# *De novo* evolution of macroscopic multicellularity

**G. Ozan Bozdag [1,†,*], Seyed Alireza Zamani-Dahaj[2,3,*], Penelope C. Kahn[1,4], Thomas C. Day[3], Kai Tong[1,2], Aishwarya H. Balwani[5], Eva L. Dyer[5], Peter J. Yunker[2,†] and William C. Ratcliff[1,†]**

[*]These authors contributed equally to this work.

[†]Corresponding authors: ozan.bozdag@gmail.com, ratcliff@gatech.edu, peter.yunker@gatech.edu

[1]School of Biological Sciences, Georgia Institute of Technology, Atlanta, GA, USA

[2]Interdisciplinary Graduate Program in Quantitative Biosciences, Georgia Institute of Technology, Atlanta, GA, USA

[3]School of Physics, Georgia Institute of Technology, Atlanta, GA, USA

[4]Department of Zoology, University of British Columbia, Vancouver, British Columbia, Canada

[5]School of Electrical & Computer Engineering, Georgia Institute of Technology, Atlanta, GA, USA

## Abstract

The evolution of large organismal size is fundamentally important for multicellularity, creating new ecological niches and opportunities for the evolution of increased biological complexity. Yet little is known about how large size evolves, particularly in nascent multicellular organisms that lack genetically-regulated multicellular development. Here we examine the interplay between biological and biophysical drivers of macroscopic multicellularity using long-term experimental evolution. Over 600 daily transfers (~3,000 generations), multicellular snowflake yeast evolved macroscopic size, becoming ~$2{\cdot}10^4$ times larger (~mm scale) and $10^4$-fold more biophysically tough, while remaining clonal. They accomplished this through sustained biophysical adaptation, evolving increasingly elongate cells that initially reduced the strain of cellular packing, then facilitated branch entanglement so that groups of cells stay together even after many cellular bonds fracture. Four out of five replicate populations show evidence of predominantly adaptive evolution, with mutations becoming significantly enriched in genes affecting cell shape and cell-cell bonds. Taken together, this work shows how selection acting on the emergent properties of simple multicellular groups can drive sustained biophysical adaptation, an early step in the evolution of increasingly complex multicellular organisms.

## Introduction

Size plays a fundamental role in the evolution of multicellularity, as it allows organisms to explore novel ecological niches (*1*), affords protection from the external environment (*2*), and underlies the evolution of cellular differentiation (*3-7*). The evolution of macroscopic size has been hypothesized to be a key driver of increased organismal complexity, as large size creates an evolutionary incentive to solve challenges of nutrient and oxygen transportation that are otherwise inescapable consequences of diffusion limitations (*8*, *9*). However, little is known about how nascent multicellular organisms, consisting of small groups of undifferentiated cells, evolve to form biomechanically tough, macroscopic multicellular bodies, and whether selection for size itself can be a driver of multicellular innovation (*10*).

The evolution of macroscopic size presents a fundamental challenge to nascent multicellular organisms, requiring the evolution of biophysical solutions to evolutionarily novel stresses that act over previously-unseen multicellular size scales (*11-15*). Indeed, while prior work with yeast and algae have shown that novel multicellularity is relatively easy to evolve *in vitro*, these organisms remain microscopic, typically growing to a maximum size of tens to hundreds of cells (*16-20*). Extant multicellular organisms have evolved a number of developmentally-regulated mechanisms for increasing biophysical toughness, or reducing the rate at which stress accumulates (*21-23*). In the absence of genetically-regulated multicellular development, however, we do not know how, or even whether, nascent multicellular organisms can evolve the increased biophysical toughness required for the evolution of macroscopic size.

Here we examine the interplay between biological, biophysical, and environmental drivers of macroscopic multicellularity using long-term experimental evolution. We subject snowflake yeast (*18*), a model of undifferentiated multicellularity, to 600 rounds (~3,000

generations) of daily selection for increased size. Because oxygen is thought to have played a key role in the evolution of macroscopic multicellularity, we evolved snowflake yeast with either anaerobic or aerobic metabolism. All five of our replicate anaerobic populations evolved macroscopic size, while all aerobic populations remained microscopic throughout the duration of the experiment, supporting the hypothesis that selection for efficient aerobic metabolism can be a powerful constraint on the evolution of multicellularity (*24*). Macroscopic size evolved through two sequential steps in all five replicate populations. First, snowflake yeast increased the length of their constituent cells, which delays organismal fracture caused by packing-induced strain (*15*). Next, they evolved to entangle branches of connected cells such that a single cell-cell separation no longer causes multicellular fracture, evolving to become ~$2{\cdot}10^5$ times larger, forming millimeter-scale groups of clonal cells. Together these adaptations increased the toughness of individual clusters by more than $10^4$-fold, transforming the initial snowflake yeast ancestor, which was ~100 times weaker than gelatin, to a multicellular organism with similar materials properties (*i.e.,* strength and toughness) to wood. Sequencing reveals evidence for sustained adaptive evolution, with a high proportion of nonsynonymous mutations occurring in genes affecting cell shape and cell-cell bonds.

## Results

We began our experiment by engineering a unicellular isolate of *S. cerevisiae* strain Y55 to grow with the snowflake phenotype by deleting the *ACE2* open reading frame (*25*). To examine the effect of oxygen on the evolution of size, we generated a spontaneous petite mutant incapable of respiration. We generated five replicate populations of each genotype, selecting for larger size every 24 h by selecting for increasingly rapid sedimentation (see Methods for details). We evolved these 15 populations over 600 rounds of growth and settling selection (~3,000

generations), archiving all populations at -80°C every 10-15 days. We refer to the five replicate anaerobic, mixotrophic, and obligately aerobic populations as PA, PM, and PO 1-5, respectively.

All five populations of anaerobic snowflake yeast evolved macroscopic size, with individual clusters visible to the naked eye (Fig. 1A-B). In contrast, snowflake yeast capable of metabolizing oxygen remained microscopic (Fig. 1C & Fig. S1), a result consistent with recent work showing that competition for limiting oxygen imposes a powerful constraint on the evolution of macroscopic multicellular size (*24*). The anaerobically-evolved populations increased their mean cluster radius from 16 µm to 434 µm, a ~$2 \cdot 10^4$-fold increase in volume (Fig. 1D, $p < 0.0001$; $F_{5,\ 13321} = 2100$, Dunnett's test in one-way ANOVA). This corresponds to an estimated increase from ~100 cells per cluster to ~450,000 (comparing average cluster volumes, accounting for changes in mean cell volume and cellular packing density within clusters). The largest clusters of 600-day evolved snowflake yeast are over a millimeter in diameter (Fig. 1A), which is comparable to the size of an adult *Drosophila* (*26*).

The evolution of macroscopic multicellularity in the anaerobic snowflake yeast populations occurred via two distinct phases (Fig. 1D). All five replicate populations evolved a ~2-fold increase in mean cluster radius (from 16.7 µm to 33.6 µm; $p < 0.0001$, $F_{5,\ 27005} = 4518$, Tukey's HSD, one-way ANOVA) over the first 100 days of selection, with little among-population variation (means range from 33.0 to 33.9 µm in the five replicate populations). At this point, they entered a period of stasis (defined as <10% change in mean size per 50-day sampling interval), waiting anywhere from 50 days in population PA5 to 250 days in population PA1 before rapidly evolving further size increases. The wide among-population variance in the duration of evolutionary stasis suggests that the evolution of macroscopic size was more challenging than the initial evolution of increased size, which was highly parallel.

