

RESEARCH ARTICLE

# Biophysics and population size constrains speciation in an evolutionary model of developmental system drift

Bhavin S. Khatri<sup>1†\*</sup>, Richard A. Goldstein<sup>2</sup>

**1** The Francis Crick Institute, London, United Kingdom, **2** Division of Infection & Immunity, University College London, London, United Kingdom

† Current address: Department of Life Sciences, Imperial College London, Silwood Park, Ascot, United Kingdom

\* [bhavin.khatri@physics.org](mailto:bhavin.khatri@physics.org)



**OPEN ACCESS**

**Citation:** Khatri BS, Goldstein RA (2019) Biophysics and population size constrains speciation in an evolutionary model of developmental system drift. *PLoS Comput Biol* 15(7): e1007177. <https://doi.org/10.1371/journal.pcbi.1007177>

**Editor:** Ville Mustonen, University of Helsinki, FINLAND

**Received:** November 30, 2018

**Accepted:** June 13, 2019

**Published:** July 23, 2019

**Copyright:** © 2019 Khatri, Goldstein. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All relevant data are within the manuscript and its Supporting Information files.

**Funding:** RAG was supported by the Medical Research Council under grant U117573805 and BSK by The Francis Crick Institute which receives its core funding from Cancer Research UK, the UK Medical Research Council and the Wellcome Trust. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

## Abstract

Developmental system drift is a likely mechanism for the origin of hybrid incompatibilities between closely related species. We examine here the detailed mechanistic basis of hybrid incompatibilities between two allopatric lineages, for a genotype-phenotype map of developmental system drift under stabilising selection, where an organismal phenotype is conserved, but the underlying molecular phenotypes and genotype can drift. This leads to number of emergent phenomenon not obtainable by modelling genotype or phenotype alone. Our results show that: 1) speciation is more rapid at smaller population sizes with a characteristic, Orr-like, power law, but at large population sizes slow, characterised by a sub-diffusive growth law; 2) the molecular phenotypes under weakest selection contribute to the earliest incompatibilities; and 3) pair-wise incompatibilities dominate over higher order, contrary to previous predictions that the latter should dominate. The population size effect we find is consistent with previous results on allopatric divergence of transcription factor-DNA binding, where smaller populations have common ancestors with a larger drift load because genetic drift favours phenotypes which have a larger number of genotypes (higher sequence entropy) over more fit phenotypes which have far fewer genotypes; this means less substitutions are required in either lineage before incompatibilities arise. Overall, our results indicate that biophysics and population size provide a much stronger constraint to speciation than suggested by previous models, and point to a general mechanistic principle of how incompatibilities arise the under stabilising selection for an organismal phenotype.

## Author summary

The process of speciation is of fundamental importance to the field of evolution as it is intimately connected to understanding the immense bio-diversity of life. There is still relatively little understanding of the underlying genetic mechanisms that give rise to hybrid incompatibilities with results suggesting that divergence in transcription factor DNA binding and gene expression play an important role. A key finding from the field of evo-

**Competing interests:** The authors have declared that no competing interests exist.

devo is that organismal phenotypes show developmental system drift, where species maintain the same phenotype, but diverge in developmental pathways; this is an important potential source of hybrid incompatibilities. Here, we explore a theoretical framework to understand how incompatibilities arise due to developmental system drift, using a tractable biophysically inspired genotype-phenotype for spatial gene expression. Modelling the evolution of phenotypes in this way has the key advantage that it mirrors how selection works in nature, i.e. that selection acts on phenotypes, but variation (mutation) arise at the level of genotypes. This results, as we demonstrate, in a number of non-trivial and testable predictions concerning speciation due to developmental system drift, which would not be obtainable by modelling evolution of genotypes or phenotypes alone.

## Introduction

The detailed genetic mechanisms by which non-interbreeding species diverge is still poorly understood. Darwin, inspired by John Herschel, called it that “mystery of mysteries” [1]; he struggled to understand how natural selection could give rise to hybrid inviability or infertility between populations without producing such incompatibilities within the populations. A solution to this problem was conceived independently by Dobzhansky, Muller and Bateson in which cross-mating would combine alleles at different loci that are incompatible due to epistatic interactions (Dobzhansky Muller incompatibilities, DMI) [2–4]. Consider a common ancestor with alleles  $ab$  across two loci, which after a period of allopatric divergence give rise to two lineages which have fixed genotypes  $Ab$  and  $aB$ , respectively. Interbreeding between these two populations would result in the heterozygotic hybrid genotype  $AaBb$ , combining the potentially incompatible  $A$  and  $B$ , a combination that could not arise in either population mating separately.

Assuming that any combination of alleles that have not been “tested” by the process of evolution represents a potential incompatibility, Orr predicted that the number of incompatibilities between pairs of alleles in a sufficiently large genome would increase with the number of substitutions separating the two lineages ( $K$ ) as  $K(K-1) \sim K^2$  [5]. Similarly, the number of untested combinations involving  $n$  loci would increase as  $\sim K^n$ , suggesting that, with evolutionary time, potential incompatibilities would become increasingly dominated by more complex epistatic interactions [5]. This would occur, firstly, because there are a larger number of combinations, and secondly, because there are more ways for separate lineages to evolve around incompatible genotypes when there is a larger number of loci. It is unclear, however, how the simplistic assumptions of this model fare with increased biological realism. Not all possible untested hybrids are equally likely to result in real incompatibilities. Selection acting on each separate lineage affects the substitutions that occur and their likely contributions to reproductive isolation. In particular, evolutionary constraints have a strong effect on the development of more complex DMIs, making it uncertain whether their role is as important as suggested by Orr’s combinatorial argument. This highlights the need for considering more realistic models that better capture the salient aspects of the underlying biology, whilst remaining sufficiently simple for tractable evolutionary modelling and simulation.

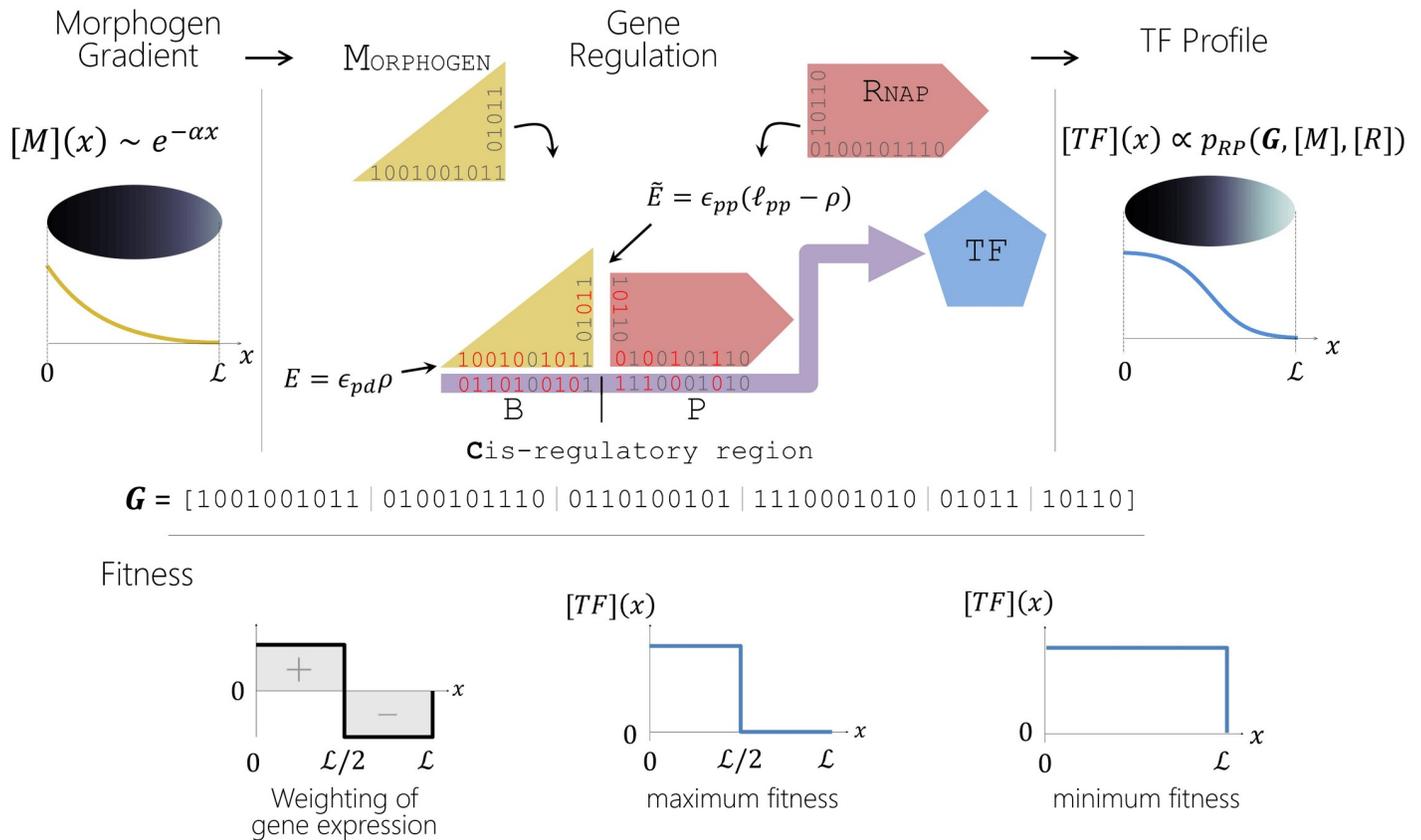
In recent years a form of epistasis has been described in a number of organisms whereby closely related species have similar organismal phenotypes but are produced by very different developmental mechanisms [6–9]. This cryptic “developmental system drift” [10, 11] could be an important source of hybrid incompatibilities that cause reproductive isolation [12, 13]. Developmental system drift is an example of a more general characteristic of biological systems

where many genotypes can correspond to the same phenotype; this redundancy of the mapping from genotype to phenotype results in a number of non-trivial behaviours which do not arise on fitness landscapes which consider evolution of phenotypes or genotypes independently [14–25]. The degree of redundancy can be represented as the “sequence entropy”, corresponding to the log of the number of genotypes corresponding to a given phenotype, in analogy to the similar expression in statistical mechanics [17, 25–27].

To explore the role of developmental system drift on speciation, we examine the growth of Dobzhansky-Muller incompatibilities using a simple genotype-phenotype map that models the development of spatial patterning of gene expression. The model, introduced by [17] allows for cryptic genetic variation and changes in molecular phenotypes while maintaining organismal phenotype under stabilising selection. In addition, we introduce a novel computational method to decompose hybrid DMIs so we can examine the behaviour of the fundamental pair-wise and higher order incompatibilities. We show that including biologically relevant elements gives rise to a number of novel phenomenon that could not arise with models based only on the fitness of genotypes or phenotypes. Our results show that small populations develop hybrid incompatibilities more quickly, due to the pressure of sequence entropy in small populations meaning the common ancestor harbours on average a larger drift load. For large populations, we find hybrid incompatibilities arise more slowly, with a growth law characteristic of a sub-diffusion of the hybrid binding energies, indicative of kinetic traps in the molecular substitution process due to roughness to the fitness landscape [17]. Strikingly, we find that for moderate population sizes it is the molecular phenotypes under weakest selection that give rise to earliest incompatibilities, since in the common ancestor they are more likely to be already maladapted. Finally, we find that unlike Orr’s prediction that complex DMIs should be abundant, pair-wise interactions between loci dominate the growth of DMIs, showing that biophysics provides a stronger constraint than pure combinatorics.

