Toward Open-Ended Fraternal Transitions in Individuality

Matthew Andres Moreno and Charles Ofria

BEACON Center, Michigan State University, East Lansing, MI 48824
mmore500@msu.edu

Abstract

The emergence of new replicating entities from the union of simpler entities represent some of the most profound events in natural evolutionary history. Such transitions in individuality are essential to the evolution of the most complex forms of life. As such, understanding these transitions is critical to building artificial systems capable of open-ended evolution. Alas, these transitions are challenging to induce or detect, even with computational organisms. Here, we introduce the DISHTINY (DIStributed Hierarchical Transitions in IndividualitY) platform, which provides simple cell-like organisms with the ability and incentive to unite into new individuals in a manner that can continue to scale to subsequent transitions. The system is designed to encourage these transitions so that they can be studied: organisms that coordinate spatiotemporally can maximize the rate of resource harvest, which is closely linked to their reproductive ability. We demonstrate the hierarchical emergence of multiple levels of individuality among simple cell-like organisms that evolve parameters for manually-designed strategies. During evolution, we observe reproductive division of labor and close cooperation among cells, including resource-sharing, aggregation of resource endowments for propagules, and emergence of an apoptosis response to somatic mutation. Many replicate populations evolved to direct their resources toward low-level groups (behaving like multi-cellular individuals) and many others evolved to direct their resources toward high-level groups (acting as larger-scale multi-cellular individuals).

Introduction

Artificial life researchers design systems that exhibit properties of biological life in order to better understand their dynamics and, often, to apply these principles toward engineering applications such as artificial intelligence (Bedau, 2003). Studies of evolution have been of particular interest to the community, especially in regard to how organisms are produced with increasing sophistication and complexity (Goldsby et al., 2017). This particular issue is often described as "open-ended evolution." Although precise definitions and measures of open-ended evolution are still being established, this term is generally understood to refer to evolving systems that exhibit the continued production of novelty (Taylor et al., 2016). Evolutionary transitions in individuality, which are key to the complexification and diversification of biological life (Smith and Szathmary, 1997), have been highlighted as key research targets with respect to the question of open-ended evolution (Ray, 1996; Banzhaf et al., 2016). In an evolutionary transition of individuality, a new, more complex replicating entity is derived from the combination of cooperating replicating entities that have irrevocably entwined their long-term fates (West et al., 2015). In particular, we focus on fraternal transition in individuality, events where closely-related kin come together or stay together to form a higher-level organism (Queller, 1997). Eusocial insect colonies and multicellular organisms exemplify this phenomenon (Smith and Szathmary, 1997). Like the definition of open-ended evolution, the notion of what constitutes an evolving individual is not concretely established. Commonly indicated features include: close coordination and cooperation, reproductive division of labor, reproductive bottlenecks, and loss of ability to replicate independently (Ereshefsky and Pedroso, 2015; Bouchard, 2013).

Our appreciation of fraternal transitions in individuality benefits from experimental work probing the origins of multicellularity. In the biological domain, Ratcliff et al. have demonstrated evolution of multicellularity in yeast, deriving fraternal clusters of cells that cling together in order to maximize their settling rate (Ratcliff et al., 2012). The contributions of Goldsby and collaborators are particularly notable among computational Artificial life work on the origins of multicellularity. Their evolutionary experiments track a population composed of demes, distinct spatial domains inhabited by clonal colonies of cells. Two distinct types of reproduction occur: (1) cells reproduce within demes and (2) deme reproduction, where a target deme is sterilized then re-innoculated with genetic material from the parent deme. With such methods, Goldsby et al. have studied division of labor (Goldsby et al., 2010, 2012), the origin of soma (Goldsby et al., 2014), and the evolution of morphological development (Goldsby et al., 2017). We aspire to complement deme-based approaches with a framework where higher level individuality unfolds via cellular reproductions within a single unified space. In particular, we are interested

in the potential for such a system to undergo nested hierarchical transitions.

Major challenges in studying evolutionary transitions in individuality include (1) determining the environmental conditions that will promote such a transition and then (2) recognizing that a transition has occurred. In order to begin exploring transitions in individuality, we must devise a system in which we expect such transitions to occur repeatably and in a detectable manner. Once we can consistently induce and observe evolutionary transitions in individuality, we may subsequently proceed to relax aspects of such a system to explore in greater detail what conditions are necessary to induce transitions and how transitions can be detected. For now, we will focus on these initial goals in the context of fraternal transitions in individuality.

To this end, we introduce the DISHTINY (DIStributed Hierarchical Transitions in IndividualitY) platform, which seeks to achieve the evolution of transitions in individuality by explicitly registering organisms in cooperating groups that coordinate spatiotemporally to maximize the harvest of a resource. Detection of such a transition in DISHTINY is accomplished by identifying resource-sharing and reproductive division of labor among organisms registered to the same cooperating group. We designed this system such that hierarchal transitions across an arbitrary number of levels of individuality can be selected for and meaningfully detected. We have focused this system on a rigid form of major transition using simple organisms, but the underlying principles can be applied to a wide range of artificial life systems. Furthermore, DISHTINY is decentralized and amenable to massive parallelization via distributed computing. We believe that such scalability — with respect to both concept and implementation — is an essential consideration in the pursuit of artificial systems capable of generating complexity and novelty rivaling that of biological life via open-ended evolution (Ackley and Cannon, 2011; Ackley, 2016).

Methods

In order to demonstrate that the DISHTINY platform selects for detectable hierarchical transitions in individuality, we performed experiments where cell-like organisms evolved parameters to control manually designed behaviors such as resource-sharing, reproductive decision-making, and apoptosis. We will first cover the design of the DISHTINY platform and then describe the simple cell-like organisms we used to evaluate the platform.

DISHTINY

DISHTINY allows cell-like organisms to replicate across a toroidal grid. Over discrete timesteps ("updates"), the cells can collect a continuous-valued resource. Once sufficient resource has been accrued, cells may pay 8.0 resource to place a daughter cell on an adjoining tile of the toroidal grid (i.e., reproduce), replacing any existing cell already there. As cells reproduce, they can choose to include offspring in the parent's cooperating "signaling channel" group or force offspring to create a new cooperating "signaling channel" group.

As shown at the top of Figure 1, resources appear at a single point then spread outwards update-by-update in a diamond-shaped wave, disappearing when the expanding wave reaches a predefined limit. Cells must be in a costly "activated" state to collect resource as it passes. The cell at the starting position of a resource wave is automatically activated, and will send the activate signal to neighboring cells on the same signaling channel. The newly activated cells, in turn, activate their own neighbors registered to the same signaling channel. Neighbors registered to other signaling channels do not activate. Each cell, after sending the activation signal, enters a temporary quiescent state so as not to reactivate from the signal. In this manner, cells sharing a signaling channel activate in concert with the expanding resource wave. As shown Figure 1a, b, the rate of resource collection for a cell is determined by the size and shape of of its same-channel signaling network; small or fragmented same-channel signaling networks will frequently miss out on resource as it passes by.

Each cell pays a resource cost when it activates. This cost is outweighed by the resource collected such that cells that activate in concert with a resource wave derive a net benefit. Recall, though, that resource waves have a limited extent. Cells that activate outside the extent of a resource wave or activate out of sync with the resource wave (due to an indirect path from the cell that originated the signal) pay the activation cost but collect no resource. Cells that frequently activate erroneously use up their resource and die. In our implementation, organisms that accrue a resource debt of -11 or greater are killed. This erroneous activation scenario is depicted in Figure 1c.

In this manner, "Goldilocks" — not to small and not too big — signaling networks are selected for. Based on a randomly chosen starting location, resource wave starting points (seeds) are tiled over the toroidal grid such that the extents of the resource waves touch, but do not overlap. All waves start and proceed synchronously; when they complete, the next resource waves are seeded. This process ensures that selection for "Goldilocks" same-channel signaling networks is uniformly distributed over the toroidal grid.