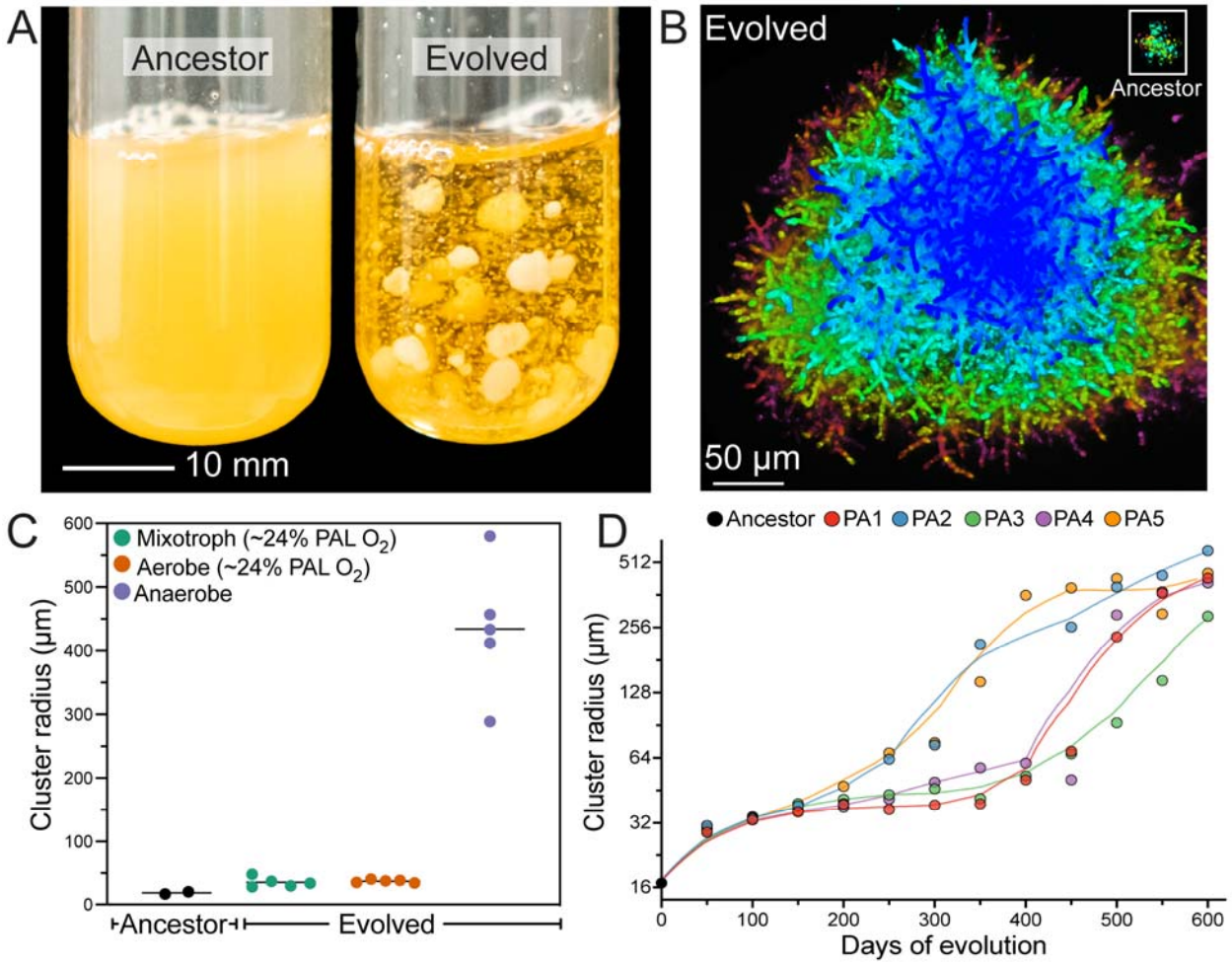



**Figure 1. Evolution of macroscopic multicellularity in five replicate snowflake yeast populations.** Individual snowflake yeast are visible to the naked eye after 600 days of evolution. Snowflake yeast are shown as they appear in our evolution experiment (A), or via confocal microscopy (B, both images shown at same magnification with color representing depth in the *z* plane). (C) Only the anaerobic populations (PA1-PA5) evolved macroscopic multicellularity after 600 transfers; mixotrophic (PM1-PM5) and obligately aerobic (PO1-PO5) snowflake yeast remained microscopic despite daily selection for larger size. (D) Within anaerobic populations, the evolution of macroscopic multicellularity occurred via two distinct phases. The early evolution of larger size, up to a weighted mean radius of ~32 µm, was highly parallel. The time required before breaking through to larger size, however, varied widely between populations. In C and D, the data points show the average radius of the cross-sectional area for snowflake yeast in each population, weighted by cluster biomass (which represents the average size multicellular group any given cell will be in, see Methods for details). See Fig. S1 for additional data on the evolution of cluster size in oxygen-using populations and Fig. S2 for full cluster size distributions for the 600-day anaerobic populations (PA1-PA5). Lines in (D) are Lowess smoothing curves, intended to aid the eye.

As with their ancestor, macroscopic snowflake yeast grow via mother-daughter cell adhesion (*i.e.*, budded daughter cells do not separate after cell division), forming a branched, tree-like structure (Fig. 2 A&B). When compressed, they fracture into small modules that resemble their snowflake yeast ancestor (Fig. 2 A&B). Throughout the experiment, snowflake yeast cells evolve to be more elongate, increasing in average aspect ratio (ratio of length to width) from ~1.2 to ~2.7 (Fig. 2C&D; $F_{5, 1993} = 206.2$, $p < 0.0001$, Dunnett's test after one-way ANOVA; Fig. S3). Initially, cluster size was a roughly linear function of aspect ratio (Fig. 2E inset), but this relationship changes once they evolve macroscopic size (Fig. 2E).

Prior work has shown that the evolution of more elongate cells increases the size to which snowflake yeast grow by decreasing the density of cellular packing (*i.e.*, their packing fraction) in the cluster interior, which reduces cell-cell collisions that drive multicellular fracture (*15*, *27*). We examined this relationship over the course of our experiment in replicate population two (PA2), which was one of the first lineages to evolve macroscopic size. As predicted by a 3D biophysical simulation (Fig. 2F), cellular elongation decreased the packing fraction of multicellular groups- but only initially, from aspect ratio ~1.2-2. Beyond this, clusters with more elongate cells actually became more densely packed, and experimentally-measured packing fraction became increasingly divergent from model predictions (Fig. 2F). This divergence suggests that this lineage evolved a novel biophysical mechanism for increased multicellular toughness, capable of withstanding growth to large size and a high cellular packing fraction.

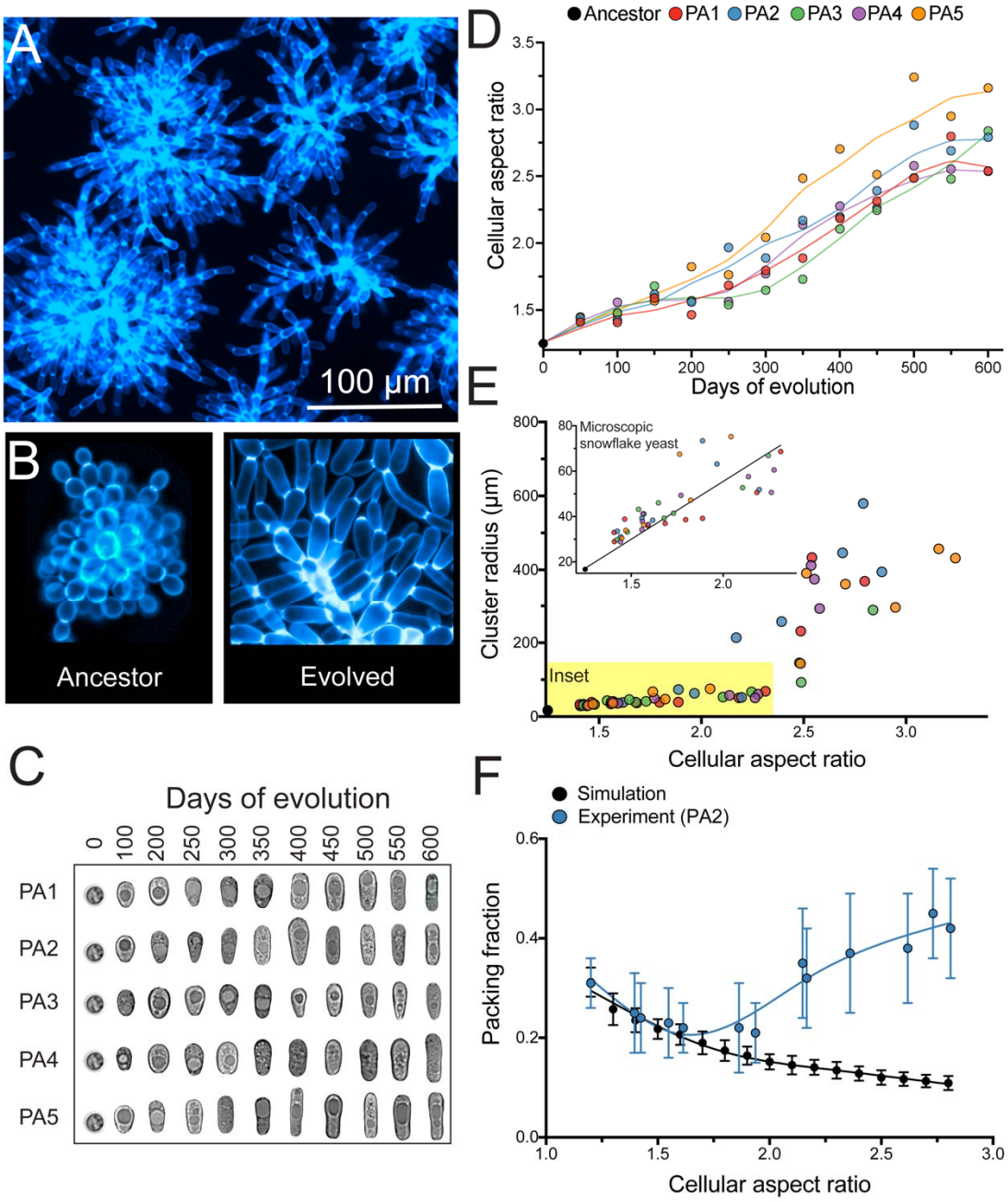



**Figure 2. Evolution of novel cell morphology.** (A) When compressed, macroscopic snowflake yeast fracture into modules that resemble their ancestor, but (B) retain the same underlying branched growth form (cell walls are stained with calcofluor-white). (C) and (D) show the parallel evolution of elongated cell shape, resulting in an increase in average aspect ratio from ~1.2 to ~2.7 (for each point in D, 453 cells were measured on average). (E) Early in their evolution (aspect ratio 1-2.3), cluster size (weighted mean radius) is an approximately linear function of cellular aspect ratio (inset; $p < 0.0001$, $y = 41.1x - 27.8$, $r^2 = 0.72$). This relationship does not hold for higher aspect ratios. (F) A biophysical model of snowflake yeast predicts that increasing cellular aspect ratio should decrease cellular packing fraction (black points). We see a close correspondence with these predictions for low aspect ratios, but our experimental data diverges from model predictions for aspect ratios beyond 2. Each datapoint in (F) reports the mean of 15 snowflake yeast clusters or 25 replicate simulations, ± one standard deviation.

The simplest way that snowflake yeast could evolve to become macroscopic is to become adhesive, forming large aggregates, which is a common strategy in yeast (i.*e.*, via flocculation (*28*)). To determine if clusters of macroscopic yeast grow via aggregation, or if they develop as a single clonal lineage, we labeled a single-strain isolate taken from PA2, t600 with either GFP or RFP. If adhesive aggregation were responsible for their large size, we would expect to see chimeric groups composed of both red and green cells (which are genetically identical, other than the fluorescent marker) (*28*). After five days of co-culture and 24h of growth, however, all multicellular clusters (n=70; Fig. S4) remained monoclonal. This is unlikely to occur with aggregation. If we conservatively assume each macroscopic snowflake yeast cluster we measured were the result of just a single fusion event, occurring with equal probability between two groups of red and green cells, then the binomial probability of finding no chimeric groups in our sample would be $10^{-6}$. Floc-like aggregation thus does not explain the evolution of macroscopic size in snowflake yeast.

To examine how changes in the topology of snowflake yeast may underlie their increased size, we imaged clusters via Serial Block Face Scanning Electron Microscopy (SBF-SEM). This technique enables us to image the interior of clusters that are difficult to resolve with light-based microscopy, ultimately allowing us to map their internal architecture with nanometer precision (*29*). Surprisingly, individual macroscopic snowflake yeast were not composed of a single topologically-connected component, like their ancestors. Instead, they contained numerous disconnected branches of cells, suggesting that the cluster remained intact even when cell-cell connections were severed (Fig. 3A). Within macroscopic clusters, separate branches contact, intercalate, and even wrap around each other (Fig. 3A). As these clusters are densely packed, moving one component would require moving many other components as well. Based on these

observations, we hypothesized that branches are entangled, in a manner reminiscent of physical gels (*30*) and entangled granular materials (*31*). Entanglement would provide a mechanism for branches of cells to remain in the same, densely packed group even after cell-cell bonds break.

