

Evolution, Symbiosis, and Autopoiesis in the Game of Life

Peter D. Turney*

Submitted to: Artificial Life. **Category:** Article. **Date:** September 16, 2020.

Abstract

Recently we introduced a model of symbiosis, *model S*, based on the evolution of seed patterns in Conway's Game of Life. In the model, the fitness of a seed pattern is measured by one-on-one competitions in the *immigration game*, a two-player variation of the Game of Life. This article examines the role of autopoiesis in determining fitness in model S. We connect our research on evolution, symbiosis, and autopoiesis with research on *soups* in the Game of Life community. A soup is a random initial pattern in a cellular automaton, such as the Game of Life. When a game begins, there is usually a flare of rapid change in the soup, resembling a fire spreading through a forest. Eventually the fire burns down and the remaining patterns are called *ash*. Ashes are stable, oscillating, and flying patterns, studied intensively by the Game of Life community for insights into the game. Ashes are instances of autopoietic structures. We use the *apgsearch* software (Ash Pattern Generator Search), developed for the study of ash, to analyze autopoiesis in model S. We find that the fitness of evolved seed patterns in model S is highly correlated with the diversity and quantity of autopoietic structures (ash).

Keywords: Symbiosis, autopoiesis, Game of Life, cellular automata, diversity, productivity.

* Ronin Institute, 127 Haddon Place, Montclair, NJ 07043-2314, USA,
peter.turney@ronininstitute.com, 819-661-4625.

1 Introduction

In previous work [27], we introduced *model S*, a computational model of the evolution of symbiosis. In model S, seed patterns compete for survival in the *immigration game*, a two-player game invented by Don Woods [13, 31]. The winner of the immigration game is the seed pattern that grows the most. Woods created the immigration game by adding a scoring mechanism to John Conway’s Game of Life (GoL) [5], converting GoL into a three-state cellular automaton. Model S adds an external evolutionary algorithm to the immigration game. The algorithm evolves a population of seed patterns, selecting seeds according to their ability to win immigration games. The main result of our past work is evidence for the hypothesis that symbiosis promotes fitness improvements in the immigration game [27].

In the course of our research with model S, it became apparent that *autopoiesis* plays an important role in determining the fitness of seed patterns. The term *autopoiesis* was introduced by Humberto Maturana and Francisco Varela [20, p. 78]:

“An autopoietic machine is a machine organized (defined as a unity) as a network of processes of production (transformation and destruction) of components which: (i) through their interactions and transformations continuously regenerate and realize the network of processes (relations) that produced them; and (ii) constitute it (the machine) as a concrete unity in space in which they (the components) exist by specifying the topological domain of its realization as such a network.”

More succinctly, Randall Beer [2, p. 185] states that “... an autopoietic system consists of a network of processes that produce components whose interactions serve to generate and maintain the very network of processes that produced them.” The purpose of the current article is to investigate autopoiesis in model S.

We began our experiments with model S with some knowledge of autopoiesis, but no expectation that the immigration game would have a connection with autopoiesis. Over the course of many experiments, we became aware that model S was evolving seed patterns that produced many autopoietic structures. In retrospect, this should not have been surprising. Given that fitness in model S is measured by relative growth, we should expect that processes that self-generate and self-maintain will be more fit than other processes, other things being equal.

Beer has shown, in a series of articles [1, 2, 3], that GoL is an excellent platform for modeling autopoiesis. The classes of structures in GoL that are most interesting to Beer, regarding autopoiesis, are *still lifes*, *oscillators*, and *spaceships* [3, p. 3]:

“[...] GoL physics exhibits a very strong decay to quiescence. As we will see, persistent GoL patterns that resist this decay do so only by virtue of their self-sustaining (autopoietic) organization. And

persistent patterns certainly abound in GoL, as is visually obvious to any observer [...] Consider [...] a *block*, a *blinker*, and a *glider*. We will focus our analysis on these three patterns because (1) they represent three distinct classes within GoL (called *still lifes*, *oscillators*, and *spaceships*, respectively); (2) they are each the simplest members of their respective classes; (3) they are each the most common members of their respective classes. Because such compact and spatiotemporally coherent patterns are resistant to the general decay to quiescence, we tentatively interpret them as the GoL analogues of biological organisms.”

These three classes of patterns (*still lifes*, *oscillators*, and *spaceships*) were named in Gardner’s first article on the Game of Life in 1970 [5], before the term *autopoiesis* was introduced in 1973 [20]. The GoL community recognized from the beginning that these three classes are of special significance. They are structures that naturally emerge in the game and capture the attention of GoL players. To a seasoned GoL player, these structures indeed seem to be “analogues of biological organisms”.

A popular activity in the GoL community is to create a random *soup* and run it to quiescence [10]. A *soup* is a random initial seed pattern. When the soup reaches quiescence, what remains is called *ash*. Ashes are defined as still lifes, oscillators, and spaceships [11]. A list of common ashes, with animated illustrations, is available on the GoL website [7].

All ashes (still lifes, oscillators, and spaceships) are instances of autopoiesis, but there are patterns in GoL that could be considered as instances of autopoiesis, although they are not ashes. For example, *puffers* can be considered as autopoietic structures [12]. Puffers are not counted as ash, because they grow without limit. A puffer looks like a train blowing smoke out its stack. The train part of the puffer is stable, but the smoke part grows indefinitely. These kinds of partly stable structures rarely occur in soups and they do not occur in our experiments with model S seed patterns, so we can ignore them in this article.

There have been several automated censuses of ashes, counting the number of times each ash type appears in soups [10]. Adam Goucher has been collecting census information since 2015, using his *apgsearch* software (Ash Pattern Generator Search) [8]. A recent report notes [9], “The results are obtained by evolving random soups of size 16×16 with density 0.5 in an infinite planar universe; as of April 27, 2019, at least 18,928,504,510,982 soups have been investigated by the census’s participants, yielding a total of at least 413,493,174,300,923 objects of 159,347 distinct types.”

In the following experiments, we use a slightly modified version of Goucher’s *apgsearch* to identify and count the ash that is generated by evolved seed patterns from model S [28]. The modification allows us to gather a census of ash for a single, specific seed pattern, instead of accumulating counts from a large number of random soups, as the original version of *apgsearch* does.

Model S [28] uses the Golly cellular automata software [26] to run immigration games [13, 31]. In model S, Golly is controlled by Python scripts that implement evolution outside of Golly. A seed pattern is represented in Python by a binary matrix. The initial population consists of 200 random seed patterns. The algorithm for evolution is a steady-state model with a constant population size [32, 33], in which each new child replaces the least fit member of the population. Members of the population are chosen for reproduction using tournament selection [33].

Model S is constructed with four layers, each subsequent layer building on the previous layers. The purpose of having four layers is to measure each additional layer's contribution to the fitness of the population. This provides some insight into the strengths and weaknesses of each layer.

Layer 1 implements a simple form of asexual reproduction, with a fixed genome size (that is, a fixed binary matrix size). A member of the population is selected for reproduction using tournament selection. The chosen seed pattern is mutated by randomly flipping some of the bits in the binary matrix and it then competes in a series of one-on-one immigration games with the other member of the population. Its fitness is the average number of games it wins. Layer 2 implements a slightly more sophisticated asexual reproduction. In addition to mutation by flipping bits (in layer 1), the binary matrix is allowed to grow or shrink by adding or subtracting a row or column to or from the matrix. Layer 3 selects two seeds from the population and then combines them with genetic crossover. The resulting child is then passed on to layer 2 for mutation. Layer 4 adds symbiosis to model S. Two seeds are selected from the population and fused together, side-by-side, creating a new symbiotic genome. This new fused seed is treated as a whole; that is, selection shifts from the level of the two parts to the level of the whole.

Our previous work with model S [27] demonstrated that layer 4, the symbiotic layer, when combined with the other three layers, resulted in a significant increase in the fitness of the population. In that prior work, we did not analyze the specific means by which symbiosis achieved greater fitness, except to provide evidence that the increased fitness was not merely due to increased size. We conjectured that autopoiesis was somehow responsible [27], but we did not know how to test this hypothesis at that time. The solution became clear when we discovered Goucher's apgsearch software [8].

The left column of Figure 1 shows four initial seed patterns, randomly sampled from the final generations of four different runs of model S. The right column shows the final ash that results from each seed. The first row presents a run of model S with layer 1 alone; the other layers are turned off. Each following row shows the effect of adding another layer to the model.

Insert Figure 1 here.

Table 1 summarizes the types and quantities of ash for the four different runs of model S in Figure 1. The statistics were generated by our modified version of Goucher’s apgsearch software [28]. We could easily have generated Table 1 by hand, using the list of common objects on Catagolue [7] as a guide, but manual labeling does not scale up for a thorough statistical analysis of the data from each generation of many runs of Model S.

Insert Table 1 here.

