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: stuart.west@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: (1) when selfish genetic elements have a greater impact at the individual level, they are more likely to be suppressed, and suppression spreads more quickly; (2) selection on selfish genetic elements leads to higher levels of distortion, making them more likely to be suppressed. Overall, our results suggest that selfish genetic elements will either have a minor impact at the individual level, or tend to be suppressed, and that selection on selfish genetic elements will often drive their own suppression.

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?

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 the 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].

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. These 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 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

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 ($0 \leq t(k) \leq 1$) 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 $1-c_{drive}(k)$ ($0 \leq c_{drive}(k) \leq 1$). 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:

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

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} ($0 \leq c_{sup} \leq 1$)[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

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

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

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

206 stronger of the two distorters is dominant, but found similar results when assuming
207 additivity (Appendix 5). We assume that the similarity in coding sequence and
208 regulatory control means that the original distorter and the mutant are both
209 suppressed by the same suppressor allele, at the same cost (c_{sup})[49]. In Appendix 5,
210 we write the recursions that detail the generational frequency changes in the
211 different possible gametes (y_0/+, y_0/*sup*, y_1/+, y_1/*sup*, y_2/+, y_2/*sup*).
212
213 We found that stronger mutant distorters ($\hat{k}>k$) will invade from rarity when the
214 marginal increase in offspring they are propagated into exceeds the marginal
215 increase in offspring they are lost from as a result of reduced fitness ($\Delta t(1-$
216 $c_{drive}(\hat{k}))>\Delta c_{drive}$, where Δ denotes marginal change ($\Delta t=t(\hat{k})-t(k)$; $\Delta c_{drive}=c_{drive}(\hat{k})-$
217 $c_{drive}(k)$)). Consequently, if distortion is initially low, and successive mutant distorters
218 are introduced, each deviating only slightly from the distorters from which they are
219 derived ("δ-weak selection"[42]), invading distorters will approach a 'target' strength,
220 denoted by k_{target}. The target strength is that at which the marginal benefit of
221 transmission is exactly counterbalanced by the marginal individual cost of reduced
222 offspring, which occurs when $\frac{dt}{dk}(1-c_{drive})=\frac{dc_{drive}}{dk}$. The target strength of distortion
223 (k_{target}) will therefore be greater if increased trait distortion (k) leads to a low rate of
224 decrease in marginal transmission benefit $\left(-\frac{d^2t}{dk^2}\right)$ relative to the rate of increase in
225 marginal individual cost $\left(\frac{d^2c_{drive}}{dk^2}\right)$ (Fig. 2b). If mutations are larger (strong selection),
226 invading distorters may overshoot the target strength of distortion ($\hat{k}>k_{target}$). Weaker
227 mutant distorters ($\hat{k}<k$) are recessive so cannot invade from rarity.
228

229 As evolution on the distorter increases the level of distortion, it makes it more likely
230 that the distorter reaches the critical level of distortion where suppression will be
231 favoured. When this is the case ($c_{sup}<c_{drive}(k_{target})$), the distorter spreads to high
232 frequency, which then causes the suppressor to increase in frequency, reversing the
233 direction of selection on the distorter, towards non-distortion (y_0), resulting in zero
234 trait distortion at equilibrium ($k^*=0$) (Fig. 2a; Appendix 6). Suppression only fails to
235 spread if the individual fitness cost associated with suppression is greater than the
236 individual fitness cost associated with the target trait distortion ($c_{sup}>c_{drive}(k_{target})$; Fig.
237 2a). Given that the individual fitness cost of pre-translational suppression at a single
238 locus is likely to be low, then any non-negligible distorter is likely to be suppressed.
239
240 Overall, our results suggest that selection on distorters will tend to drive the eventual
241 suppression of those distorters. In Appendix 7 we developed an agent-based
242 simulation, which allowed us to continuously vary the level of both distortion and
243 suppression, and obtained results in close agreement (Fig. 2a; Supplementary
244 Information 3, Fig. S2).
245

246 Specific Models

247 We then tested the robustness of our above conclusions by developing models for
248 three different biological scenarios: a sex ratio distorter on an X chromosome (X
249 driver); an imprinted gene that is only expressed when maternally inherited; and a
250 gene for the production of a public good by bacteria, which is encoded on a mobile
251 genetic element. We examined these cases because they are different types of

252 distortion, involving different selection pressures, in very different organisms. We
253 obtained qualitatively similar results in all three cases.
254
255 In all of our specific models, we assume that the suppressor: is dominant; is only
256 expressed in the presence of the distorter (facultative); completely suppresses the
257 distorter; and may incur a fitness (viability) cost to the individual when it is
258 expressed, independent of distorter strength, denoted by c_{sup} ($0 \leq c_{sup} \leq 1$)[43,44]. These
259 assumptions fit well to a molecularly characterised suppressor ("*nmy*") of a sex ratio
260 distorter ("*Dox*")[39,40]; and more generally to suppressors that act pre-
261 translationally[45,46]. We also relax a simplifying assumption of our illustrative model,
262 by allowing the transmission benefit and individual fitness cost of trait distortion to
263 vary with the population frequency of the distorter.
264

265 Sex Ratio Distortion

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

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 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. 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 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 $R_m b>c$; paternally imprinted genes when $R_p b>c$; and unimprinted genes when $((R_p+R_m)/2)b>c$[18,56,57]. Consequently, if $R_m b>c>((R_p+R_m)/2)b$, then altruistic helping is 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