Following prior work in entangled chains and knotted strings (*31*, *32*), we used our SBF-SEM dataset to quantify branch entanglement in snowflake yeast by analyzing chain topology and geometry. Specifically, we constructed the convex hull of each connected component within a sub-volume, which denotes the smallest convex polyhedron containing this component (see Fig. S5A). If a cell from one connected component overlaps with the convex hull of a second, then the two can be considered entangled. By percolating entanglement among adjacent connected components throughout the sub-volume, we can measure the extent to which the cluster's biomass is mutually entangled (Fig. 3B, Fig. S5B). For entanglement to underlie macroscopic size, the largest entangled component (consisting of many entangled pieces) must be able to resist mechanical stress, meaning that there must be an entangled component that spans the vast majority of the cluster (*33*). In analyses of 10 randomly selected sub-volumes from different macroscopic snowflake yeast clusters from population PA2, t600, we found that the largest entangled component contains 93% +/- 2% of all connected components. This observation supports the hypothesis that entanglement between cell branches can prevent cluster fracture in the event that a cell-cell bond fails.

As a further test, we investigated the mechanics of macroscopic snowflake yeast. Strain stiffening is a signature of entangled chains (*31*). When compressed, the effective stiffness of entangled chains increases with increased strain. This enables entangled materials to withstand stress orders-of-magnitude greater than their non-entangled counterparts, a property necessary for achieving macroscopic size (*31*). Conversely, as the ancestor is not entangled, it is not

expected to exhibit strain-stiffening behavior. We measured the stress response of 10 macroscopic snowflake yeast clusters under uniaxial compression using a macroscopic mechanical tester (Zwick Roell Universal Testing Machine). We repeated the same experiment for 10 ancestral snowflake yeast clusters using an atomic force microscope (AFM Workshop LS-AFM). The stress-strain plot for the ancestor is linear ($r^2$ = 0.97 +/- 0.02, average and standard deviation of the regression for 10 samples, Fig. 3C inset), clusters fracture at stress as low as 240 Pa and have toughness as low as 8.9 J/m$^3$ (*15*) In contrast, macroscopic snowflake yeast clusters have a convex stress-strain curve (Fig. 3C), can support stresses at least as large as ~7 MPa without failing, and have toughness greater than 0.6 MJ/m$^3$. Thus, entanglement both enables separate branches within macroscopic snowflake yeast to stay together and allows them to endure the large stresses necessary for growth to macroscopic size.

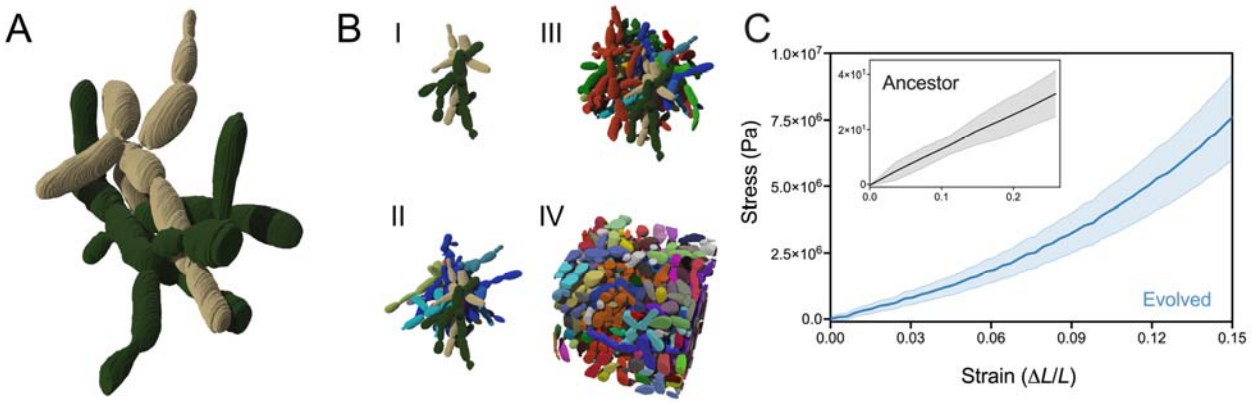



**Figure 3. Branch entanglement underlies the evolution of macroscopic size.** (A) Shown are two entangled components (green and tan), obtained via SBF-SEM imaging. (B) Branch entanglement is pervasive in macroscopic snowflake yeast. Starting with the two-component sub-volume in (A), we percolated entanglement by adding on adjacent entangled components in four steps (I-IV). (C) Stress vs. strain plot for macroscopic snowflake yeast (PA2, t600) clusters in blue and the ancestor in grey (ancestor shown again in inset with a rescaled *y* axis). Macroscopic snowflake yeast experience strain stiffening, a hallmark of entangled systems, while the ancestor's stress-strain plot is linear, which is expected for non-

entangled systems. The shaded area shows one standard deviation based on 10 repeated measurements for each.

To uncover the genomic basis of cell-level changes driving the multicellular adaptation, we sequenced the genomes of a single strain from each of the five populations (PA1-PA5) that independently evolved macroscopic multicellularity after 600 transfers (Fig. 4A&B). 4/5 had more nonsyonymous mutations than expected by chance, indicating that selection has played an important role in allele frequency change within these populations. Specifically, the ratio of nonsynonymous to synonymous mutations, relative to neutral expectations, was 2.1, 5.6, 0.76, 2.56, and 2, for isolates from PA1-PA5, respectively. PA3, which had a dN/dS ratio of just 0.76, evolved macroscopic size only near the end of the experiment, later than the other populations (Fig. 1D).

Over ~3,000 generations snowflake yeast evolve dramatically more elongate cells (Fig. 2C&D), which plays a central role in the evolution of increased cluster size (Fig. 2E) and biophysical toughness (Figs. 2F, 3, & S5). Gene Ontology (GO) terms known to affect cell length were significantly enriched, namely genes of the cell cycle (*34*) (23 / 102 nonsynonymous mutations, $p = 4.68\times10^{-10}$) and filamentous growth (6 / 102 nonsynonymous mutations, $p = 0.023$). In addition, we found 12 nonsynonymous mutations in genes with known roles in cellular budding (Fig. 4E), which includes eight genes that have previously been shown to increase the size of bud scars (*AKR1*, *ARP5*, *CLB2*, *GIN4*, *PRO2*, *RPA49*, *RSC2*, *PHO81*) (*35*, *36*). Mutations arose in two of these genes in different populations (*i.e.*, *PHO81* in populations PA1 and PA5, and *GIN4* in populations PA2 and PA3, Fig. 4C), indicating parallel evolution. All else equal, larger bud necks should increase the amount of cell wall shared between mother and daughter cells, potentially increasing the strength of cell-cell connections.

We tested the effects of cell lengthening mutations (*akr1*Δ and *arp5*Δ) and a putative cell-cell connection strengthening mutation (*gin4*Δ) in the *ace2Δ* ancestor. Together, these mutations led to a 7.7-fold increase in average radius (>100-fold greater volume; Fig. 4D; $p < 0.0001$, $t = 46.39$, df = 366.1, Welch's *t*-test), confirming that mutations affecting cell length and cell-connection strength can underlie the emergence of dramatically increased multicellular size in snowflake yeast.

In addition to the GO term analysis, we examined the potential phenotypic targets of novel mutations using the CellMap database (*37*), a method that maps core biological processes in yeast by testing the phenotypic effects of nearly a million pairwise genetic interactions (*38*). These results support our original analysis, showing that interactions between bioprocesses affecting mitosis, cell polarity (*e.g.*, the cell cycle) and the cell wall (Fig. 4F) evolved within macroscopic snowflake yeast.

