The Evolution of Genetic Robustness for Cellular Cooperation in Early Multicellular Organisms

Katherine G. Skocelas[1,2,3] Austin J. Ferguson[1,2,3] Clifford Bohm[3,4] Katherine Perry[1,3]
Rosemary Adaji[3,5] Charles Ofria[1,2,3]

[1]Department of Computer Science [2]Program in Ecology, Evolution, and Biology
[3]BEACON Center for the Study of Evolution in Action [4]Department of Integrative Biology [5]Department of Epidemiology
Michigan State University, East Lansing, MI 48824
kgs@msu.edu

Abstract

The major evolutionary transition to multicellularity shifted the unit of selection from individual cells to multicellular organisms. Constituent cells must regulate their growth and cooperate to benefit the whole organism, even when such behaviors would have been maladaptive were they free living. Mutations that disrupt cellular cooperation can lead to various ailments, including physical deformities and cancer. Organisms therefore employ mechanisms to enforce cooperation, such as error correction, policing, and genetic robustness.

We built a simulation to study this last mechanism under a range of evolutionary conditions. Specifically, we asked: How does genetic robustness against cellular cheating evolve in multicellular organisms? We focused on early multicellular organisms (with only one cell type) where cells must control their growth to avoid overwriting each other. In our model, unrestrained cells will outcompete restrained cells within an organism, but restrained cells alone will result in faster reproduction for the organism. Ultimately, we demonstrate a clear selective pressure for genetic robustness in multicellular organisms and show that this pressure increases with the total number of cells in the organism.

Introduction

Multicellular organisms have needed to coordinate cellular activity since the origin of multicellularity three and a half billion years ago (Callier, 2019). Within these organisms, cells must cooperate with copies of themselves for the higher-level organism to function and reproduce (Smith and Szathmary, 1997; Calcott and Sterelny, 2011; Queller, 2000). The clonal nature of cells in an organism ensures that kin selection aligns cellular and organismal goals, but mutants can arise that do not engage in the cooperative behaviors. If those mutants also replicate more rapidly, they can disrupt cooperation, reducing the fitness of the organism, possibly leading to cancer or death.

A variety of techniques are used by multicellular organisms to prevent defection from cooperation. These include *policing* (monitoring cellular behavior and punishing or killing cells that fail to cooperate properly), *apoptosis* (cellular suicide to avoid engaging in harmful behaviors), *error correction* (repairing mutations before they can cause harm), and genetic robustness (reducing the probability of harmful effects from mutations that do occur). Here, we focus on this final technique and examine the selective pressures by which multicellular organisms evolve robustness to mutational effects (by means of simple redundancy) in order to preserve cooperation.

Genetic robustness preserves phenotypic traits despite mutational disruption, typically via redundancy or compensatory processes (De Visser et al., 2003), and is prevalent throughout nature. Knockout experiments in yeast, for example, have demonstrated that up to half of all genes can be individually deactivated with minimal impact on fitness (Thatcher et al., 1998). Significant attention has therefore been given to the study of genetic robustness (Kitano, 2004; Lauring et al., 2013; Masel and Siegal, 2009; Lenski et al., 2006), including the role of redundancy (Gu et al., 2003; Láruson et al., 2020).

In designing our system, we choose to focus on one of the most simple forms that cellular cooperation takes: avoiding killing neighboring cells during proliferation. Indeed, inhibition of cellular replication is a hallmark of a major evolutionary transition, as a collection of cells begins to act as a higher-level individual by sacrificing their immediate replication potential for the good of the whole (Calcott and Sterelny, 2011).

Each cell division in a multicellular organism has a chance of incurring a mutation that reduces the daughter cell's ability to inhibit its own proliferation. As such, evolution selects for robustness in cooperative reproductive strategies, often with genetic redundancies to ensure that cells maintain their limits on cellular proliferation. All else equal, organisms that experience more cell divisions in their lifetimes (via having more cells or living longer) are more likely to accumulate mutations that cause cells to proliferate out of control. As such, we expect larger and longer-lived organisms to have greater selective pressures for redundancies in cellular controls.

Indeed, this type of loss of inhibited cellular proliferation is often studied in the context of cancer. As described by Nunney (1999), "Cancer occurs because the genetic control

of cell growth is vulnerable to somatic mutations [...], particularly in large, continuously dividing tissues." Despite undergoing more cell divisions in their lifetimes, however, larger and longer-lived species do not typically exhibit proportionally increased rates of cancer, a phenomenon known as Peto's paradox (Peto, 1977). In fact, Nunney (1999) statistically demonstrated that the mechanisms to avoid mutations or cancer-causing mutational effects in small and large species are so different that the mechanisms to prevent cancer in one would not be evolutionarily stable in the other. Mechanisms used by smaller organisms would be ineffective in larger organisms, while mechanisms in larger organisms would be too expensive to sustain in smaller organisms.

Does this effect occur in more primitive circumstances? We simulated asexual multicellular organisms composed of cells as they would have existed shortly after the transition to multicellularity, before any developmental processes or division of labor evolved. Cells must inhibit their growth to avoid killing (overwriting) neighboring cells, which would reduce organism fitness. The only mechanism that evolution has to work with is genetic robustness, implemented as redundancy of the inhibitory behavior. We ask: Does the selective pressure for inhibited cellular proliferation during cellular reproduction increase with organism size? Specifically, do larger organisms evolve more redundancy for inhibited cellular proliferation on average?