348 distortion is weak (low k), such that the cost of trait distortion is less than the cost of
349 suppression, suppressors are not favoured (Equation S8). Given that the distorter
350 suppression cost is likely to be low, these distorters that evade suppression will have
351 relatively little impact at the individual level. When distortion is strong (high k), then
352 suppression is favoured, and so there is no influence on the individual trait value[58].
353 Consequently, the extent to which altruistic investment deviates from the individual
354 optimum of zero investment will be small or zero (Fig. 3bi). Finally, when we allowed
355 the imprinted gene to evolve, we found that cooperative distortion increased until it
356 reached the value (k_{target}) at which the marginal benefit of cooperation is cancelled
357 out by the marginal cost of cooperation ($\frac{\partial b_{GI}}{\partial k} = \frac{\partial c_{GI}}{\partial k}$; Equation S10), increasing the
358 likelihood of suppression (Fig. 3bii).
359
360 Although there have been no direct tests, our predictions are consistent with data on
361 imprinted genes. There is no evidence that traits influenced by imprinted genes
362 deviate significantly from individual level optima under normal development[14].
363 Significant deviation is only observed when imprinted genes are deleted, implying
364 that imprinted trait distorters are either suppressed, or counterbalanced by
365 oppositely imprinted genes pulling the trait in the opposite direction[32,58]. Furthermore,
366 although many different parties (coreplicons) have vested interests in genomic
367 imprinting, our analysis suggests why the unimprinted majority could win control[59].
368 This could help explain both why imprinting appears to be relatively rare within the
369 genome[18,55,60], and why imprints are removed and re-added every generation in
370 mice, handing control of genomic patterns of imprinting to unimprinted genes[18,59,61].
371

Horizontal Gene Transfer and Public Goods

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

396 weak (low k), suppressors are not favoured, but the distorter has relatively little
397 impact at the individual level (Fig. 3ci). When distortion is strong (high k), then
398 suppression is favoured, and so there is no influence on the individual trait value[67].
399 Consequently, the extent to which public goods investment deviates from the
400 individual optimum of zero investment will be relatively small or zero (Fig. 3ci).
401 Finally, when we allowed the mobile distorter to evolve, we found that higher levels
402 of public goods investment (high k_{target}) are favoured, leading to selection for
403 suppression (Fig. 3cii). We lack empirical data that would allow us to test our model
404 of mobile public goods genes.
405

406 # Discussion

407 We have found that the individual level consequences of selfish genetic elements
408 ('*distorters*') will be either small or non-existent. If distorters lead to only small
409 distortions of traits, then they will not be suppressed, but they will only have small
410 effects on traits (Figs. 1 & 3). Specifically, distorters will only remain unsuppressed if
411 trait distortion compromises individual fitness less than suppression of the trait
412 distorter does ($c_{drive}(k)<c_{sup}$). Given that the individual fitness cost of pre-translational
413 suppression at a single locus is likely to be low, we can say that trait distortion
414 conferred by unsuppressed distorters is likely to be relatively negligible.
415

416 However, if distorters lead to large distortions of traits then this selects for their
417 suppression, and so they will have no net effect on traits at the individual level (Figs.
418 1 & 3). Stronger distorters will also be suppressed more quickly (Fig. 1b).
419 Furthermore, selection on distorters favours higher levels of distortion, which will

420 render them more likely to be ultimately suppressed. Consequently, the evolution of
421 distorters will often drive their own demise (Figs. 2 & 4). These results suggest that
422 even if there is substantial potential for genetic conflict, distorters will have relatively
423 little influence at the individual level, in support of Leigh's[27] parliament of the genes
424 hypothesis.
425
426 We have assumed that there will be a much greater number of genes where
427 suppression is favoured, and so it is relatively easy for a suppressor to be reached
428 by mutation[31]. We have therefore examined, given the potential for suppression,
429 what direction would we expect natural selection to take on average. We are not
430 claiming that appreciable trait distortion will never evolve, and there are clearly cases
431 where it has[14]. A number of biological details will matter for different systems,
432 including: (i) whether a cost of suppression is incurred when the distorter is not
433 present; (ii) whether distorters can outstrip suppressors in coevolutionary arms
434 races, leading to distorters going unsuppressed for extended time periods[31,54,68]; (iii)
435 the rate at which trait-distorting selfish genetic elements arise by mutation, relative to
436 the rate at which their suppressors arise by mutation. Our results are supported by
437 cases where appreciable distortion is only revealed in hybrid crosses, implying that
438 they are generally suppressed[53,69].
439
440 We emphasise that when the assumption of individual fitness maximisation is made
441 in behavioural and evolutionary ecology, it is not being assumed that natural
442 selection produces perfect fitness maximisers[5]. Many factors could constrain
443 adaptation, such as genetic architecture, mutation and phylogenetic constraints[70,71].

Instead, the assumption of fitness maximisation is used as a basis to investigate the selective forces that have favoured particular traits (adaptations). The aim is not to test if organisms maximise fitness, or behave 'optimally', but rather to try to understand the selective forces favouring particular traits or behaviours[2].

To conclude, debate over the validity of assuming individual level fitness maximisation has revolved around whether selfish genetic elements are common or rare[4,19,20,23,72]. We have shown that that even if selfish genetic elements are common, they will tend to be either weak and negligible, or suppressed. This suggests that even if there is the potential for appreciable genetic conflict, individual level fitness maximisation will still often be a reasonable assumption. This allows us to explain why certain traits, especially the sex ratio, have been able to provide such clear support for both individual level fitness maximisation and genetic conflict.

