# Spatiotemporal distribution of the glycoprotein pherophorin II reveals stochastic geometry of the growing ECM of *Volvox carteri*

**Benjamin von der Heyde[a,1], Anand Srinivasan[b,1], Sumit Kumar Birwa[b], Eva Laura von der Heyde[a], Steph S.M.H. Höhn[b,2], Raymond E. Goldstein[b,2], and Armin Hallmann[a,2]**

This manuscript was compiled on May 23, 2025

**The evolution of multicellularity involved the transformation of a simple cell wall of unicellular ancestors into a complex, multifunctional extracellular matrix (ECM). A suitable model organism to study the formation and expansion of an ECM during ontogenesis is the multicellular green alga *Volvox carteri*, which, along with the related volvocine algae, produces a complex, self-organized ECM composed of multiple substructures. These self-assembled structures primarily consist of hydroxyproline-rich glycoproteins, a major component of which is pherophorins. To investigate the geometry of the growing ECM, we fused the *yfp* gene with the gene for pherophorin II (PhlI) in *V. carteri*. Confocal microscopy reveals PhlI:YFP localization at key ECM structures, including the boundaries of compartments surrounding each somatic cell and the outer surface of the organism. Image analysis during the life cycle allows the stochastic geometry of growing compartments to be quantified; their areas and aspect ratios exhibit robust gamma distributions and exhibit a structural transition from a tight polygonal to a looser acircular packing geometry with stable eccentricity over time, evoking parallels and distinctions with the behavior of hydrated foams. These results provide quantitative insight into a general, open question in biology: How do cells collectively produce a complex structure external to themselves in a robust and accurate manner?**

Extracellular Matrix | *Volvox carteri* | Geometry

## Significance Statement

The extracellular matrix (ECM) plays many important structural, developmental and physiologic roles in animals, fungi, plants and algae, and was particularly important in evolutionary transitions from unicellular to multicellular organisms. As the ECM is by definition external to cells, there must necessarily be an aspect of self-assembly involved in its generation, yet little is known regarding such processes even in the simplest multicellular species. Here we report the development of a transgenic strain of the multicellular green alga *Volvox carteri*, in which a key ECM protein, pherophorin II, is fused with yellow fluorescent protein, allowing the first quantitative study of stochastic ECM geometry and growth dynamics in this system, which reveals behavior that is reminiscent of yet distinct from the hydration of foams.

Throughout the history of life, one of the most significant evolutionary transitions was the formation of multicellular eukaryotes. In most lineages that evolved multicellularity, including animals, fungi and plants, the extracellular matrix (ECM) has been a key mediator of this transition by connecting, positioning, and shielding cells (1–4). The same holds for multicellular algae such as *Volvox carteri* (Chlorophyta) and its multicellular relatives within volvocine green algae which developed a remarkable array of advanced traits in a comparatively short amount of evolutionary time—oogamy, asymmetric cell division, germ-soma division of labor, embryonic morphogenesis and a complex ECM (5–8)—rendering them uniquely suited model systems for examining evolution from a unicellular progenitor to multicellular organisms with different cell types (5, 6, 8–10). In particular, *V. carteri*'s distinct, multilayered ECM makes it a model organism for investigating mechanisms underlying ECM growth and the dynamics of its structures alongside the positions of the cells that secrete its components. Building on recent protocols for stable expression of fluorescently labeled proteins in *V. carteri* (11–13) we present here a new transgenic strain revealing localization of the glycoprotein pherophorin II and the first *in vivo* study of the stochastic geometry of a growing ECM.

*V. carteri* usually reproduces asexually (Fig. 1A). Sexual development is triggered by exposure to heat or a species-specific glycoprotein sex inducer, which results in development of sperm-packet-bearing males and egg-bearing females (14–17). In the usual asexual development, *V. carteri* consists of $\sim$ 2000 biflagellated somatic cells resembling *Chlamydomonas* in their morphology, arranged in a monolayer at the surface of a sphere, and $\sim$16 much larger, nonmotile, asexual reproductive cells (gonidia) that constitute the germline, lying just below the somatic cell layer (5–7, 18, 19). The somatic cells are specialized for photoreception and motility, and perform ECM biosynthesis jointly with the gonidia; for phototaxis, they must be positioned correctly within the ECM at the surface of the organism (14, 20, 21).

The ECM of *V. carteri* has been studied in the past decades from the perspective of structure and composition, developmental, mechanical properties, cellular interactions, molecular biology and genetics, and evolution (11, 14, 20–22). In the ontogenesis of *V. carteri*, ECM biosynthesis in juveniles begins while inside

Author affiliations: [a]Department of Cellular and Developmental Biology of Plants, University of Bielefeld, Universitätsstr. 25, 33615 Bielefeld, Germany; [b]Department of Applied Mathematics and Theoretical Physics, Centre for Mathematical Sciences,, University of Cambridge, Wilberforce Road, Cambridge CB3 0WA, United Kingdom

All authors designed research, performed research, analyzed data and wrote the paper.

The authors declare no competing interests.

[1] B.v.d.H. and A.S. contributed equally to this work.

[2]Emails: sh753@cam.ac.uk, reg53@cam.ac.uk & armin.hallmann@uni-bielefeld.de

DRAFT

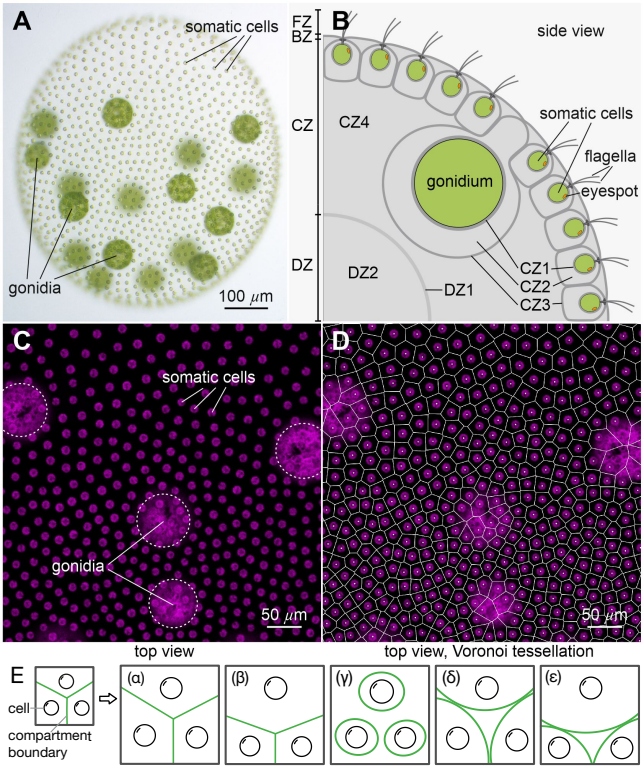


**Fig. 1.** Phenotype, ECM architecture, and cell distribution of *V. carteri*. (A) Adult female (wild-type). Somatic cells each bear an eyespot and two eukaryotic flagella (motile cilia), while deeper-lying gonidia are flagella-less. All cells have a single large chlorophyll-containing chloroplast (green). (B) Zones of the ECM, as described in text. (C) Image utilizing autofluorescence of the chlorophyll (magenta) to define cell positions. (D) Semi-automated readout of somatic cell positions in (C) followed by a Voronoi tessellation gives an estimate of the ECM neighborhoods of somatic cells. (E) Hypothetical expansion dynamics ($\alpha$-$\varepsilon$) of a compartmentalized ECM, e.g. the *Volvox* CZ2/3 (panel B). Factors include wall merging & adhesion (exemplified by $\alpha, \delta$), permeability ($\gamma$), and cell-cell heterogeneous ECM production ($\beta, \varepsilon$).

their mothers and only after all cell divisions and the process of embryonic inversion are completed. Cells of the juveniles then continuously excrete large quantities of ECM building blocks which are integrated into their ECM, producing an enormous increase in the size; within 48 hours the volume increases ∼3000-fold as the diameter increases from ∼ 70 $\mu$m to ∼ 1 mm, raising the question of how ECM geometry and relative cell positions transform during growth. In the adult, the ECM accounts for ∼ 99 % of the organism's volume and consists of morphologically distinct structures with a defined spatial arrangement. Based on electron microscopy, a nomenclature was established (20) that defines four main ECM zones: flagellar zone (FZ), boundary zone (BZ), cellular zone (CZ) and deep zone (DZ), which are further subdivided (Fig. 1B). The CZ3 forms the ECM compartment boundaries of individual cells and the BZ constitutes the outer surface of the organism. CZ3 and BZ show a higher electron density than the ECM within the compartments or the ECM in the interior below the cell layer (CZ4 and DZ2) (6, 7, 20). CZ3 and BZ therefore appear to consist of a firmer, more robust ECM, while CZ4 and DZ2 appear to be more gelatinous.

