to distorter loss and suppressor equilibration.

Appendix 4: Non-equilibrium trait distortion

We consider a distorter that is suppressed and therefore purged at equilibrium ($c_{drive}>c_{sup}$), and ask to what extent it can contribute to individual trait distortion in the

bioRxiv preprint doi: https://doi.org/10.1101/526970; this version posted March 13, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

565 period after its initial invasion but before its eventual loss (non-equilibrium). We
566 introduce the distorter (y_1) and suppressor (*sup*) from rarity and numerically iterate
567 our recursions until the distorter has been purged from the population (or a cap of
568 20,000,000 generations has been reached). We vary parameters between $0 \leq t \leq 1$,
569 $c_{sup} < c_{drive} \leq 1$, $0 \leq c_{sup} \leq 1$.
570
571 We find that a higher cost of trait distortion (c_{drive}) relative to suppression (c_{sup}) leads
572 to shorter non-equilibrium maintenance of the distorter in the population. This is
573 because the cost of trait distortion relative to suppression mediates selection on the
574 suppressor (Appendix 2). We find that a higher transmission bias (t) leads to longer
575 non-equilibrium maintenance of the distorter in the population, but this effect is
576 diluted as the cost of trait distortion (c_{drive}) is increased relative to suppression (c_{sup})
577 (Supplementary Information 2, Fig. S1). Stronger distorters (with higher k, leading to
578 higher c_{drive} and t) are therefore generally suppressed and purged more rapidly than
579 weaker distorters (Fig. 1b). Exceptions are distorters that reduce individual fitness
580 relatively negligibly after the point (k) at which suppression is favoured, such that
581 $\frac{dt}{dk} / \frac{dc_{drive}}{dk}$ is very high for values of k satisfying $c_{sup} < c_{drive}(k)$.
582
583 **Appendix 5: Invasion of a mutant distorter**
584 We ask when a mutant distorter (y_2) will invade against a resident distorter (y_1) that
585 is unsuppressed and at fixation ($k \neq \hat{k}$). We write recursions detailing the generational
586 frequency changes in the six possible gametes, y_0/+, y_0/*sup*, y_1/+, y_1/*sup*, y_2/+,
587 y_2/*sup*, with current generation frequencies denoted respectively by x_1, x_2, x_3, x_4, x_5,
588 x_6, and next generation frequencies denoted with an appended dash (’):

bioRxiv preprint doi: https://doi.org/10.1101/526970; this version posted March 13, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

$$\bar{w}\,x_1' = x_1x_1+x_1x_2+(1-t(k))(1-c_{drive}(k))x_1x_3+((1-c_{sup})/2)x_1x_4+(1-t(\hat{k}))(1-c_{drive}(\hat{k}))\,x_1x_5+((1-c_{sup})/2)x_1x_6+((1-c_{sup})/2)x_2x_3+((1-c_{sup})/2)x_2x_5 \quad (A2)$$

$$\bar{w}\,x_2' = x_1x_2+((1-c_{sup})/2)x_1x_4+((1-c_{sup})/2)x_1x_6+x_2x_2+((1-c_{sup})/2)x_2x_3+(1-c_{sup})x_2x_4+((1-c_{sup})/2)x_2x_5+(1-c_{sup})x_2x_6$$

$$\bar{w}\,x_3' = (1+t(k))(1-c_{drive}(k))x_1x_3+((1-c_{sup})/2)x_1x_4+((1-c_{sup})/2)x_2x_3+(1-c_{drive}(k))x_3x_3+(1-c_{sup})x_3x_4+(1+t(k)-t(\hat{k}))(1-c_{drive}(\max(k,\hat{k})))x_3x_5+((1-c_{sup})/2)x_3x_6+((1-c_{sup})/2)x_4x_5$$

$$\bar{w}\,x_4' = ((1-c_{sup})/2)x_1x_4+((1-c_{sup})/2)x_2x_3+(1-c_{sup})x_2x_4+(1-c_{sup})x_3x_4+((1-c_{sup})/2)x_3x_6+(1-c_{sup})x_4x_4+((1-c_{sup})/2)x_4x_5+(1-c_{sup})x_4x_6$$

