Article

De novo evolution of macroscopic multicellularity

https://doi.org/10.1038/s41586-023-06052-1

Received: 1 September 2021

Accepted: 5 April 2023

Published online: 10 May 2023

Check for updates

G. Ozan Bozdag[1,6]✉, Seyed Alireza Zamani-Dahaj[2,3,6], Thomas C. Day[3], Penelope C. Kahn[1,4], Anthony J. Burnetti[1], Dung T. Lac[1], Kai Tong[1,2], Peter L. Conlin[1], Aishwarya H. Balwani[5], Eva L. Dyer[5], Peter J. Yunker[2]✉ & William C. Ratcliff[1]✉

While early multicellular lineages necessarily started out as relatively simple groups of cells, little is known about how they became Darwinian entities capable of sustained multicellular evolution[1–3]. Here we investigate this with a multicellularity long-term evolution experiment, selecting for larger group size in the snowflake yeast (*Saccharomyces cerevisiae*) model system. Given the historical importance of oxygen limitation[4], our ongoing experiment consists of three metabolic treatments[5]—anaerobic, obligately aerobic and mixotrophic yeast. After 600 rounds of selection, snowflake yeast in the anaerobic treatment group evolved to be macroscopic, becoming around 2×10^4 times larger (approximately mm scale) and about 10^4-fold more biophysically tough, while retaining a clonal multicellular life cycle. This occurred through biophysical adaptation—evolution of increasingly elongate cells that initially reduced the strain of cellular packing and then facilitated branch entanglements that enabled groups of cells to stay together even after many cellular bonds fracture. By contrast, snowflake yeast competing for low oxygen[5] remained microscopic, evolving to be only around sixfold larger, underscoring the critical role of oxygen levels in the evolution of multicellular size. Together, this research provides unique insights into an ongoing evolutionary transition in individuality, showing how simple groups of cells overcome fundamental biophysical limitations through gradual, yet sustained, multicellular evolution.

Organismal size has a fundamental role in the evolution of multicellularity. The evolution of larger size enables organisms to gain protection from the external environment[6] and explore new niches[7], while creating opportunities for the evolution of cellular differentiation[8,9]. Increases in organismal size have also been hypothesized to have an important role in the evolution of trade-off-breaking multicellular innovations, as a large size creates an evolutionary incentive to solve challenges of nutrient and oxygen transportation that are otherwise inescapable consequences of diffusion limitations[10,11]. 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 drive sustained multicellular evolution.

The evolution of macroscopic size presents a fundamental challenge to nascent multicellular organisms, requiring the evolution of biophysical solutions to evolutionarily new stresses that act over previously unseen multicellular size scales[12,13]. Although previous studies 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[14–16]. Extant macroscopic multicellular organisms have solved the above challenges through developmental innovation, evolving mechanisms that either reduce the accumulation of biophysical strain or increase multicellular toughness[17,18]. However, in nascent multicellular organisms that have not yet evolved coordinated morphogenesis, we do not know how, or even whether, simple groups of cells 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 subjected snowflake yeast[19], a model of undifferentiated multicellularity, to 600 rounds (about 3,000 generations) of daily selection for increased size. Furthermore, because oxygen is thought to have had a key role in the evolution of macroscopic multicellularity, we evolved snowflake yeast with anaerobic, mixotrophic or obligately aerobic metabolism. All five of our replicate anaerobic populations evolved macroscopic size, whereas all aerobic and mixotrophic populations remained microscopic throughout the experiment, supporting the hypothesis that growth under low concentrations of oxygen constrains the evolution of large multicellular size[5]. Macroscopic size convergently evolved through two key changes in all five replicate populations. First, snowflake yeast increased the length of their constituent cells, which delays organismal fracture caused by packing-induced strain[13]. Next, they evolved to entangle branches of connected cells such that breaking a single cell–cell bond

[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. [6]These authors contributed equally: G. Ozan Bozdag, Seyed Alireza Zamani-Dahaj. ✉e-mail: ozan.bozdag@gmail.com; peter.yunker@gatech.edu; ratcliff@gatech.edu

no longer causes multicellular fracture, evolving to become around 2×10^5 times larger, forming millimetre-scale groups of clonal cells. Together these traits increased the toughness of individual clusters by more than 10^4-fold, transforming the initial snowflake yeast ancestor, which was weaker than gelatin, into an organism with the strength and toughness of wood. Fitness assays, sequencing and synthetic strain constructions reveal that macroscopic multicellularity evolved through selection acting on group size, an emergent multicellular trait of mutations directly affecting cellular morphology.

The multicellularity long-term evolution experiment

In 2018, we initiated the multicellularity long-term evolution experiment (MuLTEE)[5], named after the pioneering long-term evolution experiment with *E. coli* initiated by R. Lenski[20]. The central goal of this project, which we intend to run over decadal time scales, is to observe open-ended multicellular evolution in a nascent multicellular organism. We began the MuLTEE by engineering a unicellular isolate of *S. cerevisiae* strain Y55 to grow with the snowflake phenotype by deleting the *ACE2* open reading frame, ensuring that each replicate population had the same initial mechanism of group formation[21]. To examine the effect of oxygen on the evolution of size, we initiated three treatments in an otherwise isogenic ancestor: anaerobic growth (generated by selecting for a spontaneous petite mutant incapable of respiration), mixotrophy (cultured with glucose as the primary carbon source) and obligately aerobic growth (cultured with glycerol as the primary carbon source)[5]. We refer to the five replicate populations of anaerobic, mixotrophic and obligately aerobic populations as PA1–5, PM1–5 and PO1–5, respectively. We maintained strong directional selection favouring larger cluster size throughout the experiment by selecting for increasingly rapid sedimentation before transfer to fresh medium (Methods). We evolved these 15 populations over 600 rounds of growth and settling selection (approximately 3,000 generations; Fig. 1a).

Evolution of macroscopic size

All five populations of anaerobic snowflake yeast evolved macroscopic size, with individual clusters visible to the naked eye (Fig. 1b–d and Supplementary Video 1). By contrast, snowflake yeast with the ability to metabolize oxygen remained microscopic (Fig. 1a and Extended Data Fig. 1), a result that is consistent with recent research showing that competition for scarce oxygen imposes a powerful constraint on the evolution of large multicellular size[5]. Here we focus on the evolution of macroscopic size in the five replicate anaerobic populations. Yeast in this treatment group increased their mean cluster radius from 16 μm to 434 μm, an approximately 2×10^4-fold increase in volume (Fig. 1e). This corresponds to an estimated increase from around 100 cells per cluster to about 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 macroscopic snowflake yeast are over a millimetre in diameter (Fig. 1c), which is comparable to the size of an adult *Drosophila*[22]. Much like their microscopic snowflake yeast ancestor[13,19], macroscopic snowflake yeast possess a life cycle in which groups of cells both grow in size and reproduce, generating multicellular propagules, over the course of the approximately 24 h culture period (Extended Data Fig. 2). This analysis establishes that group size is heritable, and our time-series data (Fig. 1e) show that every replicate population evolved to form larger clusters at each 200-day sampling interval, strongly suggesting that larger size is an adaptive trait evolving in response to settling selection. To test this hypothesis, we performed a fitness assay competing the microscopic ancestor against each day 600 (t600) PA1–5 population, under our standard selective conditions of growth and settling selection. The 600-day evolved macroscopic snowflake yeast were far more fit than their ancestor (mean daily selection rate constant = 2.5; Fig. 1f), increasing from a mean starting frequency of 52% to 99.9% over just 3 days.

As with their ancestor, macroscopic snowflake yeast grow through incomplete mother–daughter cellular separation, forming a branched tree-like structure (Fig. 2a,b). When compressed, macroscopic clusters fracture into small modules that resemble microscopic snowflake yeast in terms of branching morphology (Fig. 2a,b). However, their cellular morphology changed markedly. Throughout the experiment, snowflake yeast cells evolve to be more elongate across all five replicate populations, increasing in average aspect ratio (ratio of length to width) from ~1.2 to ~2.7 (Fig. 2c,d and Extended Data Fig. 3b). Even in macroscopic snowflake clusters, cell size and shape did not depend on location in the cluster (that is, interior or exterior; Extended Data Fig. 4). Initially, cluster size was an approximately linear function of cellular aspect ratio (Fig. 2e (inset)), but this relationship changes once they evolve macroscopic size (Fig. 2e).

Previous research showed that the evolution of more-elongate cells increases the size to which microscopic snowflake yeast grow by decreasing the density of cellular packing (that is, their packing fraction) in the cluster interior, which reduces cell–cell collisions that drive multicellular fracture[13]. To establish a null expectation for the effect of cell aspect ratio on cluster packing fraction, we simulated the growth of individual clusters from a single cell using an experimentally validated model[13] (Methods). In these simple 3D simulations, the cellular packing fraction decreased monotonically with increasing cellular aspect ratio (Fig. 2f). We next examined this relationship over the course of our long-term experiment in replicate population two (PA2), which was one of the first lineages to evolve macroscopic size. As predicted by our simulation, cellular elongation decreased the packing fraction of microscopic multicellular groups—but only initially, from an aspect ratio of around 1.2 to 2.0. Beyond this, clusters with more elongate cells actually became more densely packed, and the experimentally measured packing fraction became increasingly divergent from the model predictions (Fig. 2f). This divergence suggests that this lineage evolved a biophysical mechanism for increased multicellular toughness, with the ability to withstand growth to macroscopic size and a high cellular packing fraction.

The simplest way that snowflake yeast could evolve to become macroscopic is to become adhesive, forming large aggregates of many separate snowflake yeast clusters. Indeed, aggregation is a common mechanism of group formation in yeast (that is, through flocculation[6]), and this would explain the modular structure of macroscopic snowflake yeast (Fig. 2a). To determine whether clusters of macroscopic yeast form through aggregation, or whether they develop as a single clonal lineage, we labelled a single-strain isolate (taken from PA2 after 600 days of selection), 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 fluorescent subclusters. However, after five rounds of co-culture, all of the multicellular clusters ($n = 70$; Extended Data Fig. 6) remained monoclonal. This is unlikely to occur with aggregation. If we conservatively assume that each macroscopic snowflake yeast cluster that we measured was 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 therefore does not explain the evolution of macroscopic size in snowflake yeast.

Evolution of branch entanglement

To examine how changes in the topology of macroscopic snowflake yeast may underlie their increased size, we imaged clusters using serial block face scanning electron microscopy (SBF-SEM). This technique is capable of imaging the interior of macroscopic clusters that are difficult to resolve using light-based microscopy, allowing us to map their

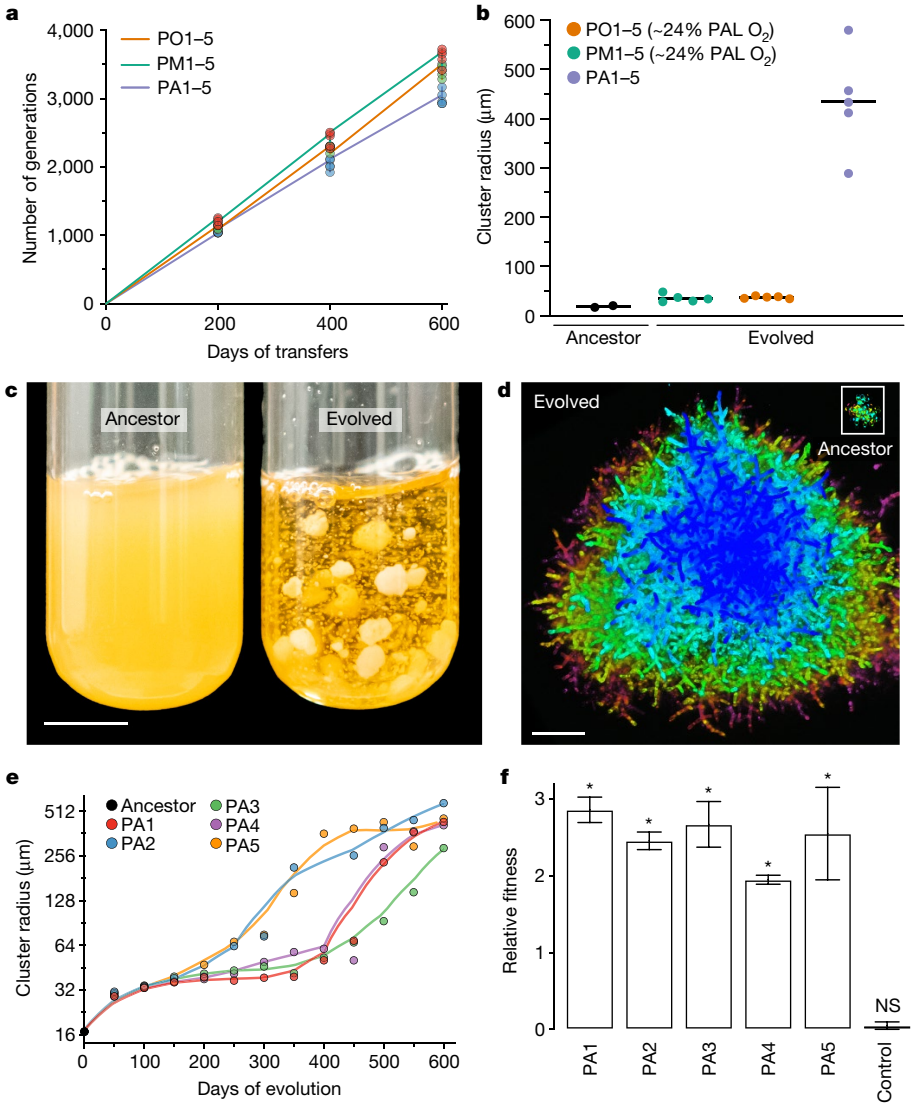


Fig. 1 | Evolution of macroscopic multicellularity in five replicate snowflake yeast populations. a, We selected for larger size over 600 daily transfers, representing about 3,000 generations. **b**, Only the anaerobic populations (PA1–5) evolved macroscopic size over this time. PAL, present atmospheric level. **c**, Individual snowflake yeast clusters from t600 are visible to the naked eye. Scale bar, 10 mm. **d**, Representative clusters of evolved and ancestral genotypes (the ancestor is shown at the top right) shown under the same magnification (colour represents depth in the *z* plane). Scale bar, 50 µm. **e**, Temporal dynamics of size evolution in the anaerobic treatment (PA), showing a considerable increase in the mean cluster radius ($P < 0.0001$; $F_{5,13321} = 2{,}100$, one-way analysis of variance (ANOVA) with Dunnett's test, comparing t600 to t*t*0). **f**, Macroscopic snowflake yeast were considerably more fit, calculated as the per-day selection rate constant[20] compared with their microscopic ancestor ($F_{5,12} = 39.5$, $P < 0.0001$ (asterisks), one-way ANOVA with Dunnett's test, comparing each evolved isolate with their ancestor). NS, not significant. For each genotype, we performed three replicate fitness assays. Error bars represent s.d. In **a**, the number of generations was estimated by measuring the average daily dilution factor for 15 populations, each with 3 technical replicates, across 3 timepoints, resulting in a total of 45 samples. In **b** and **e**, data points represent the biomass-weighted mean radius (Methods) calculated by measuring the size of an average of 1,150 snowflake yeast clusters per sample population (3 ancestors + 3 treatment groups × 5 replicate populations × 12 time points = 183 samples). See Extended Data Fig. 1 for additional data on the evolution of cluster size in oxygen-using populations (PM and PO) and Extended Data Fig. 3a for cluster size distributions for the 600-day anaerobic populations (PA1–5). The lines in **e** are Lowess-smoothed curves for visual purposes. The data underlying **a**, **b**, **e** and **f** are available at GitHub (https://github.com/ozanbozdag/de_novo_evolution_of_macroscopic_multicellularity).

internal architecture at nanometre precision[23]. 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. Furthermore, we found that individual macroscopic snowflake yeast were not composed of a single topologically connected component like their ancestors. Instead, they contained disconnected branches of cells, suggesting that the cluster remained intact even when cell–cell connections were broken (Fig. 3a). On the basis of these observations, we hypothesized that branches of macroscopic clusters are entangled in a manner that is reminiscent of physical gels[24] and entangled granular materials[25]. Entanglement would provide a mechanism for branches of cells to remain in the same, densely packed group even after cell–cell bonds break.

