EXPLORING THE ECOLOGICAL AND EVOLUTIONARY CONSEQUENCES OF CLONAL AND AGGREGATIVE DEVELOMPENT DURING THE TRANSITION TO MULTICELLULARITY

A Dissertation
Presented to
The Academic Faculty

by

Jennifer Taryn Pentz

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of Biological Sciences

Georgia Institute of Technology
December 2019

EXPLORING THE ECOLOGICAL AND EVOLUTIONARY CONSEQUENCES OF CLONAL AND AGGREGATIVE DEVELOMPENT DURING THE TRANSITION TO MULTICELLULARITY

Approved by:

Dr. William C Ratcliff, Advisor
School of Biological Sciences
Georgia Institute of Technology

Dr. J Todd Streelman
School of Biological Sciences
Georgia Institute of Technology

Dr. Sam P Brown
School of Biological Sciences
Georgia Institute of Technology

Dr. Peter J Yunker
School of Physics
Georgia Institute of Technology

Dr. R Frank Rosenzweig
School of Biological Sciences
Georgia Institute of Technology

Date Approved: August 8, 2019

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my advisor, Dr. William C. Ratcliff for his continued support of my Ph.D. It has been an honor being his first Ph.D student. I greatly appreciate all his time, ideas, motivation, and overall enthusiasm for science to make my experience at Georgia Tech positive, productive, and stimulating. I could not have imagined have a better advisor for my Ph.D study. Secondly, I would also like to thank my Ph.D thesis committee: Dr. Sam Brown, Dr. Todd Streelman, Dr. Frank Rosenzweig, and Dr. Peter Yunker, for their insightful contributions and questions that encouraged me to broaden my perspectives on my research. The members of the Ratcliff lab, both past and current, have contributed immensely to my scientific career at Georgia Tech, continuously inspiring my curiosity and critical thinking. I thank them for their constant encouragement, support, and scientific collaborations. They have also been a great source of friendship. I would also like to thank other friends and that I have had the good fortune to know while at Georgia Tech. I am grateful for all the adventures that we experienced together. Finally, I would like to thank my family for all their love and encouragement. For my parents who have always supported and encouraged my love and pursuit of science. For my brother and sisters who were always there when I needed someone to talk to or to celebrate holidays with. Thank you.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SUMMARY

Multicellular organisms form groups in one of two basic ways: cells can 'stay together' due to incomplete separation following cellular division (clonal development), or cells can 'come together' via aggregation (aggregative development). Multicellularity has evolved multiple times via both routes, but all 'complex multicellularity' (*e.g.*, plants, animals, fungi) has only evolved in lineages that develop clonally. Evolutionary theory predicts that clonal development may be superior to aggregation because groups formed this way have little among-cell genetic conflict, thereby aligning the fitness interests of lower-level units (cells), increasing the potential for groups to undergo an 'evolutionary transition in individuality' (ETI). ETIs are characterized by a hierarchical shift in the level at which heritable variation in fitness is expressed (*e.g.*, from cells to the multicellular group). In this dissertation, I compare clonal and aggregative development in a simple yeast (*Saccharomyces cerevisiae*) model system. First, I performed a selection experiment using wild-isolated aggregative yeast (termed flocs) with daily selection for rapid sedimentation in liquid medium. Clonally-developing yeast (termed 'snowflake yeast') arose and displaced flocs, and invading snowflake yeast showed higher fitness than their floc counterparts. Next, I engineered snowflake and floc yeast from a common unicellular ancestor, so these two strains only differ in their mode of cluster development. In monoculture, floc yeast were superior to snowflake yeast, growing faster and forming larger clusters that settling more rapidly. Yet, in direct competition, snowflake yeast exploit flocs, becoming disproportionately represented within fast-settling groups. Modeling suggests that 'choosy' flocs that exclude snowflake yeast would have the highest fitness,

but such a strain would not be able to invade from rare. Finally, I performed a long-term evolution experiment to compare the dynamics of multicellular adaptation in floc and snowflake yeast by selecting for increasingly large cluster size, a multicellular trait. Our environment introduces two important life history traits that affect fitness, growth (cell level) and settling (cluster level), and evolved floc and snowflake yeast exhibited fitness gains in these two opposing traits, respectively. Furthermore, snowflake yeast were enriched with mutations that decrease fitness at the single-cell level, but may be beneficial at the cluster-level. Over evolutionary time, this could result in cells becoming interdependent parts of a new multicellular individual. Taken together, these results show that non-clonal cellular binding may be beneficial in environments favoring rapid multicellular group formation, but this paves the way for persistent evolutionary conflict. Conversely, simple clonal multicellular life cycles increase the efficacy of cluster-level adaptation relative to cell-level, which can potentiate an ETI and establish the emergent multicellular cluster as the new level of biological organization. These results highlight the critical role early multicellular life cycles play in driving – or constraining – this major evolutionary transition.

CHAPTER 1. INTRODUCTION

The transition from uni- to multicellularity was transformative for life on earth. It was an important innovation that allowed for increases in biological complexity[1-7]. The first evidence of this transition dates to 3.5 billion-years-ago, believed to be prokaryotic filamentous Cyanobacteria-like organisms[8-10]. Multicellular eukaryotes may have evolved about 1.2 billion years ago[11], with a major burst in diversification around 600-700 million years ago coinciding with increasing atmospheric and oceanic oxygen[6,9,12,13]. Despite the ecological importance of this major transition, the first steps in this transition remain difficult to study in natural systems, mainly because all known multicellular lineages are ancient and most early multicellular forms have been lost to extinction.

Multicellular organisms have independently evolved from unicellular ancestors at least 25 times in the history of life, the most recent of which was in the green and brown algae 200 mya[8]. This transition involved a fundamental shift in biological organization as individual cells, formally organisms in their own right, evolve to become integral parts of a larger organism[14,15]. This creates the potential for conflict among individual cells making up the multicellular organism, who must give up their own reproductive potential to become parts of a larger organism[6,14,16,17]. A big question in evolutionary biology is how the fitness interests of lower and higher-level units become aligned, and how selection acting at the lower level, increasing the potential for the evolution of cheaters, is kept at bay so the transition becomes established. Furthermore, after the transition is established, how do incipient multicellular organisms increase in biological complexity?

The transition from uni- to multicellular life constitutes a major transition in evolution characterized by an evolutionary transition in individuality (ETI)[6,14]. Evolutionary individuals are integrated and indivisible wholes. They are a unit of selection that must satisfy Darwin's properties of heritability and variation in fitness[18-20]. During an ETI, once-autonomous individuals become so integrated that the group becomes an individual at a new, higher level, forming a new entity with a single fitness and evolutionary fate. An ETI involves a shift in the level of selection from the lower-level units to the emerging higher-level unit. In the transition to multicellularity, this shift occurs from single cells as the unit of selection to the group as the unit of selection. During this process, individual cells must relinquish their own potential for reproduction and multiply in favor of the multicellular organism. Darwinian properties of heritable variation in fitness (*e.g.*, a multicellular life cycle[21,22]) must emerge at the level of the multicellular organism[23-26], enhancing the evolvability at the new level of biological organization[27]. The properties of multicellular group formation may be integral for the ability of the emergent multicellular organism to undergo an ETI[21].

Multicellular organisms may develop clonally or via aggregation. Most extant obligate multicellular organisms, or organisms that complete their life cycle as a multicellular organism, develop clonally, while a few groups of terrestrial or semi-terrestrial microorganisms, such as *Dictyostelium* species[28], are facultatively multicellular and exhibit aggregative multicellularity only under certain environmental conditions[2,8,29,30]. Evolutionary theory predicts that clonal development may be more common than aggregative development because it limits among-cell genetic variation, thus preventing genetic conflict from eroding multicellular complexity[27]. Clonal development

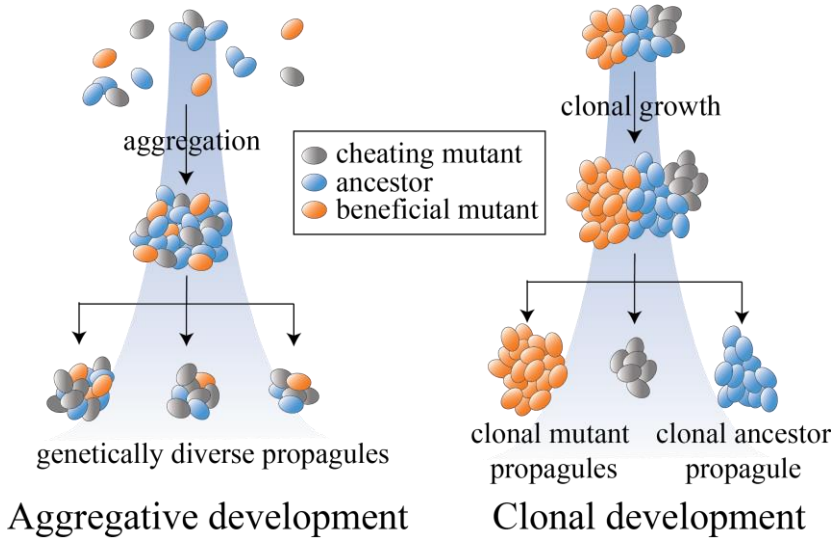

Figure 1 - Consequences of aggregative and clonal multicellular group formation. Clonal development limits among-cell genetic variation within multicellular clusters, preventing prevent the spread of cheating lineages within multicellular organisms. Additionally, clonal development increases the scope for among-cluster selection as genetic variation that arises gets partitioned among clusters in the next generation.

from a unicellular spore or zygote[31] limits genetic variation because the entire multicellular organism is derived from a single cell (Figure 1). Furthermore, any genetic variation that has arisen gets partitioned among offspring in the next generation, an important implication for the regulation of defectors. If a defector evolves within the multicellular organism, any benefits experienced by defecting will be curtailed when defectors make up the whole organism in the next generation. In contrast to clonal development, defectors that arise in aggregating multicellular organisms have a greater opportunity to exploit cooperators, reaping the benefits for their own personal gain, allowing them to spread and persist in the population (Figure 1)[8,17,30,32].

Perhaps an even greater benefit of clonal development is the potential for the emergence of multicellular evolvability. Clonal multicellular life cycles that frequently pass through a unicellular genetic bottleneck[31] result in genetic variation being partitioned

among offspring, increasing the scope for among-cluster selection relative to within-cluster selection[17,33,34] (Figure 1). This can lead to potentially high levels of multicellular heritability such that the new emergent multicellular entity is the fitness-maximizing agent and the focal point of selection. In aggregative life cycles, genetic variation arises through the association of genetically-distinct lineages, resulting in more within-cluster genetic variation, increasing the scope for within-cluster selection that may erode the transition to a new level of individuality[14,20,35]. These together may allow clonal multicellular groups to become more individuated at the multicellular level, allowing for further multicellular-level adaptions that can lead to more complex and integrated multicellular organisms. My dissertation aims to address these fundamental hypotheses using the snowflake yeast model system developed in Ratcliff *et al*. (2012).

Ratcliff *et al*. (2012) evolved multicellularity *de novo* in the lab by subjecting a unicellular population of *Saccharomyces cerevisiae*, or common baker's yeast, to daily selection for rapid sedimentation in liquid media. After 60 days of selection, all 10 replicate populations starting from a single clone evolved to form multicellular clusters, termed 'snowflake' yeast due to their distinctive branching pattern[36,37]. This branching pattern is a result of the way snowflake yeast form multicellular clusters. Clusters develop due to failed septum degradation after cytokinesis, resulting in daughter cells adhering to their mother cell post-division[36-38], resulting in clonal clusters.

In this system, once snowflake yeast evolve, they quickly become the unit of selection - whole clusters either settle quickly to survive the imposed selection, or get discarded. This lead to rapid, multicellular-level adaptation through several key innovations; increasing the average number of cells per cluster, the mass of individual cells,

and the rates of programmed cell death (PCD), or apoptosis[38]. Increased rates of apoptosis seems to have evolved due to a trade-off that emerged between the settling rate and growth rate of individual clusters[36,38]. Snowflake yeast exhibiting high rates of PCD have similarly high rates of apoptosis when cultured as single cells, suggesting that this trait co-evolves with large cluster size (See Chapter 1[39]).

However, *S. cerevisiae* also has the ability to form cellular aggregates termed 'flocs' consisting of thousands of cells, because of specific cell surface proteins forming lectin-like bonds with cell wall sugars of adjacent cells[40-44]. This is a commonly observed phenotype that has been utilized by the brewing industry due to its extremely rapid sedimentation in liquid medium after the fermentation process[44]. Despite the fast settling nature of flocculant genotypes, flocs were never observed in the settling selection experiments conducted by Ratcliff *et al.* 2012, raising the possibility that, consistent with evolutionary theory, clonal development is superior to aggregative development during the initial transition to multicellularity. My dissertation is presented in four chapters, summarized below:

Chapter 1: Apoptosis in snowflake yeast: novel trait, or side effect of toxic waste? Elevated rates of programmed cell death (apoptosis) is a key trait that evolved in large cluster-forming genotypes in Ratcliff *et al*. (2012). A 4-fold increase in the rate of apoptosis allows large clusters to make proportionally smaller propagules (less than 20% the size of their parent cluster) that are less diffusionally-limited and faster growing than their parent cluster[36]. These conditions allow apoptosis to be adaptive, increasing the growth rate of newly formed propagules enough to counteract reduced survival during settling selection[45,46]. Duran-Nebreda and Solé (2015) posited that elevated rates of apoptosis could

be a side effect of an accumulation of waste products within the largest clusters. Here we show that large snowflake yeast with elevated rates of PCD have similar levels when grown as single cells, showing that higher rates of PCD co-evolves with large cluster size. We also offer an alternative hypothesis for higher rates of PCD, showing that larger clusters have a greater fraction of old, senescent cells, which may result in clusters with increased PCD.

Chapter 2: Clonal development is evolutionary superior to aggregation in wild-collected *Saccharomyces cerevisiae*. Clonally-developing snowflake clusters arose in 10/10 replicate populations in Ratcliff *et al*. (2012), consistent with theory suggesting that clonal development is superior to aggregative development when transitioning from a unicellular to a multicellular way of life. To test this, I repeated the experiment conducted in Ratcliff *et al.* 2012 starting with five wild-isolated flocculant strains[47]. Surprisingly, snowflake yeast evolved and displaced their floc ancestors in 35 out of 40 replicate populations after 155 days of evolution. Early snowflake yeast were shown to be more fit than their floc counterparts isolated from the same population undergoing snowflake invasion, suggesting that once snowflake yeast arise in the population, they have an competitive advantage over floc yeast. This could be due to proximate effects of cluster formation, as snowflake yeast grow via deterministic rules and form larger clusters more consistently. Alternatively, snowflake yeast may possess an ultimate advantage over floc yeast, and may be more capable of multicellular-level adaptation, essential for further evolution of multicellular complexity[37].

Chapter 3: Ecological advantages and evolutionary limitations of aggregative multicellular development. The environment that snowflake yeast evolved in is characterized by periods

of growth (24 hours growing at 30°C) and selection for settling. Flocculating yeast are initially superior to snowflake yeast in this system, growing 35% faster than snowflake yeast and settling 2.5 faster, suggesting this environment favors life cycles characterized by rapid aggregation. But, when competed directly, snowflake yeast outcompete floc because they can become embedded within large chimeric aggregates. Furthermore, snowflake yeast are disproportionately represented in the aggregates, allowing snowflake yeast to drive floc yeast to extinction when in competition. Finally, using a mathematical model, we show that if a 'choosy' floc evolved (floc genotype that has a self-recognition mechanism), it would not be able to invade from rare. Permissive binding increases the opportunities for cell-cell adhesion, highlighting a trade-off of aggregative development; selection for rapid group formation favors permissive binding, but the resulting within-group genetic diversity paves the way for persistent evolutionary conflict.

Chapter 4: Developmental mode and the emergence of multicellular individuality. Multicellularity has evolved multiple times via both clonal and aggregative development[48], but 'complex multicellularity' (*e.g.*, plants, animals, fungi) has only evolved in lineages that have done so clonally. There are many benefits to clonal development that may facilitate the major evolutionary transition: clonal development may facilitate the emergence of multicellular evolvability by aligning the fitness interests of cells, thereby suppressing cheating, and by increasing the heritability of multicellular traits. In this chapter, I test the hypothesis that clonal development may facilitate the emergence of multicellular-level individuality by subjecting synthetically created snowflake (Δ*ace2*) and floc (*GAL1p-FLO1*) to selection for fast settling in liquid medium, an environment where multicellularity is adaptive. Snowflake and floc yeast responded to selection by increasing

their settling rate, but fitness increases in evolved isolates was achieved in two distinct life history traits. Snowflake yeast increased fitness during settling (multicellular-level trait), while floc yeast increased fitness during growth (unicellular-level trait). Furthermore, snowflake yeast had a larger proportion of fixed mutations that decrease fitness during growth but may increase cluster-level fitness. Our results show that, even in simple multicellular yeast, clonal development facilitates the emergence of higher-level individuality, crucial for stabilizing the first steps in this major evolutionary transition.

CHAPTER 2. APOPTOSIS IN SNOWFLAKE YEAST: NOVEL TRAIT, OR SIDE EFFECT OF TOXIC WASTE

2.1 Abstract

Recent experiments evolving *de novo* multicellularity in yeast have found that large-cluster forming genotypes also exhibit higher rates of programmed cell death (apoptosis). This was previously interpreted as the evolution of a simple form of cellular division of labor: apoptosis results in the scission of cell-cell connections, allowing snowflake yeast to produce proportionally smaller, faster-growing propagules. Through spatial simulations, Duran-Nebreda and Solé (2015) develop the novel null hypothesis that apoptosis is not an adaptation, *per se*, but is instead caused by the accumulation of toxic metabolites in large clusters. Here we test this hypothesis by synthetically creating unicellular derivatives of snowflake yeast through functional complementation with the ancestral *ACE2* allele. We find that multicellular snowflake yeast with elevated apoptosis exhibit a similar rate of apoptosis when cultured as single cells. We also show that larger snowflake yeast clusters tend to contain a greater fraction of older, senescent cells, which may explain why larger clusters of a given genotype are more apoptotic. Our results show that apoptosis is not caused by side effects of spatial structure, such as starvation or waste product accumulation, and are consistent with the hypothesis that elevated apoptosis is a trait which co-evolves with large cluster size.

2.2 Introduction

Duran-Nebreda and Solé (2015) recently reported results from a computational model examining the evolution of multicellularity in experimentally-evolved 'snowflake' yeast. In this paper, they offer an explanation for a pair of experimental results from Ratcliff *et al*. (2012): First, some yeast cells within clonal clusters undergo programmed cell death (apoptosis). Early snowflake yeast are small and exhibit little cell death, but sustained directional selection for faster settling resulted in the evolution of both greater cluster size and proportionally higher rates of apoptosis[36]. Second, the rates of apoptosis observed within snowflake yeast clusters increased nonlinearly with cluster size. Duran-Nebreda and Solé are able to recapitulate these observed dynamics in a simulation model in which apoptosis is induced by the build-up of waste products, such as acetate. All else being equal, they hypothesize that larger clusters experience greater diffusional gradients, causing internal cells to be exposed to sufficient quantities of waste products to induce cellular suicide. Elevated apoptosis in large clusters is thus, in their view, caused by a pre-existing genotype by environment interaction that emerges once large cluster size evolves[49].

This intriguing hypothesis contrasts with the hypothesis put forward by Ratcliff *et al*. (2012), which posits that apoptosis in large-bodied snowflake yeast is an evolved trait independent of cluster size. Over the course of the experiment, elevated rates of apoptosis evolved in multiple large cluster-forming genotypes from separate replicate populations, but remained low in lineages that formed small clusters. Large cluster size and apoptosis were shown to segregate independently after selfing sex, suggesting that cluster size and apoptosis are inherited independently. Ratcliff *et al*. (2012) hypothesized that the

congruence of these two traits in experimental populations was likely the result of coevolution, not a direct effect of cluster size on apoptosis[36].

Why might cellular suicide be beneficial to snowflake yeast that form large clusters? Ratcliff *et al*. (2012) hypothesized that apoptosis helps mitigate the slower growth rates of large, diffusionally-limited clusters by allowing them to produce proportionally smaller propagules. Apoptosis severs cell-cell connections[36], resulting in branch scission and the production of a propagule[50]. Production of proportionally smaller propagules is adaptive as long as their faster growth more than compensates for their reduced survival during settling selection at the end of the 24 h culture cycle. Indeed, evolutionary simulations show that this criteria is often met, as elevated apoptosis readily evolves *in silico*[46,51] . The distinction between these hypotheses is of crucial importance for properly interpreting Ratcliff *et al*. (2012)'s experimental results: if elevated apoptosis is an adaptation to large size, not its side-effect, it demonstrates that selection can readily favor the evolution of altruistic cellular traits in simple clusters of cells. Such selection underpins the evolution of cellular division of labor[32,36,52-57], and is thus of fundamental importance for understanding the evolution of increased multicellular complexity. In this paper, we perform a direct experimental test of Duran-Nebreda & Solé, 2015's hypothesis.

Early (isolated after 7 days of settling selection) and late (from 60 days of selection) snowflake yeast were obtained from the experiment conducted in Ratcliff et al. (2012). These yeast gained the ability to form clusters through the inactivation of the trans-acting transcription factor *ACE2*[50], which regulates the expression of a number of enzymes involved in mother-daughter cell separation after mitosis[58-60]. As a result, daughter cells remain attached to mother cells, creating the branched 'snowflake' phenotype. The 7-day

yeast form relatively small clusters with little apoptosis, while the 60-day yeast form large clusters with elevated apoptosis (Ratcliff *et al*., 2012). To test Duran-Nebreda and Solé's hypothesis that stronger diffusional gradients in larger clusters is the proximate cause of increased apoptosis, we conducted an experiment to examine apoptosis in the absence of any multicellular structure. We created unicellular derivatives of the 7- and 60-day snowflake yeast by complementing the evolved *ace2* strains with a functional copy of the ancestral *ACE2* allele. The resulting unicellular strains differ only at a single locus (*ACE2*) from their snowflake ancestors, but do not experience an accumulation of waste products due to multicellular structure. 60-day snowflake yeast retained their elevated apoptosis even when grown as single cells, demonstrating that apoptosis is not simply a side-effect of diffusional limitation in large clusters. To explain the nonlinear increase in apoptosis and cluster size observed *among clonal clusters* of the 60-day snowflake yeast strain (Ratcliff et al., 2012, Figure 5A), we simulate snowflake yeast growth and reproduction, and show that larger clusters tend to develop a core of old, potentially apoptosis-prone cells, while small propagules are typically composed of young and healthy cells. We find that older yeast are disproportionately apoptotic, supporting our hypothesis that large clusters contain more apoptotic cells because they contain more aged, senescent cells, not due to the accumulation of toxic metabolites.

2.3 Methods

2.3.1 Creating unicellular yeast from snowflake genotypes

We obtained unicellular derivatives of isolates from early (7 days of settling selection) and late (60 days of settling selection) snowflake yeast from the experiment

conducted in Ratcliff *et al*. (2015). Briefly, we replaced a single copy of the naturally-evolved nonfunctional *ace2* in each isolate with the ancestral *ACE2* allele fused with the antibiotic resistance gene *KANMX4*[50] using the LiAc/SS-DNA/PEG method of yeast transformation[61]. Transformants were then plated on solid Yeast Peptone Dextrose medium (YPD; per liter: 20 g dextrose, 20 g peptone, 10 g yeast extract and 15 g agar) with 200 mg/L of the antibiotic G418. For each transformant, the insertion location of the transformation sequence was confirmed by PCR.

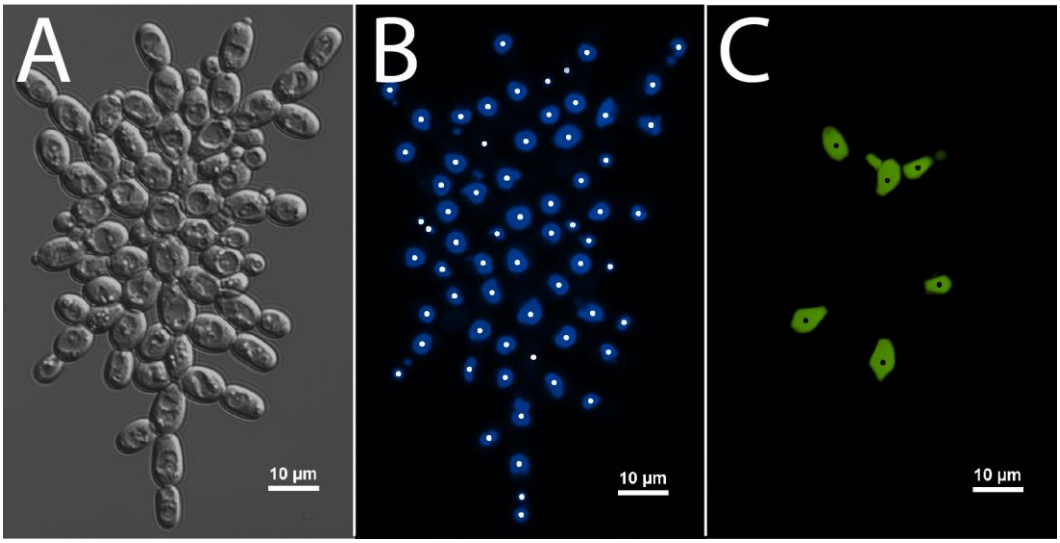

Figure 2 - Measuring apoptosis in snowflake yeast. Pictured is the same snowflake cluster imaged with A) Differential Interference Contrast microscopy, B) the blue fluorescent vacuole stain CellTracker™ Blue CMAC, and C) the green fluorescent apoptosis stain dihydrorhodamine 123. We obtained counts of all cells by counting the number of vacuoles (white dots marking cells in B) and apoptotic cells (black dots marking cells in C) by using the spot detection algorithm in NIS-Elements v. 4.30. The same method was used to measure apoptosis in unicellular strains.

2.3.2 Measuring apoptosis in yeast

We measured the rate of apoptosis in early and late isolates of multicellular snowflake yeast as well as their respective unicellular derivatives, along with unicellular

controls evolved without selection for faster settling. These latter strains were grown under the same conditions as snowflake yeast (daily 1:100 dilution into 10 mL YPD, 250 RPM shaking at 30°C), but did not experience any settling selection. To measure apoptosis, we developed a novel method to count the fraction of apoptotic cells directly. Uni- and multicellular yeast were dual-labeled with the green-fluorescent reactive oxygen species stain dihydrorhodamine 123 (DHR 123; 2.5 mg/mL in 95% ethanol) as well as a blue-fluorescent vacuole stain, CellTracker^{TM} Blue CMAC, allowing direct measurement of each (Figure 2B and C). Before measuring apoptosis, all strains were struck out on solid YPD from -80ºC stocks, then grown in 10 mL liquid YPD for 24 h at 30°C, 250 RPM shaking. Ten replicate populations of each yeast strain were then initiated with a 1:100 dilution into 10 mL YPD then cultured for 12 h. After 12 h of growth, yeast were stained with a 1% solution of DHR 123 [62] and a 0.1% solution of CellTracker^{TM} Blue CMAC, and incubated in the dark for two hours. Yeast were then double-washed with carbon free yeast nitrogen base (YNB) buffer (Sigma Aldrich, 6.7 g/L) and unicellular strains diluted 10-fold, while multicellular strains were diluted two-fold. 5 µL of this culture was placed between a slide and a 25 × 25 mm coverslip. Placing such a small volume of media on the slide tends to flatten multicellular clusters (Figure 2A), helping keep all cells in focus (Figure 2B and C). Large-field mosaic images were collected for each isolate by combining 36 separate images (each collected at 100 × magnification) using a Nikon Eclipse Ti inverted microscope with a computer-controlled Prior stage, resulting in a 10185 × 7665 pixel composite image. This technique results in large sample sizes, is sensitive to dim fluorescence, and minimizes sample bias arising from single images in which yeast touching the edge of the field of view are discarded (this disproportionately affects large

clusters). Counts of all cells (CellTracker™ Blue CMAC; white dot labels in 1B) and apoptotic cells (DHR 123; black dot labels in Figure 2C) were acquired using the spot detection algorithm in the NIS-Elements v. 4.30 software package.