Simulation

To test our hypothesis, we simulated evolving populations of multicellular organisms in a system that we named Primordium. Computer simulations of multicellular organisms have often been utilized to study the relationship between intra- and inter-cellular competition (Goldsby et al., 2014a,b; Pfeiffer and Bonhoeffer, 2003; Moreno and Ofria, 2022; Rose et al., 2020). We focused solely on organism growth (tissue accretion), avoiding complexities associated with more evolved multicellular life, such as specialized cell types and developmental patterns. Each organism begins as a single cell near the middle of a square grid; as soon as the grid is filled, the organism reproduces. As such, the size of the grid represents the organism's body size, and the speed of filling the grid is the organism's fitness. Within an organism, each cell attempts to replicate into a random neighboring grid position. An unrestrained cell will always place its offspring in the selected grid position, even if doing so replaces and kills an existing cell. A restrained cell will abort replication rather than replace a neighbor, since replacement would restart replication from that position. In practice, an organism consisting of restrained cells reproduces up to 20% faster than an organism with all unrestrained cells. Organisms that begin with a restrained cell but develop unrestrained cells during their lifetime (due to somatic mutations) end up with intermediate fitness values depending on how early unrestrained cells arose.

A cell's level of restraint is determined by its genome of zeros and ones. Organisms also have a sequestered germ cell whose genome is used to produce its initial somatic cell and is inherited by any offspring organisms. A cell is restrained if the number of ones in its genome is greater than a *restraint threshold* (60% of the genome in this work). The number of ones beyond the restraint threshold is termed the *restraint value* of that cell. A somatic mutation may induce a bit flip when a cell replicates, changing the daughter cell's restraint value. Likewise, germ-level mutations may occur when a whole organism reproduces. Since the restraint value of an organism's germ cell is inherited, it is under selection at the population level; we term this value the *restraint buffer* of the organism since it sets the initial cell's restraint value. A higher restraint buffer means that more somatic mutations can be sustained by cells (on average) before restraint is lost. (For full details on Primordium, see Methods below).

Results overview

In experiments with Primordium, we demonstrated that larger organisms have a stronger selective pressure for high restraint buffers, but many complicating factors exist. These include that mutation-selection balance can have a profound effect on the outcome of evolution, especially when restraint mechanisms require substantial genetic material. While larger organisms benefit more from a higher restraint buffer, the fitness benefit they gain declines with each additional step in restraint, making it more difficult for evolution to achieve. Thus, while we see that increasing organism size heightens selective pressure for large restraint buffers, we also find that it can easily be countered by mutational drift or the effects of a noisy fitness function. We conclude that selection technically favors higher levels of restraint in larger organisms, but other factors may prevent evolution from realizing those levels. Only under perfectly idealized conditions are we able to observe a positive relationship between organism size and restraint buffer value that continues into the largest sizes, indicating that additional factors may need to be present in nature.

Methods

Below, we detail the implementation of Primordium (Ferguson et al., 2022), the experimental methods used for data collection, and the statistical analyses performed on the results.

The Primordium evolution system

Primordium is a digital system that simulates an evolving, well-mixed population of multicellular organisms, useful for investigating how organism size influences the evolution of cellular restraint. Each organism maintains a sequestered germ cell that is used to initialize its body (with a single somatic cell, or 'soma') and for the genetic material passed to its offspring. At birth, the initial soma is placed near

the center of a two-dimensional square grid (see Figure 1). Cells copy themselves into neighboring grid positions and are subject to somatic mutations with each replication. An organism reproduces when its entire grid is filled with cells. The population of organisms is kept at a constant size, with new offspring always replacing a random existing organism; this replacement is the only mechanism for organism death. As such, organisms that reproduce faster are more likely to produce offspring before being overwritten, creating a selective pressure for organisms that fill with cells as rapidly as possible.

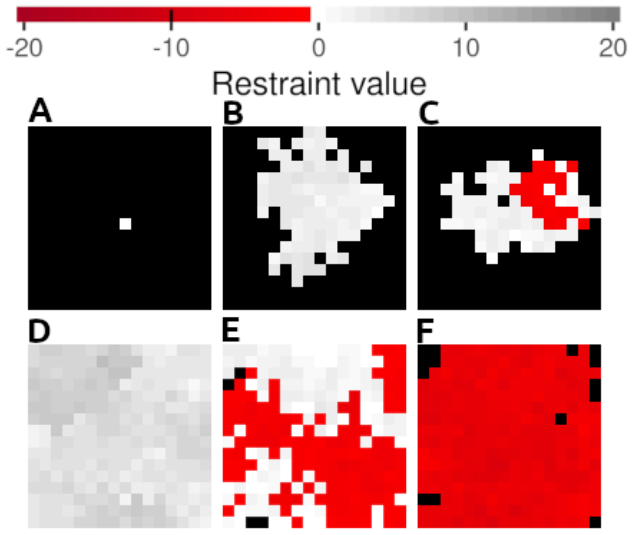

Figure 1: A visualization of six example 16x16 organisms. The color scale for cells is shown above, and empty cells are shown in black. Organism A is a brand new orgnaism. Organisms B and C have the same restraint buffer (3) and have both grown for 300 updates, but unrestrained cells have only appeared in organism C. The second row shows three organisms with restraint buffers of 5 (D), 2 (E), and -5 (F). All three organisms have received 2,732 updates, which is how long it took for organism D to fill and thus reproduce.