### A simple genotype-phenotype map to model developmental system drift

The genotype-phenotype map we use is a modification of the one described in [17]. The evolutionary task set for the gene regulation module is to turn an exponentially decaying morphogen gradient ( $M$ ) across a field of cells in an embryo into a sharp step function profile of a downstream transcription factor  $T$  with its transition at the mid-point of the embryo, as shown in Fig 1. This is accomplished by having the morphogen and an RNA Polymerase  $R$  bind to two adjacent non-overlapping binding sites in the cis-regulatory region ( $C$ ) region of the transcription factor, the promoter  $P$ , and a single binding site  $B$  adjacent to it; transcription occurs whenever the polymerase binds to the promoter, although both proteins can bind to both binding sites dependent on their binding affinities. Binding to the regulatory region is cooperative due to stabilising interactions between the two proteins when bound at the two adjacent sites. The sequences of  $M$  and  $R$  at the DNA binding sites are represented by binary strings of length  $\ell_{pd} = 10$ . The corresponding DNA binding sites  $B$  and  $P$  are also represented by binary strings of the same length. Interactions between a pair of proteins are similarly represented by binary strings of length  $\ell_{pp} = 5$ . We assume an exponential morphogen concentration profile  $[M](x, \alpha)$ , as a function of the position of embryonic cells,  $x$ ; the decay rate of the morphogen  $\alpha$  is represented as a continuous variable, with a relative probability of mutation corresponding to an effective string of length  $\ell_\alpha = 10$  bases. This results in a genome  $G$ , of total length  $\ell_G = 60$ . Protein-DNA and protein-protein binding strengths are determined by the number of mismatches between corresponding strings on the two interacting molecules, where for protein-DNA binding the cost of a mismatch is  $\epsilon_{pd} = 2k_B T$  and for protein-protein interactions  $\epsilon_{pp} = 1k_B T$ , where  $k_B T$  is Boltzmann’s constant multiplied by room temperature



**Fig 1. An overview of the genotype-phenotype map.** The gene regulatory module has as input a morphogen concentration  $[M](x)$  that varies approximately exponentially across a 1-dimensional embryo of length  $L$ , and outputs a transcription factor  $[T](x)$ . Spatial gene regulation of  $T$  is achieved via a simple Hamming model of protein DNA binding as shown. Fitness is determined by weighting any gene expression in the anterior half of the embryo as positive, and any in the posterior half as negative (as shown bottom left). We show idealised expression profiles that give maximum fitness when gene expression is confined only to the anterior half (bottom middle), and minimum fitness when there is no discrimination between anterior and posterior (bottom right); these translate to log-fitnesses of  $F = 0$  and  $F = -\infty$ , respectively (see Methods).

<https://doi.org/10.1371/journal.pcbi.1007177.g001>

(298K). We assume that there is a fixed concentration  $[R]$  of polymerase, in each cell. We then follow [28] and assume that the concentration of the transcription factor in a cell at position  $x$  ( $[T](x)$ ) is simply proportional to the probability of the polymerase being bound to the promoter, where this is calculated using standard methods of equilibrium statistical mechanics allowing for all configurations of protein species bound at these two binding sites, as well as none being bound (see Methods for details). The fitness contribution  $F$  of the overall patterning phenotype ranges from 0 to  $\kappa_F$  depending on how well expression of the transcription factor is confined to the anterior half of the embryo, as shown in Fig 1 (bottom left), where  $\kappa_F$  is a measure of the relative contribution of this trait to the fitness of the organism. We define a population-scaled fitness contribution  $2N_e\kappa_F$ , where  $N_e$  is the effective population size; for  $2N_e\kappa_F < 1$  the effects of selection are weak, and are conversely strong when  $2N_e\kappa_F > 1$ . We also assume that there is a boundary at  $F = F^*$ , below which the organism is unviable. We simulate evolution as continuous time Markov process. After evolving a single population for a given number of generations, we form two replicates of the population that evolve independently, representing the process of allopatric speciation. At various time points following this imposed isolation, we consider the fitness and viability of various outcrossings between the two populations. A DMI occurs when the fitness contribution of a particular hybrid drops below  $F^*$ .

## Results

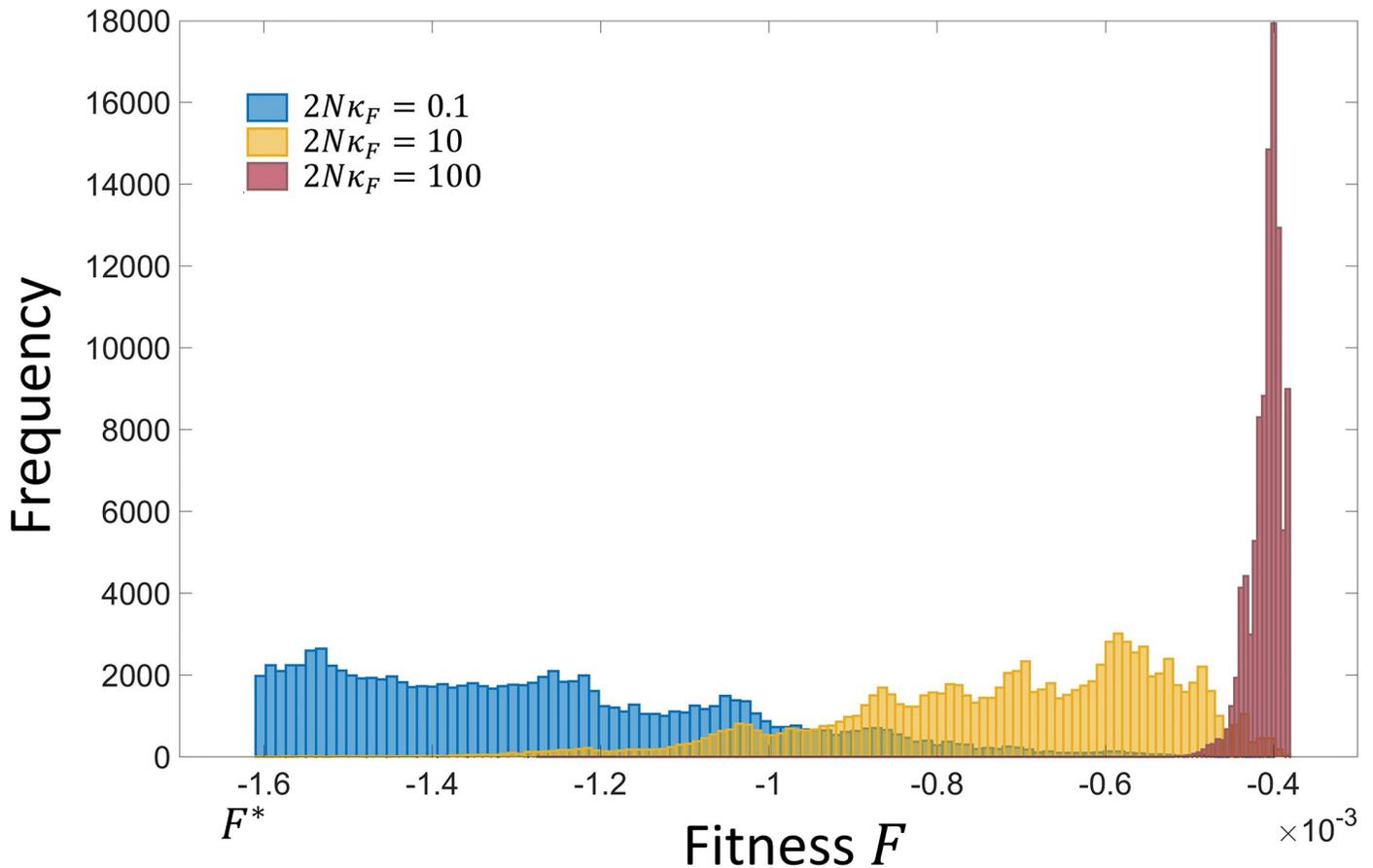
### Evolutionary properties of genotype-phenotype map on each lineage

The properties of a similar genotype-phenotype map have been previously explored [17]. An important property of this genotype-phenotype map is that only a single mechanism of patterning is found, in which the polymerase ( $R$ ) binds with intermediate affinity to the promoter ( $P$ ) but with high affinity to the morphogen ( $M$ ), while the morphogen binds to the morphogen binding site ( $B$ ) only above a critical morphogen concentration. This results in a spatial switch once the morphogen falls below this concentration; evolution then fine tunes the relationship between the protein-DNA binding energies, the protein-protein binding energy and the steepness of the morphogen gradient  $\alpha$  to turn off transcription at the mid-point of the embryo. Despite a single global solution there are many different combinations of the protein-DNA and protein-protein binding energies and  $\alpha$  that give good patterning, and each of these correspond to many possible genotypes ( $G$ ). Of the different possible binding energies, we find that  $E_{MB}$ ,  $E_{RP}$ ,  $\tilde{E}_{RM}$  (binding energy of  $M$  to  $B$ ,  $R$  to  $P$ , and  $R$  to  $M$ , respectively, calculated by the equivalent of Eq 3 in the Methods) are under strong selection, whilst the other possible binding energies are essentially neutral with weak selective effects. At large population sizes it is found that the evolutionary dynamics exhibits what is known as quenched-disorder in statistical physics, where energy phenotypes that are less constrained take different *random* values between independent evolutionary runs with no further substitutions; this indicates an underlying roughness to the fitness landscape and that these weakly selected traits are trapped in a local optimum [17].

A key property determining the rate at which incompatibilities arise is the distribution of common ancestor phenotypes as a function of the population-scaled fitness contribution  $2N_e\kappa_F$ , as shown in Fig 2. For a given value of  $\kappa_F$ , we see that for large population sizes ( $2N_e\kappa_F \gg 10$ ) the distribution is what we would expect from conventional evolutionary theory on a fitness landscape with a fitness maximum. In contrast, as the population size is decreased, we find the distribution shifts to lower fitness values to the point when selection is weak ( $2N_e\kappa_F \leq 1$ ) the distribution is poised at the inviability boundary. This effect arises due to genetic drift at low population sizes pushing populations towards marginally fit phenotypes that correspond to the largest number of genotypes, that is, with the largest sequence entropy.

### Decomposition of DMIs

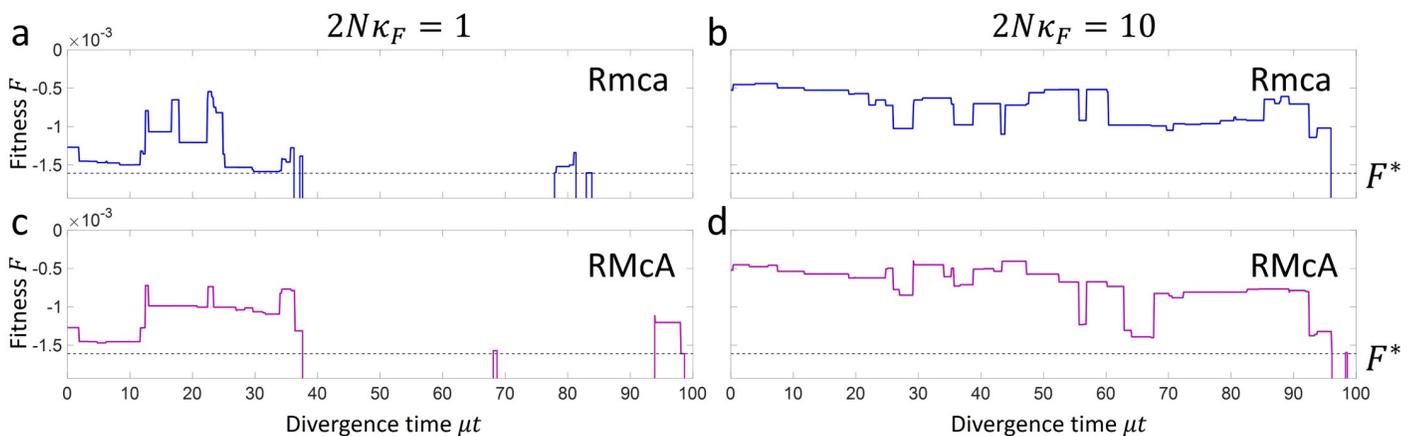
Our genome is composed of 4 loci: 1) the  $R$  locus corresponding to the polymerase sequence, 2) the Morphogen ( $M$ ) locus, 3) the  $C$  locus which corresponds to the sequences for the cis-regulatory region of the transcription factor and 4) the  $\alpha$  locus, which is the morphogen gradient steepness  $\alpha$ . Hybrids between the two lineages are constructed by independent re-assortment of these loci assuming complete linkage within each locus and no linkage between them. We define a hybrid genotype by a 4 letter string where each letter corresponds to one of the loci defined above and takes one of two cases correspond to whether the allele is from the 1st line or 2nd line; for example, the hybrid  $rMCa$  corresponds to  $R$  locus having an allele from the 1st lineage,  $M$  locus with the allele from the 2nd lineage, the transcription factor (cis-regulatory)  $C$  locus from the 2nd lineage and  $\alpha$  locus the allele from the 1st lineage. Note that the underlying sequence of each hybrid changes as different substitutions are accepted in each lineage; the notation only refers to alleles fixed at any point in time. We can represent all combinations of the four loci drawn from the two parents ( $RMCA$ ,  $RMCa$ ,  $RMcA$ , etc.) as points on a four-dimensional Boolean hypercube. In total there are  $2^4 - 2 = 14$  hybrids.