The ECM of volvocine algae predominantly consists of hydroxyproline-rich glycoproteins (HRGPs), which are also a major component of the ECM of embryophytic land plants (14, 21, 23–25), but contains no cellulose. In *V. carteri* and other volvocine algae, a large family of HRGPs, the pherophorins, are expected to constitute the main building material of the ECM (17, 22, 26–29). Just in *V. carteri*, 118 members of the pherophorin family were discovered in the genome (11, 30); pherophorin genes are typically expressed in a cell type-specific manner (31) that is constitutive or induced either by the sex inducer or wounding (17, 22, 26, 28, 29, 32). The dumbbell-like pherophorin tructure has two globular domains separated by a rod-shaped, highly proline-rich one that varies considerably in length (14, 21, 28, 29). These prolines undergo post-translational modification to hydroxyprolines; pherophorins are also strongly glycosylated (14, 21, 22, 33, 34). For two pherophorins the polymerization into an insoluble fibrous network was shown in vitro (26); some have already been localized in the ECM or in ECM fractions (11, 13, 22, 27, 29).

To investigate the ECM, we determined that pherophorin II (PhII) was the appropriate one to label fluorescently, as it is firmly integrated into the ECM. It is a 70 kDa glycoprotein with constant, weak background expression but also strong and immediate inducibility by the sex inducer (17, 35, 36). Because it can only be extracted under harsh conditions, it is thought to be a component of the insoluble part of the somatic CZ (17, 35, 36). We fused one of the nine (28, 30) gene copies of PhII with the *yfp* gene. The corresponding DNA construct was stably integrated into the genome by particle bombardment. The expression of fluorescent PhII:YFP was confirmed through confocal laser scanning microscopy (CLSM); it was found in the boundary zone and compartment boundaries of the cellular zones.

We suggest that a detailed study of the stochastic geometry of a growing ECM made possible by this strain of *Volvox* will provide insight into a more general question in biology: How do cells robustly produce structures external to themselves? There is no physical picture of how the intricate geometry of the *Volvox* ECM arises through what must be a self-organized process of polymer crosslinking (37) and hydration. While information on a structure's growth dynamics can be inferred from its evolving shape, as done for animal epithelial cells (38), this connection has not been made for *Volvox*. The dynamics of the compartmental geometry (Fig. 1E) give clues as to wall material properties (including permeability and adhesion) and by extension the "autonomy" (degree of spatial constraint) of ECM deposition during growth. Some hypothetical regimes (Fig. 1E($\beta$ and $\varepsilon$) provide readouts of cell-to-cell fluctuations in ECM production and expansion of the local region, making the PhII:YFP-stained CZ3 a valuable tool through which to understand spheroid morphogenesis. Some evoke visual analogies to two-dimensional dry (1E($\alpha$–$\beta$)) and wet (1E($\gamma-\varepsilon$)) foams, observed for more than a century in tissues (39, 40), yet relatively unexplored in an ECM.

Recently (41), the somatic cells of *V. carteri* were located via their chlorophyll autofluorescence, from which the ECM neighborhoods were found by a surface Voronoi tessellation, as in Figs. 1C,D. While the cells appear to be arranged in a quasi-regular pattern, their Voronoi areas exhibited a broad, skewed k-gamma distribution. Yet, the validity of the Voronoi model for actual ECM structures—such as the CZ3—during development has remained uncertain. And while such universal size and shape distributions have been

DRAFT

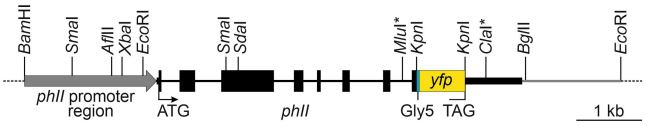


**Fig. 2.** Schematic diagram of the transformation vector pPhII-YFP. The vector carries the complete *phII* gene including its promoter region. Exons (introns) are shown as black boxes (thin lines). Directly upstream of the TAG stop codon, a 0.7 kb fragment coding for a flexible penta-glycine spacer (cyan) and the codon-adapted coding sequence of *yfp* (mVenus) (yellow) were inserted in frame using an artificially introduced *Kpn*I site. The *Kpn*I was inserted before into the section between *Mlu*I and *Cla*I (asterisks). The vector backbone (dashed lines) comes from pUC18. For details, see Methods; the sequence of the vector insert is in SI Appendix, Fig. S3.

identified in granular materials (42), mathematical constructs (43), epithelial tissues, and inert jammed matter (44), it is unknown whether this stochastic geometry also arises in a compartmentalized ECM. Subsequent work suggested that gamma-distributions may arise from bursty cellular ECM production (45). To investigate this and to answer the more fundamental questions raised above, we analyzed the geometry of the structure illuminated by PhII:YFP.

Our results show, first and foremost, that the compartmental geometry remains a connected and nearly space-filling structure throughout development; its geometry provides a meaningful readout of the spheroid's developmental dynamics during expansion. We find that this structure does not simply dilate during growth (globally 1E($\alpha$) or locally 1E($\beta$)) but instead undergoes a foam-like structural transition (Fig. 1E($\varepsilon$)) from an initial configuration that closely resembles Voronoi geometry. During this transition, we find that shape features (Table 1) exhibit the same stochastic geometry alluded to above: a cell-scale randomness in size and shape which strikingly coexists with robust organism-scale symmetries and anterior-posterior differentiation.

## Results

**A. Vector construction and generation of transformants expressing PhII:YFP.** The pherophorin II gene (*phII*) was cloned from *V. carteri* genomic DNA including its promoter, 5' and 3' UTRs and all seven introns (Fig. 2). A sequence coding for a flexible penta-glycine spacer and the codon-adapted coding sequence of *yfp* were inserted directly upstream of the *phII* stop codon to produce a *phII-yfp* gene fusion. The obtained vector pPhII-YFP (Fig. 2) was sequenced and then used for stable nuclear transformation of the nitrate reductase-deficient *V. carteri* recipient strain TNit-1013 by particle bombardment. To allow for selection, the non-selectable vector pPhII-YFP was co-transformed with the selectable marker vector pVcNR15, which carries an intact *V. carteri* nitrate reductase gene and, thus, complements the nitrate reductase deficiency of strain TNit-1013. Screening for transformants was then achieved by using medium with nitrate as nitrogen source. The transformants were investigated for stable genomic integration of the vector and, via confocal microscopy, for expression of the fluorescent protein at sufficient levels throughout their life cycle.