2.3.3 Measuring the age of apoptotic yeast

We measured the age of non-apoptotic and apoptotic cells by dual-labeling unicellular derivatives of early and late snowflake yeast with DHR 123 (to determine which cells are undergoing apoptosis) and the chitin-binding stain calcoflour (1 mg/mL in water). Calcoflour labels "bud scars" left on the mother cell after mitotic division, allowing us to determine the replicative age of the cell[63]. Unicellular derivatives of both 7- and 60-day snowflake yeast were grown and stained with DHR 123 as described above. After the 2 h incubation to label apoptotic cells, yeast were stained with a 1% solution of calcoflour and incubated in the dark for an additional 5 minutes. Yeast were then double-washed with YNB and diluted 10-fold. 5 µL was then placed between a slide and a 25 × 25 mm coverslip and imaged at 400 × magnification. 6 µm thick z-stack images of 100 apoptotic and 100 non-apoptotic cells were made by imaging 6 slices with a 1 µm step size. Bud scar counting was performed on the maximum intensity projections. To improve the efficiency of our counts, we processed these images using a custom Python script that subtracted background fluorescence and increased the contrast of bud scar edges.

2.3.4 Graph-based simulation

Cellular growth and scission were stochastically simulated using the Gillespie algorithm[64]. Given a current state of the system, represented by a population of graphs, the algorithm proceeds by randomly sampling the time the next event occurs, Δt, and by

randomly sampling which event occurs (cell reproduction via node growth or propagule production via edge scission). Because we do not have a precise understanding of how snowflake yeast clusters reproduce, we modeled branch scission with two separate approaches. With weighted reproduction, edge scission depends on the amount of strain associated with that node, or its weight w_i, calculated as the total number of downstream path-connected cells (*e.g.*, total number of genealogical descendants connected to each cell; see edge values in Figure 4A). This weighting parameter increases linearly with the number of attached cells, so scission tends to occur towards the center of the cluster. In contrast, with unweighted separation, scission occurs at a constant rate throughout the cluster, so edges tend to separate close to the cluster periphery where most nodes are located. Cell growth events occur at a constant rate, g_j=1, and edge removal events occur either at a constant rate or at a rate proportional to the current weight, $r(w_i)$. The time the next event occurs is obtained by sampling from an exponential probability density function (PDF), $PDF(\Delta t) = \lambda e^{-\lambda \Delta t}$, parametrized by the total sum of rates across nodes and edges in the population, where:

$$\lambda = \sum_{i=edges} r(w_i) + \sum_{j=nodes} g_j \tag{1}$$

The probability of choosing an event is the ratio of its rate to the total rate. In the unweighted model, edges were removed at a constant rate, $r_i = \boldsymbol{\phi}$, while in the weighted model edges were removed in a weight (w_i)-dependent rate, $r_i(w_i) = \boldsymbol{\rho}\, w_i$.

2.4 Results and Discussion

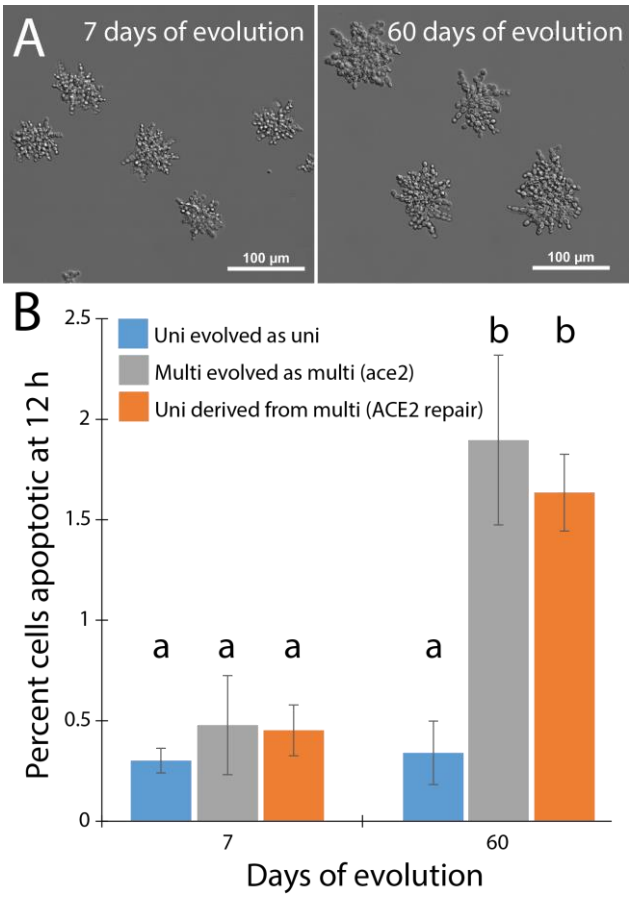

Figure 3 - Elevated apoptosis is not a consequence of cluster spatial structure. A) Snowflake yeast from the same population at different time points. Cluster size increased between 7 and 60 days. B) 60-day snowflake yeast show a four-fold increase in rate of apoptosis over 7-day snowflake yeast (grey bars; 1.90% vs. 0.48% apoptotic, respectively). Unicellular derivatives (transformed with a functional *ACE2* allele) maintained this elevated rate of apoptosis (orange bars; 1.61%). A control line in which we evolved unicellular yeast without selection for faster settling (it stayed unicellular) maintained its low ancestral rate of apoptosis (0.34% vs. 0.30% for 60- and 7-day isolates, respectively). Shown are means ± the standard deviation of 10 biological replicates.

Growth form (unicellular vs. multicellular) did not directly affect apoptosis. Early (7 day) snowflake yeast exhibited similar low rates of apoptosis (0.48% and 0.45%) when grown as snowflakes or unicells, respectively (Figure 3B). The 60-day snowflake yeast strain evolved a four-fold increase in apoptosis (1.90%), an elevated rate that was maintained in the unicellular form (1.61%). As a control, we examined apoptosis in a unicellular yeast

transferred for either 7 or 60 days without settling selection, and which remained unicellular. This control maintained its low ancestral rate of apoptosis (0.30% vs. 0.34% for day 7 and day 60 strain, respectively). Differences in apoptosis were assessed via a full factorial two-way ANOVA, with days of evolution and treatment as independent variables and percent apoptosis as the dependent variable ($F_{5,54}$=92.56, p<0.0001, r^2=0.9). Between-group differences were assessed with Tukey's HSD; significance at α=0.05 is denoted by different letters (a or b) in Figure 3B. Overall, only 60-day snowflake yeast possessed greater rates of apoptosis, and this did not depend on whether they were grown as snowflake clusters or as single cells. These results demonstrate that the elevated rates of apoptosis observed in late snowflake yeast isolates cannot be attributed to diffusional gradients imposed by multicellularity, and supports the hypothesis that apoptosis is a trait co-evolving with increased cluster size.

If apoptosis is not caused by waste product accumulation, then how do we explain the nonlinear increase in apoptosis seen among larger clusters of a single genotype (Figure 5a of Ratcliff *et al*. 2012)? We hypothesize that elevated apoptosis is due to the accumulation of old, senescent cells in large clusters, which itself is due to the geometry of snowflake yeast and their mode of reproduction. Snowflake yeast clusters reproduce by branch scission. Liberated branches act as propagules, which grow in size until they are large enough to produce their own offspring[36]. One effect of this mode of reproduction is that propagules tend to contain many young, recently-produced cells. In contrast, the cells in the interior of the cluster are less likely to be released by branch scission, and thus may stay in the cluster center for considerably longer than those on the periphery. We modeled this effect quantitatively. The structure of a snowflake yeast cluster can be represented as

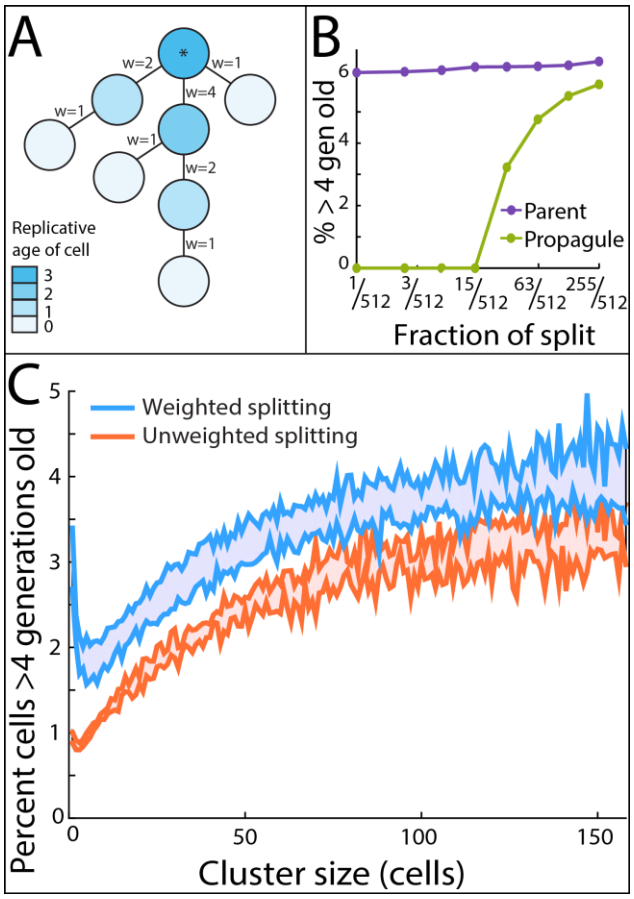

Figure 4 - Larger clusters contain a greater fraction of old cells. A) A snowflake yeast cluster represented by a graph of nodes and edges. Nodes represent component cells in the cluster and edges represent physical connections between a pair of cells. The root cell (denoted by *) is the oldest cell in the cluster, with younger cells being closer to the periphery. Edge weights *w* depend on the number of attached cells. B) Larger propagules tend to contain a higher proportion of old cells. This is due to the geometric structure of snowflake yeast: propagules that separate from the parent cluster contain mostly young cells, skewing the age distribution of old clusters and increasing the proportion of old cells they contain. C) We simulated a population of snowflake yeast growing for nine divisions with two models: in the first the per-edge probability of scission depends linearly on the number of attached descendent cells (weighted splitting; see *w* values in A), while in the second it is position independent (unweighted splitting). This figure shows only common size classes, (*i.e.*, those that occurred in at least 10% or the simulations). For both weighted and unweighted splitting, we plot the minimum and maximum percentage of cells older than 4 generations obtained in all simulations (10^3 runs for each parameter, ϕ and ρ).

a graph with a tree-like structure, in which cells are nodes and cell-cell connections are edges (Figure 4A)[46,50]. Clusters reproduce when an edge is severed. We considered 512-

cell snowflake yeast clusters (9 generations of growth from a single cell), and calculated the fraction of cells older than four generations in both propagules and parent clusters with varying reproductive asymmetry (Figure 4B). Parent clusters always contained a larger fraction of old cells; an effect that was enhanced by greater reproductive asymmetry. Apoptosis is a well-known consequence of aging in yeast[65,66], so the difference in mean age of cells in propagules relative to large, mature clusters may explain the higher rates of apoptosis observed in the latter.

We next simulated the growth and reproduction of snowflake yeast over a period of time corresponding to nine cellular generations and examined the relationship between cluster size and cellular age. We simulated our dynamic model across a range of values for ϕ [.04 .0425 .045 .0475 .05] and ρ [.01 .0125 .015 .0175 .02], with 10^3 simulations per parameter value. Our results confirm a positive relationship between cluster size and proportion of older cells (Figure 4C); these results were similar with both the weight dependent and independent models of edge scission. Indeed, the ~2% increase in cells over four generations old between small and large clusters (Figure 4C) is similar to the ~2% difference in apoptosis seen in small vs. large snowflake yeast clusters of the 60-day strain (Ratcliff et al., 2012, Figure 5A). The increase in old-aged single cells in the weighted model (Figure 4C) is due to a statistical anomaly: edge scission is rare in large clusters for peripheral cells with w=1, as they have low weight, so many singletons were the result of a split in an early round of growth when clusters contained just a few cells. Because cellular reproduction is stochastic, many singletons at the end of the simulation are old cells which failed to reproduce.

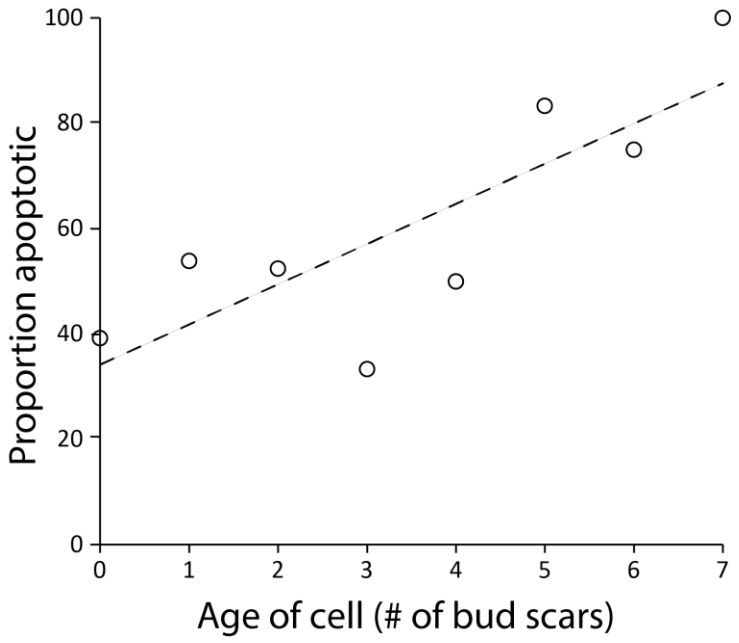

Figure 5 - Old-aged cells are disproportionately apoptotic. We measured the age of 100 apoptotic and 100 non-apoptotic cells. Old cells were more likely to be apoptotic than young cells (y = 7.6x + 34.2, p=0.02, r^2=0.61, linear regression).

To determine if apoptosis is an age-dependent process in our 60-day snowflake yeast, we measured the replicative age of 100 apoptotic and 100 non-apoptotic cells by counting bud scars. Older cells were increasingly apoptotic (Figure 5; p=0.0217, r^2=0.61, linear regression). To summarize, we found that snowflake yeast that evolved elevated rates of apoptosis after 60 days of selection continue to display this trait even when genetically altered to grow as single cells, disproving the hypothesis that side effects of cluster spatial structure (*e.g.*, waste product accumulation or starvation) directly cause this phenotype in our well-characterized lineage of snowflake yeast. We hypothesize that the positive within-genotype relationship between cluster size and fraction of cells undergoing apoptosis reported in Ratcliff *et al*. (2012) is due instead to differences in the age of the cells within these clusters. Larger clusters tend to contain a higher percentage of old cells than smaller clusters, and old cells are more likely to undergo apoptosis.

While our experimental results contradict Duran-Nebreda and Solé's (2015) main conclusions, we'd like to note an important limitation of our paper: we do not (and indeed cannot) claim that elevated apoptosis *never* evolves as a consequence of waste product accumulation in large clusters. Indeed, this mechanism may be important in independently evolved lineages of snowflake yeast, and should be considered a viable hypothesis to be tested in future experiments examining novel experimental lineages. More generally, nascent multicellular organisms are expected to be heavily affected by physical affects arising from their spatial structure [46,50,67,68]. Within-cluster gradients of nutrients, waste products, and even previously-evolved signaling molecules (*e.g.*, quorum sensing molecules) can readily cause clonal cells within clusters to differentiate phenotypically[69-73], providing raw material upon which multicellular-level selection may act. Physical models like Duran-Nebreda and Solé (2015) are thus an especially valuable tool in understanding how and why simple clusters of cells may evolve increased multicellular complexity.

2.5 Author contributions

J.P. and W.C.R. designed the experiments and analyzed the data. B.T. wrote the model. W.C.R., J.P., and B.T. wrote the paper. The authors have no conflicts of interest to declare.

2.6 Acknowledgements

The authors would like to thank Joshua Weitz for helpful feedback on the manuscript. This work was supported by NASA Exobiology Award #NNX15AR33G.

CHAPTER 3. CLONAL DEVELOPMENT IS EVOLUTIONARYILY SUPERIOR TO AGGREGATION IN WILD-COLLECTED *SACCHAROMYCES CEREVISIAE*

3.1 Abstract

The vast majority of multicellular organisms develop clonally via 'staying together' after mitotic reproduction. Evolutionary theory predicts that cells staying together provides several key advantages over multicellular construction via cells 'coming together', but little empirical work has directly compared these developmental modes. In our previous work evolving multicellularity *de novo* in the yeast *Saccharomyces cerevisiae,* cells evolved to form clonal clusters exclusively through post-division adhesion of mitotically-produced cells, a result that reflects the strong bias towards clonal development in extant multicellular taxa. An equally parsimonious explanation, however, is that cluster development through incomplete cell separation is simply easier to evolve than the production of the adhesive compounds required for aggregation. To disentangle these hypotheses we repeated the experiment of Ratcliff *et al*. (2012), selecting for rapid settling through liquid medium. Instead of using a unicellular ancestor, however, we started our experiment with five wild strains of yeast capable of aggregating into clusters via flocculation. Clonally-developing 'snowflake' yeast evolved and invaded 36/40 experimental populations within 155 transfers, and competition experiments revealed that invading snowflake yeast were substantially more fit than their floc contemporaries. These results support the hypothesis that clonal development is evolutionarily superior to

aggregation, and demonstrate that ‘snowflake’ yeast can readily evolve in diverse, wild-collected yeast strains.

3.2 Introduction

The evolution of multicellular organisms from unicellular ancestors is considered a ‘major transition’ in the history of life on earth. As such, it was one of a few innovations that allowed for the evolution of increased complexity[74]. The transition from uni- to multicellularity has occurred at least 25 times in separate lineages[75]. This transition involved a fundamental shift in biological organization, as individual cells, formally organisms in their own right, evolve to become integral parts of a new, higher-level organism. A key step in the evolution of multicellularity was a transition to larger size, which necessitated the formation of simple cellular clusters[2,76-78]. Selection must then shift from the single cell level to the cluster level, resulting in clusters that are themselves Darwinian individuals[74,79-81]. Construction of an organism from lower-level units that are fully capable of Darwinian evolution is potentially problematic, however, as it may result in evolutionary conflict between the lower- and higher-level units (*i.e.*, cells and multicellular organisms). The potential for conflict is especially strong when selection for multicellular-level functionality results in reproductive altruism among cells (*e.g.*, differentiated somatic cells that are at a reproductive dead-end)[52,81,82]. This raises a fundamental question in evolutionary biology: How do the fitness interests of lower and higher-level units become aligned, limiting the negative consequences of evolutionary conflict? Similarly, how are lower-level units (cells) de-Darwinized, limiting the potential for among-cell selection to undermine multicellular-level selection?

Not all multicellular organisms are constructed in the same manner. There are two basic modes of body formation: potentially unrelated cells either 'come together' to form a body, or cells 'stay together' after reproduction, which results in clonal development if the life cycle includes a genetic bottleneck[2,30]. Multicellular organisms have evolved via both routes. For example, the myxobacteria are a group of soil-dwelling bacteria that exhibit a social foraging behavior, coming together to form swarms that increase their feeding efficiency[83,84]. However, the vast majority of independent transitions to multicellularity have occurred via staying together[2], suggesting that this is the superior mode of multicellular development.

Evolutionary theory predicts that multicellular development via cells staying together should provide several key advantages over cells coming together[30,75]. This mode of development limits among-cell genetic variation, especially if the life cycle includes a genetic bottleneck[31]. Limiting genetic variation among lower-level units has several benefits: 1) It eliminates the potential for evolutionary conflict, since there is little standing genetic diversity within a higher-level unit for selection to act on[17,31,81,85]. 2) High among-cell genetic relatedness favors the evolution of traits that increase the fitness of the cluster, even if this reduces the fitness of individual cells. This facilitates the evolution of cellular division of labor[85-87]. 3) Any genetic variation that arises due to mutation gets partitioned among multicellular offspring, allowing selection to act on the multicellular-level phenotypic effects of *de novo* mutation. This also allows selection to act against mildly deleterious alleles[75] that would be hard to select against in chimeric organisms.

Despite the clear predictions of evolutionary theory, it has been difficult to test the hypothesis that multicellular development via staying together should be superior to

coming together. This is largely due to a lack of model systems in which both modes of development can be induced. The yeast *Saccharomyces cerevisiae* can form clusters either by incomplete cell separation after mitosis, producing 'snowflake yeast' , or by coming together through adhesive glycoprotein production, a process known as flocculation[88]. In prior experiments, Ratcliff *et al*. (2012) found that snowflake yeast evolved in 10/10 replicate populations selected for faster settling. This raises the possibility that staying together is superior to coming together in this yeast model system, but it is also possible that this trait simply evolves more readily than flocculation. Here we repeat the experiment of Ratcliff *et al.*, 2012, but rather than starting with a unicellular yeast, we start with wild-collected highly flocculent strains. Our experiment thus 'stacks the deck' in favor of floc, as all yeast start out with the ability to form a cluster via aggregation. If staying together is adaptive, it will need to evolve *de novo* and invade a population of aggregative yeast.

3.3 Methods

3.3.1 Strains, culture conditions, and selection regime

We used five field-isolated flocculating unicellular *S. cerevisiae:* strains YJM450, YJM454, YPS1000-1, YPS1009-2, and M5-2. Diploids were generated by streaking a and α mating type haploids on YPD agar plates (per liter: 20 g dextrose, 20 g peptone, 10 g yeast extract, 15 g agar). Single strains were isolated through three rounds of single-colony bottlenecking, and diploidy confirmed by tetrad formation after 4 d of shaking incubation at 30°C in sporulation media (per liter: 20 g potassium acetate, 2.2 g yeast extract, 870 mg synthetic amino acid mix, 0.5 g glucose). A single clone of each strain was then used to start eight replicate populations (40 populations total). Yeast were grown in 10 mL liquid

YPD in 25 × 150 mm tubes for 24 h at 30°C, with 250 rpm shaking. Every 24 h, the populations were subjected to settling selection by centrifuging at 100 g for 10 seconds (selection protocol described fully in Ratcliff *et al*., 2012). Whole populations were cryogenically preserved every 7 d at -80°C.

3.3.2 Constructing fluorescently labeled yeast

Single-strain isolates were obtained from a single population of strain M5-2 (replicate 2) after 60 and 120 days. Each strain was transformed to express either the green fluorescent protein yeGFP or red fluorescent protein dTomato constitutively under the TEF2 promoter, using the LiAc/SS-DNA/PEG method of transformation[61]. Transformed strains were then imaged with a SPOT Flex 64 MP camera on an Olympus IX 70 microscope at 10 and 20 X magnification.

3.3.3 Measuring yeast phenotypes

The predominant phenotype (e.g., snowflake or floc) was determined microscopically after 7, 28, 60, 91, 120, and 155 days of evolution. After 24 h of growth in liquid YPD (30°C, 250 rpm shaking), 10 µL of culture was placed on a slide under a 25 × 25 mm cover slip and imaged at 10x and 40x magnification. Snowflake yeast develop by post-division adhesion of cells, as opposed to floc's adhesive aggregation, allowing us to differentiate phenotype via cluster morphology. The numerically dominant phenotype was scored at each time point.

3.3.4 Relative fitness of early snowflake yeast vs. floc yeast

To measure the relative fitness of early snowflake yeast, we isolated a pair of snowflake and floc yeast strains from the 10 populations that contained both phenotypes at transfer 60 (4 from M5-2, 4 from YPS1009-2, 1 from YJM454, and 1 from YPS1000-1, respectively). All strains were grown in liquid YPD for 24 h, then snowflake/floc pairs were diluted 1:200 into 10 mL liquid YPD. These 10 strain pairs were competed over 5 rounds of selection for both growth and settling (100 x *g* for 10 s). For each pair, three replicate competition tubes were established. Snowflake and floc colonies have a distinct morphology (i.e., smooth colonies for floc and rough colonies for snowflake; see Figure 2a&b). We therefore used plate counts to determine fitness. Each competition tube was plated out at 1:10,000 dilution onto YPD plates after 24 h of growth (pre settling selection), and again after five rounds of growth and settling selection. Colonies were counted on digital images (in ImageJ) made from each plate after two days of growth at 30°C, taken with a Pentax K10D DSLR with a SMC Pentax-D FA 1:2.8 macro lens. Relative fitness was calculated using the ratio of Malthusian growth parameters[89].

3.4 Results and Discussion

Prior experiments selecting for rapid settling of yeast through liquid medium[36,58,90] found that clusters rapidly evolved, but these developed strictly via cells 'staying together' after reproduction, producing the 'snowflake' phenotype. It is unclear, however, if this is because snowflake yeast are actually superior to floc yeast that develop via aggregation, or if snowflake yeast simply arise more readily via mutation. By repeating our prior experiment with five wild-collected highly flocculent *Saccharomyces cerevisiae* strains, we sought to determine if snowflake yeast could invade populations of aggregative yeast.

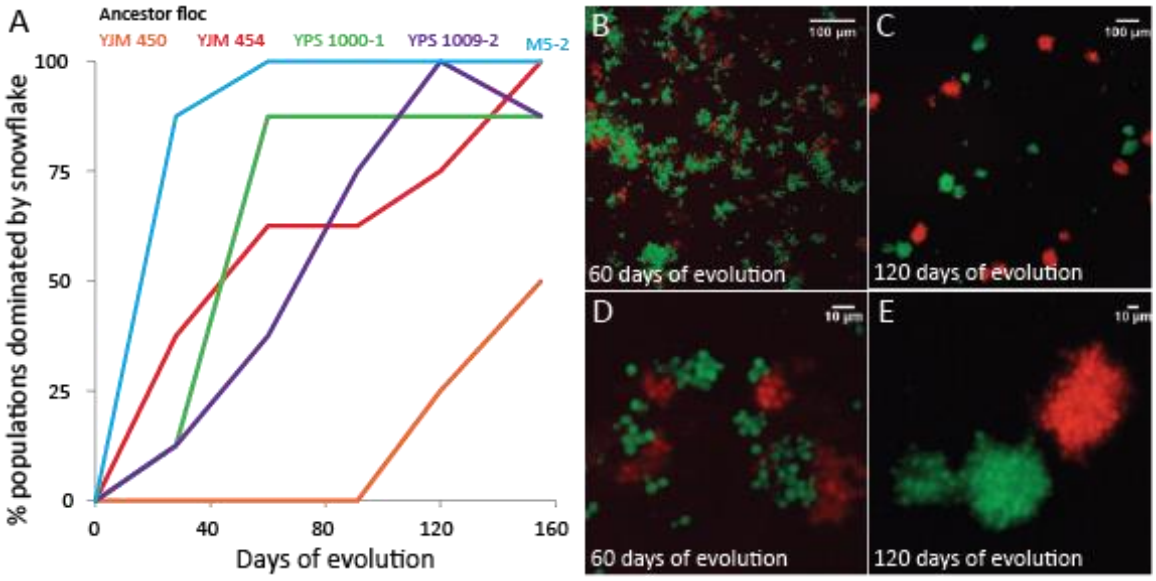

Figure 6 - Snowflake yeast evolve and displace floc in all five genetic backgrounds. A) Shown are the percentage of the eight replicate populations of each starting strain that are dominated by snowflake yeast, as determined by cluster morphology. Snowflake yeast develop by cells staying together after reproduction, while floc develop by the coming together of adhesive cells. Morphological differences between clusters are readily apparent in flocculant yeast (B,D) and snowflake yeast (C,E). These yeast were isolated from the same replicate population (M5-2) at either 60 or 120 days of evolution. In each case, a single isolate is labeled with either either the green-fluorescent yeGFP or red-fluorescent dTomato.