Following previous studies in entangled chains and knotted strings[25,26], we used our SBF-SEM dataset to quantify branch entanglement in macroscopic snowflake yeast by analysing chain topology and geometry. Specifically, we constructed the convex hull of each connected component within a subvolume, which denotes the smallest convex polyhedron containing this component (Extended Data

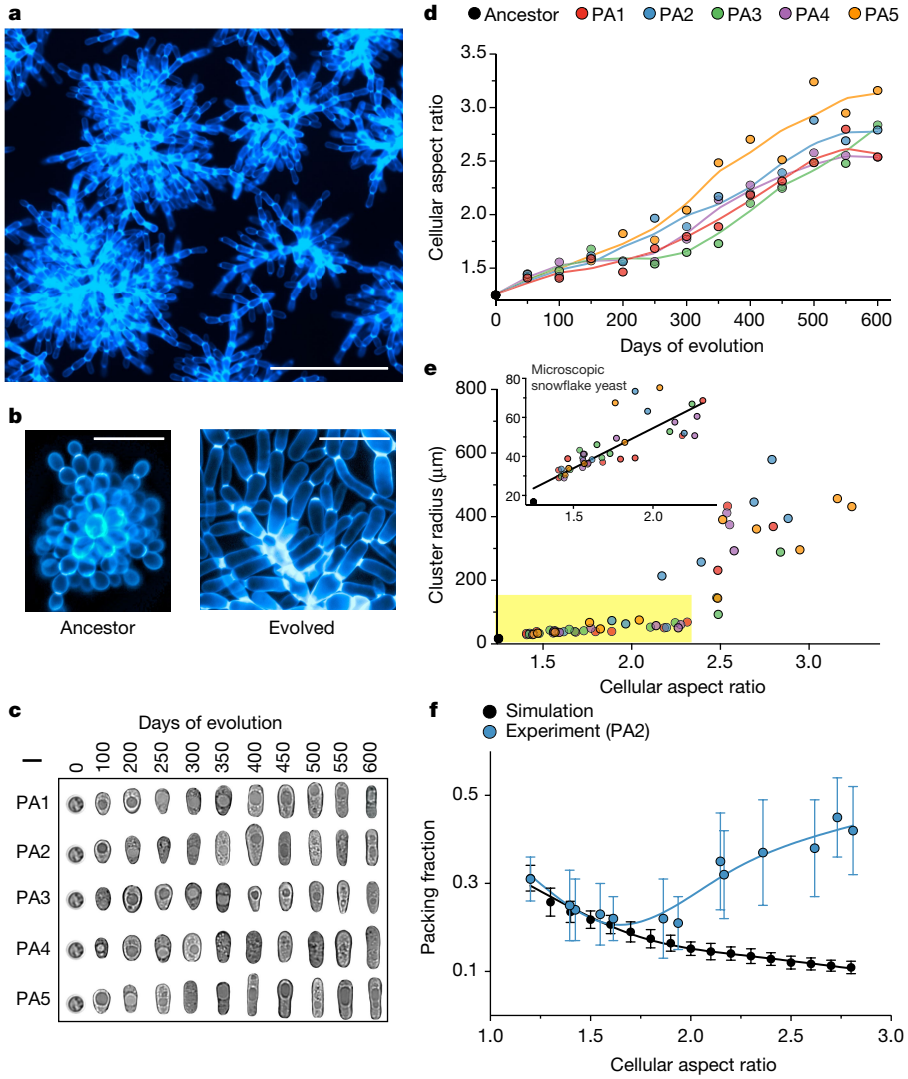


Fig. 2 | Evolution of new cell morphology. a, When compressed, macroscopic snowflake yeast fracture into modules. **b**, Macroscopic snowflake yeast grow through incomplete mother–daughter cellular separation, retaining the branched tree-like growth form of their microscopic ancestor. Cell walls are stained with calcofluor white in **a** and **b**. Scale bars, 100 μm (**a**) and 20 μm (**b**). **c**,**d**, Images (**c**) and quantification (**d**) of the parallel evolution of elongated cell shape (the ancestral genotype is the same in all replicate populations), resulting in an increase in the average aspect ratio from about 1.2 to 2.7 ($F_{5,1,993} = 206.2$, $P < 0.0001$, one-way ANOVA, comparing t600 and t0; Dunnett's test was applied, comparing each t600 population to the ancestor ($P < 0.0001$)). An expanded version of **c** is shown in Extended Data Fig. 5. For **c**, scale bar, 5 μm. **e**, Early in their evolution (aspect ratio, 1–2.3 (yellow shaded region)), cluster size (weighted mean radius) is an approximately linear function of cellular aspect ratio (inset; $P < 0.0001$, $y = 41.1x - 27.8$, $r^2 = 0.72$, linear regression analysis). This relationship does not hold beyond an aspect ratio of around 2.5. **f**, A biophysical model of snowflake yeast predicts that an increasing cellular aspect ratio should decrease the cellular packing fraction (black points). We observed a close correspondence with these predictions for low aspect ratios, but our experimental data diverge from model predictions for aspect ratios beyond 2. For each point in **d**, 453 cells per population were measured on average (1 ancestor +5 replicates × 12 timepoints, each separated by 50 days = 61 samples). Each data point in **f** reports the mean of 15 snowflake yeast clusters or 25 replicate simulations. Error bars represent s.d.

Fig. 7a). If a cell from one connected component overlaps with the convex hull of a second, then the two can be considered to be entangled. By percolating entanglement among adjacent connected components throughout the subvolume, we can measure the extent to which the cluster's biomass is mutually entangled (Fig. 3b and Extended Data Fig. 7b). 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[27]. In analyses of ten randomly selected subvolumes from different macroscopic snowflake yeast clusters from population PA2 t600 macroscopic yeast, we found that the largest entangled component contained 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 fractures.

Entanglement increases multicellular toughness

As a further test, we investigated the mechanics of macroscopic snowflake yeast. Entangled materials are known to exhibit two key mechanical signatures: strain stiffening and high material toughness[25,28]. Strain stiffening describes the fact that, when compressed, the effective stiffness of entangled chains increases with increased strain. By efficiently distributing stress across constituent bonds, entangled materials can withstand stress orders-of-magnitude greater than their non-entangled counterparts[28,29]. As the microscopic ancestor is presumably not

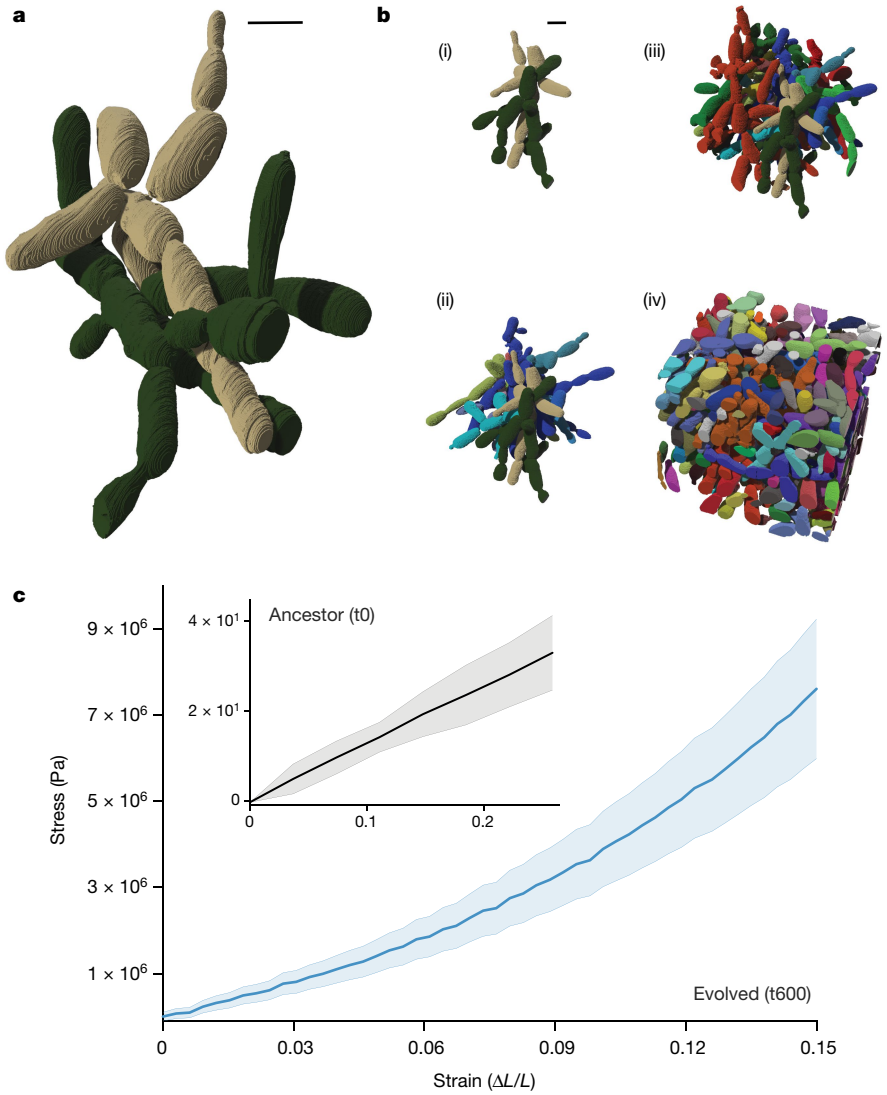


Fig. 3 | Branch entanglement underlies the evolution of macroscopic size. **a**, Two entangled components (green and tan), obtained using SBF-SEM imaging. **b**, Branch entanglement is pervasive in macroscopic snowflake yeast. Starting with the two-component subvolume in **a**, we percolated entanglement by adding on adjacent entangled components in four steps (i–iv). For **a** and **b**, scale bars, 5 µm. **c**, Stress versus strain plot for macroscopic snowflake yeast (PA2, t600) clusters in blue and the ancestor in grey (the ancestor is shown again in the inset with a rescaled *y* axis). Macroscopic snowflake yeast experience strain stiffening, a hallmark of entangled systems, whereas the ancestor's stress–strain plot is linear (that is, $r^2 = 0.97 \pm 0.02$), which is expected for non-entangled systems. The shaded area shows s.d. based on ten repeated measurements for each.

entangled, it should not exhibit strain-stiffening behaviour or possess high toughness. We measured the mechanical stress response of ten macroscopic snowflake yeast clusters under uniaxial compression using a macroscopic mechanical tester (Zwick Roell Universal Testing Machine). We repeated the same experiment for ten ancestral microscopic snowflake yeast clusters using an atomic force microscope (AFM Workshop LS-AFM). The stress–strain plot for the microscopic ancestor is linear (Fig. 3c (inset)), clusters fracture at stress as low as 240 Pa and have toughness as low as 8.9 J m^{-3} (ref. 13). By contrast, macroscopic snowflake yeast clusters have a convex stress–strain curve (Fig. 3c), can support stresses at least as large as around 7 MPa without failing and have toughness greater than 0.6 MJ m^{-3}. Thus, entanglement enables separate branches within macroscopic snowflake yeast to stay together and endure the stresses necessary for growth to large size.

To rule out alternative hypotheses, we performed additional measurements on macroscopic snowflake yeast (PA2 t600) and their microscopic ancestor. First, we measured the stiffness of individual cells to determine whether they were becoming tougher. We did not detect a change between individual cells of the ancestor and the t600 macroscopic strain (0.018 N m^{-1} and 0.019 N m^{-1} for the ancestor and evolved strain, respectively; Extended Data Fig. 8a). To determine whether, in the absence of macroscopic entangled structures, individual PA2 t600 modules still show strain-stiffening behaviour, we crushed macroscopic clusters into smaller, microscopic branches before compressing them. These small groups displayed a linear stress–strain curve like their unevolved microscopic ancestor. In the absence of their macroscopic phenotype, the modules of the PA2 t600 yeast do not behave like an entangled material (Extended Data Fig. 8b).