**B. *In vivo* localization of pherophorin II.** As expected and detailed below, there are continuous changes in the amount of ECM expansion over the life cycle. While we primarily

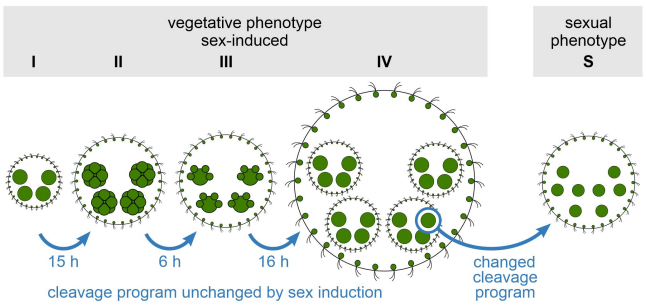


**Fig. 3.** Stages of development of *V. carteri* females. Localization of PhII and quantification of ECM features were studied in 5 stages, four with vegetative (asexual) phenotype (I-IV) and one with the fully developed sexual phenotype (S). In stages I-IV, incubation with the sex inducer was short enough to increase PhII expression for adequate localization without changing the vegetative cleavage program.

examine the parental somatic cell layer and surrounding ECM, the developmental stage of the next generation is used for a precise definition of five key stages used in the comparative analyses (Fig. 3). Stage I: Freshly hatched young adults of equivalent circular radius $R = 106 \pm 6\,\mu$m (SI Appendix, §2.B), containing immature gonidia. Stage II ($\sim$ 15 h post hatching (hph)): Middle-aged adults with $R = 221 \pm 22\,\mu$m containing early embryos (4-8 cell stage). Stage III ($\sim$ 21 hph): Older middle-aged adults with $R = 244 \pm 15\,\mu$m containing embryos before inversion. Stage IV ($\sim$ 36 hph): Old adults with $R = 422 \pm 6\,\mu$m containing fully developed juveniles. Stage S: Sexually developed adult females with $R = 265 \pm 29\,\mu$m bearing egg cells. Since expression of PhII is induced by the sex-inducer protein (17, 35, 36), in stages I to IV with vegetative phenotype, it was added 24 hours before microscopy to increase PhII expression; after such a short incubation with the sex inducer, the females show unchanged cleavage programs, unaltered growth dynamics and vegetative phenotypes (SI Appendix, Fig. S12). To obtain a changed cleavage program and a fully developed sexual phenotype (S), the females were sexually induced 72 hours before microscopy.

**C. Pherophorin II is localized in the compartment borders of individual cells.** As expected, PhII:YFP is only found in the extracellular space within the ECM. It is detectable at all developmental stages after embryonic inversion, which marks the beginning of ECM biosynthesis (34, 46), and in the ECM of organisms with the phenotypes of both vegetative and sexual development. At a first glance, in a top view, PhII:YFP appears to form a polygonal pattern at the surface of each post-embryonic spheroid, with a single somatic cell near the center of each compartment (Fig. 4). PhII:YFP is also found in the ECM compartment boundaries (CZ3) of the gonidia, which are located below the somatic cell layer (Figs. 1B and 4). These observations hold at all stages after embryonic inversion, even as the shape of the compartment boundaries varies. Cross sectional views of CZ3 compartments along the radial axis (Fig. 4D and SI Appendix, Fig. S13) prove that in all stages (I to IV) the CZ3 compartments are consistently arranged next to each other and in a single layer, with no deep overlaps or other 3D complications as seen in some epithelia (47). We discuss cross-sectional shapes of the somatic CZ3 compartments in stages I to IV and corresponding estimations

DRAFT

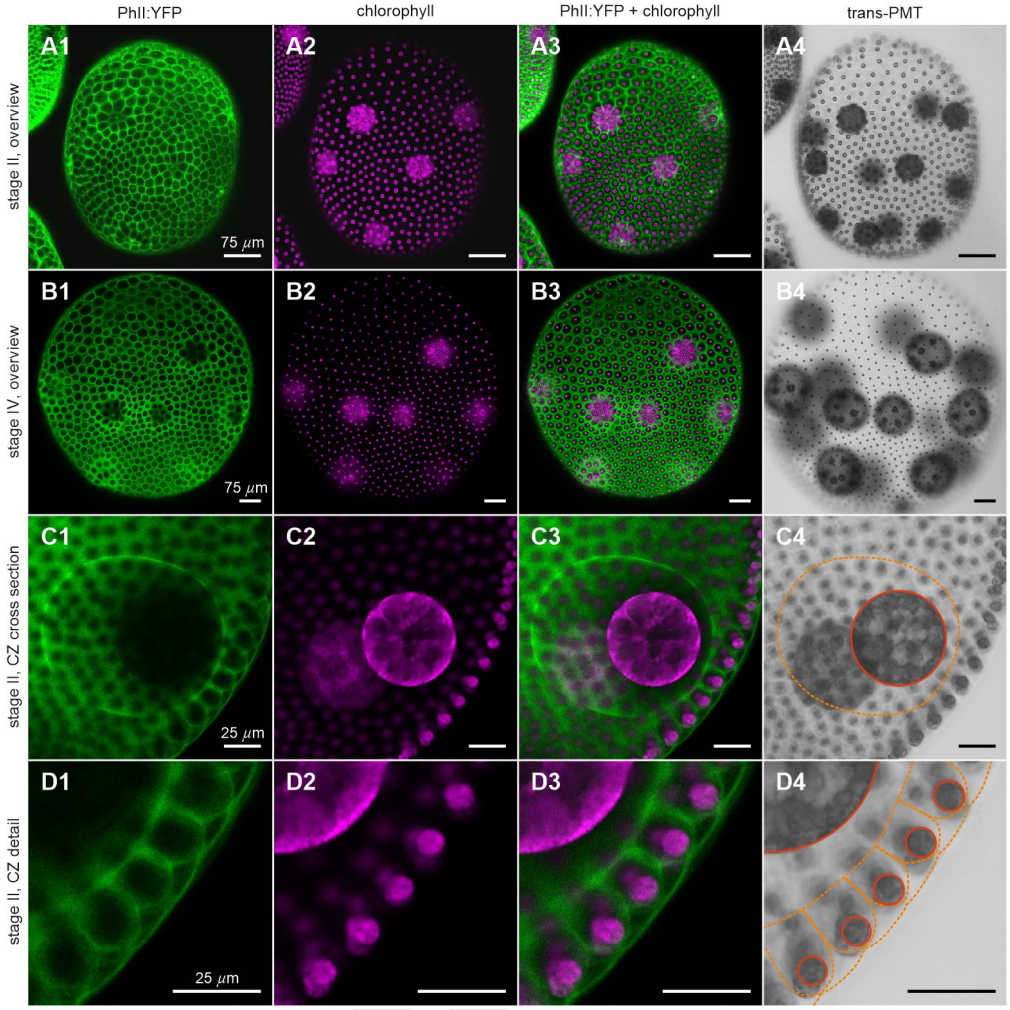


**Fig. 4.** Localization of PhII:YFP in whole, middle-aged (early stage II) and old adults (stage IV) in top view and magnified cross section through CZ. Sexually induced transformants expressing the *phII:yfp* gene under the control of the endogenous *phII* promoter were analyzed in vivo for the localization of the PHII:YFP fusion protein. PhII:YFP is located in the CZ3 of both somatic cells and gonidia as well as in the BZ. Top views of whole organism: (A) stage II immediately before the first cleavage of the gonidia inside; and (B) stage IV before hatching of the fully developed juveniles. (C) Cross section through CZ in early stage II. (D4) Magnified view of the outer most ECM region. Column 1: YFP fluorescence of the PhII:YFP protein (green). Column 2: Chlorophyll fluorescence (magenta). Column 3: Overlay of YFP and chlorophyll fluorescence. Column 4: Transmission-PMT (trans-PMT). PhII:YFP-stained ECM boundaries are highlighted in orange and cell boundaries in red.

of compartment volumetric growth in the Discussion §7 and SI Appendix (Figs. S13, S14, and Table S4).

On closer inspection in top view (Fig. 5), we see that: i) the somatic ECM compartments are a mix of hexagons, heptagons, pentagons, other polygons, circles and ovals, ii) the angularity of the compartment boundaries changes during expansion of the organism, i.e. the compartments become increasingly circular (less polygonal), and iii) each cell builds its own ECM compartment boundary. The observed localization of PhII is shown schematically in Fig. 12.