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 $x_1^*=0$, $x_2^*=0$, $x_3^*=1$, $x_4^*=0$:

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

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 ($x_1 \rightarrow f$, $x_3 \rightarrow 1-f$, $\{x_2, x_4\} \rightarrow 0$). 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 period after its initial invasion but before its eventual loss (non-equilibrium). We introduce the distorter (y_1) and suppressor (*sup*) from rarity and numerically iterate our recursions until the distorter has been purged from the population (or a cap of 20,000,000 generations has been reached). We vary parameters between $0 \leq t \leq 1$, $c_{sup} < c_{drive} \leq 1$, $0 \leq c_{sup} \leq 1$.

We find that a higher cost of trait distortion (c_{drive}) relative to suppression (c_{sup}) leads to shorter non-equilibrium maintenance of the distorter in the population. This is because the cost of trait distortion relative to suppression mediates selection on the suppressor (Appendix 2). We find that a higher transmission bias (t) leads to longer

non-equilibrium maintenance of the distorter in the population, but this effect is diluted as the cost of trait distortion (c_{drive}) is increased relative to suppression (c_{sup}) (Supplementary Information 2, Fig. S1). Stronger distorters (with higher k, leading to higher c_{drive} and t) are therefore generally suppressed and purged more rapidly than weaker distorters (Fig. 1b). Exceptions are distorters that reduce individual fitness relatively negligibly after the point (k) at which suppression is favoured, such that $\frac{dt}{dk}/\frac{dc_{drive}}{dk}$ is very high for values of k satisfying $c_{sup}<c_{drive}(k)$.

Appendix 5: Invasion of a mutant distorter

We ask when a mutant distorter (y_2) will invade against a resident distorter (y_1) that is unsuppressed and at fixation ($k\neq\hat{k}$). We write recursions detailing the generational frequency changes in the six possible gametes, $y_0/+$, y_0/sup, $y_1/+$, y_1/sup, $y_2/+$, y_2/sup, with current generation frequencies denoted respectively by x_1, x_2, x_3, x_4, x_5, x_6, and next generation frequencies denoted with an appended dash (’):

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

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

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

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

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

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

$\bar{w}$ 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 $\Delta t=t(\hat{k})-t(k)$, $\Delta c_{drive}=c_{drive}(\hat{k})-c_{drive}(k)$.

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

559 the equilibrium level of distortion, but rather, the minimum equilibrium level of
560 distortion ($k^*>k_{target}$) (Supplementary Information 3, Fig. S2b).
561
562 We could alternatively have assumed that an individual's trait is distorted according
563 to the average strength of its alleles (additive gene interactions), rather than
564 according to the stronger (higher-k) allele (dominance). Such an assumption leads to
565 a single invasion criterion for a mutant distorter, regardless of whether the mutant
566 distorter is stronger or weaker than the resident distorter, given by:
567 $\Delta t(2-c_{drive}(k)-c_{drive}(\hat{k}))>\Delta c_{drive}$. In the absence of suppression, this leads to an
568 equilibrium level of distortion (k^*), that holds even under strong selection, that
569 satisfies $2\frac{dt}{dk}(1-c_{drive})=\frac{dc_{drive}}{dk}$.
570
571 **Appendix 6: Equilibrium distorter and suppressor frequencies (long term**
572 **evolution)**
573 We ask what equilibrium state will arise after the invasion of a mutant distorter. We
574 assume that the mutant distorter (y_2) is introduced from rarity when the resident
575 distorter (y_1) has reached the population frequency given by q. We numerically
576 iterate Equations (A2), over successive generations, until equilibrium has been
577 reached. At equilibrium, for all parameter combinations (q, $t(k)$, $t(\hat{k})$, c_{sup}, $c_{drive}(k)$,
578 $c_{drive}(\hat{k})$), the resident distorter (y_1) is lost from the population (x_3,x_4=0), with either
579 the mutant distorter (y_2) and non-suppressor (+) at fixation (x_5^*=1), or the non-
580 distorter at fixation alongside the suppressor at an internal equilibrium ($x_1^*+x_2^*$=1).
581 The latter scenario arises if the mutant distorter triggers suppressor invasion
582 ($c_{sup}<c_{drive}(\hat{k})$). This equilibrium arises because mutant distorter-presence gives the

583 suppressor (*sup*) a selective advantage, leading to high suppressor frequency, which
584 in turn reverses the selective advantage of distortion, leading to distorter (y_1,y_2) loss
585 and suppressor equilibration.
586
587 **Appendix 7: Agent-based simulation**
588 We construct an agent-based simulation to ask what level of trait distortion evolves
589 when continuous variation is permitted at distorter and suppressor loci. We model a
590 population of N=2000 individuals and track evolution at two autosomal loci: a
591 distorter locus (L1) and a suppressor locus (L2). Each individual has two alleles at
592 the distorter locus, with strengths denoted by k_a and k_b, and two alleles at the
593 suppressor locus, with strengths denoted by m_a and m_b (diploid). Strengths can take
594 any continuous value between zero and one. We assume that, for both loci, the
595 strongest (highest value) allele within an individual is dominant. The absolute fitness
596 of an individual with at least one active meiotic driver ($\max(k_a,k_b)>0$) is: 1-
597 $c_{drive}(\max(k_a,k_b))(1- \max(m_a,m_b))-c_{sup}\max(m_a,m_b)$, and the absolute fitness of an
598 individual lacking an active distorter ($\max(k_a,k_b)=0$) is 1. The function
599 $c_{drive}(\max(k_a,k_b))$ is given an explicit form in simulations (Supplementary Information
600 3, Fig. S2).
601
602 Each generation, there are N breeding pairs. To fill each position in each breeding
603 pair, individuals are drawn from the population, with replacement, with probabilities
604 given by their fitness (hermaphrodites). Breeding pairs then reproduce to produce
605 one offspring, before dying (non-overlapping generations). Alleles at the suppressor
606 locus (L2) are inherited in Mendelian fashion. Alleles at the distorter locus may drive,