Finally, we performed a positive control—we created persistently entangled groups experimentally. As described in Extended Data Fig. 6, growth in well-mixed liquid medium prevents the formation of chimeric groups through entanglement. Yet, if entanglement is critical for multicellular toughness, enabling fractured branches to remain in the same group, then chimeric clusters held together only by entanglement should be possible to grow under the right environmental conditions. We allowed GFP- and RFP-tagged versions of PA2 t600 to

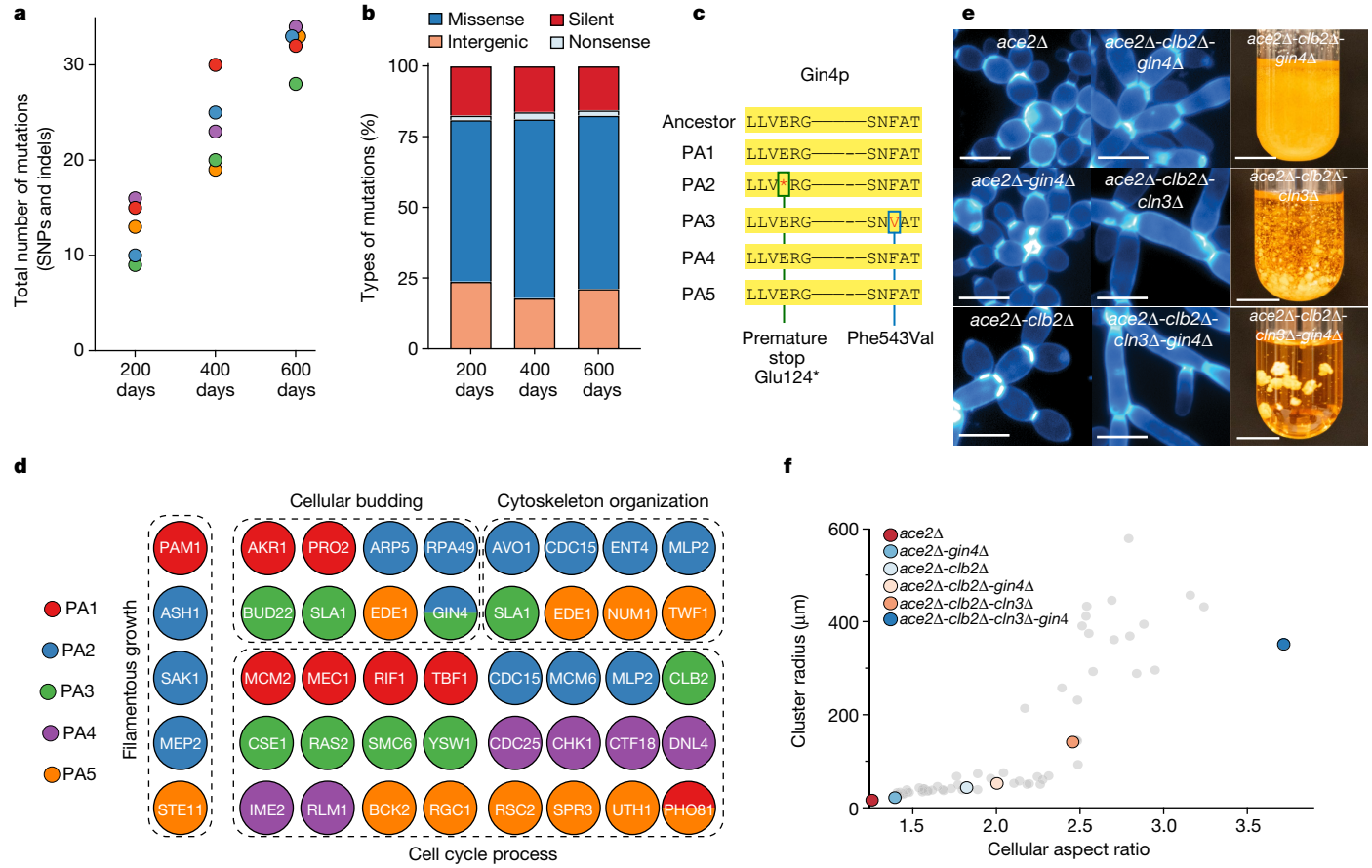


Fig. 4 | Whole-genome sequencing reveals the dynamics of molecular evolution and the genetic basis of cell-level and cluster-level changes. **a**,**b**, The number (**a**) and types (**b**) of mutations in evolved single strains from each population. SNPs, single-nucleotide polymorphisms; indels, small insertion–deletion mutations. **c**, Parallel evolution of *GIN4*, a kinase of which the loss of function increases bud neck size (further details are available at GitHub; https://github.com/ozanbozdag/de_novo_evolution_of_macroscopic_multicellularity); *GIN4* was mutated in two independent populations. **d**, Macroscopic snowflake yeast were enriched in mutations affecting cell cycle progression, cytoskeleton and filamentous growth. We also found mutations affecting budding (that is, the location of buds on the cell surface and bud neck size). **e**, Representative images of cells from strains used to re-engineer macroscopic size. Scale bars, 10 μm. **f**, Engineered strains recapitulated the evolutionary trajectory established over 600 rounds of selection. With a cellular aspect ratio of below around 2.5, snowflake yeast remained microscopic, greatly increasing in size beyond this threshold (experimentally evolved strains described in Fig. 1e are shown in grey to facilitate direct comparison). Scale bars, 10 mm. In **f**, the mean cluster size and aspect ratio of six genotypes were calculated by measuring an average of 1,132 multicellular clusters and 1,205 individual cells per genotype (further details are available at GitHub; https://github.com/ozanbozdag/de_novo_evolution_of_macroscopic_multicellularity).

grow at high density on solid medium for 48 h, then cultured these yeast in liquid medium for two rounds of growth and settling selection. These yeast readily formed and maintained chimeric groups. Specifically, 30% (31 out of 101) of the clusters of the macroscopic genotype were still chimeric, with visibly entangled branches of green and red yeast (Extended Data Fig. 9). By contrast, only 1 out of 110 clusters of the ancestral genotype were chimeric when cultured under the same conditions ($\chi^2 = 36.1$, d.f. = 1, $n = 211$, $P < 0.0001$). Taken together, this experiment shows that entanglement enables evolved snowflake yeast to remain intact, even when constituent branches lack continuous mother–daughter cellular bonds (that is, red and green branches are not attached to each other by permanent bonds).

Genetic basis of multicellular evolution

To uncover the genetic basis of the cell-level changes underlying multicellular evolution, 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). Over around 3,000 generations, snowflake yeast in our anaerobic treatment evolved substantially more-elongate cells (Fig. 2c,d), which has a central role in the evolution of increased cluster size (Fig. 2e) and biophysical toughness (Figs. 2f and 3). Gene Ontology (GO) terms associated with cell length were significantly enriched, namely genes of the cell cycle[30] (29 out of 123 mutations; $P = 0.02$) and filamentous growth (7 out of 123 mutations). Moreover, we found 11 nonsynonymous mutations in genes with known roles in cellular budding (Fig. 4d), including eight genes that have previously been shown to increase the size of the bud neck (*AKR1*, *ARP5*, *CLB2*, *GIN4*, *PRO2*, *RPA49*, *RSC2* and *PHO81*)[31,32]. Mutations arose in two of these genes in different populations (that is, *PHO81* in populations PA1 and PA5, and *GIN4* in populations PA2 and PA3; Fig. 4c), indicating parallel evolution. Larger bud scars should increase the amount of cell wall connecting cells, increasing the strength of the bond and toughness of the group. In our ancestral strain Y55, *gin4Δ*

cells formed bud necks that had a 1.65-fold larger cross-sectional area ($P < 0.0001$, $U = 14$, Mann–Whitney U-test). Consistent with this, we found that PA2 t600 macroscopic snowflake yeast evolved to form bud necks that had a 2.4-fold larger cross-sectional area, and bud scars that had a 5.8-fold greater three-dimensional volume compared with their microscopic ancestor (Extended Data Fig. 10).

Recapitulating macroscopic size through engineering

As a final proof of principle, we set out to show that cellular elongation, aided by increased cell–cell bond strength, is sufficient to underpin the origin of macroscopic size in snowflake yeast. Starting with the microscopic *ace2Δ* ancestor, we deleted the cyclins *CLN3* and *CLB2* to artificially increase the cellular aspect ratio, and deleted *GIN4* to increase the bud scar size and, therefore, strength. *CLN3*, although not present in evolved isolates, has a large, well-understood phenotypic effect on cell shape. Deleting *CLN3* and *CLB2* increased the cellular aspect ratio by 21% and 45%, respectively, in single mutants, and 95% in the double mutant, with *GIN4* deletion further increasing the aspect ratio of each genotype in addition to its effects on bud scars (Fig. 4e,f). Our results mirror those from our evolution experiment: strains with an aspect ratio of less than 2.5 were clearly microscopic, with an increasing aspect ratio resulting in a gradual increase in group size. At an aspect ratio of around 2.5, *ace2Δ-clb2Δ-cln3Δ* yeast were at the threshold of macroscopic size, but still quite a bit smaller than our t600 isolates. The *ace2Δ-clb2Δ-cln3Δ-gin4Δ* mutant, with an aspect ratio of 3.7, formed well-developed macroscopic clusters (Fig. 4f).

Discussion

Here we show that snowflake yeast, a model system of undifferentiated multicellularity, are capable of sustained multicellular evolution. Over 600 daily rounds of settling selection, they gradually evolved larger size, eventually forming macroscopic clusters containing hundreds of thousands of clonal cells. They achieved this considerable increase in size by evolving highly elongate cells that become entangled within the cluster interior. This critical innovation enables 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 around 100-fold weaker than gelatin[33] to having the strength and toughness of wood[34]. Sequencing revealed an enrichment in mutations affecting the cell cycle and budding—traits that increase cell length and the amount of cell wall material at the point of attachment. Engineered strains with mutations increasing cell length and bud scar size recapitulated our evolutionary progression from microscopic to macroscopic size.

In our system, new multicellular traits arise as an emergent property of changes in cell-level traits. Two cell-level innovations appear to have played a key role in the evolution of macroscopic size: more-elongate cells and larger-bud scars. Increased cell length initially reduces the 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. Larger bud scars increase the amount of shared cell wall connecting cells, which, all else equal, should increase multicellular toughness by strengthening cell–cell bonds[35]. While the evolution of larger, tougher multicellular groups necessarily has underlying cell-level causation, these group- and cell-level traits are distinct and non-commensurable (that is, group size and toughness cannot be measured at the single-cell level). This demonstrates that snowflake yeast are evolving under multi-level selection 2 (MLS2), a shift in evolutionary dynamics that is critical for the transition multicellular individuality, as it enables groups as a whole, and not just their constituent members, to gain adaptations[36].

Entanglement is a common mechanism through which filamentous materials can solidify. It can operate on nearly any length scale, ranging from nanoscale polymers[37] and nanofibres[38] to macroscopic staples[39] and beaded chains[40]. Relatively little is known about the role of entanglement in the materials properties of macroscopic biological structures, although recent research has shown that California blackworm collectives are entangled, and can vary their degree of entanglement to solidify and melt their groups in response to environmental change[41]. Macroscopic multicellularity has evolved repeatedly in fungi[42] and, although previous research is lacking formally examining whether the cells of fungal fruiting bodies and lichen thalli are physically entangled, they are generally composed of densely packed overlapping hyphae, strongly suggesting entanglement[43–45]. The prevalence of entanglement in superficially different systems is probably 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. Although further research will be required to test this hypothesis, the relative ease with which multicellular fungi form entangled structures may have facilitated the highly convergent evolution of macroscopic multicellularity within this clade[42], enabling different fungal lineages to independently evolve robust multicellular structures.

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. However, these features 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[46]. Although 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, can form topologically structured multicellular groups with permanent cell–cell bonds[47], suggesting that early animals may have possessed a similar mode of growth.

Despite 600 rounds of selection for increased size, our mixotrophic and obligately aerobic lineages remained microscopic (Fig. 1b and Extended Data Fig. 1), increasing their radius by less than twofold. Extending previous research examining the role of oxygen diffusion in the evolution multicellular size[5], these results highlight the importance of environmentally dependent trade-offs on the evolution of multicellularity. Oxygen can serve as a resource, enabling increased cellular growth by increasing ATP yields from metabolism[48] and enables growth on non-fermentable carbon sources[49]. For simple, diffusion-limited organisms such as snowflake yeast, low concentrations of oxygen create a cost to large size by reducing the proportion of cells in the group that have access to it—a cost that our anaerobic populations did not face (for details, see ref. 5).

During the evolutionary transition to multicellularity, groups of cells must become Darwinian entities capable of adaptation[3]. This requires that they reproduce, and have heritable variation in traits that affect fitness[50]. For groups of cells to become more than simply the sum of their parts, adaptation must take place in multicellular traits that are distinct from those of their constituent cells (that is, they must evolve under MLS2)[36]. Using long-term experimental evolution, we show that even simple groups of cells, initially differing from their unicellular ancestor by a single mutation, have an innate capacity for sustained multicellular evolution. In response to selection on group size—a broadly important trait for simple multicellular organisms[7]—snowflake yeast evolved to form radically larger and tougher multicellular groups by leveraging the emergent biophysical properties of altered cellular morphology. These results demonstrate how selection on group size can drive multicellular adaptation and biophysical innovation, and highlight the surprising ease with which evolutionary transitions in Darwinian individuality can occur.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-023-06052-1.