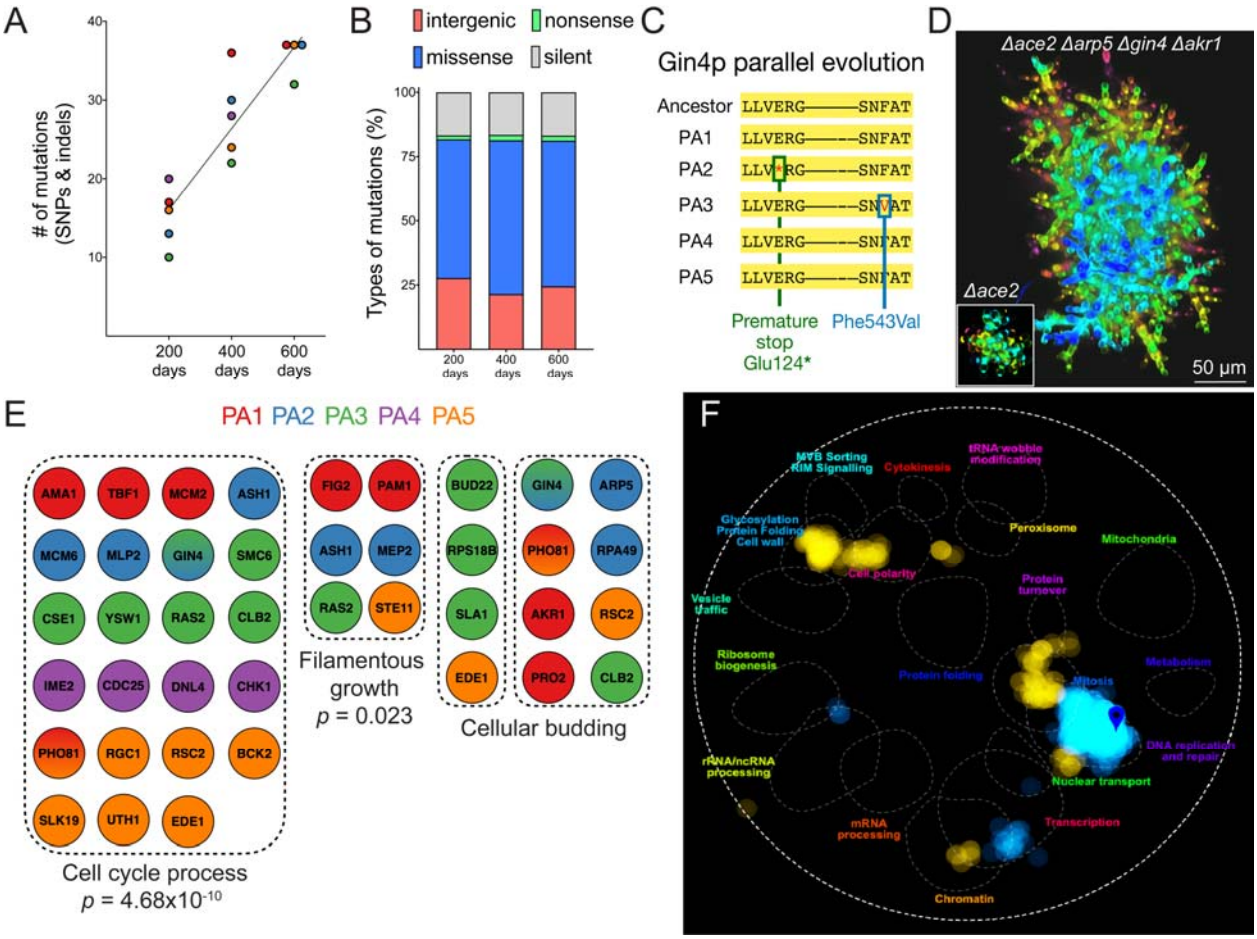



**Figure 4. Whole-genome sequencing reveals the dynamics of molecular evolution and the genetic basis of cell-level and cluster-level changes.** (A) and (B) show the number and types of mutations in evolved single strains from each population. (C) *GIN4*, a kinase controlling the size of cellular bud necks with a potential strengthening effect on connections between cells, is mutated in two independent populations. (D) Genetic engineering of cell lengthening (i.e., *AKR1* and *ARP5*) and bud-scar strengthening (i.e., *GIN4*) mutations in the ancestral background (inset) increased the weighted mean radius of snowflake yeast clusters from 16 μm to 123.5 μm (ratio of the average size of 9,964 ancestors and 367 engineered clusters). (E) Macroscopic snowflake yeast were significantly enriched in mutations affecting cell cycle progression and filamentous growth. In addition, we saw mutations affecting budding (*i.e.*, the location of buds on the cell surface, and bud neck size). While not a formal GO category (precluding a statistical test), these constituted 12 out of the 99 total genes containing non-synonymous mutations in genotypes isolated from PA1-PA5 after 600 days of evolution. (F) Spatial analysis of functional enrichment (*39*) of non-synonymous mutations show co-occurrence of mutations in genes affecting mitosis ($p = 1.11 \times 10^{-6}$) and cell polarity ($p = 1.74 \times 10^{-5}$), processes that affect cell aspect ratio and bud size / location.

## Discussion

In this paper, we show that snowflake yeast, a model system of undifferentiated multicellularity, were capable of evolving macroscopic size over 600 days of directed evolution. Macroscopic snowflake yeast are readily visible to the naked eye, containing hundreds of thousands of clonal cells. They achieved this remarkable increase in size by evolving highly elongate cells that become entangled within the cluster interior. This critical innovation allows multicellular groups to remain physically attached even when individual cellular connections are severed, increasing cluster toughness by more than 10,000 fold. As a material, snowflake yeast evolve from being ~100-fold weaker than gelatin (*40*), to having the strength and toughness of wood (*41*). 4/5 replicate populations showed evidence of predominantly adaptive evolution (dN/dS >1), and mutations were enriched in genes affecting the cell cycle and budding - traits that increase cell length and cell-cell bond strength.

To our knowledge, this is the first long-term evolution experiment to directly show how simple multicellular organisms evolve increased complexity. Snowflake yeast do not possess evolved systems of developmental control over multicellular morphology- they were synthetically created from a unicellular ancestor at the start of this experiment by deleting the *ACE2* transcription factor, and have had no prior opportunity to gain multicellular adaptations. Instead, they demonstrate how, prior to the evolution of development, novel, heritable multicellular traits can arise as an emergent property of changes in the traits of constituent cells. Two cell-level innovations appear to have played a key role in the evolution of macroscopic size: more elongate cells and larger bud scars, which increase the strength of cellular connections. Increased cell length initially reduces strain generated from cellular packing, which is the

primary manner in which size increased early in the experiment, and may underlie entanglement by facilitating cellular intercalation.

Entanglement is a common mechanism through which filamentous materials can solidify. It can operate on nearly any length scale, ranging from nanoscale polymers (*42*) and nanofibers (*43*), to macroscopic staples (*44*) and beaded chains (*45*). The prevalence of entanglement in superficially different systems is likely due to its simplicity and efficacy; if pairs of constituents are easily entangled, large mutually-entangled clusters readily form, greatly increasing the strength and toughness of the material. While much remains to be discovered about the importance of entanglement in biological processes, remarkably tough biological materials often evolve from densely-packed filamentous material (*e.g.*, articular cartilage (*46*), rhinoceros horn (*47*, *48*), and trunks supporting massive arboreal fungal (*49*) and plant bodies (*50*)).

Our results depend on the fact that snowflake yeast grow as topologically-structured groups with permanent cellular bonds, and we would not necessarily expect similar biophysical exaptation in organisms with alternative means of group formation. These features, however, make it well suited as a model system for the lineages that have ultimately evolved complex multicellularity. Of the five lineages that independently evolved complex multicellularity (fungi, animals, plants, red algae, and brown algae), all but animals possess permanent cell-cell bonds, and early multicellular lineages in each are thought to have started out as simple, topologically-structured networks (*51*). While animals do not currently have permanent cell-cell bonds, little is known about their ancestral mode of cellular adhesion. Indeed, their closest living relatives, the choanoflagellates, form topologically structured multicellular groups with permanent cell-cell bonds (*52-54*), suggesting that early animals may have possessed a similar mode of growth.

Despite daily selection for increased size, macroscopic multicellularity only evolved in our anaerobic, rather than aerobic lineages (Fig. 1C & Fig. S1). This is further evidence that, rather than stimulating the evolution multicellularity, the first oxygenation of Earth's atmosphere may have acted a powerful constraint on the evolution of large size. Oxygen can increase cellular growth by increasing ATP yield from metabolism (*55*), and allowing growth on non-fermentable carbon (*56*). When oxygen availability within an organism is limited by diffusion, however, large size can be maladaptive by reducing the ability of interior cells to utilize this valuable resource (*24*) - this is a constraint that anaerobic organisms simply do not face. The positive effect of oxygen on multicellular size (*10*) may only occur once it reaches a sufficiently high concentration, as happened around 0.6 Ga (*57*).

Broadly speaking, our results demonstrate that rapid multicellular innovation is possible, even in the absence of genetically-regulated multicellular development. The evolution of macroscopic size required the evolution of new physical mechanisms increasing organismal toughness, which itself required fundamental changes in the shape and behavior of the cells within the organism. Selection acting on the emergent properties of multicellular groups thus created ample opportunity for sustained adaptive evolution underlying the origin of novel multicellular phenotypes, providing the first *in vitro* experimental demonstration of how selection for larger size can itself be a powerful driver of increased organismal complexity.

## Materials and Methods

**Long-term evolution experiment.** To generate our ancestral snowflake yeast for the long-term evolution experiment, we started with a unicellular diploid yeast strain (Y55). In this yeast, we replaced both copies of the *ACE2* transcription factor using a *KANMX* resistance marker (ace2::*KANMX*/ace2::*KANMX*) and obtained a snowflake yeast clone (see (*25*) for a detailed description of strains and growth conditions, including measurements of oxygen concentrations in growth media). When grown in YEPD media (1% yeast extract, 2% peptone, 2% dextrose), these yeast are mixotrophic, both fermenting and respiring. When grown in YEPG media, which is the same as YEPD but with the dextrose replaced by 2.5% glycerol, these yeast are incapable of fermentation and are obligately aerobic. From this initial clone of *ace2Δ* snowflake yeast, we selected a randomly produced 'petite' ($p^-$) mutant. Due to a large deletion in its mitochondrial DNA (identified via sequencing), this snowflake yeast is unable to respire and is therefore metabolically 'anaerobic', and was cultured in YEPD. We evolved five replicate populations of mixotrophic snowflake yeast (referred to as populations PM1-PM5), obligately aerobic (PO1-PO5) and anaerobic (PA1-PA5) snowflake yeast in 10 mL of culture media, growing them in 25 x 150 mm culture tubes for 24 hours at 30□ with 225 rpm shaking. We used settling selection to select for larger cluster size. Once per day, after 24h of growth, we transferred 1.5 ml of culture into 1.5 mL Eppendorf tubes, let them settle on the bench for 3 minutes, discarded the top 1.45 mL of the culture, and only transferred the bottom 50 µl of the pellet into a new 10 mL of culture media for the next round of growth and settling selection. Once the anaerobic populations (PA1-PA5) had started to evolve visibly larger clusters with all biomass settling to the bottom of the tube in under a minute, we decreased the length of gravitational selection to 30 seconds, thus keeping them under directional selection for increased size. The timing of this change

corresponded to ~350 days for PA2 and PA5 and ~500 days for PA1, PA3, and PA4. We used wide-bore filtered pipette tips (Thermo Scientific) for our daily transfers. In total, we applied 600 rounds (days) of growth and settling selection. We archived a frozen glycerol stock of each population every 10-15 transfers.

**Measuring cluster size.** We used a standard visualization protocol to measure the size of snowflake yeast from each population over the 600-day evolution experiment. To prepare yeast for imaging, we revived evolved frozen cultures for each population in 50-day intervals (12 for each of the 15 replicate populations). We then inoculated each sample into fresh media and brought them to equilibrium over a five-day culture process, performing daily settling selection prior to transfer to fresh media. After 5 transfers, we took a random 1mL subsample of each 24-hour culture, placing them in 1.5 ml Eppendorf tubes. We added 0.5 ml of sterile water to each well of 12-well culture plates, then gently vortexed each snowflake yeast sample and diluted them into the water (1,000-fold dilution for microscopic populations, and 100-fold dilution for macroscopic populations). We shook each well plate gently to disperse the yeast clusters evenly over the bottom of each well. We then imaged each well using a 4X Nikon objective, capturing the cross-sectional area of clusters without disrupting their 3D structure. Next, we used ImageJ-Fiji to calculate the cross-sectional area of each cluster, converting pixels to microns by including a physical 100 μm scale bar in each image.