We found that snowflake yeast invaded readily, and in head-to-head competition possessed a substantial fitness advantage over their floc competitors.

Snowflake yeast rapidly appeared in our experimental populations. Within just 28 transfers, snowflake yeast had evolved *de novo,* and risen to high frequency (>50%) in 12/40 experimental populations. Over the course of the experiment (155 days), snowflake yeast evolved in 36/40 populations, generally driving their floc ancestors to extinction Figure 6a). Indeed, floc regained numerical dominance in just one population taken over by snowflake yeast (Figure 6a).

To measure the relative fitness of invading snowflake yeast, we isolated pairs of floc and snowflake yeast from the 10 populations undergoing invasion at 60 transfers. We

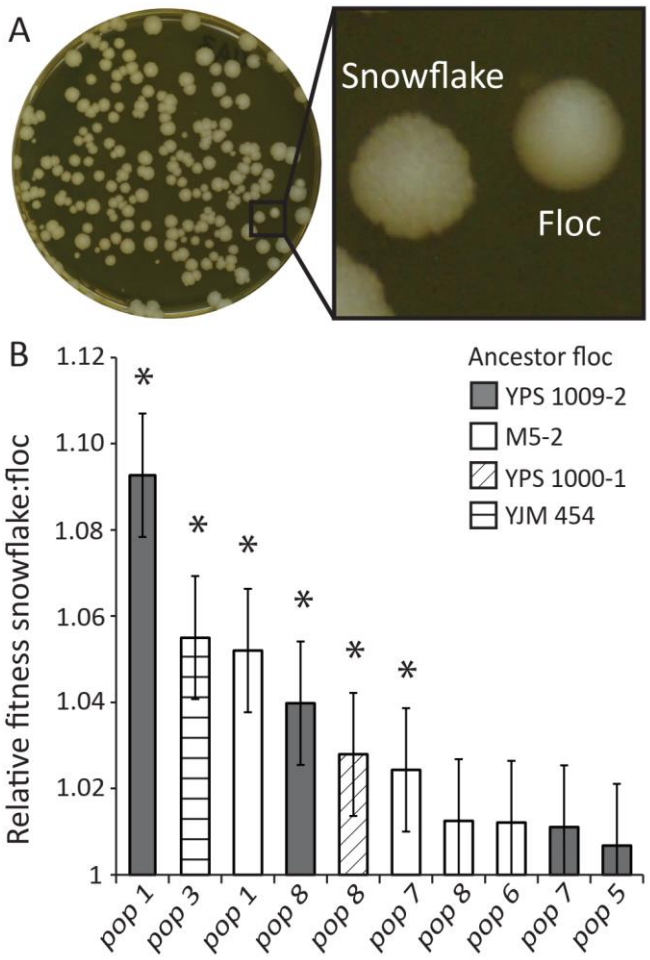

Figure 7 - Fitness of early snowflake strains. A) Snowflake and floc colonies have a distinct morphology, allowing us to use plate counts to determine fitness. B) Early snowflake yeast have a significant fitness advantage in 6/10 pairs isolated from the same population at transfer 60. Error bars are the 95% confidence intervals of least-squares means, derived from a 1-way ANOVA. Shown in (A) is a competition plate of a YPS 1009-2 floc-snowflake pair at a 1:10,000 dilution. Asterisks denote significance at the 0.05 level.

grew these pairs in isolation, and then inoculated competition populations with a 50:50 ratio (biomass) of each strain and measured relative fitness over 5 transfers. Invading snowflake yeast were more fit than their floc counterparts in all cases, but this was significant for only 6/10 strain pairs (Figure 7c; $F_{9,40} = 14.543$, $P < 0.0001$; ANOVA).

So why are snowflake yeast superior to floc yeast? We propose both proximate and ultimate hypotheses. A key aspect of fitness in our experimental system is the ability to rapidly settle to the bottom of the test tube, and for this task snowflake yeast may possess a direct advantage. While floc yeast are certainly able to settle quickly, they rely on a stochastic process for cluster formation. Specifically, clusters are formed through the collision and adhesion of sticky cells, and fragment when shear forces separate cells. In contrast, snowflake yeast form through a far more deterministic process, with daughter cells adhering to their parents after mitotic reproduction. Clusters only fragment when the among-cell tension exceeds cell-cell adhesive strength. Cluster size, and therefore settling speed[38], of individual clusters is heritable and is consistent for different genotypes of snowflake yeast[36,91]. Alternatively, floc's stochastic aggregative method of cluster formation may result in clusters that are more variable in size and settling speed. As a result, they may lack the ability to consistently produce large, fast settling clusters, allowing invasion by snowflake yeast.

Snowflake yeast may also possess an ultimate advantage, and be more capable of multicellular adaptation. Snowflake yeast develop clonally, while floc yeast form genetically chimeric clusters (Figure 6b,d). Since snowflake yeast clusters are clonal, there is little potential for the evolution of "cheating" lineages that realize a cell-level fitness benefit at the expense of cluster-level fitness[31]. The chimeric clusters formed by flocculation may result in the evolution of selfish lineages of cells, interfering with multicellular adaptation[74,92].

Even if cheating is not an issue, the high relatedness among cells in snowflake clusters should favor the evolution of reproductive altruism and cellular division of

labor[85,93]. The ability for clusters to partition tasks amongst cells may offer many benefits[74,94]. Snowflake yeast evolved a simple among-cell division of labor in response to a growth rate-settling rate trade-off. Large clusters grow less quickly than small clusters, due to limited nutrient availability to internal cells. Large snowflake yeast evolved an elevated rate of apoptosis, resulting in the production of proportionally smaller clusters that grow more rapidly[36]. Apoptotic cells leave no direct descendants, making this trait costly to individual cells expressing it, but increases the fecundity and fitness of the multicellular cluster [95]. This highlights an important shift in the level of selection from the unicellular to multicellular level, which may facilitate the subsequent evolution of multicellular complexity[74,79,81]. Genetically chimeric floc clusters are less individuated at the multicellular level, which may limit their capacity for multicellular adaptation[80,81].

Our results raise an obvious question: If snowflake yeast are actually superior to floc yeast, then why are floc more common in nature? We can think of two possible reasons. Flocculation provides protection from environmental stressors (like alcohol[96] and antibiotics[97]. Unlike snowflake yeast, flocculation can provide a fitness advantage even if opportunities for growth are limited. A rare floc strain can still join a group (and obtain a benefit of stress protection) as long as it produces adhesive glycoproteins[88]. In contrast, snowflake yeast clusters must grow large in order to gain the benefits of size, which requires a relatively resource-rich environment. Floc yeast should be far better at dispersing, as single cells and small clusters readily break away from a larger group. This both increases the number of propagules formed by a single genotype, and may also increase the distance of dispersal of each propagule. If dispersal is important to fitness, floc yeast may possess a substantial advantage.

3.5 Conclusion

‘Staying together’ and ‘coming together’ describe the two known modes of multicellular development. Evolutionary theory predicts that staying together is superior to coming together, and here we report the first empirical test directly comparing these two modes of development. Consistent with theory, we find that yeast that form clonal multicellular clusters (snowflake yeast) possess a striking fitness advantage over those that form chimeric aggregates (floc yeast). We hypothesize that the superiority of snowflake yeast could be due to either proximate effects of developmental mode (a larger average cluster size with a smaller variance than floc yeast), and/or ultimate evolutionary effects (more capable of multicellular adaptation than floc yeast). Further experiments are under way testing these hypotheses.

3.6 Acknowledgements

The authors would like to thank Adaugo Asouzu for laboratory assistance and Barry Williams for providing the ancestral floc yeast isolates. This work was supported by National Science Foundation grant DEB-1051115.

3.7 Author contributions

WR and MT conceived of the project. JP and WR conducted the experiments, analyzed the data, and wrote the paper.

CHAPTER 4. ECOLOGICAL ADVANTAGES AND EVOLUTIONARY LIMITATIONS OF AGGREGATIVE MULTICELLULAR DEVELOPMENT

4.1 Abstract

All multicellular organisms develop through one of two basic routes: they either aggregate from free-living cells, creating potentially-chimeric multicellular collectives, or they develop clonally via mother-daughter cellular adhesion. While evolutionary theory makes clear predictions about trade-offs between these developmental modes, these have never been experimentally tested in otherwise genetically-identical organisms. We engineered unicellular baker's yeast (*Saccharomyces cerevisiae*) to develop either clonally ('snowflake', Δ*ace2*), or aggregatively ('floc', *GAL1*p::*FLO1*), and examined their fitness in a fluctuating environment characterized by periods of growth and selection for rapid sedimentation. When cultured independently, aggregation was far superior to clonal development, providing a 35% advantage during growth, and a 2.5-fold advantage during settling selection. Yet when competed directly, clonally-developing snowflake yeast rapidly displaced aggregative floc. This was due to unexpected social exploitation: snowflake yeast, which do not produce adhesive FLO1, nonetheless become incorporated into flocs at a higher frequency than floc cells themselves. Populations of chimeric clusters settle much faster than floc alone, providing snowflake yeast with a fitness advantage during competition. Mathematical modeling suggests that such developmental cheating may be difficult to circumvent; hypothetical 'choosy floc' that avoid exploitation by maintaining clonality pay an ecological cost when rare, often leading to their extinction.

Our results highlight the conflict at the heart of aggregative development: non-specific cellular binding provides a strong ecological advantage – the ability to quickly form groups – but this very feature leads to its exploitation.

4.2 Introduction

The evolution of complex life on Earth has occurred through key steps in which formerly autonomous organisms evolve to become integral parts of a larger, higher-level organism[14,98-101]. These have been termed Major Transitions in Evolution[101], or Evolutionary Transitions in Individuality[14,102], one example of which is the transition from uni- to multi-cellularity. Multicellularity has evolved in at least 25 times in organisms as diverse as bacteria[2,103], archaea[104], and among deeply-divergent lineages of eukaryotes[75,105].

There are two basic modes of multicellular development. Cells can 'stay together' after mitotic division, resulting in clonal development if the life cycle includes a genetic bottleneck[2,30]. Alternatively, potentially unrelated cells can 'come together' via aggregation, which occurs in a few groups of terrestrial microorganisms[106,107]. Clonal development is thought to possess several advantages over aggregation for multicellular construction. First, under clonal development, cells comprising the multicellular organism have a high degree of genetic relatedness[32], which aligns the fitness interests of individual cells, facilitating the evolution of cooperative traits (*e.g.*, division of labor). Additionally, clonal development limits the potential for evolutionary conflict, as there is little standing genetic variation within an organism for selection to act on[17,31,34,81]. Through the same mechanism, clonal development stifles opportunities for the evolution of parasitic cell

lineages that infiltrate and exploit functional organisms[108]. Second, organismal clonality facilitates cluster-level selection. Genetic uniformity among the cells in a group results in a direct correspondence between emergent multicellular traits and heritable information (primarily genes) responsible for generating these traits[52,109]. Variation in the identity and frequency of different genotypes of cells within aggregates across multicellular generations undermines the heritability of emergent multicellular traits. Further, clonal development facilitates the shift from selection acting among cells to whole groups, simultaneously minimizing within-group genetic variation (thus largely preventing within-group selection) and maximizing between-group genetic variation[31]. Perhaps because of these benefits, the majority of independently-evolved multicellular lineages develop clonally.

Yet aggregative development possesses a unique (but largely unappreciated) advantage: multicellular bodies can form far more rapidly[30]. If a group is formed via the 'staying together' of cells after division, then its formation occurs by growth, causing the time required for body formation to scale with cellular generation time and organism size. In contrast, aggregation can occur far more rapidly. For example, aggregation of *Dictylostelium* into a multicellular mound can occur just 4-6 hours after starvation[110], and flocculation of yeast can occur in seconds[111]. Indeed, aggregative development is common in organisms that rapidly switch from unicellular to multicellular life history strategies upon sudden environmental change (*e.g.*, starvation in *Dictylostelium discoideum*[112] and *Myxococcus xanthus*[113]). Aggregation may also bring together cells with complementary properties, taking advantage of mutualistic interactions[114-116], but the evolutionary stability of this interaction generally requires a mechanism to limit social exploitation, such as host sanctions[117,118] or partner fidelity across generations[119].

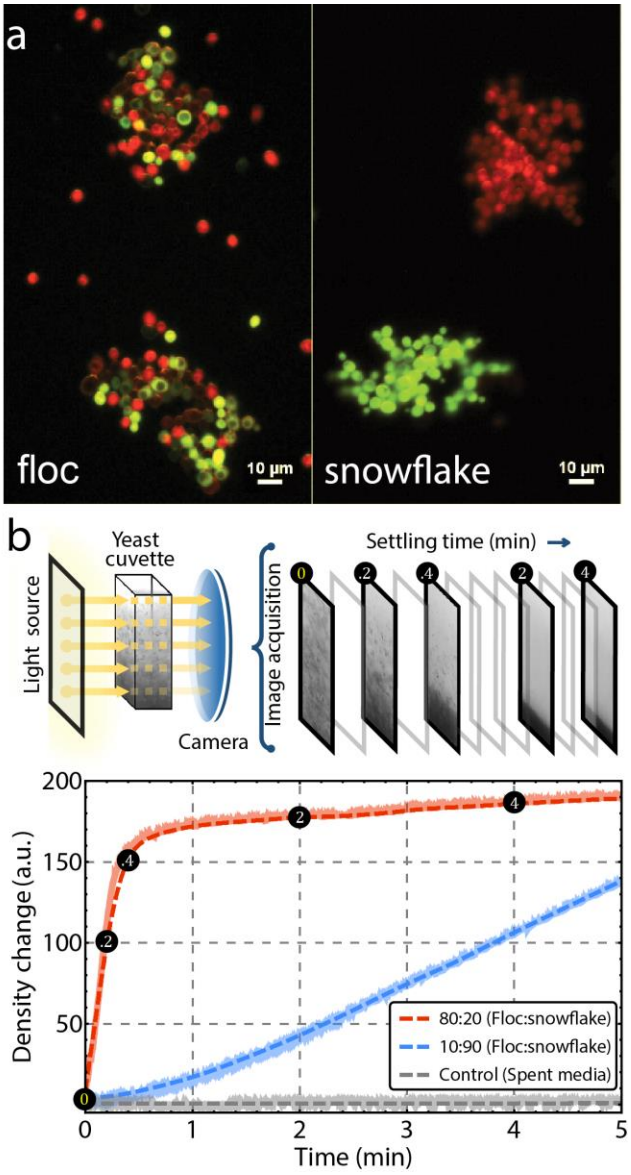

Figure 8 - Synthetic yeast system to study clonal and aggregative multicellular development. a) Synthetically created floc and snowflake yeast (*FLO1* insert and *ace2* knockout, respectively) labeled with either a red or green fluorescent marker. Both strains were created from the same unicellular ancestor. Flocs may be genetically-diverse, while snowflake yeast form clonal clusters. b) Settling rate was measured using high-resolution video acquisition of back-illuminated yeast cultures over 5 minutes of settling. Individual pixel intensities, which correlate to yeast density, were used to measure the rate of density change (see Figure 17; Appendix Movie 1). Raw density data (shadowed lines) was smoothed with a Savitzky-Golay smoothing function (dashed line) and the maximum slope of these dynamics is calculated as the settling rate. Shown are the density dynamics of fast (80:20 Floc:Snowflake) and slow (10:90 Floc:Snowflake) co-cultures, as well as a cell-free control where no density change is expected.

The origin of complex life cannot be understood in the absence of evolutionary mechanism. It thus is imperative that we understand how basic mechanisms of multicellular development effect the subsequent evolution of multicellular complexity. Mathematical modeling[17,30,52,109,120-123] and experiments in diverse systems[40,47,56,108,124,125] have generated consistent and robust predictions for the evolutionary consequences of variation in developmental mode. Yet because no model organisms develop through both routes, no experiments have directly compared ecological vs. evolutionary trade-offs between aggregative and clonal development. Here, we circumvent this historical limitation by engineering unicellular yeast (*Saccharomyces cerevisiae*) so that they form multicellular groups via either clonal development or aggregation.

The yeast *S. cerevisiae* can aggregate to form large clumps consisting of thousands of cells termed 'flocs'. Aggregation occurs via a lectin-like bonding between cell surface FLO proteins and cell wall sugars in adjacent cells[40,43]. Flocs preferentially form among mutual FLO^+ cells; FLO^- cells tend to be excluded from the group[126]. However, genetically-diverse strains can join a floc if they are FLO^+ (Figure 8a). In contrast, 'snowflake yeast' develop clonally, forming multicellular groups as a consequence of failed septum degradation after cytokinesis[37]; Figure 8a). When a cell-cell connection is severed, the group produces a viable propagule. This propagule experiences a single-cell genetic (but not physiological) bottleneck, as the most basal cell in the propagule is the mitotic parent of every cell in the group[37].

Engineered isogenic floc and snowflake yeast were constructed from a common unicellular ancestor. They were grown in a fluctuating environment, 24 hours of shaking incubation followed by selection for rapid sedimentation, that favors a rapid transition from

unicellularity (providing the highest growth rates) to multicellularity (increasing survival during settling selection). Aggregation was a superior strategy in monocultures: floc yeast, which spend most of the growth phase as unicells or small groups, grew 35% faster than snowflake yeast, but rapidly formed large flocs during settling selection, settling 2.5 times as fast as snowflake yeast. Yet in competition, snowflake yeast rapidly outcompete floc, the result of an unexpected social interaction. Despite being FLO^-, snowflake yeast embed themselves within floc clusters, making up a disproportionately high fraction of the biomass within flocs. Spatial analysis of chimeric aggregates demonstrates that snowflake yeast are uniformly, not randomly, distributed within the floc, suggesting a simple physical interaction between floc and snowflake is necessary for the formation of chimeric aggregate clusters. In principle, this parasitism could be prevented if floc evolved a partner choice mechanism, excluding heterospecific genotypes. We examined the invasion of such a 'choosy' floc genotype using mathematical modeling. In our model, selective binding is ecologically costly, as there is an advantage for individual cells to form groups with as many other cells as possible (this way they form the largest groups). Rare choosy floc is therefore unable to invade permissive floc, snowflake yeast, or a population consisting of both. Because choosy floc's aggregative performance is strongly frequency dependent, it should perform poorly (relative to a permissive floc) in genetically-diverse populations. This ecological cost may limit the evolution of strong kin recognition during aggregative development, paving the way for persistent evolutionary conflict.

4.3 Results

There are two important life history traits that affect fitness in our fluctuating environment: growth during 24 h batch culture and settling rate during settling

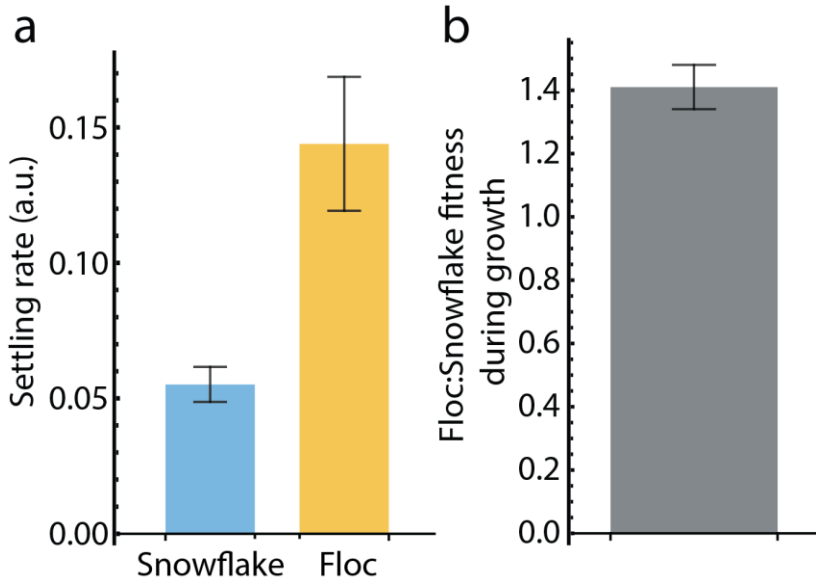

Figure 9 - Aggregative floc yeast are more fit than clonally-developing snowflake yeast in an environment favoring rapid group formation. Floc yeast are superior in two important life history traits that affect fitness in our experimental system. a) Floc yeast settle 2.5 times faster than snowflake yeast (t_8 = 9.82, p<0.0001, two-tailed t-test). Error bars are standard deviation (n=8). b) Floc yeast outcompete snowflake yeast over one 24 h growth period. Fitness was measured as the ratio of Malthusian growth parameters[127] for one 24 h period. Error bar is standard deviation (n=5).

selection[36,38,45]. To measure settling rate, we developed a novel method to quantify the dynamical effects of aggregation and settling in real time (Figure 8b, Figure 17; Appendix Movie 1; see methods section for details). Floc yeast are superior in both traits. First, floc yeast settle 2.5 times as fast as snowflake yeast, rapidly forming large aggregates during settling selection (Figure 9a; Appendix Movie 2; t_8 = 9.82, p<0.0001, two-tailed t-test). In direct competition, floc yeast outcompete snowflake yeast over one 24 h growth cycle (Figure 9b). This is likely a consequence of nutrient and oxygen limitation in snowflake clusters, which, in contrast to floc yeast, are always multicellular.

Co-culturing floc and snowflake yeast introduced markedly different behaviors. The settling rates of mixed populations increased dramatically (Figure 10a), and was

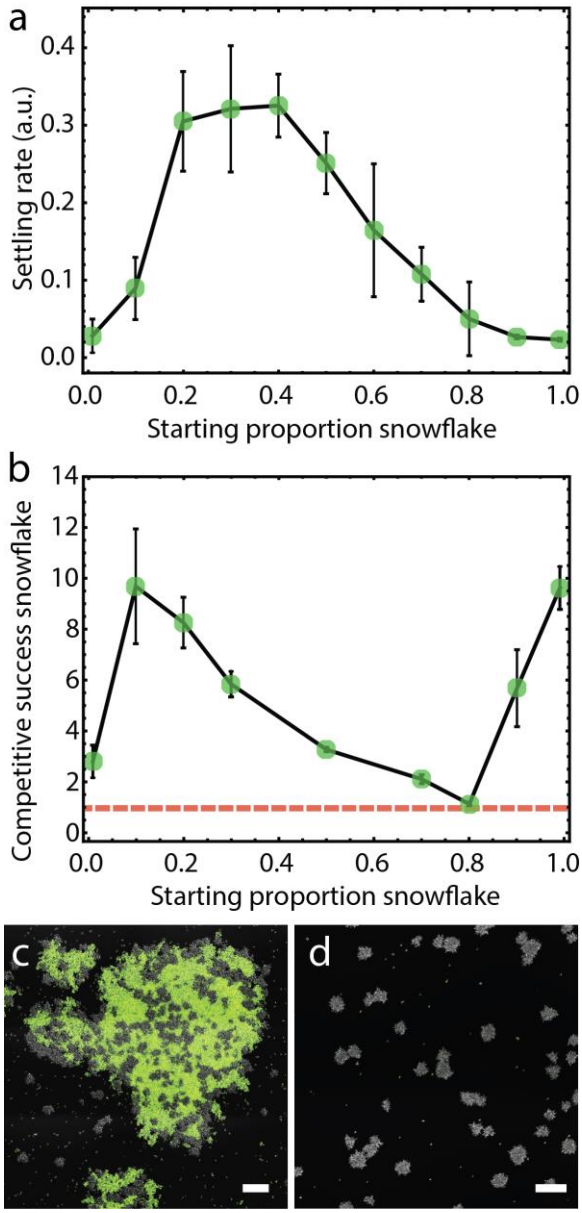

Figure 10 - Co-culturing floc and snowflake yeast. a) Mixed populations settle more rapidly than snowflake yeast or floc alone. Settling occurs the most rapidly at intermediate frequencies (20-50%; $F_{10,33} = 25.5$; $P < 0.0001$; ANOVA followed by Tukey's HSD). Error bars are standard deviation of four biological replicates, settling rate units are arbitrary. b) We measured the competitive success of snowflake yeast across two rounds of growth and settling. Snowflake yeast were more fit than floc at all genotype frequencies. Error bars are standard deviation of five biological replicates. (c). Snowflake yeast form chimeric aggregates with floc. Shown are snowflake yeast and GFP-tagged floc yeast starting at an initial inoculation ratio of 30:70 snowflake:floc-GFP (c) or 99:1 (d). Note that floc are below the concentration threshold required for aggregation, existing as unicells. Scale bars are 100 µm.

highest when snowflake yeast were at an intermediate frequency (20-50%; $F_{10,33} = 25.5$; $p < 0.0001$; ANOVA, pairwise differences assessed with Tukey's HSD with $\alpha = 0.05$). To examine the effects of co-culture on fitness, we performed a series of competition experiments (two rounds of growth and settling) across a range of starting snowflake frequencies, from 1% to 99%. Surprisingly, snowflake yeast were more fit than floc in all competitions, and their fitness was highly frequency-dependent. When snowflake yeast were rare (starting at 1% of the initial culture biomass), they had a small competitive advantage over floc (Figure 10b). This increased dramatically when they were more common (10-20% of initial starting biomass), declining until snowflake yeast reached 80%. Flocculation was impeded when snowflake yeast constituted >80% of the population, allowing multicellular snowflake yeast to compete against largely unicellular floc, causing their relative fitness to again increase dramatically (Figure 10b&d). These dynamics appear to be the result of an unexpected interaction: when mixed together, snowflake yeast and floc form chimeric clusters during the settling phase of the experiment (Figure 10c).

To determine which phase of the periodic environmental regime (*i.e.*, growing vs. settling) favored snowflake yeast during competition with floc, we measured snowflake yeast competitive success across one culture cycle. Consistent with earlier experiments (Figure 9c), snowflake yeast lost to floc over one 24 h growth cycle (Figure 11). Snowflake yeast fitness during growth was negative frequency dependent ($y = -0.005x + 0.91$, $p<0.0001$, linear regression). This is likely a consequence of overall nutrient consumption rates. When slower-growing snowflake yeast make up a larger fraction of the population, they consume resources less quickly, extending the time over which their floc competitors

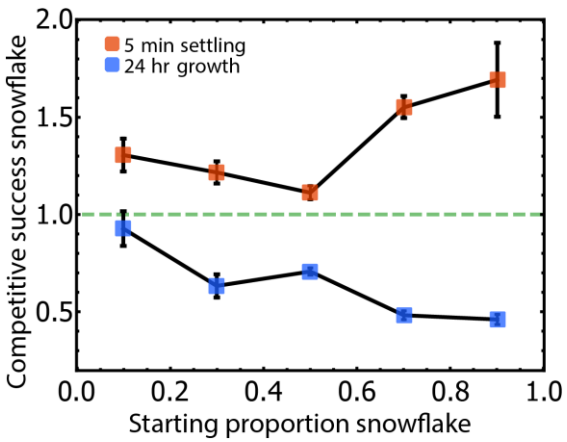

Figure 11 - Snowflake yeast outcompete floc during settling selection when forming chimeric aggregates. We examined the competitive success of snowflake yeast in competition with floc during both growth (over 24 h of culture) and settling selection (5 minutes at 1 g). Snowflake yeast had lower fitness at all starting genotype frequencies during the growth phase of the culture, yet had higher fitness during settling selection. This is in stark contrast to what we observe in pure culture, where floc yeast settle 2.5 times as quickly as snowflakes (Figure 1b). Error bars are standard deviation of five biological replicates.

can compound their growth rate advantage. In contrast to growth, however, snowflake yeast possessed an advantage during settling selection (Figure 11).