Figure 1 and Table 1 suggest that symbiosis (layer 4) in model S results in evolved seed patterns that create diverse types and large quantities of ash, compared with the other layers. The remainder of the paper will argue that this is generally true; the trend in Table 1 is not merely a chance occurrence. Statistical analysis of the data supports the hypothesis that symbiosis yields highly fit seed patterns by evolving seeds that yield diverse types and large quantities of ash.

The main contributions of this article are (1) a model that brings together evolution, symbiosis, autopoiesis, and diversity, (2) linking the work on autopoiesis in the artificial life and theoretical biology communities with the work of Game of Life enthusiasts on exploring soups and ashes, (3) evidence that the evolution of symbiosis not only increases the fitness of a population, it also increases the diversity and quantity of autopoietic structures in the population; indeed, it seems that natural selection leads to the increase in autopoietic diversity and quantity, which *causes* increasing fitness, and (4) we have made the software available for replicating and extending the work presented here [28].

We review related work in Section 2. In Section 3, we briefly summarize our past work with model S. Section 4 discusses the relation between research on autopoiesis and research on ashes and soups. The core of the paper is Section 5, which presents various results that show how fitness in model S is related to the diversity and productivity of autopoietic structures. Section 6 discusses future work and limitations of the analysis. The article concludes in Section 7.

2 Related Work

Conway’s Game of Life was introduced to the world by Gardner in 1970 [5]. His introductory article mentioned *still lifes*, *oscillators*, and *spaceships*, the three classes of objects that are now covered by the term *ash* [11]. The phrase *primordial soup* was used by Wainwright in 1971 to describe a random initial state in GoL [6]. The GoL community generates random soups and runs them to quiescence, in order to find interesting ashes.

Varela and Maturana introduced *autopoiesis* to the world in 1973 [20]. McMullin [22] reports that they had coined the word *autopoiesis* in 1971 and they were planning at that time to make a cellular automaton computer simulation of autopoiesis, inspired by the work of von Neumann [30] and Conway [5]. Their cellular automaton model was published in 1974 [22, 29]. It used a simulated chemistry in a two-dimensional grid space. A cell in the space could be empty or it could contain one of three types of particles. The rules of the artificial chemistry were designed to allow particles to connect and form a kind of cell membrane. The membrane was capable of self-maintenance and self-construction.

Beer has made a strong case for GoL as a model of autopoiesis [1, 2, 3]. At first, he focused on *gliders* as instances of autopoiesis [1, 2], but later he expanded his studies to *still lifes*, *oscillators*, and *spaceships* [3]. This expansion connects the research in autopoiesis to the investigations of the GoL community, exploring soups and ashes.

The term *symbiosis* is derived from the Ancient Greek word συμβίωσις (*sumbīōsis*), meaning *living together*. There are many different types of symbiosis [4]. Martin and Schwab write [19, p. 32], “Confusion has afflicted the definition of *symbiosis* for over 130 years.” They propose a set of terms based on whether the two species in a symbiotic relationship experience a beneficial effect (+), a harmful effect (-), or a neutral effect (0). This yields six types of symbiosis: *neutralism* (0/0), *antagonism* (-/-), *amensalism* (0/-), *agonism* (+/-), *commensalism* (+/0), and *mutualism* (+/+) [19, see their Figure 10]. They also distinguish *endosymbiosis* (one organism lives within the tissues of the other) and *ectosymbiosis* (one organism lives on the body surface of the other, including the digestive tract).

Maynard Smith and Szathmáry, studying the major transitions in evolution, found that several of the major transitions involved symbiosis with a shift in the level of selection [21]. The prototypical example is the transition from prokaryotes to eukaryotes [17, 18, 21]. This is an instance of endosymbiosis where the two organisms became so tightly bound together that they cannot live apart, so natural selection must select them as a whole; it is no longer possible to select only one of the two parts.

In our past work [27], we introduced a model of symbiosis in GoL, called *model S*. In addition to the usual genetic operators (such as *mutation*, *crossover*, and *selection*), we added a *fusion* operator, which joins two seed patterns together to make a new seed pattern. Fusion is a simple joining of two binary matrices side-by-side, where a binary matrix represents a seed pattern. Fusion can be viewed as a model of endosymbiosis with a shift in the level of selection from the parts to the whole. The main result of this work was the observation that symbiosis by fusion yields a substantial increase in the fitness of the population, even though fusion is rare: Layer 4 with symbiotic mutualism achieved high fitness with an average of only 5.4 instances of symbiosis over the course of a run with 20,000 births [27, see Table 4].

The current article brings together the work of the Game of Life community on soups and ashes [8, 9, 10], the research in theoretical biology on autopoiesis [2, 3, 29], the research of the evolutionary computation community on models of evolution [24, 32, 33], and the research in theoretical biology on symbiosis [4, 17, 18].

3 Evolution of Symbiosis in Model S

The Game of Life is a solitaire game in which the player creates an initial seed pattern in an infinite grid of square cells [5, 23]. The rules of GoL determine how the pattern changes over time. When GoL was first introduced, the rules were applied manually, but it was not long before computers were used to apply the rules. In model S, we view the initial seed pattern as analogous to a *genome* and the rules determine how the *phenome* develops from the genome (*ontogeny* — the development of an organism over its lifespan).

Each cell in the GoL grid is either *dead* (state 0) or *alive* (state 1). The state of a cell depends on the state of its eight nearest neighbours in the grid. The rules for GoL are concisely expressed as B3/S23: A cell is *born* (it switches from state 0 to state 1) if it has exactly three live neighbours (B3). A cell *survives* (it stays in state 1) if it has either two or three live neighbours (S23). Otherwise the cell remains dead or it dies (state 0).

Time passes in a series of discrete intervals. The states of the cells in the grid at time $t = 0$ constitute the initial seed pattern for the game. The states at time t uniquely determine the new states at time $t + 1$. With each time step, all of the cells in the grid are updated.

The immigration game is a two-player game in which the two players each create an initial seed pattern [13, 31]. The rules for life and death are B3/S23, as in GoL, except that there are two *alive* states in the immigration game (states 1 and 2). Conventionally, the background (state 0) is usually white and the living states (states 1 and 2) are usually red and blue. There are two new rules for determining colour: (1) Live cells do not change colour unless they die (switch to state 0). (2) When a new cell is born, it takes the colour of the majority of its live neighbours. Since birth requires three live neighbours, there is always a clear majority.

In model S, fitness is determined by one-on-one competitions between two seed patterns in the immigration game. The seeds are represented as binary matrices when they are not competing. When it is time to compete, one seed is assigned red (state 1 remains state 1) and the other seed is assigned blue (state 1 switches to state 2). The two competitors are randomly situated in a toroidal grid with some space separating them. The game runs until a time limit is reached. The winner is the seed that grows the most. The growth of the red seed is the number of cells in state 1 at the end of the game minus the number of cells in state 1

at the beginning of the game. Likewise, the growth of the blue seed is the increase in blue states. If a seed shrinks instead of growing, its score is zero. Ties are allowed.

In these competitions, the size of the toroidal grid and the time limit for the game are both functions of the sizes of the seeds. Larger seeds are allowed larger grids and longer time limits. The reasoning is that any fixed limit on space and time would set a bound on the possible evolution of the seeds over many generations, which would be less interesting than unbounded growth. The motivation for a toroidal grid is to limit the space for growth, so that the two seeds are forced to interact. Without a limit on space, the two seeds might simply avoid each other, which would be less interesting than interaction. The toroid limits the space for any given competition, but the toroid increases in size from one competition to another, if the seed patterns increase in size. Likewise, the time limit is fixed for any given competition, but can increase from one competition to another.

In the immigration game, if states 1 and 2 were coloured black, the game would appear to be exactly the same as the Game of Life. The purpose of the two colours is to score the two players, to convert GoL from a solitaire game into a two-player competitive game. In the immigration game, we can choose to be red-blue colour blind, and view the game as GoL. In the work we describe in this article, model S measures the fitness of seeds by competitions in the three-state immigration game (two seeds at a time), but apgsearch analyzes the ashes in the two-state GoL (one seed at a time).

The fitness of a seed in the immigration game is the fraction of games that it wins, playing against every seed in the population. The initial population consists of 200 randomly generated seed patterns. Tournament selection randomly chooses two seeds from the population and the fitter of the two is selected for reproduction. The type of reproduction depends on which layers are enabled. The new child replaces the least fit member of the population, so the population remains at 200 for the entire run of model S. When 200 children have been born, we say that one generation has passed. We run model S for 100 generations, resulting in 20,000 births (100 generations with 200 births = 20,000 births; the generation count goes from 0 to 100, but generation 0 is randomly generated, not born from parents).

The fitness in model S is an internal, relative fitness. The fitness of a seed is relative to the fitness of the other seeds in the current population. With each new birth, the fitness of every seed in the population is updated. It follows that the average fitness (the average fraction of games won) is always 0.5. This means that we cannot compare the fitness of two seeds that are sampled from different populations. We argued in our prior work [27] that an internal, relative fitness measure is more likely to result in open-ended evolution [25] than an external, absolute fitness measure, although we do not yet have a proof for this hypothesis.