Cells control the size and shape of their same-channel signaling group through strategic reproduction. Three choices are afforded: whether to reproduce at all, where among the four adjoining tiles of the toroidal grid to place their offspring, and whether the offspring should be registered to the parent's signaling channel or be given a random channel ID (in the range 1 to 2^{22}). No guarantees are made about the uniqueness of a newly-generated channel ID, but chance collisions are rare.

Hierarchical levels are introduced into the system through

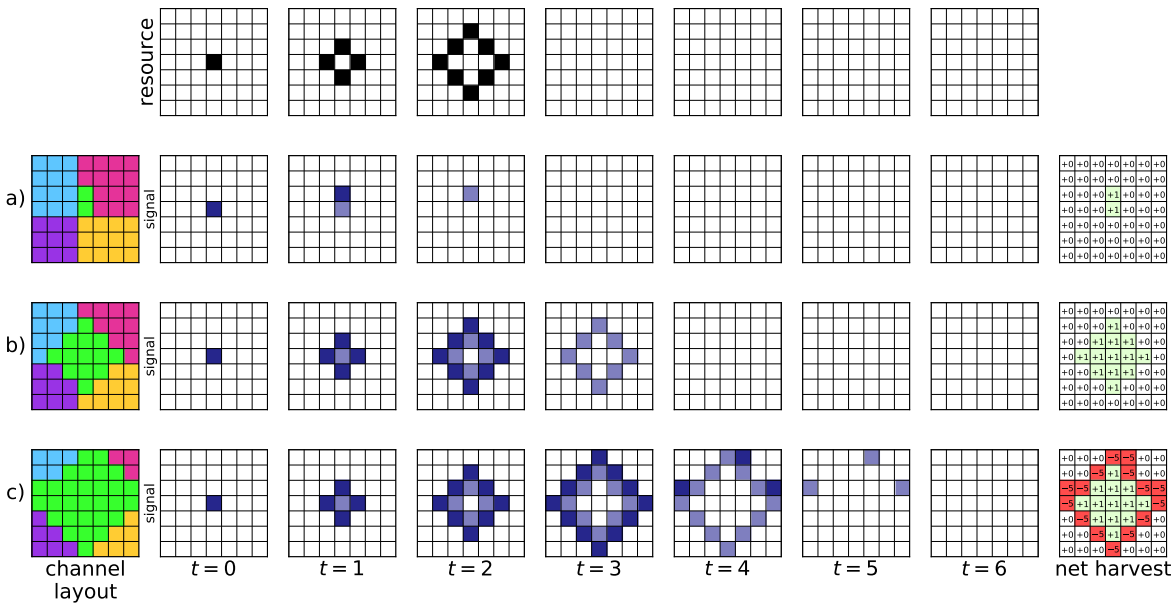


Figure 1: **Activation signaling, and net resource collection for three different-sized same-channel networks during a resource wave event.** At the top, a resource wave is depicted propagating over three updates and then ceasing for four updates (left to right). In row a, a small two-cell channel-signaling group (far left, in green) is activated; tracking the resource wave (top) yields a small net resource harvest (far right). In row b, an intermediate-sized 13-cell channel-signaling group yields a high net resource harvest. Finally, in row c, a large 29-cell channel-signaling group incurs a net negative resource harvest. In rows a, b, and c, dark purple indicates the active state, light purple indicates the quiescent state, and white indicates the ready state.

multiple separate, but overlaid, instantiations of this resource wave/channel-signaling scheme. We refer to each independent resource wave/channel-signaling system as a "level." In our experiments, we allowed two resource wave/channel-signaling levels, identified here as level one and level two. On level one, resource waves extended a radius of three toroidal tiles. On level two they extended a radius of 24 toroidal tiles. On both levels, activated cells netted $+1.0$ resource from a resource wave, but suffered an activation penalty of -5.0 if no resource was available. Due to the different radii of resource waves on different levels, level one selects for small same-channel signaling networks and level two selects for large same-channel signaling networks.

Cells were marked with two separate channel IDs, one for level one and another for level two. We enforced hierarchical nesting of same-channel signaling networks during reproduction: daughter cells may inherit neither channel ID, just the level-two channel ID, or both channel IDs. Daughter cells may not inherit only the level-one channel ID while having a different level-two channel ID. The distribution of IDs across the level-two and level-one channels can be envisioned by analogy to political countries and territories. Each country (i.e., level-two channel network) may have one or many territories (i.e., level-one channel network). However, no territory spans more than one country. Figure 2 depicts hierarchically nested channel states at the end of three evolutionary runs.

Channel IDs enable straightforward detection of an evolutionary transition in individuality. Because common channel IDs may only arise systematically through inheritance, common channel IDs indicate a close hereditary relationship in addition to a close cooperative relationship. Because new channel IDs arise first in a single cell, same-channel signaling networks are reproductively bottlenecked, ensuring meaningful reproductive lineages at the level of the same-channel signaling network. To recognize an evolutionary transition in individuality, we therefore evaluate

1. Do cells with the same channel ID choose to share resources (e.g., cooperate)?

2. Is there division of reproductive labor between members of the same channel (e.g., do cells at the interior of a network cede reproduction to those at the periphery?)

If these conditions are met among cells sharing the same level-one channel, a first-level transition in individuality may have occurred. Likewise, if these conditions are met among cells sharing the same level-two channel, a second-level transition in individuality may have occurred. In either case, observation of altruistic behavior, such as an apoptosis response to mutation, would further evidence a transition.

Organisms

We performed our experiments using cell-like organisms composed of 15 floating-point parameters, each controlling a specific strategy component pertinent to transitions in individuality (i.e., reproductive division of labor, resource pooling, apoptosis, propagule generation, and propagule endowment). These particular cell-like organisms are in no way inherent to the DISHTINY platform, but were merely developed to study transitions using as simple a model system as feasible. On reproduction, we applied mutation to each parameter independently with probability 0.00005.

The **aversion parameters** ($\boldsymbol{A_1}$ and $\boldsymbol{A_2}$) allow cells to avoid reproducing over neighbors sharing the same signaling channel. Specifically, they control the probability that a cell declines to supplant a neighbor sharing the same level-one (A_1) or level-two (A_2) channel ID. If a cell declines to place its offspring in all four adjoining tiles, it does not reproduce. Mutation is performed by a redraw from the uniform distribution $U(-0.5, 1.5)$ clamped to the range $[0, 1]$.

The **resource allocation parameters** control the proportion of resources that go to the cell's stockpile ($\boldsymbol{P_c}$), its level-one channel's resource pool ($\boldsymbol{P_1}$), or its level-two channel's resource pool ($\boldsymbol{P_2}$). These parameters are initialized by a draw from $U(-1.0, 2.0)$ clamped to the range $[0, 1]$ and mutated by addition of a normal value drawn from $N(0.0, 0.2)$ with the result clamped to the range $[0, 1]$. The set P_c, P_1, P_2 is always normalized to sum to 1.

Channel resource pools are identical to an organism's individual stockpile, except that any deficit is distributed evenly among the individual organism's stockpile. On every update, cells can spend from their individual stockpile to reproduce or from a channel pool, with priority given to cells nearest to the centroid of that pool's members. As such, pool-funded reproduction fills in a same-channel signaling network from the inside out and help produce diamond-shaped same-channel signaling networks. (Distance is measured using the taxicab metric.)