meaning the parental allele of strength k_a is inherited, rather than the allele of strength k_b, with the probability $(1+(t(k_a)-t(k_b))(1-\max(m_a,m_b)))/2$. The transmission bias function, t, is given an explicit form in simulations (Supplementary Information 3, Fig. S2). Each generation, distorter and suppressor alleles have a 0.01 chance of mutating to a new value, which is drawn from a normal distribution centred around the pre-mutation value, with variance 0.2, and truncated between 0 and 1. We track the population average distorter strength, denoted by E[k], and suppressor strength, denoted by E[m], over 20,000 generations. We see that, allowing for continuous variation at the distorter and suppressor loci, if the cost of suppression (c_{sup}) is not excessively high, trait distortion at equilibrium is either low or nothing (Fig. 2a; Fig. S2b).

Acknowledgements

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

Author Contributions

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

Competing Interests

The authors declare no competing interests.

Data Accessibility

The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

1. Fisher, R. A. *The Genetical Theory of Natural Selection*. (Oxford University Press, 1930).
2. Gardner, A. Adaptation as organism design. *Biology Letters* **5,** 861–864 (2009).
3. Hamilton, W. D. The genetical evolution of social behaviour. I & II. *Journal of Theoretical Biology* **7** (1964).
4. Grafen, A. The formal darwinism project in outline. *Biol Philos* **29,** 155–174 (2014).
5. Parker, G. A. & Maynard Smith, J. Optimality theory in evolutionary biology. *Nature Publishing Group* **348,** 27–33 (1990).
6. Frank, S. A. *Foundations of Social Evolution*. (Princeton University Press, 1998). doi:10.1038/sj.hdy.6885351
7. Stephens, D. W. & Krebs, J. R. *Foraging Theory*. (Princeton University Press, 1986).
8. Stearns, S. C. *The Evolution of Life Histories*. (Oxford University Press on Demand, 1992).

651 9. West, S. *Sex Allocation*. (Princeton University Press, 2009).
652 doi:10.1515/9781400832019
653 10. Davies, N. B., Krebs, J. R. & West, S. A. *An Introduction to Behavioural*
654 *Ecology*. (John Wiley & Sons, 2012).
655 11. Westneat, D. & Fox, C. W. *Evolutionary Behavioral Ecology*. (Oxford
656 University Press on Demand, 2010).
657 12. Bourke, A. F. G. *Principles of Social Evolution*. (OUP Oxford, 2011).
658 13. Werren, J. H., NUR, U. & Wu, C. I. Selfish Genetic Elements. *Trends in*
659 *Ecology & Evolution* **3,** 297–302 (1988).
660 14. Burt, A. & Trivers, R. *Genes in Conflict*. (Harvard University Press, 2006).
661 doi:10.4159/9780674029118
662 15. Gardner, A. & Úbeda, F. The meaning of intragenomic conflict. *Nature Ecology*
663 *& Evolution* **1,** 1–9 (2017).
664 16. Ågren, J. A. & Clark, A. G. Selfish genetic elements. *PLoS Genet* **14,**
665 e1007700–20 (2018).
666 17. Hamilton, W. D. Extraordinary sex ratios. A sex-ratio theory for sex linkage and
667 inbreeding has new implications in cytogenetics and entomology. *Science* **156,**
668 477–488 (1967).
669 18. Haig, D. *Genomic Imprinting and Kinship*. (Rutgers University Press, 2002).
670 19. Haig, D. Genetic dissent and individual compromise. *Biol Philos* **29,** 233–239
671 (2014).
672 20. Bourke, A. F. G. The gene's-eye view, major transitions and the formal
673 darwinism project. *Biol Philos* **29,** 241–248 (2014).

674 21. Cosmides, L. M. & Tooby, J. Cytoplasmic inheritance and intragenomic
675 conflict. *Journal of Theoretical Biology* **89,** 83–129 (1981).
676 22. West, S. A. & Gardner, A. Adaptation and Inclusive Fitness. *Current Biology*
677 **23,** R577–R584 (2013).
678 23. Ågren, J. A. Selfish genetic elements and the gene's-eye view of evolution.
679 *Curr Zool* **62,** 659–665 (2016).
680 24. Charnov, E. L. *The Theory of Sex Allocation*. (Princeton University Press,
681 1982).
682 25. Jaenike, J. Sex Chromosome Meiotic Drive. *Annual Review of Ecology and*
683 *Systematics* **32,** 25–49 (2001).
684 26. Leigh, E. G., Herre, E. A. & Fischer, E. A. Sex allocation in animals.
685 *Experientia* **41,** 1265–1276 (1985).
686 27. Leigh, E. G. *Adaptation and diversity*. (1971).
687 28. Leigh, E. G. When does the good of the group override the advantage of the
688 individual? *Proc Natl Acad Sci USA* **80,** 2985–2989 (1983).
689 29. Leigh, E. G. How does selection reconcile individual advantage with the good
690 of the group? *Proc Natl Acad Sci USA* **74,** 4542–4546 (1977).
691 30. Strassmann, J. E. & Queller, D. C. The Social Organism: Congresses, Parties,
692 and Committees. *Evolution* **64,** 605–616 (2010).
693 31. Queller, D. C. & Strassmann, J. E. Evolutionary Conflict. *Annu. Rev. Ecol.*
694 *Evol. Syst.* **49,** annurev–ecolsys–110617–062527–21 (2018).
695 32. Gardner, A. & Ross, L. Mating ecology explains patterns of genome
696 elimination. *Ecol Lett* **17,** 1602–1612 (2014).

