

Supporting Information

“Sexual Reproduction Reshapes the Genetic Architecture of Digital Organisms”

Dusan Misevic, Charles Ofria, and Richard E. Lenski

Proceedings of the Royal Society London B

Phenotypes in Avida. As described in the main text, when an organism performs one of nine basic logic operations on one or two random 32-bit strings and then outputs the bitwise-correct result, it obtains additional energy that accelerates the execution of its genomic program. The logic rules for all the operations are presented in Table S1 below. For example, if bit A = 1 and bit B = 0, then (A OR B) = 1. These rules are defined on single-bit inputs; for an organism to be rewarded for any operation, it must correctly perform it on all 32 bits of the number strings.

Consider an organism that obtained the following two inputs and then output the string below.

Input A: 0101010111000000011101010101100

Input B: 10000110101000111101010110011110

Output : 1101011111100011111111110111110

The organism would receive the energy reward for performing the OR operation, because it correctly calculated the OR function for all 32 pairs of the corresponding bits in Input A and Input B and output the correct result.

The ability to perform a logic operation is scored as a phenotypic trait in the GP maps. Fig. 1c (main text) shows two different organisms that can perform eight of the nine rewarded operations. In these examples, both the sexual and asexual organisms have evolved the ability to perform the NOT, NAND, AND, ORN, OR, ANDN, NOR, and EQU operations, but neither one can perform the XOR operation that is represented by the penultimate column.

Resolving genome-length issues. In the main text, we reported that sexual populations evolved larger genomes, on average, than did asexual populations. This difference was driven by a subset of populations that had evolved very large genomes. Out of the 100 populations with each reproductive mode, 35 sexual populations evolved genomes that averaged at least twice the ancestral length of 50 instructions, while only 9 asexual populations did so. These very large genomes typically evolved via genome doublings, and sexual reproduction in Avida evidently increased the rate of genome doubling events, created more favorable genetic combinations in larger genomes, or both.

To determine whether genome doublings and the resulting difference in average genome length between evolved sexual and asexual populations might bias or confound the comparisons of their genetic architectures, we performed two supplementary analyses described below. The first analysis excluded all those populations that evolved average genome lengths twice or more the ancestral length. The second analysis used 100 additional populations with each reproductive mode, in which genome doublings were prevented during the experiment itself.

Both supplementary analyses eliminated the greater average genome length of the sexual populations and, in fact, they reversed the difference such that the asexual populations had on average significantly longer genomes. Both also eliminated the higher mean fitness values of the sexual populations relative to the asexual populations. However, both supplementary analyses confirmed all the differences in genetic architecture between sexual and asexual populations: (i) sexual genomes were more modular (higher PM and FM); (ii) sexual populations were more robust to the effects of single mutations (lower α); and (iii) asexual populations had stronger net directional epistasis tending to alleviating interactions (lower β).

Reduced data set. We excluded all populations with average genome length ≥ 100 instructions, which is twice the ancestral length. We repeated the original analyses on this reduced data set, which included 65 sexual and 91 asexual populations. The differences between sexual and asexual populations in PM , FM , α , and β remained significant and in the same direction as in the full data set (Table S2 below). However, average genome length was now significantly greater in the asexual organisms, the opposite outcome to the full data set, so that genome length cannot be driving these differences in genetic architecture.

Genome doublings prevented. Analysis of the reduced data set shows that the larger tail of long genomes in the sexual populations was not responsible for the evolved differences in genetic architecture between sexual and asexual organisms. We sought further confirmation by evolving an additional 200 populations with the experimental conditions identical to the original runs in all but one respect: genome doublings were prevented from ever occurring in these experiments by imposing a 10% limit on the difference in genome length between parents and offspring. Whenever a divide instruction (sexual or asexual) was about to be executed, the Avida program checked the genome length of the incipient offspring. If the potential genome was more than 10% longer (or shorter) than its parent's genome, the divide instruction was skipped and the execution proceeded to the next instruction in the genome. Thus, no offspring were ever produced with genome doublings or other radical changes in genome length. Very large genomes did not evolve in these modified runs and, in fact, sexual populations had significantly shorter genomes, on average, than did asexual populations. However, the genetic architectural features PM , α , and β differed significantly and in the same direction as observed in the original experiments (Table S3 below). The difference in FM was no longer significant but remained in the same direction.

Table S1. Truth table for nine logic operations. NOT is performed only on Input A, while the other eight operations are functions of both inputs.

Input		Logic operation								
A	B	NOT	NAND	AND	ORN	OR	ANDN	NOR	XOR	EQU
0	0	1	1	0	1	0	0	1	0	1
0	1	1	1	0	0	1	0	0	1	0
1	0	0	1	0	1	1	1	0	1	0
1	1	0	0	1	1	1	0	0	0	1

Table S2. Comparisons of properties between sexual and asexual evolved populations in the reduced data set. The p values are based on two-tailed t -tests.

response variable	mean sexual (\pm s.d.)	mean asexual (\pm s.d.)	p
\log_{10} fitness	4.835 (1.230)	5.033 (1.104)	0.303
\log_{10} genome length	1.752 (0.100)	1.826 (0.059)	< 0.001
physical modularity, PM	0.775 (0.057)	0.689 (0.066)	< 0.001
functional modularity, FM	0.803 (0.084)	0.761 (0.077)	0.002
average effect of single point mutations on fitness, α	0.724 (0.173)	0.792 (0.176)	0.019
net directional epistasis, β	0.904 (0.051)	0.854 (0.088)	< 0.001

Table S3. Comparisons of properties between sexual and asexual evolved populations with genome doublings prevented. The p values are based on two-tailed t -tests.

response variable	mean sexual (\pm s.d.)	mean asexual (\pm s.d.)	p
\log_{10} fitness	4.678 (1.107)	4.944 (1.242)	0.112
\log_{10} genome length	1.703 (0.013)	1.845 (0.061)	< 0.001
physical modularity, PM	0.717 (0.064)	0.694 (0.073)	0.021
functional modularity, FM	0.771 (0.078)	0.761 (0.080)	0.356
average effect of single point mutations on fitness, α	0.765 (0.162)	0.827 (0.183)	0.004
net directional epistasis, β	0.901 (0.056)	0.828 (0.107)	< 0.001