Channel cap parameters $\boldsymbol{C_1}$ and $\boldsymbol{C_2}$ regulate the size of same-channel signaling networks. When an organism reproduces, it checks the size of its level-one signaling network against C_1 and the size of its level-two signaling group against C_2. If neither cap is met or exceeded, then the organism will produce an offspring sharing both of its channel IDs. If only the C_1 cap is exceeded, then the organism will produce an offspring with new level-one channel ID but identical level-two channel ID. Finally, if the C_2 cap is exceeded, then the organism will produce an offspring with new IDs for both channels. For level-one caps, these parameters are initialized by a draw from $U(0.0, 16.0)$. For level-two caps, these parameters are initialized by a draw from $U(0.0, 128.0)$. Both are mutated by addition of a value drawn from $N(0.0, 24.0)$ with the result clamped to be non-negative.

The **endowment parameters** $\boldsymbol{E_c}$, $\boldsymbol{E_1}$, and $\boldsymbol{E_2}$ determine the amount of resource provided to offspring. This endowment is paid as an additional cost by the cell stockpile (or same-channel resource pool) funding a reproduction. The full amount of the received endowment is divided between the daughter cell's stockpile, level-one same-channel resource pool, and level-two same-channel resource pool according to the offspring's resource allocation parameters. $\boldsymbol{E_c}$ is the endowment amount paid to an offspring that shares both channel IDs of the parent; $\boldsymbol{E_1}$ is the endowment paid to an offspring that shares just the level-two channel ID of the parent; and $\boldsymbol{E_2}$ is the endowment paid to an offspring that shares neither the level-one nor the level-two channel ID of the parent. Endowed resources help new-channel propagules to rapidly grow their signaling network in order to begin collecting resource at a rate competitive to other well-established same-channel signaling networks. In order that adequate resource remain to ensure parental stability, endowment was paid out only after twice the endowment amount had been accrued (leaving an amount of resource equal to the endowment remaining with the parent). Cell level endowments are initialized by a draw from $U(0.0, 5.0)$. Level-one endowments are initialized by a draw from $U(0.0, 80.0)$. Level-two endowments are initialized by a draw from $U(0.0, 405.0)$. All endowments are mutated by addition of a value drawn from $N(0.0, 10.0)$ with the result clamped to be non-negative

Parameters $\boldsymbol{M_c}$, $\boldsymbol{M_1}$, and $\boldsymbol{M_2}$ control the **apoptosis response to mutation**. Each time that a mutation occurs during reproduction, the mutated offspring attempts suicide with probability M_c if it shares both channel IDs of its parent, probability M_1 if it shares just the level-two channel ID of its parent, and probability M_2 if it shares neither channel ID of the parent. The M_x value applied is from the offspring's genotype after mutation. Attempted suicide succeeds 90% of the time. This capacity enables first- or second-level individuals to combat somatic mutation. Initialization and mutation each of these parameters is performed by a redraw from the distribution $U(-0.5, 1.5)$ clamped to the range $[0, 1]$.

Finally, parameters $\boldsymbol{S_1}$ and $\boldsymbol{S_2}$ **fine-tune site choice for offspring placement**. If an organism is placing an offspring with identical channel IDs, with probability S_1 the four possible sites for offspring placement are considered in order of increasing distance from the centroid of the parent's level-one signaling network. If an organism is placing an offspring with identical level-two channel ID but differ-

ent level-one channel ID, with probability S_2 the four possible sites for offspring placement are considered in order of increasing distance from the centroid of the parent's level-two same-channel signaling network. Otherwise, the four possible sites for offspring placement are considered in a random order. Initialization and mutation are performed by a draw from the distribution $U(-0.5, 1.5)$ clamped to the range $[0, 1]$.

Treatments

Our standard treatment was designed to assess the evolutionary trajectories of populations in DISHTINY. We seeded each tile on the 120×120 toroidal grid with a randomized organism and ran the simulation for 20 million updates. In order to facilitate turnover, we culled the population intermittently. Starting at update 500,000, and every 50,000 updates thereafter, we randomly selected second-level channel IDs and killed all cells with that channel ID, continuing until at least 5% of grid tiles were empty. We performed 50 replicates within this treatment. On average, each cellular generation took just over 500 updates. Across all successive 10,000 update segments of all replicates, the mean number of cellular generations elapsed per 10,000 updates was 19.2 with a standard deviation of 2.7 cellular generations per 10,000 updates.

In order to detangle the impact of same-channel signaling networks with respect to kin recognition versus cooperation to increase resource collection rate, we performed control evolutionary trials where same-channel signaling networks did not affect cellular resource collection rate. Under control conditions, same-channel signaling networks just helped cells recognize other related cells. In our implementation, this treatment corresponded to a constant per-update inflow of 0.02 resource units into all cells. All cells were activated (in order to take up the resource) at all updates and no cost for activation was assessed. We chose this resource inflow rate in order to approximately match the cellular generation rate of the control treatment to that of the standard treatment. In control runs, each cellular generation took around 450 updates. Across all successive 10,000 update segments of all replicates, the mean number of cellular generations elapsed per 10,000 updates was 22.0 with a standard deviation of 2.0 cellular generations per 10,000 updates. Due to checkpoint-restart failures on our compute cluster, control experiments were curtailed at 3 million updates. All other aspects of control runs, including culling and the functionality of all lifestyle parameters, were otherwise identical to standard conditions. We performed 50 replications of the control treatment.

In standard evolutionary runs, we observed a spectrum of evolved resource-caching strategies. To assess the relative fitness of these evolved organisms, we ran competitions between the most common genotype from three standard evolutionary runs. The first genotype allocated resource exclusively to its first-level same-channel resource pool (i.e., $P_1 = 1.0$), the second split resource evenly between its first-level and second-level resource pool (i.e., $P_1 = P_2 = 0.5$), and the third allocated resource primarily to the second-level resource pool (i.e., $P_2 > P_1$). (No most-common genotypes allocated resource exclusively to the second-level resource pool.) We seeded each competition with three copies of each genotype, uniformly spaced over the 120×120 toroidal grid with random arrangement. Each competition lasted 2 million updates. We performed 50 runs in this experiment.

Implementation

We performed our computational experiments at the Michigan State University High Performance Computing Center. Each replicate of standard evolutionary experiments required approximately six days of compute time to reach 20 million updates. Each replicate of control evolutionary experiments expended approximately two days of compute time to reach 3 million updates. Control runs were somewhat slower than standard runs, perhaps due to increased computational overhead associated with bookkeeping for the larger same-channel groups that evolved under the control conditions. Each replicate of competition experiments consumed approximately ten hours of compute time. For standard evolutionary experiments, data processing required approximately four hours of compute time per run. Other data processing was computationally negligible.

We implemented our experimental system using the Empirical library for scientific software development in C++, available at `https://github.com/devosoft/Empirical`. The code used to perform and analyze our experiments, our figures, data from our experiments, and a live in-browser demo of our system is available via the Open Science Framework at `https://osf.io/ewvg8/`.