697 33. Crow, J. F. Why is mendelian segregation so exact? *Bioessays* **13,** 305–312
698 (1991).
699 34. Ridley, M. *Mendel's Demon*. 176–177 (Weidenfeld & Nicolson, 2000).
700 35. Prout, T., Bundgaard, J. & Bryant, S. Population genetics of modifiers of
701 meiotic drive I. The solution of a special case and some general implications.
702 *Theoretical Population Biology* **4,** 446–465 (1973).
703 36. Caubet, Hatcher, MocquardRigaud. Genetic conflict and changes in
704 heterogametic mechanisms of sex determination. *Journal of Evolutionary*
705 *Biology* **13,** 766–777 (2000).
706 37. Randerson, J. P. The evolutionary dynamics of male-killers and their hosts.
707 *Heredity* **84,** 1–9 (2000).
708 38. Haig, D. Intragenomic politics. *Cytogenet Genome Res* **113,** 68–74 (2006).
709 39. Tao, Y., Masly, J. P., Araripe, L., Ke, Y. & Hartl, D. L. A sex-ratio meiotic drive
710 system in Drosophila simulans. I: an autosomal suppressor. *PLoS Biol* **5,** e292
711 (2007).
712 40. Tao, Y. *et al.* A sex-ratio meiotic drive system in Drosophila simulans. II: an X-
713 linked distorter. *PLoS Biol* **5,** e293 (2007).
714 41. Qiu, S. A computational study of off-target effects of RNA interference. *Nucleic*
715 *Acids Research* **33,** 1834–1847 (2005).
716 42. Wild, G. & Traulsen, A. The different limits of weak selection and the
717 evolutionary dynamics of finite populations. *Journal of Theoretical Biology* **247,**
718 382–390 (2007).
719 43. Vaz, S. C. & Carvalho, A. B. Evolution of autosomal suppression of the sex-
720 ratio trait in Drosophila. *Genetics* **166,** 265–277 (2004).

721 44. Hall, D. W. Meiotic drive and sex chromosome cycling. *Evolution* **58,** 925–931
722 (2004).
723 45. Aravin, A. A., Hannon, G. J. & Brennecke, J. The Piwi-piRNA pathway
724 provides an adaptive defense in the transposon arms race. *Science* **318,** 761–
725 764 (2007).
726 46. Doron, S. *et al.* Systematic discovery of antiphage defense systems in the
727 microbial pangenome. *Science* **359,** eaar4120–13 (2018).
728 47. Rood, E. S. & Freedberg, S. Intragenomic conflict produces sex ratio dynamics
729 that favor maternal sex ratio distorters. *Ecol Evol* **6,** 8085–8093 (2016).
730 48. Unckless, R. L., Larracuente, A. M. & Clark, A. G. Sex-ratio meiotic drive and
731 Y-linked resistance in Drosophila affinis. *Genetics* **199,** 831–840 (2015).
732 49. Atlan, A., Mercot, H., Landre, C. & Montchamp-Moreau, C. The Sex-Ratio Trait
733 in Drosophila simulans: Geographical Distribution of Distortion and
734 Resistance. *Evolution* **51,** 1886 (1997).
735 50. Carvalho, A. B., Sampaio, M. C., Varandas, F. R. & Klaczko, L. B. An
736 experimental demonstration of Fisher's principle: Evolution of sexual
737 proportion by natural selection. *Genetics* **148,** 719–731 (1998).
738 51. Capillon, C. & Atlan, A. Evolution of Driving X Chromosomes and Resistance
739 Factors in Experimental Populations of Drosophila simulans. *Evolution* **53,** 506
740 (1999).
741 52. Bastide, H. *et al.* Rapid Rise and Fall of Selfish Sex-Ratio X Chromosomes in
742 Drosophila simulans: Spatiotemporal Analysis of Phenotypic and Molecular
743 Data. *Molecular Biology and Evolution* **28,** 2461–2470 (2011).

