LETTERS

Sexual reproduction selects for robustness and negative epistasis in artificial gene networks

Ricardo B. R. Azevedo[1], Rolf Lohaus[1], Suraj Srinivasan[1], Kristen K. Dang[2] & Christina L. Burch[3]

The mutational deterministic hypothesis for the origin and maintenance of sexual reproduction posits that sex enhances the ability of natural selection to purge deleterious mutations after recombination brings them together into single genomes[1]. This explanation requires negative epistasis, a type of genetic interaction where mutations are more harmful in combination than expected from their separate effects. The conceptual appeal of the mutational deterministic hypothesis has been offset by our inability to identify the mechanistic and evolutionary bases of negative epistasis. Here we show that negative epistasis can evolve as a consequence of sexual reproduction itself. Using an artificial gene network model[2,3], we find that recombination between gene networks imposes selection for genetic robustness, and that negative epistasis evolves as a by-product of this selection. Our results suggest that sexual reproduction selects for conditions that favour its own maintenance, a case of evolution forging its own path.

A century of genetic research has revealed two general properties of spontaneous mutations with detectable effects on fitness: most of them are deleterious, and they frequently interact with each other[4,5]. Many types of interactions are possible, including directional epistasis, in which the average effect of spontaneous mutations changes in the presence of other mutations in the genome[6]. Directional epistasis can be either negative (synergistic) or positive (antagonistic), depending on whether the average effect of mutations becomes more or less harmful, respectively, as the number of other mutations in the genome increases (Fig. 1). Directional epistasis holds particular interest for evolutionary biologists because it is expected to determine the outcome of multiple evolutionary processes, notably the evolution of sex and recombination[1]. Empirical studies on a variety of organisms have reported every conceivable form of directional epistasis: negative[7–9], positive[6,10] and no significant directional epistasis[11,12]. These mixed results have not helped to clarify either the mechanistic or evolutionary causes of directional epistasis[13].

In contrast, evolutionary simulations using computational models of RNA secondary structure[14], viral replication[15] and artificial life[14] have demonstrated that the average strength and direction of epistasis can be shaped by natural selection. One mechanism by which epistasis evolves in these models[13] is through a negative correlation among genotypes between the extent of genetic robustness (or genetic canalization, measured as the insensitivity of a phenotype to mutation) and the direction of epistasis. As a consequence, selection for higher robustness produces a correlated response in the strength of epistasis in all three models, towards either weaker positive or stronger negative epistasis[14,15]. The repeatability of this result in models of different biological systems suggests that the strength and direction of epistasis observed in living organisms depend on their history of selection for genetic robustness.

Theory predicts that traits can evolve to be robust to genetic perturbations (that is, mutation and recombination) under a variety of selective regimes[16–18], as long as the following two conditions are met: genes must interact to determine the trait[17–19], and the population must contain sufficient genetic variation[18]. Whereas the former condition is inherent to particular organisms, the latter condition will depend on population genetic parameters such as the mutation and recombination rates. Experimental tests of these predictions using computational models confirm that high mutation rates, such as those experienced by RNA viruses, favour the evolution of genetic robustness[2,3,18,20]. Sexual reproduction (that is, increased recombination) is also expected to impose stronger selection for genetic robustness than asexual reproduction[21,22], but this hypothesis has never been tested experimentally[21].

To test this hypothesis, and to determine whether the evolution of genetic robustness is accompanied by the evolution of negative epistasis, we return to the computational model of genetic networks used in two previous studies[2,3]. We chose this model primarily because it explicitly incorporates one of the key characteristics required for the evolution of robustness[17–19]—genetic interactions. Furthermore, empirical data from biological systems has consistently suggested that extant gene networks are robust to changes in biochemical rate parameters and levels of gene activity[19,23]. Previous work with this model has shown that genetic robustness (again, measured as robustness to mutation) evolves readily if networks are subjected to selection for the production of a stable gene expression pattern[2,3]. Here we explore the extent to which recombination contributed to the evolution of genetic robustness in this model, and ask whether recombination, through its effect on robustness[2,21,22], can cause the direction of epistasis to evolve.