In order to evaluate the results from running model S, we need to create an external, absolute fitness measure that will allow us to compare seeds from different populations. This external fitness measure has no impact on the course of evolution in model S; it is only applied after a run of model S, in order to allow us to interpret the results of the run. The external, absolute measure we use is fraction of immigration games that a seed wins when competing against seeds of the same size (the same number of rows and columns) and the same density (the same number of living cells in the initial seed pattern) [27]. This is analogous to the sport of boxing, where fighters are matched by weight. A given seed's competitors are generated by copying the given seed and then randomly shuffling the cells in the copied seed. This ensures that the winner is winning due to the structure of the seed (the specific pattern of living and dead cells), not the size or density of the seed.

Figure 2 shows the external, absolute fitness curves for the four layers of model S. Each of the four curves is the average of 12 separate runs of model S. For each generation (each 200 births), we sample the top 50 most fit seeds (based on their internal, relative fitness) and calculate their external, absolute fitness. We call these seeds the *elite* seeds of the given generation. Figure 2 is copied from our past work with model S [27]. We saved all of the elite seeds from our past work, which allows us to use these seeds again in the current article. Figure 2 presents the external, absolute fitness of the same seeds as we will be using later in Section 5.

Insert Figure 2 here.

All four layers in Figure 2 are statistically significantly different from each other in their external, absolute fitness, except for layers 2 and 3, where the difference is not significant. In particular, layer 4, the symbiotic layer, is significantly more fit than the other three layers. See our prior work for the details of the statistical analysis [27].

Figure 3 shows the areas (rows \times columns) of the elite seeds for the four layers of model S. Each of the four curves is the average of 12 separate runs of model S. Comparing Figure 2 and Figure 3, we can see that area is highly *correlated* with fitness, but increasing area does not *cause* increasing fitness. Increasing area is *necessary* for increasing fitness, because increasingly complex structures require larger seeds to store increasing amounts of information. But increasing area is *not sufficient* for increasing fitness, because a specific kind of structure is required for fitness; that is, a structure that grows quickly and robustly during the immigration game.

Insert Figure 3 here.

In our previous paper [27], introducing model S, we presented three different arguments in support of the claim that increasing area does not cause increasing fitness. We will only mention one of those arguments here: Note that the external, absolute fitness is defined by competitions between seeds that are matched in size and density. Therefore area cannot explain why the average symbiotic seed in the final generation of layer 4 has a fitness of 93.6%. If area were the deciding factor, the fitness would have to be 50%, because the two competitors have the same area. The only difference between the symbiotic seed and its competitor is that the competitor has been shuffled, which damages its structure but preserves its area. The structure is what determines the winner, not area.

4 Soups, Ashes, and Autopoiesis

In the context of the Game of Life, the term *primordial soup* was used by Robert Wainwright in 1971, according to Martin Gardner [6, p. 116]:

“Wainwright has made a number of intriguing investigations. He filled a 120-by-120 square field with 4,800 randomly placed bits (a density of one-third) and tracked their history for 450 generations, by which time the density of this primordial soup, as Wainwright calls it, had thinned steadily to one-sixth. Whether it would eventually vanish or, as Wainwright says, percolate at a constant minimum density is anybody's guess. At any rate, during the 450 generations 42 short-lived gliders were formed.”

More recently, Andrzej Okrasinski made a census of the ashes arising from random soups from 2003 to 2008 [14], Achim Flammenkamp made a census in 2004 [15], Nathaniel Johnston made a census from 2009 to 2011 [16], and Adam Goucher has been making a census from 2015 to the present [9]. Clearly the study of ashes from random soups is a popular activity in the GoL community. Currently Goucher's apgsearch is the most popular tool for running random soups and classifying their ashes [8].

Apgsearch is currently at version 5.0, which is written in C++. Apgsearch was designed for distributed computation, with many different users on the Internet running random soups and uploading their ash censuses to a central server. Over 100 users have contributed to the census of ash objects [9].

We used apgsearch version 1.1, which is written in Python and uses Golly, since our model S is also written in Python and uses Golly. We modified this version of apgsearch to analyze the ash from a single seed pattern, rather than aggregating ash statistics from very large samples of soups. The Python version of apgsearch is slower than the C++ version, but speed was not an issue for our analysis of model S seeds.

Apgsearch has discovered 159,347 distinct types of ash objects so far [9]. Beer has studied three of these objects (a block, a blinker, and a glider) in depth [3], to determine whether they satisfy Maturana and Varela’s two conditions for autopoiesis, the *closure condition* and the *boundary condition* [3, p. 2]:

“The *closure condition* demands that the network of processes must produce the components whose interactions generate and maintain that very same network. The *boundary condition* demands that the spatial boundary that distinguishes an autopoietic system from its background must itself be generated and maintained by the network of processes and in turn must play a central role in enabling those same processes.”

Beer finds that all three objects satisfy the conditions for autopoiesis. Since blocks, blinkers, and gliders are the simplest members of their respective classes (still lifes, oscillators, and spaceships), it seems likely that all three classes satisfy the conditions for autopoiesis. That is, it is likely that all 159,347 distinct types of ash objects from Goucher’s census satisfy the conditions for autopoiesis.

5 Fitness and Autopoiesis

In this section, we present three new analyses of the data from our previous work with model S [27]. First, we apply the modified apgsearch to the stored elite seeds from our past work, in order to count the number of structures that are produced (the *quantity* of ashes). Second, we count the number of types of autopoietic structures the seeds produce (the *diversity* of ashes). Third, we *shuffle* the seeds before we run apgsearch, to see whether the quantity and diversity depend on the size of the seeds or the structure of the seeds (shuffling changes structure but has no effect on size). The results show that the quantity and diversity of ash for a given layer of model S is highly correlated with the fitness of that layer. Shuffling greatly reduces the quantity and diversity of ashes, indicating that the quantity and diversity of ash is strongly affected by the structure of the seeds.

5.1 Productivity of the Layers of Model S

The *productivity* of a seed pattern is the quantity of ash it generates. For example, the initial seed pattern for layer 4 in Figure 1 yields 106 objects, as we can see in Table 1. For each of the four layers of model S, Figure 4 shows the average quantity of ash produced by seed patterns in each generation. Each point on the curve for a layer in Figure 4 is the average productivity of 600 seed patterns (50 elite seeds \times 12 runs of model S).

Insert Figure 4 here.

If we compare the productivity curves of the four layers in Figure 4 with the fitness curves in Figure 2, we see the same general pattern in both figures: The productivity and fitness of layer 1 (the uniform asexual layer) flattens out early. Layer 2 (the variable asexual layer) and layer 3 (the sexual layer) have similar levels of productivity and fitness, with a slight advantage to layer 2. Both layers 2 and 3 are more productive and fit than layer 1. Layer 4 (the symbiotic layer) quickly moves to the top. Figure 4 shows that the symbiotic layer is much more productive than the other layers and Figure 2 shows it is much more fit.

Table 2 focuses on the final generation of the runs (generation 100), comparing the final fitness levels of the four layers (Figure 2) with the productivity levels (Figure 4). The correlation between fitness and productivity in the final generation is high (0.706) and statistically significant.

Insert Table 2 here.

The average seed pattern in the final generation of layer 4 (the symbiotic layer) generates 73.1 ashes (see Table 2 or Figure 4) and has an area of 94.1 (see Figure 3). Therefore the average ash generated by the final generation of symbiotic seeds is 0.78 ashes per area (73.1 ashes / 94.1 area).

Catagolue reported a census on April 27, 2019 of 18,928,504,510,982 random soups, resulting in a total of 413,493,174,300,923 ashes [9]. This yields an average of 21.8 ashes per random soup. Each Catagolue soup begins with a random 16×16 binary matrix, which has an area of 256 cells. Therefore the average ash generated by random soups is 0.09 ashes per area (21.8 ashes / 256 area).

The average final generation seed pattern of layer 4 has an area that is 37% (94.1 area / 256 area) of the area of Catagolue's random soups and a productivity that is 335% (73.1 ashes / 21.8 ashes) of the productivity of Catagolue's soups. Therefore the evolved seeds of layer 4 are 9.1 times more productive than random soups (335% / 37%). This shows that the evolved seeds in the final generation of layer 4 are much more productive than random seeds. However, model S was not designed to produce ash. It was designed to play the immigration game. With hindsight, it makes sense that winning the immigration game requires evolution to generate seeds that maximize the production of autopoietic structures.

5.2 Diversity of the Layers of Model S

The *diversity* of a seed pattern is the number of types of ashes it generates. For example, the initial seed pattern for layer 4 in Figure 1 yields 10 types of ashes, as we can see in Table 1. Figure 5 shows the average diversity of ashes produced by seed patterns in each generation of model S, for each of the four layers. Each point on the curve for a layer in Figure 5 is an average of 600 seed patterns (50 elite seeds × 12 runs of model S).

Insert Figure 5 here.