744 53. Blows, M. W., Berrigan, D. & Gilchrist, G. W. Rapid evolution towards equal
745 sex ratios in a system with heterogamety. *Evol Ecol Res* 277–283 (1999).
746 54. Hornett, E. A. *et al.* Evolution of Male-Killer Suppression in a Natural
747 Population. *PLoS Biol* **4,** e283–6 (2006).
748 55. Peters, J. The role of genomic imprinting in biology and disease: an expanding
749 view. *Nature Publishing Group* **15,** 517–530 (2014).
750 56. Úbeda, F. & Gardner, A. A model for genomic imprinting in the social brain:
751 juveniles. *Evolution* **64,** 2587–2600 (2010).
752 57. Queller, D. C. Theory of genomic imprinting conflict in social insects. *BMC*
753 *Evol Biol* **3,** 15 (2003).
754 58. Wilkins, J. F. & Haig, D. Genomic imprinting of two antagonistic loci.
755 *Proceedings of the Royal Society B: Biological Sciences* **268,** 1861–1867
756 (2001).
757 59. Burt, A. & Trivers, R. Genetic conflicts in genomic imprinting. *Proceedings of*
758 *the Royal Society B: Biological Sciences* **265,** 2393–2397 (1998).
759 60. Gregg, C. *et al.* High-Resolution Analysis of Parent-of-Origin Allelic Expression
760 in the Mouse Brain. *Science* **329,** 643–648 (2010).
761 61. Kafri, T., Gao, X. H. & Razin, A. Mechanistic Aspects of Genome-Wide
762 Demethylation in the Preimplantation Mouse Embryo. *Proc Natl Acad Sci USA*
763 **90,** 10558–10562 (1993).
764 62. West, S. A., Diggle, S. P., Buckling, A., Gardner, A. & Griffin, A. S. The Social
765 Lives of Microbes. *Annu. Rev. Ecol. Evol. Syst.* **38,** 53–77 (2007).
766 63. Frank, S. A. A general model of the public goods dilemma. *Journal of*
767 *Evolutionary Biology* **23,** 1245–1250 (2010).

768 64. Nogueira, T. *et al.* Horizontal Gene Transfer of the Secretome Drives the
769 Evolution of Bacterial Cooperation and Virulence. *Current Biology* **19,** 1683–
770 1691 (2009).
771 65. Smith, J. The social evolution of bacterial pathogenesis. *Proceedings of the*
772 *Royal Society B: Biological Sciences* **268,** 61–69 (2001).
773 66. Johnson, L. J. The Genome Strikes Back: The Evolutionary Importance of
774 Defence Against Mobile Elements. *Evolutionary Biology* **34,** 121–129 (2007).
775 67. Mc Ginty, S. É. & Rankin, D. J. The evolution of conflict resolution between
776 plasmids and their bacterial hosts. *Evolution* **66,** 1662–1670 (2012).
777 68. Dyson, E. A. & Hurst, G. Persistence of an extreme sex-ratio bias in a natural
778 population. *Proc Natl Acad Sci USA* **101,** 6520–6523 (2004).
779 69. Corbett-Detig, R., Medina, P., Frérot, H., Blassiau, C. & Castric, V. Bulk pollen
780 sequencing reveals rapid evolution of segregation distortion in the male
781 germline of Arabidopsis hybrids. *Evolution Letters* **129,** 1393–11 (2019).
782 70. Williams, G. C. *Adaptation and Natural Selection*. (Princeton University Press,
783 1966). doi:10.2307/j.ctv39x5jt
784 71. Maynard Smith, J. *et al.* Developmental Constraints and Evolution: A
785 Perspective from the Mountain Lake Conference on Development and
786 Evolution. *The Quarterly Review of Biology* **60,** 265–287 (1985).
787 72. Gardner, A. & Grafen, A. Capturing the superorganism: a formal theory of
788 group adaptation. *Journal of Evolutionary Biology* **22,** 659–671 (2009).

Figure Legends

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$k^{1.5}$.

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

812 on the graph corresponds to a different parameter regime. Along the *x* axis, we vary
813 the rate of increase in the marginal individual cost of trait distortion $\left(\frac{d^2 c_{drive}}{dk^2}\right)$ relative
814 to the rate of decrease in the marginal transmission benefit $\left(-\frac{d^2 t}{dk^2}\right)$, which
815 determines the target level of trait distortion (k_{target}) (Supplementary Information 1).
816 Along the *y* axis, we vary the cost of suppression (c_{sup}). We see that, when there is a
817 low target level of distortion (left hand side of the heat map), suppressors fail to
818 invade, leading to a low level of trait distortion at equilibrium. When there is a high
819 target level of distortion and a relatively low cost of suppression (bottom right hand
820 side), suppressors invade, leading to no trait distortion at equilibrium. When there is
821 a high target level of distortion and a relatively high cost of suppression (top right
822 hand side), suppressors do not invade, leading to appreciable trait distortion at
823 equilibrium.
824
825 **Figure 3. Specific Biological Scenarios.** We consider three biological scenarios:
826 (a) sex ratio distortion by an X driver; (b) cooperative investment by an imprinted
827 gene, in which the cost of cooperation is assumed to be $c=k$ and the benefit is $b=k^{\alpha}$;
828 (c) cooperative public goods investment by a mobile gene, in which the cost of
829 cooperation is assumed to be $c=k$ and the benefit is $b=8k^{\alpha}$. Part (i) plots equilibrium
830 trait values after short term coevolution between distorters of different strengths (*k*)
831 and their suppressors. In part (i), the cost of suppression is taken to be $c_{sup}=0.05$;
832 double female mating ($\lambda=2$) is assumed in (ai); and slightly decelerating returns on
833 cooperation ($\alpha=0.9$) is assumed in (bi) and (ci). Part (i) shows that trait distortion will
834 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.

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

858 be suppressed and individual fitness maximisation (the black cross) is regained

859 (dashed arrows).