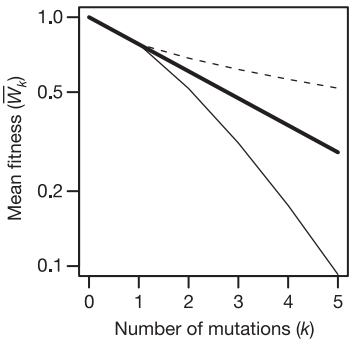

Figure 1 | Types of directional epistasis for deleterious mutations. Three hypothetical relationships between fitness (log scale) and number of deleterious mutations are plotted. All relationships depicted have the same mutational robustness ($\overline{W}_1 = 0.78$) but different directions of epistasis: negative epistasis (plain line, concave downwards; $1 - \beta < 0$), no directional epistasis (bold, straight line; $1 - \beta = 0$) and positive epistasis (dashed line, concave upwards; $1 - \beta > 0$).

[1]Department of Biology and Biochemistry, University of Houston, Houston, Texas 77204-5001, USA. [2]Department of Biomedical Engineering, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599-7575, USA. [3]Department of Biology, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599-3280, USA.

Briefly, the model[2,3] represents individuals as networks of N interacting transcriptional regulators (Fig. 2a). The genotype of an individual is represented by an $N \times N$ matrix R, whose elements r_{ij} describe the regulatory effect of the product of gene j on the expression of gene i (Fig. 2b). The number of regulatory interactions is determined by a connectivity parameter (c) that specifies the proportion of non-zero matrix elements. This matrix of regulatory relationships acts on gene expression patterns, which are represented by a state vector $\mathbf{S}(t)$, whose elements $s_i(t)$ describe the expression states of genes $i = 1, 2, \ldots, N$ at time t. The expression state of a gene can vary continuously between complete repression, $s_i(t) = -1$, and complete activation, $s_i(t) = 1$. Gene expression states change over time according to the following equation:

$$s_i(t+1) = f\left[\sum_{j=1}^{N} r_{ij} s_i(t)\right] \quad (1)$$

where $f(x)$ is a sigmoidal filter function[2] that determines how the total regulatory input influences gene expression (see Supplementary Fig. 1a). Development is modelled as the progression from an initial gene expression state to an equilibrium gene expression pattern (Fig. 2c; see Methods). In this model, genotypes that achieve any stable, fixed-point equilibrium expression pattern are considered developmentally stable[2], and therefore viable. Genotypes that do not achieve a stable equilibrium (for example, oscillatory gene expression) are considered inviable[2].

The mechanistic underpinnings of this model allow *a priori* predictions about the evolution of genetic robustness. In this model, the reproductive success of a viable genotype (that is, its fitness) is the proportion of its offspring that are also viable. When offspring are produced asexually, and differ from their parents only by mutations, the fitness of a genotype is given by:

$$W_{\text{asex}} = \sum_{k=0}^{cN^2} \phi_k \overline{W}_k \quad (2)$$

where $\phi_k = \mu^k e^{-\mu}/k!$ is the Poisson probability that offspring acquire k mutations when the mutation rate per individual network per generation is μ, and $\overline{W}_k$ is the mean fitness of the genotype after the addition of k mutations. For mutation rates $\mu \leq 0.1$, terms for $k > 1$ can be effectively ignored, so that $W_{\text{asex}} \approx 1 - \mu + \mu\overline{W}_1$. $\overline{W}_1$ is our measure of mutational robustness, the probability that a genotype with one mutation is viable. Therefore, given a sufficiently high mutation rate, asexually reproducing networks are expected to evolve mutational robustness. In contrast, when offspring are produced via sexual reproduction, and differ from their parents primarily as a result of recombination, the fitness of a genotype is given by $W_{\text{sex}} = W_{\text{asex}}(1 - L)$, where L is the probability that a mating with a random individual in the population will result in an inviable offspring, a measure of the recombination load[24]. Therefore, sexual populations should experience selection for two distinct types of genetic robustness: mutational robustness and recombinational robustness.