1. Libby, E. & Rainey, P. B. A conceptual framework for the evolutionary origins of multicellularity. *Phys. Biol.* **10**, 035001 (2013).
2. Michod, R. E. in *The Evolution of Multicellularity* (eds Herron M. D., Conlin P. L. & Ratcliff W. C.) 25–52 (CRC Press, 2022).
3. Buss, L. W. *The Evolution of Individuality* Vol. 796 (Princeton Univ. Press, 2014).
4. Knoll, A. H. The multiple origins of complex multicellularity. *Ann. Rev. Earth Planet. Sci.* **39**, 217–239 (2011).
5. Bozdag, G. O., Libby, E., Pineau, R., Reinhard, C. T. & Ratcliff, W. C. Oxygen suppression of macroscopic multicellularity. *Nat. Commun.* **12**, 2838 (2021).
6. Smukalla, S. et al. *FLO1* is a variable green beard gene that drives biofilm-like cooperation in budding yeast. *Cell* **135**, 726–737 (2008).
7. Tong, K., Bozdag, G. O. & Ratcliff, W. C. Selective drivers of simple multicellularity. *Curr. Opin. Microbiol.* **67**, 102141 (2022).
8. Bonner, J. T. Perspective: the size-complexity rule. *Evolution* **58**, 1883–1890 (2004).
9. Willensdorfer, M. Organism size promotes the evolution of specialized cells in multicellular digital organisms. *J. Evol. Biol.* **21**, 104–110 (2008).
10. Knoll, A. H. & Hewitt, D. in *The Major Transitions in Evolution Revisited* (eds Calcott B. & Sterelny K.) 251–270 (MIT Press, 2011).
11. Bonner, J. T. *Why Size Matters: From Bacteria to Blue Whales* (Princeton Univ. Press, 2011).
12. Boudaoud, A. An introduction to the mechanics of morphogenesis for plant biologists. *Trends Plant Sci.* **15**, 353–360 (2010).
13. Jacobeen, S. et al. Cellular packing, mechanical stress and the evolution of multicellularity. *Nat. Phys.* **14**, 286–290 (2018).
14. Boraas, M. E., Seale, D. B. & Boxhorn, J. E. Phagotrophy by a flagellate selects for colonial prey: a possible origin of multicellularity. *Evol. Ecol.* **12**, 153–164 (1998).
15. Koschwanez, J. H., Foster, K. R. & Murray, A. W. Sucrose utilization in budding yeast as a model for the origin of undifferentiated multicellularity. *PLoS Biol.* **9**, e1001122 (2011).
16. Herron, M. D. et al. De novo origins of multicellularity in response to predation. *Sci. Rep.* **9**, 2328 (2019).
17. Westbrook, J. W. 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. *Am. Nat.* **177**, 800–811 (2011).
18. Prakash, V. N., Bull, M. S. & Prakash, M. Motility-induced fracture reveals a ductile-to-brittle crossover in a simple animal's epithelia. *Nat. Phys.* **17**, 504–511 (2021).
19. Ratcliff, W. C., Denison, R. F., Borrello, M. & Travisano, M. Experimental evolution of multicellularity. *Proc. Natl Acad. Sci. USA* **109**, 1595–1600 (2012).
20. 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. *Am. Nat.* **138**, 1315–1341 (1991).
21. Ratcliff, W. C., Fankhauser, J. D., Rogers, D. W., Greig, D. & Travisano, M. Origins of multicellular evolvability in snowflake yeast. *Nat. Commun.* **6**, 6102 (2015).
22. Heinrich, E. C., Farzin, M., Klok, C. J. & Harrison, J. F. The effect of developmental stage on the sensitivity of cell and body size to hypoxia in *Drosophila melanogaster*. *J. Exp. Biol.* **214**, 1419–1427 (2011).
23. Denk, W. & Horstmann, H. Serial block-face scanning electron microscopy to reconstruct three-dimensional tissue nanostructure. *PLoS Biol.* **2**, e329 (2004).
24. Edwards, C. E., Mai, D. J., Tang, S. & Olsen, B. D. Molecular anisotropy and rearrangement as mechanisms of toughness and extensibility in entangled physical gels. *Phys. Rev. Mater.* **4**, 015602 (2020).
25. Brown, E., Nasto, A., Athanassiadis, A. G. & Jaeger, H. M. Strain stiffening in random packings of entangled granular chains. *Phys. Rev. Lett.* **108**, 108302 (2012).
26. Raymer, D. M. & Smith, D. E. Spontaneous knotting of an agitated string. *Proc. Natl Acad. Sci. USA* **104**, 16432–16437 (2007).
27. Wilhelm, J. & Frey, E. Elasticity of stiff polymer networks. *Phys. Rev. Lett.* **91**, 108103 (2003).
28. Kim, J., Zhang, G., Shi, M. & Suo, Z. Fracture, fatigue, and friction of polymers in which entanglements greatly outnumber cross-links. *Science* **374**, 212–216 (2021).
29. Tauber, J., Rovigatti, L., Dussi, S. & Van Der Gucht, J. Sharing the load: stress redistribution governs fracture of polymer double networks. *Macromolecules* **54**, 8563–8574 (2021).
30. Sheu, Y.-J., Barral, Y. & Snyder, M. Polarized growth controls cell shape and bipolar bud site selection in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **20**, 5235–5247 (2000).
31. 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. *Curr. Genet.* **55**, 365–380 (2009).
32. Sopko, R. et al. Mapping pathways and phenotypes by systematic gene overexpression. *Mol. Cell* **21**, 319–330 (2006).
33. Lee, D. H., Tamura, A., Arisaka, Y., Seo, J.-H. & Yui, N. Mechanically reinforced gelatin hydrogels by introducing slidable supramolecular cross-linkers. *Polymers* **11**, 1787 (2019).
34. Gerhards, C. *Effects of Type of Testing Equipment and Specimen Size on Toughness of Wood* Vol. 97 (Forest Products Laboratory, 1968).
35. Jacobeen, S. et al. Geometry, packing, and evolutionary paths to increased multicellular size. *Phys. Rev. E* **97**, 050401 (2018).
36. Okasha, S. Multilevel selection and the major transitions in evolution. *Phil. Sci.* **72**, 1013–1025 (2005).
37. Graessley, W. W. *The Entanglement Concept in Polymer Rheology* (Springer, 1974).
38. Chen, W., Yu, H., Li, Q., Liu, Y. & Li, J. Ultralight and highly flexible aerogels with long cellulose I nanofibers. *Soft Matter* **7**, 10360–10368 (2011).
39. Gravish, N., Franklin, S. V., Hu, D. L. & Goldman, D. I. Entangled granular media. *Phys. Rev. Lett.* **108**, 208001 (2012).
40. Zou, L.-N., Cheng, X., Rivers, M. L., Jaeger, H. M. & Nagel, S. R. The packing of granular polymer chains. *Science* **326**, 408–410 (2009).
41. Ozkan-Aydin, Y., Goldman, D. I. & Bhamla, M. S. Collective dynamics in entangled worm and robot blobs. *Proc. Natl Acad. Sci. USA* **118**, e2010542118 (2021).
42. Nagy, L. G. in *The Evolution of Multicellularity* 279–300 (CRC Press, 2022).
43. Moore, D. *Fungal Biology in the Origin and Emergence of Life* (Cambridge Univ. Press, 2013).
44. García-Segovia, P., Andrés-Bello, A. & Martínez-Monzó, J. Rehydration of air-dried Shiitake mushroom (*Lentinus edodes*) caps: comparison of conventional and vacuum water immersion processes. *LWT Food Sci. Technol.* **44**, 480–488 (2011).
45. Roth, R., Wagner, R. & Goodenough, U. Lichen 3. Outer layers. *Algal Res.* **56**, 102332 (2021).
46. Yanni, D. et al. Topological constraints in early multicellularity favor reproductive division of labor. *eLife* **9**, e54348 (2020).
47. Larson, B. T. et al. Biophysical principles of choanoflagellate self-organization. *Proc. Natl Acad. Sci. USA* **117**, 1303–1311 (2020).
48. Hinkle, P. C., Kumar, M. A., Resetar, A. & Harris, D. L. Mechanistic stoichiometry of mitochondrial oxidative phosphorylation. *Biochemistry* **30**, 3576–3582 (1991).
49. Raymond, J. & Segrè, D. The effect of oxygen on biochemical networks and the evolution of complex life. *Science* **311**, 1764–1767 (2006).
50. Lewontin, R. C. The units of selection. *Annu. Rev. Ecol. Syst.* **1**, 1–18 (1970).

Methods

Long-term evolution experiment

To generate our ancestral snowflake yeast for the MuLTEE, we started with a unicellular diploid yeast strain (Y55). In these yeast, we replaced both copies of the *ACE2* transcription factor[21] using a *KANMX* resistance marker (*ace2::KANMX/ace2::KANMX*) and obtained a snowflake yeast clone (see ref. 5 for a detailed description of strains and growth conditions, including measurements of oxygen concentrations in growth medium). When grown in YEPD medium (1% yeast extract, 2% peptone, 2% dextrose), these yeast are mixotrophic, both fermenting and respiring. When grown in YEPG medium, which is the same as YEPD medium 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. Owing to a large deletion in its mitochondrial DNA (identified by sequencing), this snowflake yeast is unable to respire and is therefore metabolically anaerobic, and was cultured in YEPD medium.

We evolved five replicate populations of mixotrophic (referred to as populations PM1–PM5), obligately aerobic (PO1–PO5) and anaerobic (PA1–PA5) snowflake yeast in 10 ml of culture medium, growing them in 25 × 150 mm culture tubes for 24 h at 30°C with 225 rpm shaking. We used settling selection to select for larger cluster size. Once per day, after around 24 h of growth, we transferred 1.5 ml of culture into 1.5 ml Eppendorf tubes, let them settle on the bench for 3 min, 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 medium 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 s, therefore keeping them under directional selection for increased size. The timing of this change corresponded to around 350 days for PA2 and PA5 and about 500 days for PA1, PA3 and PA4. We used wide-bore filtered pipette tips (Thermo Fisher 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 at −80 °C every 10–20 transfers.

Measuring cluster size

We developed a standard visualization protocol to be able to measure the size of both microscopic and macroscopic snowflake yeast from each population over the 600-day evolution experiment. To prepare yeast for imaging, we revived evolved frozen cultures for each population at 50-day intervals (12 for each of the 15 replicate populations). We next inoculated each sample into 10 ml fresh medium and brought them to equilibrium over a 5-day culture process, performing daily settling selection before transfer into fresh medium. After five transfers, we pipetted a random 1 ml subsample of each 24 h 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 ×4 Nikon objective, capturing the cross-sectional area of clusters without disrupting their 3D structure. Next, we used ImageJ to calculate the cross-sectional area of each cluster, converting pixels to micrometres by including a physical 100 µm scale bar in each image.

Calculating the weighted average cluster size

The distribution of cluster size across various isolates is not consistent—microscopic populations are unimodal, whereas macroscopic populations contain a substantial 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 effect on sample statistics, such as average cluster size. Owing to their skewed size distribution, 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 Extended Data Fig. 3a 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 (micrometre) radius to be consistent with the units that have historically been used in the palaeontological literature documenting the evolution of macroscopic multicellular organisms. Python code and raw data used in calculating weighted biomass mean averages are available in our public data repository (https://github.com/ozanbozdag/de_novo_evolution_of_macroscopic_multicellularity).

Assessing fitness

We measured the relative fitness of the evolved macroscopic populations (PA1–5) in competition against the ancestor in liquid culture under the same conditions as in our evolution experiment. To differentiate competing strains, we used an ancestral snowflake yeast strain carrying a hemizygous red fluorescent protein (*ura3::dTOMATO/URA3*). Before coculturing these strains, we first grew evolved populations and the ancestral strain in separate cultures overnight. We next mixed the two types in a 1.5 ml microcentrifuge tube to start the competition assay in fresh 10 ml YEPD cultures. We examined the fitness of PA1–5 t600, as well as an ancestor:ancestor control, with three replicate competitions per treatment. We grew these competition cultures for 24 h in 10 ml YEPD (conditions as described in the evolution experiment), followed by 3 min of settling selection in 1.5 ml Eppendorf tubes. We then transferred the bottom 50 µl into a fresh culture tube for the next round of growth and settling, repeating the same procedure for three rounds across the fitness assay. The initial frequency of the evolved populations ranged between 35% and 70% and, after three rounds of growth and settling selection, they reached a range between 99.8% and 99.9%.

To calculate the relative fitness of the evolved populations against the common ancestor, we calculated their selection rate constant, as described previously[20]. To do so, we estimated the initial and final cellular density of yeast by measuring the cross-sectional area of the evolved and ancestral snowflake yeast clusters at the beginning and end of the fitness assay using a Nikon Eclipse Ti inverted microscope at ×100 magnification. We next calculated the selection rate by dividing the estimated density of the evolved populations at the end and beginning of the competition assay, followed by subtracting the natural log of this value from that of the ancestral strain[20]. Finally, we confirmed that the expression of the hemizygous dTOMATO construct did not impose a significant fitness cost on the ancestral snowflake yeast strain through coculturing yeast with and without the constructs (that is, ancestor:ancestor control). The results are reported in Fig. 1f, in which the right column labelled as control shows no significant difference ($P = 0.22$, $t = 1.7$, d.f. = 2, one sample t-test). Image analysis was performed using ImageJ (v.2.3.0).

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 × 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 with settling selection. On the final day, we transferred 100 µl of each culture into tubes with fresh YEPD and incubated them for 12 h. We next stained samples in calcofluor-white by incubating them in the dark for 30 min (at a final concentration of 5 µM) before imaging (×40 objective, ultraviolet light excitation of a blue fluorescent cell wall stain, imaged on the Nikon Ti-E system). We measured the aspect ratio of individual cells within snowflake yeast clusters using ImageJ, analysing an average of 453 cells per population.

Simple biophysical model examining packing fraction as a function of aspect ratio

We simulated the growth of snowflake yeast from a single cell. Cells were modelled as prolate ellipsoids with one long (major) axis and two equal shorter (minor) axes. Clusters started as a single cell and were grown for nine cellular generations. New cells first emerged from their mother's distal pole; subsequent cells emerged with a polar angle of 45° and a random azimuthal angle. If adding a new cell would cause too much overlap with existing branches, the new cell was deleted and the mother cell lost its chance to reproduce that generation. We simulated the growth of 50 clusters of each genotype, which varied in their cellular aspect ratio, defined as the ratio of the major axis to minor axis length, ranging from 1.2 to 2.8. We then calculated each simulated cluster's packing fraction by fitting a convex hull to the cluster and measuring the ratio of the total volume to the volume specifically occupied by cells. The MATLAB code to grow snowflake yeast using this protocol is provided in Supplementary File 1.

Testing aggregative versus clonal development

To determine whether macroscopic snowflake yeast aggregate or develop clonally (Extended Data Fig. 6), we isolated a single genotype from PA2, t600 (strain GOB1413-t600), and engineered it to constitutively express either green or red fluorescent proteins. To do that, we amplified the *prTEF_GFP_NATMX* construct from the pFA6a-eGFP plasmid and the *prTEF_dTOMATO_NATMX* construct from the pFA6a-tdTomato plasmid. We then separately replaced the *URA3* open reading frame with *GFP* or *dTOMATO* constructs in an isogenic single-strain isolate according to the LiAc transformation protocol[51]. We selected transformants on nourseothricin sulfate (Gold Biotechnology) YEPD plates and confirmed green or red fluorescent protein activity of transformed macroscopic clusters by visualizing them under the 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 5 days, transferring 1% of the population to fresh medium 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, enabling us to count the number of snowflake yeast clusters that were green, red or chimeric.

We examined the potential for entanglement alone to allow for persistent interactions among disconnected components (Extended Data Fig. 9) by crushing GFP- and RFP-tagged macroscopic snowflake yeast (PA2 t600) into smaller groups, and then growing a mixture of them on the surface of agar plates for 48 h, potentially allowing branches of adjacent genotypes to entangle through growth. We then scraped these populations and grew them in 10 ml YEPD (yeast-extract, peptone, dextrose) medium with shaking at 250 rpm for two 24 h rounds of growth and settling selection. We imaged the resulting clusters under wide-field microscopy (Nikon Ti-E), taking pictures of individual clusters under the bright-field, green and red channels, collecting data for a random sample of clusters (101 for the PA2 t600 and 110 for the ancestor control). We analysed the images in ImageJ, and any clusters that contained both green and red cells were scored as chimeric. The images shown in Extended Data Fig. 9 were taken on the Nikon AR1 confocal microscope, enabling a better view of the 3D structure of chimeric intercalation.