**Calculating the weighted average cluster size.** The distribution of clusters size across various isolates are not consistent- some are unimodal, while others contain a large number of small groups that may only contain a few cells. Even when these small groups constitute a trivial amount of the population's biomass, variation in their abundance can have a large impact on sample statistics, like average cluster size. Because of this variation, mean size is an unreliable

and often uninformative measure of the central tendency of the cluster size distribution, and does not accurately describe how cells are distributed across different cluster size classes. To account for this, we calculated the distribution of cellular biomass over the range of cluster sizes, and found the mean of this biomass distribution (which is the same as weighting mean cluster size by its biomass). This weighted mean cluster size represents the expected size group any given cell will be in (see Fig. S2 for a visual representation), and is an accurate measure of changes in the distribution of cellular biomass across different cluster sizes over evolutionary time. Rather than presenting the weighted mean group size as a volume, we transformed these into an average radius to be consistent with the units that have historically been used in the paleontological literature discussing the evolution of macroscopic multicellular organisms.

**Aspect ratio data collection and analysis.** To measure the evolution of cellular aspect ratio in populations PA1-PA5 over the 600-day evolution experiment, we first inoculated 61 samples (1 ancestor + 5 replicates x 12 time points, each separated by 50 days) and grew them overnight in shaking incubation as described above. Following the same growth protocols as in our cluster size measurements, we grew these samples for five consecutive days. On the final day, we transferred 100 µl of each culture into tubes with fresh YEPD and incubated them for 12 hours. Next, we stained samples in calcofluor-white by incubating them in the dark for 30 minutes (at a final concentration of 5 µM) prior to imaging (40x objective, UV excitation of blue fluorescent cell wall stain, imaged on a Nikon Ti-E). We measured the aspect ratio of individual cells within snowflake yeast clusters on ImageJ-Fiji, analyzing an average of 453 cells per population.

**Testing aggregative vs. clonal development.** To determine if macroscopic snowflake yeast are aggregative, we isolated a single genotype from PA2, t600 (strain GOB1413-600), and engineered it to constitutively express either green or red fluorescent proteins. To do that, we

amplified the prTEF_GFP_NATMX construct from a pFA6a-eGFP plasmid and the prTEF_dTOMATO_NATMX construct from a pFA6a-tdTomato plasmid. We then separately replaced the *URA3* open reading frame with *GFP* or *dTOMATO* constructs in an isogenic single strain isolate following the LiAc transformation protocol (*58*). We selected transformants on Nourseothricin Sulfate (Gold Biotechnology Inc., U.S.) YEPD plates and confirmed green or red fluorescent protein activity of transformed macroscopic clusters by visualizing them under a Nikon Eclipse Ti inverted microscope. To test whether they grow clonally or aggregatively, we first inoculated *GFP* or *dTOMATO* expressing clones individually overnight. We then mixed the two cultures in equal volume and diluted 100-fold into a 10 mL fresh culture. We co-cultured this mixed population for five days, transferring 1% of the population to fresh media every 24 h. Finally, we washed this culture in 1 mL sterile water and visualized 70 individual clusters under both red and green fluorescent channels, allowing us to count the number of snowflake yeast clusters that were green, red, or chimeric.

**DNA extraction and genome sequencing.** To extract DNA for whole-genome sequencing, we isolated clones from each of the evolved replicate populations of anaerobic yeast (*i.e.*, PA1-PA5) and their common ancestor after 200, 400, and 600 days of evolution. We inoculated these 16 samples in YEPD for 12 hours and extracted their genomic DNA using a commercially available kit (Amresco, Inc. VWR USA). We measured the concentration of DNA with a Qubit fluorometer (Thermo Fisher Scientific, Inc.). We prepared genomic DNA libraries for the 16 samples using NEBNext Ultra DNA Library Prep Kit for Illumina (New England Biolabs, Inc). We quantified the quality of the genomic DNA library using the Agilent 2100 Bioanalyzer system located at the Genome Analysis Core Laboratories at Georgia Institute of Technology (Agilent Technologies, Inc). Finally, whole genomes were sequenced using the HiSeq 2500

platform (Illumina, Inc) by the Genome Analysis Core Center located in the Petit Institute, Georgia Tech. As a result, we obtained paired 150 bp (R1 & R2) FASTQ reads from two lanes (L1 & L2).

**Bioinformatic analysis.** For our bioinformatic analysis, we used the bash command-line interface on a Linux platform. To identify *de novo* mutations (single nucleotide changes, or 'SNPs,' and small insertion/deletions, or 'indels') in the ancestral and evolved genomes, we first filtered out low-quality reads using a sliding window approach on Trimmomatic (v0.39). We aligned reads to the yeast reference genome (S288C, SGD) using an algorithm in the BWA software package (*i.e.*, BWA-MEM) (*59*). Next, we used the genome analysis toolkit (GATK) to obtain and manipulate .bam files (*60*). Duplicate reads were marked using the Picard - Tools (MarkSuplicates v2.18.3). We called SNPs using two different tools, i.e., GATK4 HaplotypeCaller (v4.0.3.0) and FreeBayes (v1.2.0) (*60, 61*). We validated SNP calls by comparing results obtained by two independent tools. For indels, we used the output from HaplotypeCaller. To filter variants according to their quality/depth scores and generate an overview of the variant calling step's statistical outcome, we used VCFTOOLS (v0.1.16) (*62*). Finally, after manually checking each variant call by visualizing SAM files and VCF files on Integrative Genomics Viewer (IGV) (*63*), we extracted *de novo* variants by making a pairwise comparison of each VCF file of the evolved genomes against the VCF file of the ancestral genome by using bcftools-isec (v1.10) (*64*). Lastly, we annotated evolved mutations using SnpEff (v4.3T) (*65*).

We calculated dN/dS ratios by employing the counting method (*66*). Briefly, we counted the number of synonymous and nonsynonymous mutations in each population. Because analyses of dN/dS ratios typically use mutations that are fixed in the population (*67*), we examined

variants that arose at t400 and persisted to t600. We then counted the number of synonymous and nonsynonymous mutations in each genome, and the number of synonymous and nonsynonymous sites in the coding regions of the genes that contained a mutation, finding the dN/dS ratio, relative to neutral expectations, as *dN/dS = (Nd/Nc) / (Sd/Sc)* (see Source Data). Here *Nd* and *Sd* are the number of nonsynonymous and synonymous variants that arose; *Nc* and *Sc* are the number of nonsynonymous and synonymous sites in the sequences of genes containing a mutation.

To search for gene ontology (GO) term enrichment for *de novo* mutations, we generated a combined list of synonymous and nonsynonymous mutations within gene coding regions. We then searched for significantly enriched gene ontology terms using the process option on the yeast genome database (*68*).

**Genetically engineering large-sized snowflake yeast.** To genetically engineer a snowflake yeast with cell lengthening (*AKR1* and *ARP5*) and bud-scar strengthening (*GIN4*) mutations, we first deleted each gene separately in the ancestral background (i.e., *ace2Δ*). By inducing meiosis in sporulation media (2 % KAc), we generated F1 spores, dissected tetrads on a tetrad scope (Singer Instruments, Somerset, UK), and as a result of auto-diploidization, obtained homozygous deletion mutants for each mutant. By two rounds of sporulation, mating, and analysis of dissected tetrads based on their drug resistance phenotype, we brought all three mutations under the *ace2Δ* background, generating a quadruple mutant (*ace2Δ*, *gin4Δ*, *arp5Δ*, *akr1Δ*). Finally, we quantified the cellular aspect ratio and cluster size (cross-sectional area) of this mutant by imaging uncompressed clusters under a Nikon Eclipse Ti inverted microscope.

**Specimen preparation for Serial Bulk Faced Scanning Electron Microscopy (SBF-SEM)**. We fixed snowflake yeast in 2% formaldehyde (fresh from paraformaldehyde (EMS)) containing 2 mM $CaCl_2$, incubating at 35°C for 5 minutes followed by 2-3 hours on ice. Next, we incubated these yeast for an hour in a solution of 1.5% potassium ferrocyanide, 0.15M cacodylate buffer, 2 mM $CaCl_2$, and 2% aqueous osmium tetroxide. This last step was performed on ice and under vacuum. Finally, we washed our yeast and incubated them in thiocarbohydrazide solution (10 g / L double-distilled water) for 60 minutes at 60°C, followed by en bloc uranyl acetate and lead aspartate staining (*69, 70*).

**SBF-SEM.** We imaged fixed yeast on a Zeiss Sigma VP 3View. This system has Gatan 3View SBF microtome installed inside a Gemini SEM column. For this work, yeast clusters that were embedded in resin were typically imaged at 2.5 keV, using 50-100 nm cutting intervals, 50 nm pixel size, beam dwell time of 0.5-1 µsec and a high vacuum chamber.

**SEM Image analysis.** Images were initially in.dm3 format, which we converted to .tiff using GMS3 software. We then cleaned the images and passed them through a gaussian filter in Python. Using the interactive learning and segmentation toolkit (ilastik), we segmented images into 3 parts: live cells, dead cell debris, and background. We then imported segmented HDF5 files in python. First, we identified connected cells using the nearest neighbor algorithm to identify connected cells. We call a set of connected cells inside a sub-volume a connected component. Then, using a 3D extension of the gift-wrapping algorithm, we extracted the convex hull of each connected component.

**Visualization of SEM images.** After segmenting images as described above, we created a mesh of individual cells by dilating binarized images. After creating the surface mesh of each

individual cell using the mesh tool in Mathematica 12, we imported whole sub-volumes in Rhino6. Then we manually identified cell-to-cell connections and colored each connected component differently.

**Volume fraction data collection and analysis.** We measured the packing fraction (proportion of the cluster volume that is cellular biomass) by measuring the number of cells within a cluster, their size, and the volume of the cluster, following the protocol described in Zamani et al. (2021) (*71*).

**Mechanical testing.** To test the response of ancestral clusters to uniaxial compression we submerged individual clusters under water, and then compressed them using a Puima Chiaro nanoindenter (Optics11, 19.5 um spherical glass probe). For mechanical measurements of macroscopic snowflake yeast, we used a Zwick Roell Universal Testing Machine (UTM) with 5 N probe. As above, individual clusters were extracted from the growth tube and placed on the testing stage while submerged under water.