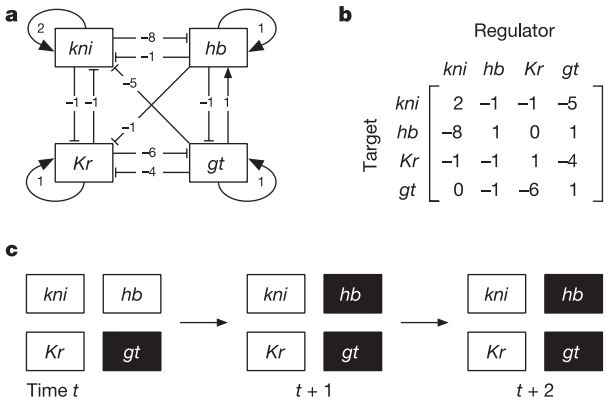

Figure 2 | Application of our network model to the gap gene system of *Drosophila melanogaster*. **a**, Network representation of the regulatory interactions between four gap genes[29] (*gt*, *giant*; *hb*, *hunchback*; *kni*, *knirps*; *Kr*, *Krüppel*). Activations and repressions are denoted by arrows and bars, respectively. Numbers indicate the relative interaction strengths[30]. **b**, Interaction matrix (R) representing the network in **a**. The element in row i and column j (r_{ij}) denotes the regulatory effect of the product of gene j on the expression of gene i. **c**, Graphical representation of the gene expression states of each gap gene over three successive time steps. For the purpose of this illustration we consider gene i to be ON (filled box) if $s_i(t) > 0$, and OFF (open box) if $s_i(t) \leq 0$. The change in gene expression pattern matches events at ~80% anterior–posterior position in the *Drosophila* embryo between early and mid cleavage cycle 14A (ref. 29). Successive iterations beyond the $t + 1$ step do not change the gene expression pattern, the hallmark of a stable equilibrium.

In order to explore the effect of sexual reproduction on the evolution of genetic robustness, we first investigated the behaviour of this model using conditions that are known to produce robustness to mutation[2] ($\mu = 0.1$, $N = 10$ genes, $c = 0.75$; see Methods), varying only the reproductive mode from sexual to asexual. Fifty clonal populations of 500 individuals were founded by different randomly generated, viable genotypes. Each population was subjected to selection for the ability to produce a stable gene expression pattern and allowed to evolve separately via sexual and asexual reproduction. We monitored evolution until an equilibrium level of mutational robustness was achieved. Contrary to earlier claims[2,3], our simulations show that sexual reproduction has a substantive effect on the evolution of mutational robustness (Fig. 3a). Although mutational robustness increased in asexual populations, it reached a significantly lower equilibrium value than in sexual populations (paired t-test: $t = 31.0$, 49 degrees of freedom (d.f.), $P < 0.0001$). An investigation of epistasis in these evolved populations revealed that sexual reproduction also had a qualitative effect on the evolution of directional epistasis ($t = 23.6$, 49 d.f., $P < 0.0001$). The magnitude of epistasis evolved regardless of reproductive mode, but the direction of epistasis only changed when reproduction was sexual. At equilibrium, asexual populations exhibited average positive epistasis of a reduced magnitude, whereas sexual populations exhibited negative epistasis.