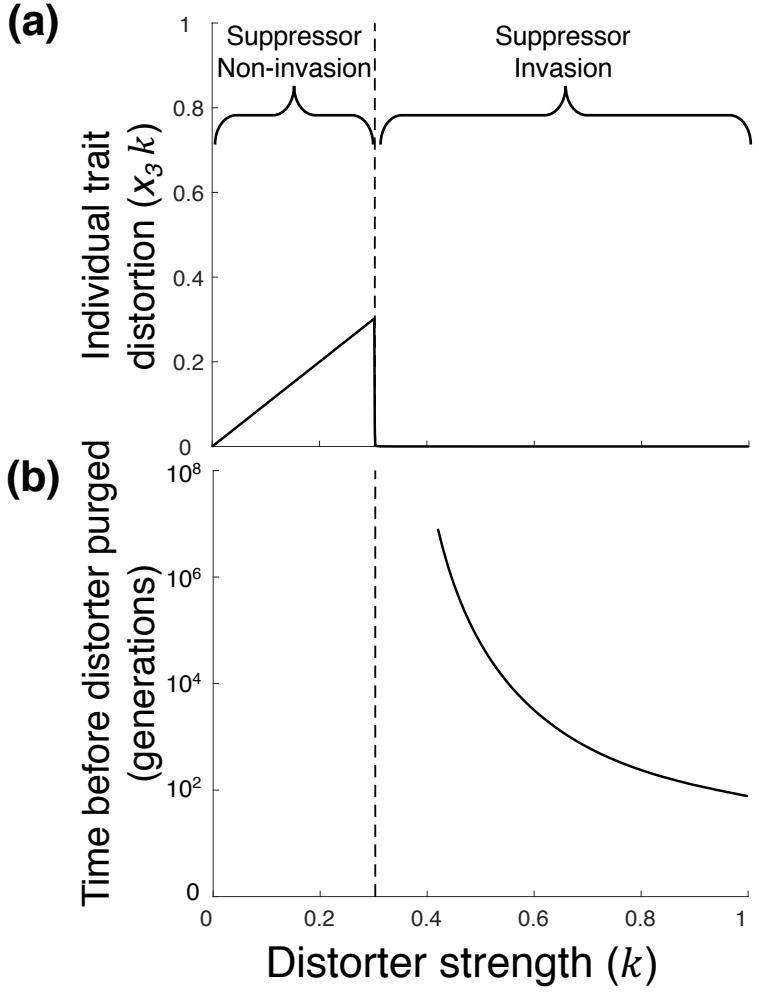
(a)
Suppressor
Non-invasion
Suppressor
Invasion
Individual trait
distortion ($x_3\,k$)
1
0.8
0.6
0.4
0.2
0
(b)
Time before distorter purged
(generations)
10^8
10^6
10^4
10^2
0
0
0.2
0.4
0.6
0.8
1
Distorter strength (k)

Figure 1

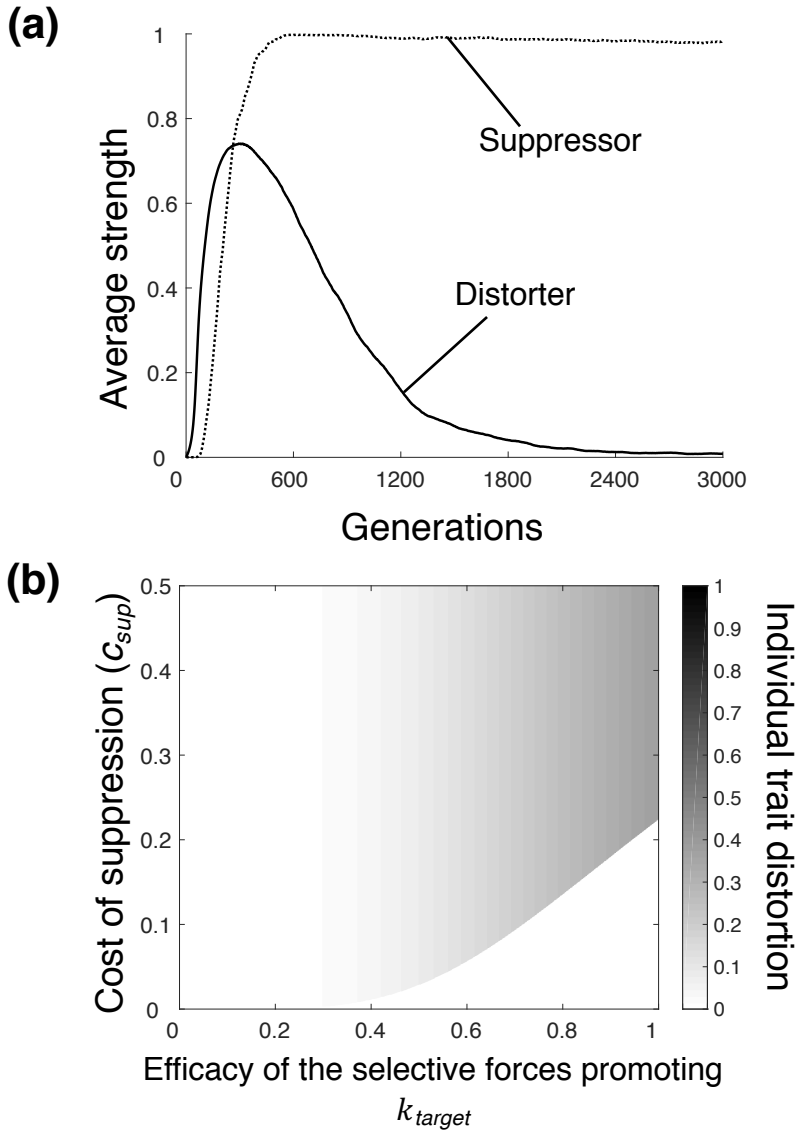
(a)
Average strength
Generations
Suppressor
Distorter
0
0.2
0.4
0.6
0.8
1
0
600
1200
1800
2400
3000
(b)
Cost of suppression (c_{sup})
Efficacy of the selective forces promoting
k_{target}
Individual trait distortion
0
0.1
0.2
0.3
0.4
0.5
0
0.2
0.4
0.6
0.8
1
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1