## Acknowledgements

We thank Jennifer T. Pentz for teaching us Illumina library preparation, Shweta Biliya at the High Throughput DNA Sequencing Core at Georgia Tech for sequencing the genomes of evolved strains, and Kingsley A. Boateng at Core Facilities at the Carl R. Woese Institute for Genomic Biology for the SEM imagn. Siyi Cao helped us with microscopy during the early stages of this project. We thank Christian Orlic and all members of the Ratcliff group for insightful comments on the manuscript. This work was supported by NIH grants R35-GM138030-01 to W.C.R. and R35-GM138354-02 to P.J.Y., and a Packard Fellowship for Science and Engineering to W.C.R.

## Author Contributions

G.O.B, A.Z., P.J.Y and W.C.R. conceived of the project. G.O.B., A.Z., and P.K. designed and conducted the experiments. A.Z., T.D. and P.J.Y performed the yeast biophysical simulations. G.O.B., A.Z., W.C.R. and P.Y. analyzed the data. G.O.B and A.Z. made the figures. G.O.B, W.C.R. and P.J.Y. wrote the first draft of the paper, all authors contributed to revision. E.L.D., K.T., and A.H.B. assisted G.O.B. and A.Z. with image analysis.

**Competing interests:** the authors have no competing interests to declare.

# References


1. J. T. Bonner, *The evolution of complexity by means of natural selection*. (Princeton University Press, 1988).
2. S. Smukalla *et al.*, FLO1 is a variable green beard gene that drives biofilm-like cooperation in budding yeast. *Cell* **135**, 726-737 (2008).
3. G. Bell, A. O. Mooers, Size and complexity among multicellular organisms. *Biological Journal of the Linnean Society* **60**, 345-363 (1997).
4. J. T. Bonner, Perspective: the size-complexity rule. *Evolution* **58**, 1883-1890 (2004).
5. M. Willensdorfer, Organism size promotes the evolution of specialized cells in multicellular digital organisms. *Journal of evolutionary biology* **21**, 104-110 (2008).
6. C. A. Solari, J. O. Kessler, R. E. Goldstein, A general allometric and life-history model for cellular differentiation in the transition to multicellularity. *The American Naturalist* **181**, 369-380 (2013).
7. R. Fisher, J. Shik, J. Boomsma, The evolution of multicellular complexity: the role of relatedness and environmental constraints. *Proceedings of the Royal Society B* **287**, 20192963 (2020).
8. A. H. Knoll, D. Hewitt, Phylogenetic, functional and geological perspectives on complex multicellularity. *The major transitions in evolution revisited*, 251-270 (2011).
9. J. T. Bonner, *Why size matters: from bacteria to blue whales*. (Princeton University Press, 2011).
10. A. H. Knoll, The multiple origins of complex multicellularity. *Annual Review of Earth and Planetary Sciences* **39**, 217-239 (2011).
11. A. Boudaoud, An introduction to the mechanics of morphogenesis for plant biologists. *Trends in plant science* **15**, 353-360 (2010).
12. S. Höhn, A. R. Honerkamp-Smith, P. A. Haas, P. K. Trong, R. E. Goldstein, Dynamics of a Volvox embryo turning itself inside out. *Physical review letters* **114**, 178101 (2015).
13. D. Bi, X. Yang, M. C. Marchetti, M. L. Manning, Motility-driven glass and jamming transitions in biological tissues. *Physical Review X* **6**, 021011 (2016).
14. M. Delarue *et al.*, Self-driven jamming in growing microbial populations. *Nature physics* **12**, 762-766 (2016).
15. S. Jacobeen *et al.*, Cellular packing, mechanical stress and the evolution of multicellularity. *Nature physics* **14**, 286-290 (2018).
16. M. E. Boraas, D. B. Seale, J. E. Boxhorn, Phagotrophy by a flagellate selects for colonial prey: a possible origin of multicellularity. *Evolutionary Ecology* **12**, 153-164 (1998).
17. J. H. Koschwanez, K. R. Foster, A. W. Murray, Sucrose utilization in budding yeast as a model for the origin of undifferentiated multicellularity. *PLoS Biol* **9**, e1001122 (2011).
18. W. C. Ratcliff, R. F. Denison, M. Borrello, M. Travisano, Experimental evolution of multicellularity. *Proceedings of the National Academy of Sciences* **109**, 1595-1600 (2012).
19. W. C. Ratcliff *et al.*, Experimental evolution of an alternating uni-and multicellular life cycle in Chlamydomonas reinhardtii. *Nature communications* **4**, 1-7 (2013).
20. M. D. Herron *et al.*, De novo origins of multicellularity in response to predation. *Scientific reports* **9**, 1-9 (2019).
21. J. W. Westbrook *et al.*, What makes a leaf tough? Patterns of correlated evolution between leaf toughness traits and demographic rates among 197 shade-tolerant woody species in a neotropical forest. *The American Naturalist* **177**, 800-811 (2011).

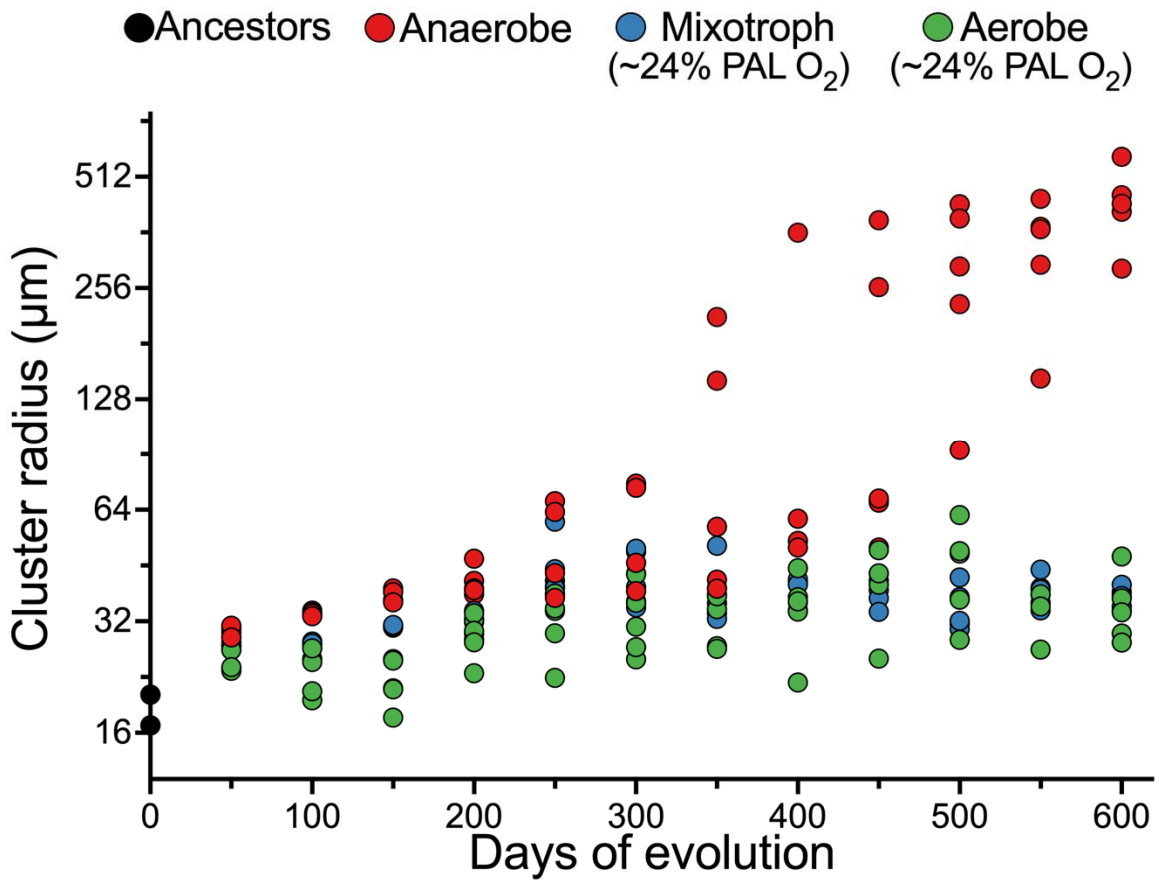
22. S. P. Veres, J. M. Lee, Designed to fail: a novel mode of collagen fibril disruption and its relevance to tissue toughness. *Biophysical journal* **102**, 2876-2884 (2012).
23. V. N. Prakash, M. S. Bull, M. Prakash, Motility-induced fracture reveals a ductile-to-brittle crossover in a simple animal's epithelia. *Nature Physics* **17**, 504-511 (2021).
24. G. O. Bozdag, E. Libby, R. Pineau, C. T. Reinhard, W. C. Ratcliff, Oxygen suppression of macroscopic multicellularity. *Nature Communications* **12**, 1-10 (2021).
25. W. C. Ratcliff, J. D. Fankhauser, D. W. Rogers, D. Greig, M. Travisano, Origins of multicellular evolvability in snowflake yeast. *Nature communications* **6**, 1-9 (2015).
26. E. C. Heinrich, M. Farzin, C. J. Klok, J. F. Harrison, The effect of developmental stage on the sensitivity of cell and body size to hypoxia in Drosophila melanogaster. *Journal of Experimental Biology* **214**, 1419-1427 (2011).
27. S. Jacobeen *et al.*, Geometry, packing, and evolutionary paths to increased multicellular size. *Physical Review E* **97**, 050401 (2018).
28. J. T. Pentz *et al.*, Ecological advantages and evolutionary limitations of aggregative multicellular development. *Current Biology* **30**, 4155-4164. e4156 (2020).
29. W. Denk, H. Horstmann, Serial block-face scanning electron microscopy to reconstruct three-dimensional tissue nanostructure. *PLoS Biol* **2**, e329 (2004).
30. C. E. Edwards, D. J. Mai, S. Tang, B. D. Olsen, Molecular anisotropy and rearrangement as mechanisms of toughness and extensibility in entangled physical gels. *Physical Review Materials* **4**, 015602 (2020).
31. E. Brown, A. Nasto, A. G. Athanassiadis, H. M. Jaeger, Strain stiffening in random packings of entangled granular chains. *Physical review letters* **108**, 108302 (2012).
32. D. M. Raymer, D. E. Smith, Spontaneous knotting of an agitated string. *Proceedings of the National Academy of Sciences* **104**, 16432-16437 (2007).
33. J. Wilhelm, E. Frey, Elasticity of stiff polymer networks. *Physical review letters* **91**, 108103 (2003).
34. Y.-J. Sheu, Y. Barral, M. Snyder, Polarized growth controls cell shape and bipolar bud site selection in Saccharomyces cerevisiae. *Molecular and cellular biology* **20**, 5235-5247 (2000).
35. M. Watanabe, D. Watanabe, S. Nogami, S. Morishita, Y. Ohya, Comprehensive and quantitative analysis of yeast deletion mutants defective in apical and isotropic bud growth. *Current genetics* **55**, 365-380 (2009).
36. R. Sopko *et al.*, Mapping pathways and phenotypes by systematic gene overexpression. *Molecular cell* **21**, 319-330 (2006).
37. M. Usaj *et al.*, TheCellMap. org: a web-accessible database for visualizing and mining the global yeast genetic interaction network. *G3: Genes, genomes, genetics* **7**, 1539-1549 (2017).
38. M. Costanzo *et al.*, A global genetic interaction network maps a wiring diagram of cellular function. *Science* **353**, (2016).
39. A. Baryshnikova, in *Computational Cell Biology*. (Springer, 2018), pp. 249-268.
40. D. H. Lee, A. Tamura, Y. Arisaka, J.-H. Seo, N. Yui, Mechanically reinforced gelatin hydrogels by introducing slidable supramolecular cross-linkers. *Polymers* **11**, 1787 (2019).
41. C. Gerhards, *Effects of type of testing equipment and specimen size on toughness of wood*. (Forest Products Laboratory, 1968), vol. 97.
42. W. W. Graessley, in *The entanglement concept in polymer rheology*. (Springer, 1974), pp. 1-179.
43. W. Chen, H. Yu, Q. Li, Y. Liu, J. Li, Ultralight and highly flexible aerogels with long cellulose I nanofibers. *Soft matter* **7**, 10360-10368 (2011).
44. N. Gravish, S. V. Franklin, D. L. Hu, D. I. Goldman, Entangled granular media. *Physical review letters* **108**, 208001 (2012).