Inside an organism, each cell has a replication timer that starts when the cell is born. A cell's replication time is calculated as a base number of cycles (100) plus a noise factor (uniform between zero and 50 cycles) to stagger replication. When the timer elapses, the cell randomly chooses an offspring position from the eight options in its Moore neighborhood. If the selected position is empty, the cell replicates into it. If the position is occupied, however, the behavior depends on whether the parent cell's genome encodes for restraint. If the cell is *restrained*, the offspring cell is discarded, leaving the neighbor unaffected; if the cell is *unrestrained*, the neighboring cell is overwritten by the offspring, which then starts replicating itself from the beginning. In either case, the parent cell's timer is then reset to zero and a new replication time is calculated.

Each genome is a bitstring (length 100 by default) that determines if its cell is restrained. A *restraint threshold* indicates the minimum number of ones in the bitstring required for the cell to be restrained (here, 60% of genome length). While restrained behavior is determined merely by whether the number of ones in a genome is greater than the threshold, we also measure a cell's *restraint value* as the difference between these values. For restrained cells, this value indicates how many restraint-reducing mutations can be sustained before restraint is lost. For unrestrained cells, restraint value is negative, indicating how far away from restraint the cell is. Note that the data in a cell's genome is fully encapsulated in its restraint value, as order does not matter. Mutations do, however, function as they would in a bitstring; for example if a length 100 genome has 70 ones, that corresponds to a 30% chance of a mutation increasing the number of ones and a 70% chance of it decreasing.

Given that the default restraint threshold requires at least 60% ones in a genome, somatic mutations will reduce restraint on average; as such, mutation accumulation during soma growth moves toward unrestrained behavior. Since restrained cells never overwrite their neighbors, they are at a competitive disadvantage against any unrestrained cells that may appear.

Organism reproduction also has some probability of mutation when the parent organism's sequestered germ cell is passed to the child organism. This probability is parameterized as the germ mutation rate; we used, by default, a 2% chance per organism reproduction event. If a mutation occurs, a single bit in the germ cell's genome is flipped, increasing or decreasing the offspring's germ cell restraint value by one. To create a distinction between the cellular and organismal levels, we record the restraint value of the germ cell as the organism's *restraint buffer*. The restraint buffer provides an indication of how large the organism can grow before mutation accumulation starts producing unrestrained cells. Thus, different restraint buffers can result in different reproduction times for organisms and is the only factor subject to evolution in between-organism competition. Larger organisms undergo more cell replications and therefore require higher restraint buffers to avoid unrestrained cells arising within their lifetime.

Simulation controls for infinite population sizes and genome lengths

To exclude the possibility that our results were caused by insufficient population size or genome length, we performed controls to simulate infinite-size populations and infinite-length genomes.

In a finite genome, a bit flip from a zero to a one would decrease the number of zeros and thus decrease the likelihood that subsequent mutations would hit a zero. We simulated an infinite genome by removing this feedback on sub-

sequent probabilities; specifically, we locked both germ and somatic mutations to a 60% probability of reducing restraint, which is the same probability as finite genomes at the restraint threshold. We also removed all limits on restraint values and restraint buffers.

To simulate an infinite population, we converted the principles of Primordium into a population genetics model. For each replicate, we first used Primordium to gather the distribution of organism fitness at each restraint buffer value. We then applied an iterative formula to determine the restraint buffer distribution after a given number of generations (based on models of asexual haploids in discrete, non-overlapping generations (Crow and Kimura, 1970)). Each generation was represented by the portion of the population with each restraint buffer. The weighted proportion of offspring was initially determined as the current proportion times its (empirically measured) expected fitness. After normalizing these values (such that all proportions again add up to one), we accounted for mutations, moving an appropriate portion of each group to a restraint buffer category of plus or minus one. Full details of the model can be found in Section 10 of the supplement (Ferguson et al., 2022).

Experiment design

In our baseline experiment for this work, we examined the effect of six organism sizes (16x16, 32x32, 64x64, 128x128, 256x256, and 512x512) on the evolution of restraint using all default parameters. We conducted more limited studies at organism size 1024x1024, but these experiments proved too slow to include for all results and are included in supplement Sections 2 and 9 (Ferguson et al., 2022). The initial experiment produced results partially opposing our hypothesis, so we then conducted additional experiments to tease apart the underlying dynamics. Specifically, we analyzed the importance of the germ mutation rate, somatic mutation rate, genome length, and population size.

We experimentally determined the default values for each parameter by conducting preliminary experiments where we swept each value (data available in supplement Sections 3-6 (Ferguson et al., 2022)). By default, all of our experiments consisted of a population of 200 multicellular organisms evolving for 10,000 generations with 100-bit genomes, and each treatment was replicated 100 times with different random number seeds. The restraint threshold was always set such that 60% of the genome must be ones in order to confer restraint. All organisms in the population begin with a restraint buffer of zero (*i.e.*, evolution always begins at the restraint threshold).

With Primordium's default parameters, there is a 50% chance that a single somatic mutation will occur when an individual cell replicates and a 2% chance of a single germ mutation when a whole organism reproduces. Both types of mutations toggle a single bit and will therefore either increase or decrease the restraint value of the genome by one.

To make the computational costs feasible, we pre-generated the replication time data for organisms before evolution. We simulated 100 organisms at each possible combination of restraint buffer value and organism size to produce a distribution of the time required for the organism to reproduce. During evolution, each time an organism is born, we pull from these distributions to determine when it will reproduce instead of simulating each individual cell. Comparison experiments demonstrated that this limited number of fitness samples has no qualitative effects on the overall evolution of restraint (see supplement Section 5 (Ferguson et al., 2022)).