Sample preparation for SBF-SEM

For specimen preparation for SBF-SEM (Fig. 3a,b and Extended Data Fig. 7), we fixed snowflake yeast in 2% formaldehyde (fresh from paraformaldehyde (EMS)) containing 2 mM $CaCl_2$, incubating at 35 °C for 5 min followed by 2–3 h on ice. We next incubated these yeast samples for 1 h in a solution of 1.5% potassium ferrocyanide, 0.15 M cacodylate buffer, 2 mM $CaCl_2$ and 2% aqueous osmium tetroxide. This last step was performed on ice and under a vacuum. Finally, we washed our yeast and incubated them in thiocarbohydrazide solution (10 g l^{-1} double-distilled water) for 60 min at 60 °C, followed by en bloc uranyl acetate and lead aspartate staining[52,53].

SBF-SEM

We imaged fixed yeast on the Zeiss Sigma VP 3View system. This system has a 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, a beam dwell time of 0.5–1 µs 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 three parts: live cells, dead cell debris and background. We then imported segmented HDF5 files in Python. First, we identified connected cells using the nearest-neighbour algorithm to identify connected cells. We call a set of connected cells inside a subvolume 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 subvolumes in Rhino6. We then manually identified cell-to-cell connections and coloured 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 according to a previously described protocol[54].

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 µm spherical glass probe). For mechanical measurements of macroscopic snowflake yeast, we used a Zwick Roell Universal Testing Machine (UTM) with a 5 N probe. As above, individual clusters were extracted from the growth tube and placed onto the testing stage while submerged under water. Two different machines were used for these two different mechanical measurements due to the orders-of-magnitude change in size and toughness achieved by the evolved snowflake yeast. But, while the two different instruments have different measurement precisions, the mechanical measurements between the two remain comparable. This is because stress–strain relationships reflect material properties, like Young's modulus, that are invariant to the size of the sample. Thus, given the same material, the two machines would return the

same stress–strain relationship up until the nanoindenter's maximum stress limit is reached.

Preparing glass slides with attached cells

We coated glass slides with concanavalin A to make a sticky glass surface to which individual cells could adhere. We started by preparing a 10 mg ml^{-1} solution of concanavalin A dissolved in sterile deionized water, which can be stored at −20 °C. This stock solution was diluted 1:10, and then 200 µl of diluted solution was pipetted onto a glass slide in a sterile environment. The slide was allowed to incubate for 5 min at room temperature, then washed twice with sterile deionized water and left to aspirate dry in the hood. Cell cultures were inoculated (100 µl) onto the glass surface and left to settle for 5 min.

AFM measurements

Before measuring the properties of individual cells using the atomic force microscope, we restored ACE2 functionality to increase the frequency of single cells available for mechanical testing. To do this, we reinserted a single copy of the ancestral *ACE2* allele fused with the antibiotic-resistance gene *HYGNT1* into the genome of the PA ancestor and PA1 t600 isolate under the control of its native promotor using the LiAc/SS-DNA/PEG method of yeast transformation[51]. Transformants were then plated on YEPD agar plates (1% yeast extract, 2% peptone, 2% dextrose, 1.5% agar) supplemented with 200 mg l^{-1} of the antibiotic hygromycin B (Enzo Life Sciences). All atomic force measurements used an atomic force microscope from Asylum Research that was integrated with an inverted optical microscope (Nikon). For single-cell measurements, we used a silicon nitride cantilever with a nominal stiffness of 0.06 N m^{-1} with an attached borosilicate glass bead with diameter 2 µm (Novascan Technologies). The cantilever was measured by thermal analysis to have a stiffness of 0.0593 N m^{-1}. For cluster-level measurements, we used tipless, aluminium-coated cantilevers with a rectangular shape (length 225 µm, width 40 µm) that have a nominal stiffness of 30 N m^{-1} (AppNano). For measurements, either single cells or entire clusters were visually aligned with the cantilever probe, which was then moved at a velocity of 1 µm s^{-1} to compress the cell or cluster with increasing force.

Chitin staining protocol

We stained cells with calcofluor white according to the following protocol. First, we mixed 500 µl of cell culture from the ancestor and PA2 t600 strains into the same tube. We then sampled 150 µl (containing both the ancestor and t600 yeast clusters) from the mixed culture. We removed the supernatant through an iterated process of centrifugation and pipetting medium removal. We then diluted 15 µl of 1 mg ml^{-1} calcofluor solution into 500 µl 1× phosphate-buffered saline solution (PBS) and mixed with the yeast pellet. We incubated the sample in the dark at room temperature for 25 min, then we removed the calcofluor media by centrifugation and pipetting. Finally, we added 200 µl 1× PBS on top of the pellet. A total of 20 µl of this cell suspension was pipetted onto a clean glass slide and covered with a coverslip for microscopy analysis.

Single-cell and bud scar confocal microscopy

We used the Nikon A1R confocal microscope equipped with a ×60 oil-immersion objective to obtain *z*-stack images of individual cells stained with calcofluor white. To track the location and size of bud scars, we wrote a MATLAB script to extract the brightest calcofluor signals, as the chitinous bud site region makes bud scars brighter than the other portions of the cell wall. Brightness isosurfaces isolated the bud scars themselves, and the brightness of each voxel contained within the isosurface was recorded to track the density of chitin. Next, the isosurface points were rotated to the *x*–*y* plane by finding the principal axes of the shape through principal component analysis. The rotated surface points were then used to calculate the height and cross-sectional area of the bud scars.

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 (that is, PA1–PA5) and their common ancestor after 200, 400, and 600 days of evolution. To pick clonal isolates, we diluted populations of snowflake yeast clusters in 1.5 ml tubes and plated them at a density of 100–200 colonies per plate. We next restreaked those initial single colonies onto fresh plates, therefore ensuring that each colony on a plate results from a single snowflake yeast cluster. As snowflake yeast grow clonally, we expected that these isolates would represent only a single clone of cells, with no more variation than would be expected from any single cell isolate that grew into a population, generating de novo mutation along the way (subsequent analysis of the genomes confirmed this—we never saw evidence of >1 genotype present in any isolate). We inoculated these 16 samples in YEPD for 12 h and extracted their genomic DNA using a commercially available kit (Amresco). We measured the concentration of DNA with a Qubit fluorometer (Thermo Fisher Scientific). We prepared genomic DNA libraries for the 16 samples using NEBNext Ultra DNA Library Prep Kit for Illumina (New England Biolabs). 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). Finally, whole genomes were sequenced using the HiSeq 2500 platform (Illumina) by the Genome Analysis Core Center located in the Petit Institute, Georgia Tech. As a result, we obtained paired 150 bp (R1 and R2) FASTQ reads from two lanes (L1 and L2) (raw reads are publicly available).

Bioinformatics analysis

For our bioinformatic analysis, we used the bash command-line interface on a Linux platform. To identify de novo mutations (SNPs and indels) in the ancestral and evolved genomes, we first filtered out low-quality reads using a sliding-window approach in Trimmomatic (v.0.39). We aligned reads to the yeast reference genome (S288C, SGD) using an algorithm in the BWA software package (BWA-MEM)[55]. Next, we used the genome analysis toolkit (GATK) to obtain and manipulate .bam files[56]. Duplicate reads were marked using the Picard tool (MarkDuplicates v.2.18.3). We called SNPs using two different tools—GATK4 HaplotypeCaller (v.4.0.3.0) and FreeBayes (v.1.2.0)[56,57]. 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 (v.0.1.16)[58]. Finally, after manually checking each variant call by visualizing SAM files and VCF files on Integrative Genomics Viewer (IGV)[59], 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 (v.1.10)[60]. Finally, we annotated evolved mutations using SnpEff (v.4.3T)[61] (VCF files are available in our public data repository).

To search for GO term enrichment for de novo mutations presented in Fig. 4d (https://github.com/ozanbozdag/de_novo_evolution_of_macroscopic_multicellularity), we generated a combined list of synonymous and nonsynonymous mutations within gene coding regions. We then searched for enriched GO terms using GO Term Finder and GO Slim Mapper on the yeast genome database[62].

Genetically engineering macroscopic snowflake yeast

To genetically engineer snowflake yeast strains with cell lengthening (*CLB2* and *CLN3*) and bud-scar strengthening (*GIN4*) mutations that are shown in Fig. 4, we used homozygous unicellular (GOB76) and multicellular strains (GOB8). To engineer each gene deletion, we first amplified hygromycin-, nourseothricin- and G418-resistance marker cassettes from plasmids and individually transformed them into yeast cells (see Supplementary Table 1 for the genotypes, plasmids and primers). We next induced sporulation of individual heterozygous mutant strains

(2% potassium acetate), dissected tetrads and obtained homozygous deletions through auto-diploidization. By subsequent rounds of sporulation and intertetrad mating on appropriate multidrug plates, we generated the double-, triple- and quadruple-mutant strains presented in Fig. 4e,f. Finally, we quantified the cellular aspect ratio and cluster size (cross-sectional area) of these mutants by imaging clusters under the Nikon Eclipse Ti inverted microscope, using the same methods as previously described.

Life-cycle experiment

To characterize the life cycle of the ancestral (microscopic) and evolved (macroscopic, PA2-t600) snowflake yeast, we inoculated both strains starting from frozen glycerol stocks (Extended Data Fig. 2). We then grew them under the conditions of the evolution experiment for five rounds of growth and settling selection. Applying a 5 day growth and settling selection brings cultures to an equilibrium, reflecting the physiology and size distribution observed during the evolution experiment. On the final day, we sampled from the growing cultures at 0, 3, 6, 12 and 24 h and measured the number, size and volume of cultures using the same methods described in the 'Measuring cluster size' section.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

Underlying data used to generate figures and raw data are available at GitHub (https://github.com/ozanbozdag/de_novo_evolution_of_macroscopic_multicellularity). Raw Illumina sequencing reads are available at the NIH Sequence Read Archive under accession number PRJNA943273. All microscopy images used to generate data are archived in the Ratcliff laboratories Dropbox and are available on request.

Code availability

Codes used in this study are available at GitHub (https://github.com/ozanbozdag/de_novo_evolution_of_macroscopic_multicellularity). Code for the simple 3D biophysical simulation are provided in Supplementary File 1; these simulations were adapted from ref. 35 (further information can be requested from T.C.D. or P.J.Y.).

51. Gietz, R. D. & Schiestl, R. H. High-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method. *Nat. Protoc.* **2**, 31–34 (2007).
52. Deerinck, T. J. et al. High-performance serial block-face SEM of nonconductive biological samples enabled by focal gas injection-based charge compensation. *J. Microsc.* **270**, 142–149 (2018).
53. Ngo, H. T. & Yin, C. S. *Luteimonas terrae* sp. nov., isolated from rhizosphere soil of Radix ophiopogonis. *Int. J. Syst. Evol. Microbiol.* **66**, 1920–1925 (2016).
54. Zamani-Dahaj, S. A. et al. Spontaneous emergence of multicellular heritability. Preprint at *bioRxiv* https://doi.org/10.1101/2021.07.19.452990 (2021).
55. Li, H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. Preprint at https://arxiv.org/abs/1303.3997 (2013).
56. McKenna, A. et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **20**, 1297–1303 (2010).
57. Garrison, E. & Marth, G. Haplotype-based variant detection from short-read sequencing. Preprint at https://arxiv.org/abs/1207.3907 (2012).
58. Danecek, P. et al. The variant call format and VCFtools. *Bioinformatics* **27**, 2156–2158 (2011).
59. Thorvaldsdóttir, H., Robinson, J. T. & Mesirov, J. P. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Briefings Bioinform.* **14**, 178–192 (2013).
60. Danecek, P. et al. Twelve years of SAMtools and BCFtools. *Gigascience* **10**, giab008 (2021).
61. Cingolani, P. 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).
62. Cherry, J. M. et al. Saccharomyces Genome Database: the genomics resource of budding yeast. *Nucleic Acids Res.* **40**, D700–D705 (2012).

Acknowledgements We thank J. T. Pentz for teaching us Illumina library preparation; S. Biliya at the High Throughput DNA Sequencing Core at Georgia Tech for sequencing the genomes of evolved strains; K. A. Boateng at Core Facilities at the Carl R. Woese Institute for Genomic Biology for the SEM imaging; S. Cao for helping with microscopy during the early stages of this project; and C. Orlic, L. Nagy, and all of the members of the Ratcliff group for comments on the manuscript. This work was supported by NIH grants R35-GM138030 to W.C.R. and R35-GM138354 to P.J.Y., Human Frontiers in Science Grant RGY0080/2020 to W.C.R., and a Packard Fellowship for Science and Engineering to W.C.R.