$$\bar{w}\,x_5' = (1+t(\hat{k}))(1-c_{drive}(\hat{k}))x_1x_5+((1-c_{sup})/2)x_1x_6+((1-c_{sup})/2)x_2x_5+(1-t(k)+t(\hat{k}))(1-c_{drive}(\max(k,\hat{k})))x_3x_5+((1-c_{sup})/2)x_3x_6+((1-c_{sup})/2)x_4x_5+(1-c_{drive}(\hat{k}))x_5x_5+(1-c_{sup})x_5x_6$$

$$\bar{w}\,x_6' = ((1-c_{sup})/2)x_1x_6+((1-c_{sup})/2)x_2x_5+(1-c_{sup})x_2x_6+((1-c_{sup})/2)x_3x_6+((1-c_{sup})/2)x_4x_5+(1-c_{sup})x_4x_6+(1-c_{sup})x_5x_6+(1-c_{sup})x_6x_6.$$

$\bar{w}$ is the average fitness of individuals in the current generation, and equals the sum of the right-hand side of the system of equations. The mutant distorter can invade when the equilibrium given by $x_1^*=0$, $x_2^*=0$, $x_3^*=1$, $x_4^*=0$, $x_5^*=0$, $x_6^*=0$ is unstable, which occurs when the leading eigenvalue of the Jacobian stability matrix for this equilibrium is greater than one. Testing for stability in this way, we find that, if the mutant distorter is weaker than the resident, it can never invade. If the mutant distorter is stronger than the resident, it invades from rarity when $\Delta t(1-c_{drive}(\hat{k}))>\Delta c_{drive}$, where $\Delta t=t(\hat{k})-t(k)$, $\Delta c_{drive}=c_{drive}(\hat{k})-c_{drive}(k)$.

bioRxiv preprint doi: https://doi.org/10.1101/526970; this version posted March 13, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

The implication is that, if trait distortion is initially low, and mutant distorters are successively introduced, each deviating only very slightly from the resident distorter from which they are derived, such that $\hat{k}=k\pm\delta$, where δ is very small ("δ-weak selection"[42]), then distorters will approach a 'target' strength at which $\frac{dt}{dk}(1-c_{drive})=\frac{dc_{drive}}{dk}$. In the absence of suppression, this target (k_{target}) is the equilibrium level of distortion ($k^*=k_{target}$). However, if mutant distorters (y_2) are allowed to deviate appreciably from residents (y_1) (strong selection), then distorters may invade even if they overshoot the target ($\hat{k}>k_{target}$). In the absence of suppression, k_{target} is then not the equilibrium level of distortion, but rather, the minimum equilibrium level of distortion ($k^*>k_{target}$) (Supplementary Information 3, Fig. S2b).

We could alternatively have assumed that an individual's trait is distorted according to the average strength of its alleles (additive gene interactions), rather than according to the stronger (higher-k) allele (dominance). Such an assumption leads to a single invasion criterion for a mutant distorter, regardless of whether the mutant distorter is stronger or weaker than the resident distorter, given by: $\Delta t(2\text{-}c_{drive}(k)\text{-}c_{drive}(\hat{k}))>\Delta c_{drive}$. In the absence of suppression, this leads to an equilibrium level of distortion (k^*), that holds even under strong selection, that satisfies $2\frac{dt}{dk}(1-c_{drive})=\frac{dc_{drive}}{dk}$.

Appendix 6: Equilibrium distorter and suppressor frequencies (long term evolution)

bioRxiv preprint doi: https://doi.org/10.1101/526970; this version posted March 13, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

We ask what equilibrium state will arise after the invasion of a mutant distorter. We assume that the mutant distorter (y_2) is introduced from rarity when the resident distorter (y_1) has reached the population frequency given by q. We numerically iterate Equations (A2), over successive generations, until equilibrium has been reached. At equilibrium, for all parameter combinations (q, $t(k)$, $t(\hat{k})$, c_{sup}, $c_{drive}(k)$, $c_{drive}(\hat{k})$), the resident distorter (y_1) is lost from the population (x_3,x_4=0), with either the mutant distorter (y_2) and non-suppressor (+) at fixation (x_5*=1), or the non-distorter at fixation alongside the suppressor at an internal equilibrium (x_1*+x_2*=1). The latter scenario arises if the mutant distorter triggers suppressor invasion (c_{sup}<$c_{drive}(\hat{k})$). This equilibrium arises because mutant distorter-presence gives the suppressor (*sup*) a selective advantage, leading to high suppressor frequency, which in turn reverses the selective advantage of distortion, leading to distorter (y_1,y_2) loss and suppressor equilibration.