**Fig 2. Histogram of the fitness of populations over single long runs as a function of the population scaled fitness contribution  $2N\kappa_F$ .**

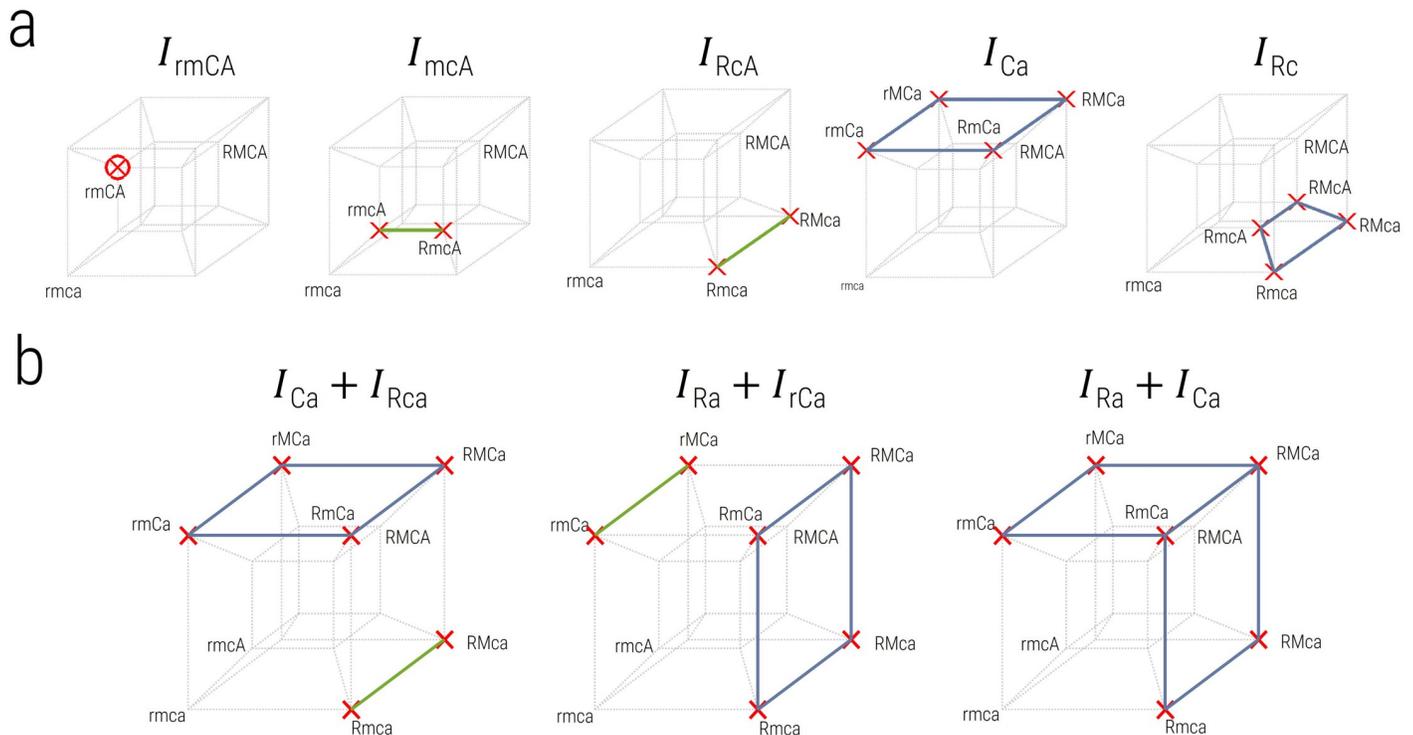
<https://doi.org/10.1371/journal.pcbi.1007177.g002>

In Fig 3, we plot a typical time series of how the fitness of two different hybrids ( $R_{mca}$ (a & b) and  $R_{McA}$ (c & d)) changes over a divergence time  $\mu t$  separating a pair of lineages, for  $2N\kappa_F = 1$  (a & c) and  $2N\kappa_F = 10$  (b & d), where  $\mu = \ell_G \mu_0$  is the mutation rate for all base pairs in all loci.  $2N\kappa_F > 1$  indicates strong selection, whilst  $2N\kappa_F \leq 1$  indicates weak selection where



**Fig 3. Plot of the times series of two hybrids  $R_{mca}$ (a & b) and  $R_{McA}$ (c & d) at population scaled fitness contributions of  $2N\kappa_F = 1$  (a & c) and  $2N\kappa_F = 10$  (b & d). Times when the fitness of a hybrid drop below the critical value  $F^*$  (dashed line) correspond to DMIs.**

<https://doi.org/10.1371/journal.pcbi.1007177.g003>

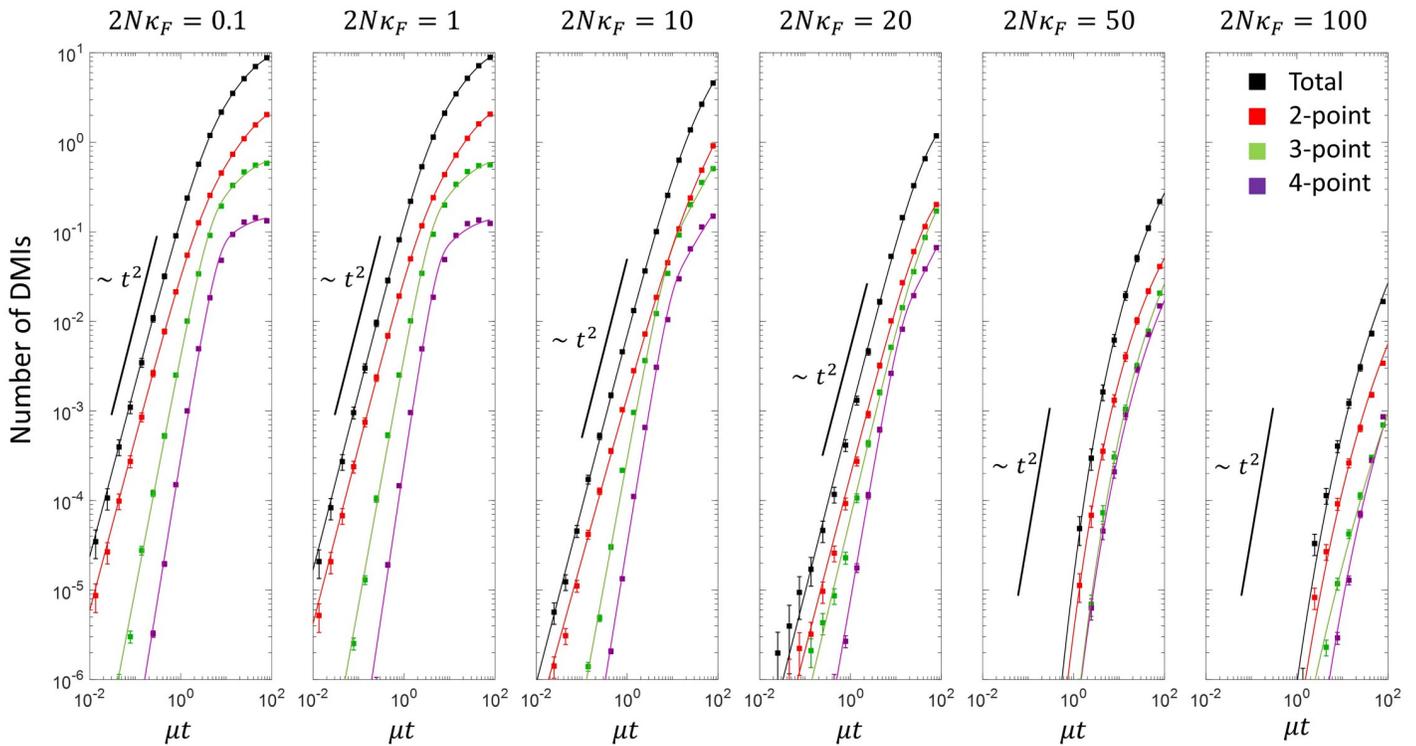


**Fig 4. Decomposition of hybrid DMIs on a Boolean hyper 4-cube.** Each point on the 4-cube represent each possible hybrid genotype across 4 loci, including the genotype of each parental lineage, where each red cross represents an incompatible hybrid genotype. As shown in a) the fundamental types of DMIs, where a blue square or face identify a subspace of genotypes that correspond to a 2-way DMI, a single green edge or line corresponds to a 3-way DMI and a red open circle corresponds to a single isolated 4-way DMI. b) A more complicated pattern of hybrid DMIs and their decomposition into fundamental types.

<https://doi.org/10.1371/journal.pcbi.1007177.g004>

genetic drift dominates (For reference, in human populations it has been estimated that  $\approx 20 - 30\%$  of mutations are weakly selected [29, 30], compared to in *Drosophila*  $< 10\%$  [30]). We see that the fitness of hybrids generally decreases in a stochastic fashion; when the log-fitness of a hybrid drops below the threshold  $F^*$  (indicated by the dashed line), a DMI arises as is indicated by a vertical log-fitness line ( $F = -\infty$ ) for that hybrid. As can be seen in Fig 3, at any given time a changing subset of the fourteen possible hybrids might be incompatible.

**Total number and type of DMIs.** To decompose the pattern of hybrid DMIs into fundamental pair-wise, 3-way and 4-way incompatibilities, we use a parsimonious method (Material & Methods) which finds the minimum total number of fundamental DMIs types that can explain the pattern of DMIs observed. As shown in Fig 4a, on a 4-dimensional Boolean hypercube, a pair-wise incompatibility is represented by a face of the cube (blue square), a 3-way incompatibility by a single edge (green line) and a 4-way incompatibility by a single point (red circle). This arises because, for example, as in Fig 4a, a 2-point DMI  $I_{Ca}$  means any genotype which features this sub-genotype  $Ca$  must, by definition, also be a DMI; the alleles on the remaining loci not involved in the DMI, can take any value and so form a 2D subspace, which is a whole face of the Boolean 4-cube. Similarly, a 3-point DMI, such as  $I_{mcA}$  as shown, constrains the sub-genotype of 3 loci to be incompatible and the remaining locus can take any value forming an edge ( $r \rightarrow R$  in the example) of a 4D Boolean hypercube. A 4-point DMI must be a single point in a 4D Boolean hypercube (e.g.  $I_{rmCA}$  as shown). In Fig 4b, we show an example where the pattern of hybrid incompatibilities, shown by red crosses, can be explained in three different ways, each with only 2 DMIs: 1)  $I_{Ca} + I_{RcA}$ , 2)  $I_{Ra} + I_{rCa}$ , 3)  $I_{Ca} + I_{Ra}$ ; two is the minimum number of DMIs, as there is no way to explain the pattern seen with a single



**Fig 5. Plot of the total number of DMIs vs divergence time, together with their decomposition into the total number of 2-way, 3-way, 4-way DMIs, for various scaled population sizes.** For  $2N\kappa_F \leq 20$  the solid lines correspond to fits of the simulation data to Eq 1, while for  $2N\kappa_F \geq 50$  correspond to fits to Eq 2.

<https://doi.org/10.1371/journal.pcbi.1007177.g005>

pair-wise, 3-way or 4-way DMI. We assume each of these is, a priori, equally likely and so the total number of pair-wise incompatibilities is calculated as an average over the different ways we can explain the observed pattern of hybrid DMIs; in the example in Fig 4b, we therefore have a total number of 2-way DMIs of  $n_2 = 4/3$  and 3-way DMIs  $n_3 = 2/3$ , giving a total number of  $n = n_2 + n_3 = 2$  DMIs, which is the minimum number of DMIs needed to explain the pattern of hybrid incompatibilities.

Using this method, we plot the total number of each type of DMI versus divergence time in Fig 5, where the panels correspond to different population scaled fitness contributions from  $2N_e\kappa_F = 0.1$  to  $2N_e\kappa_F = 100$ . We see clearly that as the population size is decreased the rate at which incompatibilities arise increases. This effect results from the shift in distribution of fitness to poorer values for smaller populations in Fig 2; this means that the common ancestor is more likely to be slightly maladapted (higher mutational load) for smaller populations, and so fewer substitutions are required before hybrid incompatibilities arise. This effect was also found for simple models of protein-DNA binding [26, 27], but arose as a result of direct selection for high binding affinity, which also corresponds to low sequence entropy as there are exponentially fewer ways of binding with no or only a few mismatches between protein and DNA. In this work there is only selection on the overall patterning phenotype, yet remarkably we find analogous behaviour due to the effect of sequence entropy of this phenotype.