Especially in the early stages I and II, the fluorescent ECM compartment borders are predominantly pentagonal or hexagonal in top view (Figs. 4A and 5A) with a rough ratio of 1:2 (SI Appendix, Fig. S4). By stage IV they become more rounded and the interstices we term "extracompartmental ECM spaces" between the compartments increase in size and number (Figs. 4B and 5B). Since the compartment boundaries of adjacent cells are close to each other in early stages, the two compartments appear to be separated by a single wall. Later, when the boundaries are more circular, it becomes apparent that it is a double wall; each somatic cell produces its own boundary (Figs. 5A,B). Figs. 5A1-B1 show that the surface coverage of the CZ by compartments changes from $\sim 100\%$ to $\sim 80\%$ while the structure remains singly connected. The dynamics follow scenario Fig. 1E($\varepsilon$): the CZ3 is geometrically coupled to expansion of the spheroid. This coupling forms the basis for interpreting aspects of development, such as the degree of AP differentiation (Figs. 4A1-B1), directly from CZ3 geometry, as we explore in §D.

As shown, PhII:YFP is a component of the CZ3 of both somatic cells and gonidia (Figs. 1B, 4 and 5). It seems to be a firmly anchored building block there, as the observed structure is highly fluorescent and yet sharply demarcated from other non-fluorescent adjacent ECM structures. If the PhII:YFP protein was prone to diffusion, one would expect a brightness gradient starting from the structures. There are no interruptions in the fluorescent labeling of boundaries associated with passages between neighboring compartments.

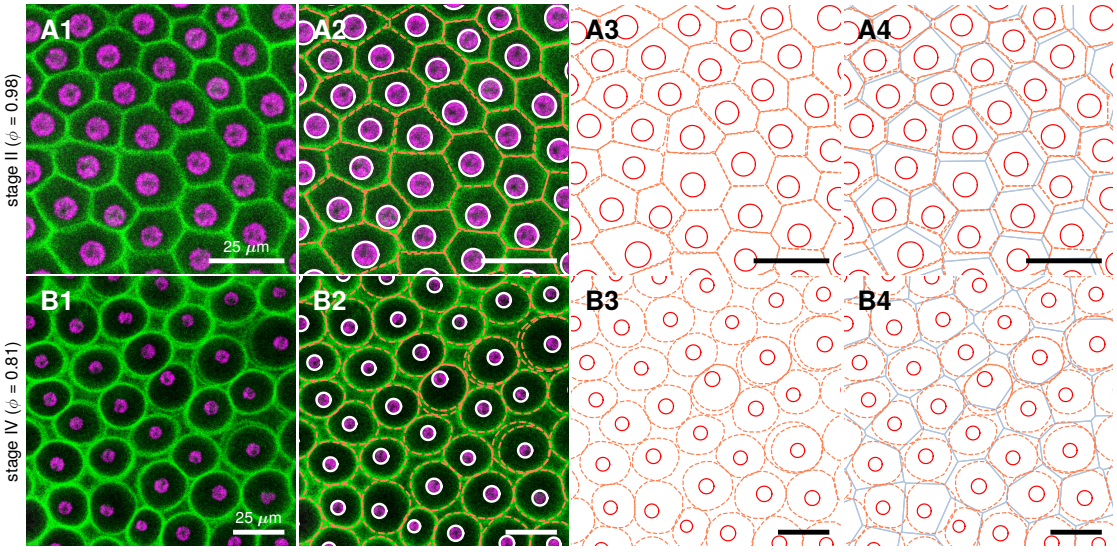


**Fig. 5.** As in Fig. 4, but a close-up of PhII:YFP localization in top view. (A) Stage II. Magnified view of the somatic CZ3 compartments (1), identified by orange in column 2 along with cell boundaries (white), shown together in columns 3 and 4 with underlaid Voronoi tessellation (blue) and cell boundaries in red. (B) In stage IV the CZ3 compartments become more bubble shaped and individual CZ3 walls separate from the neighbors, leaving extracompartmental ECM space. Double-walls also appear, highlighted in B3. The surface covering fraction by compartments in these top view sections is denoted by $\phi$, which decreases from approximately 100% in stage II (A1) to 80% in stage IV (B1).

As the boundaries are so close together in early stages, only the thickness of the double wall can be determined and then halved; we estimate stage II single wall thickness as $1.6 \pm 0.4\,\mu$m and a similar value of $1.9 \pm 0.5\,\mu$m in stage IV.

A comparison of the PhII:YFP-stained structures with descriptions of ECM structures from electron microscopy shows that the PhII:YFP location corresponds exactly to that of the ECM structure CZ3 (20). This applies to the entire period from the beginning of ECM biosynthesis after embryonic inversion to the maximally grown old adult.

The relatively regular pattern of compartments (Fig. 5A) is disturbed by the gonidia (later embryos) which, being far larger than somatic cells, are pushed under the somatic cell layer. The CZ3 compartments of the deeper gonidia nevertheless extend to the surface (Fig. 6). The surrounding somatic CZ3 compartments are elongated in the direction of the gonidial CZ3 protrusions to the surface (Fig. 6). Since the PhII:YFP-stained CZ3 structure completely encloses each cell, these walls cannot be impermeable; ECM proteins must pass through them, as growth of the spheroid requires the cell-free areas outside the ECM compartments of individual cells to increase immensely in volume. This applies in particular to the deep zone, but also to the areas between the compartments. All ECM material required for this increase can only be produced and exported by the cells and must then pass through the ECM compartment borders of cells.

**D. Quantification of surface somatic CZ3 geometry.** The localization of PhII at the CZ3 compartment boundaries allows us to carry out the first quantitative analyses of their surface geometry, both along the posterior-anterior (PA) axis and through the life cycle stages. A semi-automated image analysis (SI Appendix, §2) reveals geometric features described in Table 1 and Fig. 7. A total of 29 spheroids across five developmental stages were analyzed: 7 in stage I, 5 each in stages II, III, and IV of the asexual life cycle, and 7 in the sexual life cycle (stage S). The mounting procedure (Materials and Methods) induces some elastic distortion in top view; we report here the original features, discussing distortion correction in the SI Appendix. The results below (see Figs. 8 and 9) do not meaningfully differ with correction.

The PA axis of *V. carteri* spheroids is their swimming direction, and along this axis the distance between the somatic cells and the size of their eyespots decreases toward the posterior pole (Fig. 1A), as noted previously ((6), p. 229). Offspring (gonidia, embryos, and daughter spheroids) are mainly located in the posterior hemisphere. We approximate the PA axis by a line passing through the center of the spheroid and normal to the best-fit line passing through manually identified juveniles in the anterior (Fig. 7A). This estimated PA axis is typically well-approximated by the elliptical major axis of the spheroid (SI Appendix, Fig. S5).

The metrics shown in Fig. 7 (SI Appendix, §2) are derived from the top-viewed compartment shape outline or from geometric moments of area. The matrix $\mathsf{M}_2$ of $2^{\text{nd}}$ moments

$$\mathsf{M}_2 = \iint_{\text{cz3}} (\boldsymbol{x} - \boldsymbol{x}_{\text{cz3}}) \otimes (\boldsymbol{x} - \boldsymbol{x}_{\text{cz3}}) d^2\boldsymbol{x}, \qquad [1]$$

and its normalization $\mathsf{\Sigma} = \mathsf{M}_2/a_{\text{cz3}}$ are interpretable as elastic strain tensors with respect to unit-aspect-ratio shapes. The eigenvalues $\lambda_{\max}, \lambda_{\min}$ (whose square roots define the major, minor axis lengths) of $\mathsf{\Sigma}$ are principal stretches of this deformation, yielding the aspect ratio and other quantities defined in Table 1. Overall, we measure changes in the moments of area (SI Appendix, §2.B) to quantify ECM geometry during growth; the $0^{\text{th}}$ gives the area increase, the $1^{\text{st}}$ quantifies migration of compartment centroids with respect to cells, the $2^{\text{nd}}$ gives the change in eccentricity (Table 1). The sum of second moments reveals changes in crystallinity of the entire CZ3 configuration, as described in §G.