At the beginning of the simulation, all organisms are given a generation value of zero. When an organism reproduces, the offspring's generation is set to its parent's generation plus one. Every time the average generation value of organisms in the population surpasses a whole number, we collect the average restraint buffer of all organisms in the population. Examining these values from preliminary data, we determined that populations had stabilized by the time the average generation crossed 10,000 (*i.e.*, running the simulation longer produced no additional changes in the evolved restraint buffer values). Most analyses focused on the average restraint buffers at the end of 10,000 generations of evolution. Additionally, we analyzed the pre-generated replication times for organisms under various configurations to determine how different parameters affected fitness.

Statistics and data availability

All statistics were calculated by first performing a Kruskal-Wallis test to determine if significant variation existed across treatments. When significance was indicated, we determined which treatments were significantly different with a pairwise Wilcoxon Rank-Sum test and Bonferroni-Holm corrections for multiple comparisons. Statistics have been included in the figures where appropriate, and all statistics are available in the supplement (Ferguson et al., 2022). Primordium was developed with the Empirical C++ library for scientific software (Ofria et al., 2020). Analyses and visualizations were conducted using R version 3.6.3 (R Core Team, 2020) and the ggplot2 library (Wickham, 2016). All source code, analyses, and other supplemental material can be found on GitHub (Ferguson et al., 2022).

Results & Discussion

For our baseline analysis, we examined the evolved restraint buffer for a range of organism sizes. Larger multicellular organisms, by definition, undergo more cellular replication on average. As such, a higher restraint buffer is required to avoid unrestrained cells appearing during organism growth and slowing reproduction. Indeed, we found that (on average) the evolved restraint buffer initially increases with organism size. A turning point emerges at size 128x128, however, beyond which the evolved restraint buffer decreases, in opposition to our expectation (see Figure 2).

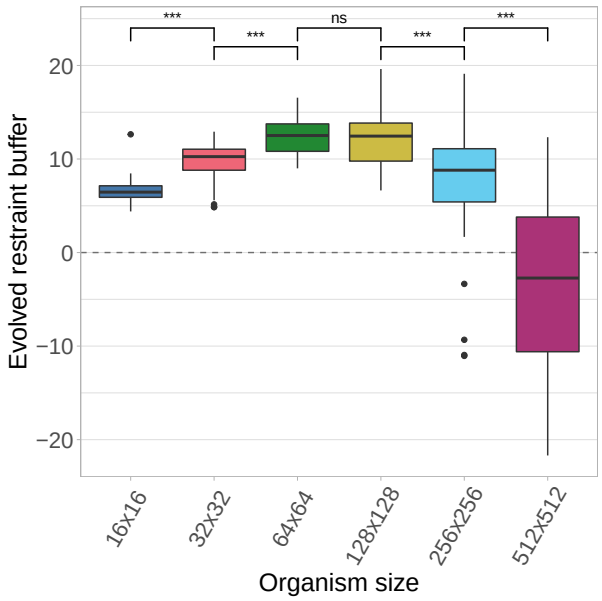

Figure 2: Boxplots show the evolved restraint buffers of populations with varying organism sizes and 100-bit genomes. Each boxplot represents 100 independent replicates. Each replicate is summarized as the average restraint buffer of all organisms at the end of 10,000 generations. Colors are unique for each organism size and consistent across all figures. Statistics between adjacent organism sizes are shown (for all figures: 'ns' for $p > 0.05$, '*' for $p \leq 0.05$, '**' for $p \leq 0.01$, '***' for $p \leq 0.001$)

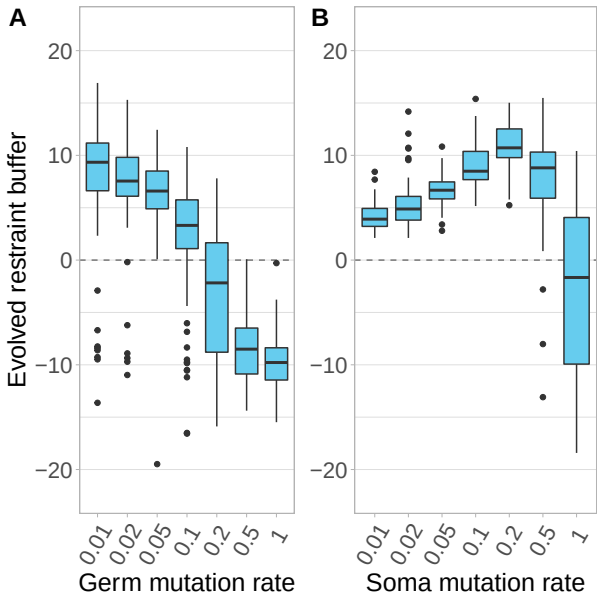

Figure 3: Boxplots show the evolved restraint buffers of populations of 256x256 organisms evolved under varying A) germ and B) somatic mutation rates. Both mutation rates are quantified per-genome. Each boxplot represents 100 independent replicates. Each replicate is summarized as the average restraint buffer of all organisms at the end of 10,000 generations.

Due to the simplicity of this system, there are only two key pressures acting on the population: selection and mutational drift. Selection acts on both the population level (improved organism restraint can speed up reproduction) and the cellular level (unrestrained cells spread faster). Mutational pressures are always biased toward an equal balance of zeros and ones, which is below the 60% ones needed for restraint. At the population level, this biased mutational pressure acts counter to selection, pulling down germ restraint buffers. Mutational pressure from somatic mutations provides a selective advantage for organisms with a higher restraint buffer, as their cells remain restrained for longer. Where mutation-selection balance leaves a genome at the end of evolution depends on the strength of each pressure and the nuances of their interactions. Below, we disentangle how these pressures interact across levels to create the observed peak and subsequent decline in evolved restraint values.