If we compare the diversity curves of the four layers in Figure 5 with the fitness curves in Figure 2, we see the same general pattern in both figures: Layer 1 (the uniform asexual layer) flattens out early. Layer 3 (the sexual layer) rises above layer 1, but eventually flattens out. Layer 2 (the variable asexual layer) rises above layer 3. Layer 4 (the symbiotic layer) quickly moves to the top, above the other three layers.

This qualitative similarity between fitness (Figure 2) and diversity (Figure 5) is confirmed by measuring the correlations. Table 3 compares the fitness in the final generation with the diversity in the final generation. The correlation between fitness and diversity is 0.834. The correlation is high and it is statistically significant.

Insert Table 3 here.

We know from Section 5.1 that the evolved seed patterns in layer 4 are much more productive than random soups (there is a large *quantitative* difference between the number of seeds generated in layer 4 and the number of seeds generated in random soups), but is there a difference in the *types* of seeds in layer 4 and the *types* of seeds in random soups (is there a *qualitative* difference in the distribution of seed types)? Table 4 suggests that there is no qualitative difference in the types.

Insert Table 4 here.

Table 4 shows all of the types of ash found in each layer of model S in the final generation. Focusing on layer 1 as an example (the first column in Table 4), there are seven types of ash, ranked in descending order of their frequencies in their layer in model S; that is, the *block* is the most frequent ash found in layer 1 and the *boat* is the least frequent ash found in layer 1. For each ash type in the first column, there is a corresponding number in the second column that gives the rank of the ash type in Catagolue [7] (from a census in July 15, 2017). For example, the *block* is ranked 1 in Catagolue and the *block* is also first in layer 1. However, the *pond* is third in layer 1, yet it is ranked 9 in Catagolue.

For most entries in Table 4, the rank of an ash type that occurs in model S is approximately the same as its rank in Catagolue. For example, the *boat* is the seventh ash type in layer 1 (see column 1) and it is the sixth ash type in Catagolue (see column 2). The *tub* is the eighth ash type in layer 4 (see column 7) and it is the eighth type in Catagolue (see column 8).

Although the ranks of ash types in model S are mostly similar to their ranks in Catagolue, there are a few outliers. We have highlighted in italics one outlier in each of the four layers: the *pond* in layers 1 and 3 (ranked third in model S but ranked ninth in Catagolue), the *pulsar* in layer 2 (ranked tenth in model S but ranked twenty-first in Catagolue), and the *loop* in layer 4 (ranked twelfth in model S but ranked forty-ninth in Catagolue).

We believe these apparent outliers in Table 4 are simply random noise. That is, it seems there is nothing special about the distribution of types that we see in Table 4; the relative frequencies of the types in model S are approximately what we would expect from Catagolue [7]. It appears that, when it comes to winning the immigration game, it's not the *type* of ash that matters; it's the *quantity* of ash — any ash is good ash. However, we discuss this topic further in Section 6.

5.3 Structure versus Area

In our past work with model S [27], we argued that, although the fitness of a seed pattern (its tendency to win immigration games) is *correlated* with the area of the seed pattern (compare Figure 2 and Figure 3), area does not *cause* fitness, area merely provides room for complex structure (area *allows* fitness), and fitness is the result of a certain type of complex structure (the type that wins games). We provided three lines of evidence to support this claim [27]. As one example, consider how external, absolute fitness is measured in Section 3, as illustrated in Figure 2. The external, absolute fitness of a seed is the fraction of contests that it wins when it competes against a shuffled version of itself. The average symbiotic seed in the final generation of layer 4 has an external, absolute fitness of 93.6%, which means that it wins 93.6% of the games that it plays against a shuffled version of itself. The shuffled version has the same area (the same number of rows and columns and the same number of living cells) as the evolved seed, but it lacks the structure that would allow it to win games.

In this section, we argue that the specific kind of complex structure that enhances fitness is structure that produces large quantities of ash. As evidence for this claim, we show that shuffling seeds results in a large decrease in the quantity of ash that they produce. This suggests that the property that evolution selects in model S is productivity of ash: It seems that increasing ash productivity *causes* increasing fitness.

Figure 6 shows the result of shuffling on *productivity* (the quantity of ash produced by seeds). Compare Figure 6 (the productivity of shuffled seeds) with Figure 4 (the productivity of normal seeds). Clearly shuffling substantially reduces the productivity of the seeds. These are stored seeds from our previous experiments with model S [27], so the seeds in Figure 4 are identical to the seeds in Figure 6, except that the seeds in Figure 6 were shuffled before calculating their productivity. In both cases, the productivity for a given generation and layer is the average of 600 values (12 runs \times 50 elite seeds).

Insert Figure 6 here.

Table 5 compares the productivity of the seeds before shuffling (the *intact* column) and after shuffling (the *shuffled* column). The productivity of the shuffled seeds ranges from 7% to 14% of the productivity of the intact seeds. This is evidence that the specific kind of structure that yields fitness is also the specific kind of structure that yields productivity. By design, shuffling has no impact on area and density, so the loss in productivity that we see here has nothing to do with area or density; it can only be caused by the change in structure.

Insert Table 5 here.

Figure 7 shows the impact of shuffling on *diversity* (the number of types of ash produced by seeds). Compare Figure 7 (the diversity of shuffled seeds) with Figure 5 (the diversity of normal seeds). Shuffling greatly reduces the diversity of ash.

Insert Figure 7 here.

Table 6 compares the diversity of the seeds before shuffling (*intact*) and after shuffling (*shuffled*). The diversity of the shuffled seeds ranges from 18% to 33% of the diversity of intact seeds. Shuffling clearly reduces diversity, although the reduction is likely to be a consequence of the reduced productivity of the shuffled seeds. As the number of ashes decreases, the number of types of ashes also tends to decrease, so the reduced diversity is likely a side effect of reduced productivity.

Insert Table 6 here.

The results in this section support the hypothesis that ash productivity requires a specific kind of structure, which seems to be the same kind of structure that causes increasing fitness; that is, the experiments suggest that increasing ash productivity is the cause of increasing fitness in model S. Increasing ash diversity appears to be a side effect of increasing ash productivity, but this hypothesis requires further investigation.

6 Future Work and Limitations

In Section 5.2, we noted that *ponds*, *pulsars*, and *loops* might be random outliers or they could be the result of evolutionary selection in model S. It seems possible that they are selected because they are more stable than other objects, so they can survive close contact with other objects while playing the immigration game, whereas other objects are destroyed. We plan to test the stability of these objects, relative to other objects, and to look for other properties that might explain why *ponds*, *pulsars*, and *loops* seem to be favoured.

There are objects known as *eaters* that are able to survive close contact with other objects. The most common *eater* is object 17 in the list of common objects [7], yet it is not one of the 19 objects generated in layer 4 (see Table 4). This suggests that *ponds*, *pulsars*, and *loops* are merely random outliers, rather than particularly stable objects. On the other hand perhaps they have another property, other than stability, which is favoured by evolutionary selection. We leave these questions for future work.

The score in the immigration game is based on growth, which is measured by the increase in the number of live cells of a given colour (red or blue). In Section 5.1, we counted the number of ashes, which is not the same as counting the increase in live cells. Selection in model S might be biased towards producing either many small objects or a few large objects. Many small objects may be less vulnerable in the immigration game than a few large objects. We plan to look for a bias towards smaller objects or possibly towards larger objects. With random soups, there is a natural bias towards smaller objects, because they have higher probabilities of being randomly generated than larger objects. To see whether the immigration game prefers smaller objects, we need to take into account this natural bias for smaller objects.

One possibility would be to alter the immigration game so that the winner is determined by the number of ashes, instead of the number of living cells. If this alteration changes the distribution of ash types, then we will learn something about how the method of scoring affects the kinds of ash we see.

7 Conclusion

When we began our work with model S [27], our objective was to see whether symbiosis could contribute to fitness increase in the immigration game. We did not have any specific expectations about the kinds of structures that evolution might produce in model S. After many experiments with evolving seed patterns in model S, we gradually became aware that autopoiesis seems to play a role in determining fitness, but we did not follow up on this observation in our previous article [27]; it was only mentioned as a possibility for future work.

The purpose of the work presented here has been to closely examine the role of autopoiesis in the evolution of seed patterns in model S. We believe that the results presented here are strong evidence that

increasing instances of autopoiesis cause increasing fitness in the evolving population. In particular, Section 5.1 support the claim that adding symbiosis to the evolutionary process results in a large increase in external, absolute fitness (Figure 2), which is due to a large increase in the quantity of autopoietic structures (Figure 4).

Section 5.2 suggests that the types of ash generated by seed patterns in model S are much like the types of ash generated by random soups (Table 4). The increase in diversity seems to be a side effect of the increase in productivity. However, there are some outliers (*ponds*, *pulsars*, and *loops*) that suggest there may be some differences in the types of ashes that occur random soups, compared with evolved seed patterns. The differences may be due to specific qualities of some types of ash (such as stability), or they might be due to variation in the size of autopoietic structures (maybe it is best for a seed to produce many small autopoietic structures, rather than a few large autopoietic structures), or the differences might be merely random variations. We plan to explore this in future work.