Appendix 7: Agent-based simulation

We construct an agent-based simulation to ask what level of trait distortion evolves when continuous variation is permitted at distorter and suppressor loci. We model a population of N=2000 individuals and track evolution at two autosomal loci: a distorter locus (L1) and a suppressor locus (L2). Each individual has two alleles at the distorter locus, with strengths denoted by k_a and k_b, and two alleles at the suppressor locus, with strengths denoted by m_a and m_b (diploid). Strengths can take any continuous value between zero and one. We assume that, for both loci, the strongest (highest value) allele within an individual is dominant. The absolute fitness of an individual with at least one active meiotic driver (max(k_a,k_b)>0) is: 1-

bioRxiv preprint doi: https://doi.org/10.1101/526970; this version posted March 13, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

660 $c_{drive}(\max(k_a,k_b))(1- \max(m_a,m_b))-c_{sup}\max(m_a,m_b)$, and the absolute fitness of an
661 individual lacking an active distorter ($\max(k_a,k_b)=0$) is 1. The function
662 $c_{drive}(\max(k_a,k_b))$ is given an explicit form in simulations (Supplementary Information
663 3, Fig. S2).
664
665 Each generation, there are N breeding pairs. To fill each position in each breeding
666 pair, individuals are drawn from the population, with replacement, with probabilities
667 given by their fitness (hermaphrodites). Breeding pairs then reproduce to produce
668 one offspring, before dying (non-overlapping generations). Alleles at the suppressor
669 locus (L2) are inherited in Mendelian fashion. Alleles at the distorter locus may drive,
670 meaning the parental allele of strength k_a is inherited, rather than the allele of
671 strength k_b, with the probability $(1+(t(k_a)-t(k_b))(1-\max(m_a,m_b)))/2$. The transmission
672 bias function, t, is given an explicit form in simulations (Supplementary Information 3,
673 Fig. S2). Each generation, distorter and suppressor alleles have a 0.01 chance of
674 mutating to a new value, which is drawn from a normal distribution centred around
675 the pre-mutation value, with variance 0.2, and truncated between 0 and 1. We track
676 the population average distorter strength, denoted by $E[k]$, and suppressor strength,
677 denoted by $E[m]$, over 20,000 generations. We see that, allowing for continuous
678 variation at the distorter and suppressor loci, if the cost of suppression (c_{sup}) is not
679 excessively high, trait distortion at equilibrium is either low or nothing (Fig. 2a; Fig.
680 S2b).
681

682 # References

bioRxiv preprint doi: https://doi.org/10.1101/526970; this version posted March 13, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

683 1. Fisher, R. A. *The Genetical Theory of Natural Selection*. (Oxford University
684 Press, 1930).
685 2. Gardner, A. Adaptation as organism design. *Biology Letters* **5,** 861–864
686 (2009).
687 3. Hamilton, W. D. The genetical evolution of social behaviour. I & II. *Journal of*
688 *Theoretical Biology* **7** (1964).
689 4. Grafen, A. The formal darwinism project in outline. *Biol Philos* **29,** 155–174
690 (2014).
691 5. Parker, G. A. & Maynard Smith, J. Optimality theory in evolutionary biology.
692 *Nature Publishing Group* **348,** 27–33 (1990).
693 6. Frank, S. A. *Foundations of Social Evolution*. (Princeton University Press,
694 1998). doi:10.1038/sj.hdy.6885351
695 7. Stephens, D. W. & Krebs, J. R. *Foraging Theory*. (Princeton University Press,
696 1986).
697 8. Stearns, S. C. *The Evolution of Life Histories*. (Oxford University Press on
698 Demand, 1992).
699 9. West, S. *Sex Allocation*. (Princeton University Press, 2009).
700 doi:10.1515/9781400832019
701 10. Davies, N. B., Krebs, J. R. & West, S. A. *An Introduction to Behavioural*
702 *Ecology*. (John Wiley & Sons, 2012).
703 11. Westneat, D. & Fox, C. W. *Evolutionary Behavioral Ecology*. (Oxford
704 University Press on Demand, 2010).
705 12. Bourke, A. F. G. *Principles of Social Evolution*. (OUP Oxford, 2011).