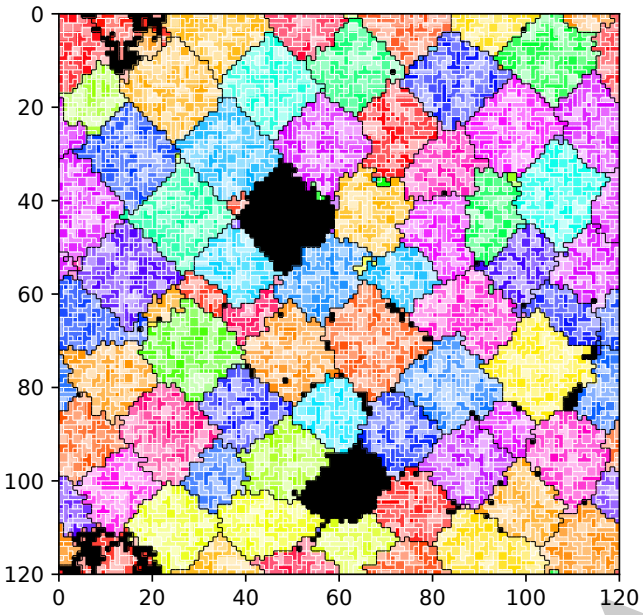
Results and Discussion

Standard Evolutionary Experiments

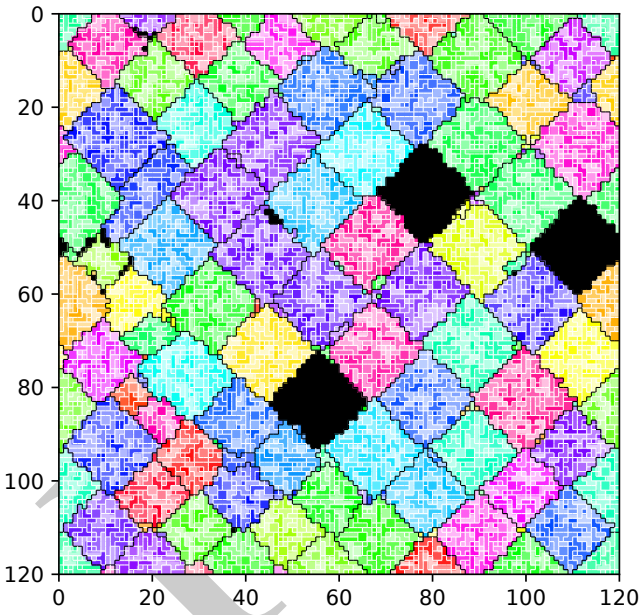
A spectrum of resource allocation strategies ranging from purely allocation to level-one same-channel resource pools to primarily allocation to level-two same-channel resource pools were observed at the conclusion of different runs of our evolutionary simulation (mean cellular generation 37,168 with standard deviation 4,684). We interpret these outcomes as ranging between individuality at the level of first-level same-channel groups to individuality at the level of second-level same-channel groups. Figure 2 shows the level-one and level-two signaling networks at the end of runs where first-, split-, and second-level resource allocation evolved, respectively. First-level allocators form somewhat irregular level-two amalgamations of diverse level-one networks. Second-level allocators form highly regular diamond-shaped level-two signaling networks. Split-allocation individuals exhibit a level-two phenotype of intermediate regularity. Figure 3 shows a time series of signal-

	Competitors			Mean Dominant ($\pm S.D.$)			Pop Mean ($\pm S.D.$)		Control Pop Mean ($\pm S.D.$)
	$P_1 = 1.0$	$P_2 = P_1$	$P_2 > P_1$	$P_1 = 1.0$	$1.0 > P_1 > P_2$	$P_2 \geq P_1$	*all*	*all*	*all*
Cell Gen.	29920	33852	47507	30841 ± 3183	35346 ± 3444	39315 ± 3346	6670 ± 729	6069 ± 672	6626 ± 377
Upd.	20M	20M	20M	20M	20M	20M	3.3M	3M	3M
n	1	1	1	9	7	34	50	50	50
A_1	0.00	0.00	0.89	0.23 ± 0.35	0.50 ± 0.47	0.57 ± 0.46	0.53 ± 0.37	0.53 ± 0.35	0.56 ± 0.34
A_2	1.00	1.00	1.00	1.00 ± 0.00	1.00 ± 0.00	1.00 ± 0.00	1.00 ± 0.00	1.00 ± 0.00	0.99 ± 0.01
P_c	0.00	0.00	0.00	0.00 ± 0.00	0.00 ± 0.00	0.03 ± 0.05	0.02 ± 0.03	0.02 ± 0.02	0.00 ± 0.00
P_1	1.00	0.50	0.00	1.00 ± 0.00	0.60 ± 0.07	0.28 ± 0.16	0.42 ± 0.25	0.42 ± 0.24	0.56 ± 0.37
P_2	0.00	0.50	1.00	0.00 ± 0.00	0.40 ± 007	0.69 ± 0.14	0.56 ± 0.24	0.56 ± 0.24	0.44 ± 0.37
C_1	3.13	3.45	2.04	3.90 ± 0.60	3.38 ± 0.33	3.03 ± 0.69	3.21 ± 0.63	3.21 ± 0.60	28.6 ± 21.7
C_2	233.2	238.6	290.2	230.6 ± 71.1	192.7 ± 45.3	271.6 ± 73.6	201.5 ± 58.1	195.8 ± 55.3	484.0 ± 123.5
E_c	0.87	0.14	4.20	0.29 ± 0.37	0.44 ± 0.59	0.21 ± 0.75	1.14 ± 1.07	1.21 ± 1.05	1.50 ± 1.08
E_1	33.4	11.7	4.80	47.2 ± 21.7	21.3 ± 12.0	4.62 ± 7.05	18.1 ± 16.2	19.2 ± 15.9	28.9 ± 22.2
E_2	341.4	397.4	321.1	231.2 ± 94.3	283.1 ± 57.0	325.4 ± 68.9	303.0 ± 66.5	302.7 ± 65.2	317.3 ± 66.3
M_c	0.11	1.00	0.66	0.33 ± 0.41	0.74 ± 0.31	0.67 ± 0.35	0.39 ± 0.32	0.39 ± 0.31	0.18 ± 0.23
M_1	0.00	1.00	0.40	0.52 ± 0.41	0.65 ± 0.46	0.68 ± 0.38	0.52 ± 0.37	0.51 ± 0.35	0.48 ± 0.33
M_2	0.00	0.44	1.00	0.45 ± 0.39	0.52 ± 0.37	0.50 ± 0.42	0.47 ± 0.33	0.47 ± 0.32	0.53 ± 0.36
S_1	0.00	1.00	1.00	0.65 ± 0.38	0.55 ± 0.40	0.47 ± 0.42	0.39 ± 0.36	0.40 ± 0.34	0.47 ± 0.34
S_2	0.00	0.01	0.46	0.51 ± 0.43	0.35 ± 0.39	0.45 ± 0.39	0.47 ± 0.34	0.46 ± 0.34	0.55 ± 0.34

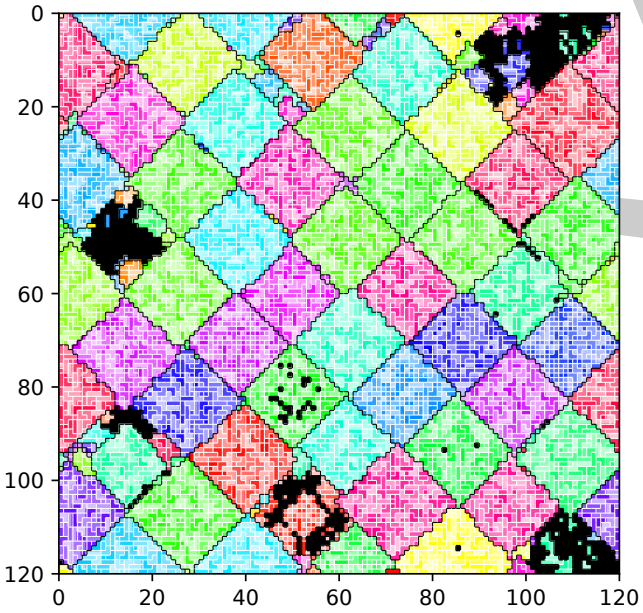
Table 1: The leftmost two table segments enumerate genotypes used as seeds for competition experiments ("Competitors") and the mean values of the most abundant genotype at the end of evolutionary runs ("Mean Dominant"), both partitioned by resource-caching strategy. The rightmost table segments enumerate the population mean genotype values for standard evolutionary trials ("Pop Mean") and control treatments ("Control Pop Mean"), matched at both absolute update count and (approximately) elapsed cellular generations.