Acknowledgments

Thanks to Andrew Trevorrow, Tom Rokicki, Tim Hutton, Dave Greene, Jason Summers, Maks Verver, Robert Munafo, Brenton Bostick, and Chris Rowett, for developing *Golly*. Thanks to Adam Goucher for developing *apgsearch*.

References

1. Beer, R. D. (2004). Autopoiesis and cognition in the game of life. *Artificial Life*, 10(3): 309-326.
2. Beer, R. D. (2014). The cognitive domain of a glider in the game of life. *Artificial life*, 20(2): 183-206.
3. Beer, R. D. (2015). Characterizing autopoiesis in the game of life. *Artificial Life*, 21(1): 1-19.
4. Douglas, A. (2010). *The Symbiotic Habit*. Princeton, New Jersey: Princeton University Press.
5. Gardner, M. (1970). Mathematical Games: The fantastic combinations of John Conway's new solitaire game 'life'. *Scientific American*, 223(4): 120-123.
6. Gardner, M. (1971). Mathematical Games: On cellular automata, self-reproduction, the Garden of Eden, and the game 'life'. *Scientific American*, 224(2): 112-117.
7. Goucher, A. P. (2020). *Most common objects on Catagolue*. https://www.conwaylife.com/wiki/Most_common_objects_on_Catagolue.
8. Goucher, A. P. (2020). *Apsearch*. <https://www.conwaylife.com/wiki/Apgsearch>.
9. Goucher, A. P. (2020). *Catagolue*. <https://www.conwaylife.com/wiki/Catagolue>.
10. LifeWiki. (2020). *Soup*. <https://www.conwaylife.com/wiki/Soup>.
11. LifeWiki (2020). *Ash*. <https://www.conwaylife.com/wiki/Soup#Ash>.

12. LifeWiki. (2020). *Puffer*. <https://www.conwaylife.com/wiki/Puffer>.
13. LifeWiki. (2020). *Colourised Life: Immigration*. https://www.conwaylife.com/wiki/Colourised_Life#Immigration.
14. LifeWiki (2020). *Andrzej Okrasinski's census*. https://www.conwaylife.com/wiki/Andrzej_Okrasinski%27s_census.
15. LifeWiki (2020). *Achim Flammenkamp census*. https://www.conwaylife.com/wiki/Achim_Flammenkamp%27s_census.
16. LifeWiki (2020). *The online Life-like CA soup search*. https://www.conwaylife.com/wiki/The_Online_Life-Like_CA_Soup_Search.
17. Margulis, L. (1970). *Origin of eukaryotic cells*. New Haven, CT: Yale University Press.
18. Margulis, L. (1981). *Symbiosis in cell evolution*. San Francisco, CA: W. H. Freeman.
19. Martin, B. D., & Schwab, E. (2013). Current usage of symbiosis and associated terminology. *International Journal of Biology*, 5(1): 32-45.
20. Maturana, H. R., & Varela, F. J. (1973). Autopoiesis: The organization of the living. In H. R. Maturana and F. J. Varela (Eds.), *Autopoiesis and cognition* (pp. 59–138). Boston: Reidel.
21. Maynard Smith, J., & Szathmáry, E. (1995). *The major transitions in evolution*. Oxford, UK: Oxford University Press.
22. McMullin, B. (2004). Thirty years of computational autopoiesis: A review. *Artificial Life*, 10(3): 277-295.
23. Poundstone, W. (2013). *The recursive universe: Cosmic complexity and the limits of scientific knowledge*. Mineola, NY: Dover Publications.
24. Simon, D. (2013). *Evolutionary optimization algorithms*. Hoboken, NJ: John Wiley & Sons.
25. Taylor, T., Bedau, M., Channon, A., Ackley, D., et al. (2016). Open-ended evolution: Perspectives from the OEE workshop in York. *Artificial Life*, 22(3): 408-423.
26. Trevorrow, A., Rokicki, T., Hutton, T., Greene, D., Summers, J., Verver, M., Munafò, R., Bostick, B., & Rowett, C. (2020). *Golly 3.4 for Windows*. <http://golly.sourceforge.net/>.
27. Turney, P. D. (2020). Symbiosis promotes fitness improvements in the Game of Life. *Artificial Life*, 26(3): 1-30.
28. Turney, P. D. (2020). *Source code for model S and measuring autopoiesis*. <https://github.com/pdturney/autopoiesis>.
29. Varela, F. J., Maturana, H. R., & Uribe, R. (1974). Autopoiesis: The organization of living systems, its characterization and a model. *BioSystems*, 5: 187–196.
30. von Neumann, J. (1966). *Theory of self-reproducing automata*. Urbana, IL: University of Illinois Press. Edited and completed by Arthur W. Burks.

31. Wainwright, R. T. (1971). *Lifeline Volume 2*. http://www.conwaylife.com/wiki/Lifeline_Volume_2.
32. Whitley, D., & Kauth, J. (1988). GENITOR: A different genetic algorithm. In *Proceedings of the Rocky Mountain Conference on Artificial Intelligence*. Denver, CO.
33. Whitley, D. (1989). The GENITOR algorithm and selective pressure: Why rank-based allocation of reproductive trials is best. In D. Schaffer (ed.), *Third International Conference on Genetic Algorithms* (pp. 116-121). San Francisco, CA: Morgan Kaufmann.

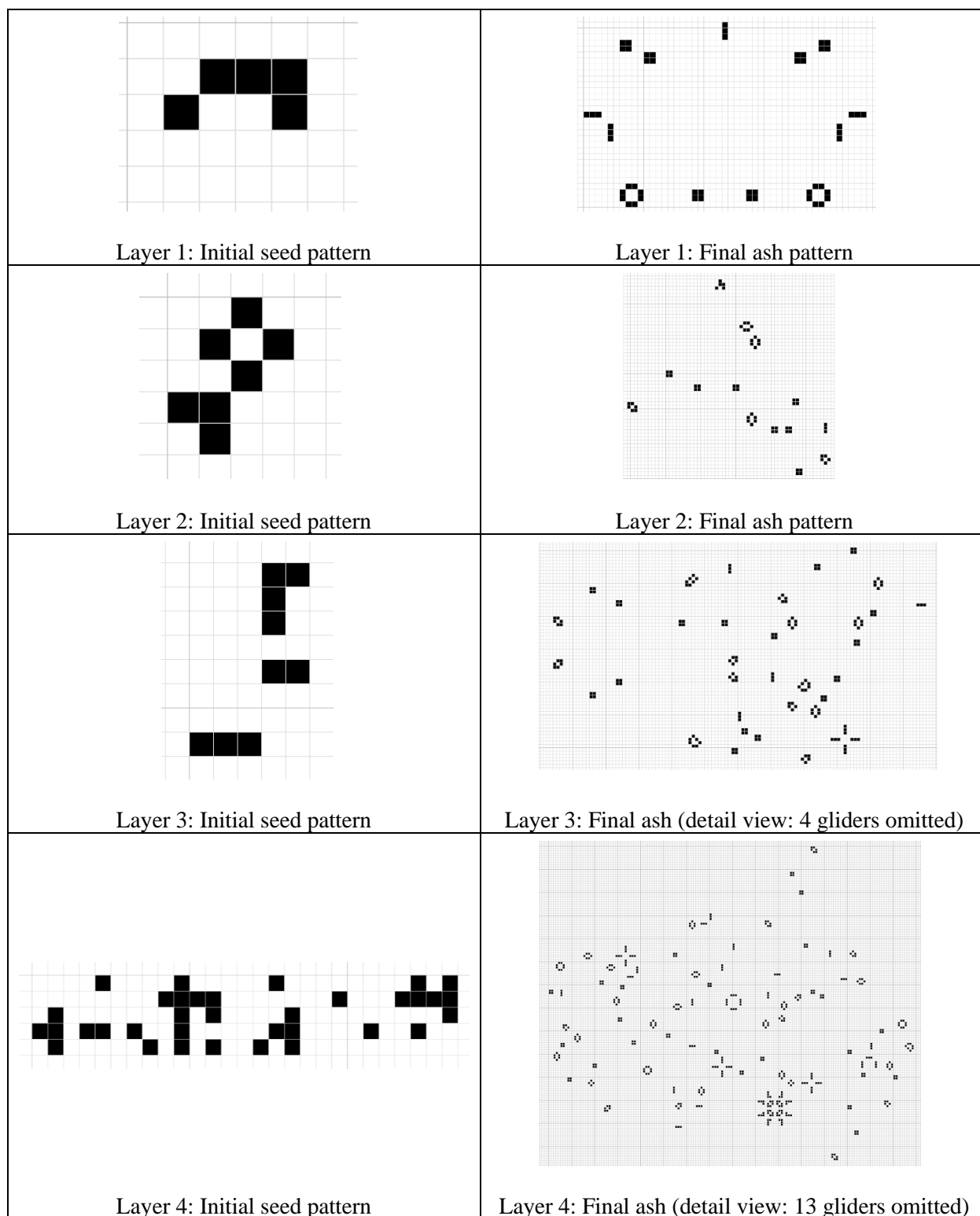


Figure 1. These images show examples of initial evolved seed patterns for each layer on the left and the corresponding final patterns on the right. The final patterns continue to change, but they are composed entirely of ashes. Gliders are omitted in layers 3 and 4 because they have moved far from the core ashes.