bioRxiv preprint doi: https://doi.org/10.1101/526970; this version posted March 13, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

13. Werren, J. H., NUR, U. & Wu, C. I. Selfish Genetic Elements. *Trends in Ecology & Evolution* **3,** 297–302 (1988).

14. Burt, A. & Trivers, R. *Genes in Conflict*. (Harvard University Press, 2006). doi:10.4159/9780674029118

15. Gardner, A. & Úbeda, F. The meaning of intragenomic conflict. *Nature Ecology & Evolution* **1,** 1–9 (2017).

16. Ågren, J. A. & Clark, A. G. Selfish genetic elements. *PLoS Genet* **14,** e1007700–20 (2018).

17. Hamilton, W. D. Extraordinary sex ratios. A sex-ratio theory for sex linkage and inbreeding has new implications in cytogenetics and entomology. *Science* **156,** 477–488 (1967).

18. Haig, D. *Genomic Imprinting and Kinship*. (Rutgers University Press, 2002).

19. Haig, D. Genetic dissent and individual compromise. *Biol Philos* **29,** 233–239 (2014).

20. Bourke, A. F. G. The gene's-eye view, major transitions and the formal darwinism project. *Biol Philos* **29,** 241–248 (2014).

21. Cosmides, L. M. & Tooby, J. Cytoplasmic inheritance and intragenomic conflict. *Journal of Theoretical Biology* **89,** 83–129 (1981).

22. West, S. A. & Gardner, A. Adaptation and Inclusive Fitness. *Current Biology* **23,** R577–R584 (2013).

23. Ågren, J. A. Selfish genetic elements and the gene's-eye view of evolution. *Curr Zool* **62,** 659–665 (2016).

24. Charnov, E. L. *The Theory of Sex Allocation*. (Princeton University Press, 1982).

bioRxiv preprint doi: https://doi.org/10.1101/526970; this version posted March 13, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

730 25. Jaenike, J. Sex Chromosome Meiotic Drive. *Annual Review of Ecology and*
731 *Systematics* **32,** 25–49 (2001).
732 26. Leigh, E. G., Herre, E. A. & Fischer, E. A. Sex allocation in animals.
733 *Experientia* **41,** 1265–1276 (1985).
734 27. Leigh, E. G. *Adaptation and diversity*. (1971).
735 28. Leigh, E. G. When does the good of the group override the advantage of the
736 individual? *Proc Natl Acad Sci USA* **80,** 2985–2989 (1983).
737 29. Leigh, E. G. How does selection reconcile individual advantage with the good
738 of the group? *Proc Natl Acad Sci USA* **74,** 4542–4546 (1977).
739 30. Strassmann, J. E. & Queller, D. C. The Social Organism: Congresses, Parties,
740 and Committees. *Evolution* **64,** 605–616 (2010).
741 31. Queller, D. C. & Strassmann, J. E. Evolutionary Conflict. *Annu. Rev. Ecol.*
742 *Evol. Syst.* **49,** annurev–ecolsys–110617–062527–21 (2018).
743 32. Gardner, A. & Ross, L. Mating ecology explains patterns of genome
744 elimination. *Ecol Lett* **17,** 1602–1612 (2014).
745 33. Crow, J. F. Why is mendelian segregation so exact? *Bioessays* **13,** 305–312
746 (1991).
747 34. Ridley, M. *Mendel's Demon*. 176–177 (Weidenfeld & Nicolson, 2000).
748 35. Prout, T., Bundgaard, J. & Bryant, S. Population genetics of modifiers of
749 meiotic drive I. The solution of a special case and some general implications.
750 *Theoretical Population Biology* **4,** 446–465 (1973).
751 36. Caubet, Hatcher, MocquardRigaud. Genetic conflict and changes in
752 heterogametic mechanisms of sex determination. *Journal of Evolutionary*
753 *Biology* **13,** 766–777 (2000).