(a) Mean P_c=0.03, P_1=0.75, P_2=0.23; cell gen. 29920



(b) Mean P_c=0.03, P_1=0.51, P_2=0.49; cell gen. 33852



(c) Mean P_c=0.08, P_1=0.01, P_2=0.90; cell gen. 47507

Figure 2: End state of same-channel signaling networks in replicates where resource was exclusively allocated to first-level channel pools (2a), was split evenly between first- and second-level channel pools (2b), and was primarily allocated to second-level channel pools (2c). Level-one channels are coded by color saturation and level-two channels are coded by color hue. A single cell-like organism occupies each grid tile except for black tiles, which are empty.

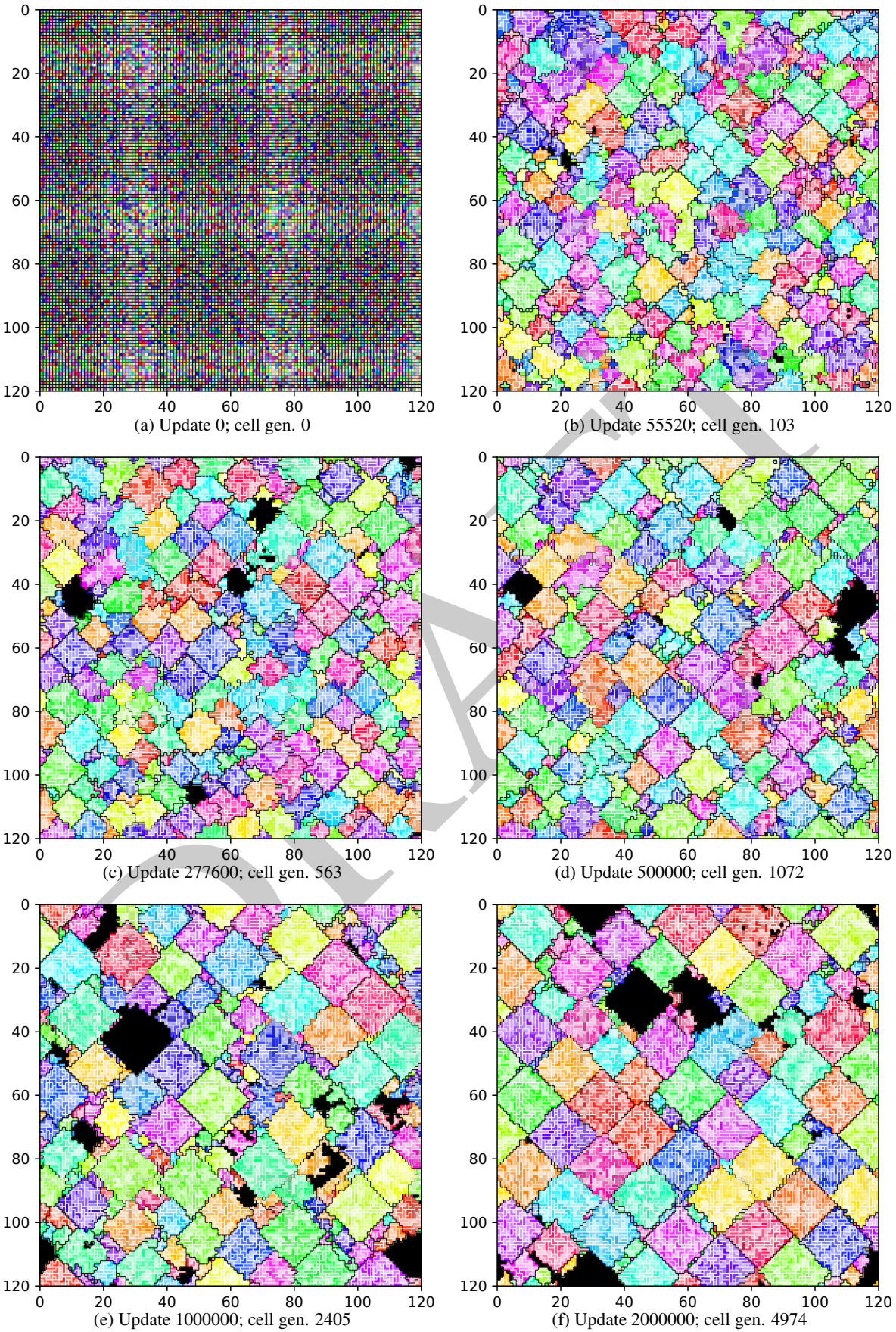


Figure 3: Progression of of same-channel level-one and level-two signaling networks states in an evolutionary run where level-two resource sharing evolved. Level-one channels are coded by color saturation and level-two channels are coded by color hue. A single cell-like organism occupies each grid tile except for black tiles, which are empty.

ing network snapshots in an evolutionary run where second-level individuality evolved.

Table 1 summarizes most-common genotypes observed at the end of our evolutionary simulations. In the standard treatment, all evolved genotypes had A_2 fixed at 1.0. So, reproduction over cells sharing the same level-two channel was universally avoided; genotypes evolved so that cells declined to reproduce when they were located at the interior of level-two same-channel signaling networks.

However, a variety of resource-caching strategies evolved. Most-abundant genotypes at the end of nine evolutionary runs exclusively cached resource in organisms' level-one signaling network's pool (i.e., $P_1 = 1.0$). We observed strategies where resource was primarily, but not entirely, cached in an organism's level-one signaling network pool (i.e., $1.0 > P_1 > P_2$) as the most-abundant genotype at the end of seven evolutionary runs. In one run, the most-abundant final genotype split resources evenly between an organism's level-one and level-two signaling network pool ($P_1 = P_2 = 0.5$). Finally, we observed strategies where resource was primarily, but not entirely, cached in an organism's level-two signaling network pool (i.e., $1.0 > P_2 > P_1$) as the most-abundant genotype at the end of 33 evolutionary runs.

We suspect that a trade-off between growth rate and long-term stability prompted the universal allocation of at least some resource to level-one pools and/or cell stockpiles. Cell- and level-one resource caching might function something like saving for a rainy day. Because reproduction over level-two channel-mates was universally avoided, cells and level-one same-channel networks situated at the interior of a larger level-two same-channel network do not expend their resource pools unless that larger level-two same-channel network is damaged, exposing them to directly-adjacent cells of a different level-two channel. Thus, resource accumulates in cell stockpiles and level-one pools until the level-two same-channel network comes under stress. Split allocation might also represent hedging against defection of a second-level channel-mate by via somatic mutation.

Indeed, we did observe selection for apoptosis in the 41 replicates where the dominant genotype employed second-level resource caching. In these replicates, the average population mean value of M_c was 0.68 with standard deviation 0.33, significantly greater than the value $M_c = 0.5$ we would expect in the absence of a selective pressure on apoptosis response to mutation ($p < 0.001$, bootstrap test).