Why does sexual reproduction cause the evolution of increased mutational robustness? Sexual reproduction is not expected to increase the strength of selection for mutational robustness directly. However, it is expected to select for recombinational robustness, and this could cause a correlated response in mutational robustness. We devised two experiments to test this hypothesis. In the first experiment, we investigated whether the effect of sexual reproduction on mutational robustness depended on the high mutation rate ($\mu = 0.1$). Theory predicts that mutational robustness will evolve through the direct action of selection only if μ is greater than the reciprocal of the effective population size[7] (that is, $\mu > 0.002$ in our simulations). Thus, we tested our hypothesis by re-running the simulations at a mutation rate of $\mu = 0.002$ (Fig. 3a). At this low mutation rate, mutational robustness failed to evolve in asexual populations within 50,000 generations. However, mutational robustness did increase significantly in sexual populations within 20,000 generations. The inability of asexual populations to respond to selection for mutational robustness confirms that selection acting directly on mutational robustness is ineffective when $\mu = 0.002$. Thus, the mutational robustness that evolved in these sexual populations did not evolve through the direct action of selection. Rather, it must have evolved as a correlated response to selection for recombinational robustness, the only other source of selection in these simulations.

In the second experiment we constructed genetically variable populations (see Supplementary Methods) and allowed them to evolve in the absence of new mutations ($\mu = 0$), that is, in the absence of selection for mutational robustness. In this experiment,

sexual populations showed significant increases in mutational robustness, whereas asexual populations did not (Fig. 3b). These results confirm our hypothesis because sexual and asexual populations differed only by the presence and absence, respectively, of selection for recombinational robustness. By manipulating the amount of genetic variation present in the founder populations, we also showed that the evolutionary response in mutational robustness increased with the strength of selection for recombinational robustness (that is, with the magnitude of the recombination load, L; Supplementary Fig. 2).

Directional epistasis evolved in both of these experiments (Fig. 3) in a similar manner to the initial simulations. Conditions that showed no change in mutational robustness also showed no change in directional epistasis. However, conditions that caused an evolutionary response in mutational robustness also caused the evolution of negative, or less positive, epistasis. Taken together, these results confirm that mutational robustness and negative epistasis both evolved in response to selection for recombinational robustness.

The most likely explanation for the evolution of negative epistasis in these simulations is that epistasis evolved as a correlated response to selection for genetic robustness. The direction of epistasis was negatively correlated with mutational robustness among a random sample of viable gene networks (Supplementary Fig. 3). Similar correlations were found in digital organisms and RNA secondary structure[14], supporting the theoretical prediction[14,25] that it is impossible to change genetic robustness and the direction of epistasis independently.

Although we recognize that our model describes a simplified view of transcriptional regulation, it captures an important feature of real genetic regulatory systems: genetic interactions are abundant, causing mutations to have different effects depending on the genetic background in which they arise. We propose that sexual reproduction will favour the evolution of increased genetic robustness and, therefore, negative epistasis in any system with two key properties: numerous genetic interactions and abundant genetic variation—both known requirements for the evolution of genetic robustness[17–19]. Consistent with this proposal, the parameters that determine the number of genetic interactions (connectivity and gene number) and the amount of genetic variation (mutation rate, population size and the strength of stabilizing selection) all influenced whether negative epistasis evolved in our simulations (Fig. 3; see also Supplementary Figs 4–7). In contrast, the network topology, the shape and variance of the mutational distribution, and environmental stochasticity did not qualitatively affect the outcome (Supplementary Figs 1, 5, 8 and 9). Most notably, negative epistasis failed to evolve in networks that were both small and sparsely connected (Supplementary Figs 5 and 6). However, the requirements for the evolution of negative epistasis were not too restrictive. Sexual reproduction produced negative epistasis even in small networks as long as they were sufficiently connected, and in sparsely connected networks as long as they contained a sufficient number of genes (Supplementary Figs 5 and 6). Stabilizing selection acting on the gene expression pattern also prevented the evolution of negative epistasis, but only when it was exceptionally strong (Supplementary Fig. 7). The wealth of genetic interactions in the transcriptional networks of real organisms[26] and the abundance of genetic variation in natural populations[27] suggest that negative epistasis will evolve in many sexually reproducing organisms.