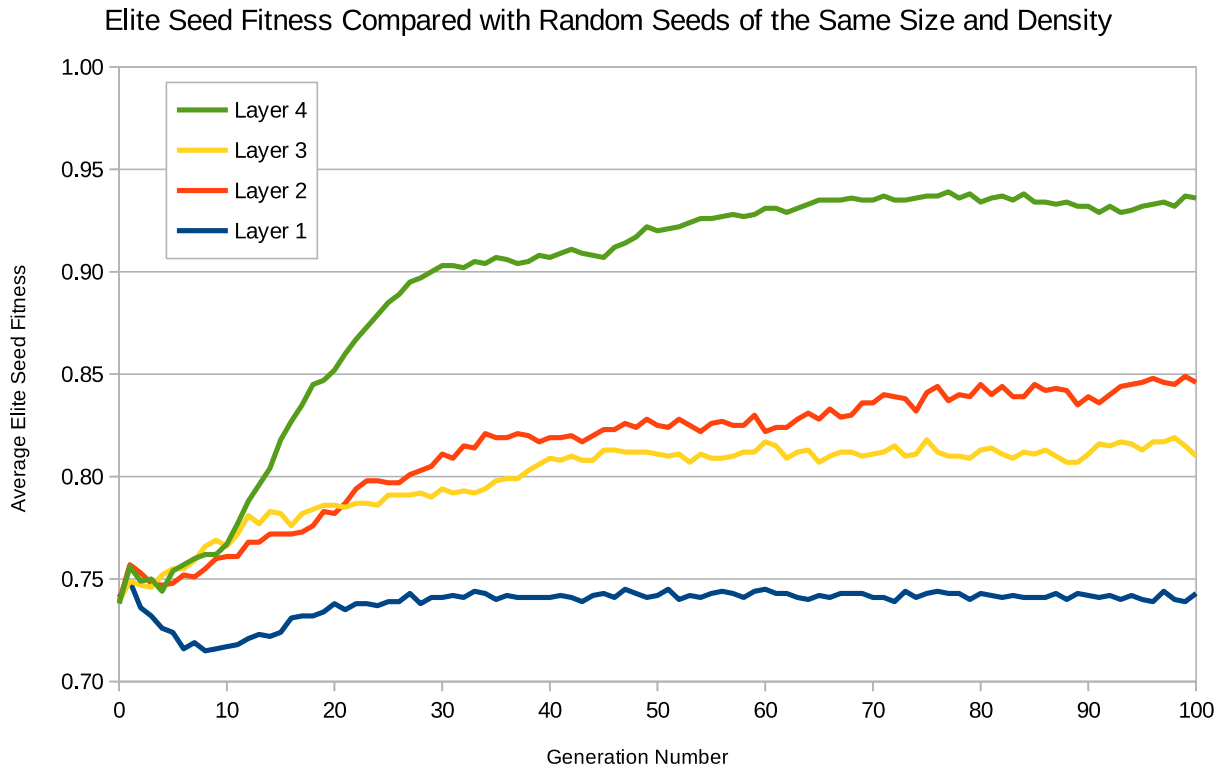


Figure 2. Each curve in this figure (that is, each layer) is the average of 12 separate runs of model S. The curve for a layer represents the average fitness of the top 50 most fit seeds in the population (the elite seeds). The fitness of an evolved seed is the fraction of immigration games that it wins when competing against randomly generated seeds with the same size (the same width and height) and the same density (the same number of live cells). For example, the figure shows that the average elite seed in layer 4 wins 93.6% of its competitions in the final generation (generation 100). Since the competition is matched by size and density, the difference in the fitness of evolved seeds and random seeds can only be due to difference in their structure (their specific pattern of ones and zeros). (This figure is copied from our earlier paper on model S [27].)

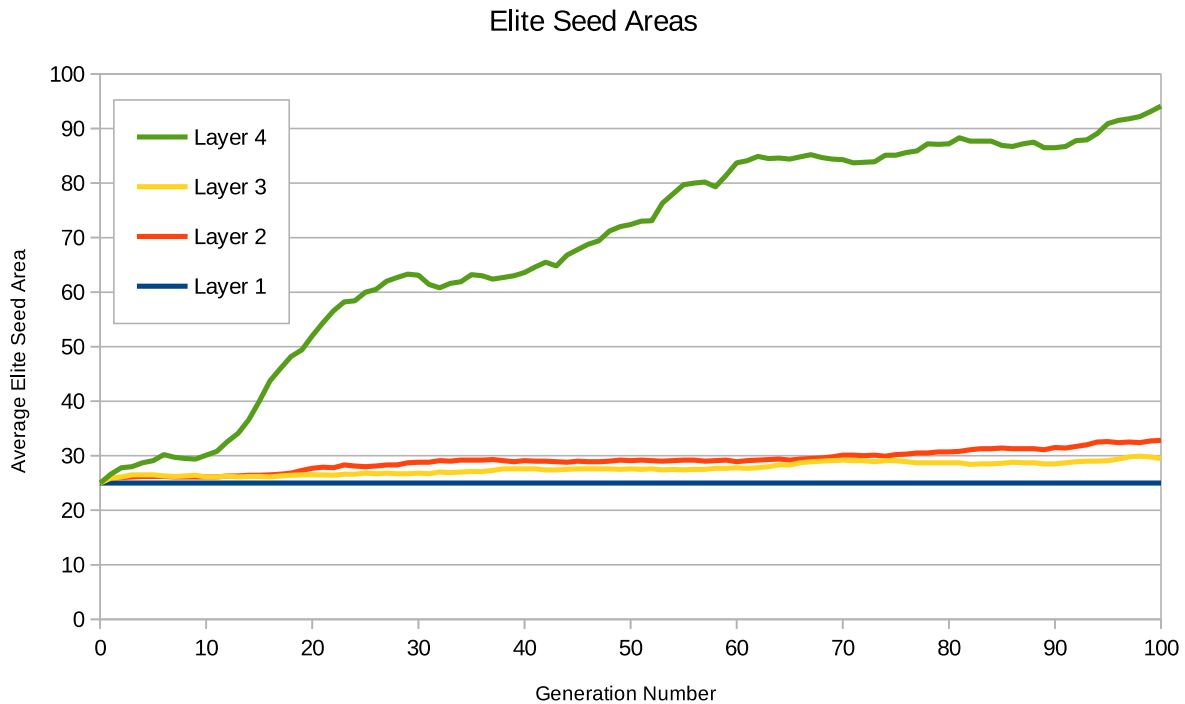


Figure 3. The curves in this figure represent the area of the elite seeds, averaged over 12 runs of model S. The area of a binary seed matrix is the number of columns times the number of rows. The similarity of this figure with Figure 2, which shows the average fitness of the seeds, might suggest that the increase in area causes the increase in fitness, but we argue that this is not the case. Increasing area is necessary for increasingly complex and effective structures, but increasing area cannot increase fitness without the specific structural patterns that cause diversity and productivity in ash. (This figure is copied from our earlier paper on model S [27].)