One way that snowflake yeast could gain an advantage during settling selection is if they are over-represented in large, fast-settling chimeric aggregates. This would be unexpected, as *FLO1* yeast preferentially adhere to other floc cells, efficiently excluding non-flocculating unicells from flocs[126] (Figure 18). We imaged co-cultures in which snowflake yeast were either rare (20% initial biomass; Figure 12a) or common (80% initial biomass; Figure 12b). Surprisingly, snowflake yeast were overrepresented in chimeric flocs (*i.e.*, groups larger than the largest individual snowflakes; Figure 12c) at both genotype frequencies.

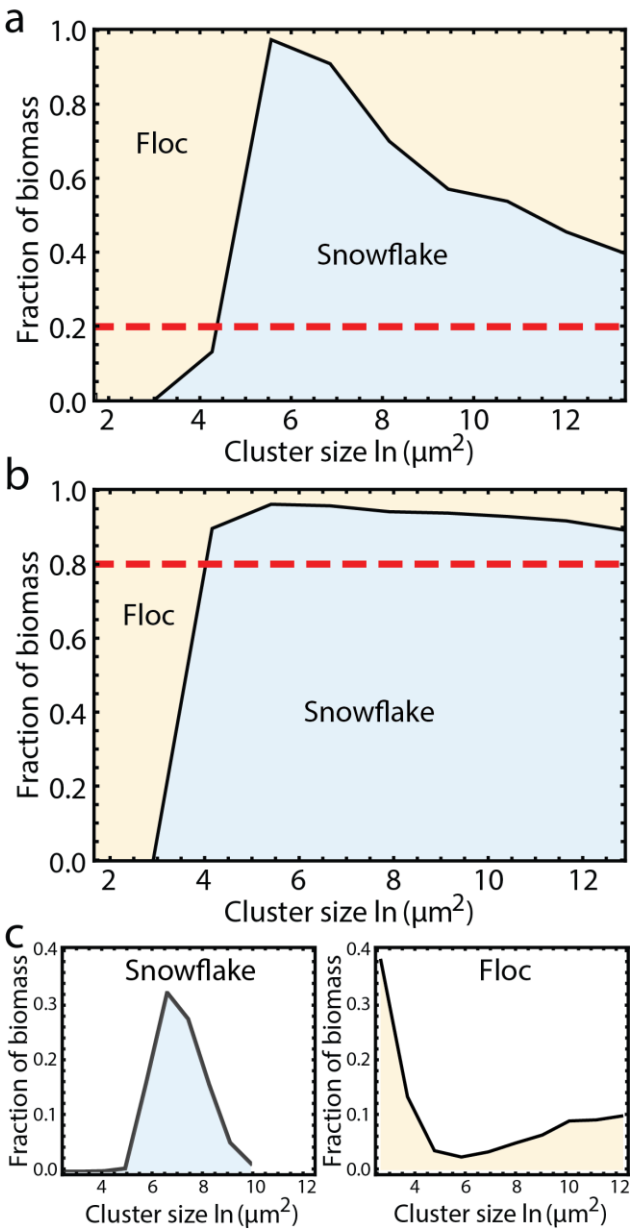

Figure 12 - Snowflake yeast are overrepresented in large chimeric aggregates. Snowflake yeast constitute a larger fraction of the biomass within large flocs than is expected by their overall population frequency (dashed line). Shown are snowflake yeast at 20% (a) and 80% (b) overall frequency. c) Size distributions of pure snowflake and floc cultures.

One feature of chimeric aggregates that stands out is the appearance of a relatively uniform distribution of snowflake yeast within the aggregate (Figure 19). We rarely see large patches of pure floc cells, and never see large patches of just snowflake yeast. To

quantify the spatial distribution of snowflake yeast within chimeric aggregates, we first measured the spatial autocorrelation function (Moran's *I*). We found that the correlation length is similar in size to the cluster radius (14.1 +/- 0.2 μm, 14.2 +/- 0.2 μm, 13.9+/- 0.1 μm, 11.7+/-1.3 μm for 30%, 20%, 10%, and 1% snowflake yeast, respectively). We next measured the pair correlation function, *g(r)*, which measures the probability of finding two clusters separated by a given distance (Figure 20), normalized by a random distribution at the same density. We find that the distribution of snowflake yeast clusters is highly structured within aggregates. Clusters are unlikely to be found very close to each other; specifically, clusters are less likely to be found with a center-to-center separation less than or equal to 1.3 times their diameter than expected by random chance. Relatedly, clusters are more likely to be found with center-to-center separations between 1.3-1.9 times their diameter than one expected by random chance. Thus, the distribution of clusters within an aggregate is more evenly dispersed than would be expected by a random mixing of genotypes. This even dispersal suggests that snowflake yeast are capable of binding to a patch of floc cells, but not a patch of snowflake yeast, during aggregate formation. Consistent with this hypothesis, floc appear to act as an adhesive, binding together snowflakes (Figure 21). We do not see any evidence of direct snowflake-snowflake adhesion. This analysis shows that snowflake yeast join chimeric aggregates more efficiently than floc yeast, despite the fact that floc yeast can stick to both floc and snowflake yeast, while snowflake yeast can only stick to floc yeast.

A classic solution to social conflict in aggregating multicellular organisms is kin recognition, allowing individuals to prevent cheating by only joining groups with close relatives[128-131]. Here, we examine whether a self-recognition mechanism would help

flocculating yeast outcompete snowflake yeast by constructing a mathematical model (see Methods). Briefly, we assume that there are three types of yeast: a snowflake yeast strain (*S*), a "choosy floc" (*C*) that uses a self-recognition mechanism to adhere just to clonemates, and a "permissive floc" (*P*) that has no such self-recognition mechanism, adhering to both permissive floc and snowflake yeast. We simplify our analysis by focusing strictly on the role of self-recognition in the formation of groups. Thus, we assume that after some initial period of population growth, there is an aggregation phase in which cells stop reproducing and the flocculating yeast aggregate to form groups. Rather than modeling the complex dynamics of group size and shape during settling selection, we make the simplifying assumption that only the largest groups survive. While floc yeast rapidly form groups, increasing in size as a function of time (Figure 22), snowflakes themselves do not change in size (as there is no growth; Figure 22), though they may join aggregates with permissive floc. When floc are growing at higher density, it takes less time to form groups that can outcompete snowflake yeast during settling selection (Figure 22).

We consider all pair-wise competitions between permissive floc, choosy floc, and snowflake yeast for different starting genotype frequencies (Figure 13a-c). For each competition, we simulate the aggregation process and then select 10% of the population from the largest groups (selection that is roughly analogous to the experimental protocol). Snowflake yeast are overrepresented within large, fast-settling flocs (recapitulating our experimental data; Figure 10), allowing them to outcompete permissive floc regardless of their starting frequency. We also find that the largest chimeric aggregates, representing the fastest-settling aggregates, form with intermediate frequencies of snowflake yeast (peaking

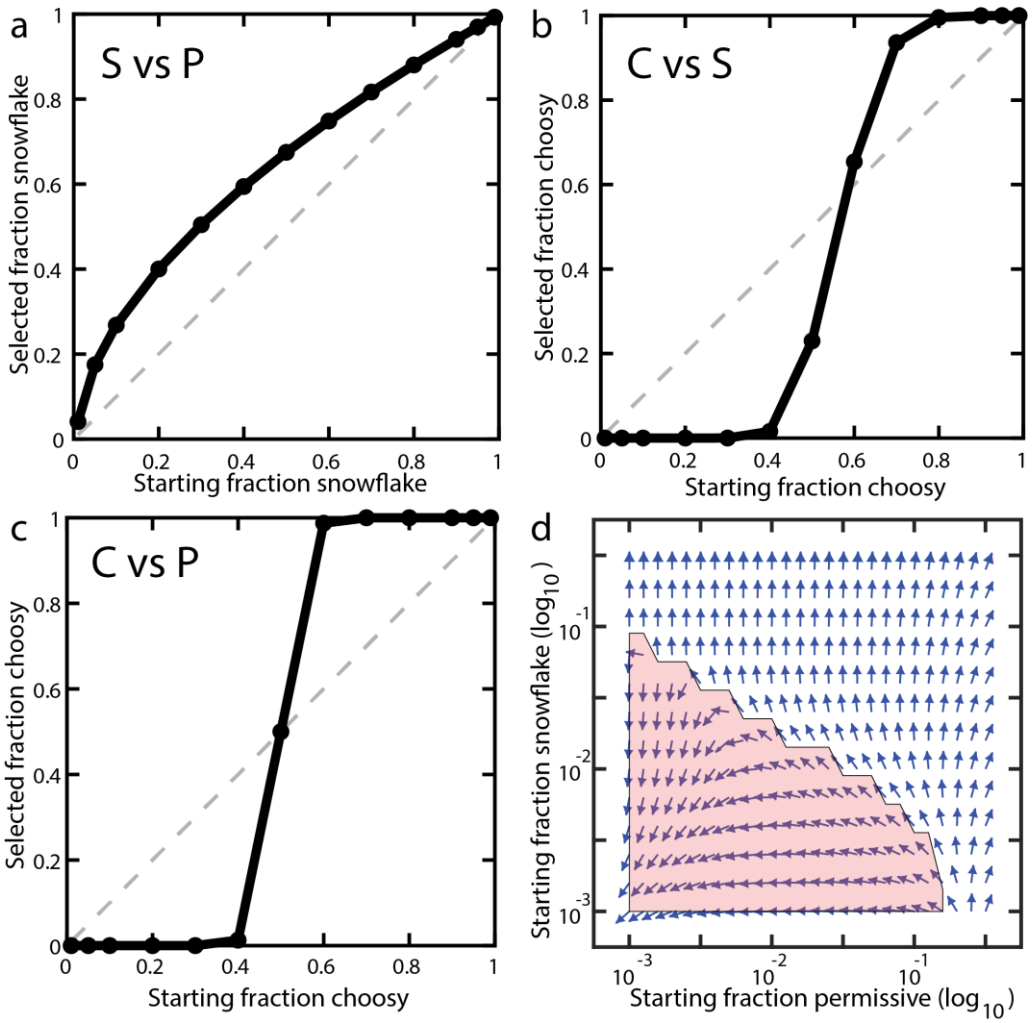

Figure 13 - Modeling the dynamics of kin recognition in floc yeast. a) Snowflake yeast, *S*, were capable of displacing permissive floc, *P*, at all frequencies during settling selection. In contrast, the fitness of choosy floc, *C*, (b) and permissive floc (c) were both strongly positively frequency dependent. d) Phase portrait showing the changes in *P* and *S* after one round of settling selection in competition with *C*. Arrows show the direction of change in proportion of *S* and *P* as a function of different starting frequencies. When *P* and *S* start out above a critical threshold, they displace *C*; otherwise, *C* displaces them (red highlight).

at 40% S; Figure 24). This is similar to our experimental data (Figure 10a), where the fastest-settling aggregates are also found at intermediate frequencies (20%-50%). In contrast, if snowflake and choosy floc compete, then choosy floc increases in abundance whenever it is more than ~60% of the population (though the precise frequency depends

on model parameters, like density, aggregation time, and binding probability; Figure 23, Figure 25, Figure 26). Thus, neither snowflake yeast nor choosy floc can invade each other when rare. Finally, since permissive and choosy floc behave the same in the absence of snowflake (they do not co-aggregate), their dynamics are entirely frequency dependent and neither can invade from rare.

In a three-way competition, snowflakes can invade populations of choosy floc with the help of permissive floc (Figure 13d; see results from longer durations of aggregation in Figure 26). By forming large, fast-settling chimeric aggregates, mixtures of snowflake and permissive floc can outcompete choosy floc (Figure 13d; Figure 26). Of course, this is an unstable alliance, as the exploitation of permissive floc will ultimately result in a monotypic population of snowflake yeast (Figure 13a&d; Figure 26). Sometimes, however, this social exploitation of floc is costly for snowflake yeast. When snowflake and permissive floc are below the threshold required to displace choosy floc, exploitation of permissive floc results in a rapid deterioration of their ability to make large chimeric aggregates, to the detriment of both snowflake and permissive floc (Figure 13d; Figure 26).

A simple extrapolation of our model highlights the cost of kin discrimination during aggregative development. Consider a genetically-diverse population of aggregative organisms, each of which only adheres to clonemates. Because aggregation rate is frequency and density dependent (Figure 13, Figure 22, Figure 23), any genotypes that are locally rare will be unable to rapidly form large groups, as they will be capable of interacting with only a small fraction of the population. Strict kin recognition during aggregation therefore undermines the ecological advantage of aggregation. This is even

more of a problem if the benefits of aggregation require that a size threshold be met (*e.g.*, enough individuals to form a multicellular fruiting body[132]).

4.4 Discussion

Development is a fundamental aspect of multicellularity, orchestrating the pattern of cellular behaviors that give rise to multicellular phenotype and influencing a lineage's evolutionary potential. Despite significant theoretical work, the lack of appropriate model systems has limited our ability to directly test the role of developmental mechanism on the subsequent evolution of multicellularity. We circumvent this limitation by engineering aggregative and clonal development from an isogenic unicellular yeast ancestor (Figure 8a).

We grew our yeast under conditions in which selection favored a rapid transition from a unicellular to multicellular stage, the type of environment that is thought to favor aggregative multicellularity[30]. The advantage that aggregative floc yeast showed in monoculture (Figure 9) evaporated once they were competed directly with clonally-developing snowflake yeast (Figure 10), the result of a wholly unexpected social exploitation. Snowflake yeast, which do not produce adhesive Flo1 proteins, embed themselves within large floccy aggregates at a higher frequency than the floc genotype (Figure 10c&d, 5; Figure 20, Figure 21). As a result of this social exploitation, snowflake yeast rapidly displace floc (Figure 10b). This result is even more striking in light of prior work in flocculating yeast, where Smukalla *et al* (2008) show that *FLO1* acts as a greenbeard gene, excluding unicellular *FLO1*$^{-}$ competitors from the floc. This is thought to be a consequence of preferential binding between *FLO1*$^{+}$ cells, leading to phase

separation[126]. In our case, the ability for *FLO1*$^{-}$ snowflake yeast to co-aggregate with floc appears to arise as a consequence of their branchy structure, allowing them to become entangled within a floc. Our results also provide context for understanding the results of a prior experiment, in which five wild isolates of flocculating yeast were evolved with daily settling selection. Here, snowflake yeast arose *de novo* and largely displaced their floc ancestors in 35/40 replicate populations[47].

Self / nonself recognition systems play a key role during the evolution of multicellularity, limiting the potential for within-organism genetic conflict[129,130,133]. This may be especially important in lineages that develop aggregatively, as they are more likely to form genetically-diverse multicellular groups. Kin-recognition mechanisms have evolved independently in cellular slime molds[129,133] and *Myxococcus* bacteria[130,134], both of which develop via aggregation. We explored the evolution of self-recognition in our system using a mathematical model. We considered our standard 'permissive floc', which binds to other permissive floc or snowflake yeast, and 'choosy floc', which only attaches to clonemates. While it might seem like choosy floc (which axiomatically cannot experience social conflict) would always be at an advantage, this was not true. Permissive binding increases opportunities for cell-cell adhesion, increasing aggregation speed and group size. Indeed, our experiments show striking support for this hypothesis: floc that formed chimeric aggregates were capable of settling much faster than floc alone (Figure 10a). In our model, choosy floc pay an ecological cost when rare, as it can only bind a small fraction of the cells in the population, forming small groups. This strong positive-frequency dependent selection makes it difficult for choosiness to arise from a population of permissive ancestors, a cost which is compounded if the population is composed of

multiple choosy genotypes, each of which is only capable of adhering to clonemates. Consistent with this hypothesis, kin discrimination systems in extant aggregative organisms are quite permissive: wild-collected isolates readily form genetic chimeras[133,135,136], sometimes (but not always[116,137]) resulting in social cheating[130,138,139].

Our results highlight a fundamental trade-off faced during aggregative development: selection for rapid group formation often favors permissive binding, but the resulting high within-group genetic diversity lays the foundation for persistent evolutionary conflict. This has important implications for the evolution of multicellular complexity, as the resulting genetic conflict can undermine multicellular adaptation[125]. Indeed, aggregation is relatively uncommon among independently-evolved multicellular lineages[48,107], and all known examples of independently evolved 'complex multicellularity' (*i.e.*, metazoans, land plants, mushroom-forming fungi, brown algae, and red algae[105]) develop clonally. In the context of major evolutionary transitions, aggregation appears to be self-limiting, the evolutionary potential of aggregative lineages constrained by an ecological imperative for effective group formation.

4.5 Methods

4.5.1 Strain construction

All strains used in this study are listed in Table 1. We constructed snowflake and flocculating genotypes from a single clone of the initially unicellular *S. cerevisiae* strain Y55. Snowflake yeast were made as in[50], but we replaced the *ACE2* ORF with *HYGMX*. Flocculating yeast were made by amplifying the *KAN-GAL1*p*::FLO1* cassette from DNA template from *S. cerevisiae* strain KV210[44,140] and replacing the *URA3* ORF in our

ancestral strain. *ura3*Δ*::KAN-GAL1*p*::FLO1*/*ura3*Δ*::KAN-GAL1*p*::FLO1* diploids were obtained by autodiploidization of single spores collected via tetrad dissection onto Yeast Peptone Dextrose plates (YPD; per liter: 20 g dextrose, 20 g peptone, 10 g yeast extract, 15 g agarose) then replica plated onto YPD + 200 mg/L G418. Transformants were confirmed by PCR as well as phenotype when grown in YPGal medium (per liter: 20 g galactose, 20 g peptone, 10 g yeast extract). For microscopy and competition experiments, strains were tagged with green and red fluorophores. To do this, plasmids pFA6a-TEF2Pr-eGFP-ADH1-Primer-NATMX4 and pFA6a-TEF2Pr-dTomato-ADH1-Primer-NATMX4 were amplified and inserted into the *LYS2* locus, and transformants were confirmed via fluorescent microscopy. All transformations were done using the LiAc/SS-DNA/PEG method of transformation[61].

Table 1 – Strains used in this study.

Strain	**Relevant Genotype**	**Reference**
Snowflake	Δ*ace2*::HYGMX	This study
Floc	Δ*ura3::KAN-GAL1*p*::FLO1*	This study
Snowflake-GFP	Δ*lys2*::*TEF2*p-yeGFP	This study
Snowflake-RFP	Δ*lys2*::*TEF2*p-dTomato	This study
Floc-GFP	Δ*lys2*::*TEF2*p-yeGFP	This study
Floc-RFP	Δ*lys2*::*TEF2*p-dTomato	This study

4.5.2 Competitive success assay

To determine if snowflake yeast had a competitive advantage over floc yeast, we competed snowflake and floc starting at a range of initial genotypic frequencies (0-100%

snowflake in 10% increments) over two days of daily selection for fast settling for 5 min on the bench as in Ratcliff *et al.*, 2012. Snowflake and flocculating yeast were grown up in a mixture of galactose and glucose (YPGal+Dex; per liter; 18 g galactose, 2 g dextrose, 20 peptone, 10 g yeast extract) for 24 h at 30°C, shaking at 250 rpm. This concentration of galactose and glucose was used because it yielded clusters of similar size after 24 h of growth in snowflake and floc yeast (mean floc log(volume) = 12.5, mean snowflake log(volume) = 11.5, $t(2) = -0.39$, $p = 0.73$). Five replicates of 500 µL of each starting genotypic frequency was mixed from overnight cultures and 100 µL of this culture was diluted into 10 mL YPGal+Dex for the competition experiment. The remaining 400 µL was used to measure the initial count of snowflake and floc yeast. To do this, EDTA (50 mM, pH 7) was used to de-flocculate cells to run through a CyFlow® Cube8 flow cytometer where two distinct peaks corresponding to unicellular floc cells and snowflake cultures could be counted. Counts of unicellular floc and snowflake yeast were obtained for time 0 and after three days of competition. The competitive success of snowflake yeast was calculated as the ratio of snowflake to floc yeast after competition relative to before competition using the equation (2):

$$Competitive\ success = \frac{f_2(1 - f_{1)}}{f_1(1 - f_2)} \quad (2)$$

where f_1 is the frequency of snowflake yeast before competition and f_2 is the frequency of snowflake yeast after competition[124]. This fitness measure is simple and general (*i.e.*, it doesn't assume any underlying model of population dynamics, like exponential growth), and accommodates different starting frequencies.

4.5.3 Measuring settling rate

Unlike snowflake yeast, floc yeast form groups as they are settling, so we needed to measure the properties of flocs during the process of settling directly. To do this, we developed a novel, robust, high-throughput method of measuring the settling speed of yeast populations. Various methods to measure aggregation and settling in yeast exist[111,141-144], but most of them introduce experimental variables that limit their relevance to our system[141], and no method is considered standard in yeast research in general[141,144,145]. Importantly, most of them lack the temporal resolution needed (seconds) to capture the fast-settling profiles of some of our strains. In our method, we placed the yeast in back-illuminated cuvettes, and used high-speed high-resolution video acquisition (24 fps, 3840 x 2160 pixels, Sony a7R II, 90 mm macro lens) to capture changes in pixel densities over the settling time (Figure 8b). Our method relies in the fact that settling and flocculation produce optically denser regions, relative to the initial density distribution (Figure 17; Appendix Movie 1), thus allowing us to measure the rate of this density changes. We pre-processed our raw density data with a Savitzky-Golay smoothing function in order to preserve the signal over the noise without sensibly changing the shape of the dynamics (Figure 27). We then calculated a characteristic settling rate, as the maximum slope in the density dynamics. We validated our method by quantifying the percentage of biomass settled at 5 min in floc and snowflake cultures, showing that, as expected, a higher settling rate indicate a higher proportion of settled cells (Figure 28).

4.5.4 Competitive success during growth and settling

There are two important life history traits in our experimental system: growth rate and settling rate[36,45]. We measured the competitive success of snowflake yeast during both stages. Snowflake and floc yeast were grown separately for 24 h in YPGal+Dex. As above, five replicates of 500 µL of various starting genotypic frequencies (10-90% snowflake in 20% increments) were mixed from overnight cultures and 100 µL was used to dilute into fresh YPGal+Dex and the remaining 400 µL was used to calculate initial snowflake and unicellular floc counts as described above. To measure snowflake competitiveness during growth, 500 µL of each culture was deflocculated using EDTA and snowflake and floc counts were measured on the flow cytometer after 24 h or growth at 30°C, shaking at 250 rpm. To measure competitive success over one round of settling selection, 2 mL of each snowflake/floc co-culture was aliquoted into 2 mL microcentrifuge tubes. 500 µL was then aliquoted into 1 mL microcentrifuge tubes and deflocculated to obtain pre-selection snowflake and floc concentrations as described above. The remaining 1.5 mL was allowed to settle on the bench for 5 min, after which the top 1.4 mL was discarded. The remaining 100 µL was deflocculated and post-selection snowflake and floc counts were obtained via flow cytometry.

4.5.5 Examining the composition of aggregates

The composition of snowflake and floc yeast within large chimeras was measured by fluorescent microscopy, using a Nikon Eclipse Ti inverted microscope with a computer-controlled Prior stage. Specifically, snowflake and floc-GFP were grown for 24 h in YPGal+Dex. Four replicates of snowflake and floc co-cultures with differing amounts of starting snowflake (20% or 80%, respectively) were inoculated into fresh medium and grown for another 24 h. 10 µL of this culture was placed between a slide and a 25 x 25 mm

coverslip and the whole coverslip was imaged by combining 150 separate images at 100 x magnification yielding a 42456 x 42100 pixel (1.78 billion pixels; 1.23 x 1.22 cm) composite image. The percentage of biomass in different cluster size classes belonging to either snowflake or floc yeast was calculated using a custom Python script. “Large flocs” were considered to be anything larger than the largest snowflake clusters (Figure 12c).

4.5.6 Mathematical modeling

We consider a settling competition between snowflake clusters and flocculating cells. If flocculation, settling, and reproduction all occur together we might expect a complicated set of dynamics resulting from the interplay between these processes. We simplify our analyses by focusing strictly on aggregation. We assume that aggregation and settling happen after the primary growth phase and occur faster than reproduction such that the populations of cells are large as a result of several generations of reproduction in media. Furthermore, we consider aggregation and settling as two separate processes. Thus, we assume that there is some time in which cells aggregate and afterwards the groups are exposed to settling selection. This assumption allows us to focus on modeling the dynamics of aggregation and circumvent explicit spatial models that would be required to consider the dynamic interactions between aggregation and settling via centrifugation. We model the dynamics of aggregation using a system of differential equations, where a snowflake cluster composed of i cells is denoted as S_i, a floc of i choosy cells is C_i,and a floc of i permissive cells is P_i. (equations 3-6):

$$\frac{dC_i}{dt} = \sum_{j=1}^{\frac{i}{2}} p(i-j,j)C_j C_{(i-j)} - \sum_{j=1}^{N-i} p(i,j)\left(1 + \delta_{i,j}\right)C_i C_j \tag{3}$$

$$\frac{dS_i}{dt} = -\sum_{j=1}^{N-i} p(i,j)S_i P_j \tag{4}$$

$$\frac{dP_i}{dt} = \sum_{j=1}^{i-1} p(j,i-j)S_j P_{i-j} + \sum_{j=1}^{\frac{i}{2}} p(j,i-j)P_j P_{(i-j)} - \sum_{j=1}^{N-i} p(i,j)P_i S_j - \sum_{j=1}^{N-i} p(i,j)\,(1+\delta_{i,j})P_i P_j \tag{5}$$

$$\frac{dn_i}{dt} = \sum_{j=1}^{i-1} p(j,i-j)S_j P_{i-j}(j + \frac{n_{i-j}}{P_{i-j}}) + \sum_{j=1}^{\frac{i}{2}} p(j,i-j)P_j P_{i-j}\left(\frac{n_j}{P_j} + \frac{n_{i-j}}{P_{i-j}}\right) - \sum_{j=1}^{N-i} p(i,j)P_i S_j\left(\frac{n_i}{P_i} + j\right) - \sum_{j=1}^{N-i} p(i,j)\left(1+\delta_{i,j}\right)P_i P_j\left(\frac{n_i}{P_i} + \frac{n_j}{P_j}\right) \tag{6}$$

Choosy floc clusters C_i can bind to former larger flocs but we assume each cluster reaches a maximum size N (N=1000 in our computations). We also ignore group fragmentation. Thus: $C_i + C_j \xrightarrow{p(i,j)} C_{i+j}$, where $i + j \leq N$ and $p(i,j)$ is the probability of a successful binding that depends on the size of two flocs. Specifically:

$$p(i,j) = (i^{1/3} + j^{1/3})^3 \quad (7)$$

This function depends on many factors including the geometry of the two clusters, the probability of collision, the probability of a collision resulting in binding, etc. We assume that it is a simple function of the radii of the two clusters: $p(i,j) = (r_i + r_j)^3$ where r_i and r_j are the radii of C_i and C_j and the radii can be approximated by considering the clusters as spheres. Thus, if the volume of a single cell is $\frac{4}{3}\pi r^3$, then the volume of C_i is $(i)\frac{4}{3}\pi r^3$ which makes the radius of C_i equal to $i^{1/3}r$. We consider r =1 to simply the calculations. The $\delta_{i,j}$ term accounts for the extra loss if two identically-sized flocs interact, i.e. if two C_i bind then the loss is double that of a C_i binding a C_j where $i \neq j$.