45. L.-N. Zou, X. Cheng, M. L. Rivers, H. M. Jaeger, S. R. Nagel, The packing of granular polymer chains. *Science* **326**, 408-410 (2009).
46. R. Gottardi *et al.*, Supramolecular organization of collagen fibrils in healthy and osteoarthritic human knee and hip joint cartilage. *PloS one* **11**, e0163552 (2016).
47. M. Ryder, Structure of rhinoceros horn. *Nature* **193**, 1199-1201 (1962).
48. T. L. Hieronymus, L. M. Witmer, R. C. Ridgely, Structure of white rhinoceros (Ceratotherium simum) horn investigated by X-ray computed tomography and histology with implications for growth and external form. *Journal of Morphology* **267**, 1172-1176 (2006).
49. R. Honegger, D. Edwards, L. Axe, C. Strullu-Derrien, Fertile Prototaxites taiti: a basal ascomycete with inoperculate, polysporous asci lacking croziers. *Philosophical Transactions of the Royal Society B: Biological Sciences* **373**, 20170146 (2018).
50. T. Speck, I. Burgert, Plant stems: functional design and mechanics. *Annual review of materials research* **41**, 169-193 (2011).
51. D. Yanni *et al.*, Topological constraints in early multicellularity favor reproductive division of labor. *Elife* **9**, e54348 (2020).
52. B. T. Larson *et al.*, Biophysical principles of choanoflagellate self-organization. *Proceedings of the National Academy of Sciences* **117**, 1303-1311 (2020).
53. T. Brunet *et al.*, Light-regulated collective contractility in a multicellular choanoflagellate. *Science* **366**, 326-334 (2019).
54. M. J. Dayel *et al.*, Cell differentiation and morphogenesis in the colony-forming choanoflagellate Salpingoeca rosetta. *Developmental biology* **357**, 73-82 (2011).
55. P. C. Hinkle, M. A. Kumar, A. Resetar, D. L. Harris, Mechanistic stoichiometry of mitochondrial oxidative phosphorylation. *Biochemistry* **30**, 3576-3582 (1991).
56. J. Raymond, D. Segrè, The effect of oxygen on biochemical networks and the evolution of complex life. *Science* **311**, 1764-1767 (2006).
57. T. W. Lyons, C. T. Reinhard, N. J. Planavsky, The rise of oxygen in Earth's early ocean and atmosphere. *Nature* **506**, 307-315 (2014).
58. R. D. Gietz, R. H. Schiestl, High-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method. *Nature protocols* **2**, 31-34 (2007).
59. H. Li, Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv preprint arXiv:1303.3997*, (2013).
60. A. McKenna *et al.*, The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome research* **20**, 1297-1303 (2010).
61. E. Garrison, G. Marth, Haplotype-based variant detection from short-read sequencing. *arXiv preprint arXiv:1207.3907*, (2012).
62. P. Danecek *et al.*, The variant call format and VCFtools. *Bioinformatics* **27**, 2156-2158 (2011).
63. H. Thorvaldsdóttir, J. T. Robinson, J. P. Mesirov, Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Briefings in bioinformatics* **14**, 178-192 (2013).
64. P. Danecek *et al.*, Twelve years of SAMtools and BCFtools. *Gigascience* **10**, giab008 (2021).
65. P. Cingolani *et al.*, A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. *Fly* **6**, 80-92 (2012).
66. M. W. Hahn, *Molecular population genetics*. (Oxford University Press, 2018).
67. S. Kryazhimskiy, J. B. Plotkin, The population genetics of dN/dS. *PLoS genetics* **4**, e1000304 (2008).
68. J. M. Cherry *et al.*, Saccharomyces Genome Database: the genomics resource of budding yeast. *Nucleic acids research* **40**, D700-D705 (2012).

69. T. J. Deerinck *et al.*, High-performance serial block-face SEM of nonconductive biological samples enabled by focal gas injection-based charge compensation. *Journal of microscopy* **270**, 142-149 (2018).
70. H. T. Ngo, C. S. Yin, Luteimonas terrae sp. nov., isolated from rhizosphere soil of Radix ophiopogonis. *International journal of systematic and evolutionary microbiology* **66**, 1920-1925 (2016).
71. S. A. Zamani-Dahaj *et al.*, Spontaneous emergence of multicellular heritability. *bioRxiv*, (2021).

675 **Supplementary Figures**

676




677

678 **Figure S1**. Temporal dynamics of size evolution in each population and treatment group. Data
679 points show the weighted average radius of cluster size for the whole population (see Methods
680 for details).

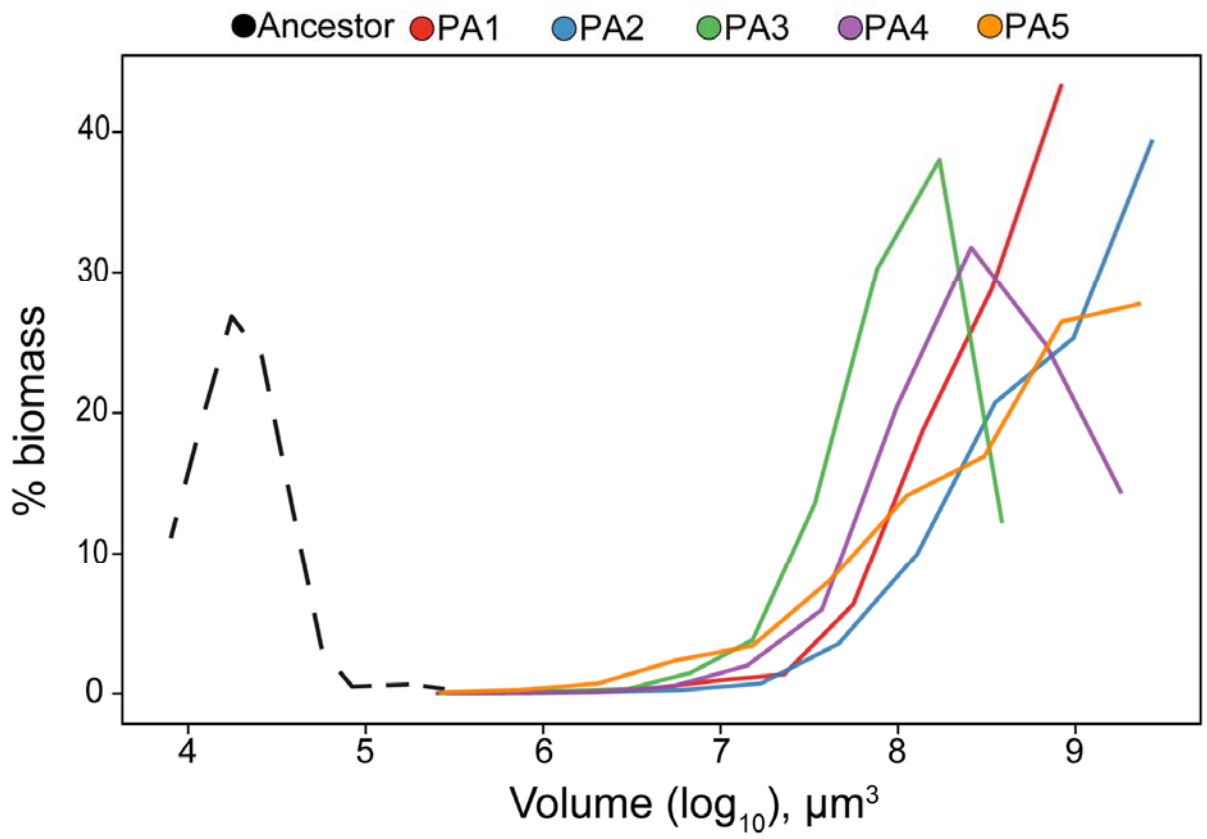
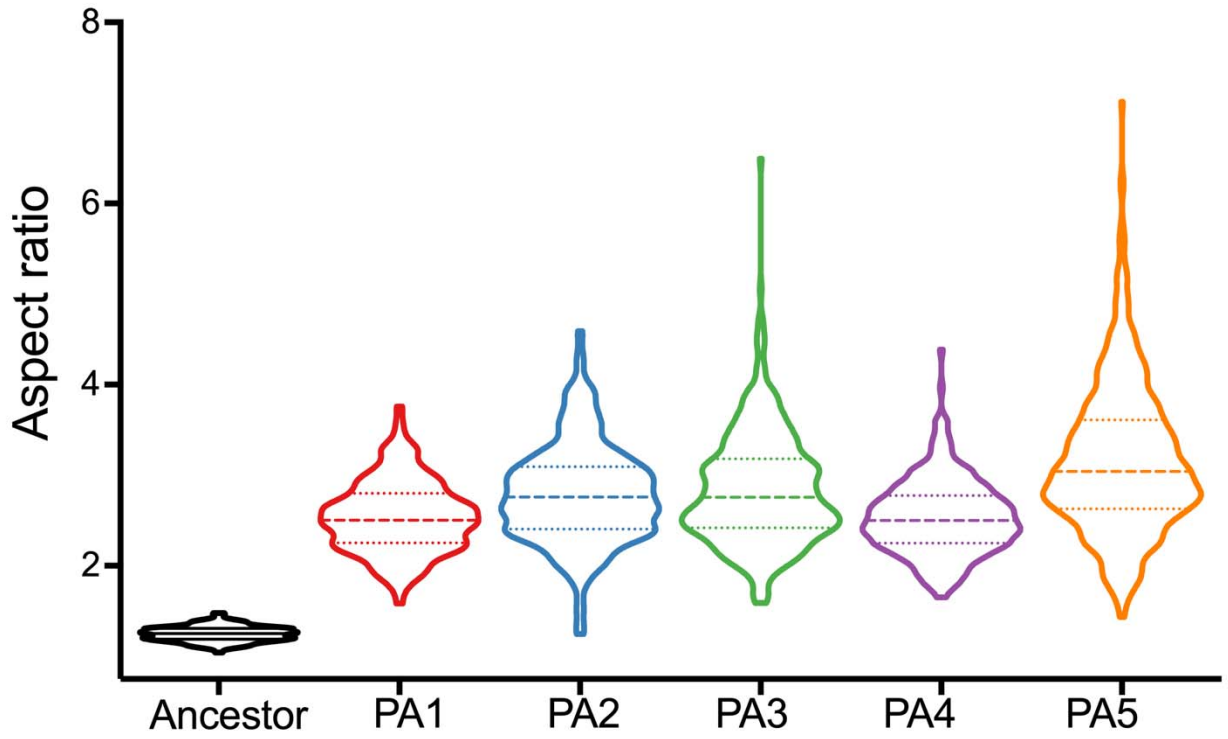