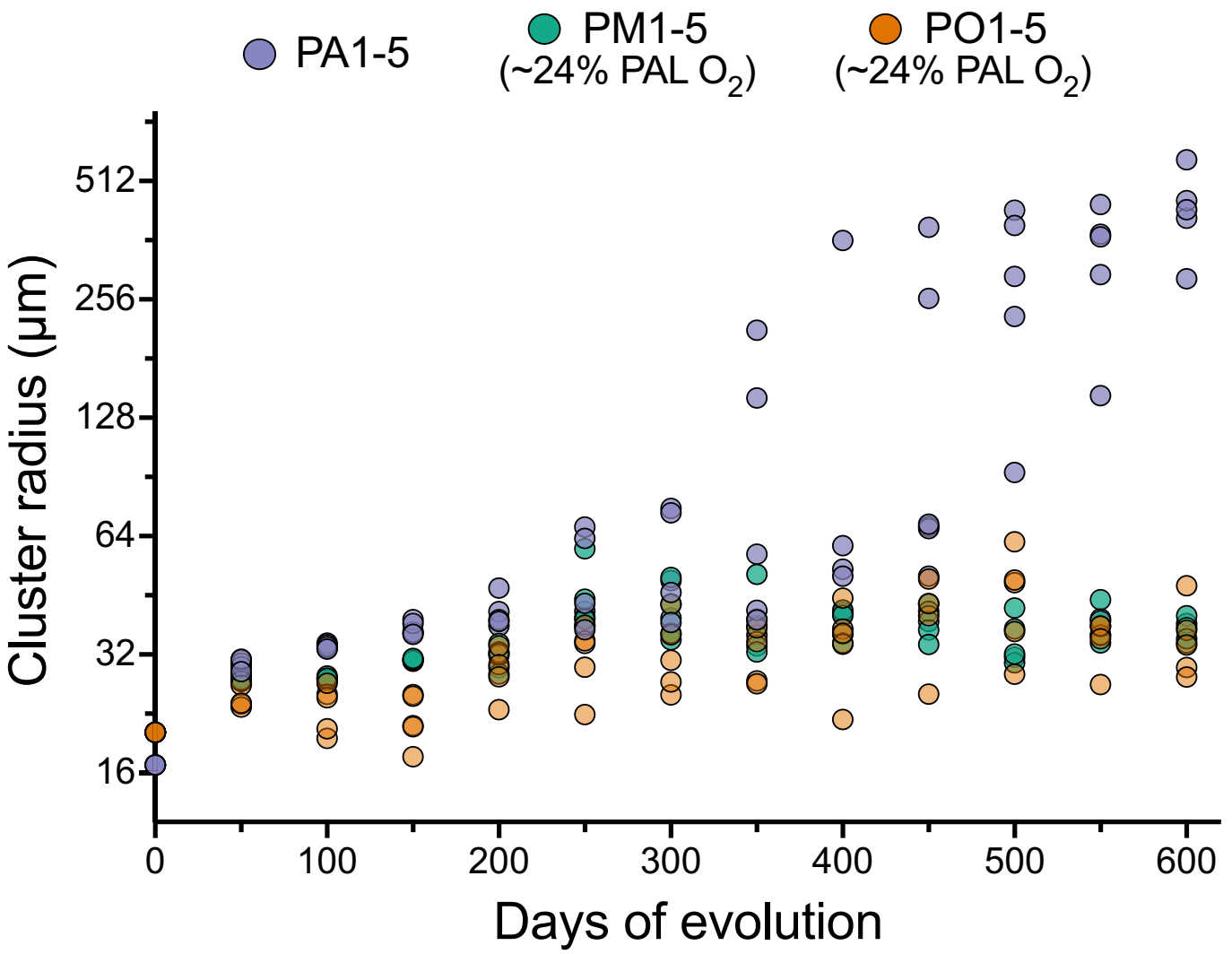
Author contributions G.O.B., S.A.Z.-D., P.J.Y. and W.C.R. conceived the project. G.O.B. and W.C.R. designed the MuLTEE. G.O.B. performed the evolution experiment. G.O.B., S.A.Z.-D., P.C.K. and T.C.D. designed and collected data. S.A.Z.-D. generated SBF-SEM images. S.A.Z.-D., T.C.D. and P.J.Y. performed the yeast biophysical simulations. E.L.D. and A.H.B. assisted G.O.B. and S.A.Z.-D. with image analysis. A.J.B. genetically engineered large snowflake yeast. K.T. performed life-cycle experiments. D.T.L. measured the number of generations. P.L.C. performed unicellular reversion experiments. G.O.B., S.A.Z.-D., T.C.D., W.C.R. and P.Y. analysed the data. G.O.B. made the figures. G.O.B., W.C.R. and P.J.Y. wrote the first draft of the paper, and all of the authors contributed to the revision.

Competing interests The authors declare no competing interests.

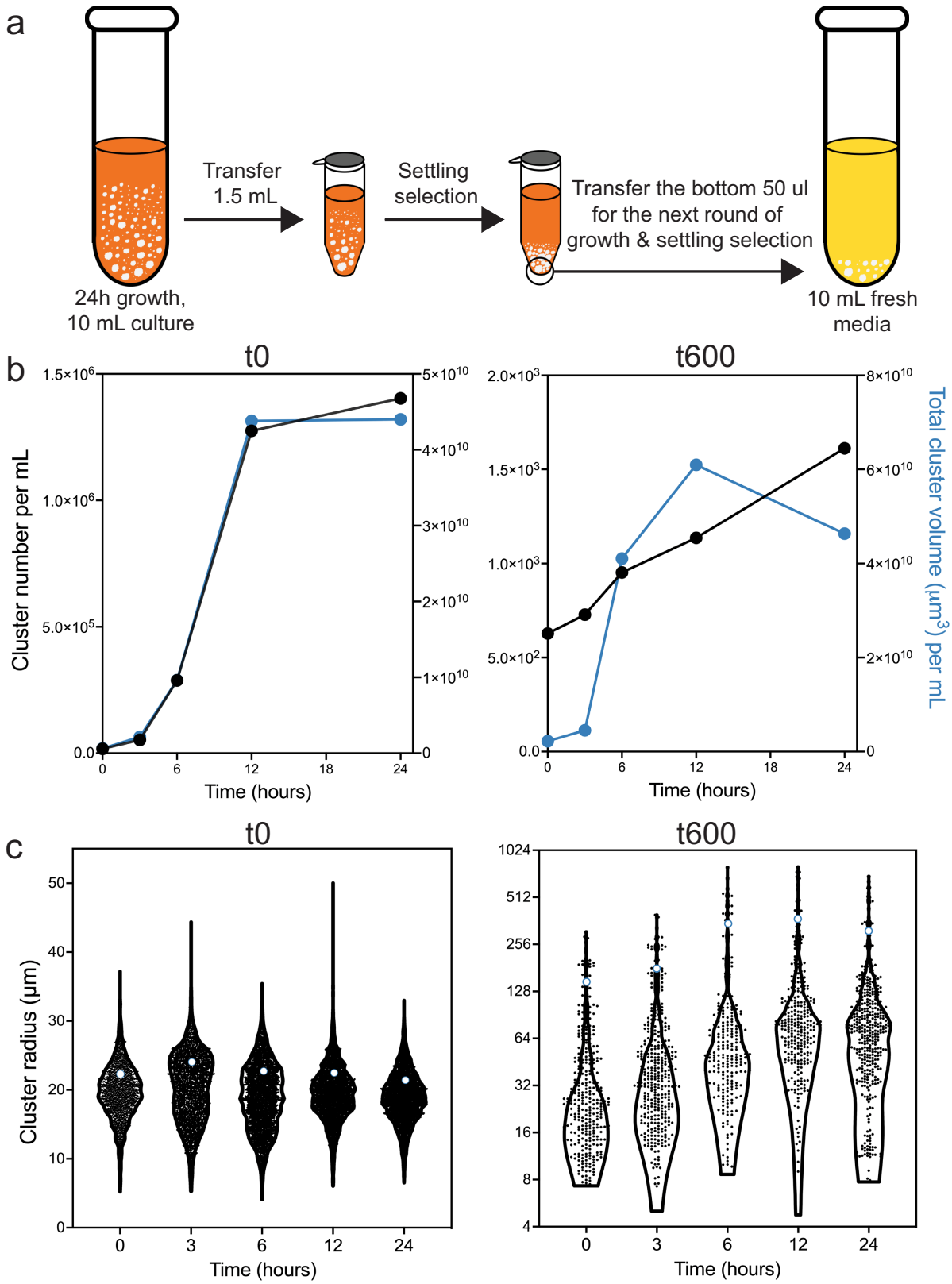
Additional information
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41586-023-06052-1.

Correspondence and requests for materials should be addressed to G. Ozan Bozdag, Peter J. Yunker or William C. Ratcliff.

Peer review information *Nature* thanks Shiladitya Banerjee, Omaya Dudin and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

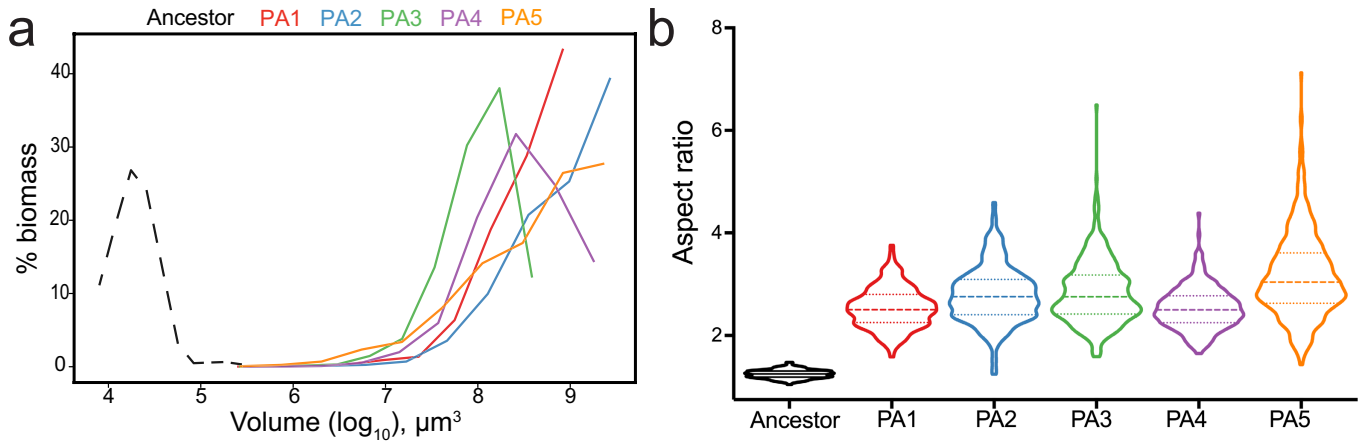


Extended Data Fig. 1 | Temporal dynamics of size evolution in each population and treatment group. Data points show the weighted average radius of cluster size for the entire population. This was calculated by measuring the size of an average of 1150 snowflake yeast clusters per sample population (3 ancestors + 3 treatment groups x 5 replicate populations x 12 time points = 183 samples, all data is publicly available under the raw data file). Please see the Methods section for details on how weighted average radius was calculated.

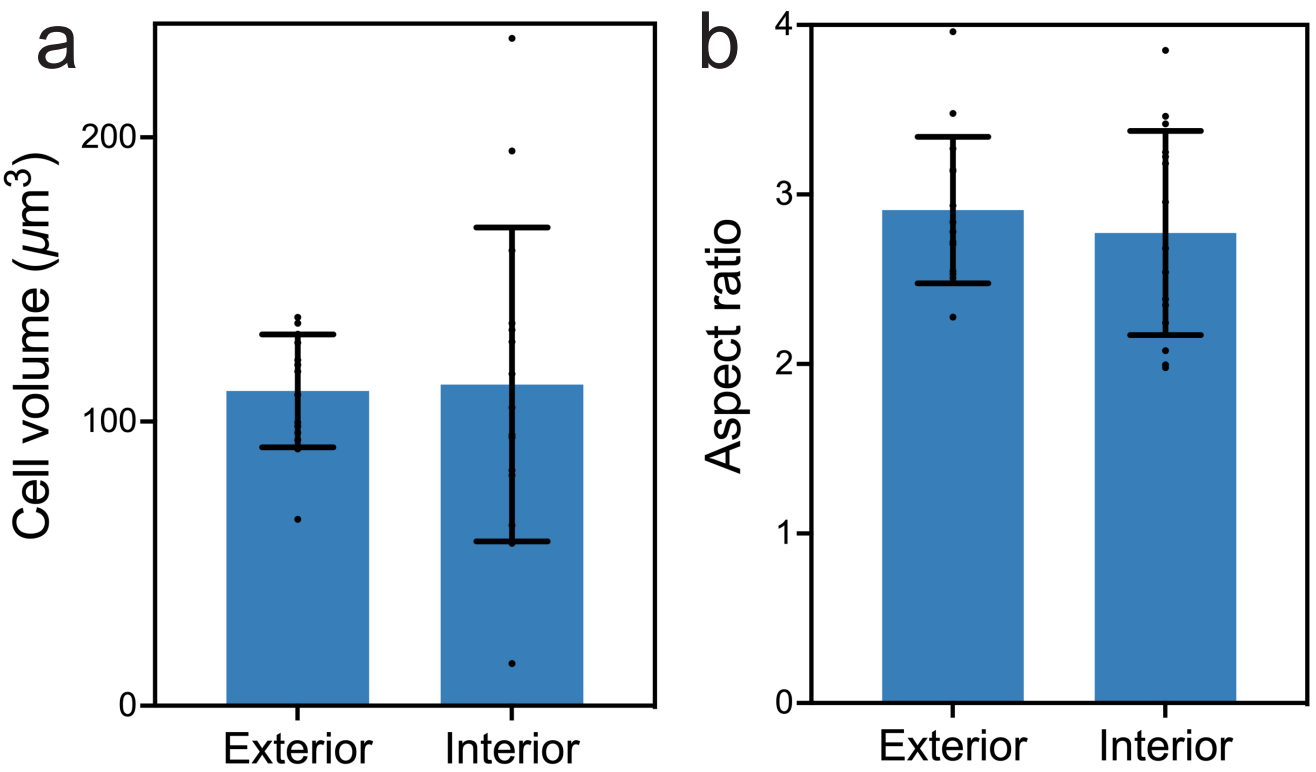


Extended Data Fig. 2 | See next page for caption.