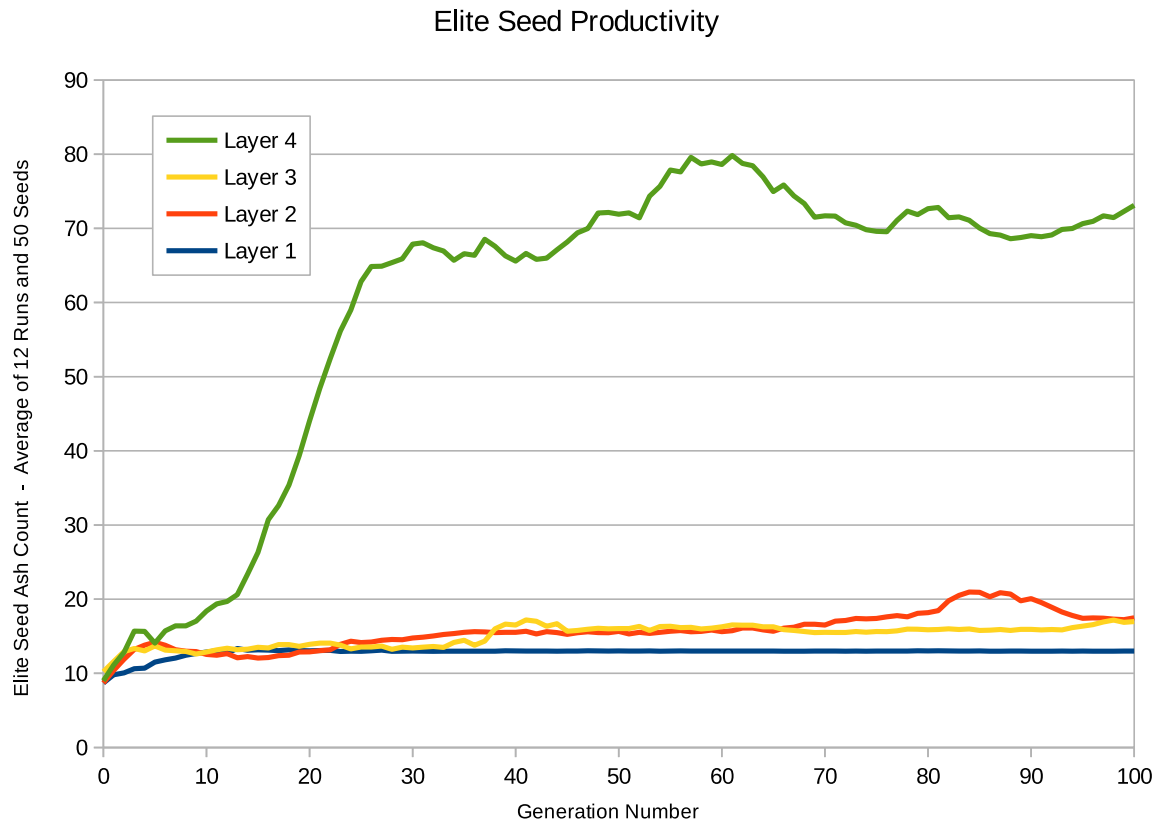


Figure 4. The curve for a layer represents the number of ashes it yields for each generation. In the final generation, the average elite seed in layer 4 generates 73.1 ashes with 7.9 different types. The curves for fitness in Figure 2 and the curves for diversity in Figure 3 have the same general shape as the curves for productivity in Figure 5. The curves suggest that increasing productivity of ash could be a cause of increasing fitness.

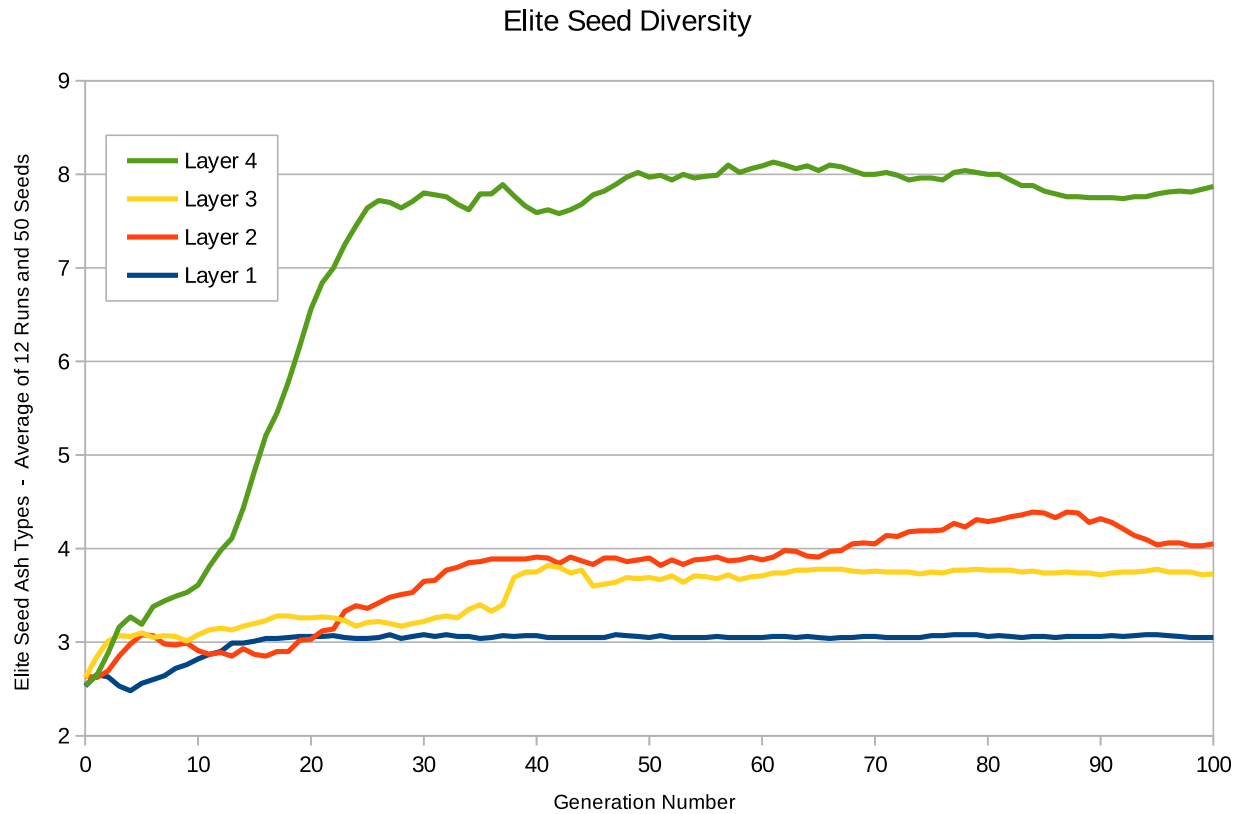


Figure 5. The curve for a layer represents the average number of types of ash generated by the elite seeds. For example, the average elite seed in layer 4 generates 7.9 different types of ash in the final generation (generation 100). The number of types of ash generated by a seed may be viewed as a measure of the diversity of strategies that the seed employs in order to win immigration games. The curves for fitness in Figure 2 have the same general shape as the curves for diversity in Figure 3. This suggests that increasing diversity in ash could be a cause of increasing fitness.

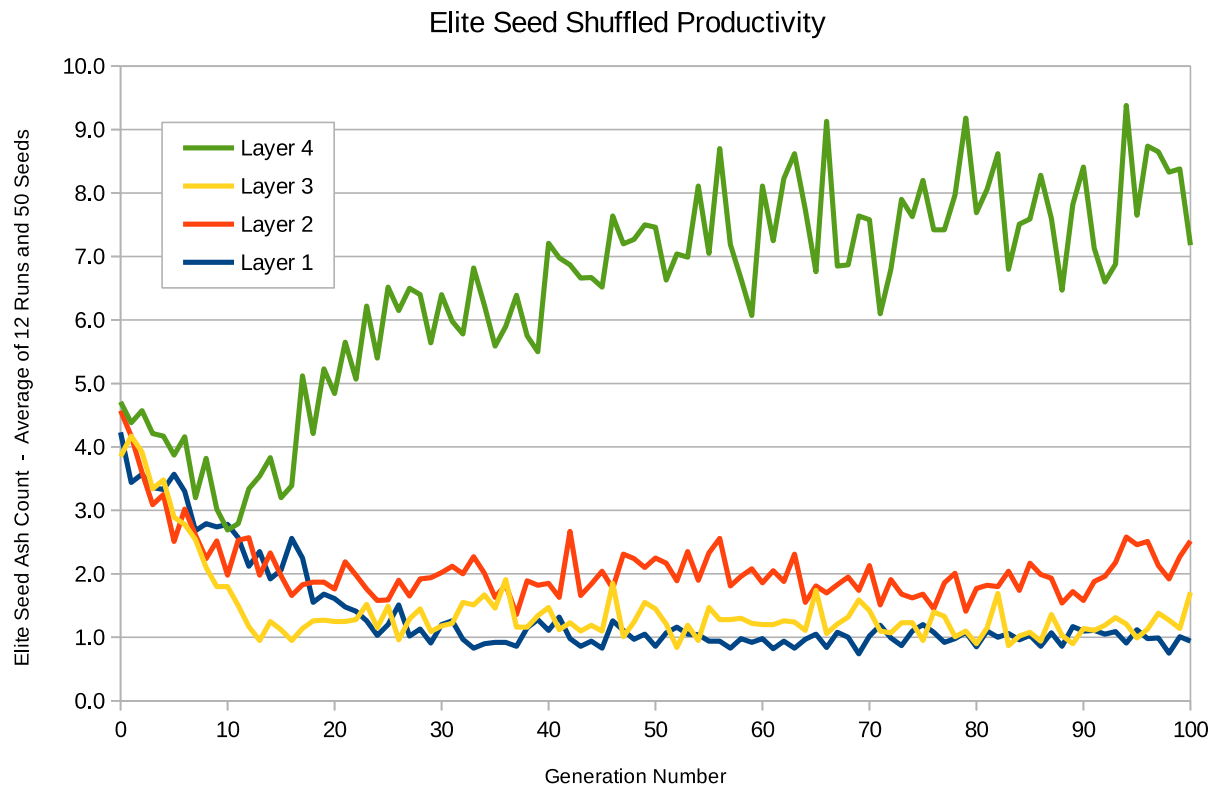


Figure 6. These curves show the impact of shuffling on productivity. Comparing this figure with Figure 5, we can see that shuffling causes a large drop in the productivity of the seeds (the number of ashes they generate). Since the only difference between Figure 5 and Figure 6 is the shuffling, the drop in productivity can only be due to the loss of structure.