We use P_i to denote permissive flocs. Since P cells can bind to either its own cells or snowflake cells, a P_i cluster may be composed of k floc cells and $i - k$ snowflake cells for any $k \geq 1$. We assume that there are a large number of clusters and cells and track the number of snowflake cells in P_i clusters for each size i, which we denote n_i. This assumption corresponds to treating the aggregative mixture as a classic tank mixing problem.

In all competitions except for Figure 25 and Figure 26, we assume an initial inoculum of 1000 concentration units that is split between C, S, and P. The initial distribution of S_i is fit to a lognormal distribution that matches empirical data (Figure 22). This distribution only changes in the presence of permissive floc. The winner of the settling competition is determined by solving equations (3-6s) for some time t and selecting the largest 10% of the population, using group size as a proxy for settling speed. This is

analogous to our experimental system, where 10% of fastest-settling yeast biomass gets passaged to the next tube following settling selection. For *C* cells, as time increases, more of the distribution is represented in the largest fractions ($\approx i = N$; Figure 22). Thus, the amount of *C* cells in the top 10% of possible clusters size increases with time, but levels out for longer *t* (Figure 22).

The mathematical model captures a single round of aggregation and selection without regard to how populations grow in between selective events. In cases where we consider multiple rounds of aggregation and selection (Figure 25), we do not use any explicit models of population growth. Rather, we multiply the final proportions of cells after selection by the inoculum size and use that as the input to the next iteration of aggregation and selection. This bypasses the possibility that different population growth dynamics might alter the proportions of cell types. In addition, we also assume that the *P* and *S* cells dissociate from their mixed groups and begin the next aggregation and selection phase as separate entities.

4.6 Author contributions

J.P., W.C.R. and P.M.Z designed the experiments. J.P. conducted the experiments. J.P. and P.M.Z. designed the settling rate analysis pipeline. All authors contributed to data analysis and interpretation. E.L. performed the modeling. P.Y. performed the spatial analysis of snowflake yeast within chimeric aggregates. J.P. and W.C.R. wrote the first draft of the paper, all authors contributed to revision. This work was supported by NASA Exobiology grant no. NNX15AR33G to W.C.R. and E.L., NSF grant no. IOS-1656549 to W.C.R. and

P.J.Y., NSF grant no. IOS-1656849 to E.L., and a Packard Foundation Fellowship for Science and Engineering to W.C.R.

4.7 Acknowledgements

We would like to thank Kevin Verstrepen for providing us with strain *S. cerevisiae* KV210, which contained the *GALp::FLO1* construct, and Daniel Weissman for insightful feedback on the manuscript.

CHAPTER 5. DEVELOPMENTAL MODE AND THE EMERGENCE OF MULTICELLULAR INDIVIDUALITY

5.1 Abstract

The evolution of multicellularity was transformative for life on Earth and is marked by and evolutionary transition in individuality (ETI; from the cells to the multicellular group). However, it is still unclear how multicellular groups emerge with the potential of undergoing an ETI. Early multicellular life cycles may be integral to this process. Multicellular organisms can develop by 'staying together' following cellular division (clonal development) or via the aggregation of individual cells (aggregative development). Clonal development begets clusters with minimal genetic conflict, while aggregation results in genetically-diverse multicellular collectives. Clonal development may be superior to aggregative development because it aligns the fitness interests of cells, but this is hard to test experimentally due to a lack of a suitable model. Here, we circumvent this limitation by synthetically creating a yeast system capable of both clonal and aggregative development. We subjected clonally-reproducing (termed 'snowflake' yeast) and aggregative (termed 'flocs') yeast to daily selection for rapid sedimentation in liquid medium, an environment where multicellularity is adaptive. Snowflake and floc yeast responded to selection by increasing their settling rate, but fitness increases in evolved isolates was achieved in two distinct life history traits. Snowflake yeast increased fitness during settling (multicellular-level trait), while floc yeast increased fitness during growth (unicellular-level trait). Furthermore, snowflake yeast had a larger proportion of fixed mutations that decrease fitness during growth but may increase cluster-level fitness. Our

results show that, even in simple multicellular yeast, clonal development facilitates the emergence of higher-level individuality, crucial for stabilizing the first steps in this major evolutionary transition.

5.2 Introduction

The evolution of multicellularity gave rise to a remarkable diversity of multicellular forms and life cycles. Generally, there are two basic routes to forming a multicellular body. Individual cells can 'stay together' by disrupting separation following cellular division, forming clonal clusters that exhibit little within-group genetic variation. Alternatively, cells may live solitarily for most of their life cycle but occasionally 'come together', or aggregate, to form a potentially-diverse multicellular collective in response to some stimuli, such as starvation[2,30,84,106,146]. Multicellularity has evolved multiple times via both routes[2,8,107], but 'complex multicellularity'[105] (*e.g.*, plants, animals, fungi) has only evolved in lineages that have made the transition to multicellularity clonally[48,105].

The evolution of multicellularity is one of the major transitions in evolution[101], which are characterized by evolutionary transitions in individuality (ETI). During an ETI, there is a shift in the hierarchical level at which selection via natural selection operates (i.e., heritable variation that affects fitness)[14,24,101]. During the transition to multicellularity, clonal multicellular life cycles may potentiate the ability for simple multicellular groups to undergo an ETI by increasing the efficacy of group-level selection relative to cell-level selection. Specifically, clonal development limits the potential for evolutionary conflict, as there is little standing within-group genetic variation for selection to act on[8,108]. Second, any variation that arises due to mutation gets partitioned among multicellular offspring,

allowing selection to act on the group-level effects of *de novo* mutations, increasing the scope for group-level adaptation. This also limits the potential for genetic conflict by suppressing the proliferation of 'cheating' cellular lineages[8,125]. Despite these clear predictions, they are difficult to test experimentally due to a lack of a suitable model system that is capable of both developmental modes. We circumvent this limitation by creating an isogenic yeast system capable of both clonal and aggregative development.

Clonal 'snowflake yeast' were made by knocking out *ACE2* in a unicellular ancestor, producing clonal clusters[37,147]. Floc yeast were created by placing the dominant *FLO1* gene under transcriptional control of the inducible *GAL1* promoter, leading to strong flocculation and the formation of chimeric clusters when grown in galactose-containing medium[126,147]. Snowflake and floc genotypes were created from a single homozygous-diploid unicellular ancestor, producing strains that differ only in these two genes, and thus only differing in their mode of cluster formation[147]. Here, we use experimental evolution to directly compare these simple multicellular life cycles when evolved in an environment that selects for large size, which is thought to be one of the earliest benefits of multicellularity[2]. Additionally, this environment has previously been shown to select for clonally-developing snowflake yeast[36].

Isogenic snowflake and floc yeast were evolved in an environment characterized by periods of growth (selection for rapid growth) and settling (selection for large size)[36]. Both snowflake and floc yeast exhibited significant increases in mean settling rate over the course of the experiment. Furthermore, snowflake and floc yeast showed increases in fitness relative to their ancestor over multiple days of competition, but this fitness increase was obtained via two different important life history traits that affect fitness in this system.

Floc yeast increased fitness during growth but not during settling, despite considerable increases in settling rate and flocculation. Conversely, snowflake yeast exhibited elevated fitness during settling, but not during growth. This suggests that simple clonal multicellular life cycles may potentiate a shift in selection from the lower-level units to the emerging, higher-level multicellular individual, an important step during the evolution of multicellularity[14,27,81,102].

5.3 Methods

5.3.1 Strains and media

Table 2 – Strains used in this study.

Strain	**Relevant Genotype**	**Reference**
Snowflake	Δ*ace2*::HYGMX	Pentz *et al*. (2019)
Floc	Δ*ura3::KAN-GAL1*p*::FLO1*	Pentz *et al*. (2019)
Snowflake-GFP	Δ*lys2*::*TEF2*p-yeGFP	Pentz *et al*. (2019)
Floc-GFP	Δ*lys2*::*TEF2*p-yeGFP	Pentz *et al*. (2019)

Strains used in this study are listed in Table 2 and the construction of these is described in Pentz *et al*. (2019). Briefly, flocculant yeast were created by replacing *URA3* ORG with the *KAN-GAL1p::FLO1* cassette[126] and snowflake yeast were created by replacing the *ACE2* ORF with the drug maker, *KANMX*. These genotypes were created from the same homozygous diploid unicellular ancestor (Y55), so these strains differ only in their mode of cluster formation. All experiments were performed in rich medium composed of a mix of glucose and galactose (YPGal+Dex; per liter; 18 g galactose, 2 g

glucose, 20 g peptone, 10 g yeast extract), shaking at 250 rpm at 30°C. These growth conditions yield clusters of similar size after 24 h of growth.

5.3.2 Experimental evolution

Twenty replicate populations of both snowflake and floc yeast were initiated into 10 mL of YPGal+Dex from a single clone and grown overnight. Every 24 h, each population (40 populations total) were subjected to daily selection for settling for 5 minutes on the bench as described in Ratcliff *et al*. (2012). Our selection regime yielded ~3 generations per day for snowflake populations or ~5 generations for floc populations, declining as more biomass was being transferred over evolutionary time (Figure 29). Every 7 days, whole populations where cryogenically stored at -80°C.

5.3.3 Measuring settling rate

To explore the dynamics of multicellular adaptation in floc and snowflake yeast populations, we measured the settling rate of each population every 7 days over the 24-week experiment. We thawed cryogenically-stored whole populations and subsequently inoculated 100µL into 10 mL of YPGal+Dex and grew them for 24 h. Then, 100 µL of overnight cultures was inoculated into fresh YPGal+Dex media and grown for an additional 24 h. We measured the settling rate of populations as described in Pentz *et al*. (2019). Briefly, high-speed high-resolution videos of yeast populations settling in back-illuminated cuvettes were recorded using a Sony a7RII and 90 mm macro lens (24 fps, 3840 x 2160 pixels). Then, custom scripts were used to determine the rate of yeast biomass displacement, or settling rate, based on changes in pixel densities over settling time[147].

5.3.4 Fitness competitions

To determine the fitness of evolved populations, a representative genotype was isolated from each population at the end of the experiment (24 weeks) by three rounds of single-colony selection on YPGal+Dex agarose plates (YPGal+Dex with 15 g/L agarose). Our selection regime is characterized by fluctuating periods of selection for growth and selection for rapid settling, or large size[36,147]. Thus, it is important to measure fitness in both of these important life history traits. To do so, we quantified the fitness of evolved isolates relative to their ancestor over one round of growth and one round of settling selection. Specifically, to initiate competitions, we inoculated 10 mL cultures of YPGal+Dex with isolates from each population as well as a GFP-marked ancestor and grown for 24 h. Then, we mixed each of the evolved isolates in equal volumes with its marked ancestor (floc or snowflake), and 100 µL of this mixture was inoculated into 10 mL of YPGal+Dex to start competitions. Counts of the GFP-tagged ancestor and evolved isolate were obtained via flow cytometry using a CyFlow® Space flow cytometer where GFP and non-GFP can be distinguished using the FL1 fluorescence channel. Prior to running on the flow cytometer, floc competitions were deflocculating using 50 mM EDTA (pH 7). Counts were obtained at time 0 and after 24 h of growth to determine the fitness of the evolved isolate over one period of growth. To measure fitness over one round of settling selection, 2 mL of the overnight mixed culture was aliquoted into a microcentrifuge tube, and 500 µL was used to determine pre-selection counts. The remaining 1.5 mL was used to perform one round of settling selection (5 min on the bench), after which the top 1.4 mL was discarded. The remaining bottom 100 µL was used to determine post-selection counts. Finally, counts were obtained after three days of competition to calculate fitness over

multiple rounds of growth and selection for settling. In all competitions, relative fitness was calculated using the ratio of Malthusian growth parameters[127]. Relative fitness was normalized to the fitness of the ancestral strain for each environment (growth, settling, and three days) used for the experiment.

5.3.5 Genomic DNA preparation

To determine the genetic basis of observed fitness differences, we performed whole-genome sequencing of 24-week evolved isolates and the starting ancestral genotypes. Yeast were streaked out for single colonies from -80°C glycerol stocks. Single colonies were grown overnight in 10 mL YPGal+Dex and genomic DNA was isolated from 1 mL aliquotes using the VWR® Life Science Yeast Genomic DNA Purification Kit (VWR 89492-616, us.vwr.com).

5.3.6 Whole-genome sequencing

DNA libraries were prepared using the NEBNext® Ultra™ II FS DNA Library Prep Kit for Illumina (www.neb.com) and were sequenced on an Illumina HiSeq 2500. Paired-end 150bp reads were used for all samples. Mean coverage across the genome was 200X for evolved isolate DNA and 50X for ancestor DNA.

5.3.7 Sequencing analysis

DNA sequences were quality trimmed using Trimmomatic[148] and then aligned to the S288C reference genome R64-2-1 using the Burrows-Wheeler Aligner[149]. Duplicates were marked using SAMBLASTER[150] then converted to a BAM file, then sorted and indexed. Variants were called using the Genome Analysis Tool Kit (GATK)

HaplotypeCaller[151]. SNPs and INDELs were first filtered based on read depth and quality using vcffilter (https://github.com/vcflib/vcflib). Variants were removed with a read depth less than 10 and a quality score less than 20. Then, bcftools isec (https://github.com/samtools/bcftools) was used to filter out variants shared between the ancestor and evolved isolates, accounting for variants called due to aligning *S. cerevisiae* strain Y55 used in our experiments to the S288C reference genome. Finally, bcftools isec was used again to identify unique variants for each evolved isolate. Variants were manually validated using the Integrated Genomics Viewer (IGV)[152].

5.4 Results

5.4.1 Dynamics of adaption in response to continued selection for settling

Our selection regime involved selection for rapid sedimentation in liquid media daily[36], and continued selection can rapidly lead to cluster-level adaption[36-38], increasing cluster size that accelerates the settling speed of individual clusters[38]. Previously, we have quantified the effect of settling selection on snowflake yeast by using a variety of tools (microscopy, flow cytometry, etc..[36-38]) that cannot be used for floc genotypes because flocs are forming aggregates dynamically. So, we created a method to measure the settling rate (an important life history trait in our system) in real time[47]. We measured the settling speed (see methods) of whole populations of snowflake and floc yeast (20 replicates each) each week over the course of the experiment (Figure 14). In both developmental modes, the settling rate of the population significantly increased after 24-weeks (floc: $F_{23,456}$ = 16.57, p<0.0001; snowflake: $F_{23,456}$ = 13.65, p<0.0001, ANOVA, pairwise differences assessed with Tukey's post-doc HSD with α=0.05) although floc yeast exhibited a larger

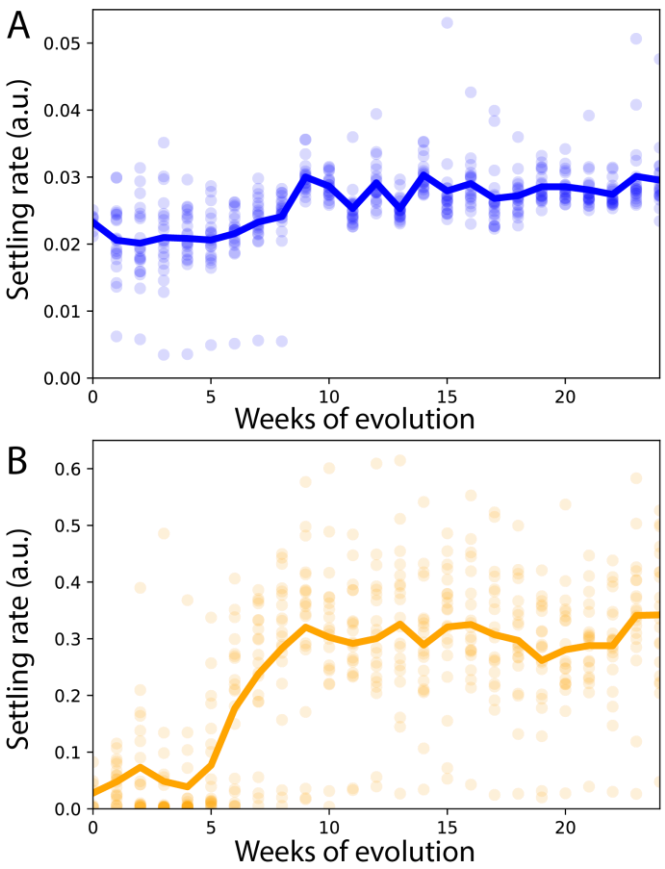

Figure 14 - Dynamics of settling rate adaption. The mean settling rate of 20 replicate populations of snowflake yeast (A) and floc yeast (B) over 24-weeks of continued selection for rapid settling. Both snowflake and floc yeast significantly increased their mean settling rate after 24 weeks, but floc yeast exhibited a larger fold increase relative to snowflake yeast (12-fold vs. 1.3-fold increase, respectively). Each circle represents and independently-evolved population.

fold increase relative to their ancestor than snowflake yeast (12-fold increase vs 1.3-fold increase after 24 weeks, respectively).

Phenotypically, one major driver of increased settling in floc populations may be increased flocculation efficiency, resulting in a larger ratio of floc:planktonic cells (Figure 30). This could lead to increased floc size[111] or a greater number of flocs forming that could survive settling selection, resulting in the dramatic increase in settling rate observed in floc

populations. Snowflake yeast exhibit phenotypic changes consistent with previous experiments: increased cluster size, cells/cluster, and individual cell size[36,38] (Figure 31).

5.4.2 Snowflake and floc yeast gain fitness in two different, important lift history traits

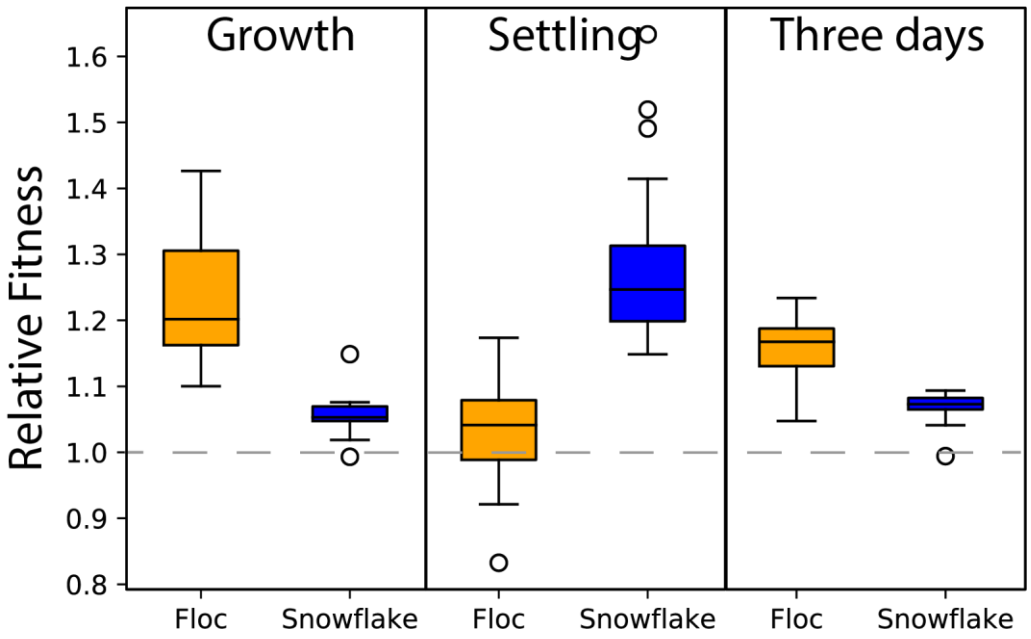

Figure 15 - Fitness of evolved isolates during growth, settling, and three days of competition. We measured the relative fitness of 24-week evolved isolates (20 floc and 20 snowflake) to their GFP-marked ancestor over one growth cycle (first panel), one round of settling selection (second panel), 3-days of competition in their natural environment (periods of growth and settling, third panel). Both floc and snowflake yeast increased their relative fitness over three days of competition. But, floc yeast increased fitness during the growth phase, while snowflake yeast increased fitness during settling. Each box plot represents the distribution of relative fitness for the 20 evolved isolates. For each isolate, we performed four replicate competitions.

Next, we asked how continued selection for large size affects the fitness of floc vs. snowflake yeast in two important life history traits that affect fitness in our system[36,38,45]. Our selection environment is characterized by fluctuating periods of growth and selection for settling, which introduces fitness trade-offs between growth rate and cluster size[36,45]. Trade-offs in life history traits are thought to be particularly important during an ETI[57].

Consequently, in addition to quantifying the fitness of evolved isolates over multiple rounds of competition, we measured the fitness of evolved isolates over one round of growth and one round of settling[47]. Both floc and snowflake populations increased in fitness over three rounds of competition (Figure 15), but this fitness increased is gained in distinct life history traits. Floc yeast increase in fitness during the growth phase of the experiment relative, while no fitness gain was seen over one round of settling (Figure 15; $F_{5,114}$ = 34.4, p<0.0001, pairwise differences assessed with Tukey's post-hoc HSD). Alternatively, fitness increases in snowflake yeast were mostly achieved by increasing fitness during settling as opposed to growth (Figure 15; $F_{5,114}$ = 34.4, p<0.0001, pairwise differences assessed with Tukey's post-hoc HSD). Altogether, this may suggest that selection may be acting more strongly at the group-level in clonally-reproducing snowflake yeast while acting at the single-cell level in aggregative floc yeast. This is a notable result, as this is consistent with evolutionary theory that suggests that clonal development may provide several key advantages to emergent multicellular organisms that may facilitate the evolution of multicellular-level traits[8].

5.4.3 Snowflake yeast accumulate a larger proportion of growth-inhibiting mutants

Whole genome-sequencing revealed that snowflake and floc did not differ significantly in the number of fixed mutations, with an average of 4 or 6 mutations per population, respectively (p=0.06, two-tailed t-test; Figure 16A). However, over the course of our 24-week selection experiment, snowflake yeast underwent ~400 generations of evolution while floc yeast experienced ~700 generations, suggesting that snowflake yeast may be accumulating mutations at a quicker rate than flocculating yeast. All mutants are listed in Table 3 (Snowflake) and Table 4 (Floc).

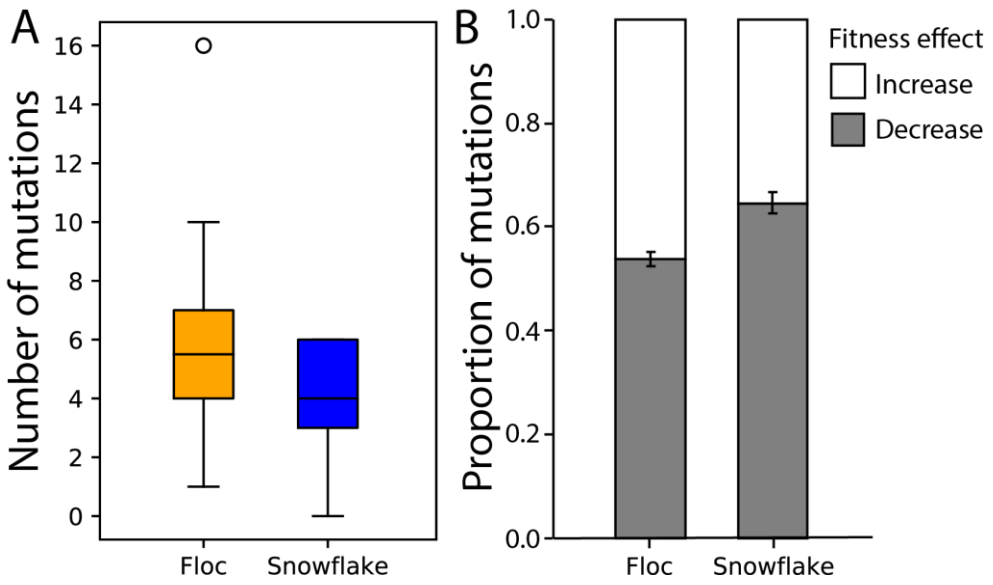

Figure 16 - Whole genome sequencing of 24-week evolved floc and snowflake populations. A) After 24 weeks of selection for fast settling, floc and snowflake do not show significant differences in the number of fixed mutations (p=0.06, two-tailed t-test), but this represents ~700 generations in floc populations and ~400 generations in snowflake populations. B) Snowflake yeast show a larger proportion of mutations that decrease fitness during growth.

During the transition from uni- to multicellularity, selection must shift from the lower to the emergent higher-level unit. The fitness of evolved isolates (Figure 15) may suggest that snowflake yeast might be experiencing selection at the cluster-level, which would lead to mutations that increase fitness at the cluster-level at the expense of fitness at the unicellular-level. So, we examined the fitness effect of each mutation in floc and snowflake yeast by comparing mutations against reported fitness effects of gene-deletion mutants when grown in galactose medium[153]. We found that snowflake had significantly more mutations that decrease fitness during growth (65% in snowflake populations vs. 54% in floc populations). Over time, this could result in the de-Darwinization of cells, reinforcing the multicellular transition[19].

5.5 Discussion

The transition from uni- to multicellularity requires Darwinian properties (*e.g.*, heritable variation in fitness) to emerge at the level of the multicellular group[14,21,24]. This is a daunting task, as selection acting at the level of individual cells may prevent this transition from becoming established, thwarting further multicellular-level adaptation[8,154]. Here, we show that) the properties of simple multicellular life cycles may either promote or hinder the emergence of cluster-level selection in multicellular yeast.

Consistent with previous experiments, clonally-reproducing snowflake yeast exhibited a rapid shift in the level of selection, from the unicell to the multicellular group[36,38,109]. Increases in mean settling rate over 24-weeks of continued selection for large size were likely due to traits previously shown to have cluster-level benefits in snowflake yeast, including larger, more elongated cells[38,155]. Furthermore, evolved isolates showed greater fitness gains during settling, a multicellular-level trait, relative to growth. It is important to note that settling speed might be a function of growth rate in snowflake yeast (faster-growing cells beget larger clusters), but this may not account for the overall increase in fitness during settling. Snowflake yeast had a higher proportion of fixed mutations that exhibit fitness defects during growth, some of which may be beneficial at the group-level. For example, loss-of-function mutations in *AYR1* decrease fitness during growth in galactose[153], but result in elongated buds[156], potentially increasing the packing fraction, and thus the size, of snowflake clusters[155].

Aggregative floc yeast showed a greater fold-increase in mean settling rate over time relative to snowflake yeast, showing signs of increased flocculation efficiency, which leads to faster-settling clusters[157,158]. Yet, evolved isolates showed no fitness gain during settling. In fact, some strains lose to their ancestor over one round of selection for settling.

But, all evolved floc populations showed a significant increase in growth rate relative to their ancestor. This is consistent with the benefits and costs of chimeric aggregates in other systems[84,159]. Chimerism may provide several benefits, such as bringing together beneficial interactions between cells, such as metabolic cross-feeding[160]. Additionally, chimeric aggregates may experience many size-related benefits (*e.g.*, protection from environment stresses)[159,161,162]. In the slime mold *Dictyostelium discoideum*, non-clonal binding may lead to larger aggregate slugs that migrate faster and further when cell density is limited[159], increasing dispersal distance[161]. Our selection environment favors a rapid transition from uni- to multicelluarity due to a fluctuating environment between growth and settling[47]. Indeed, ancestral floc yeast show an advantage over snowflake yeast in monoculture, growing faster and setting more rapidly[147]. Increased growth rates observed in floc populations may increase the efficacy of flocculation, maximizing cell density earlier in the culture cycle. This may lead to increased random interactions between cells, increasing the probability that such interactions will overcome repulsions of like-charges in yeast cells walls, resulting in successful flocculation[111,163], but further experiments are needed to explore this.