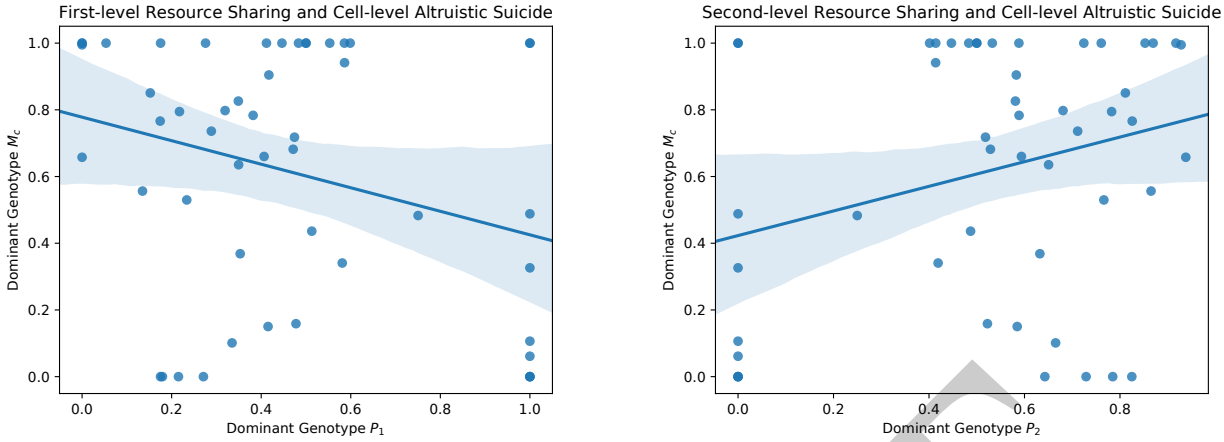
To assess whether heavy second-level resource allocators, which we characterize as higher-level individuals, were more likely to employ apoptosis to mitigate somatic mutation, we examined the relationship between first- and second-level resource pooling and cellular apoptosis at the conclusion of our 50 replicate evolutionary trials. We observed a significant negative correlation between dominant genotype P_1 and M_c ($p < 0.05$; bootstrap test; Figure 4a) and a significant positive correlation between dominant genotype P_2 and M_c ($p < 0.05$; bootstrap test; Figure 4b). This result suggests that second-level individuals, in particular, relied on apoptosis to mitigate somatic mutation.

We also assessed whether higher-level individuals provided larger resource endowments to their second-level propagules (offspring sharing neither the level-one nor the level-two channel ID with the parent). We examined the relationship between first and second-level resource pooling and dominant genotype second-level propagule endowment at the conclusion of our 50 replicate evolutionary trials. We observed a significant negative correlation between dominant genotype P_1 and E_2 ($p < 0.05$; bootstrap test) and a significant positive correlation between dominant genotype P_2 and E_2 ($p < 0.05$; bootstrap test). Second-level individuals might provide larger endowments to propagules simply due to a greater capacity to collect resource or perhaps because of stronger selection for well-endowed offspring when competing against other second-level individuals.

This result prompts the reverse question: do lower-level individuals provide larger resource endowments to first-level propagules (offspring that do not share level-one channel ID with the parent but may or may not share level-two channel ID with the parent)? Indeed, we observed a significant positive correlation between first-level resource sharing and first-level endowment ($p < 0.0001$; bootstrap test) and a significant negative correlation between second-level resource sharing and first-level endowment ($p < 0.0001$; bootstrap test). Cells that pool resource with their smaller level-one same-channel group tend to invest more heavily into the direct offshoots of their level-one same-channel group than cells that pool resource with their larger level-two same-channel group. This observation suggests that, although cells do not directly displace their level-one channel-mates, competitive dynamics between may be at play.

Competition Experiments

Next, we wanted to compare first-, second-, and split-level allocators to determine which genotype was the most fit. We ran competition experiments between dominant genotypes from evolutionary runs representative of each of these strategies. To prevent further evolution, we disabled mutation for these experiments. To represent first-level allocators, we selected randomly from the nine pure first-level allocator dominant genotypes we observed. To represent the split-level allocators, we selected the single dominant genotype where resource was partitioned exactly evenly between first- and second-level channel pools. To represent second-level allocators, we selected the dominant genotype with the largest second-level allocation proportion. Table 1 enumerates the three representative genotypes used. Figure 5 shows a time series of signaling network snapshots in an competition experiment run. Colonies of each genotype can be seen to grow from each seed and then clash, ultimately yielding a



(a) Correlation plot of dominant genotype P_1 and dominant genotype M_c.

(b) Correlation plot of dominant genotype P_2 and dominant genotype M_c.

Figure 4: Plots of dominant resource caching strategies and dominant apoptosis strategies. A bootstrapped 95% confidence interval for the fit is shaded. Both correlations are statistically significant ($p < 0.05$; bootstrap test).

population dominated by second-level allocators.

Indeed, the second-level resource caching strategy became most abundant in all 50 trials. Across the 50 replicates, at update 1.5 million (cellular generation 3489 with standard deviation 40) the second-level resource caching strategy constituted 90.2%, with standard deviation 3.8%, of the competing population of cells. In the absence of mutation, second-level allocators tend to exhibit greater fitness than split- and first-level allocators ($p < 0.0001$; two-tailed exact test).

In competition experiments, however, higher-level individuals likely benefited from elimination of somatic mutation. To assess the relative fitness of first- and second-level individuals without mutation disabled, we examined the relationship between first- and second-level resource pooling and the rate of cellular reproduction at the end of each of the 50 replicate evolutionary trials performed. We observed a significant negative correlation between mean P_1 and cellular reproduction rate ($p < 0.0001$; bootstrap test; Figure 6a) and a significant positive correlation between mean P_2 and cellular reproduction rate ($p < 0.0001$; bootstrap test; Figure 6b). This result suggests that second-level allocators tend to collect resource more effectively than split- and first-level allocators.

Control Evolutionary Experiments

Under control conditions where resource was distributed evenly to all cells regardless of same-channel group configuration, split-level resource caching also evolved. Split-level allocation was the most common strategy at update 3 million in all replicates. Strategies where resource was primarily, but not entirely, cached in an organism's level-one signaling network pool (i.e., $1.0 > P_1 > P_2$) were most-abundant at the end of 33 evolutionary runs and strategies where resource was primarily, but not entirely, cached in an organism's level-two signaling network pool (i.e., $1.0 > P_2 > P_1$) were most-abundant at the end of 17 evolutionary runs. As shown in Table 1, the average population mean of P_1 is greater in the control treatment than in the standard treatment at time-points matched by absolute elapsed update count and approximate elapsed cellular generations, but this difference is not statistically significant.

Consistent with the standard treatment, we observed strong selection against direct reproductive competition between channel-mates at update 3 million in the control treatment. Nearly all most-common genotypes completely avoided reproducing over level-two channel-mates (i.e., $A_2 = 1.0$), except for a single most-common genotype where a very slim probability of reproducing over level-two channel-mates was allowed ($A_2 = 0.996$).

The emergence of resource-sharing and competition avoidance under control conditions suggests kin recognition alone can prompt some aspects of higher-level individuality. However, we observed selection *against* the apoptosis response to mutation, M_c, under control conditions. Across 50 replicates of the control treatment, the average population mean value of M_c was 0.18 with standard deviation 0.23 — significantly less than the value $M_c = 0.5$ expected without selective pressure against apoptosis response to mutation ($p < 0.0001$, two-tailed t test). Indeed, population mean M_c for control runs was also significantly reduced compared to the standard treatment at time-points matched by absolute elapsed update count ($p < 0.001$; two-tailed t test) and by approximate elapsed cellular generations ($p < 0.001$; two-