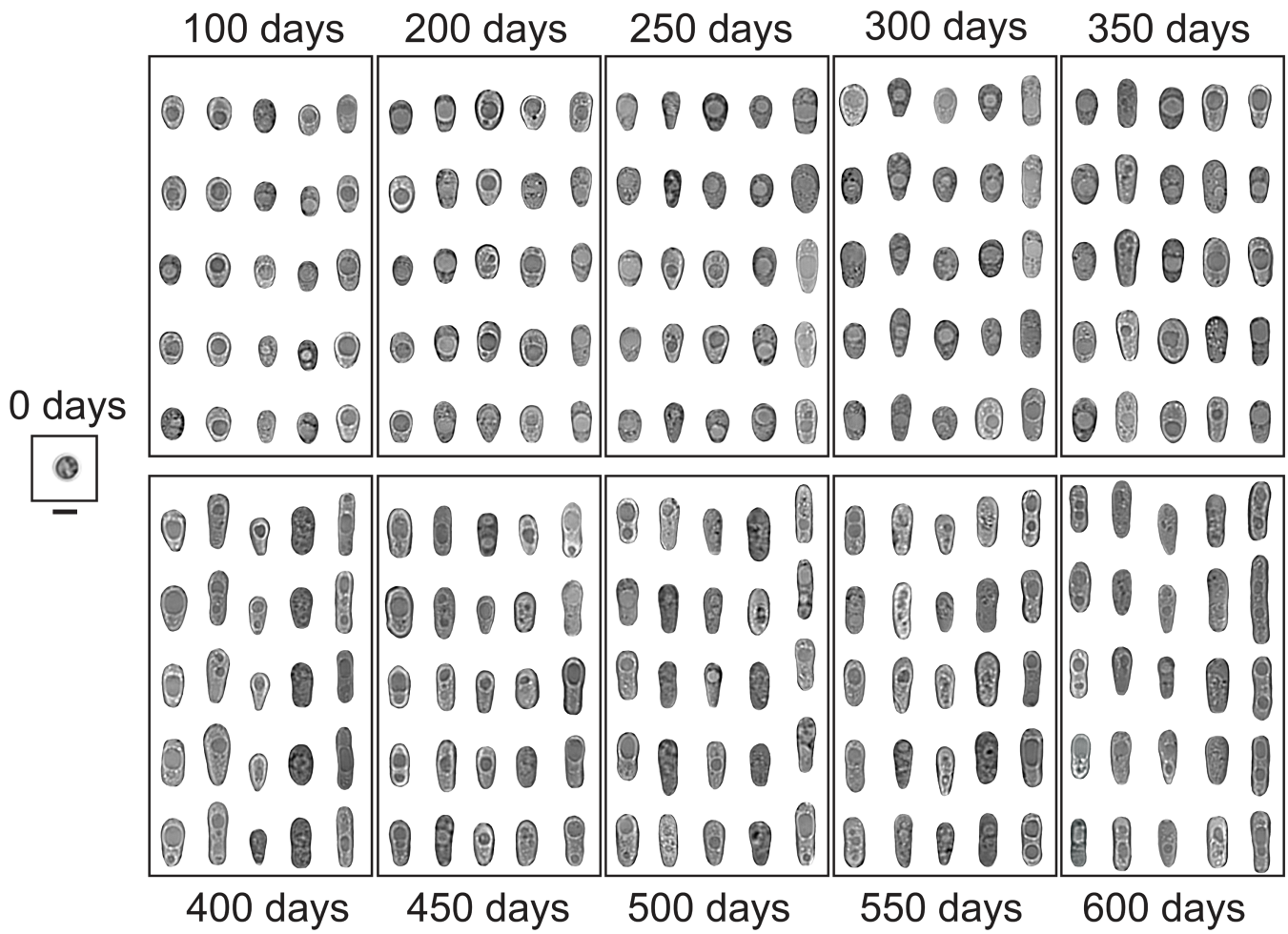
Extended Data Fig. 2 | Characterizing the life-cycle of the ancestral (microscopic) and evolved (macroscopic) snowflake yeast. a, During the ~24-hour growth cycle, snowflake yeast compete for growth and reproduction in 10 mL of YPED (250 RPM at 30 °C). At the end of the growth phase, we select for larger group size via settling selection. While there is a theoretical maximum survival rate of 15% (that is, if all of the cells survived settling selection), we only transfer the bottom 50 µl of pellet biomass regardless of how many cells settle, creating an arms race that favours the fastest groups within the population. Our measurements of the number of cellular generations per day in Fig. 1a suggests about 3% of the cells survive from one day to the next on average. **b**, Both the microscopic (ancestral) and macroscopic (t600) snowflake yeast clusters have a life cycle, reproducing during the growth phase. **c**, Consistent with entanglement producing tough groups, macroscopic snowflake yeast release mostly microscopic propagules, possibly from branch tips at the exterior of the group, where the opportunity for entanglement is minimal. Despite the presence of many small propagules, most of the biomass in the population is contained within macroscopic clusters. The open circles represent the biomass-weighted mean size, which is the average sized group the mean cell finds itself in. A total of 14,313 clusters were analysed for the t0 time point, and 1,603 clusters were analysed for the t600 time point, across 0, 3, 6, 12, and 24-hour time points.



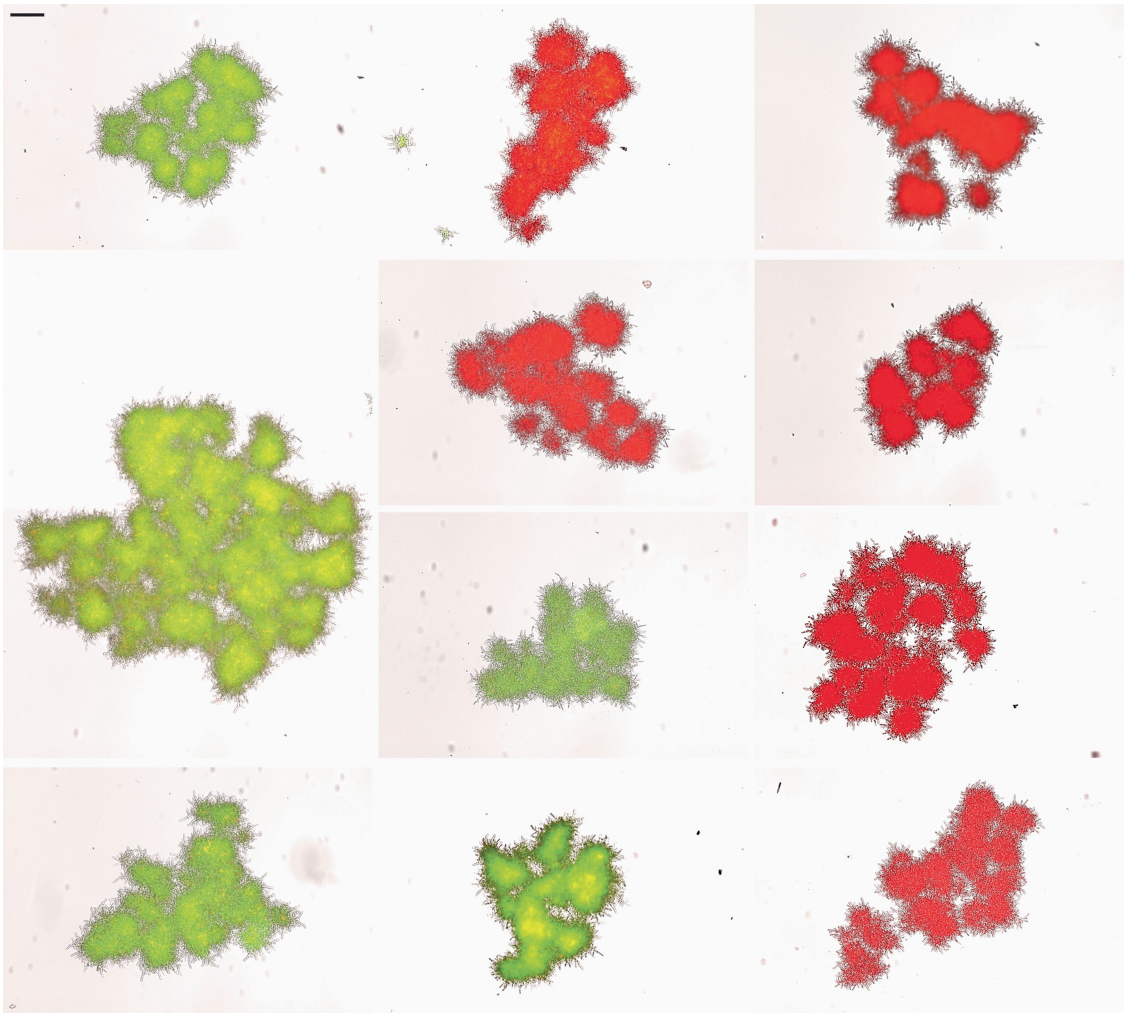
Extended Data Fig. 3 | Cluster size and aspect ratio distribution. a, Biomass distribution as a function of cluster size for the ancestral snowflake yeast (dotted line) and 600 day evolved populations of PA1-PA5. The 'weighted mean size' used in Figs. 1, 2 and 4 is the mean of the biomass distribution. **b**, Distribution of aspect ratios for ancestral and 600-day evolved populations of anaerobic snowflake yeast.



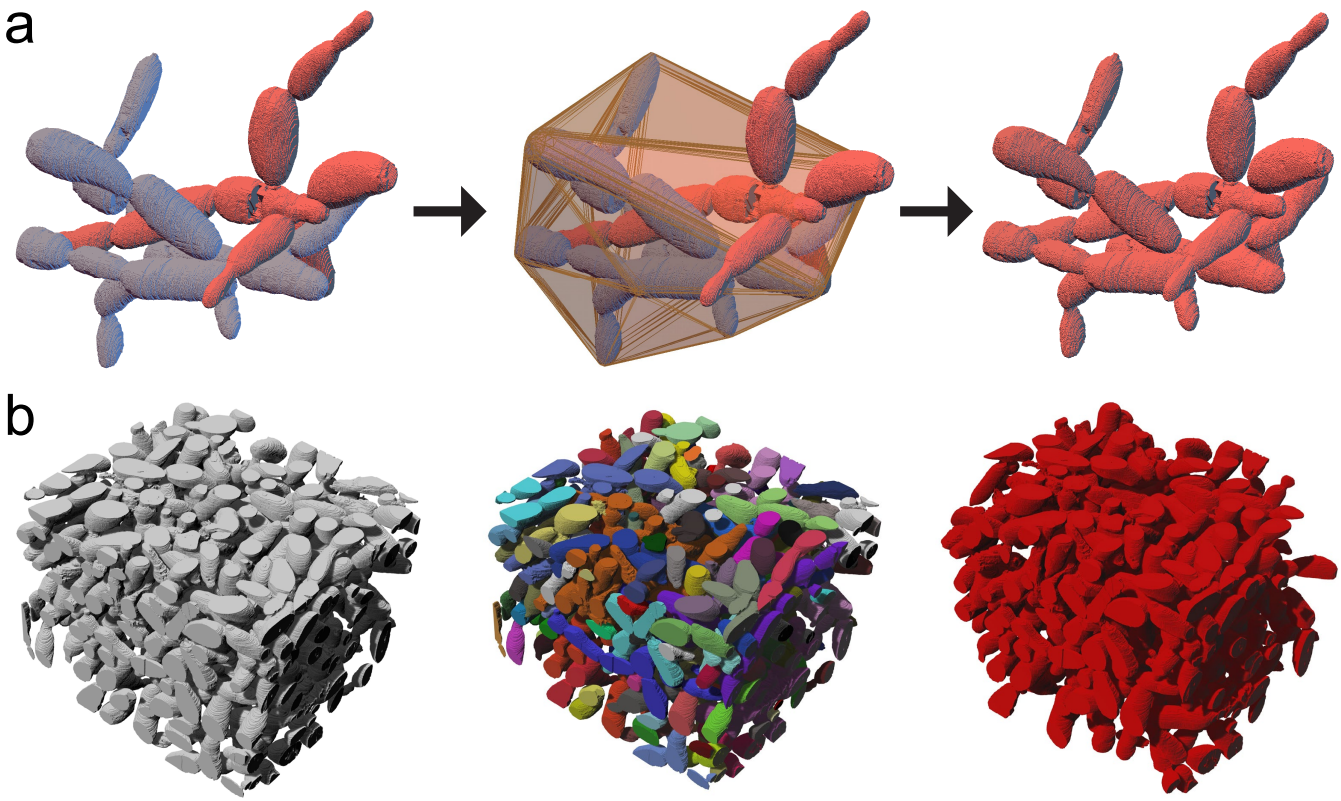
Extended Data Fig. 4 | Cell shape is not substantially affected by location within macroscopic yeast. a and **b** show cell volume and cell shape (aspect ratio) measured for 10 cells from the interior of a macroscopic cluster and 10 cells from the exterior of a cluster (measured in t600 macroscopic clusters). Average cell volume for exterior and interior are 110.8 μm^3 and 113.1 μm^3 ($p = 0.88$, $t = 0.15$ df = 17.55, Welch's t-test), and average cell shape for exterior and interior are 2.9 and 2.8 ($p = 0.51$, $t = 0.68$, df = 14, Welch's t-test). Individual measurements are marked as points, the mean and one standard deviation are indicated by the bar plot.



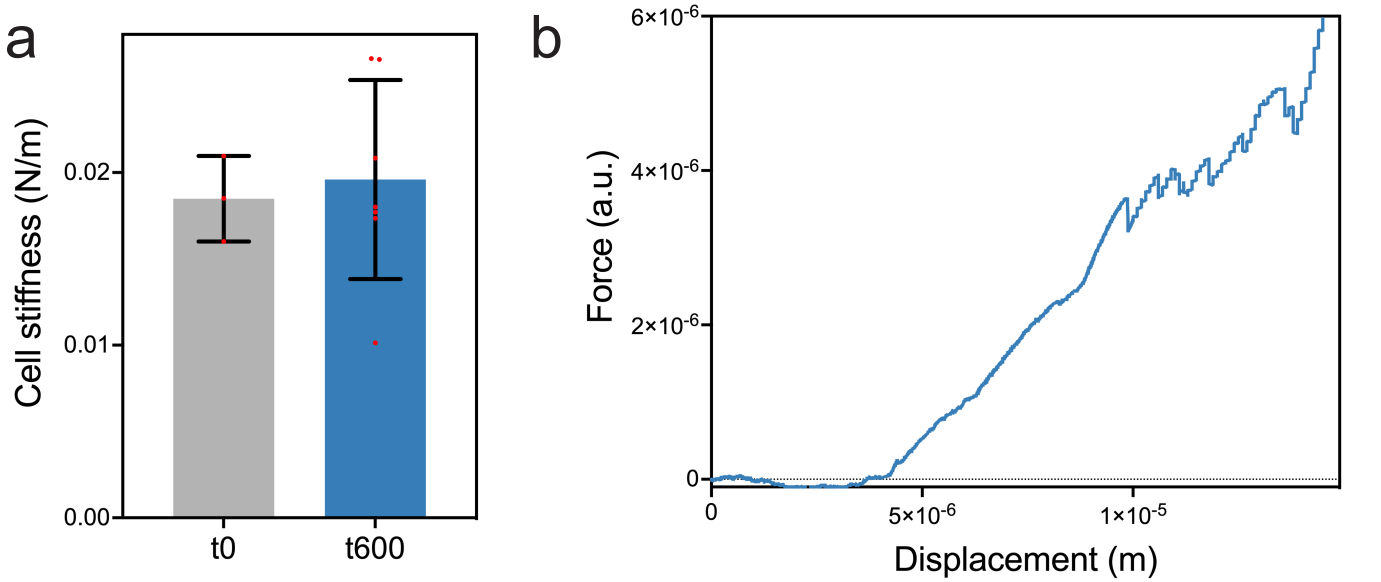
Extended Data Fig. 5 | Parallel evolution of elongated cell shape across all five replicates of each PA population. For each evolutionary time point and population, five different cells are shown (organized vertically from left to right: PA1 on the further left and PA5 on the further right in each box). Scale bar is 5 μm (under the ancestral cell). This is a more detailed version of the plot shown in Fig. 2c.