Mutation rates for germ and somatic cells push evolution in opposing directions

Mutations to organism germ cells provide variation for evolution to act upon. If the mutation rates for these cells are too high, however, they create a strong mutational pressure for less restrained organisms, overwhelming selection. Figure 3A shows how evolution is less capable of producing high restraint buffers as germ mutation rates increase; 256x256 organisms are used for illustrative purposes, but a qualitatively similar effect can be seen at all organism sizes (see Section 4 of the supplement (Ferguson et al., 2022)).

Conversely, resistance to somatic mutations is why higher levels of restraint are valuable in the first place. If the somatic mutation rate drops too low, then a high restraint buffer is no longer valuable and is not selected by evolution. Figure 3B demonstrates this effect by measuring evolved restraint buffers in tests of 256x256 organisms at various somatic mutation rates. If the somatic mutation rate is too high (above 0.2), however, we see that selection is no longer able to counter mutational decay and the organisms evolve smaller restraint buffers than they did at lower somatic mutation rates. Data for all organism sizes are available in Section 3 of the supplement (Ferguson et al., 2022))

In considering how these effects play out in nature, germ cells are sequestered in most organisms in order to keep their mutation rates low. Somatic cells, on the other hand, do all of the dirty work for the body and thus tend to be subjected to mutation rates two orders of magnitude higher (Milholland et al., 2017). As such, the combination of mutation

rates that we use as our default parameters appear to be realistic for natural evolution.

Longer genomes reduce mutational pressure

In our baseline experiment, each genome consists of 100 bits, with 60 bits (60%) needing to be ones for a cell to be restrained. As such, a cell with a restraint value of zero (exactly 60 ones) would have a 60% chance of a mutation eliminating its restraint and a 40% chance of increasing it. A restraint value of 10 would shift these values to a 70% chance of reduction and only a 30% chance of increase.

We examined the comparable situation in genomes of other lengths, from 25 bits to 400 bits in length, each with a 60% threshold for restraint. A cell with a length 400 genome and a restraint value of 10 has 250 ones in it, and thus only a 62.5% chance of a restraint-reducing mutation and a 37.5% chance of a mutation increasing restraint. At the other extreme, a cell with a length 25 genome and a restraint buffer of 10 consists of all ones, and so it has a 100% chance of a mutation reducing its restraint.

What effect should a longer genome have on evolution? At the population level, it should reduce the incremental mutational pressure against restraint, and thus allow higher restraint buffers to evolve in organisms. Within cells, however, a longer genome also decreases the probability of restraint-reducing somatic mutations, reducing the rate at which restraint decays, and thus weakening selection for high restraint buffers. We created an additional control that isolated the change in mutational pressure, but the results were qualitatively identical to those described below (see Section 8 of supplement for details (Ferguson et al., 2022)).

Figure 4A shows that, when we evolve 256x256 organisms with differing genome lengths, longer genomes result in the evolution of higher restraint buffers. The range of evolved restraint buffers also increases with genome length, because all genome lengths have results that extend to the median of the genome, which is at a lower restraint buffer in longer genomes. Indeed, examining length-400 genomes at all organism sizes, we see in Figure 4B that restraint buffers now peak at organisms of size 256x256, but drop again for the largest organisms.

To ensure these trends were not the result of insufficient genome length, we repeated the experiment with infinite genomes (*i.e.*, all mutations have a 60% chance of decreasing restraint and restraint buffers are not capped). These results were qualitatively the same. While the "peak" organism size increased, the constant mutational pressure was still strong enough to cause a downturn in evolved restraint at the largest organism sizes (see Section 9 of supplement (Ferguson et al., 2022)).

Natural genomes are, of course, not infinite, yet they are far longer than the finite genomes used in this study, though only a small portion of a natural genome is related to cellular restraint. The system we use is also far simpler than the

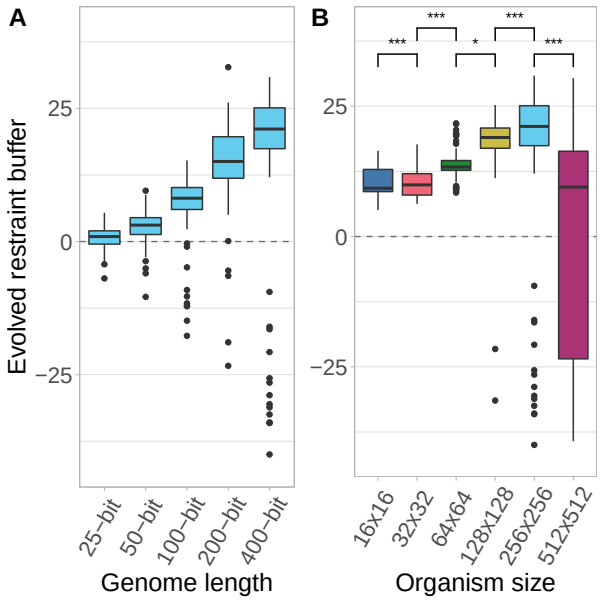

Figure 4: Subplot A shows the evolved restraint buffer for each genome length with a 256x256 organism, while subplot B shows the evolved restraint buffer at each organism size for a 400-bit genome. Each boxplot represents 100 independent replicates. Each replicate is summarized as the average restraint buffer of all organisms at the end of 10,000 generations.

complex regulatory networks involved in real-world organisms, and in future work it will be important to examine how those complications play out.