**Figure S2**. Biomass distribution as a function of group size for the ancestral snowflake yeast (dotted line) and 600 day evolved populations of PA1-PA5. The 'weighted mean size' used in Figures 1, 2 and 4 is the mean of the biomass distribution.

686




687

688 **Figure S3**. Distribution of aspect ratios for ancestral and 600-day evolved populations of

689 anaerobic snowflake yeast.

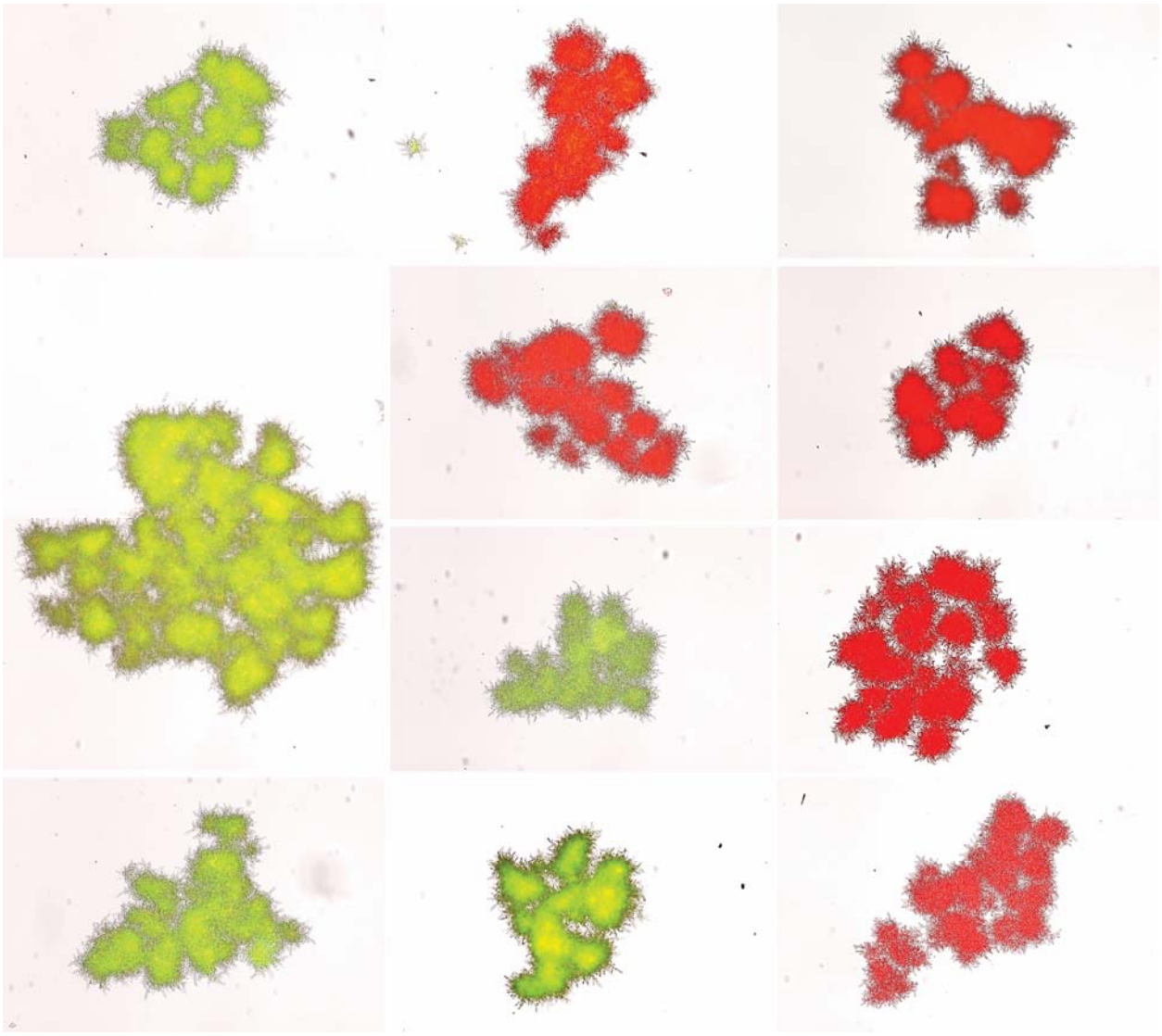


**Figure S4**. Macroscopic snowflake yeast are monoclonal, growing via permanent mother-daughter cellular bonds, not aggregation. We co-cultured GFP and RFP-tagged genotypes of a macroscopic single strain isolate (PA2, strain ID: GOB1413-600) for 5 days, then imaged 70 clusters on a Nikon Ti-E. Shown are a composite of 11 individual clusters, which all remain entirely green or red. Individual clusters were compressed with a coverslip for imaging, resulting in their fragmentation into multiple modules.

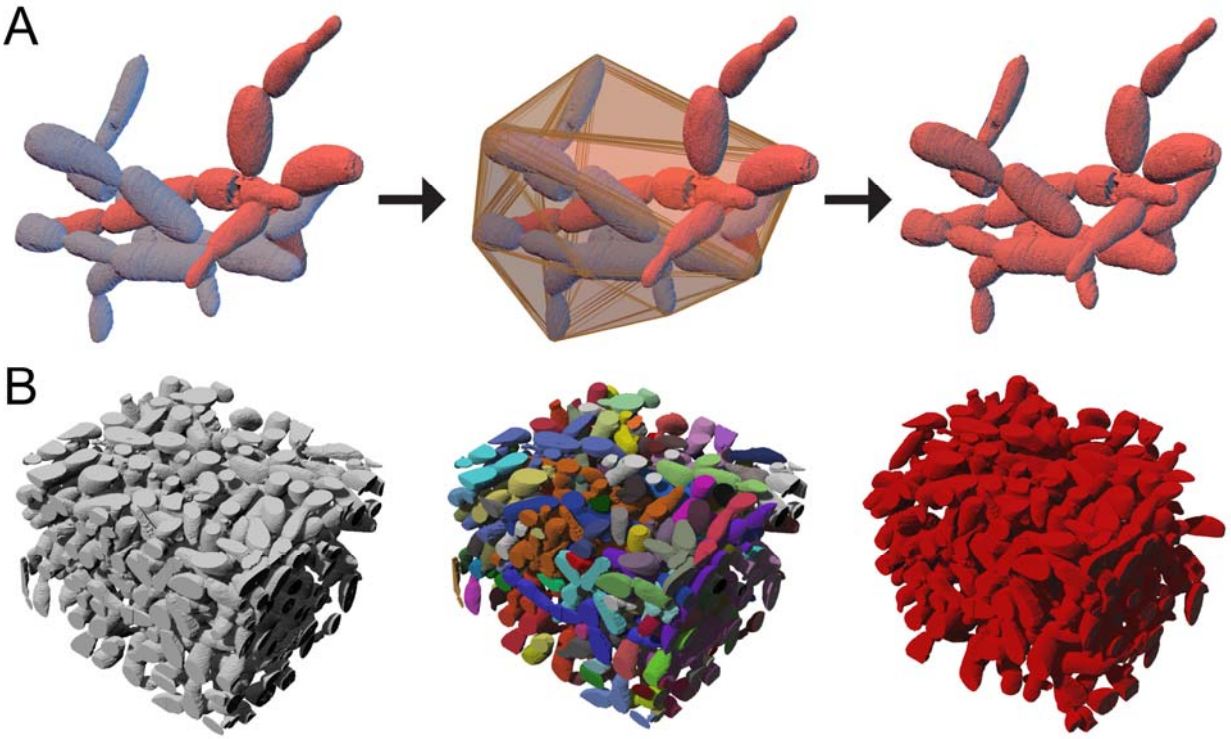



**Figure S5**. Quantifying entanglement via analysis of the topology and geometry of a snowflake yeast cluster. (A) We measured entanglement of individual components by fitting a convex hull around each component, and determining whether the other component overlaps with the space bounded by this convex hull. Here we just show the convex hull for the blue component, which overlaps with the red component. These components are thus part of the same entangled component. (B) Using this approach, we identified the components within a sub-volume of a macroscopic snowflake yeast, and used a percolation analysis to examine the fraction of the biomass that is part of the same entangled component (colored in red).