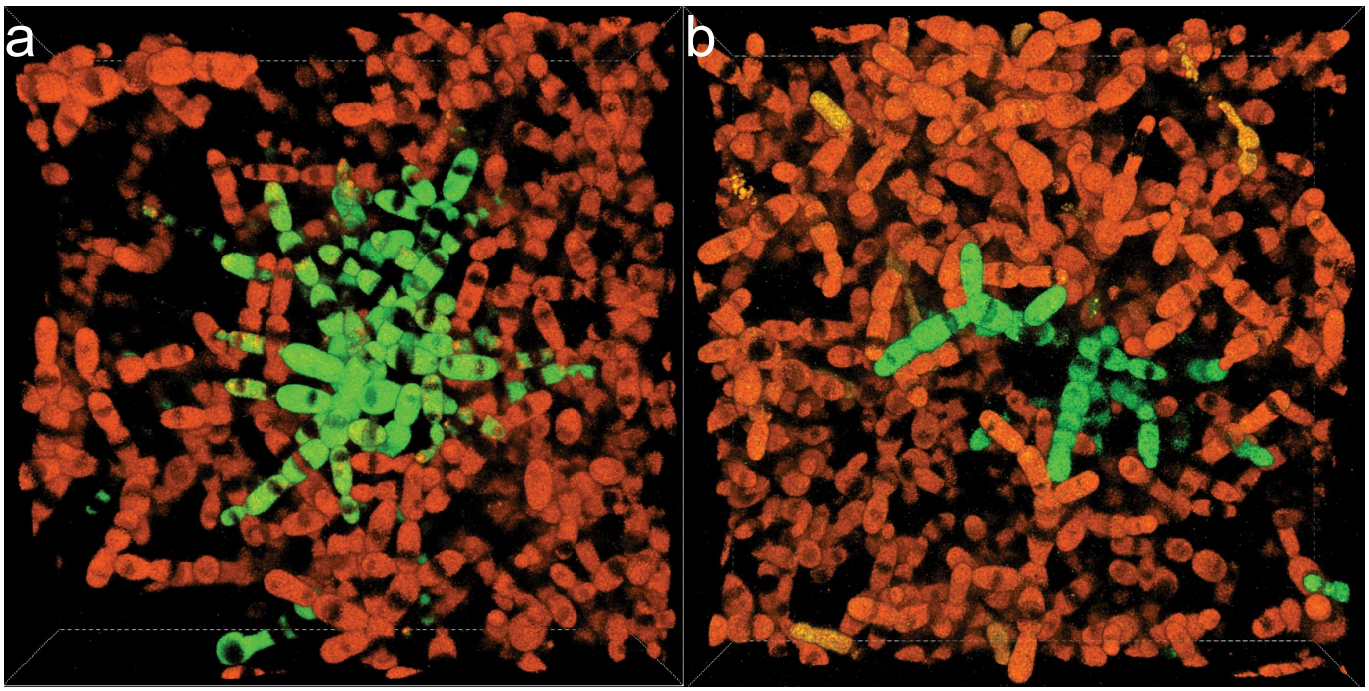
Extended Data Fig. 6 | Macroscopic snowflake yeast are monoclonal, growing via permanent mother-daughter cellular bonds, not aggregation. We co-cultured GFP and RFP-tagged genotypes of a macroscopic single strain isolate (PA2, strain ID: GOB1413-600) for 5 days, then imaged 70 clusters on a Nikon Ti-E. 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. Scale bar (top-left) is 100 µm.



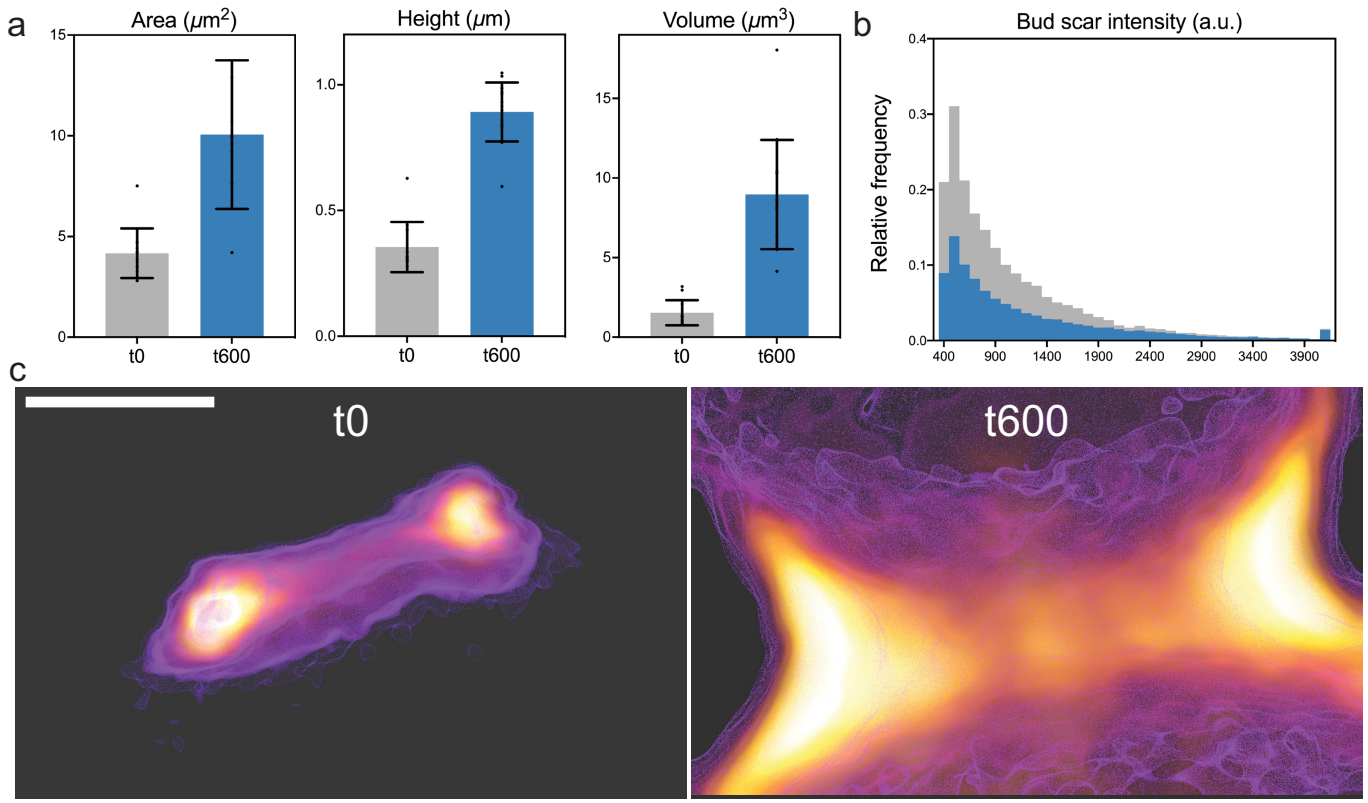
Extended Data Fig. 7 | Quantifying entanglement via analysis of the topology and geometry of a snowflake yeast cluster. a, We measured entanglement of individual components by fitting a convex hull around each component, and determining whether the other component overlaps with the space bounded by this 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 (coloured in red).



Extended Data Fig. 8 | Cell stiffness and stress-strain curve. a, Individual cells do not change their stiffness over 600 rounds of selection (average cell stiffness for the ancestor and t600 isolates are 0.019 and 0.020, respectively. $p = 0.77$, t = 0.31, df = 8, Welch's unequal variances *t*-test). Single-cell stiffness values measured from atomic force microscopy (AFM) of individual cells. Error bars are one standard deviation. **b**, Macroscopic snowflake yeast fractured into small modules prior to compression do not show strain stiffening behaviour. Shown here is an AFM trajectory of cantilever deflection vs displacement for one t600 cluster that has been crushed into small, unentangled pieces.



Extended Data Fig. 9 | Representative confocal images show chimeric clusters that are formed after growth in liquid culture followed by entanglement on agar plates. a,**b**, Close-up view of clusters, highlighting the tangled red and green branches. Each frame is 139.64 x 139.64 x 34.50 µm in X, Y, and Z axes, respectively.



Extended Data Fig. 10 | Dimensions of bud scars connecting cells in microscopic, ancestral (t0, grey) and macroscopic, evolved snowflake yeast clusters (PA2 t600, blue). Macroscopic t600 yeast had 2.4x larger bud scar cross-sectional area (**a**; $p < 0.001$, $t = 5.3$, df = 24, t-test), 2.8x greater bud scar height (**b**; $p < 0.001$, $t = 12.5$, df = 24, t-test), resulting in bud scars with 5.8-fold greater volume (**c**; $p < 0.001$, $t = 7.3$, df = 24, t-test) than the microscopic ancestor. Error bars are one standard deviation. **b**, Histogram of pixel intensities for bud scars stained with chitin stain calcofluor white, isolated from ancestor (t0, microscopic) and t600 (macroscopic) bud scars. The t600 strain has a 27% higher mean fluorescence intensity, suggesting that they may have evolved moderately higher chitin density in the bud scar. **c**, The size differences in bud scars is readily visible. Shown are the side view of buds from the ancestor (left) and t600 evolved (right), imaged at the same microscope settings. The scale bar is 0.5 µm.

nature portfolio

Corresponding author(s): G. Ozan Bozdag, Peter J. Yunker, William C. Ratcliff

Last updated by author(s): Mar 17, 2023

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed
☐	☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒ The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐ A description of all covariates tested
☒	☐ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒ For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐ Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Multicellular (or cluster) size, aspect ratio, and volume fraction data were collected using a Nikon Eclipse Ti inverted microscope (NIS-Elements version 4.30.01). 3D structure of the macroscopic snowflake yeast was resolved using a Serial Bulk Faced Scanning Electron Microscope (imaged on a Zeiss Sigma VP 3View system). Mechanical testing (presented in Fig3c) data were collected using a Puima Chiaro nanoindenter (Optics11, 19.5 um spherical glass probe; for ancestral clusters) or a Zwick Roell Universal Testing Machine (UTM; for macroscopic clusters). Whole genomes were sequenced using an Illumina HiSeq 2500 platform.
Data analysis	bwa-mem (BWA version 0.1.17), GATK4 HaplotypeCaller (v4.0.3.0), VCFTOOLS (version 0.1.17), SnpEff (v4.3T), Trimmomatic (v0.39), BCFtools-isec (v1.10), and Linux bash shell were used in sequence analysis and data parsing. Prism GraphPad version 9.5.0 for Mac OS was used for statistical analysis. Prism GraphPad version 9.5.0 for Mac OS was used for regression analyses. NIS-Elements version 4.30.01 and ImageJ (2.3.0) built-in macro functions were used for image analysis (measuring the cross-sectional area of snowflake yeast and aspect ratio of single cells; Fig 1e&f and Fig2d). An online tool (at yeast genome.org) was used to perform GO-term analysis (presented in Fig 4d). A custom Python code was used to calculate the average multicellular size presented in Fig 1e - we provide code used to generate plots in a public repository (GitHub). Mathematica 12 and Rhino6 were used to analyze Serial Bulk Faced Scanning Electron Microscope images. A new MATLAB code was provided as Supplementary File 1.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

All source data, raw data, and code are available at https://github.com/ozanbozdag/de_novo_evolution_of_macroscopic_multicellularity. Microscopy images are archived at Ratcliff Lab Dropbox account and are available upon request.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☐ Life sciences ☐ Behavioural & social sciences ☒ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Ecological, evolutionary & environmental sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	We conducted an evolution experiment over 600 days to select microscopic snowflake yeast clusters for large multicellular size. To investigate the effects of oxygen availability, we evolved these yeast under conditions with relatively low oxygen levels (PM1-5 and PO1-5 evolved in 25% PAL oxygen) or under anaerobic conditions (PA1-5), making a total of 3 x 5 = 15 replicate populations. All five replicate populations that were not constrained by limiting oxygen evolved macroscopic multicellular size, while populations evolving under a quarter of modern-day oxygen levels remained microscopic. This study examines this transition from microscopic to macroscopic size through phenotypic analysis, 3D microscopy, biophysical testing, and genome sequencing.
Research sample	We employed diploid Saccharomyces cerevisiae as our model organism. To initiate our experiment with a multicellular phenotype, we deleted the ACE2 transcription factor from the yeast genome. This organism is a well-established eukaryotic model system, making it genetically and phenotypically accessible.
Sampling strategy	We began our evolution experiment with five replicate populations and three metabolic treatment groups. To analyze evolutionary trends, we simply studied whole population samples. In that, to measure multicellular size we collected data from 1150 multicellular clusters per population, and to measure cellular aspect ratio, we collected data from an average of 453 individual cells per population, across 12 time points x 5 replicate populations. For genome sequencing, we isolated a single clonal snowflake yeast cluster from these populations and sequenced those isolates.
Data collection	GOB performed the long-term evolution experiment and created a -80 archive of each population at intervals of approximately 10-15 days. These populations can be revived for further examination. The data were collected by GOB, AZ, PK, TCD, DTL, and AJB without any manipulation by the experimenter. Genome sequencing data was generated using the Illumina HiSeq-2500 platform at the IBB sequencing core facility located at GA-Tech. Cluster size and aspect ratio data were collected using a Nikon Eclipse Ti inverted microscope at the Ratcliff Lab. Mechanical testing using UTM was performed at the Yunker Lab. AZ collected SEM images at the Core Facilities at the Carl R. Woese Institute for Genomic Biology for the SEM imaging. Microscopy images are available upon request (archived in our institute's Dropbox account).
Timing and spatial scale	Long-term evolution experiment (corresponding to the first 600 days) was performed between 2017-2018. Multicellular size, aspect ratio, SEM images, mechanical testing, and genomics data collection and analysis were performed in 2019-2021. Revision experiments (Fig1a&f, 4f, Extended data figures 2, 4, 8, 9, and 10) performed in 2022.
Data exclusions	No data were excluded from the analysis. All source data used to generate figures and all raw data used to calculate average values are shared in GOB's GitHub account.
Reproducibility	All five populations (PA1-5) evolved macroscopic size mainly driven by cell-level changes, such as cellular elongation. This parallel trend (5/5) not only demonstrates the reproducibility of our data but also provides confidence that, if these populations were to be re-evolved, the same outcome (macroscopic size) would arise. Mechanical testing was performed on 10 replicate clusters across various time points, and no failed attempts were encountered in the measurement or replication of these analyses. We have kept a frozen record of all those strain isolates and populations in our -80 freezer, which are available for further analysis, such as replicating our results. Therefore, frozen samples can be revived to reproduce the data presented in the paper.
Randomization	To ensure masked sample collection, we assigned random labels to samples at all stages of data collection, including cluster size and aspect ratio measurements, generation time measurements, and genomic DNA library preparation.

Blinding	Blinding was not necessary for our study as we did not need to make any decisions based on phenotypic analysis. We collected multicellular size and aspect ratio data by sampling a significant portion of the clusters from the original populations.

Did the study involve field work? ☐ Yes ☒ No

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☐	☒ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals	Other organisms are used: Saccharomyces cerevisiae - a non-pathogenic microbial eukaryote. Stain: Y55.
Wild animals	n/a
Field-collected samples	n/a
Ethics oversight	n/a

Note that full information on the approval of the study protocol must also be provided in the manuscript.