**E. Surface CZ3 geometry along PA axis during the life cycle.**

DRAFT

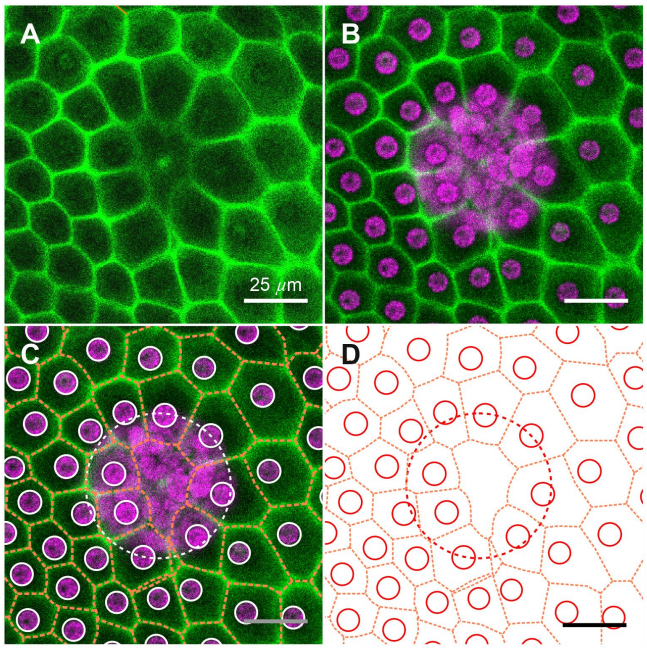


**Fig. 6.** Magnified top view of CZ3 above a reproductive cell in early stage II. (A) PhII:YFP fluorescence (green). (B) overlay with chlorophyll fluorescence (magenta). (C) overlay with highlighted CZ3 (orange) and cell boundaries (white). (D) as in (C) with only cell boundaries (red) and CZ3 (orange).

**Table 1. Metric definitions.**

| metric | symbol | definition | units |
|---|---|---|---|
| Somatic cell area | $a_{\text{cell}}$ | - | $\mu$m$^2$ |
| Somatic cell centroid | $\boldsymbol{x}_{\text{cell}}$ | - | $\mu$m |
| CZ3 compartment area | $a_{\text{cz3}}$ | - | $\mu$m$^2$ |
| CZ3 compartment centroid | $\boldsymbol{x}_{\text{cz3}}$ | - | $\mu$m |
| CZ3 compartment perimeter | $\ell_{\text{cz3}}$ | - | $\mu$m |
| CZ3 covariance matrix | $\Sigma$ | Eq. (1) | $\mu$m$^2$ |
| Aspect ratio | $\alpha$ | $\sqrt{\lambda_{\max}/\lambda_{\min}}$ | unitless |
| Circularity | $q$ | $\sqrt{4\pi a_{\text{cz3}}}/\ell_{\text{cz3}}$ | unitless |
| Somatic cell offset vector | $\Delta\boldsymbol{x}$ | $\boldsymbol{x}_{\text{cell}} - \boldsymbol{x}_{\text{cz3}}$ | $\mu$m |
| Somatic cell offset | $r$ | $\|\Delta\boldsymbol{x}\|$ | $\mu$m |
| Somatic cell offset (whitened) | $\tilde{r}$ | $\sqrt{\Delta\boldsymbol{x}\cdot\Sigma^{-1}\Delta\boldsymbol{x}}$ | unitless |
| Voronoi error | $e_V$ | vor $\cap$ cz3 / vor $\cup$ cz3 | unitless |

***E.1. Anterior CZ3 compartment areas expand toward end of life cycle.*** Fig. 8A1-2 shows that somatic cell area increases modestly, by $\sim 10\%$, along the PA axis at all stages. In contrast, the CZ3 compartment area grows substantially along this axis, from $\sim 44\%$ in stage I to $\sim 130\%$ in stage IV (SI Appendix, Table S2). Moreover, the slope increases after the equatorial region in all stages, most prominently in stage IV (panel B1). Cell and compartment areas also increase by life cycle stage as shown in Fig. 8A-B. Somatic cell areas double from I-II, growing merely $\sim 15\%$ afterwards, whereas compartment areas expand primarily after III, with a $\sim 160\%$ increase occurring from III-IV (SI Appendix, Table S2).

Fig. 8A2-G2 shows distributions of the metrics; apart from cell area, all exhibit positive skew and exponential tails which suggest good fits with gamma-type distributions (41),

$$p_{\lambda,k}(x) = \lambda^k x^{k-1} e^{-\lambda x}/\Gamma(k), \qquad [2]$$

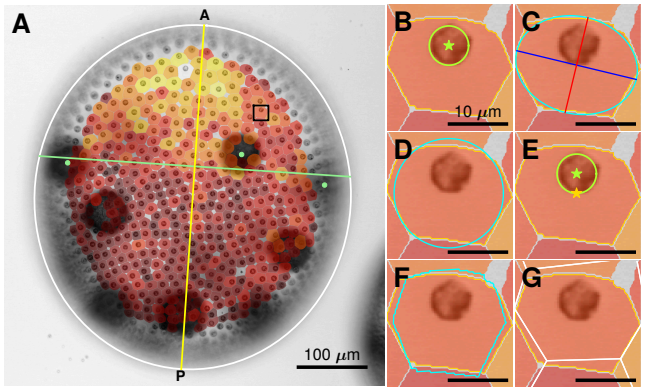


**Fig. 7.** Geometric features of cell/compartments pairs. (A) Trans-PMT image of stage III spheroid in top view, with elliptical outline (white), estimated PA axis (yellow) that is orthogonal to line through gonidia (green dots). Overlaid are segmentations of CZ3 compartments, colored dark to light by size. (B-G) Schematics of geometric features computed from cell (green) and compartment (yellow) boundaries, as indicated: (B) $a_{\text{cell}}, a_{\text{cz3}}$, (C) aspect ratio $\alpha$ and corresponding ellipse (cyan) with major and minor axes in blue and red, (D) deviation from a circle of the same area (cyan), (E) offset of cell area center of mass (green star) from compartment area center of mass (yellow star), (F) whitening transform of the compartment area (cyan), and (G) Voronoi tessellation (white) error $e_V$.

where $x$ is suitably standardized. This skew should be considered when making mean-based comparisons across life cycle stages. The long left tails of cell area reflect the persistence of small somatic cells throughout the life cycle, confirmed by inspection in the chlorophyll signal. Lastly, the cell size distribution primarily translates rightward in time, while the compartment area size distribution simultaneously translates and stretches, indicating an increase in polydispersity.

***E.2. CZ3 compartment areas transition from tighter polygonal to looser elliptical packing.*** Panels D1-2 in Fig. 8 show that while there is no apparent trend in the circularity of CZ3 compartment areas along the PA axis, the average circularity increases from stages I to IV. Since extracompartmental ECM space appears as compartment areas increase in circularity both effects correlate with enlargement of the spheroid. Fig. 8D2 shows that circularity increases in mean while decreasing in variance, suggesting a relaxation process by which compartments of a particular aspect ratio but different polygonal initial configurations relax to a common elliptical shape with the same aspect ratio. This is also apparent by

| Stage | Parent radius ($\mu$m) | Offspring radius ($\mu$m) | Somatic cell radius ($\mu$m) | Parental ECM volume change (est., mm$^3$) | Parental ECM growth rate (est., mm$^3$/h) |
|---|---|---|---|---|---|
| I | $106 \pm 6$ | $16 \pm 2$ | $2.8 \pm 0.5$ | $\downarrow$ | $\downarrow$ |
| II | $221 \pm 19$ | $29 \pm 2$ | $3.9 \pm 0.4$ | 0.039 | 0.0026 |
| III | $244 \pm 13$ | $30 \pm 4$ | $4.0 \pm 0.3$ | 0.015 | 0.0025 |
| IV | $422 \pm 6$ | $79 \pm 6$ | $4.2 \pm 0.4$ | 0.229 | 0.0143 |
| S | $265 \pm 26$ | $15 \pm 1$ | $4.1 \pm 0.3$ | n/a | n/a |

**Table 2. Summary of estimated volumetric growth by stage. Values reported are mean $\pm$ standard deviation. Estimated ECM volume is volume of spheroid minus that estimated of juveniles and somatic cells, as explained in SI Appendix, Table S1. Values in final two columns represent changes with respect to preceding stage.**