**Pair-wise incompatibilities dominate reproductive isolation.** As is clearly shown by Fig 5, pair-wise incompatibilities dominate for all population sizes, though at larger population sizes the difference is smaller, and particularly at shorter times. These results should be contrasted with the prediction of Orr that 2-way: 3-way: 4-way incompatibilities should arise in

the ratio 12 : 24 : 14 (for  $L = 4$  loci); so for example, there should be double the number of 3-way incompatibilities to pair-wise.

It is possible that this bias could arise by the development of a random set of incompatibilities that would preferentially decompose into sets of 2-way DMIs. In order to consider this possibility, we randomly assigned a set of incompatibilities and analysed them using our decomposition procedure. Contrary to the observed results, we find that this decomposition results in the largest number of 4-way DMIs, followed by 3-way and then 2-way DMIs (S1 Text). This is as expected as when the probability that a hybrid is incompatible is small, by random chance, we should expect to see isolated incompatible hybrid genotypes, which would be explained by 4-way DMIs; the fact that our results show that 2-way DMIs are dominant, even at early times, means that whole faces of the Boolean hypercube are found to incompatible, which is not likely to arise in a random model. Hence overall, these results show that contrary to Orr’s prediction that higher order DMIs should be easier to evolve, higher order DMIs evolve more slowly and are in smaller number compared to pair-wise DMIs.

**Quantification of power law growth of DMIs for small populations.** We find that for small population sizes 2-way, 3-way and 4-way DMIs all increase as a power law at small times, indicated by a straight line on a log-log plot, in agreement with [5, 31], who predicted that  $n$ -way DMIs should increase as  $\sim t^n$ . To more quantitatively assess the exponent, we fit the data for  $2N\kappa_F \leq 20$  using the phenomenological equation:

$$I(t) = \frac{I_0 t}{T + t} (1 - \exp - (t/\tau)^{\gamma-1}) \tag{1}$$

which has the asymptotic form of  $I(t) \sim t^\gamma$  for  $t \ll \tau$  and  $t \ll T$  and an opposite limit of  $I(t \rightarrow \infty) = I_0$ . We see that for the total number of DMIs and for 2-way, 3-way and 4-way DMIs, this form fits the data well at intermediate and small population sizes. We tabulate the power law exponent derived from these fits in Table 1. We see that the total number of DMIs and 2-way DMIs have a power law exponent close to  $\gamma = 2$ , which is consistent with the Orr model. The 4-way DMIs have a larger exponent than the 3-way DMIs, which is larger than the 2-way DMIs, also in agreement with Orr’s model, although the values of the exponent are substantially lower than  $n$  for the 3-way and 4-way DMIs.

**Sub-diffusive growth of DMIs for large populations.** For large population sizes there is no clear power law, which is again consistent with previous simulations [27] and also theoretical calculations [26]; when high fitness corresponds to high binding affinity, so that the common ancestor distribution is peaked away from the inviability boundary, large populations have a small drift load, meaning that DMIs arise when hybrid energy traits diffuse to the boundary. One such analytically tractable model was investigated in [26] and predicted that the number of DMIs is a complementary error function, which has an asymptotic form  $I(t) \sim \frac{\sqrt{4\mu t}}{K^*} e^{-(K^*)^2/4\mu t}$ , which due to the essential singularity as  $t \rightarrow 0$  has the property of negative curvature on a log-log plot. However, neither this form nor its multidimensional

**Table 1. Table of values of the exponent  $\gamma$  characterising the power law of growth of DMIs at short times and small scaled population sizes.**

$2N\kappa_F$		0.1	1	10	20
Total	$\gamma$	$1.94 \pm 0.05$	$1.99 \pm 0.05$	$1.95 \pm 0.05$	$2.00 \pm 0.12$
2-way		$1.93 \pm 0.05$	$1.98 \pm 0.06$	$1.84 \pm 0.06$	$1.97 \pm 0.13$
3-way		$2.66 \pm 0.13$	$2.81 \pm 0.17$	$2.76 \pm 0.24$	$2.14 \pm 0.15$
4-way		$3.17 \pm 0.19$	$3.43 \pm 0.41$	$3.14 \pm 0.24$	$2.93 \pm 0.19$

<https://doi.org/10.1371/journal.pcbi.1007177.t001>

generalization fit the simulation data well (not shown). A functional form that is a good fit to the data at large populations sizes is

$$I(t) = \frac{\sqrt{4(\mu t)^\beta}}{K^*} e^{-(K^*)^2/4(\mu t)^\beta}, \tag{2}$$

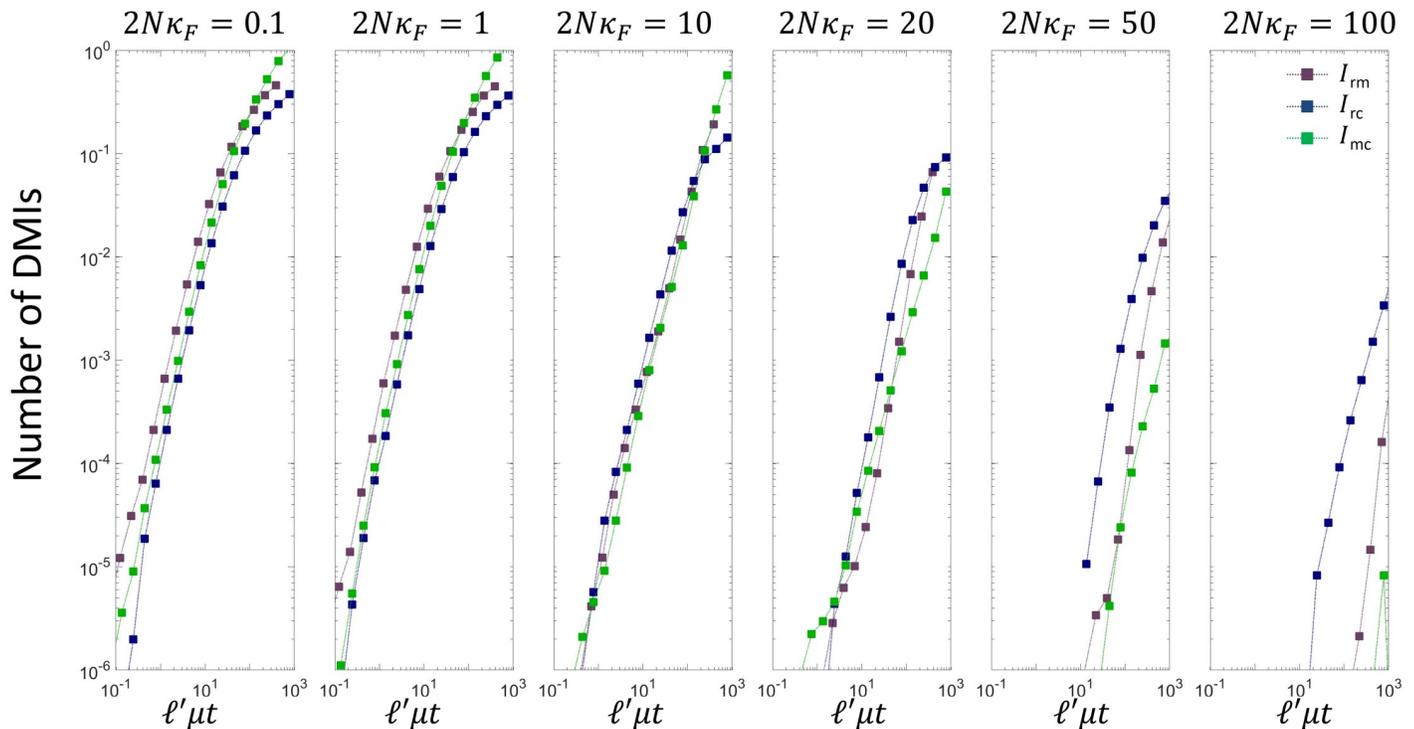
which arises when considering fractional Brownian processes with exponent  $\beta$  [32]; normal diffusion or Brownian motion arises when  $\beta = 1$ , while  $\beta < 1$  corresponds to *subdiffusive* behaviour, while  $\beta > 1$  is *superdiffusive*. It is clear by examining the exponent  $\beta$  in Table 2 from fits of the data in Fig 5 at large population size ( $2N\kappa_F \geq 50$ ) that the DMIs arise as a result of a sub-diffusive process, where 2-way, 3-way and 4-way DMIs have an exponent  $\beta \approx 1/3$  for  $2N\kappa_F = 50$  and  $\beta \approx 1/4$  for  $2N\kappa_F = 100$ . The most likely mechanism that would give rise to sub-diffusive behaviour is a broad spectrum of times between substitutions; even though in the simulations the kinetic Monte Carlo scheme is based on a Poisson process for a given genotypic state  $G$ , the distribution of rates could vary significantly as populations explore the fitness landscape. This would be consistent with the results in [17], which reveal the underlying fitness landscape of this spatial patterning genotype-phenotype map to be rough, which could lead to broad distribution of substitution rates in each lineage and effective sub-diffusive behaviour of the hybrids. Finally as expected the average number of substitutions needed at large population sizes is large, with values of  $K^*$  ranging from 6 to 9, and increases with increasing population size, as expected; it also increases very moderately with increasing  $n$ , which would be consistent with an increase in dimensionality of higher order DMIs. Interestingly, these values of  $K^*$  would indicate that the fraction of viable genotypes is very small,  $\sim 2^{K^*} / 2^{\ell_G} \sim 10^{-13}$ , where  $\ell_G = 60$  is the number of effective binary sites in this genotype-phenotype map.

**Molecular phenotypes under weakest selection dominate DMIs at larger population sizes.** The DMI decomposition algorithm allows examination of the behaviour of different types of pair-wise, 3-way and 4-way DMIs. The different pair-wise DMIs can easily be identified with the different molecular binding energy phenotypes that combine together to form the overall organismal phenotype. The plot in Fig 6 shows how the number of DMIs grows for the three most dominant pair-wise DMIs  $I_{r_m}$ ,  $I_{r_c}$  and  $I_{m_c}$  (pair-wise DMIs involving the  $\alpha$  locus are orders of magnitude smaller—see S1 Text), as a function of the scaled divergence time  $\ell' \mu t$ , where  $\ell' = 5$  for  $I_{r_m}$  and  $\ell' = 10$  for  $I_{r_c}$  and  $I_{m_c}$ , cancelling how mutations are more likely to hit longer regions of the genome [27]. Strikingly, we find that contrary to what might seem intuitive, it is the molecular interactions under weakest selection that dominate the number of DMIs, particularly for larger population sizes. For  $2N\kappa_F \gg 50$ , we see at early times  $I_{r_c} > I_{r_m} > I_{m_c}$ ,

**Table 2. Table of values of the parameters characterising the sub-diffusive growth of DMIs for large scaled population sizes;  $\beta = 1$  corresponds to normal diffusive motion,  $\beta < 1$  to sub-diffusion and  $\beta > 1$  super-diffusion, while  $K^*$  corresponds roughly to the number of substitutions required to reach the inviable region.**

$2N\kappa_F$		50	100
Total	$\beta$	0.47 ± 0.03	0.33 ± 0.02
	$K^*$	6.58 ± 0.12	7.37 ± 0.26
2-way	$\beta$	0.32 ± 0.01	0.25 ± 0.01
	$K^*$	6.71 ± 0.07	7.51 ± 0.18
3-way	$\beta$	0.33 ± 0.01	0.22 ± 0.01
	$K^*$	7.53 ± 0.10	8.09 ± 0.27
4-way	$\beta$	0.31 ± 0.02	0.25 ± 0.02
	$K^*$	7.50 ± 0.28	8.66 ± 0.39

<https://doi.org/10.1371/journal.pcbi.1007177.t002>



**Fig 6. Plot of the spectrum 2-way DMIs vs scaled divergence time for different scaled population sizes.** Here the divergence time for each pair-wise DMI is scaled by the number of base-pairs involved in each interaction in order to remove the effect that interactions with a larger number of base-pairs mutate at a larger rate in proportion to their length; for  $I_{rm}$ ,  $\ell' = 5$ , and for  $I_{rc}$  and  $I_{mc}$ ,  $\ell' = 10$ .

<https://doi.org/10.1371/journal.pcbi.1007177.g006>

which is the same order of increasing selective constraint on each of these molecular interactions (S1 Text). This has a simple explanation: at a sufficiently large population size such that none of the selective constraints on different molecular phenotypes is neutral, in the common ancestor the molecular interaction strengths under weakest selection will have the largest drift load and hence fewer substitutions are needed for incompatibilities to arise in hybrids. This is precisely as observed in simple models of transcription factor DNA binding, where the incompatibilities arise more quickly where binding is under weak selection for high affinity [27].