bioRxiv preprint doi: https://doi.org/10.1101/526970; this version posted March 13, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

754 37. Randerson, J. P. The evolutionary dynamics of male-killers and their hosts.
755 *Heredity* **84,** 1–9 (2000).
756 38. Haig, D. Intragenomic politics. *Cytogenet Genome Res* **113,** 68–74 (2006).
757 39. Tao, Y., Masly, J. P., Araripe, L., Ke, Y. & Hartl, D. L. A sex-ratio meiotic drive
758 system in Drosophila simulans. I: an autosomal suppressor. *PLoS Biol* **5,** e292
759 (2007).
760 40. Tao, Y. *et al.* A sex-ratio meiotic drive system in Drosophila simulans. II: an X-
761 linked distorter. *PLoS Biol* **5,** e293 (2007).
762 41. Qiu, S. A computational study of off-target effects of RNA interference. *Nucleic*
763 *Acids Research* **33,** 1834–1847 (2005).
764 42. Wild, G. & Traulsen, A. The different limits of weak selection and the
765 evolutionary dynamics of finite populations. *Journal of Theoretical Biology* **247,**
766 382–390 (2007).
767 43. Vaz, S. C. & Carvalho, A. B. Evolution of autosomal suppression of the sex-
768 ratio trait in Drosophila. *Genetics* **166,** 265–277 (2004).
769 44. Hall, D. W. Meiotic drive and sex chromosome cycling. *Evolution* **58,** 925–931
770 (2004).
771 45. Aravin, A. A., Hannon, G. J. & Brennecke, J. The Piwi-piRNA pathway
772 provides an adaptive defense in the transposon arms race. *Science* **318,** 761–
773 764 (2007).
774 46. Doron, S. *et al.* Systematic discovery of antiphage defense systems in the
775 microbial pangenome. *Science* **359,** eaar4120–13 (2018).
776 47. Rood, E. S. & Freedberg, S. Intragenomic conflict produces sex ratio dynamics
777 that favor maternal sex ratio distorters. *Ecol Evol* **6,** 8085–8093 (2016).

bioRxiv preprint doi: https://doi.org/10.1101/526970; this version posted March 13, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

778 48. Unckless, R. L., Larracuente, A. M. & Clark, A. G. Sex-ratio meiotic drive and
779 Y-linked resistance in Drosophila affinis. *Genetics* **199,** 831–840 (2015).
780 49. Atlan, A., Mercot, H., Landre, C. & Montchamp-Moreau, C. The Sex-Ratio Trait
781 in Drosophila simulans: Geographical Distribution of Distortion and
782 Resistance. *Evolution* **51,** 1886 (1997).
783 50. Carvalho, A. B., Sampaio, M. C., Varandas, F. R. & Klaczko, L. B. An
784 experimental demonstration of Fisher's principle: Evolution of sexual
785 proportion by natural selection. *Genetics* **148,** 719–731 (1998).
786 51. Capillon, C. & Atlan, A. Evolution of Driving X Chromosomes and Resistance
787 Factors in Experimental Populations of Drosophila simulans. *Evolution* **53,** 506
788 (1999).
789 52. Bastide, H. *et al.* Rapid Rise and Fall of Selfish Sex-Ratio X Chromosomes in
790 Drosophila simulans: Spatiotemporal Analysis of Phenotypic and Molecular
791 Data. *Molecular Biology and Evolution* **28,** 2461–2470 (2011).
792 53. Blows, M. W., Berrigan, D. & Gilchrist, G. W. Rapid evolution towards equal
793 sex ratios in a system with heterogamety. *Evol Ecol Res* 277–283 (1999).
794 54. Hornett, E. A. *et al.* Evolution of Male-Killer Suppression in a Natural
795 Population. *PLoS Biol* **4,** e283–6 (2006).
796 55. Peters, J. The role of genomic imprinting in biology and disease: an expanding
797 view. *Nature Publishing Group* **15,** 517–530 (2014).
798 56. Úbeda, F. & Gardner, A. A model for genomic imprinting in the social brain:
799 juveniles. *Evolution* **64,** 2587–2600 (2010).
800 57. Queller, D. C. Theory of genomic imprinting conflict in social insects. *BMC*
801 *Evol Biol* **3,** 15 (2003).