Figure 2

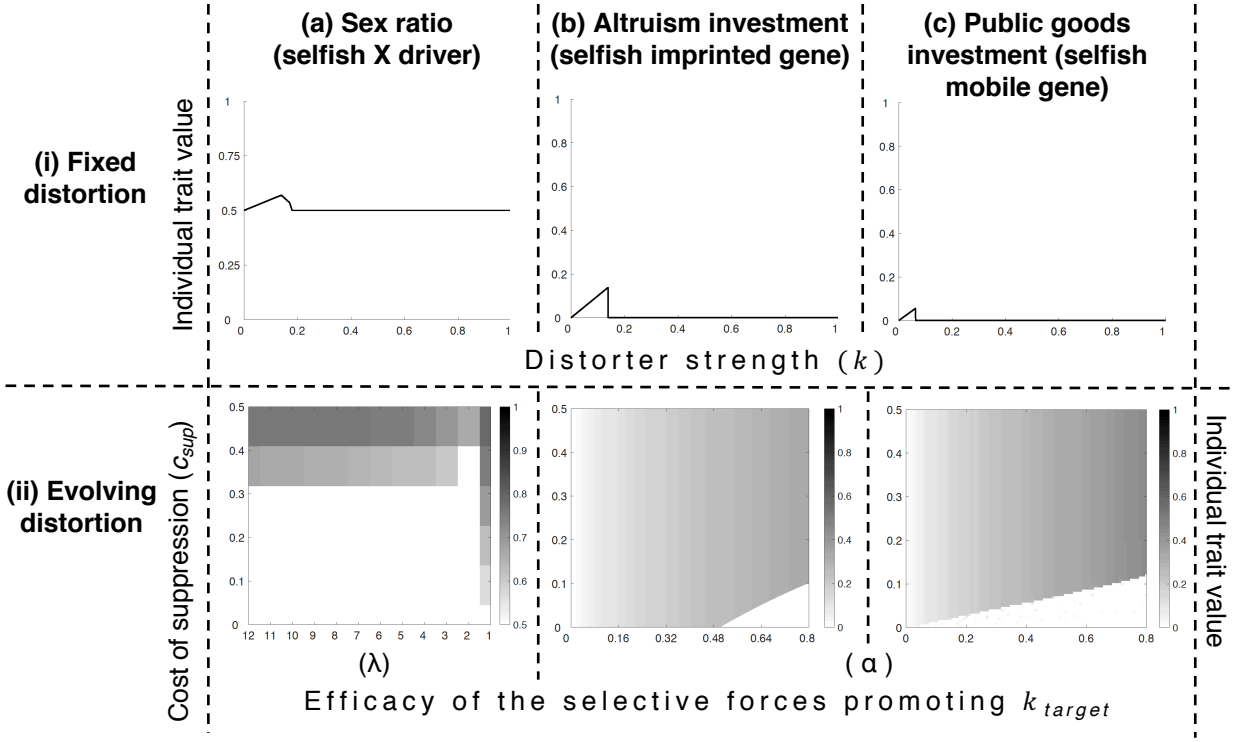
(a) Sex ratio (selfish X driver)
(b) Altruism investment (selfish imprinted gene)
(c) Public goods investment (selfish mobile gene)
(i) Fixed distortion
(ii) Evolving distortion
Individual trait value
Distorter strength (k)
Cost of suppression (c_{sup})
(λ)
(α)
Efficacy of the selective forces promoting k_{target}
Individual trait value

Figure 3

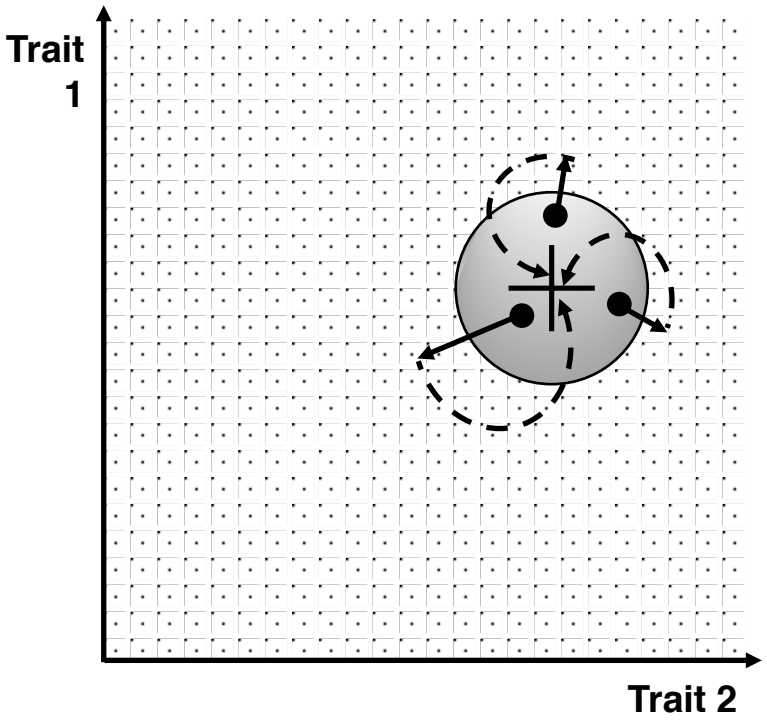

Figure 4