DRAFT

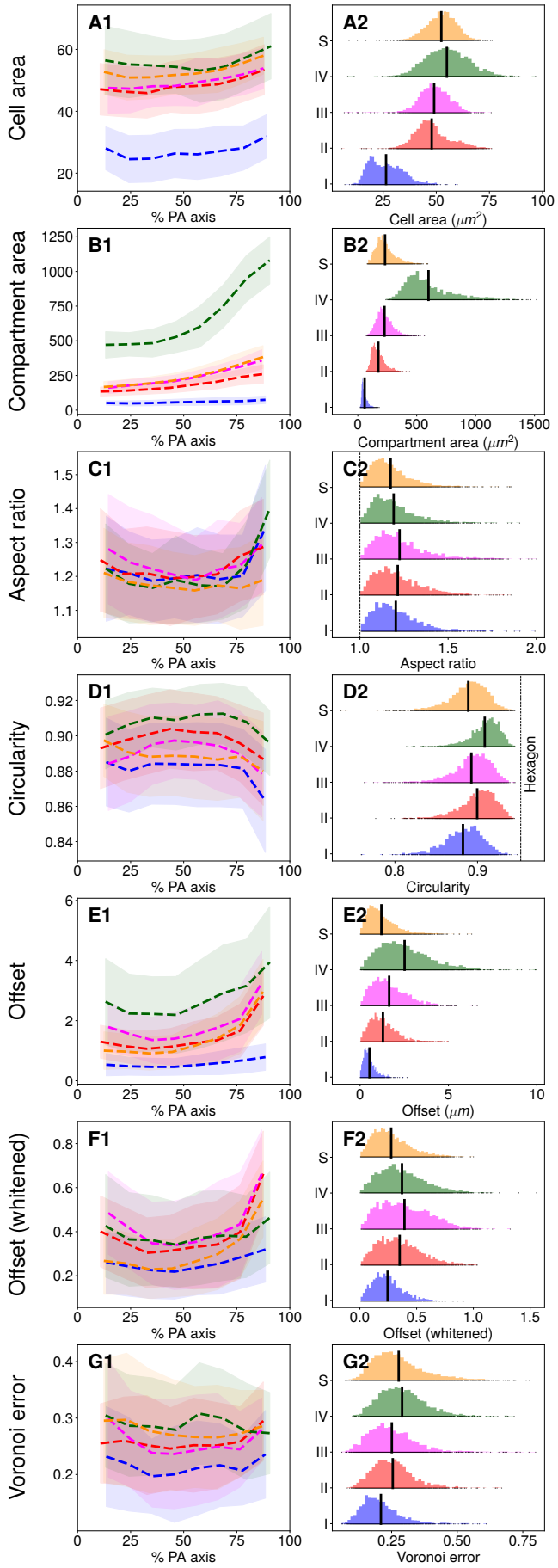


**Fig. 8.** PA axis and life cycle variation in top view. (A-G) Column 1: Computed metrics binned in 8 equally spaced segments along the PA axis. Means are shown as dashed lines with per-bin standard deviation reported by shaded segments. Colors correspond to developmental stages defined in Fig. 3. Column 2: Histograms of metrics by stage in 100 equally spaced bins, by stage, with empirical means indicated by vertical black bars. Units are noted in parentheses, and otherwise, are dimensionless. See SI Appendix, Table S2 for numerical values of empirical means. Fig. S15 and Table S5, SI Appendix show analogous data for distortion-corrected surface CZ3 geometry accounting for the mounting procedure (Materials and Methods). We find a maximum 5% increase in mean compartment area across stages I-IV, with minimal change in other CZ3 features.

the $\sim 39\%$ increase from stage I to IV in error with respect to the Voronoi tessellation (Fig. 8G2 and SI Appendix, Table S2), whose partitions are always convex polygons.

***E.3. CZ3 compartment areas enlarge anisotropically.*** While the compartment areas become more circular as they expand, the aspect ratio is independent of stage and thus of organism size. The apparent increase at extremes of the PA axis (U-shaped curves) is almost entirely accounted for by distortion correction (Fig. S15 and Table S5, SI Appendix), which still leaves the distributions (Fig. 8C2) unchanged. Fig. 8C shows that aspect ratio distributions are not only stable in mean, with less than 5% variation, but also in skewness and variance; they are gamma-distributed throughout growth with stable distribution parameters (SI Appendix, Fig. S9). Together, the stability of aspect ratio and increasing compartment area circularity during growth suggests a transition from tightly packed, polygonal compartment areas (where neighboring boundaries are closely aligned) to elliptical configurations in which neighboring boundaries are no longer in full contact. We term this process *acircular relaxation*.

To study how ECM is distributed around the somatic cells, we quantified the cellular offset during the life cycle. The absolute offset from the compartment area center of mass (Fig. 8E1-2) increases from stages I to IV, and along the PA axis, indicating a strong correlation between larger compartment areas and cellular displacements (as quantified in §F). Perhaps counterintuitively, the cellular offset vector shows no correlation with the primary elongation axis of the compartment area; their relative angle is uniformly distributed in $[0, \frac{\pi}{2}]$ in I-IV and S (SI Appendix, Fig. S10). In contrast, the whitened offset (which accounts for compartment area and anisotropy) is nearly constant in mean after from II-IV (black vertical lines in Fig. 8F). The support of the distribution does increase, albeit at a smaller rate than that of the cellular offset. Throughout this analysis of variation along the PA axis (Fig. 8), similarly sized spheroids in the asexual and sexual life cycle stages, bearing embryos or egg cells respectively, resemble each other in ECM geometry.

**F. CZ3 geometry shows feature correlations.** The analysis above indicates compelling correlations between geometric features of the CZ3 architecture during growth. Here we analyze these in more detail with pooled data from all spheroids presented in Fig. 8. Figure 9A shows an exponential increase in the compartment area $a_{\mathrm{cz3}}$ with cell area $a_{\mathrm{cell}}$ through stage III, saturating at stage IV. This quantifies the prior observation that somatic cells primarily grow before stage II, in contrast to compartment areas, which primarily grow after stage III. As expected from the PA analysis, the aspect ratio (B) is decoupled from compartment area

JAIP

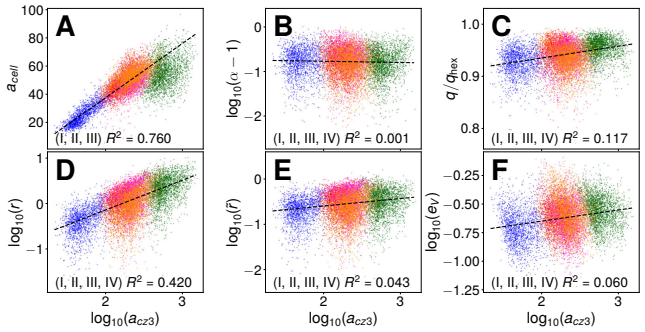


**Fig. 9.** Pair correlations of compartment features in top view. At stages I-IV (blue, red, magenta, green) and S (orange), plots show correlations between compartment area ($a_{cz3}$) and other metrics defined in Table 1. Coordinate transforms are chosen in either linear- or log-scale, with natural offsets (e.g. $q_{\rm hex}$ in panel (C), the circularity of a regular hexagon), to produce approximately equally-sized contours across life cycle stages. $R^2$ is the linear regression correlation coefficient for stages listed. Similarly as in Fig. 8, geometric distortion-corrected versions of A, B, and C may be found in SI Appendix, Fig. S16. We find that the feature correlations are essentially unchanged, with a slight improvement in cell/compartment area correlation (A).