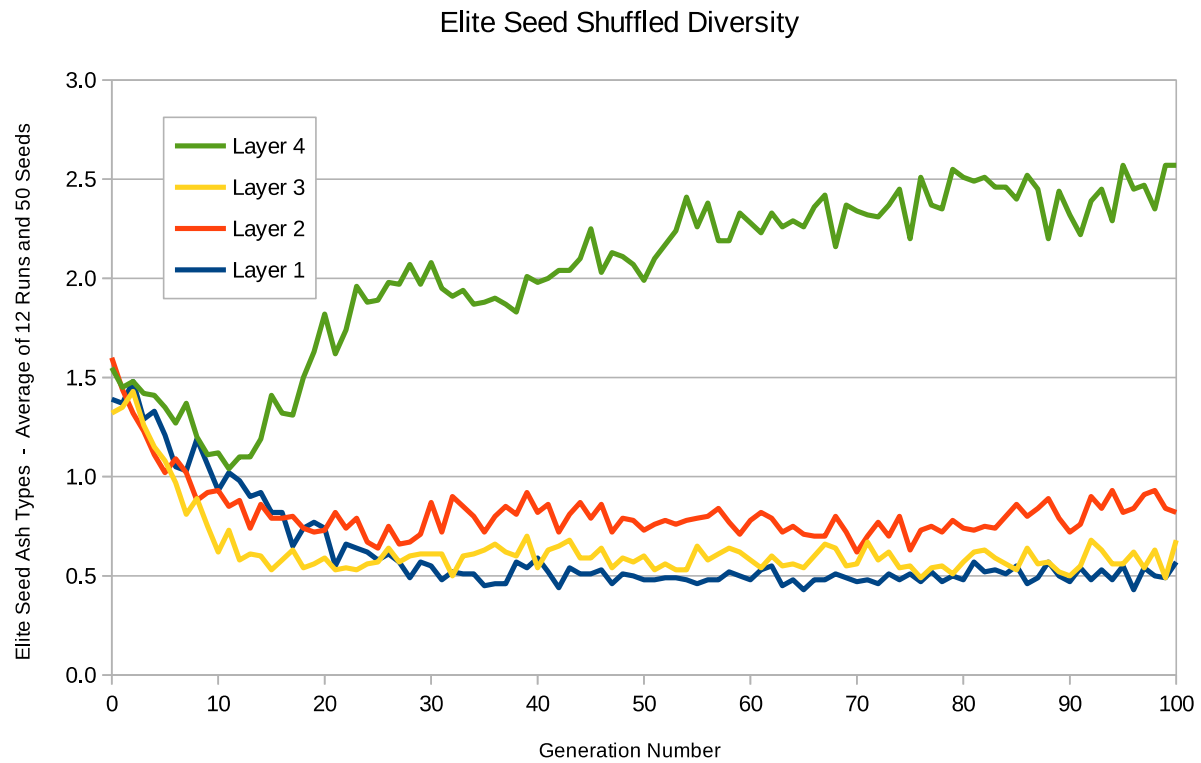


Figure 7. In this experiment, we take the same seeds as in the previous experiments, but we shuffle them before we calculate the diversity of the ash they produce (the number of distinct types of ash). The binary matrix of a seed pattern is shuffled by randomly selecting pairs of cells in the matrix and swapping their values. The result is a new matrix with exactly the same number of rows and columns and exactly the same density of ones and zeros, but a completely different pattern structure. Comparing this figure with Figure 3, the diversity of all four layers drops considerably when the seeds are shuffled. Since the only difference between Figure 3 and Figure 7 is the shuffling, the drop in diversity can only be due to the loss of structure.

Table 1. This table summarizes the statistics for the final ash patterns shown in Figure 1. For more information about the types of ashes listed in the table, see Adam Goucher's Catagolue [7].

Layer 1 ashes		Layer 2 ashes		Layer 3 ashes		Layer 4 ashes	
Type	Count	Type	Count	Type	Count	Type	Count
block	6	block	7	block	16	blinker	36
blinker	5	beehive	3	blinker	8	block	25
pond	2	blinker	1	boat	5	beehive	15
		glider	1	glider	4	glider	13
		ship	1	beehive	4	boat	6
		boat	1	ship	2	ship	4
				loaf	2	pond	3
				long boat	1	tub	2
						loaf	1
						pulsar	1
Num objects	13	Num objects	14	Num objects	42	Num objects	106
Num types	3	Num types	6	Num types	8	Num types	10

Table 2. This table presents the correlation between elite seed fitness in the final generation and elite seed productivity in the final generation. The productivity of an elite seed is measured by the number of ashes it creates. The correlation between fitness and productivity (0.706) is high and significant. We evaluate the statistical significance of the correlation using a two-tailed Student t-test for Pearson correlations. Each correlation is based on comparing two samples of 48 values each (4 layers \times 12 runs).

	Fitness in the final generation (see Figure 2)	Productivity in the final generation (see Figure 4)
Layer 1	0.743	13.0
Layer 2	0.846	17.5
Layer 3	0.810	17.0
Layer 4	0.936	73.1
Correlation	0.706	
p-value of Correlation	4.20E-12	
p-value < 0.05	Yes	

Table 3. This table presents the correlation between elite seed fitness in the final generation and elite seed diversity in the final generation. The fitness of an elite seed is measured by competitions between the elite seed and random seeds of matched size and density. The diversity of an elite seed is measured by the number of different types of ash that it yields. The correlation between fitness and diversity (0.834) is high and significant. We evaluate the statistical significance of the correlation using a two-tailed Student t-test for Pearson correlations. Each correlation is based on comparing two samples of 48 values each (4 layers \times 12 runs).

	Fitness in the final generation (see Figure 2)	Diversity in the final generation (see Figure 5)
Layer 1	0.743	3.1
Layer 2	0.846	4.0
Layer 3	0.810	3.7
Layer 4	0.936	7.9
Correlation	0.834	
p-value of Correlation	9.15E-20	
p-value < 0.05	Yes	

Table 4. This table lists the types of ash that are created in each layer of model S. For each layer, we collect all the ashes generated from the 600 seed patterns in generation 100 (50 elite seeds \times 12 runs). For each type of ash in each layer of model S, we rank the ashes in order of decreasing frequency and we show the rank of the type in Catagolue [7], where ashes are also ranked in order of decreasing frequency. In general, the rank of each ash type in model S is similar to the rank in Catagolue (we have marked some possible outliers in italics). The high correspondence between the ranks in the model S layers and the ranks in Catagolue suggests that the relative probability of a given type of ash occurring in model S is much the same as the relative probability of that type of ash occurring in any random soup.

Layer 1	Rank	Layer 2	Rank	Layer 3	Rank	Layer 4	Rank
block	1	block	1	block	1	block	1
blinker	2	blinker	2	blinker	2	blinker	2
<i>pond</i>	9	glider	4	<i>pond</i>	9	beehive	3
beehive	3	beehive	3	glider	4	glider	4
glider	4	pond	9	beehive	3	boat	6
ship	7	loaf	5	boat	6	ship	7
boat	6	ship	7	ship	7	loaf	5
		tub	8	loaf	5	tub	8
		boat	6	long boat	10	pond	9
		<i>pulsar</i>	<i>21</i>	tub	8	toad	11
		toad	11	ship-tie	12	pulsar	21
						<i>loop</i>	<i>49</i>
						half-bakery	15
						long boat	10
						beacon	13
						mango	16
						ship-tie	12
						barge	14
						integral sign	25
total types	7	total types	11	total types	11	total types	19
total freq	7,799	total freq	10,490	total freq	10,187	total freq	43,865
num seeds	600	num seeds	600	num seeds	600	num seeds	600
freq / seeds	13.0	freq / seeds	17.5	freq / seeds	17.0	freq / seeds	73.1

Table 5. Here we compare the productivity of the final generations of the four layers with and without shuffling. The rightmost column shows that shuffling greatly reduces productivity. The key factor for productivity is the specific evolved structure of the seed pattern, not the area of the pattern.

	Productivity intact (see Figure 4)	Productivity shuffled (see Figure 6)	Ratio shuffled / intact
Layer 1	13.00	0.94	7%
Layer 2	17.48	2.52	14%
Layer 3	16.98	1.71	10%
Layer 4	73.11	7.18	10%

Table 6. Here we compare the diversity of the final generations of the four layers with and without shuffling. The rightmost column shows that shuffling greatly reduces diversity. This supports the claim that, although increasing area is required for increasing diversity, the key factor for diversity is the specific evolved structure of the seed pattern.

	Diversity intact (see Figure 5)	Diversity shuffled (see Figure 7)	Ratio shuffled / intact
Layer 1	3.05	0.57	19%
Layer 2	4.05	0.82	20%
Layer 3	3.73	0.68	18%
Layer 4	7.87	2.57	33%