The selective pressure for restraint has diminishing returns

Next, we analyzed how the selective pressure at different restraint values changes with organism size. While each successive increment to the restraint buffer proved to be beneficial, we observed a diminishing return in that benefit in practice (see Figure 5). Small organisms see the greatest fitness boost just above the restraint threshold, after which the fitness advantage for additional restraint quickly diminishes once an organism's restraint buffer is high enough that unrestrained cells never appear during its lifetime. Larger organisms have a smaller initial spike, but given that they experience more cell divisions, the saturation point occurs at larger restraint buffer values. Thus, selective pressure persists longer in larger organisms. In fact, in the largest organisms (256x256 and 512x512), this saturation point is unreachable with the 100-bit genome.

While the fitness data shows that additional bits of restraint still provide a small selective advantage to larger organisms (Figure 5), the evolved restraint buffers decline for organisms larger than 128x128 (Figure 2). We know that

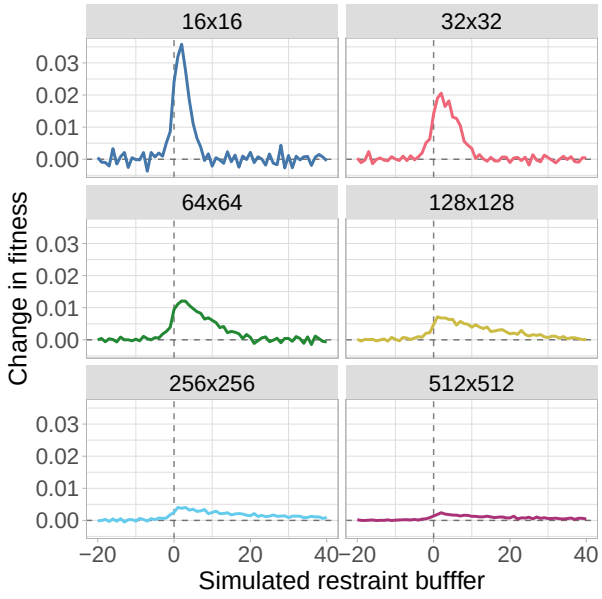

Figure 5: Lines show the average benefit of each additional bit of restraint buffer for various organism sizes. Values are calculated as the measured difference in fitness between n bits of restraint and $n-1$ bits of restraint, and each restraint buffer value was averaged over 10,000 samples. Larger values on the y-axis indicate a greater increase in benefit. Data were calculated to a restraint value of -60, but all values below -20 fluctuate around zero across all organism sizes.

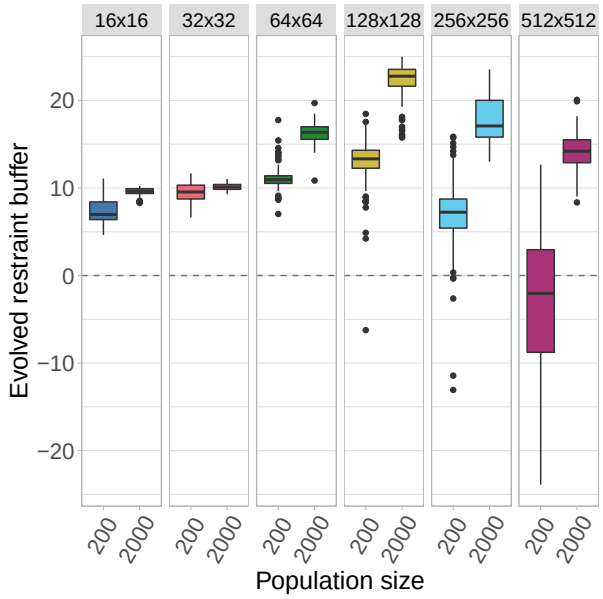

Figure 6: Boxplots show the evolved restraint buffers of populations with 100-bit genomes. Each subplot shows a particular organism size evolved with two population sizes: 200 organisms (left boxplot, default) and 2,000 organisms (right boxplot). Each boxplot represents 100 independent replicates, each summarized as the average restraint buffer of all organisms at the end of 10,000 generations. All comparisons between population sizes of 200 and 2,000 organisms are highly significant ($p \leq 0.001$).

populations are expected to evolve to the restraint value where mutational and selection pressures are in equilibrium. Thus, the selection pressure for additional bits of restraint must have become too weak to oppose the mutational pressure against additional bits. From evolutionary theory, we know that population size greatly affects selective pressure, with larger populations experiencing increased selective pressure (Gossmann et al., 2012). Thus, we replicated the original experiment with larger populations to observe the results of increased selective pressure. Indeed, Figure 6 shows that increasing the population size from 200 to 2,000 increased the evolved restraint buffer at every organism size. Even with this larger population, however, the evolved restraint buffers still decrease at sizes greater than 128x128.

Natural populations are often huge, so we extended this experiment with a population genetics model to simulate the selection pressure in an infinite population. In Figure 7A, we see that populations of size 256x256 organisms evolve more restraint than the 128x128 populations, but the 512x512 populations still evolve less restraint. Therefore we conclude that the mutational pressure is a limiting factor in the evolution of restraint in our system.