bioRxiv preprint doi: https://doi.org/10.1101/526970; this version posted March 13, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

802 58. Wilkins, J. F. & Haig, D. Genomic imprinting of two antagonistic loci.
803 *Proceedings of the Royal Society B: Biological Sciences* **268,** 1861–1867
804 (2001).
805 59. Burt, A. & Trivers, R. Genetic conflicts in genomic imprinting. *Proceedings of*
806 *the Royal Society B: Biological Sciences* **265,** 2393–2397 (1998).
807 60. Gregg, C. *et al.* High-Resolution Analysis of Parent-of-Origin Allelic Expression
808 in the Mouse Brain. *Science* **329,** 643–648 (2010).
809 61. Kafri, T., Gao, X. H. & Razin, A. Mechanistic Aspects of Genome-Wide
810 Demethylation in the Preimplantation Mouse Embryo. *Proc Natl Acad Sci USA*
811 **90,** 10558–10562 (1993).
812 62. West, S. A., Diggle, S. P., Buckling, A., Gardner, A. & Griffin, A. S. The Social
813 Lives of Microbes. *Annu. Rev. Ecol. Evol. Syst.* **38,** 53–77 (2007).
814 63. Frank, S. A. A general model of the public goods dilemma. *Journal of*
815 *Evolutionary Biology* **23,** 1245–1250 (2010).
816 64. Nogueira, T. *et al.* Horizontal Gene Transfer of the Secretome Drives the
817 Evolution of Bacterial Cooperation and Virulence. *Current Biology* **19,** 1683–
818 1691 (2009).
819 65. Smith, J. The social evolution of bacterial pathogenesis. *Proceedings of the*
820 *Royal Society B: Biological Sciences* **268,** 61–69 (2001).
821 66. Johnson, L. J. The Genome Strikes Back: The Evolutionary Importance of
822 Defence Against Mobile Elements. *Evolutionary Biology* **34,** 121–129 (2007).
823 67. Mc Ginty, S. É. & Rankin, D. J. The evolution of conflict resolution between
824 plasmids and their bacterial hosts. *Evolution* **66,** 1662–1670 (2012).

bioRxiv preprint doi: https://doi.org/10.1101/526970; this version posted March 13, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

825 68. Dyson, E. A. & Hurst, G. Persistence of an extreme sex-ratio bias in a natural
826 population. *Proc Natl Acad Sci USA* **101,** 6520–6523 (2004).
827 69. Corbett-Detig, R., Medina, P., Frérot, H., Blassiau, C. & Castric, V. Bulk pollen
828 sequencing reveals rapid evolution of segregation distortion in the male
829 germline of Arabidopsis hybrids. *Evolution Letters* **129,** 1393–11 (2019).
830 70. Maynard Smith, J. *et al.* Developmental Constraints and Evolution: A
831 Perspective from the Mountain Lake Conference on Development and
832 Evolution. *The Quarterly Review of Biology* **60,** 265–287 (1985).
833 71. Lehmann, L., Alger, I. & Weibull, J. Does evolution lead to maximizing
834 behavior? *Evolution* **69,** 1858–1873 (2015).
835 72. Gardner, A. & Grafen, A. Capturing the superorganism: a formal theory of
836 group adaptation. *Journal of Evolutionary Biology* **22,** 659–671 (2009).

837

838 Acknowledgements

839 We thank Geoff Wild, Alan Grafen, Egbert Giles Leigh, Jr., Guy Cooper, Sam Levin,
840 Asher Leeks and Matishalin Patel for helpful comments and discussion.

841

842 Author Contributions

843 TWS and SAW designed the study and wrote the paper. TWS carried out
844 mathematical analysis.

845

846 Competing Interests

bioRxiv preprint doi: https://doi.org/10.1101/526970; this version posted March 13, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

847 The authors declare no competing interests.

848

849 Data Accessibility

850 The data that support the findings of this study are available from the corresponding

851 author upon reasonable request.