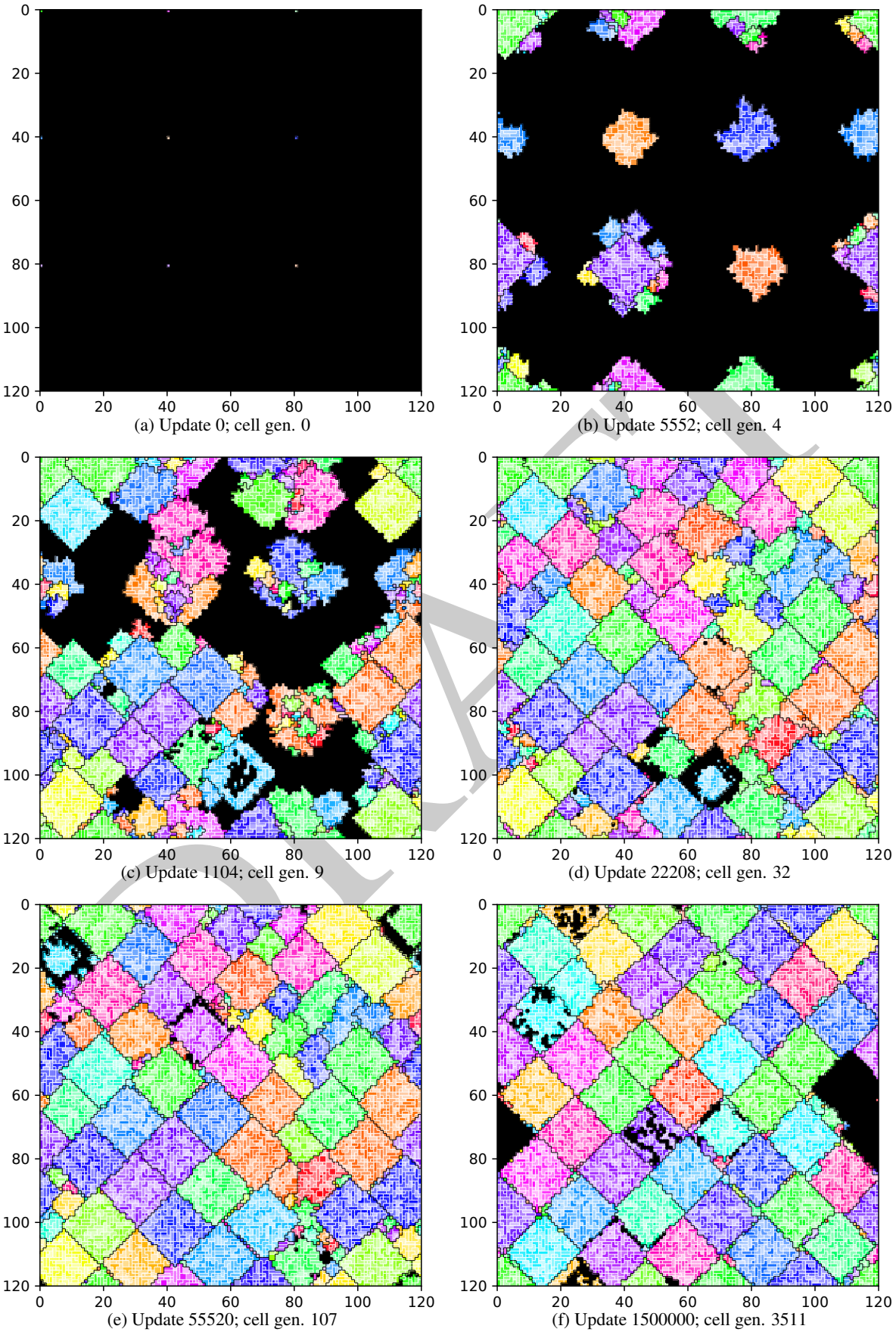
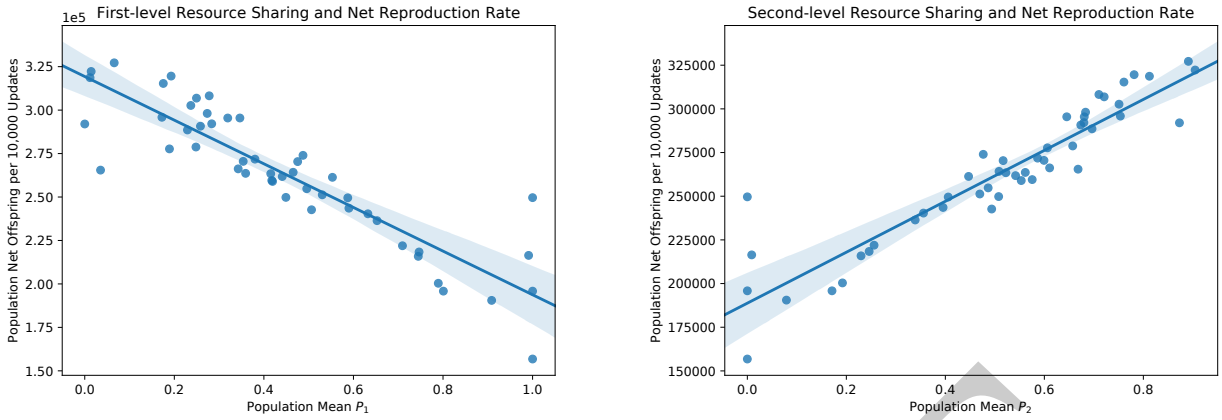


Figure 5: Progression of of same-channel level-one and level-two signaling networks states in a competition run. We seeded the grid with three copies of each of three champion genotypes from evolutionary trials. Then, with mutation disabled to prevent further evolution, the genotypes competed. Level-one channels are coded by color saturation and level-two channels are coded by color hue. A single cell-like organism occupies each grid tile except for black tiles, which are empty.



(a) Correlation plot of population mean P_1 and population net reproduction rate.

(b) Correlation plot of population mean P_2 and population net reproduction rate.

Figure 6: Mean resource caching strategies and net reproduction rate across populations. A bootstrapped 95% confidence interval for the fit is shaded. Both correlations are statistically significant ($p < 0.0001$; bootstrap test).

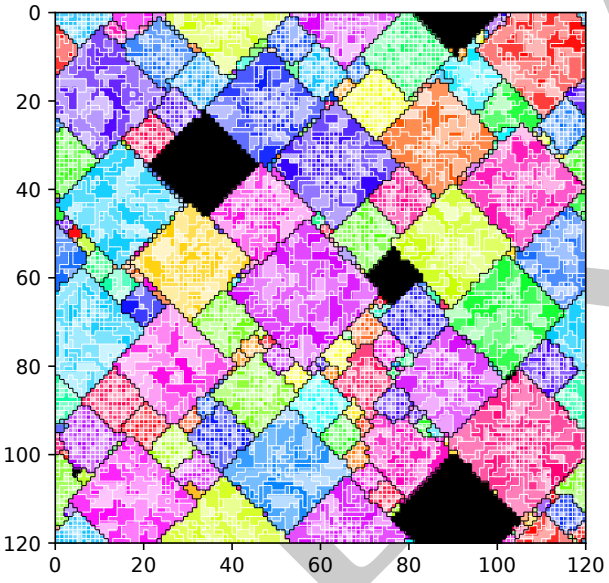


Figure 7: End state (update 3000000, cell gen. 6916) of same-channel signaling networks evolved under the control treatment. Level-one channels are coded by color saturation and level-two channels are coded by color hue. A single cell-like organism occupies each grid tile except for black tiles, which are empty.

tailed t test). Perhaps under control conditions, the apoptosis response to mutation is disfavored because kin groups stand to lose less from mutant members (i.e., the resource penalty for excessive same-channel network expansion is absent). It appears that, at least in our system, kin recognition alone does not suffice to prompt full-fledged fraternal transitions in individuality.

In the absence of resource penalties for erroneous activation under control conditions, we also observed the evolution of larger same-channel groups. At update 3 million, most-common genotypes encoded a level-two same-channel cap C_2 of 484.0 cells with standard deviation of 123.5. Compared to the standard treatment, control runs exhibited larger mean level-two same-channel caps C_2 at time-points matched by absolute elapsed update count ($p < 0.0001$; two-tailed t test) and approximate elapsed cellular generations ($p < 0.0001$; two-tailed t test). Even at 20 million updates, when evolution had elapsed around six times as many cellular generations in the standard treatment compared to the control treatment at update 3 million, mean level-two same-channel caps C_2 reached only 262.9 with standard deviation 72.2 under the standard treatment. This is significantly smaller than mean C_2 under the control treatment at update 3 million ($p < 0.0001$; two-tailed t test). Figure 7 depicts the comparatively large same-channel level two groups present at the end of a control run. Table 1 summarizes most-common genotypes observed under the control treatment.