In the converse limit, when population sizes are sufficiently small that all molecular phenotypes are neutral (up to the truncation threshold), we find only small differences between the rise of each of the pair-wise incompatibility types; this is again consistent with simple models of transcription factor DNA binding, where the phenotypic distribution of binding energies is dominated by their sequence entropy and not fitness at small population sizes and there is only a weak dependence on sequence length [27]. In the S1 Text, we show this is also true of the higher order DMIs, where it is purely the sequence-entropy constraints of each DMI type which controls the dynamics of incompatibilities and fitness has little bearing; however, as the population size is increased this degeneracy is lifted and each of the DMIs exhibits individual behaviour.

### Discussion

There is still very little understood about the underlying genetic basis that gives rise to reproductive isolation between lineages. Gene expression divergence is thought to be a strong determinant of the differences between species [33–37] with a growing body of evidence for their

direct role in speciation, particularly through transcription factors [38–42]. In particular, transcription factors mediate the control of gene expression and the ultimately body plans of organisms through complicated gene regulatory networks. Studies in *Drosophila* and nematodes have shown in closely diverged species conserved body plans yet plasticity in the underlying molecular architectures between them [7–11, 43].

This evidence suggests that divergence in the regulatory networks controlling gene expression and body plans in allopatric lineages is likely to be an important mechanism of speciation in many higher organisms, yet we do not yet have a quantitative evolutionary framework to model such processes, which can then be used to make predictions. A key challenge is to link changes at the genetic level, where mutations arise, to their outcomes on changes in phenotype, where selection actually acts. Although in the past century we have made great strides in our understanding of evolution through the modern synthesis of Fisher, Wright, Kimura and many others [44–48], this has focussed on the evolution of either genotypes or phenotypes separately. However, in recent years there has been increasing attention on the development of more realistic fitness landscapes, and how this affects the evolutionary process. For example, in biologically realistic systems, often many genotypes correspond to the same phenotype; this redundancy of the mapping from genotype to phenotype results in a number of non-trivial and emergent behaviours which do not arise on fitness landscapes which consider evolution of phenotypes or genotypes independently [14–24].

### Role of free fitness in genotype-phenotype maps

One important and well explored example is the evolution of transcription factor DNA binding [15, 16, 22, 26, 27, 49, 50], where the genotype-phenotype map from sequence to binding affinity can be explicitly enumerated under simplifying assumptions [51, 52]. These investigations show that for small populations dominated by genetic drift, evolution does not optimise fitness. Rather, there is a trade off between the high fitness of a small number of sequences that bind well and the exponentially larger number of sequences that bind less well. The result is the maximisation of a combination of fitness and the number of sequences that correspond to that phenotype. We can take advantage of analogies with statistical mechanics and represent this combination as the “free fitness”, where the log of the number of sequences is the “sequence entropy” of a phenotype [15, 26, 27, 53]. In this formulation, the effective population size is analogous to an inverse temperature for a physical system connected to a heat-bath, where decreasing population size increases the effect of drift and the importance of sequence entropy relative to fitness.

When the free fitness framework is applied to the role of transcription factor DNA binding in allopatric speciation, our previous work gave rise to a simple prediction: incompatibilities arise more quickly for smaller, drift-dominated, populations [26, 27], supporting previous computational studies by Tulchinsky et al. [54], that showed decreased hybrid fitness for smaller populations. This can be understood as a result of the greater importance of sequence entropy for small populations, resulting in common ancestors with a higher drift load, which are therefore closer to incompatible regions. As a result, fewer substitutions are required for the development of hybrid incompatibilities [26, 27]. Conversely, those transcription factor binding site pairs under weaker selection, at a fixed population size, will give rise to incompatibilities more quickly, as they are more susceptible to drift and in the common ancestor will have a larger drift load.

In this paper, we examine speciation in a more realistic genotype phenotype map. For the first time we examine how incompatibilities arise in allopatry for a simple evolutionary model of developmental system drift, where a higher level organismal spatial patterning phenotype is

maintained by stabilising selection, whilst the underlying molecular binding energy phenotypes and ultimately the sequences that determine them, the genotype, are allowed to drift in the evolutionary simulations. Earlier analyses of this model demonstrated the evolution of a number of non-trivial features such as a balance between fitness and sequence entropy deciding the course of evolution at small population sizes and a roughness to the fitness landscape for phenotypes which have high fitness [17]. Importantly here, unlike in previous works [26, 27] we do not directly select for high binding affinity, but only on the organismal level phenotype, but as we discuss, we find the same population size dependence, as well as a number of other novel phenomenon to the speciation process, which would not be obtainable by modelling selection only at the level of phenotypes or genotypes. The results show that biophysics and effective population size provide a much stronger constraint than previous simple modelling of the dynamics of hybrid incompatibilities would suggest [5, 31].

### Small populations characterised by power law growth of incompatibilities, large populations by diffusive growth

A key result we find is that small populations are characterised by a power law growth of incompatibilities with time, vs large populations a diffusive law (discussed below). Thus we suggest that empirical evidence of power law growth in incompatibilities is a signature of allopatric speciation at small population sizes. The Orr model of the growth of hybrid incompatibilities predicts that incompatibilities grow as a power law of the divergence time between allopatric lineages [5, 31], where the exponent represents the number of genes participating in the interaction (e.g. 2 for a 2-way incompatibility). The results of our model also yield this prediction, but only when populations sizes are sufficiently small. There is, however, an alternative model for the power law behaviour to the combinatoric argument made by Orr. As argued in [27], at small population sizes, where genetic drift is dominant and there is a large drift load, common ancestor populations are poised close to the incompatibility boundary and the growth of DMIs at short times is determined by the likelihood that a few critical substitutions arrive quickly, which is given by a Poisson process; if the critical number of substitutions is  $K^*$  then for short times we would expect  $P_i(t) \sim (\mu t)^{K^*}$  and so given that at least  $n$  substitutions are needed for a  $n$ -way incompatibility, we would expect  $K^* \geq n$ . In this paper, we introduced a new method to decompose DMIs into their fundamental pair-wise, 3- and 4-way incompatibilities, and find that for more complex incompatibilities (more loci involved) the larger the exponent of their power law growth. However, we find the exponents we measure for 3- and 4-way incompatibilities are smaller than the predicted exponents of 3 and 4 respectively. We suggest this could be due, as shown in the S1 Text, to the greater number of higher order DMIs arising just by chance, leading to an overestimation of 3- and 4-way DMIs at short times, where at short times a *smaller* exponent corresponds to a larger number (i.e.  $\tau^{n-1} > \tau^n$  for  $\tau < 1$ , where  $\tau$  is some dimensionless timescale).

Examining the growth of incompatibilities at large population sizes, we see there is a characteristic negative curvature on a log-log plot, predicted theoretically by [26], indicating that, as the number of substitutions needed for incompatibilities is large, the changes in the hybrid traits can be approximated by a diffusion process. However, we find that a simple model of diffusion does not fit the simulation data well; instead a model of sub-diffusion, that would arise if there are a number of kinetic traps giving a broad distribution of substitution times, does fit the data well. This is consistent with the finding that the genotype-phenotype map has a rough fitness landscape, which is only revealed at sufficiently large population sizes [17]. However, it is not clear whether we would observe quenched disorder and sub-diffusive behaviour with

more realistic biophysical models that include a larger alphabet size with 20 amino acids and 4 nucleotides.

### **Incompatibilities arise more rapidly in small populations**

We also find that incompatibilities arise more rapidly in smaller populations, which is an emergent effect due to the genotype-phenotype map, giving a bias in degeneracy of different phenotypes; lower fitness and less sharp patterning organismal phenotypes have many more sequences than higher fitness, sharper patterning, phenotypes. In smaller drift-dominated populations, this means there is bias towards phenotypes of small sequence entropy (log degeneracy) that counteracts the tendency of natural selection to favour phenotypes of high fitness. Consequently, the common ancestor in small populations is more likely to be slightly maladapted and less substitutions are needed before hybrids develop incompatibilities. These predictions are consistent with empirical evidence for an inverse correlation of speciation rates with effective population size; the net rates of diversification from phylogenetic trees [55–57] indicate smaller populations speciate more quickly, as well from inferred times for post-zygotic isolation to arise [58–60], where for example mammals and cichlids, which have effective population sizes of order  $10^4$  [61–64], develop reproductive isolation more quickly than birds, which have effective populations sizes of order  $10^6$  [65]. This model and the similar results obtained for transcription factor DNA binding [26, 27] provide a robust explanation of how stabilising selection can give rise to this population size effect in speciation, which do not require passing through fitness valleys as do models based on the founder effect [66–69].

### **Universal phenomenon? Interplay between sequence entropy and specificity of fit phenotypes**

However, the results for this genotype-phenotype map for developmental system drift are particularly noteworthy compared to the previous results on transcription factor DNA binding, as they are obtained without directly selecting for high affinity, low sequence entropy, binding phenotypes; here we only select on the organismal spatial patterning phenotype, but nonetheless we find small populations develop hybrid incompatibilities more quickly through a similar mechanism of the interplay between fitness, sequence entropy and populations size. Although studies with more complex genotype-phenotype maps will be required, we suggest this points towards a broad principle, where the specificity required of a phenotype to be functional and of high fitness will mean that it will be coded by fewer genotypes. For example, in simple models of protein stability, the empirical observation that all proteins tend to be marginally stable, can be explained by the fact that as the stability of a protein is increased the number of sequences that give that stability decreases rapidly [19]; assuming high fitness corresponds to maximum stability, this phenotype is highly specified, as only a few sequences will meet the requirement that all inter-chain interactions in the protein are favourable.

### **Molecular phenotypes under weakest selection dominate DMIs at larger population sizes**

Another property that emerges from this model not obtainable by simply modelling transcription factor DNA binding is that certain molecular phenotypes are more important than others in giving rise to incompatibilities. One particular feature of this model is that the selective constraints on the different molecular binding energy phenotypes emerge through the evolutionary process of stabilising selection on the organismal phenotype, and are not specified by the model. Most strikingly, and counterintuitively, the model predicts that molecular phenotypes

that are under the weakest selective constraints (but not strictly neutral) dominate by giving rise to the earliest incompatibilities for intermediate and large population sizes. Remarkably, here this emerges as a consequence of stabilising selection on the organismal phenotype and not due to selection imposed for good binding affinity as in previous works [27].

We note that these results have been obtained by changing the population size whilst keeping the strength of selection on the organismal trait  $\kappa_F$  fixed. It would be tempting to use these results to suggest that overall those traits in a genome under weakest selection would give rise to the earliest incompatibilities and hence dominate allopatric speciation. However, the question of how the dynamics of hybrid incompatibilities changes as the strength of selection changes is a subtle one, which we leave to future work; in this model a reduction in  $\kappa_F$  has the effect of changing the phenotypic regions of incompatibility, with non-trivial consequences. It should also be noted, the role of sequence entropy in giving a strong population size dependence to the rate of reproductive isolation; if we consider only a peaked phenotypic landscape without a sequence entropy bias, a reduction in population size would only lead to a broadening of the phenotypic distribution and not a change in the mean of the distribution, and so a much weaker effect, as the common ancestor will still be *typically* far from incompatible regions. On the other hand with strong (exponential) degeneracy biases, the mean phenotype of the common ancestor changes strongly giving the large population size effect seen, which is as demonstrated in Fig 2.