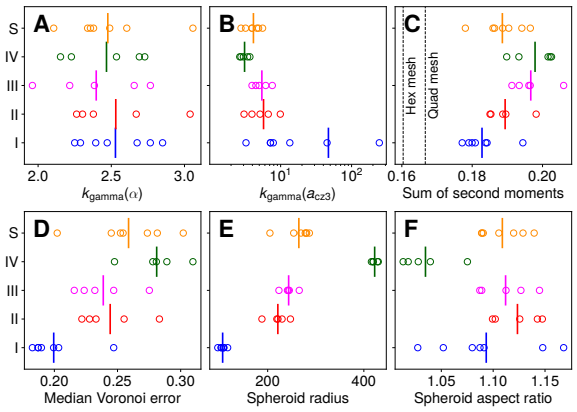


**Fig. 10. Global geometric properties of spheroids in stages I-IV, S in top view.** (A) shape parameter $k$ of gamma distribution fit to $\alpha$, (B) the same for $a_{cz3}$, (C) sum of second moments Eq. (3), (D) median $e_V$ over the whole spheroid, (E) the circular radius of each spheroid (geometric mean of major and minor axes), and (F) the aspect ratio of each spheroid (ratio of major and minor axes).

size, while the circularity (C) increases. This reinforces the conclusion that as compartment areas expand they preserve their aspect ratio while decreasing in polygonality. The cellular offset (D) reveals a power-law relationship with compartment area, which, in conjunction with the weak coupling between whitened offset and compartment area size (E), further supports that conclusion in light of a scaling argument we explain in Discussion §5.

**G. Tessellation properties change during the life cycle.** The metrics in Fig. 10 reveal clear trends by life cycle stage for the global geometry of each spheroid in top view. Panel E shows the increasing circular radius, with most of the increase occurring between stages I-II and III-IV. II-III is separated by fewer hours and occurs during the first dark phase. Stage S is sorted in size close to stage III, supporting its resemblance with stages II and III in the PA analysis.

At fixed mean, the shape parameter $k_{\rm gamma}$ of the gamma distribution, Eq. (2), is a measure of the entropy of the configuration, with high $k$ indicating an increasingly crystalline, Gaussian-distributed configuration by the central limit theorem (48). Fig. 10A confirms the stability of the aspect ratio distribution between stages, exhibiting values of $k_{\rm gamma}$ between $\sim 2-3$, similar to ranges previously reported for confluent tissues (44). Simultaneously, panel B shows that $k_{\rm gamma}$ in the distribution of $a_{\rm cz3}$ is decreasing from stages I-IV, so the configuration (primarily the anterior hemisphere, SI Appendix, Fig. S11) becomes increasingly disordered. The initial high values of $k_{\rm gamma}$ are consistent with the earlier observation that CZ3 compartments begin in a tightly-packed configuration, and as $k$ quantifies regularity we infer that both tight packing and proximity to an equal-area lattice describes the initial configuration. The values of $k_{\rm gamma}$ between 2 and 3 in Stage IV are close to those for the Voronoi tessellations (41), which is remarkable given the decreasing polygonality of the CZ3 compartments observed in §E.2.

The standardized sum of second moments, defined as

$$n \sum_{i=1}^{n} \mathrm{Tr}(\mathsf{M}_2^{(i)}) / \left( \sum_{i=1}^{n} a_{\rm cz3}^{(i)} \right)^2 , \qquad [3]$$

with $i = 1 \ldots n$ indexing CZ3 compartments per spheroid, is a space-partitioning cost that is minimized by equilateral hexagonal meshes (e.g. the surface of a honeycomb, see SI Appendix, §2.C.2). Although CZ3 compartments do not tile space due to extracompartmental spaces, Eq. (3) can nevertheless be computed for the covered area. This energetic cost for each spheroid, representing deviations of the CZ3 architecture from optimal space-partitioning, is displayed in panel C. We find that the somatic CZ3 architecture becomes *decreasingly* optimal during expansion—an observation consistent with the underlying increasing trends in cell offset and area polydispersity as quantified in panel B, §E.1, and §E.3. The metrics in B-C thus show the counterintuitive result that the CZ3 space partitioning is increasingly disordered as the global sphericity is maintained (and even improved, panel F) during the dramatic enlargement.

**H. Pherophorin II is also localized in the boundary zone.** Confocal cross sections reveal that PhII is also part of the boundary zone (BZ), the outermost ECM layer of the organism (Figs. 1B and 4C and D). The PhII:YFP-stained BZ extends as a thin $\sim 1.1 \pm 0.4\,\mu$m arcing layer from the flagella exit points of one cell to those of all neighboring cells. This shape indicates that the outer surface of the spheroid has small indentations at the locations of somatic cells, where flagella penetrate the ECM. At these points, the BZ is connected to the CZ3 of the somatic cells below. Because the BZ is thin and not flat, it is not visible in a top, cross-sectional view of a spheroid through the centers of the somatic ECM compartments (e.g. Fig. 5), and only partly visible when the focal plane cuts through the BZ. If the focal plane is placed on the deepest point of the indentations, only the areas at which the BZ is connected to CZ3 can be seen (Fig. 11). From the centers of these areas the two flagella emerge and the flagellar tunnels are seen as two black dots due to the lack of fluorescence there (Fig. 11B). A closer look at the fine structure at the BZ-CZ3 connection site shows that fiber-like structures radiate from there to the BZ-CZ3 connection sites of neighboring somatic cells.

DRAFT

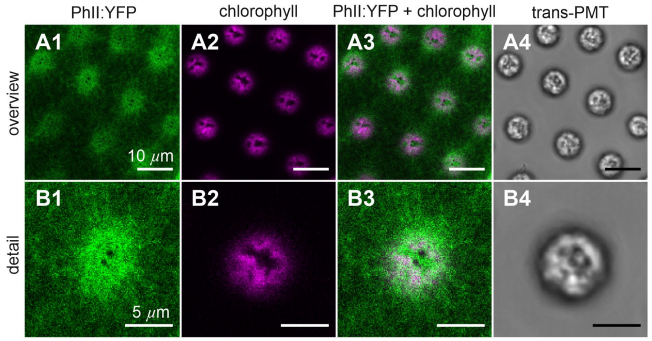


**Fig. 11.** Magnified top view in regions where the BZ is connected to CZ3 in early stage II. Imaging as in Fig. 4. Fiber-like structures radiate from these regions. Flagellar tunnels are seen in the centers of those areas as two dark dots.

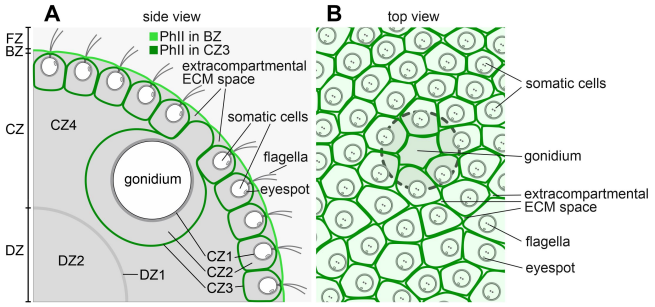


**Fig. 12.** Overview of PhII:YFP localization in the ECM in early stage II. (A) Schematic cross section, showing localization in CZ3 of both somatic cells and gonidia (dark green) and BZ (light green). (B) Schematic top view, looking through the boundary zone, showing PhII:YFP in the CZ3 and the existence of extracompartmental ECM spaces. Position of a gonidium below the somatic cell layer is indicated (dashed); the CZ3 compartment of the deeper gonidium extends to the surface where it is surrounded by ECM compartments with somatic cells.

## Discussion and Conclusions

**1. Holistic view of PhII localization.** Synthesizing the results of preceding sections, we arrive at the summary shown in Fig. 12 of the identified locations of PhII. It forms compartment boundaries (CZ3) around each somatic cell and each gonidium, and is also found in the outer border (BZ) of the spheroid; CZ3 and BZ are connected where the flagella emerge. While each compartment boundary can be assigned to the cell it encloses, and is most likely synthesized solely by that cell, PhII in the BZ is evidently formed collectively by neighboring cells. Since the compartment boundaries of somatic cells are not completely adjacent to those of neighbors, and the BZ does not rest directly on the compartment boundaries, extracompartmental ECM space remains between the CZ3 enclosures as well as between them and the BZ. The extracompartmental ECM space thus appears to be a net-like coherent space connected to CZ4.