Conclusion

Using simple organisms that evolve parameters for a set of manually-designed strategies, we have demonstrated that DISHTINY selects for genotypes that exhibit high-level individuality. We observed a spectrum of first- and second-

level individuality among evolutionary outcomes. Specifically, we observed

1. reproductive division of labor among members of the same channel (i.e., individuals enveloped in a same-channel signaling network ceded reproduction to those at the periphery),
2. cooperation between members of the same channel (i.e., pooling of resource on same-channel signaling networks),
3. reproductive bottlenecking (i.e., groups of cells sharing a channel ID descend from a single originator of that channel ID), and
4. suppression of somatic mutation via apoptosis coincident with second-level individuality.

Competition experiments revealed that second-level individuals usually outcompete lower-level individuals. The magnitude of resource endowment for propagules was also correlated with second-level individuality.

Although shifts in individuality to level-one and level-two signaling networks were both observed, the question of whether these transitions were truly hierarchical in nature is debatable. That is, it is not clear whether level-one individuality was to some extent preserved in or necessary for the emergence of level-two individuality. Given the nature of the manually-designed strategies for resource-pooling and reproductive division of labor, level-two resource pooling and division of labor could readily leapfrog over level-one resource pooling and division of labor and, in many ways, seemed to completely supersede those level-one efforts.

We believe that this is a shortcoming of the manual design of behaviors for which simple cell-like organisms evolved parameters, not the DISHTINY platform itself. We have nevertheless demonstrated that DISHTINY ultimately selects for high-level individuality. We are eager to work with more sophisticated cell-like organisms capable of arbitrary computation via genetic programming in order to pursue more open-ended evolutionary experiments. We will also test the implications of relaxing current arbitrary restrictions that artificially promote transitions, such as the hierarchical nesting of same-channel signaling networks and the explicitly-defined signaling networks themselves, leaving these details to evolution to figure out. Further work will provide valuable insight into scientific questions relating to major evolutionary transitions such as the role of pre-existing phenotypic plasticity (Clune et al., 2007; Lalejini and Ofria, 2016), pre-existing environmental interactions, pre-existing reproductive division of labor, and how transitions relate to increases in organizational (Goldsby et al., 2012), structural, and functional (Goldsby et al., 2014) complexity.

We believe that such an approach also provides a unique opportunity to fundamentally advance Artificial life with respect to open-ended evolution. Fundamental to this goal is scale. The DISHTINY platform trivially scales to select for an arbitrary number of hierarchical levels of individuality (not just the two hierarchical levels explored in these experiments). Importantly, the platform is implemented in a decentralized manner and can comfortably scale as additional computing resources are provided. Parallel computing is widely exploited in evolutionary computing, where subpopulations are farmed out for periods of isolated evolution or single genotypes are farmed out for fitness evaluation (Lin et al., 1994; Real et al., 2017). DISHTINY presents a more fundamental parallelization potential: principled parallelization of the evolving individual phenotype at arbitrary scale (i.e., a high-level individual as a large collection of individual cells on the toroidal grid). Such parallelization will be key to realizing evolving computational systems with scale — and, perhaps, complexity — approaching biological systems.

Acknowledgements

Thanks to members of the DEVOLAB, in particular Michael J. Wiser, for feedback on statistical methods and Heather Goldsby for concept and editing feedback. This research was supported in part by NSF grants DEB-1655715 and DBI-0939454, and by Michigan State University through the computational resources provided by the Institute for Cyber-Enabled Research. This material is based upon work supported by the National Science Foundation Graduate Research Fellowship under Grant No. DGE-1424871. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

References

Ackley, D. H. (2016). Indefinite scalability for living computation. In *Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence*.

Ackley, D. H. and Cannon, D. C. (2011). Pursue robust indefinite scalability. In *HotOS*.

Banzhaf, W., Baumgaertner, B., Beslon, G., Doursat, R., Foster, J. A., McMullin, B., De Melo, V. V., Miconi, T., Spector, L., Stepney, S., et al. (2016). Defining and simulating open-ended novelty: requirements, guidelines, and challenges. *Theory in Biosciences*, 135(3):131–161.

Bedau, M. A. (2003). Artificial life: organization, adaptation and complexity from the bottom up. *Trends in cognitive sciences*, 7(11):505–512.

Bouchard, F. (2013). What is a symbiotic superindividual and how do you measure its fitness. *From groups to individuals: evolution and emerging individuality*, 243.

Clune, J., Ofria, C., and Pennock, R. T. (2007). Investigating the emergence of phenotypic plasticity in evolving digital organisms. In *European Conference on Artificial Life*, pages 74–83. Springer.

Ereshefsky, M. and Pedroso, M. (2015). Rethinking evolutionary individuality. *Proceedings of the National Academy of Sciences*, 112(33):10126–10132.

Goldsby, H. J., Dornhaus, A., Kerr, B., and Ofria, C. (2012). Task-switching costs promote the evolution of division of labor and shifts in individuality. *Proceedings of the National Academy of Sciences*, 109(34):13686–13691.

Goldsby, H. J., Knoester, D. B., and Ofria, C. (2010). Evolution of division of labor in genetically homogenous groups. In *Proceedings of the 12th annual conference on Genetic and evolutionary computation*, pages 135–142. ACM.

Goldsby, H. J., Knoester, D. B., Ofria, C., and Kerr, B. (2014). The evolutionary origin of somatic cells under the dirty work hypothesis. *PLoS biology*, 12(5):e1001858.

Goldsby, H. J., Young, R. L., Hofmann, H. A., and Hintze, A. (2017). Increasing the complexity of solutions produced by an evolutionary developmental system. In *Proceedings of the Genetic and Evolutionary Computation Conference Companion*, pages 57–58. ACM.

Lalejini, A. and Ofria, C. (2016). The evolutionary origins of phenotypic plasticity. In *Proceedings of the Artificial Life Conference*.

Lin, S.-C., Punch, W. F., and Goodman, E. D. (1994). Coarse-grain parallel genetic algorithms: Categorization and new approach. In *Parallel and Distributed Processing, 1994. Proceedings. Sixth IEEE Symposium on*, pages 28–37. IEEE.

Queller, D. C. (1997). Cooperators since life began. *The Quarterly Review of Biology*, 72(2):184–188.

Ratcliff, W. C., Denison, R. F., Borrello, M., and Travisano, M. (2012). Experimental evolution of multicellularity. *Proceedings of the National Academy of Sciences*, 109(5):1595–1600.

Ray, T. S. (1996). Evolving parallel computation. *Complex Systems*, 10:229–237.

Real, E., Moore, S., Selle, A., Saxena, S., Suematsu, Y. L., Tan, J., Le, Q. V., and Kurakin, A. (2017). Large-scale evolution of image classifiers. In Precup, D. and Teh, Y. W., editors, *Proceedings of the 34th International Conference on Machine Learning*, volume 70 of *Proceedings of Machine Learning Research*, pages 2902–2911, International Convention Centre, Sydney, Australia. PMLR.

Smith, J. and Szathmary, E. (1997). *The Major Transitions in Evolution*. OUP Oxford.

Taylor, T., Bedau, M., Channon, A., Ackley, D., Banzhaf, W., Beslon, G., Dolson, E., Froese, T., Hickinbotham, S., Ikegami, T., et al. (2016). Open-ended evolution: perspectives from the oee workshop in york. *Artificial life*, 22(3):408–423.

West, S. A., Fisher, R. M., Gardner, A., and Kiers, E. T. (2015). Major evolutionary transitions in individuality. *Proceedings of the National Academy of Sciences*, 112(33):10112–10119.