### Biophysics dominates combinatorics: Pair-wise incompatibilities dominate reproductive isolation

Another finding of significance is that pair-wise or 2-way DMIs dominate compared with higher order DMIs (3- and 4-way in this model with 4 loci). This is in contrast to Orr's theoretical argument that predicts a very specific ratio of 2-way: 3-way: 4-way DMIs, equal to 12: 24: 14, which assumes that the fraction of viable paths from the common ancestor to the current day species increases as we consider higher order DMIs [5]. This argument partly rests on the assumption that the number of inviable genotypes remains fixed as a larger number of loci are considered, which would seem a very strong assumption. Despite its simplicity, the genotype-phenotype map in this paper has many of the key features required for higher levels of epistasis, with protein-DNA binding, protein-protein binding and control of the morphogen steepness, all interacting in a non-linear fashion to produce a single gene expression patterning phenotype and so there is clearly the potential for the Orr prediction to be verified; in contrast, we find the converse and our results show there is no bias towards 3-way DMIs, in fact showing instead that the ratio of 2-way to 3-way DMIs is at short times many orders of magnitude larger. This suggests, in this simple, but still relatively complex model, that biophysical constraints are far more important than a purely combinatorial argument would suggest. Evidence could be obtained from more detailed studies similar to [6, 70], where a power law with an exponent greater than 2 would indicate higher-order DMIs are dominant; currently this evidence suggests a quadratic growth law, however, a study with more time-points or species-pairs would provide more confidence. An alternative approach would be to look for linkage disequilibrium between unlinked regions of hybrid genomes, such as was found with hybrids of two species of swordtail fish [71], and though computationally challenging and requiring large numbers of parallel datasets, compare this against evidence for pervasiveness of higher order epistasis. Although recent results of [72], would seem to contradict our conclusions, their finding of extensive complex epistasis relates to higher order interactions between sites within a single locus, coding for protein stability or enzymatic activity, whereas our work relates to epistasis between multiple loci. Similarly, the results of hybrid incompatibilities within RNA molecules [73], which show a 'spiralling complexity' of DMIs would appear to be of limited biological

relevance to allopatric speciation in higher organisms, as these are related to epistasis within a single locus, which are unlikely to segregate into different recombinants in a hybrid.

### Symmetry breaking at large population sizes

Finally, for small populations we find clustering in the behaviours of growth of different types of DMIs, in particular, 3-way DMIs ([S1 Text](#)), which can be explained by the different sequence entropy constraints on different molecular phenotypes. This degeneracy is then lifted at larger population sizes and each  $n$ -way DMI takes on a different identity in their pattern of growth; this has strong analogy to physical phenomenon in statistical physics where constraints of symmetry dominate at large temperatures, in a regime where noise is important, but at smaller temperatures this symmetry is broken.

### Future directions

A clear future direction to investigate would be the effect of multiple transcription factors binding to enhancer regions to control gene expression [[74–76](#)] in large gene networks, where there is potential scope for complex epistasis across many loci coding for a large number of transcription factors. However, as our results show, despite the possibility and a prior expectation of a larger number of triplet interactions, pair-wise interactions dominate; for complex transcriptional control, if pair-wise interactions between proteins, and proteins and DNA dominate, for example in determining the binding affinity of transcriptional complexes, then our conclusions would hold. As we broaden the scope to large gene regulatory networks, there is no strong and direct empirical evidence for pervasive higher order epistasis in their function, which could give rise to higher order incompatibilities being dominant [[77](#)]. Specifically, although there is evidence that higher order incompatibilities have arisen in natural populations [[78–81](#)], nonetheless a survey of these findings suggest there is no evidence for their dominance [[81](#)] as would be predicted by Orr and would be consistent with our findings that point towards biophysics providing a stronger constraint.

### Conclusions

Overall, our results point to a basic principle, where developmental system drift or cryptic variation [[7, 10, 11, 43](#)], play a key role in speciation; basic body plans or phenotypes are conserved, but co-evolution of the components and loci of complicated gene regulatory networks can change differently in different lineages, giving incompatibilities that grow in allopatry. Here, we suggest a universal mechanism, where the rate of growth of incompatibilities is controlled by the drift load, or distribution of phenotypic values, of the common ancestor, which in turn is determined by a balance between selection pushing populations towards phenotypes of higher fitness and genetic drift pushing them towards phenotypes that are more numerous (higher sequence entropy); this basic mechanism would predict in general that traits under weaker selection will dominant the initial development of reproductive isolation. In particular, although in principle more complicated regulation could give rise to more complex patterns of epistasis [[5](#)], our findings suggest that more simple, pair-wise, incompatibilities dominate the development of reproductive isolation between allopatric lineages under stabilising selection.

### Methods

#### Details of the GP map

We model the binding energies of proteins to DNA using the “two-state” approximation [[51, 52](#)], which assumes that the binding energy of each amino acid-nucleotide interaction at the

binding interface is additive and to a good approximation controlled by the number of mismatches, which each have the same penalty in binding affinity. The various protein-DNA binding energies in the main text are then given by the Hamming distance between the respective sequences. We assume these energies  $E$  are measured relative the background free energy of specific and non-specific binding all other sites in the genome, such that the probability of a given transcription factor being bound to a single site is  $p = 1/(1+e^{(E-\mu)/k_B T})$ , where  $\mu$  is the chemical potential ( $\sim \log(\text{concentration})$ ) of the TF [17, 26, 82]. For example, the binding energy between the morphogen ( $M$ ) and the first binding site ( $B$ ) is given by

$$E_{MB} = \epsilon_{pd} \rho(\mathbf{g}_m, \mathbf{g}_B) \tag{3}$$

where  $\rho(\mathbf{g}_m, \mathbf{g}_B)$  is the Hamming distance between the protein binding sequence ( $\mathbf{g}_m$ ) for the morphogen and the sequence for a first regulatory binding site ( $\mathbf{g}_B$ ), where  $\epsilon_{pd}$  is the cost in energy for each mismatch. We assume  $\epsilon_{pd} = 2k_B T$  as a typical value for the mis-match energy, which are found to be in the range  $1-3k_B T$  [51, 52]. Similarly the co-operative protein-protein binding energy, for example between RNAP and the morphogen is  $\tilde{E}_{RM}$  is

$$\tilde{E}_{RM} = \epsilon_{pp} (\ell_{pp} - \rho(\mathbf{g}_R, \check{\mathbf{g}}_M)) \tag{4}$$

where  $\mathbf{g}_R$  is the sequence involved in protein-protein interactions for the polymerase, and  $\check{\mathbf{g}}_M$  represents the equivalent binary sequence for the morphogen, flipped about its centre, which mimics the chirality of real proteins and prevents the co-operative stability from always being maximum for homo-dimers. The parameter  $\epsilon_{pp}$  is the stability added for each favourable hydrophobic interaction between amino acids, which we assume to be  $\epsilon_{pp} = -k_B T$ . Given  $\ell_{pp} = 5$  this gives interactions consistent with typical literature values of  $-2$  to  $-7k_B T$  for hydrophobic interactions between proteins [28, 83].

The morphogen concentration profile is approximately exponential; the exact profile we use is

$$[M](x) = [M_0] \frac{\cosh(\alpha(x-L))}{\sinh(\alpha L)} \tag{5}$$

where this arises from solving the reaction-diffusion equation with reflecting boundary conditions and is valid for all  $\alpha$ .

### Probability of RNAP being bound to promoter

To calculate this probability, we use the canonical ensemble of statistical mechanics, for which the partition function  $\mathcal{Z}$  is most simply expressed in terms of a spin-like variable, which represents the occupation of each binding site,  $\sigma_j = \{0, R, M\}$ ,

$$\mathcal{Z} = \sum_{\sigma_p} \sum_{\sigma_B} e^{-(E_{\sigma_p p} + E_{\sigma_B B} - \mu_{\sigma_p} - \mu_{\sigma_B} + \tilde{E}_{\sigma_p \sigma_B})/k_B T}$$

with  $E_{0j} = 0$ ,  $\tilde{E}_{i'j} = 0$ , for either  $i = 0$  or  $i' = 0$  and  $\mu_0 = 0$ , where  $\mu_{\sigma_j} = k_B T \ln [\sigma_j]$  represents the chemical potentials of the protein species at the  $j^{\text{th}}$  binding site with  $[\sigma_j]$  being the concentration of species  $\sigma_j$  and 0 represents a free binding site. Formally this construction is known as a 3-state Pott's model. So given a 'genome'  $\mathbf{G} = [\mathbf{g}_R, \mathbf{g}_p, \mathbf{g}_M, \mathbf{g}_m, \mathbf{g}_B]$  from which the protein-DNA and protein-protein binding energies are calculated ( $E_{ij}$  and  $\tilde{E}_{i'j}$ , respectively),  $p_{RP}$  is given schematically by

$$p_{RP} = p(\sigma_B - R^*) = \frac{1}{\mathcal{Z}} ((0 - R^*) + (M - R^*) + (R - R^*)) \tag{6}$$

where, for example,

$$(M - R^*) \equiv e^{-(E_{RP} + E_{MB} - \mu_R - \mu_M + \tilde{E}_{RM})/k_B T}$$

is the Boltzmann factor for co-operative binding of the morphogen and RNAP to  $\mathcal{R}$ . Note that we ignore in the partition sum protein-protein binding when not bound to DNA, since these co-operative binding energies tend to be relatively weak compared to DNA binding.

### Monte Carlo scheme for speciation simulations

We use a kinetic Monte Carlo scheme to simulate the evolutionary process for the genome  $\mathbf{G}$  and  $\alpha$  on two independent lineages, assuming a fixed effective population size of  $N$ , and that we are in the regime of small effective population size ( $\ell_G \mu_0 N \ll 1$ , where  $\mu_0$  is the base-pair mutation rate). This means the population is represented by a single fixed sequence (or number for  $\alpha$ ) for all of the loci at each time-point, where effectively mutations are either instantly fixed or eliminated. Specifically, we use the Gillespie algorithm [84], to simulate evolution as a continuous time Markov process; at each step of the simulation the rate of fixation of all one-step mutations from the currently fixed alleles (wild type) is calculated, and one of these mutations is selected randomly in proportion to their relative rate. Time is then progressed by  $K^{-1} \ln(u)$ , where  $K$  is the sum of the rates of all one-step mutants and  $u$  is a random number drawn independently between 0 and 1, which ensures the times at which substitutions occur is Poisson distributed, as we would be expected for a random substitution process. The rates are based upon the Kimura probability of fixation [85]:

$$k = \mu_0 N \frac{1 - e^{-2\delta F}}{1 - e^{-2N\delta F}} \approx \mu_0 \frac{2N\delta F}{1 - e^{-2N\delta F}}, \tag{7}$$

where  $\delta F$  is the change of fitness of a mutation at a particular location, given by fitness function defined in the main text, and  $\mu_0 N$  is the rate at which mutations arise in the population; the latter approximation in Eq 7 assumes  $\delta F \ll 1$ . Note that although in the simulations we use the full form for the fixation probability, fitness effects are typically small ( $\delta F \ll 1$ ) in the simulations, so the substitution rates only depend on the population-scaled fitness changes  $2N\delta F$  which, for a given mutation, is proportional to  $2N\kappa_F$ . Finally, we allow continuous ‘mutations’ in the morphogen steepness parameter  $\alpha$ , chosen from a Gaussian distribution of standard deviation  $\delta\alpha = 0.5$  and assign it an 10 effective base-pairs, which are used when assigning relative weight in the kinetic Monte Carlo scheme, where the total number of base-pairs is  $\ell_G = 60$ .

We determine the Malthusian or log fitness of the spatial gene regulation, from the resulting concentration profile  $[TF](x)$  by use of a functional that promotes expression of the  $TF$  in the anterior half, whilst penalising expression in the posterior half, with truncation selection below a critical value  $F^*$ :

$$F = \begin{cases} \kappa_F \ln(\mathcal{W}) & \text{if } \kappa_F \ln(\mathcal{W}) > F^* \\ -\infty & \text{if } \kappa_F \ln(\mathcal{W}) < F^* \end{cases} \tag{8}$$

where,

$$\mathcal{W}[[TF](x)] = \frac{\int_0^{\mathcal{L}/2} [TF](x) dx - \int_{\mathcal{L}/2}^{\mathcal{L}} [TF](x) dx}{\frac{\mathcal{L}}{2} \max_x \{[TF](x)\}}. \tag{9}$$

where we use a value of  $F^* = -1.6 \times 10^{-3}$ , which corresponds a value of  $\mathcal{W} \approx 0.2$ , when  $\kappa_F = 10^{-3}$ . Strictly, an inviability on a lineage or a hybrid should correspond to  $\mathcal{W} = 0$  or  $F^* = -\infty$ ,

however, these values were chosen to so that a reasonable number of incompatibilities arise in a simulation; for comparison the typical maximum of the integral  $\mathcal{W} \approx 0.6$ . In this paper, we explore how the changing the population scaled strength of selection ( $2N\kappa_F$ ) affects the rate of reproductive isolation, by keeping  $\kappa_F$  fixed and varying  $N$  accordingly. Note that although here the exact form of the fitness is slightly different to the one used in [17], the qualitative behaviour is the same (S1 Text).

The speciation simulations consist of two replicate simulations starting with the same common ancestor and with the same fitness function. We draw the common ancestor from the equilibrium distribution for  $\mathbf{G}$  and  $\alpha$ . To do this we start from a random initial genome, and run one long simulation for 100,000 substitutions for a fixed scaled population size  $2N\kappa_F$ , in order to effectively equilibrate the system (typically 10,000 substitutions are required to adapt to an ensemble of fit states). This represents a reference equilibrium state; different random draws from the equilibrium distribution then consist of running the simulation for a further 100 substitutions.