However, there are often costs associated with chimerism. Chimeric aggregates are often easily exploited by social 'cheats'. In both *D. discoideum* and the social bacterium *Myxococcus xanthas*, chimerism often leads to developmental cheating where one clone becomes disproportionately represented in the spore population[154,164]. In yeast, flocculation also seems susceptible to cheats that decrease flocculation efficiency, because non-flocculant cells can compete with flocculant cells because FLO proteins bind to mannoproteins that are ubiquitous in yeast cell walls[165]. Thus, although evolved flocculant

yeast seem to be superior aggregators, they are exploitation by less efficient aggregators, decreasing their fitness during settling.

During the transition to multicellularity, the type of multicellular life cycle that could evolve depends on the ecological and historical constraints of the unicellular organism. For example, clonal multicellularity seems to have aquatic origins, while aggregative multicellularity has terrestrial origins[2]. Despite this, we show that in simple isogenic yeast that only differ in their mode of cluster formation, clonal development is superior to aggregation, potentiating an evolutionary transition in individuality from the cells to the emergent multicellular organism. This is an integral step during the transition to multicellularity, increasing the scope for complex, multicellular-level adaptations[8,27,30,57,101,102,105,109]. This work highlights the paramount important of early multicellular life cycles during this major transition in evolution.

APPENDIX A. SUPPLEMENTARY FIGURES AND TABLES

Appendix Movie 1. Individual pixel intensities, termed here Focal Densities (in the [0, 1] interval), were measured over 5 minutes of settling, at 24 frames per second (see Methods section for details). At all frames and for all focal densities, the absolute difference, relative to the first frame, was quantified. The total of these differences (for the entire cuvette at each frame) was expressed as the density change. We show two representative settling dynamics, corresponding to one of the fastest (80:20 Floc:Snowflake) and one of the slowest (10:90 Floc:Snowflake) settling rates. We show a cell-free negative control with no expected density changes (spent media), showing that there is no overall change in the density. The raw data is shown as shadowed lines in the graph. To correct for the noise in the raw data, we applied a Savitzky-Golay filter (see Supplementary Figure 10 for details), corresponding to the dashed lines of the graph.

Appendix Movie 2. 5 minutes of settling in a monoculture of snowflake yeast (left) and floc yeast (right).

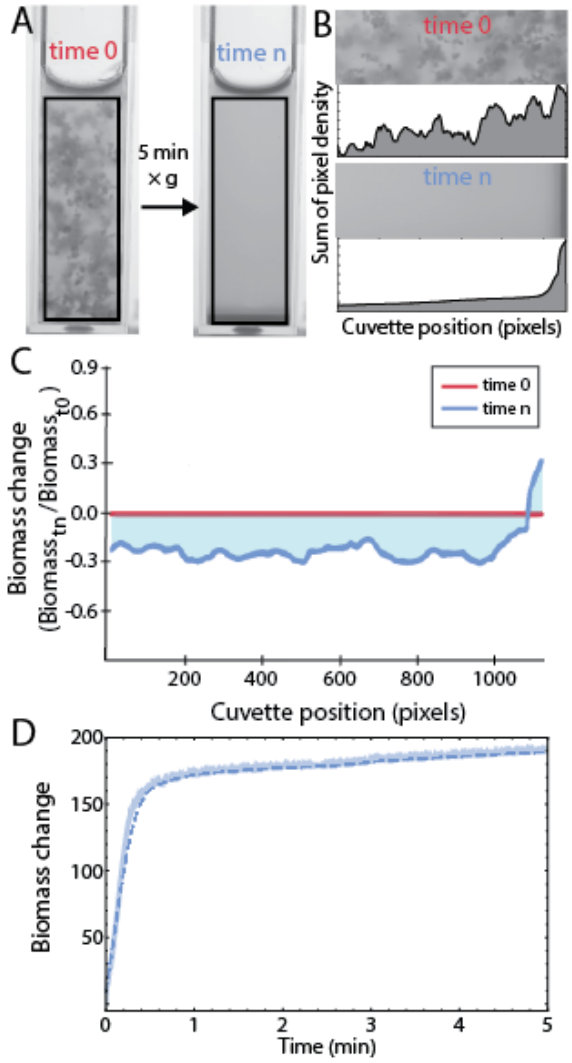

Figure 17 - Measuring settling rates of yeast populations. A) Cuvettes before (time 0) and after settling (time *n*). B) Individual images were extracted from 5 minutes of high resolution video (3840 × 2160 pixels, 24 frames per second), and cropped to remove cuvette borders (black boxes in A). Here, the cuvettes have been rotated 90° clockwise. The sum of the pixel densities across each row was calculated from pixel brightness. C) For each time *n*, the change in pixel density, or biomass, is calculates as the biomass at time *n* relative to time 0 (B) for each row along the length of the cuvette. D) The overall biomass change over the course of the 5 min timelapse is plotted as the sum of absolute biomass change along length of cuvette (the blue shaded area in C)) over time. The settling rate of the population is calculated as the maximum slope of this curve.

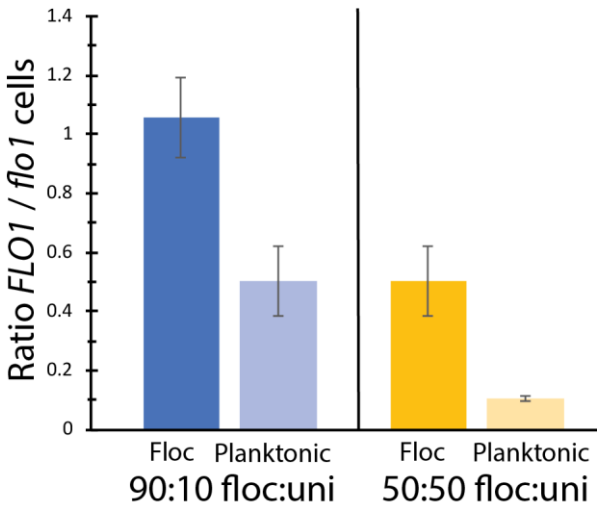

Figure 18 - Non-flocculating unicells are largely excluded from flocs. Shown are the ratio of flocculating (*FLO1*) to non-flocculating (*flo1*) cells in the flocculating and plaktonic subpopulations of co-cultures. These populations were initially inocluated at a ratio of 90:10 *FLO1:flo1* or 50:50 *FLO1:flo1* cells. *flo1* cells are preferentially excluded from flocs. Error bars are standard deviations of three biological replicates.

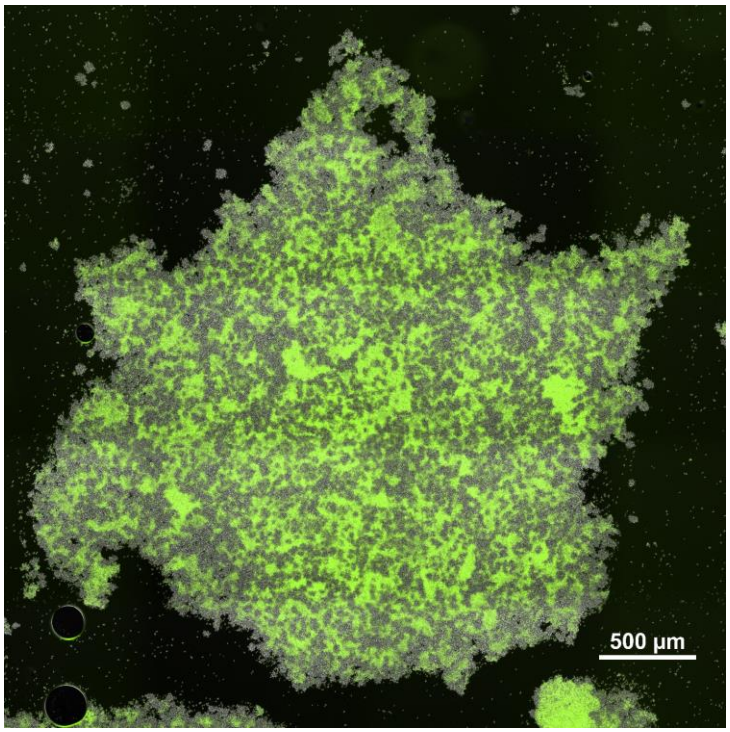

Figure 19 - Large chimeric aggregate from a population that is 70% floc, 30% snowflake yeast.

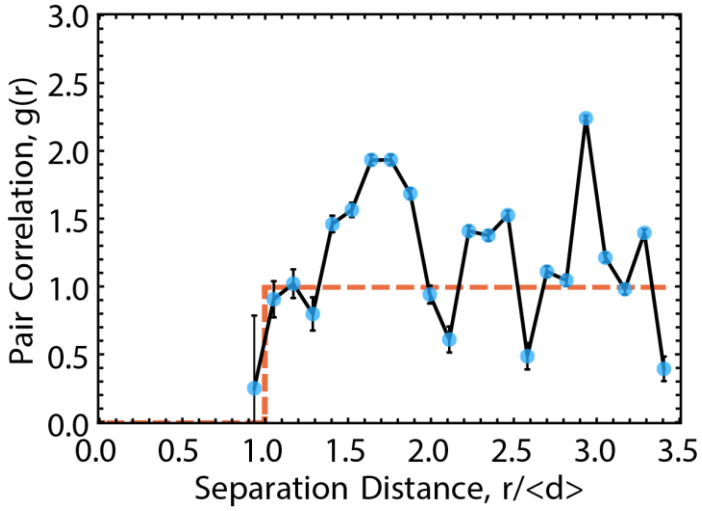

Figure 20 - Pair correlation function measurement. The pair correlation function, *g(r)*, is measured from snowflake yeast cluster positions within aggregates, taken from micrographs. The black line represents the pair correlation function measured in a sample that is 10% snowflake yeast with standard error bars. The red line represents the pair correlation function for a random distribution of clusters.

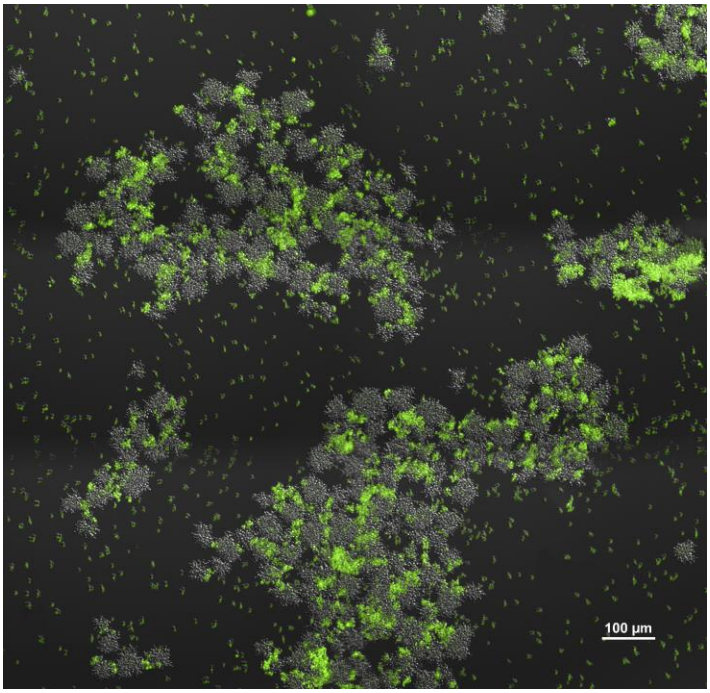

Figure 21 - Floc yeast act as an adhesive, allowing snowflake yeast to form aggregative chimeric aggregates. This is an image from a population with 30% floc, 70% snowflake yeast. All snowflake yeast within chimeric aggregates adhere via small patches of floc cells (labeled with GFP).

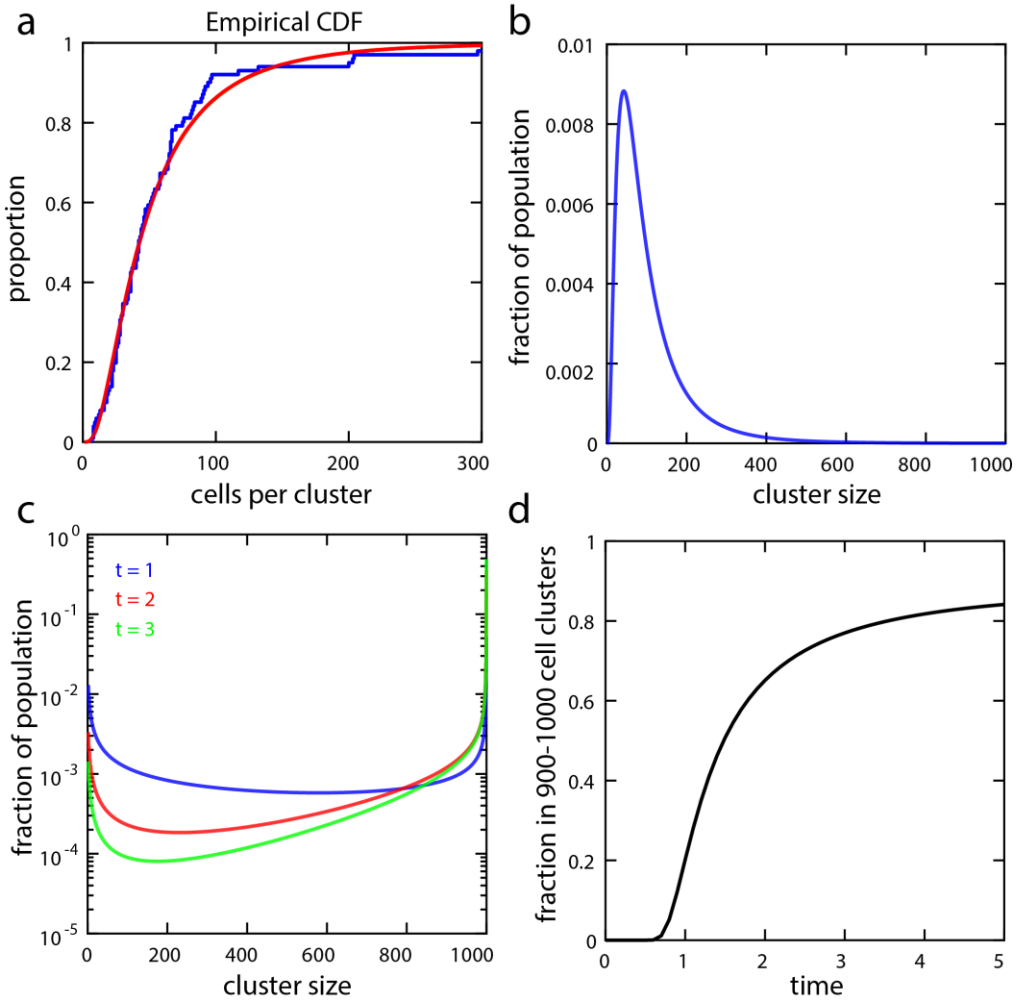

Figure 22 - Distribution of S_i and P_i as a function of *i*. a) The blue curve is the cumulative density function of the distribution of snowflake cluster sizes using empirical data and the red curve is the fit to a lognormal distribution. b) The distribution of snowflake cluster size S_i used in model simulations, which does not change over time. c) Distribution of permissive floc cluster sizes P_i for three different times. As the amount of time for aggregation increases, a greater proportion of the distribution is represented in the largest cluster size fractions. d) The fraction of P_i in the top 10% of possible clusters sizes increases with aggregation time, leveling out for larger *t* as P_i reaches the maximum cluster size.

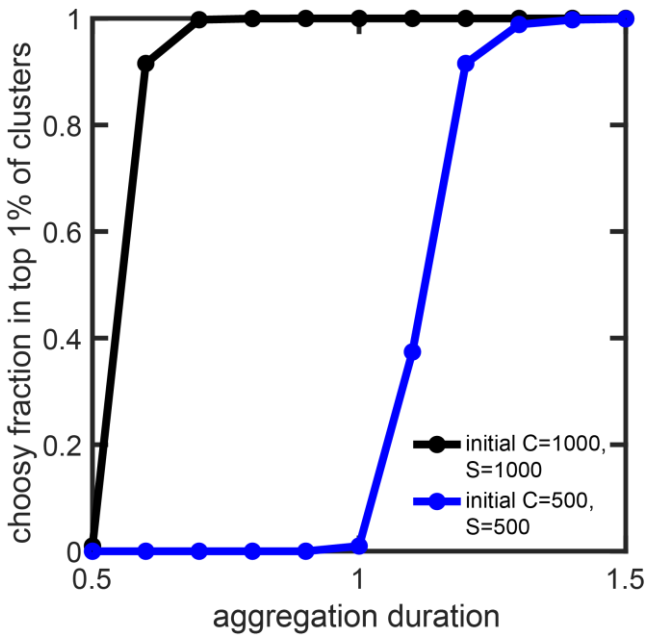

Figure 23 - Effect of population size on the aggregative fitness of choosy floc, C. Choosy floc and snowflake do not co-aggregate during competition. We can thus examine the effect of overall population size on C's aggregation by varying population size (increasing the density of C and probability they will interact and aggregate), and aggregation duration. Higher densities and longer aggregation durations favor choosy floc.

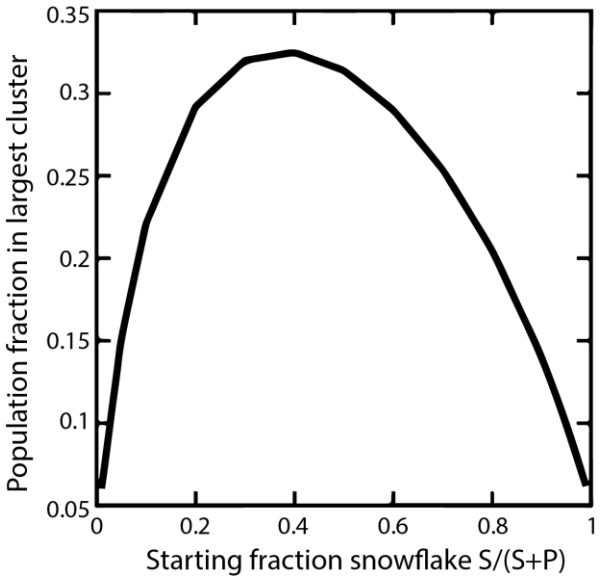

Figure 24 - Populations of snowflake and permissive floc yeast form large, fast-settling aggregates at intermediate frequencies (peaking at 40%). This is similar to experimental data showing a peak settling speed between 30 and 40% S (Figure 10a).

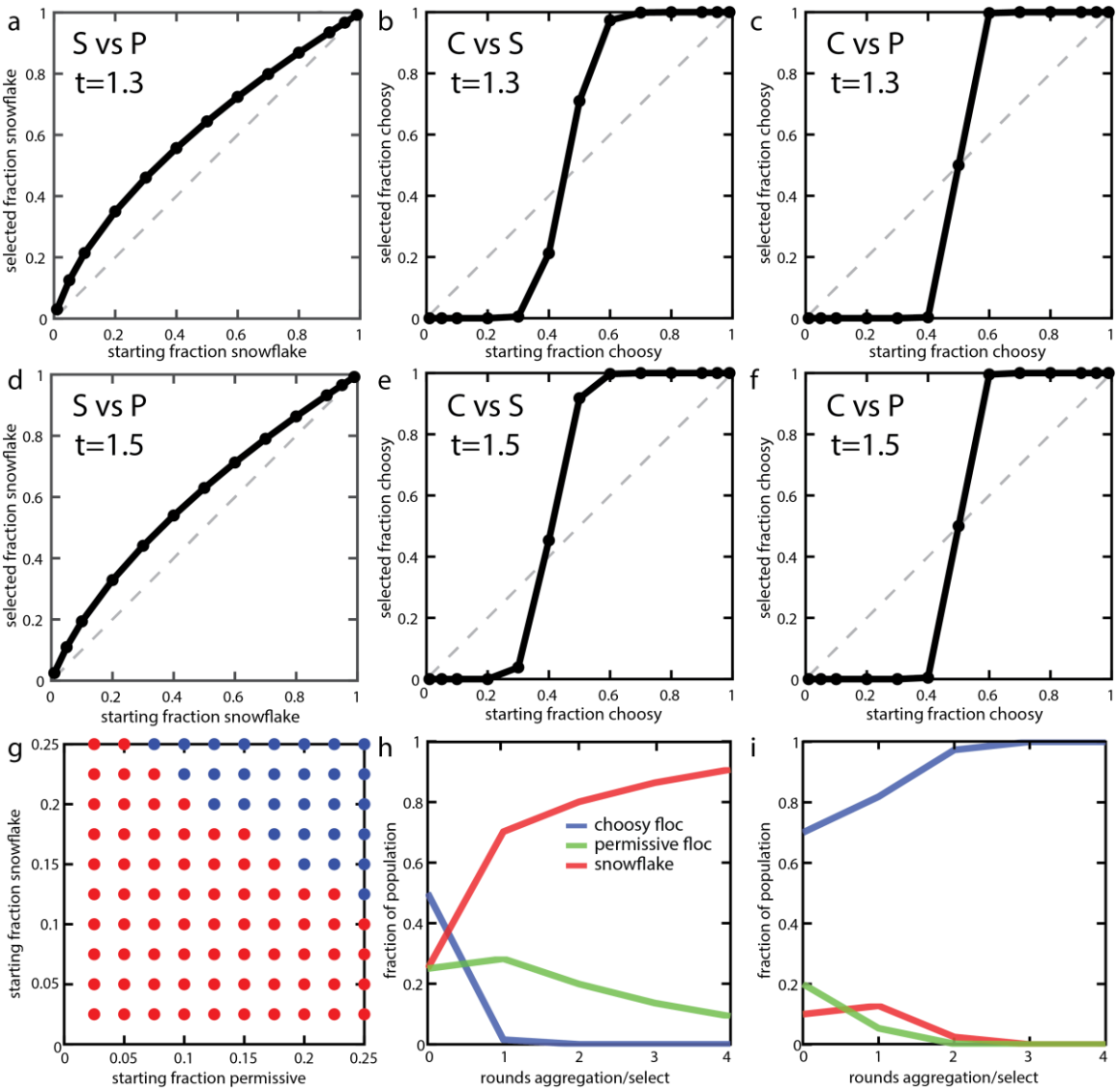

Figure 25 - Effect of aggregation time on model dynamics. We examine the dynamics of all three pairwise competitions between *S*, *P* and *C* as a function of aggregation time; *t*=1.3 (a-c) and *t*=1.5 (d-f). In general, longer aggregation times favor choosy floc, giving it a greater opportunity to form large clonal aggregates. a-f can be compared to Figure 13a-d, where *t*=1. g) Invasion phase diagram for the three-way competition, with *t*=1.5; when red, *C* increases in frequncy. Note that at this time duration, *C* can displace *P*+*S* from a lower starting frequency than when *t*=1 (Figure 6d). h & i) Competition across multiple rounds of settling selection under conditions favoring *C*, namely, long aggregation (*t*=1.5) and strong selection (biomass from the largest groups composing 1% of population's biomass survive settling selection). *C* are unable to displace *P* and *S* when starting at 50% frequency (h), but are when starting at 70% (i). Note also how over the first timestep in (h), snowflake yeast and permissive floc both increase in frequency, then snowflake yeast subsequently parasitizes the floc.

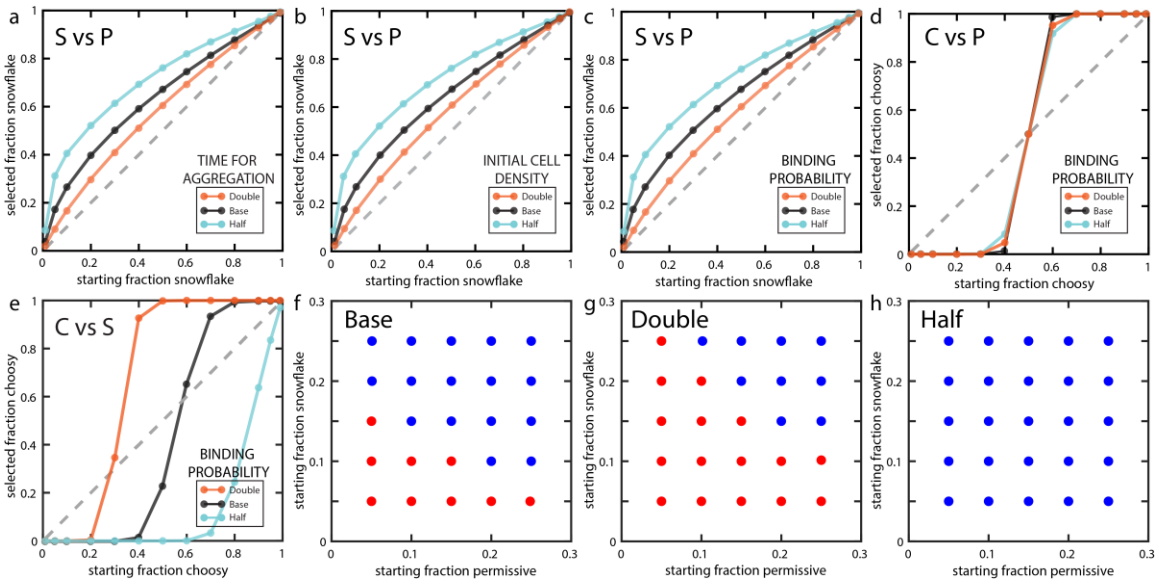

Figure 26 - Parameter sensitivity analysis. We analyzed the dependence of our model on three key parameters (aggregation time, initial cell density, and binding probability) by either doubling or halving each parameter relative to the base calculation (Figure 13). a-c) Increasing aggregation time, initial cell density, or binding probability gives an advantage to *P* cells over *S*, but *S* still increases in relative frequency at all initial starting *S* fractions. Additionally, when each of these parameters are either doubled or halved, the result is the same. Thus, our model has the same sensitivity to each of these parameters. d) *C* vs *P* yields the same result regardless of how the parameters change because *C* and *P* flocs are identical in the absence of *S*. e) *S* cannot be invaded if the binding probability (or aggregation time or initial cell density) is halved, but only wins when it starts at more than 75% of the population when doubled. f-h) Invasion diagram for three-way competition when the binding probability (or aggregation time or initial cell density) is halved or doubled. When red, *C* increases in frequency. If parameters are doubled (g), favoring flocculation types, *S*+*P* still win at some frequencies, but *S*+*P* always win when parameters are halved (h).