The evolution of negative epistasis in our simulations is remarkable because it suggests that sexual reproduction selects for conditions that favour its own maintenance. Because negative epistasis enhances the ability of natural selection to purge deleterious mutations in sexual populations, our results could explain the maintenance of sexual reproduction in the face of its numerous

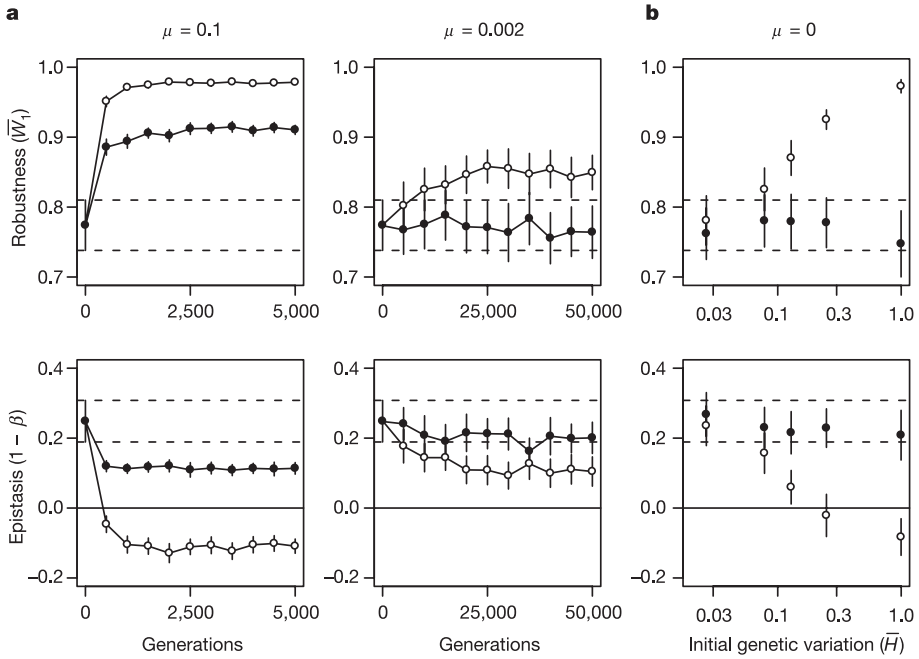

Figure 3 | Sexual reproduction selects for mutational robustness and negative epistasis. a, Selection for the ability to produce any stable gene expression pattern was imposed on 50 replicate populations subjected to high ($\mu = 0.1$) and low ($\mu = 0.002$) mutation rates. Plots show the average evolutionary responses in robustness to mutation and direction of epistasis. **b**, The 50 individuals used to found the homogeneous populations described in **a** were used to found new populations of 500 individuals with 1, 3, 5, 10 or 75 random mutations each. These populations were then allowed to evolve without the occurrence of new mutations ($\mu = 0$) until genetic variation was exhausted. We plot the robustness to mutation and the direction of epistasis at equilibrium (that is, after evolution stops) against the initial genetic variation in each treatment (see Supplementary Methods). Each population in **a** and **b** was evolved under either asexual (filled circles) or sexual (open circles) reproduction. Data are expectations and 95% confidence intervals for the median value among the 500 individuals in each population. Dashed lines indicate the 95% confidence intervals for the mean of the 50 founder networks.

costs[28]. We hypothesize that sexual reproduction enabled evolution of the robustness apparent in the developmental networks of multicellular organisms and that negative epistasis should be associated with robustness in these systems. If these hypotheses are correct, they will help to explain the prevalence of sexual reproduction among living organisms.

METHODS

Network model. Networks are generated by randomly filling the entries of the R matrix (for example, Fig. 2b) with $(1-c)N^2$ zeros and cN^2 standard normal random variates. A corresponding initial gene expression pattern, $\mathbf{S}(0)$, is created for each network by randomly setting each $s_i(0)$ to either -1 or 1. Development begins with the initial gene expression pattern, $\mathbf{S}(0)$, and proceeds through 100 iterations of equation (1). We determined that an equilibrium steady state was achieved when the following criterion was met[2]:

$$\sum_{\theta=t-10}^{t} D[\mathbf{S}(\theta),\overline{\mathbf{S}}(t)] \leq 10^{-3}, \quad \text{where } D[\mathbf{S},\mathbf{S}'] = \frac{1}{4N}\sum_{i=1}^{N}(s_i - s_i')^2$$

is a measure of the difference between the gene expression patterns $\mathbf{S}$ and $\mathbf{S}'$, and $\overline{\mathbf{S}}(t)$ is the average of the gene expression levels over the time interval from $t-10$ to t. The ability of a genotype to reach equilibrium within 100 iterations is termed developmental stability[2].