### Decomposing DMIs

Given a pattern of hybrid incompatibilities, for example, as shown in Fig 4, if there is a 2-way DMI (e.g. between  $C$  and  $\alpha$  loci, which we denote  $I_{Ca}$ ), then all four hybrid-genotypes containing this DMI (e.g.  $RMCa$ ,  $RmCa$ ,  $rMCa$ ,  $rmCa$ ) are inviable; these genotypes define a two-dimensional subspace (or face) of the hypercube. Similarly, the points (e.g.  $rmCa$ ,  $RmCa$ , which we denote  $I_{mCa}$ ) containing a 3-way DMI form a one-dimensional subspace (or line), while a 4-way DMI takes up only a single point in the four-dimensional hypercube. These different 2-way, 3-way and 4-way DMIs are the fundamental incompatibility types which we seek to explain the pattern of hybrid inviable genotypes observed, for example, as in Fig 4a.

However, this decomposition is hugely underdetermined, as there are only  $2^4 - 2 = 14$  possible hybrids (not including the well-adapted genotypes of line 1 and line 2) and a total of  $I_{max} = 3^L + 1 - 2^{L+1} = 50$ , different fundamental incompatibilities, for  $L = 4$  loci. This arises as the total number of  $n$ -point DMIs is  $(2^n - 2) \binom{L}{n}$ , as there are  $\binom{L}{n}$  combinations of  $n$  loci amongst  $L$  total loci and then considering a binary choice of alleles across both lines, there are a total of  $2^n$  allelic combinations or states, 2 of which are the fit allelic combinations where all alleles come from one lineage or the other giving  $2^n - 2$ . For example, between each pair of loci there are  $2^2 - 2 = 2$  mismatching combinations of alleles (e.g.  $rM$  and  $Rm$ ) that could give DMIs and  $\binom{L}{2} = L(L - 1)/2 = 6$  pairwise interactions. A similar argument would give a total of 24 3-way DMIs as there are  $2^3 - 2 = 6$  mismatching combinations of alleles at 3 loci (e.g., excluding  $rmc$  and  $RMC$ ) and  $\binom{L}{3} = 4$  3-way interactions and similarly,  $14 \binom{L}{4} = 14$  for 4-way interactions. In total, the max number of DMIs is  $I_{max} = \sum_{n=2}^L (2^n - 2) \binom{L}{n} = 3^L + 1 - 2^{L+1}$ , which for  $L = 4$  loci is  $I_{max} = 50$ .

The approach we take is to find only those combinations of fundamental DMIs that have the smallest total number that can explain the pattern of hybrid incompatibilities, which from a Bayesian perspective would have the smallest Occam factors [86]; for instance, as shown in Fig 4b the list of 6 incompatible hybrid genotypes  $rmCa$ ,  $rMCa$ ,  $RmCa$ ,  $RMCa$ ,  $Rmca$ ,  $RMca$ , shown by red crosses, can be explained most parsimoniously by three different minimal combinations of fundamental DMIs, each with only 2 DMIs (see main text).

### Supporting information

#### S1 Text. Supporting figures and results.

(PDF)

## Acknowledgments

We would like to thank Davis Pollock for initial discussions.

## Author Contributions

**Conceptualization:** Bhavin S. Khatri, Richard A. Goldstein.

**Formal analysis:** Bhavin S. Khatri.

**Funding acquisition:** Richard A. Goldstein.

**Investigation:** Bhavin S. Khatri.

**Methodology:** Bhavin S. Khatri, Richard A. Goldstein.

**Software:** Bhavin S. Khatri.

**Supervision:** Richard A. Goldstein.

**Validation:** Bhavin S. Khatri.

**Visualization:** Bhavin S. Khatri.

**Writing – original draft:** Bhavin S. Khatri.

**Writing – review & editing:** Bhavin S. Khatri, Richard A. Goldstein.

## References

1. Darwin CR. The Origin of Species. J. Murray, London; 1859.
2. Dobzhansky T. Studies on Hybrid Sterility. II. Localization of Sterility Factors in *Drosophila Pseudoobscura* Hybrids. *Genetics*. 1936; 21(2):113–135. PMID: [17246786](https://pubmed.ncbi.nlm.nih.gov/17246786/)
3. Muller HJ. Isolating mechanisms, evolution and temperature. *Biol Symp*. 1942; 6:71–125.
4. Bateson W. In: Seward AC, editor. Darwin and Modern Science. Cambridge University Press; 1909. p. 85–101.
5. Orr HA. The population genetics of speciation: the evolution of hybrid incompatibilities. *Genetics*. 1995; 139(4):1805–1813. PMID: [7789779](https://pubmed.ncbi.nlm.nih.gov/7789779/)
6. Matute DR, Coyne JA. Intrinsic reproductive isolation between two sister species of *Drosophila*. *Evolution*. 2010; 64(4):903–920. <https://doi.org/10.1111/j.1558-5646.2009.00879.x> PMID: [19891626](https://pubmed.ncbi.nlm.nih.gov/19891626/)
7. Verster AJ, Ramani AK, McKay SJ, Fraser AG. Comparative RNAi screens in *C. elegans* and *C. briggsae* reveal the impact of developmental system drift on gene function. *PLoS genetics*. 2014; 10(2): e1004077. <https://doi.org/10.1371/journal.pgen.1004077> PMID: [24516395](https://pubmed.ncbi.nlm.nih.gov/24516395/)
8. Wotton KR, Jiménez-Guri E, Crombach A, Janssens H, Alcaine-Colet A, Lemke S, et al. Quantitative system drift compensates for altered maternal inputs to the gap gene network of the scuttle fly *Megaselia abdita*. *eLife*. 2015; 4:e04785. <https://doi.org/10.7554/eLife.04785>
9. Crombach A, Wotton KR, Jiménez-Guri E, Jaeger J. Gap Gene Regulatory Dynamics Evolve along a Genotype Network. *Molecular Biology and Evolution*. 2016; 33(5):1293–1307. <https://doi.org/10.1093/molbev/msw013> PMID: [26796549](https://pubmed.ncbi.nlm.nih.gov/26796549/)
10. True JR, Haag ES. Developmental system drift and flexibility in evolutionary trajectories. *Evolution & development*. 2001; 3(2):109–119. <https://doi.org/10.1046/j.1525-142x.2001.003002109.x>
11. Haag ES. The same but different: worms reveal the pervasiveness of developmental system drift. *PLoS Genet*. 2014; 10(2):e1004150. <https://doi.org/10.1371/journal.pgen.1004150> PMID: [24516407](https://pubmed.ncbi.nlm.nih.gov/24516407/)
12. Johnson NA, Porter AH. Rapid speciation via parallel, directional selection on regulatory genetic pathways. *Journal of Theoretical Biology*. 2000; 205(4):527–542. <https://doi.org/10.1006/jtbi.2000.2070> PMID: [10931750](https://pubmed.ncbi.nlm.nih.gov/10931750/)
13. Johnson NA, Porter AH. Toward a new synthesis: population genetics and evolutionary developmental biology. In: *Microevolution Rate, Pattern, Process*. Springer; 2001. p. 45–58.
14. Fontana W. Modelling 'evo-devo' with RNA. *Bioessays*. 2002; 24(12):1164–1177. <https://doi.org/10.1002/bies.10190>

15. Berg J, Willmann S, Lässig M. Adaptive evolution of transcription factor binding sites. *BMC Evol Biol.* 2004; 4(1):42. <https://doi.org/10.1186/1471-2148-4-42> PMID: 15511291
16. Mustonen V, Kinney J, Callan CG, Lässig M. Energy-dependent fitness: a quantitative model for the evolution of yeast transcription factor binding sites. *Proceedings of the National Academy of Sciences.* 2008; 105(34):12376–12381. <https://doi.org/10.1073/pnas.0805909105>
17. Khatri BS, McLeish TCB, Sear RP. Statistical mechanics of convergent evolution in spatial patterning. *Proc Natl Acad Sci U S A.* 2009; 106(24):9564–9569. <https://doi.org/10.1073/pnas.0812260106> PMID: 19497876
18. Hayden EJ, Ferrada E, Wagner A. Cryptic genetic variation promotes rapid evolutionary adaptation in an RNA enzyme. *Nature.* 2011; 474(7349):92–95. <https://doi.org/10.1038/nature10083> PMID: 21637259
19. Goldstein RA. The evolution and evolutionary consequences of marginal thermostability in proteins. *Proteins: Structure, Function, and Bioinformatics.* 2011; 79(5):1396–1407. <https://doi.org/10.1002/prot.22964>
20. Schaper S, Louis AA. The arrival of the frequent: how bias in genotype-phenotype maps can steer populations to local optima. *PLoS one.* 2014; 9(2):e86635. <https://doi.org/10.1371/journal.pone.0086635> PMID: 24505262
21. Greenbury SF, Johnston IG, Louis AA, Ahnert SE. A tractable genotype-phenotype map modelling the self-assembly of protein quaternary structure. *Journal of The Royal Society Interface.* 2014; 11(95):20140249. <https://doi.org/10.1098/rsif.2014.0249>
22. Lynch M, Hagner K. Evolutionary meandering of intermolecular interactions along the drift barrier. *Proceedings of the National Academy of Sciences.* 2015; 112(1):E30–E38. <https://doi.org/10.1073/pnas.1421641112>
23. Manrubia S, Cuesta JA. Evolution on neutral networks accelerates the ticking rate of the molecular clock. *Journal of The Royal Society Interface.* 2015; 12(102):20141010. <https://doi.org/10.1098/rsif.2014.1010>
24. Greenbury SF, Schaper S, Ahnert SE, Louis AA. Genetic correlations greatly increase mutational robustness and can both reduce and enhance evolvability. *PLoS Comput Biol.* 2016; 12(3):e1004773. <https://doi.org/10.1371/journal.pcbi.1004773> PMID: 26937652
25. Goldstein RA, Pollock DD. Sequence entropy of folding and the absolute rate of amino acid substitutions. *Nature ecology & evolution.* 2017; 1(12):1923. <https://doi.org/10.1038/s41559-017-0338-9>
26. Khatri BS, Goldstein RA. A coarse-grained biophysical model of sequence evolution and the population size dependence of the speciation rate. *Journal of theoretical biology.* 2015; 378:56–64. <https://doi.org/10.1016/j.jtbi.2015.04.027> PMID: 25936759
27. Khatri BS, Goldstein RA. Simple biophysical model predicts faster accumulation of hybrid incompatibilities in small populations under stabilizing selection. *Genetics.* 2015; 201(4):1525–1537. <https://doi.org/10.1534/genetics.115.181685> PMID: 26434721
28. Shea MA, Ackers GK. The OR control system of bacteriophage lambda. A physical-chemical model for gene regulation. *J Mol Biol.* 1985; 181(2):211–230. [https://doi.org/10.1016/0022-2836\(85\)90086-5](https://doi.org/10.1016/0022-2836(85)90086-5) PMID: 3157005
29. Eyre-Walker A, Woolfit M, Phelps T. The distribution of fitness effects of new deleterious amino acid mutations in humans. *Genetics.* 2006; 173(2):891–900. <https://doi.org/10.1534/genetics.106.057570> PMID: 16547091
30. Huber CD, Kim BY, Marsden CD, Lohmueller KE. Determining the factors driving selective effects of new nonsynonymous mutations. *Proceedings of the National Academy of Sciences.* 2017; 114(17):4465–4470. <https://doi.org/10.1073/pnas.1619508114>
31. Orr HA, Turelli M. The evolution of postzygotic isolation: accumulating Dobzhansky-Muller incompatibilities. *Evolution.* 2001; 55(6):1085–1094. <https://doi.org/10.1111/j.0014-3820.2001.tb00628.x> PMID: 11475044
32. Bertin E, Bouchaud JP. Subdiffusion and localization in the one-dimensional trap model. *Physical Review E.* 2003; 67(2):026128. <https://doi.org/10.1103/PhysRevE.67.026128>
33. King MC, Wilson AC. Evolution at two levels in humans and chimpanzees. *Science.* 1975; 188(4184):107–116.
34. Wolf JB, Lindell J, Backström N. Speciation genetics: current status and evolving approaches; 2010.
35. Wray GA. The evolutionary significance of cis-regulatory mutations. *Nature Reviews Genetics.* 2007; 8(3):206–216. <https://doi.org/10.1038/nrg2063> PMID: 17304246
36. Abzhanov A, Kuo WP, Hartmann C, Grant BR, Grant PR, Tabin CJ. The calmodulin pathway and evolution of elongated beak morphology in Darwin's finches. *Nature.* 2006; 442(7102):563–567. <https://doi.org/10.1038/nature04843>