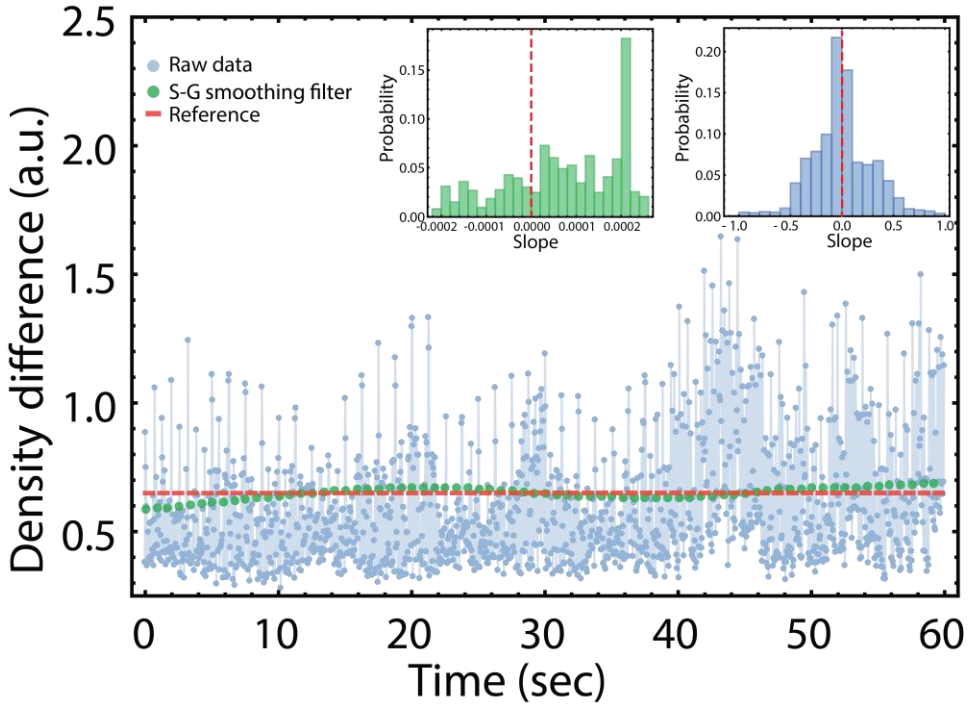

Figure 27 - Processing of raw density data. Raw data (dotted blue line) from our density measurements have small fluctuations (note the range on the y-axis, ~200 times smaller than that of Figure 8B) that nevertheless can be problematic to estimate a characteristic 'maximum slope' of the overall dynamics. We show here a negative control, spent media, where we expect no significant changes in density over time, corresponding to a characteristic slope of zero (red dashed line, intersect in Y at 0.652 corresponds to the average of the dynamics). Using a Savitzky-Golay smoothing function with a window size of 415 (various window sizes give similar results), we were able to recover a distribution of slopes very close (within 2 x 10-4) to what was expected (green histogram), compared to the raw data (blue histogram). This process yielded a high signal-to-noise ratio (28.04 compared to 2.19 of the raw data - defined here as μ/σ, the reciprocal of the coefficient of variation). Therefore we applied the same pre-processing to all our raw density data. This process was implemented with MATLAB built-in functions (MATLAB, MathWorks).

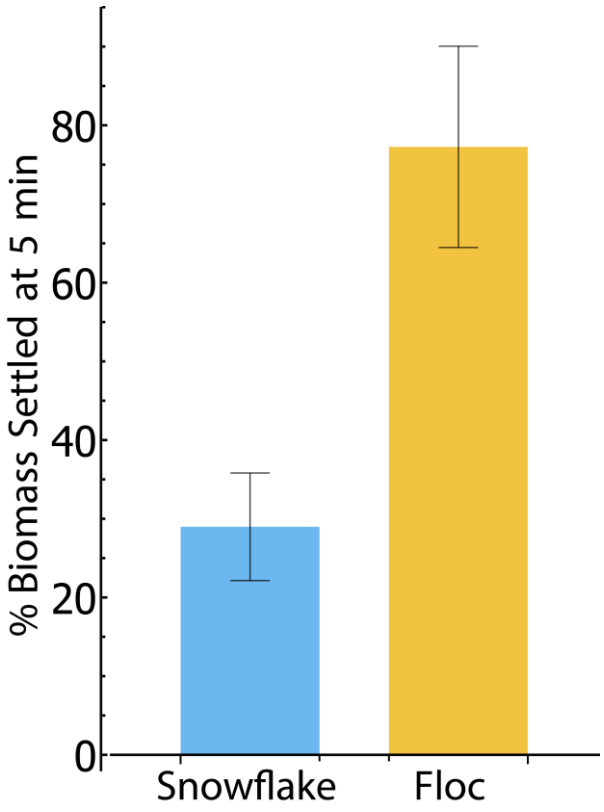

Figure 28 - Biomass measurement of settling survival. We verified that floc yesat settle more rapidly than snowflakes, as determined by our biomass-displacement method (i.e. Figure 8b, Figure 9a; Figure 17; Appendix Movie 1), by directly measuring the amount of yeast biomass to survive the 5 minute settling rate challenge. Floc settles roughly 2.65 times better over 5 minutes of settling than snowflake yeast (77% vs 29% settled, respectively; t5 = 7.44, p=0.0003, two-tailed t-test). Error bars represent standard deviation of five biological replicates.

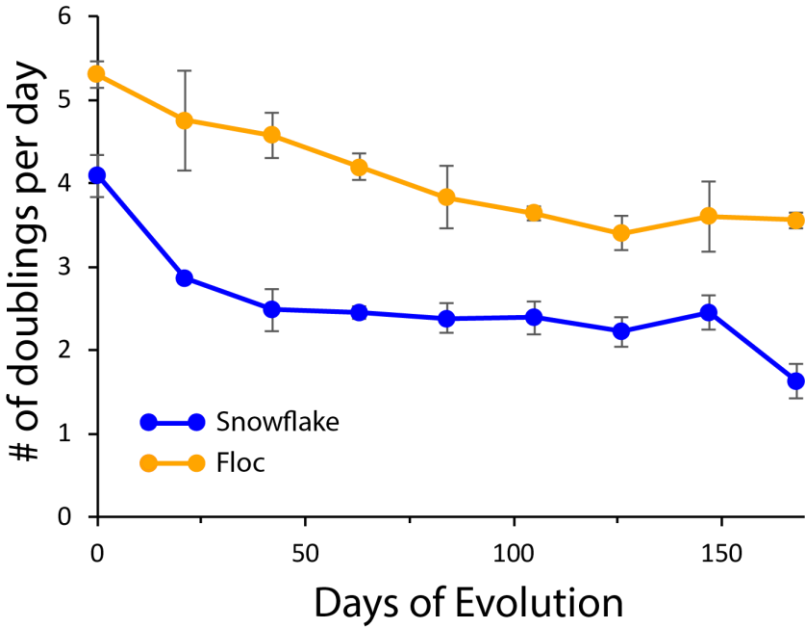

Figure 29 - Number of doublings per day in snowflake and floc yeast over the course of our 24-week selection experiment. To assay the number of generations floc and snowflake yeast experienced over the duration of our selection experiment, we measured the number of doublings per day, quantified as the log2(density after 24h/density post-selection) every three weeks. On average, snowflake and floc yeast experienced 3 or 5 doublings per day, respectively, but this decreased as more biomass was being transferred during the settling selection. Thus, over the course of 24-weeks, snowflake and floc yeast underwent ~400 or ~700 generations, respectively. Error bars represent standard deviation of three biological replicates.

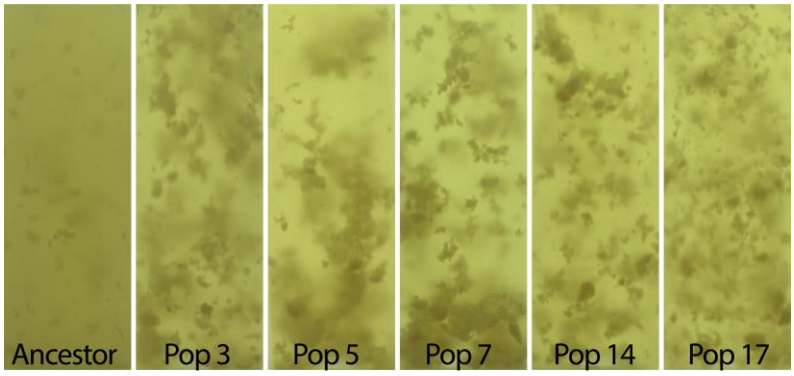

Figure 30 - Floc populations show signs of increased flocculation efficiency. Pictured are initial images of five randomly chosen floc populations from settling videos of 24-week evolved populations, as well as the ancestor. Evolved populations show signs of increased flocculation. There may be less planktonic cells (lower background optical density in evolved populations) as well more cellular biomass in flocs.

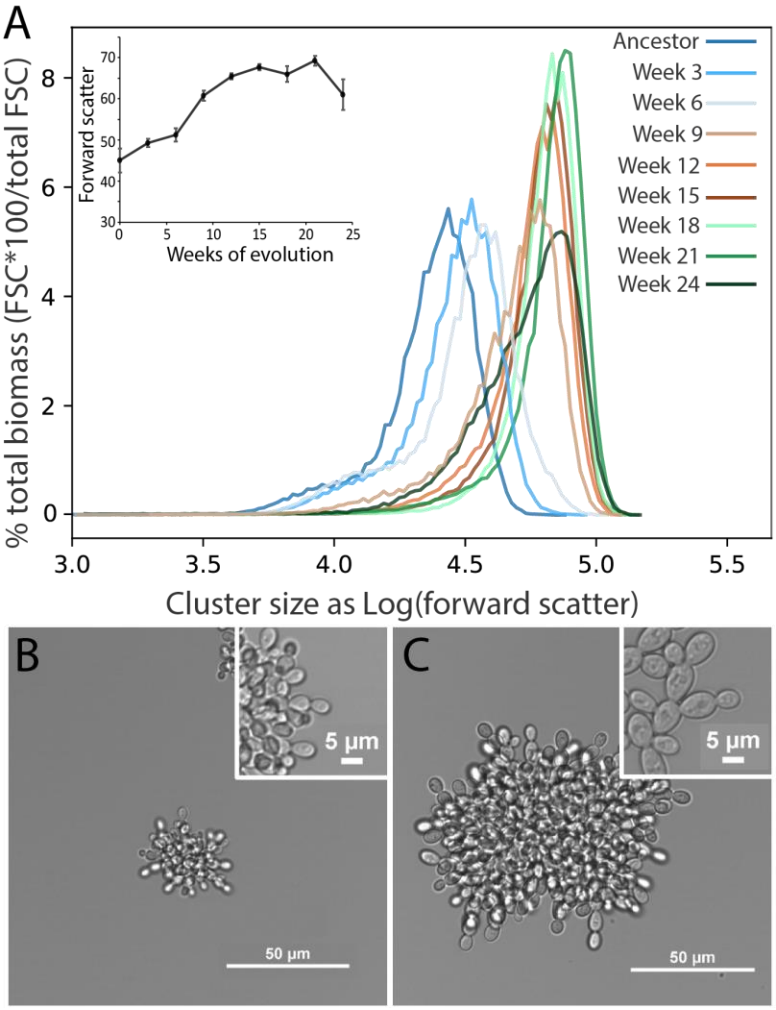

Figure 31 - Larger cluster size evolved in snowflake populations. A) Cluster size as determined by forward scatter on a flow cytometer of randomly chosen representative snowflake population. Over the course of the experiment, the size of snowflake clusters increased, increasing the settling rate of the population. Inset shows average cluster size. Error bars represent standard deviation of three biological replicates. Clusters evolved for 24 weeks (C) shows phenotypic changes from the ancestor (B) consistent with previous selection experiments; larger cluster size, more cells per cluster, and larger, more elongated cells (inset)

Table 3 – Mutations in snowflake populations.

CHROMOSOME	POSITION	REF ALLELE	ALT ALLELE	GENE	EFFECT	IMPACT	POPULATION
II	296313	G	GA	*OLA1*	upstream_gene_variant	MODIFIER	8
II	429297	A	G	*POL30*	upstream_gene_variant	MODIFIER	16
II	639665	C	G	*YBR208C*	missense_variant	MODERATE	19
III	128402	T	A	*SUF2*	upstream_gene_variant	MODIFIER	15
III	238999	C	T	*SED4*	upstream_gene_variant	MODIFIER	8
IV	224252	T	G	*RDI1*	upstream_gene_variant	MODIFIER	10
IV	621894	A	G	*SSS1*	upstream_gene_variant	MODIFIER	20
IV	689746	C	A	*VBA4*	missense_variant	MODERATE	9
IV	771130	A	G	*HOM2*	synonymous_variant	LOW	3
IV	1232488	G	A	*RGA2*	synonymous_variant	LOW	4
IV	1309183	C	A	*HKR1*	missense_variant	MODERATE	7
IV	1517254	G	A	*IRC4*	missense_variant	MODERATE	18
IV	1521621	G	T	*IRC4*	upstream_gene_variant	MODIFIER	6
IX	54713	C	T	*IMP2'*	synonymous_variant	LOW	11
IX	99968	G	GT	*CSM2*	frameshift_variant	HIGH	9
IX	126537	C	G	*AYR1*	missense_variant	MODERATE	16
IX	217265	T	C	*SEC28*	missense_variant	MODERATE	13
IX	308377	C	T	*YIL025C*	missense_variant	MODERATE	8
IX	310944	T	G	*IRR1*	upstream_gene_variant	MODIFIER	19
V	24932	T	G	*HXT13*	upstream_gene_variant	MODIFIER	15
V	94103	T	A	*ECM10*	upstream_gene_variant	MODIFIER	10
V	460550	G	A	*FTR1*	synonymous_variant	LOW	1
VI	6798	G	A	*COS4*	missense_variant	MODERATE	12
VI	222196	ACTGGCGCTGGCG	A	*RRT5*	disruptive_inframe_deletion	MODERATE	13
VII	28131	T	C	*HFM1*	missense_variant	MODERATE	4
VII	106307	T	A	*CHC1*	missense_variant	MODERATE	12
VII	231834	G	A	*TIP20*	missense_variant	MODERATE	4
VII	548606	A	C	*GSC2*	missense_variant	MODERATE	2
VII	1000522	C	T	*PUP2*	upstream_gene_variant	MODIFIER	18
VII	1015174	C	A	*APL6*	missense_variant	MODERATE	3
VIII	140815	C	T	*ARG4*	synonymous_variant	LOW	17
VIII	457627	A	G	*YHR177W*	missense_variant	MODERATE	1
VIII	471179	G	A	*GND1*	missense_variant	MODERATE	12
VIII	497318	T	C	*AIM18*	upstream_gene_variant	MODIFIER	10
X	64948	C	T	*UBP12*	synonymous_variant	LOW	16
X	161252	A	G	*MRS3*	missense_variant	MODERATE	5
X	414949	G	A	*MAD3*	upstream_gene_variant	MODIFIER	12
X	544595	A	C	*CDC8*	synonymous_variant	LOW	4
X	623894	G	C	*ADO1*	missense_variant	MODERATE	13
XI	73310	T	C	*PEX1*	missense_variant	MODERATE	18

XI	161774	A	G	*SRP102*	synonymous_variant	LOW	6
XI	166019	A	T	*RSM22*	upstream_gene_variant	MODIFIER	17
XI	322043	C	G	*YET1*	upstream_gene_variant	MODIFIER	4
XI	368876	T	G	*RGT1*	missense_variant	MODERATE	18
XI	465107	G	T	*TOF2*	upstream_gene_variant	MODIFIER	10
XII	283154	G	T	*RPL10*	missense_variant	MODERATE	1
XII	291205	TTGA	T	*GAL2*	conservative_inframe_deletion	MODERATE	3
XII	412641	A	G	*USB1*	upstream_gene_variant	MODIFIER	19
XII	423609	C	A	*YLR140W*	missense_variant	MODERATE	3
XII	430878	T	C	*ACF2*	missense_variant	MODERATE	19
XII	441099	C	T	*STM1*	synonymous_variant	LOW	4
XII	444668	G	A	*YLR152C*	synonymous_variant	LOW	17
XII	448357	T	G	*YLR152C*	upstream_gene_variant	MODIFIER	6
XII	448371	G	GTTT	*YLR152C*	upstream_gene_variant	MODIFIER	3
XII	448398	GGTT	G	*YLR152C*	upstream_gene_variant	MODIFIER	3
XII	449925	C	A	*ACS2*	upstream_gene_variant	MODIFIER	8
XII	495807	A	G	*SEC10*	synonymous_variant	LOW	18
XII	496434	C	T	*SEC10*	synonymous_variant	LOW	17
XII	497836	G	T	*SEC10*	missense_variant	MODERATE	17
XII	992571	A	C	*CRN1*	synonymous_variant	LOW	8
XIII	160670	C	T	*CMP2*	missense_variant	MODERATE	2
XIII	294231	T	C	*CLU1*	missense_variant	MODERATE	6
XIII	437898	G	T	*YMR085W*	missense_variant	MODERATE	15
XIV	263555	GA	G	*YNL205C*	upstream_gene_variant	MODIFIER	8
XIV	418983	G	A	*YNL109W*	missense_variant	MODERATE	9
XIV	752419	T	C	*YNR065C*	missense_variant	MODERATE	18
XV	56451	G	T	*CTR9*	upstream_gene_variant	MODIFIER	7
XV	112614	A	G	*YOL107W*	synonymous_variant	LOW	11
XV	408648	C	T	*YOR041C*	missense_variant	MODERATE	19
XV	715329	T	A	*SLK19*	stop_lost&splice_region_variant	HIGH	11
XV	741210	C	T	*MGM1*	missense_variant	MODERATE	9
XV	744649	A	G	*SAS5*	synonymous_variant	LOW	6
XV	984420	T	G	*REV1*	missense_variant	MODERATE	20
XVI	38344	C	T	*ACM1*	missense_variant	MODERATE	20
XVI	310399	A	C	*NAN1*	missense_variant	MODERATE	7
XVI	360726	A	G	*ELP4*	missense_variant	MODERATE	17
XVI	376144	G	A	*GLR1*	missense_variant	MODERATE	1
XVI	435991	GAAT	G	*tC(GCA)P1*	upstream_gene_variant	MODIFIER	2
XVI	494650	C	T	*TRM44*	missense_variant	MODERATE	19
XVI	513898	C	G	*ULP1*	missense_variant	MODERATE	7
XVI	617628	T	A	*ATH1*	synonymous_variant	LOW	16
XVI	639549	G	C	*ARP7*	missense_variant	MODERATE	12

Table 4 – Mutations in floc populations.

CHROMOSOME	POSITION	REF ALLELE	ALT ALLELE	GENE	EFFECT	IMPACT	POPULATION
I	203554	C	T	*FLO1*	missense_variant	MODERATE	15
I	203652	A	G	*FLO1*	missense_variant	MODERATE	5
I	203983	A	G	*FLO1*	missense_variant	MODERATE	4
I	203983	A	G	*FLO1*	missense_variant	MODERATE	7
I	204004	C	G	*FLO1*	missense_variant	MODERATE	18
I	206198	C	T	*FLO1*	synonymous_variant	LOW	11
II	40590	C	CAT	*YBL100C*	upstream_gene_variant	MODIFIER	12
II	140113	C	T	*FUI1*	missense_variant	MODERATE	9
II	350224	A	AGTGAAGAGTGGGTGAATTTTGA GATAATTGTTGGGATTCCAT	*PRP6*	upstream_gene_variant	MODIFIER	2
II	372410	G	A	*TIP1*	missense_variant	MODERATE	20
III	297939	C	A	*GIT1*	missense_variant	MODERATE	7
IV	60452	T	C	*WHI4*	upstream_gene_variant	MODIFIER	2
IV	127578	C	T	*VMA1*	synonymous_variant	LOW	11
IV	135093	G	C	*RPL41A*	upstream_gene_variant	MODIFIER	11
IV	205751	G	T	*RPO21*	stop_gained	HIGH	2
IV	208797	G	A	*RPO21*	stop_gained	HIGH	6
IV	209313	A	C	*RPO21*	missense_variant	MODERATE	1
IV	226239	G	A	*CDC53*	missense_variant	MODERATE	3
IV	364879	C	A	*MCH1*	upstream_gene_variant	MODIFIER	20
IV	466800	G	A	*SNQ2*	synonymous_variant	LOW	16
IV	638744	A	G	*GIS1*	missense_variant	MODERATE	15
IV	776219	T	C	*SSY1*	synonymous_variant	LOW	15
IV	778253	C	A	*SSY1*	synonymous_variant	LOW	18
IV	859587	T	C	*MSC2*	missense_variant	MODERATE	8
IV	871328	T	TA	*YDR209C*	frameshift_variant	HIGH	7
IV	952563	A	G	*PRP28*	upstream_gene_variant	MODIFIER	5
IV	952568	C	T	*PRP28*	upstream_gene_variant	MODIFIER	5
IV	967667	C	G	*RMD5*	missense_variant	MODERATE	2
IV	1000466	G	A	*HEL2*	synonymous_variant	LOW	18
IV	1334776	A	ACC	*PPM1*	upstream_gene_variant	MODIFIER	8
IX	20801	G	T	*IMA3*	upstream_gene_variant	MODIFIER	1
IX	28010	G	T	*YIL169C*	upstream_gene_variant	MODIFIER	12
IX	240205	C	T	*RNR3*	synonymous_variant	LOW	15
IX	375791	GATGCGACTGCAA	G,GATGCGACTGCAAATGCGACTGCAA	*DSN1*	disruptive_inframe_insertion	MODERATE	6
IX	389931	C	CGT	*FLO11*	frameshift_variant	HIGH	1
V	68602	G	C	*GLY1*	missense_variant	MODERATE	13
V	122183	T	G	*EAF5*	missense_variant	MODERATE	11
V	293276	C	A	*MOT2*	missense_variant	MODERATE	14
V	293279	G	A	*MOT2*	missense_variant	MODERATE	4

V	414874	G	A	*LCP5*	missense_variant	MODERATE	3
V	549201	G	A	*DMC1*	missense_variant	MODERATE	15
VII	209907	C	T	*ARI1*	missense_variant	MODERATE	4
VII	444974	G	A	*CWH41*	synonymous_variant	LOW	20
VII	447966	C	T	*TRP5*	synonymous_variant	LOW	3
VII	453366	G	C	*STT3*	missense_variant	MODERATE	6
VII	471469	G	A	*PDR1*	missense_variant	MODERATE	20
VIII	66056	C	G	*NPR3*	upstream_gene_variant	MODIFIER	20
VIII	234935	A	C	*YHR069C-A*	missense_variant	MODERATE	5
VIII	441894	G	T	*MTG2*	missense_variant	MODERATE	1
VIII	467593	G	T	*SVP26*	missense_variant	MODERATE	3
X	218779	A	G	*PRM10*	synonymous_variant	LOW	19
X	429348	A	G	*CYR1*	missense_variant	MODERATE	17
X	521104	C	T	*SSC1*	missense_variant	MODERATE	19
X	664912	T	G	*EFM3*	missense_variant	MODERATE	14
XI	147135	C	T	*PIR1*	upstream_gene_variant	MODIFIER	16
XI	153288	C	A	*ELF1*	stop_gained	HIGH	20
XI	239201	A	T	*SEG2*	missense_variant	MODERATE	5
XI	244896	G	A	*GFA1*	missense_variant	MODERATE	14
XI	544795	C	T	*DYN1*	missense_variant	MODERATE	15
XI	569145	G	T	*GPT2*	missense_variant	MODERATE	3
XI	619522	G	A	*RPL40B*	upstream_gene_variant	MODIFIER	8
XII	290572	A	C	*GAL2*	missense_variant	MODERATE	12
XII	290587	G	A	*GAL2*	missense_variant	MODERATE	19
XII	291254	AT	A	*GAL2*	frameshift_variant	HIGH	1
XII	291389	G	A	*GAL2*	missense_variant	MODERATE	3
XII	291478	A	AG	*GAL2*	frameshift_variant	HIGH	2
XII	291484	C	A	*GAL2*	missense_variant	MODERATE	2
XII	291711	C	A	*GAL2*	stop_gained	HIGH	13
XII	320516	G	A	*CSF1*	upstream_gene_variant	MODIFIER	14
XII	377304	G	A	*AVL9*	upstream_gene_variant	MODIFIER	4
XII	410628	T	C	*ACE2*	upstream_gene_variant	MODIFIER	7
XII	608573	C	A	*EST1*	missense_variant	MODERATE	6
XII	700451	T	A	*YLR278C*	missense_variant	MODERATE	4
XII	701491	G	A	*YLR278C*	missense_variant	MODERATE	3
XII	721436	A	T	*YLR290C*	upstream_gene_variant	MODIFIER	10
XII	948370	G	C	*BER1*	start_lost	HIGH	17
XII	948786	A	G	*BER1*	missense_variant	MODERATE	16
XII	993393	C	T	*YLR428C*	upstream_gene_variant	MODIFIER	15
XII	1054249	A	G	*PDP3*	synonymous_variant	LOW	15
XIII	71235	TA	T	*TSL1*	frameshift_variant	HIGH	1
XIII	613883	A	G	*ECM5*	missense_variant	MODERATE	19
XIII	622168	G	GTA	*YMR182W-A*	upstream_gene_variant	MODIFIER	3
XIV	16287	T	A	*AAD14*	missense_variant	MODERATE	17

XIV	21294	T	A	*PEX6*	missense_variant	MODERATE	7
XIV	182479	T	C	*RPA49*	missense_variant	MODERATE	20
XIV	368602	C	A	*NAM9*	missense_variant	MODERATE	19
XIV	490505	C	T	*RNH201*	synonymous_variant	LOW	3
XIV	528578	G	T	*VAC7*	missense_variant	MODERATE	20
XIV	572005	C	T	*YNL034W*	missense_variant	MODERATE	6
XIV	600246	G	A	*YNL018C*	missense_variant	MODERATE	6
XIV	607600	G	T	*HEF3*	missense_variant	MODERATE	4
XIV	681382	G	A	*SSK2*	missense_variant	MODERATE	13
XIV	707777	G	A	*TRM112*	upstream_gene_variant	MODIFIER	2
XIV	780666	C	A	*COS10*	missense_variant	MODERATE	15
XV	29953	C	T	*HPF1*	missense_variant	MODERATE	18
XV	29954	C	T	*HPF1*	synonymous_variant	LOW	18
XV	99957	A	T	*MSN1*	missense_variant	MODERATE	15
XV	100064	A	G	*MSN1*	missense_variant	MODERATE	12
XV	100083	A	T	*MSN1*	missense_variant	MODERATE	16
XV	100107	C	T	*MSN1*	missense_variant	MODERATE	19
XV	288702	G	A	*DIS3*	upstream_gene_variant	MODIFIER	1
XV	332907	T	A	*UTP23*	upstream_gene_variant	MODIFIER	18
XV	412133	T	C	*WHI2*	missense_variant	MODERATE	20
XV	443488	G	C	*YOR062C*	missense_variant	MODERATE	20
XV	443493	A	ATTT	*YOR062C*	disruptive_inframe_insertion	MODERATE	20
XV	613626	C	T	*RPB2*	missense_variant	MODERATE	13
XV	628280	G	A	*ISN1*	upstream_gene_variant	MODIFIER	20
XV	764790	G	T	*HER1*	missense_variant	MODERATE	2
XV	905525	G	T	*SFG1*	missense_variant	MODERATE	2
XV	956308	G	A	*TEA1*	synonymous_variant	LOW	7
XV	1030956	G	A	*MRS6*	synonymous_variant	LOW	20
XV	1059285	A	AC	*FIT2*	upstream_gene_variant	MODIFIER	3
XVI	76565	CA	C	*VIK1*	upstream_gene_variant	MODIFIER	16
XVI	445747	A	G	*YPL060C-A*	upstream_gene_variant	MODIFIER	20
XVI	451866	A	G	*PDR12*	upstream_gene_variant	MODIFIER	20
XVI	578436	G	A	*RPA135*	missense_variant	MODERATE	18
XVI	646246	T	C	*IRC16*	upstream_gene_variant	MODIFIER	20

REFERENCES

1 Bell, G. & Mooers, A. O. Size and complexity among multicellular organisms. *Biological Journal of the Linnean Society* **60**, 345-363 (1997).