Evolution. In a typical evolutionary simulation, a single random individual capable of producing a stable gene expression pattern is cloned to generate a population of 500 identical individuals. In an asexually reproducing population, offspring are generated by picking an individual at random from the population and allowing it to produce a clone of itself, such that each non-zero entry in the R interaction matrix mutates (replacement with an independent standard normal random variate) with probability $\mu/(cN^2)$. In our model, mutations should be viewed as acting on the cN^2 *cis*-regulatory elements, not the coding sequences of the N genes themselves; in addition, mutations cannot alter the number of genes, or establish new interactions between genes. Only offspring capable of producing a stable gene expression pattern survive. This process is repeated until 500 developmentally stable individuals are produced, which go on to found the following generation. In a sexual population, offspring are generated by picking two individuals at random from the population, and selecting rows of the R matrices from each parent with equal probability (analogous to free recombination between units formed by each gene and its *cis*-regulatory elements, but with no recombination within regulatory regions), while allowing each non-zero entry to mutate as above. Each selective regime was applied to a fixed panel of 50 replicate populations, each derived from a single independently generated random individual and initial gene expression pattern; simulations were run for as long as was necessary to obtain an equilibrium (that is, no significant change) in the second half of the simulation. In each simulation, all individuals experience the same initial gene expression pattern as the founder individual.

Robustness and epistasis. The mean effects of k mutations on fitness were modelled by the relationship[14,15]: $\log(\overline{W}_k) = -\alpha k^{\beta}$ (Fig. 1, equation (2)). Mutational robustness and directional epistasis were measured by $\overline{W}_1$ and $1-\beta$, respectively. To estimate these parameters for a given genotype, we generated 100 individuals with five successive rounds of random mutations each, and measured the proportion of viable genotypes, $\overline{W}_k$, with $k = 1, 2, \ldots, 5$ mutations. We modelled $\overline{W}_k$ using a generalized linear model with complementary log–log link and a binomial error structure[29]:

$$\log[-\log(\overline{W}_k)] = \log(\alpha) + \beta\log(k).$$

$\overline{W}_1$ was measured directly and β was estimated using maximum likelihood.

Received 1 October; accepted 22 November 2005.