37. Wittkopp PJ, Haerum BK, Clark AG. Regulatory changes underlying expression differences within and between *Drosophila* species. *Nature genetics*. 2008; 40(3):346–350. <https://doi.org/10.1038/ng.77> PMID: 18278046
38. Ting CT, Tsaur SC, Wu ML, Wu CI. A rapidly evolving homeobox at the site of a hybrid sterility gene. *Science*. 1998; 282(5393):1501–1504. <https://doi.org/10.1126/science.282.5393.1501> PMID: 9822383
39. Brideau NJ, Flores HA, Wang J, Maheshwari S, Wang X, Barbash DA. Two Dobzhansky-Muller genes interact to cause hybrid lethality in *Drosophila*. *science*. 2006; 314(5803):1292–1295. <https://doi.org/10.1126/science.1133953> PMID: 17124320
40. Bayes JJ, Malik HS. Altered heterochromatin binding by a hybrid sterility protein in *Drosophila* sibling species. *Science*. 2009; 326(5959):1538–1541. <https://doi.org/10.1126/science.1181756> PMID: 19933102
41. Liénard MA, Araripe LO, Hartl DL. Neighboring genes for DNA-binding proteins rescue male sterility in *Drosophila* hybrids. *Proceedings of the National Academy of Sciences*. 2016; 113(29):E4200–E4207. <https://doi.org/10.1073/pnas.1608337113>
42. Mack KL, Nachman MW. Gene Regulation and Speciation. *Trends in Genetics*. 2016;. <https://doi.org/10.1016/j.tig.2016.11.003> PMID: 27914620
43. Gavin-Smyth J, Matute DR. Embryonic lethality leads to hybrid male inviability in hybrids between *Drosophila melanogaster* and *D. santomea*. *Ecology and evolution*. 2013; 3(6):1580–1589. <https://doi.org/10.1002/ece3.573> PMID: 23789069
44. Fisher RA. *The Genetical Theory of Natural Selection*. Oxford: Oxford Univ. Press; 1930.
45. Wright S. Evolution in Mendelian Populations. *Genetics*. 1931; 16(2):97–159. PMID: 17246615
46. Kimura M. Diffusion models in population genetics. *Journal of Applied Probability*. 1964; 1(2):177–232. <https://doi.org/10.2307/3211856>
47. Crow JF, Kimura M. *An Introduction to Population Genetics Theory*. Harper & Row; 1970.
48. Ewens WJ. *Mathematical Population Genetics: 1. A Theoretical Introduction*. Springer; 2010.
49. Mustonen V, Lässig M. Evolutionary population genetics of promoters: predicting binding sites and functional phylogenies. *Proceedings of the National Academy of Sciences of the United States of America*. 2005; 102(44):15936–15941. <https://doi.org/10.1073/pnas.0505537102> PMID: 16236723
50. Haldane A, Manhart M, Morozov AV. Biophysical fitness landscapes for transcription factor binding sites. *PLoS computational biology*. 2014; 10(7):e1003683. <https://doi.org/10.1371/journal.pcbi.1003683> PMID: 25010228
51. von Hippel PH, Berg OG. On the specificity of DNA-protein interactions. *Proc Natl Acad Sci U S A*. 1986; 83(6):1608–1612. <https://doi.org/10.1073/pnas.83.6.1608> PMID: 3456604
52. Gerland U, Moroz JD, Hwa T. Physical constraints and functional characteristics of transcription factor-DNA interaction. *Proc Natl Acad Sci U S A*. 2002; 99(19):12015–12020. <https://doi.org/10.1073/pnas.192693599> PMID: 12218191
53. Iwasa Y. Free fitness that always increases in evolution. *Journal of Theoretical Biology*. 1988; 135(3):265–281. [https://doi.org/10.1016/S0022-5193\(88\)80243-1](https://doi.org/10.1016/S0022-5193(88)80243-1) PMID: 3256719
54. Tulchinsky AY, Johnson NA, Watt WB, Porter AH. Hybrid incompatibility arises in a sequence-based bioenergetic model of transcription factor binding. *Genetics*. 2014; 198(3):1155–1166. <https://doi.org/10.1534/genetics.114.168112> PMID: 25173845
55. Coyne JA, Orr HA. The evolutionary genetics of speciation. *Philos Trans R Soc Lond B Biol Sci*. 1998; 353(1366):287–305. <https://doi.org/10.1098/rstb.1998.0210> PMID: 9533126
56. Barraclough TG, Nee S. Phylogenetics and speciation. *Trends in Ecology & Evolution*. 2001; 16(7):391–399. [https://doi.org/10.1016/S0169-5347\(01\)02161-9](https://doi.org/10.1016/S0169-5347(01)02161-9)
57. Nee S. Inferring speciation rates from phylogenies. *Evolution*. 2001; 55(4):661–668. [https://doi.org/10.1554/0014-3820\(2001\)055%5B0661:ISRFP%5D2.0.CO;2](https://doi.org/10.1554/0014-3820(2001)055%5B0661:ISRFP%5D2.0.CO;2) PMID: 11392383
58. Fitzpatrick BM. Rates of evolution of hybrid inviability in birds and mammals. *Evolution*. 2004; 58(8):1865–1870. <https://doi.org/10.1554/04-190> PMID: 15446440
59. Stelkens RB, Young KA, Seehausen O. The accumulation of reproductive incompatibilities in African cichlid fish. *Evolution*. 2010; 64(3):617–633. <https://doi.org/10.1111/j.1558-5646.2009.00849.x> PMID: 19796149
60. Cooper A, Penny D. Mass survival of birds across the Cretaceous-Tertiary boundary: molecular evidence. *Science*. 1997; 275(5303):1109–1113. <https://doi.org/10.1126/science.275.5303.1109> PMID: 9027308
61. Van Oppen MH, Turner G, Rico C, Deutsch J, Ibrahim K, Robinson R, et al. Unusually fine-scale genetic structuring found in rapidly speciating Malawi cichlid fishes. *Proceedings of the Royal Society of London B: Biological Sciences*. 1997; 264(1389):1803–1812. <https://doi.org/10.1098/rspb.1997.0248>

62. Fiumera A, Parker P, Fuerst P. Effective population size and maintenance of genetic diversity in captive-bred populations of a lake Victoria cichlid. *Conservation Biology*. 2000; 14(3):886–892. <https://doi.org/10.1046/j.1523-1739.2000.97337.x>
63. Jorde LB, Bamshad M, Rogers AR. Using mitochondrial and nuclear DNA markers to reconstruct human evolution. *Bioessays*. 1998; 20(2):126–136. [https://doi.org/10.1002/\(SICI\)1521-1878\(199802\)20:2<126::AID-BIES5>3.0.CO;2-R](https://doi.org/10.1002/(SICI)1521-1878(199802)20:2<126::AID-BIES5>3.0.CO;2-R) PMID: 9631658
64. Gattepaille L, Günther T, Jakobsson M. Inferring past effective population size from distributions of coalescent times. *Genetics*. 2016; 204(3):1191–1206. <https://doi.org/10.1534/genetics.115.185058> PMID: 27638421
65. Sawai H, Kim HL, Kuno K, Suzuki S, Gotoh H, Takada M, et al. The origin and genetic variation of domestic chickens with special reference to junglefowls *Gallus g. gallus* and *G. varius*. *PloS one*. 2010; 5(5):e10639. <https://doi.org/10.1371/journal.pone.0010639> PMID: 20502703
66. Lande R. Effective deme sizes during long-term evolution estimated from rates of chromosomal rearrangement. *Evolution*. 1979; 33(1Part1):234–251. <https://doi.org/10.1111/j.1558-5646.1979.tb04678.x> PMID: 28568063
67. Lande R. Expected time for random genetic drift of a population between stable phenotypic states. *Proc Natl Acad Sci U S A*. 1985; 82(22):7641–7645. <https://doi.org/10.1073/pnas.82.22.7641> PMID: 3865184
68. Barton NH, Charlesworth B. Genetic Revolutions, Founder Effects, and Speciation. *Annual Review of Ecology and Systematics*. 1984; 15(1):133–164. <https://doi.org/10.1146/annurev.es.15.110184.001025>
69. Barton N, Rouhani S. The frequency of shifts between alternative equilibria. *Journal of theoretical biology*. 1987; 125(4):397–418. [https://doi.org/10.1016/S0022-5193\(87\)80210-2](https://doi.org/10.1016/S0022-5193(87)80210-2) PMID: 3657219
70. Moyle LC, Nakazato T. Hybrid incompatibility “snowballs” between *Solanum* species. *Science*. 2010; 329(5998):1521–1523. <https://doi.org/10.1126/science.1193063>
71. Schumer M, Cui R, Powell DL, Dresner R, Rosenthal GG, Andolfatto P. High-resolution mapping reveals hundreds of genetic incompatibilities in hybridizing fish species. *Elife*. 2014; 3:e02535. <https://doi.org/10.7554/eLife.02535>
72. Weinreich DM, Lan Y, Wylie CS, Heckendorn RB. Should evolutionary geneticists worry about higher-order epistasis? *Current opinion in genetics & development*. 2013; 23(6):700–707.
73. Kalirad A, Azevedo RB. Spiraling complexity: a test of the snowball effect in a computational model of RNA folding. *Genetics*. 2017; 206(1):377–388. <https://doi.org/10.1534/genetics.116.196030> PMID: 28007889
74. Bintu L, Buchler NE, Garcia HG, Gerland U, Hwa T, Kondev J, et al. Transcriptional regulation by the numbers: models. *Curr Opin Genet Dev*. 2005; 15(2):116–124. <https://doi.org/10.1016/j.gde.2005.02.007> PMID: 15797194
75. Spitz F, Furlong EE. Transcription factors: from enhancer binding to developmental control. *Nature Reviews Genetics*. 2012; 13(9):613–626. <https://doi.org/10.1038/nrg3207> PMID: 22868264
76. Levo M, Segal E. In pursuit of design principles of regulatory sequences. *Nature Reviews Genetics*. 2014; 15(7):453–468. <https://doi.org/10.1038/nrg3684> PMID: 24913666
77. Taylor MB, Ehrenreich IM. Higher-order genetic interactions and their contribution to complex traits. *Trends in genetics*. 2015; 31(1):34–40. <https://doi.org/10.1016/j.tig.2014.09.001> PMID: 25284288
78. Orr HA, Irving S. Complex epistasis and the genetic basis of hybrid sterility in the *Drosophila pseudoobscura* Bogota-USA hybridization. *Genetics*. 2001; 158(3):1089–1100. PMID: 11454758
79. Cabot EL, Davis AW, Johnson NA, Wu CI. Genetics of reproductive isolation in the *Drosophila simulans* clade: complex epistasis underlying hybrid male sterility. *Genetics*. 1994; 137(1):175–189. PMID: 8056308
80. Palopoli MF, Wu CI. Genetics of hybrid male sterility between *drosophila* sibling species: a complex web of epistasis is revealed in interspecific studies. *Genetics*. 1994; 138(2):329–341. PMID: 7828817
81. Fraïsse C, Elderfield J, Welch J. The genetics of speciation: are complex incompatibilities easier to evolve? *Journal of evolutionary biology*. 2014; 27(4):688–699. <https://doi.org/10.1111/jeb.12339>
82. Lässig M. From biophysics to evolutionary genetics: statistical aspects of gene regulation. *BMC Bioinformatics*. 2007; 8 Suppl 6:S7. <https://doi.org/10.1186/1471-2105-8-S6-S7> PMID: 17903288
83. Buchler NE, Gerland U, Hwa T. On schemes of combinatorial transcription logic. *Proc Natl Acad Sci U S A*. 2003; 100(9):5136–5141. <https://doi.org/10.1073/pnas.0930314100> PMID: 12702751
84. Gillespie DT. A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *Journal of computational physics*. 1976; 22(4):403–434. [https://doi.org/10.1016/0021-9991\(76\)90041-3](https://doi.org/10.1016/0021-9991(76)90041-3)

85. Kimura M. On the probability of fixation of mutant genes in a population. *Genetics*. 1962; 47:713–719. PMID: [14456043](https://pubmed.ncbi.nlm.nih.gov/14456043/)
86. MacKay DJC. IV. In: *Information Theory, Inference, and Learning Algorithms*. Cambridge University Press; 2007. p. 343.