When all other factors are controlled for, larger organisms evolve greater restraint

Even with the increased selective pressure of an infinite population size, the largest organisms are unable to overcome the mutational pressure of the finite genome to evolve more restraint than the next largest organism size. Thus, we asked if combining the infinite population with an infinite genome to neutralize mutational pressure would be sufficient to continue the expected trend. Figure 7B shows that, indeed, when both the infinite population and infinite genome controls are in place, the evolved restraint buffer continues to increase with organism size. Since this combination of controls changes the trend, we conclude that the downturn pattern must be a combination of increasing mutational pressure and decreasing selective pressure as restraint buffers increase.

In nature, populations are not unlimited, but are often much larger than the size-2,000 populations that we tested. Similarly, while biological genomes are not unlimited, they are typically significantly larger than modeled here. Thus, our system's infinite population and genome model highlights factors that are clearly important for these dynamics, but are reasonable to assume that they would exist in nature.

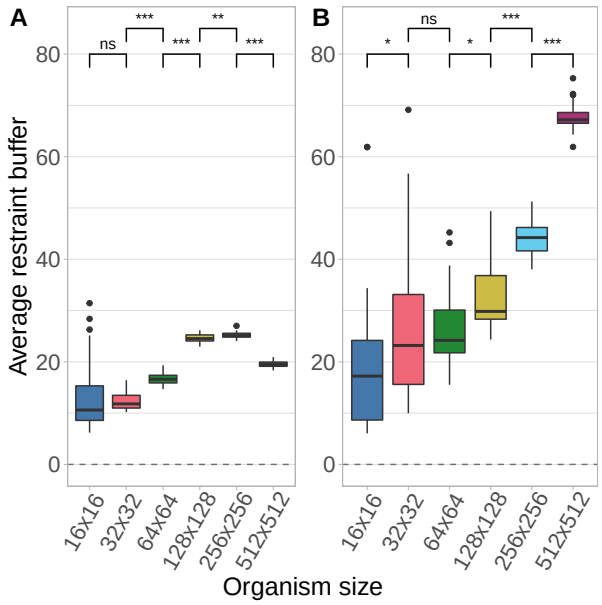

Figure 7: Boxplots show the average restraint buffer value after 10,000 updates in an infinite population model. Each boxplot represents 100 independent replicates. Subplot A shows the results for the default 100-bit genome, while subplot B shows the results for the infinite genome.

Conclusions

We have shown evidence that the pressure imbued by simple space management can select for genetic robustness to improve cooperation. Larger organisms optimize their fitness advantage if their genomes are more mutations away from an unrestrained state. After an initial transition to multicellularity, these dynamics appear sufficient to create a selective pressure for increased restraint on cellular replication without the need for developmental patterning.

While larger organism sizes do have increased selective pressure for larger restraint buffers, we have also found that this is not a strong effect. As the restraint buffer increases, so too does mutational pressure. Larger organisms can have mutational pressure so strong (and selective pressure so weak) that the evolved restraint buffer actually decreases.

In Primordium, being unrestrained results in at most a 20% fitness loss for an organism, making it possible to overcome with other pressures. More complex aspects of multicellular organisms (such as developmental patterns) would make unrestrained cellular behaviors more harmful or even fatal. As a result, "the cells of multicellular organisms, even those with body plans as simple as sponges, have evolved mechanisms to maintain appropriate numbers of cells within tissues" (DeGregori, 2011). Avoiding the turning point that we saw in Primordium required unrealistic controls, indicating that there must, indeed, be other factors in natural systems increasing the selective pressure for restraint. The dynamics we observed, however, may help explain the initial bootstrap as multicellular organisms first evolved to be large enough for developmental processes to become beneficial. The benefit to biological organisms of being larger also applies a strong indirect selective pressure for restraint. Dedicating more energy and more of their genome to restraint mechanisms imposes a cost on organisms, but helps not only in avoiding diseases such as cancer, but also facilitates achieving larger body masses. A larger body size can improve predation success, defense against predation, range of food sources, mating success, longevity, and intelligence (with increased brain size) (Hone and Benton, 2005).

From the perspective of cancer research, it is clear that Peto's paradox (*i.e.*, larger/longer-lived species are expected to exhibit proportionally higher rates of cancer, but do not in practice) is the result of many evolutionary forces and dynamics (Peto, 1977). Here, we focused only on the pressure imbued by the management of space as a resource, disallowing the evolution of most parameters (organism size, genome length, cell replication strategies, *etc.*).

Furthermore, Primordium uses binary genomes and a simple restraint mechanism. In a real-world biological system, restraint would be harder to build than destroy—worsening mutational pressures—but a smaller region of the genome would encode for it, making it a smaller mutational target.

Many of these complications would be exciting to study to identify their effects on genetic robustness in multicellular organisms. For example, we should examine the effect of a more complex genome alphabet and genes that code for functional fitness. Primordium would also be ideal to study the evolution of genome length, with large genomes allowing for increased cellular robustness, trading off against a higher mutational load.

In nature, multicellular organisms can be huge, despite undergoing vastly more cell divisions. Indeed, the largest (blue whales) coordinate quadrillions of cells. Deciphering how this is possible will allow us to not only better understand our natural world, but also give us insights on how to evolve larger and more complex artificial organisms.

Acknowledgements

We thank Heather Goldsby, Acacia Ackles, Ben Kerr, the members of the MSU Digital Evolution Lab, and prior anonymous reviewers for their contributions. This work was supported by U.S. National Science Foundation grant DEB-1655715, the BEACON Center for the Study of Evolution in Action, and the Department of Defense (DoD) through the National Defense Science & Engineering Graduate (NDSEG) Fellowship Program. Computational resources were provided by the MSU Institute for Cyber-Enabled Research.