1. Kondrashov, A. S. Deleterious mutations and the evolution of sexual reproduction. *Nature* **336,** 435–440 (1988).
2. Siegal, M. L. & Bergman, A. Waddington's canalization revisited: developmental stability and evolution. *Proc. Natl Acad. Sci. USA* **99,** 10528–10532 (2002).
3. Wagner, A. Does evolutionary plasticity evolve? *Evolution* **50,** 1008–1023 (1996).
4. Lynch, M. *et al.* Perspective: Spontaneous deleterious mutation. *Evolution* **53,** 645–663 (1999).
5. Whitlock, M. C., Phillips, P. C., Moore, F. B. G. & Tonsor, S. J. Multiple fitness peaks and epistasis. *Annu. Rev. Ecol. Syst.* **26,** 601–629 (1995).
6. Burch, C. L. & Chao, L. Epistasis and its relationship to canalization in the RNA virus phi 6. *Genetics* **167,** 559–567 (2004).
7. de Visser, J. A. G. M., Hoekstra, R. F. & van den Ende, H. An experimental test for synergistic epistasis and its application in *Chlamydomonas*. *Genetics* **145,** 815–819 (1997).
8. Mukai, T. The genetic structure of natural populations of *Drosophila melanogaster*. VII. Synergistic interaction of spontaneous mutant polygenes controlling viability. *Genetics* **61,** 749–761 (1969).
9. Whitlock, M. C. & Bourguet, D. Factors affecting the genetic load in *Drosophila*: synergistic epistasis and correlations among fitness components. *Evolution* **54,** 1654–1660 (2000).
10. Bonhoeffer, S., Chappey, C., Parkin, N. T., Whitcomb, J. M. & Petropoulos, C. J. Evidence for positive epistasis in HIV-1. *Science* **306,** 1547–1550 (2004).
11. de Visser, J. A. G. M., Hoekstra, R. F. & van den Ende, H. Test of interaction between genetic markers that affect fitness in *Aspergillus niger*. *Evolution* **51,** 1499–1505 (1997).
12. Elena, S. F. & Lenski, R. E. Test of synergistic interactions among deleterious mutations in bacteria. *Nature* **390,** 395–398 (1997).
13. Michalakis, Y. & Roze, D. Evolution. Epistasis in RNA viruses. *Science* **306,** 1492–1493 (2004).
14. Wilke, C. O. & Adami, C. Interaction between directional epistasis and average mutational effects. *Proc. R. Soc. Lond. B* **268,** 1469–1474 (2001).
15. You, L. & Yin, J. Dependence of epistasis on environment and mutation severity as revealed by in silico mutagenesis of phage T7. *Genetics* **160,** 1273–1281 (2002).
16. Kawecki, T. J. The evolution of genetic canalization under fluctuating selection. *Evolution* **54,** 1–12 (2000).
17. Rice, S. H. The evolution of canalization and the breaking of von Baer's laws: Modeling the evolution of development with epistasis. *Evolution* **52,** 647–656 (1998).
18. Wagner, G. P., Booth, G. & Bagheri-Chaichian, H. A population genetic theory of canalization. *Evolution* **51,** 329–347 (1997).
19. Nijhout, H. F. The nature of robustness in development. *Bioessays* **24,** 553–563 (2002).
20. Wilke, C. O., Wang, J. L., Ofria, C., Lenski, R. E. & Adami, C. Evolution of digital organisms at high mutation rates leads to survival of the flattest. *Nature* **412,** 331–333 (2001).
21. de Visser, J. A. G. M. *et al.* Perspective: Evolution and detection of genetic robustness. *Evolution* **57,** 1959–1972 (2003).
22. Stearns, S. C. The evolutionary links between fixed and variable traits. *Acta Paleontol. Pol.* **38,** 215–232 (1994).
23. Stelling, J., Sauer, U., Szallasi, Z., Doyle, F. J. III & Doyle, J. Robustness of cellular functions. *Cell* **118,** 675–685 (2004).
24. Charlesworth, B. & Barton, N. Recombination load associated with selection for increased recombination. *Genet. Res.* **67,** 27–41 (1996).
25. Wagner, G. P., Laubichler, M. D. & Bagheri-Chaichian, H. Genetic measurement of theory of epistatic effects. *Genetica* **102–103,** 569–580 (1998).
26. Milo, R. *et al.* Network motifs: Simple building blocks of complex networks. *Science* **298,** 824–827 (2002).
27. Barton, N. H. & Keightley, P. D. Understanding quantitative genetic variation. *Nature Rev. Genet.* **3,** 11–21 (2002).
28. Maynard Smith, J. *The Evolution of Sex* (Cambridge Univ. Press, Cambridge, 1978).
29. Crawley, M. J. *Statistical Computing* (Wiley, Chichester, 2002).
30. Jaeger, J. *et al.* Dynamical analysis of regulatory interactions in the gap gene system of *Drosophila melanogaster*. *Genetics* **167,** 1721–1737 (2004).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank T. Flatt, Y. Fofanov, F. Galis, J. Kingsolver, A. Monteiro, M. Travisano and G. Wagner for discussions. The UH, UNC and NIH (grant to C.L.B.) provided financial support.

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to R.B.R.A. (razevedo@uh.edu).