2 Bonner, J. T. The origins of multicellularity. *Integrative Biology Issues News and Reviews* **1**, 27-36 (1998).

3 Hedges, S. B., Blair, J. E., Venturi, M. L. & Shoe, J. L. A molecular timescale of eukaryote evolution and the rise of complex multicellular life. *BMC evolutionary biology* **4**, 2 (2004).

4 McShea, D. W. PERSPECTIVE METAZOAN COMPLEXITY AND EVOLUTION: IS THERE A TREND? *Evolution* **50**, 477-492 (1996).

5 McShea, D. W. & Brandon, R. N. *Biology's first law: the tendency for diversity and complexity to increase in evolutionary systems*. (University of Chicago Press, 2010).

6 Smith, J. M. & Szathmary, E. *The major transitions in evolution*. (Oxford University Press, 1997).

7 Valentine, J. W., Collins, A. G. & Meyer, C. P. Morphological complexity increase in metazoans. *Paleobiology* **20**, 131-142 (1994).

8 Grosberg, R. K. & Strathmann, R. R. The evolution of multicellularity: a minor major transition? *Annu. Rev. Ecol. Evol. Syst.* **38**, 621-654 (2007).

9 Knoll, A. H. *Life on a young planet: the first three billion years of evolution on earth*. (Princeton University Press, 2015).

10 Schopf, J. W. Microfossils of the Early Archean Apex chert: new evidence of the antiquity of life. *Science* **260**, 640-646 (1993).

11 Knoll, A. H., Javaux, E. J., Hewitt, D. & Cohen, P. Eukaryotic organisms in Proterozoic oceans. *Philosophical Transactions of the Royal Society of London B: Biological Sciences* **361**, 1023-1038 (2006).

12 Carroll, S. B. Chance and necessity: the evolution of morphological complexity and diversity. *Nature* **409**, 1102-1109 (2001).

13 King, N. The unicellular ancestry of animal development. *Developmental cell* **7**, 313-325 (2004).

14 Buss, L. W. *The evolution of individuality*. (Princeton University Press, 2014).

15 Buss, L. W. Evolution, development, and the units of selection. *Proceedings of the National Academy of Sciences* **80**, 1387-1391 (1983).

16 Frank, S. A. Perspective: repression of competition and the evolution of cooperation. *Evolution* **57**, 693-705 (2003).

17 Michod, R. E. & Roze, D. Cooperation and conflict in the evolution of multicellularity. *Heredity* **86**, 1-7 (2001).

18 Lewontin, R. C. The units of selection. *Annual review of ecology and systematics* **1**, 1-18 (1970).

19 Godfrey-Smith, P. *Darwinian populations and natural selection*. (Oxford University Press, 2009).

20 Okasha, S. *Evolution and the levels of selection*. (Oxford University Press, 2006).

21 Van Gestel, J. & Tarnita, C. E. On the origin of biological construction, with a focus on multicellularity. *Proceedings of the National Academy of Sciences* **114**, 11018-11026 (2017).

22 Bonner, J. T. *Size and cycle: an essay on the structure of biology*. Vol. 2087 (Princeton University Press, 2015).

23 Michod, R. E. *Darwinian dynamics: evolutionary transitions in fitness and individuality*. (Princeton University Press, 2000).

24 Clarke, E. Origins of evolutionary transitions. *Journal of biosciences* **39**, 303-317 (2014).

25 Griesemer, J. The units of evolutionary transition. *Selection* **1**, 67-80 (2001).

26 Jablonka, E. & Lamb, M. J. The evolution of information in the major transitions. *Journal of Theoretical Biology* **239**, 236-246 (2006).

27 Michod, R. E. & Nedelcu, A. M. On the reorganization of fitness during evolutionary transitions in individuality. *Integrative and Comparative Biology* **43**, 64-73 (2003).

28 Douglas, T. E., Queller, D. C. & Strassmann, J. E. Social amoebae mating types do not invest unequally in sexual offspring. *Journal of Evolutionary Biology* **30**, 926-937 (2017).

29 Bonner, J. T. *First signals: the evolution of multicellular development*. (Princeton University Press, 2009).

30 Tarnita, C. E., Taubes, C. H. & Nowak, M. A. Evolutionary construction by staying together and coming together. *Journal of theoretical biology* **320**, 10-22 (2013).

31 Grosberg, R. K. & Strathmann, R. R. One cell, two cell, red cell, blue cell: the persistence of a unicellular stage in multicellular life histories. *Trends in ecology & evolution* **13**, 112-116 (1998).

32 Fisher, R. M., Cornwallis, C. K. & West, S. A. Group formation, relatedness, and the evolution of multicellularity. *Current Biology* **23**, 1120-1125 (2013).

33 Bell, G. & Koufopanou, V. The architecture of the life cycle in small organisms. *Philosophical Transactions of the Royal Society of London B: Biological Sciences* **332**, 81-89 (1991).

34 Hamilton, W. D. The genetical evolution of social behaviour. II. *Journal of theoretical biology* **7**, 17-52 (1964).

35 Michod, R. E., Nedelcu, A. M. & Roze, D. Cooperation and conflict in the evolution of individuality: IV. Conflict mediation and evolvability in Volvox carteri. *BioSystems* **69**, 95-114 (2003).

36 Ratcliff, W. C., Denison, R. F., Borrello, M. & Travisano, M. Experimental evolution of multicellularity. *Proceedings of the National Academy of Sciences* **109**, 1595-1600 (2012).

37 Ratcliff, W. C., Fankhauser, J. D., Rogers, D. W., Greig, D. & Travisano, M. Origins of multicellular evolvability in snowflake yeast. *Nature communications* **6**, 6102 (2015).

38 Ratcliff, W. C., Pentz, J. T. & Travisano, M. Tempo and mode of multicellular adaptation in experimentally evolved Saccharomyces cerevisiae. *Evolution* **67**, 1573-1581 (2013).

39 Pentz, J. T., Taylor, B. P. & Ratcliff, W. C. Apoptosis in snowflake yeast: novel trait, or side effect of toxic waste? *Journal of The Royal Society Interface* **13**, 20160121 (2016).

40 Brückner, S. & Mösch, H.-U. Choosing the right lifestyle: adhesion and development in Saccharomyces cerevisiae. *FEMS microbiology reviews* **36**, 25-58 (2011).

41 Soares, E. V. Flocculation in Saccharomyces cerevisiae: a review. *Journal of applied microbiology* **110**, 1-18 (2011).

42 Bony, M., Thines-Sempoux, D., Barre, P. & Blondin, B. Localization and cell surface anchoring of the Saccharomyces cerevisiae flocculation protein Flo1p. *Journal of bacteriology* **179**, 4929-4936 (1997).

43 Guo, B., Styles, C. A., Feng, Q. & Fink, G. R. A Saccharomyces gene family involved in invasive growth, cell–cell adhesion, and mating. *Proceedings of the National Academy of Sciences* **97**, 12158-12163 (2000).

44 Verstrepen, K., Derdelinckx, G., Verachtert, H. & Delvaux, F. Yeast flocculation: what brewers should know. *Applied microbiology and biotechnology* **61**, 197-205 (2003).

45 Conlin, P. L. & Ratcliff, W. C. Trade-offs drive the evolution of increased complexity in nascent multicellular digital organisms. *Multicellularity: origins and evolution* **131** (2016).

46 Libby, E., Ratcliff, W., Travisano, M. & Kerr, B. Geometry shapes evolution of early multicellularity. *PLoS computational biology* **10**, e1003803 (2014).

47 Pentz, J. T., Travisano, M. & Ratcliff, W. C. in *Fourteenth International Conference on the Synthesis and Simulation of Living Systems.* 550-554.

48 Brunet, T. & King, N. The origin of animal multicellularity and cell differentiation. *Developmental cell* **43**, 124-140 (2017).

49 Duran-Nebreda, S. & Solé, R. Emergence of multicellularity in a model of cell growth, death and aggregation under size-dependent selection. *Journal of The Royal Society Interface* **12**, 20140982 (2015).

50 Ratcliff, W. C., Fankhauser, J. D., Rogers, D. W., Greig, D. & Travisano, M. Origins of multicellular evolvability in snowflake yeast. *Nature communications* **6** (2015).

51 Conlin, P. & Ratcliff, W. C. Trade-offs drive the evolution of increased complexity in nascent multicellular digital organisms. *Vienna Series in Theoretical Biology* (In press).

52 Libby, E. & Rainey, P. B. A conceptual framework for the evolutionary origins of multicellularity. *Physical biology* **10**, 035001 (2013).

53 Smith, J. M. & Szathmáry, E. *The Major Transitions in Evolution*. (Oxford University Press, 1995).

54 Shelton, D. E. & Michod, R. E. Group selection and group adaptation during a major evolutionary transition: insights from the evolution of multicellularity in the volvocine algae. *Biological Theory* **9**, 452-469 (2014).

55 Bourke, A. F. *Principles of social evolution*. (Oxford University Press Oxford, 2011).

56 Hammerschmidt, K., Rose, C. J., Kerr, B. & Rainey, P. B. Life cycles, fitness decoupling and the evolution of multicellularity. *Nature* **515**, 75 (2014).

57 Michod, R. E., Viossat, Y., Solari, C. A., Hurand, M. & Nedelcu, A. M. Life-history evolution and the origin of multicellularity. *Journal of theoretical Biology* **239**, 257-272 (2006).

58 Oud, B. *et al.* Genome duplication and mutations in ACE2 cause multicellular, fast-sedimenting phenotypes in evolved Saccharomyces cerevisiae. *Proceedings of the National Academy of Sciences* **110**, E4223-E4231 (2013).

59 Voth, W. P., Olsen, A. E., Sbia, M., Freedman, K. H. & Stillman, D. J. ACE2, CBK1, and BUD4 in budding and cell separation. *Eukaryotic cell* **4**, 1018-1028 (2005).

60 King, L. & Butler, G. Ace2p, a regulator of CTS1 (chitinase) expression, affects pseudohyphal production in Saccharomyces cerevisiae. *Current genetics* **34**, 183-191 (1998).

61 Gietz, R. D., Schiestl, R. H., Willems, A. R. & Woods, R. A. Studies on the transformation of intact yeast cells by the LiAc/ss-DNA/PEG procedure. *Yeast* **11**, 355-360 (1995).

62 Madeo, F. *et al.* Oxygen stress: a regulator of apoptosis in yeast. *The Journal of cell biology* **145**, 757-767 (1999).

63 Mortimer, R. K. & Johnston, J. R. Life span of individual yeast cells. (1959).

64 Gillespie, D. T. Exact stochastic simulation of coupled chemical reactions. *The journal of physical chemistry* **81**, 2340-2361 (1977).

65 Herker, E. *et al.* Chronological aging leads to apoptosis in yeast. *The Journal of cell biology* **164**, 501-507 (2004).

66 Laun, P. *et al.* Aged mother cells of Saccharomyces cerevisiae show markers of oxidative stress and apoptosis. *Molecular microbiology* **39**, 1166-1173 (2001).

67 Newman, S. A. & Bhat, R. Dynamical patterning modules: physico-genetic determinants of morphological development and evolution. *Physical Biology* **5**, 015008 (2008).

68 Drasdo, D., Hoehme, S. & Block, M. On the role of physics in the growth and pattern formation of multi-cellular systems: What can we learn from individual-cell based models? *Journal of Statistical Physics* **128**, 287-345 (2007).

69 Branda, S. S., González-Pastor, J. E., Ben-Yehuda, S., Losick, R. & Kolter, R. Fruiting body formation by Bacillus subtilis. *Proceedings of the National Academy of Sciences* **98**, 11621-11626 (2001).

70 Čáp, M., Štěpánek, L., Harant, K., Váchová, L. & Palková, Z. Cell differentiation within a yeast colony: metabolic and regulatory parallels with a tumor-affected organism. *Molecular cell* **46**, 436-448 (2012).

71 Kaiser, D. Coupling cell movement to multicellular development in myxobacteria. *Nature Reviews Microbiology* **1**, 45 (2003).

72 Miller, M. B. & Bassler, B. L. Quorum sensing in bacteria. *Annual Reviews in Microbiology* **55**, 165-199 (2001).

73 Váchová, L., Kučerová, H., Devaux, F., Úlehlová, M. & Palková, Z. Metabolic diversification of cells during the development of yeast colonies. *Environmental microbiology* **11**, 494-504 (2009).

74 Smith, J. M. & Szathmary, E. *The major transitions in evolution*. (Oxford University Press, 1995).

75 Grosberg, R. K. & Strathmann, R. R. The evolution of multicellularity: a minor major transition? *Annual Review of Ecology, Evolution, and Systematics*, 621-654 (2007).

76 Boraas, M. E., Seale, D. B. & Boxhorn, J. E. Phagotrophy by a flagellate selects for colonial prey: a possible origin of multicellularity. *Evolutionary Ecology* **12**, 153-164 (1998).

77 Kirk, D. L. A twelve-step program for evolving multicellularity and a division of labor. *BioEssays* **27**, 299-310 (2005).

78 Pfeiffer, T. & Bonhoeffer, S. An evolutionary scenario for the transition to undifferentiated multicellularity. *Proceedings of the National Academy of Sciences* **100**, 1095-1098 (2003).

79 Damuth, J. & Heisler, I. L. Alternative formulations of multilevel selection. *Biology and Philosophy* **3**, 407-430 (1988).

80 Godfrey-Smith, P. Darwinian individuals. *From groups to individuals* (2013).

81 Michod, R. E. On the transfer of fitness from the cell to the multicellular organism. *Biology and Philosophy* **20**, 967-987 (2005).

82 Herron, M. D. & Michod, R. E. Evolution of complexity in the volvocine algae: transitions in individuality through Darwin's eye. *Evolution* **62**, 436-451 (2008).

83 Olive, L. S. *The Mycetozoans*. (Academic Press Inc., 1975).

84 Velicer, G. J. & Vos, M. Sociobiology of the myxobacteria. *Annual review of microbiology* **63**, 599-623 (2009).

85 Hamilton, W. D. The genetical evolution of social behaviour. I. *Journal of theoretical biology* **7**, 1-16 (1964).

86 Bonner, J. On the origin of differentiation. *Journal of biosciences* **28**, 523-528 (2003).

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

88 Smukalla, S. *et al.* < i> FLO1</i> Is a Variable Green Beard Gene that Drives Biofilm-like Cooperation in Budding Yeast. *Cell* **135**, 726-737 (2008).

89 Lenski, R. E., Rose, M. R., Simpson, S. C. & Tadler, S. C. Long-term experimental evolution in Escherichia coli. I. Adaptation and divergence during 2, 000 generations. *American Naturalist* **138**, 1315-1341 (1991).

90 Koschwanez, J. H., Foster, K. R. & Murray, A. W. Improved use of a public good selects for the evolution of undifferentiated multicellularity. *Elife* **2** (2013).

91 Rebolleda-Gomez, M., Ratcliff, W. & Travisano, M. in *Artificial Life*. 99-104.

92 Diggle, S. P., Griffin, A. S., Campbell, G. S. & West, S. A. Cooperation and conflict in quorum-sensing bacterial populations. *Nature* **450**, 411-414 (2007).

93 Willensdorfer, M. On the evolution of differentiated multicellularity. *Evolution* **63**, 306-323 (2009).

94 Szathmáry, E., Calcott, B. & Sterelny, K. *The major transitions in evolution revisited*. (2011).

95 Libby, E., Ratcliff, W., Travisano, M. & Kerr, B. Geometry shapes evolution of early multicellularity. *arXiv preprint arXiv:1403.7556* (2014).

96 Hu, C., Bai, F. & An, L. Effect of flocculence of a self-flocculating yeast on its tolerance to ethanol and the mechanism. *Sheng wu gong cheng xue bao= Chinese journal of biotechnology* **21**, 123-128 (2005).

97 Lachance, M.-A. Yeast selection in nature. *inYeast Strain-Selection*, 21-41 (1990).

98 Bonner, J. T. *The evolution of complexity by means of natural selection*. (Princeton University Press, 1988).

99 Leigh, E. G. How does selection reconcile individual advantage with the good of the group? *Proceedings of the National Academy of Sciences* **74**, 4542-4546 (1977).

100 Maynard Smith, J. Evolutionary progress and levels of selection. *Evolutionary progress* **219**, 230 (1988).

101 Szathmáry, E. & Smith, J. M. *Major Transitions in Evolution*. (Oxford University Press Oxford, 1997).

102 Michod, R. E. Evolution of individuality during the transition from unicellular to multicellular life. *Proceedings of the National Academy of Sciences* **104**, 8613-8618 (2007).

103 Kaiser, D. Building a multicellular organism. *Annual review of genetics* **35**, 103-123 (2001).

104 Robinson, R. W. Life cycles in the methanogenic archaebacterium Methanosarcina mazei. *Applied and environmental microbiology* **52**, 17-27 (1986).

105 Knoll, A. H. The multiple origins of complex multicellularity. *Annual Review of Earth and Planetary Sciences* **39**, 217-239 (2011).

106 Bonner, J. T. *First Signals: The Evolution of Multicellular Development*. (Princeton University Press, 2000).

107 Du, Q., Kawabe, Y., Schilde, C., Chen, Z.-h. & Schaap, P. The evolution of aggregative multicellularity and cell–cell communication in the Dictyostelia. *Journal of molecular biology* **427**, 3722-3733 (2015).

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

109 Ratcliff, W. C., Herron, M., Conlin, P. L. & Libby, E. Nascent life cycles and the emergence of higher-level individuality. *Phil. Trans. R. Soc. B* **372**, 20160420 (2017).

110 Chisholm, R. L. & Firtel, R. A. Insights into morphogenesis from a simple developmental system. *Nature reviews Molecular cell biology* **5**, 531-541 (2004).

111 Stratford, M. & Keenan, M. Yeast flocculation: quantification. *Yeast* **4**, 107-115 (1988).

112 Bonner, J. T. *The cellular slime moulds*. Vol. 2nd ed (Princeton University Press, 1967).

113 Kaiser, D., Manoil, C. & Dworkin, M. Myxobacteria: cell interactions, genetics, and development. *Annual Reviews in Microbiology* **33**, 595-639 (1979).

114 Ispolatov, I., Ackermann, M. & Doebeli, M. Division of labour and the evolution of multicellularity. *Proceedings of the Royal Society of London B: Biological Sciences* **279**, 1768-1776 (2012).

115 Boles, B. R., Thoendel, M. & Singh, P. K. Self-generated diversity produces "insurance effects" in biofilm communities. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 16630-16635 (2004).

116 Pande, S. & Velicer, G. J. Chimeric Synergy in Natural Social Groups of a Cooperative Microbe. *Current Biology* (2018).

117 Visick, K. L., Foster, J., Doino, J., McFall-Ngai, M. & Ruby, E. G. Vibrio fischeri lux genes play an important role in colonization and development of the host light organ. *Journal of Bacteriology* **182**, 4578-4586 (2000).

118 Kiers, E. T., Rousseau, R. A., West, S. A. & Denison, R. F. Host sanctions and the legume–rhizobium mutualism. *Nature* **425**, 78-81 (2003).

119 Bright, M. & Bulgheresi, S. A complex journey: transmission of microbial symbionts. *Nature Reviews Microbiology* **8**, 218 (2010).

120 De Monte, S. & Rainey, P. B. Nascent multicellular life and the emergence of individuality. *Journal of biosciences* **39**, 237-248 (2014).

121 Amado, A., Batista, C. & Campos, P. R. A theoretical approach to the size-complexity rule. *Evolution* **72**, 18-29 (2018).

122 Garcia, T., Doulcier, G. & De Monte, S. The evolution of adhesiveness as a social adaptation. *Elife* **4**, e08595 (2015).

123 Biernaskie, J. M. & West, S. A. in *Proc. R. Soc. B.* 20151075 (The Royal Society).

124 Bastiaans, E., Debets, A. J. & Aanen, D. K. Experimental evolution reveals that high relatedness protects multicellular cooperation from cheaters. *Nature communications* **7** (2016).

125 Kuzdzal-Fick, J. J., Fox, S. A., Strassmann, J. E. & Queller, D. C. High relatedness is necessary and sufficient to maintain multicellularity in Dictyostelium. *Science* **334**, 1548-1551 (2011).

126 Smukalla, S. *et al.* FLO1 is a variable green beard gene that drives biofilm-like cooperation in budding yeast. *Cell* **135**, 726-737 (2008).

127 Lenski, R. E., Rose, M. R., Simpson, S. C. & Tadler, S. C. Long-term experimental evolution in Escherichia coli. I. Adaptation and divergence during 2,000 generations. *The American Naturalist* **138**, 1315-1341 (1991).

128 Strassmann, J. E., Gilbert, O. M. & Queller, D. C. Kin discrimination and cooperation in microbes. *Annual review of microbiology* **65**, 349-367 (2011).

129 Kalla, S. E., Queller, D. C., Lasagni, A. & Strassmann, J. E. Kin discrimination and possible cryptic species in the social amoeba Polysphondylium violaceum. *BMC evolutionary biology* **11**, 31 (2011).

130 Vos, M. & Velicer, G. J. Social conflict in centimeter-and global-scale populations of the bacterium Myxococcus xanthus. *Current Biology* **19**, 1763-1767 (2009).

131 Stefanic, P., Kraigher, B., Lyons, N. A., Kolter, R. & Mandic-Mulec, I. Kin discrimination between sympatric Bacillus subtilis isolates. *Proceedings of the National Academy of Sciences* **112**, 14042-14047 (2015).

132 Kim, S. K. & Kaiser, D. C-factor has distinct aggregation and sporulation thresholds during Myxococcus development. *Journal of bacteriology* **173**, 1722-1728 (1991).

133 Ostrowski, E. A., Katoh, M., Shaulsky, G., Queller, D. C. & Strassmann, J. E. Kin discrimination increases with genetic distance in a social amoeba. *PLoS biology* **6**, e287 (2008).

134 Rendueles, O. *et al.* Rapid and widespread de novo evolution of kin discrimination. *Proceedings of the National Academy of Sciences* **112**, 9076-9081 (2015).

135 Vos, M. & Velicer, G. J. Genetic population structure of the soil bacterium Myxococcus xanthus at the centimeter scale. *Applied and environmental microbiology* **72**, 3615-3625 (2006).

136 Foster, K. R., Fortunato, A., Strassmann, J. E. & Queller, D. C. The costs and benefits of being a chimera. *Proceedings of the Royal Society of London B: Biological Sciences* **269**, 2357-2362 (2002).

137 Nanjundiah, V. & Sathe, S. Social selection and the evolution of cooperative groups: the example of the cellular slime moulds. *Integrative biology* **3**, 329-342 (2011).

138 Strassmann, J. E., Zhu, Y. & Queller, D. C. Altruism and social cheating in the social amoeba Dictyostelium discoideum. *Nature* **408**, 965-967 (2000).

139 Noh, S., Geist, K. S., Tian, X., Strassmann, J. E. & Queller, D. C. Genetic signatures of microbial altruism and cheating in social amoebas in the wild. *Proceedings of the National Academy of Sciences*, doi:10.1073/pnas.1720324115 (2018).

140 Verstrepen, K. J., Jansen, A., Lewitter, F. & Fink, G. R. Intragenic tandem repeats generate functional variability. *Nature genetics* **37**, 986-990 (2005).

141 Soares, E. V. & Mota, M. Quantification of yeast flocculation. *Journal of the Institute of Brewing* **103**, 93-98 (1997).

142 Bayly, J. C., Douglas, L. M., Pretorius, I. S., Bauer, F. F. & Dranginis, A. M. Characteristics of Flo11-dependent flocculation in Saccharomyces cerevisiae. *FEMS yeast research* **5**, 1151-1156 (2005).

143 Govender, P., Domingo, J. L., Bester, M. C., Pretorius, I. S. & Bauer, F. F. Controlled expression of the dominant flocculation genes FLO1, FLO5, and FLO11 in Saccharomyces cerevisiae. *Applied and environmental microbiology* **74**, 6041-6052 (2008).

144 Di Gianvito, P., Tesnière, C., Suzzi, G., Blondin, B. & Tofalo, R. FLO 5 gene controls flocculation phenotype and adhesive properties in a Saccharomyces cerevisiae sparkling wine strain. *Scientific reports* **7**, 10786 (2017).

145 Hope, E. A. *et al.* Experimental evolution reveals favored adaptive routes to cell aggregation in yeast. *Genetics* **206**, 1153-1167 (2017).

146 Crespi, B. J. The evolution of social behavior in microorganisms. *Trends in ecology & evolution* **16**, 178-183 (2001).

147 Pentz, J. T., Márquez-Zacarías, P., Yunker, P. J., Libby, E. & Ratcliff, W. C. Ecological advantages and evolutionary limitations of aggregative multicellular development. *bioRxiv*, 255307 (2019).

148 Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114-2120 (2014).

149 Li, H. & Durbin, R. Fast and accurate long-read alignment with Burrows–Wheeler transform. *Bioinformatics* **26**, 589-595 (2010).

150 Faust, G. G. & Hall, I. M. SAMBLASTER: fast duplicate marking and structural variant read extraction. *Bioinformatics* **30**, 2503-2505 (2014).

151 McKenna, A. *et al.* The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome research* **20**, 1297-1303 (2010).

152 Robinson, J. T. *et al.* Integrative genomics viewer. *Nature biotechnology* **29**, 24 (2011).

153 Giaever, G. *et al.* Functional profiling of the Saccharomyces cerevisiae genome. *nature* **418**, 387 (2002).

154 Dao, D. N., Kessin, R. H. & Ennis, H. L. Developmental cheating and the evolutionary biology of Dictyostelium and Myxococcus. *Microbiology* **146**, 1505-1512 (2000).

155 Jacobeen, S. *et al.* Cellular packing, mechanical stress and the evolution of multicellularity. *Nature Physics* **14**, 286 (2018).

156 Watanabe, M., Watanabe, D., Nogami, S., Morishita, S. & Ohya, Y. Comprehensive and quantitative analysis of yeast deletion mutants defective in apical and isotropic bud growth. *Current genetics* **55**, 365-380 (2009).

157 Hope, E. A. & Dunham, M. J. Ploidy-regulated variation in biofilm-related phenotypes in natural isolates of Saccharomyces cerevisiae. *G3: Genes, Genomes, Genetics* **4**, 1773-1786 (2014).

158 Van Mulders, S. E. *et al.* Phenotypic diversity of Flo protein family-mediated adhesion in Saccharomyces cerevisiae. *FEMS yeast research* **9**, 178-190 (2009).

159 Foster, K. R., Fortunato, A., Strassmann, J. E. & Queller, D. C. The costs and benefits of being a chimera. *Proceedings of the Royal Society of London. Series B: Biological Sciences* **269**, 2357-2362 (2002).

160 Shitut, S., Ahsendorf, T., Pande, S., Egbert, M. & Kost, C. Nanotube-mediated cross-feeding couples the metabolism of interacting bacterial cells. *Environmental microbiology* **21**, 1306-1320 (2019).

161 Bonner, J. T., Koontz Jr, P. G. & Paton, D. Size in relation to the rate of migration in the slime mold Dictyostelium discoideum. *Mycologia* **45**, 235-240 (1953).

162 Lyons, N. A. & Kolter, R. On the evolution of bacterial multicellularity. *Current opinion in microbiology* **24**, 21-28 (2015).

163 Stratford, M. & Keenan, M. Yeast flocculation: kinetics and collision theory. *Yeast* **3**, 201-206 (1987).

164 Strassmann, J. E., Zhu, Y. & Queller, D. C. Altruism and social cheating in the social amoeba Dictyostelium discoideum. *Nature* **408**, 965 (2000).

165 Veelders, M. *et al.* Structural basis of flocculin-mediated social behavior in yeast. *Proceedings of the National Academy of Sciences* **107**, 22511-22516 (2010).