References

Calcott, B. and Sterelny, K., editors (2011). *The Major Transitions in Evolution Revisited*. Vienna Series in Theoretical Biology. MIT Press, Cambridge, MA.

Callier, V. (2019). Core Concept: Solving Peto's Paradox to better understand cancer. *Proceedings of the National Academy of Sciences*, 116(6):1825–1828.

Crow, J. and Kimura, M. (1970). An introduction to population genetics theory. Harper and Row, Publishers.

De Visser, J. A. G. M., Hermisson, J., Wagner, G. P., Meyers, L. A., Bagheri-Chaichian, H., Blanchard, J. L., Chao, L., Cheverud, J. M., Elena, S. F., Fontana, W., Gibson, G., Hansen, T. F., Krakauer, D., Lewontin, R. C., Ofria, C., Rice, S. H., Dassow, G. v., Wagner, A., and Whitlock, M. C. (2003). Perspective: Evolution and Detection of Genetic Robustness. *Evolution*, 57(9):1959–1972.

DeGregori, J. (2011). Evolved Tumor Suppression: Why Are We So Good at Not Getting Cancer? *Cancer Research*, 71(11):3739–3744.

Ferguson, A. J., Ofria, C., and Skocelas, K. (2022). Primordium Supplemental Material. GitHub. https://doi.org/10.5281/zenodo.6564985.

Goldsby, H. J., Knoester, D. B., Kerr, B., and Ofria, C. (2014a). The effect of conflicting pressures on the evolution of division of labor. *PLoS ONE*, 9(8):e102713.

Goldsby, H. J., Knoester, D. B., Ofria, C., and Kerr, B. (2014b). The evolutionary origin of somatic cells under the dirty work hypothesis. *PLoS Biology*.

Gossmann, T. I., Keightley, P. D., and Eyre-Walker, A. (2012). The effect of variation in the effective population size on the rate of adaptive molecular evolution in eukaryotes. *Genome Biology and Evolution*, 4(5):658–667.

Gu, Z., Steinmetz, L. M., Gu, X., Scharfe, C., Davis, R. W., and Li, W.-H. (2003). Role of duplicate genes in genetic robustness against null mutations. *Nature*, 421(6918):63–66.

Hone, D. W. and Benton, M. J. (2005). The evolution of large size: How does Cope's Rule work? *Trends in Ecology and Evolution*, 20(1):4–6.

Kitano, H. (2004). Biological robustness. *Nature Reviews Genetics*, 5(11):826–837.

Lauring, A. S., Frydman, J., and Andino, R. (2013). The role of mutational robustness in RNA virus evolution. *Nature Reviews Microbiology*, 11(5):327–336.

Lenski, R. E., Barrick, J. E., and Ofria, C. (2006). Balancing Robustness and Evolvability. *PLoS Biology*, 4(12):e428.

Láruson, A. J., Yeaman, S., and Lotterhos, K. E. (2020). The Importance of Genetic Redundancy in Evolution. *Trends in Ecology & Evolution*, 35(9):809–822.

Masel, J. and Siegal, M. L. (2009). Robustness: mechanisms and consequences. *Trends in Genetics*, 25(9):395–403.

Milholland, B., Dong, X., Zhang, L., Hao, X., Suh, Y., and Vijg, J. (2017). Differences between germline and somatic mutation rates in humans and mice. *Nature Communications*, 8(1):1–8.

Moreno, M. A. and Ofria, C. (2022). Exploring Evolved Multicellular Life Histories in a Open-Ended Digital Evolution System. *Frontiers in Ecology and Evolution*, 10:750837.

Nunney, L. (1999). Lineage selection and the evolution of multistage carcinogenesis. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 266(1418):493–498.

Ofria, C., Moreno, M. A., Dolson, E., Lalejini, A., Papa, S. R., Perry, K., Boyd, S., Fenton, J., Jorgensen, S., Hoffman, R., Miller, R., Edwards, O. B., Stredwick, J., Clemons, R., Vostinar, A., Moreno, R., Nitash C G, Zaman, L., Schossau, J., and Rainbow, D. (2020). Empirical. https://doi.org/10.5281/zenodo.2575606.

Peto, R. (1977). Epidemiology, multistage models, and short-term mutagenicity tests. In Hiatt HH, Watson JD, W. J., editor, *The origins of human cancer, Cold Spring Harbor conf. on cell proliferation.*, pages 1403–1428. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Pfeiffer, T. and Bonhoeffer, S. (2003). An evolutionary scenario for the transition to undifferentiated multicellularity. *Proceedings of the National Academy of Sciences*, 100(3):1095–1098.

Queller, D. C. (2000). Relatedness and the fraternal major transitions. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 355(1403):1647–1655.

R Core Team (2020). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Rose, C. J., Hammerschmidt, K., Pichugin, Y., and Rainey, P. B. (2020). Meta-population structure and the evolutionary transition to multicellularity. *Ecology Letters*, 23(9):1380–1390.

Smith, J. M. and Szathmary, E. (1997). *The Major Transitions in Evolution*. Oxford University Press, New York, NY.

Thatcher, J. W., Shaw, J. M., and Dickinson, W. J. (1998). Marginal fitness contributions of nonessential genes in yeast. *Proceedings of the National Academy of Sciences*, 95(1):253–257.

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York.