**2. Relation to earlier ECM studies by electron microscopy.** In earlier transmission electron microscopy images showing heavy metal-stained sections of the ECM, both the CZ3 and the BZ can be recognized as relatively dark structures, whereas the CZ2, CZ4 and the deep zone are very bright (20). As the degree of darkness reflects the electron density and atomic mass variations in the sample, PhII evidently forms firmer wall-like structures in CZ3 and BZ, while CZ2, CZ4 and the deep zone have a very low density and are presumably of gel-like consistency. Using quick-freeze/deep-etch electron microscopy, it was shown that the fine structure of the ECM of volvocine algae such as *Chlamydomonas* and *Volvox* resembles a three-dimensional network (49, 50). While both the CZ3 and BZ are likely dense networks with a fine pore size to the mesh, they must nevertheless allow the passage of small molecules and non-crosslinked ECM building blocks exported by cells, as evidenced by the growth during development of these compartments, the extracompartmental space, and the deep zone. Cells must also be able to absorb nutrients from the outside, which must pass through both the BZ and CZ3. The BZ may be a denser network than the CZ3, in order to prevent ECM building blocks from escaping into the environment.

**3. Mechanical implications of identified ECM structures.** As revealed by the localization of PhII:YFP and prior electron microscope studies, the BZ appears to form a dense "skin" on the outer surface of the spheroid to which the CZ3 compartments are firmly attached at the flagella exit points. This point-like attachment allows the compartments to expand in all directions during growth. As the CZ3 compartments are densely packed and attached to the BZ, they would be expected to provide rigidity as a kind of "exoskeleton" of the alga. A simple experiment shows this feature: if many *Volvox* are pressed together inside the suspending liquid medium and then released, they elastically repel each other. And while it is clear that the constant expansion of the compartments by incorporating further ECM components allows the ECM to enlarge considerably, the precise mechanism for transporting the ECM components to their destinations remains an open question.

**4. Evolution of the volvocine ECM and convergence of a monolayer epithelium-like architecture.** The ECM of *V. carteri* evolved from the (cellulose-free) cell wall of a *Chlamydomonas*-like, unicellular ancestor (7). The cell wall of extant *Chlamydomonas* species consists of an outer "tripartite" layer with a highly regular, quasi-crystalline structure and an inner more amorphous layer. In few-celled volvocine genera with a low degree of developmental complexity, such as *Pandorina*, the tripartite layer is partially split, so that its outer leaflet is continuous across the surface of the organism, while its inner leaflet still surrounds each individual cell body (51). In larger, more complex volvocine algae (*Eudorina* to *Volvox*), the entire tripartite layer is continuous over the surface of the organism. In the genus *Volvox*, the outer layer has developed even further and the tripartite layer has become part of the boundary layer, the BZ, while the inner layer has evolved species-specific (CZ3) compartments (20). The architecture of these CZ3 compartments shows certain parallels with the epidermis of most plant leaves (52) and even epithelia in animals (38), all of which possess an (initially) closely packed, polygonal architecture. Looking at the plasma membranes in animal epithelia, cellulose-based cell walls in epidermal cells of land plants and cellulose-free CZ3 structures of somatic *Volvox* cells as touching compartment boundaries, their packing geometry can be described using the same physical concepts (53). Interestingly, they also share the presence of an adjacent thin ECM layer, which represents a kind of boundary in all of them: the cuticle (secreted by

DRAFT

plant epidermis cells), the basal lamina (secreted by animal epithelial cells) and the BZ (secreted by somatic *Volvox* cells). The shared geometrical solution likely represents an example of convergent evolution driven by the common pressure to evolve a monolayer epithelium-like architecture with protective and control functions.

**5. Characteristics of the CZ3 stochastic geometry.** Examination of the stochastic geometry of the space partition formed by CZ3 walls, along with somatic cell positions, reveals four key findings. First, somatic cell growth occurs mainly between stages I-II, whereas CZ3 compartment areas in top view grow mostly during III-IV (Fig. 8). The surface areas of the compartments are well-approximated by gamma distributions with anterior/posterior hemispheres exhibiting different values of $k$ (SI Appendix, Fig. S8). Such area distributions arise in granular and cellular materials (42, 54); it is novel to observe them in intrinsic structures of an ECM, with non-stationarity in $k$ revealing A/P differentiation.

Second, the aspect ratio distributions of compartments in top view are remarkably stable throughout the life cycle (Figs. 8 and 9), and are well-approximated by gamma distributions with stable $k$ (SI Appendix, Fig. S9). Maintaining a fixed aspect ratio $\alpha \approx 1.2$ requires that each compartment area enlarges anisotropically, in strong contrast with the trend in aspect ratio of the overall spheroid, which is both lower and decreases from stages III-IV to less than 1.05 (Fig. 10F). This lack of local-global coupling suggests that compartment area anisotropy could be set by geometric constraints in the cellular configuration prior to stage I and may explain how the organism maintains (and increases) its sphericity despite the strong non-uniformity in size and shape of its compartments. Shape variability in the form of gamma-distributed aspect ratios arises in a large class of epithelial tissues and inert jammed matter (44) which exhibit deviations from optimal space-partitioning in the sense of Eq. (3). The epithelium-like architecture formed by the CZ3 robustly falls into this class.

Third, the somatic cell offset from the compartment centroid increases steadily with compartment area (Fig. 8D) while the whitened offset remains relatively constant (Fig. 8E). Both observations suggest a growth-induced deformation in which space is locally dilated (in tangent planes containing compartment surfaces) as $\mathbb{R}^2 \mapsto \rho\mathbb{R}^2, \rho \geq 1$, which indeed increases offsets while preserving whitened offsets (SI Appendix, §B.1). Such transformations also preserve aspect ratios, consistent with earlier observations. The cellular offset angle with respect to the principal stretch axis, on the other hand, is a priori unconstrained by these observations, and we find it is uniformly distributed in $[0, \pi/2]$ (SI Appendix, Fig. S10). This highlights a stochastic decoupling between cell positioning and compartment shape, much like that previously observed between compartment and spheroid principal stretches. Likewise, this may be established before stage I and later scaled by compartment growth—analogous to two points diverging on the surface of an expanding bubble. Importantly, *global* dilations of space, likening *Volvox* itself to a bubble, cannot produce the increasing compartment area polydispersity observed during the life cycle, while *local* dilations as above, likening the CZ3 'epithelium' to a spherical raft of heterogeneously inflating bubbles, can. In a continuum limit, such growth-induced local dilations may be represented by conformal maps (55).

**6. The CZ3 structural transition.** As best seen in Fig. 5, the CZ3 compartment mesh undergoes a structural transition (in top view) from a space *partition* to a *packing* whose surface covering fraction decreases significantly while remaining a connected structure. ECM-filled extracompartmental spaces expand during this transition, accompanying an apparent relaxation of the compartments from polygonal to elliptical shapes. This is reminiscent of the wet-dry transition in two-dimensional foams, wherein initially polygonal bubbles meeting at Plateau borders relax under surface tension to circular shapes as the liquid fraction (here, extracompartmental space) increases. Convergence to spherical bubbles and loss of contacts occurs at the *jamming transition*, which for 2D foams occurs at gas fraction $\phi_c \approx 0.84$ (56). While a systematic investigation of the average contact count and covering fraction $\phi$ of compartments requires comprehensive extraction of features from full 3D data (accounting for defects induced by offspring, Fig. 4B1), our preliminary results from top view (Fig. 5) indicate that the covering fraction can decrease to as little as $\phi \approx 0.81$ in stage IV. Viewing the surface CZ3 geometry as a 2D foam, then, it is perhaps surprising that the aspect ratio distribution remains so stable (Fig. 8C) with mean $\alpha \approx 1.2$ given that $\phi$ is near the jamming transition. On the other hand, a foam model for the compartment geometry may appear reasonable given the dramatic shape changes during development, but foams are vanishing-adhesion limits of more general plausible models (57) used previously for cell monolayers. There, cell-cell adhesion produces aspherical equilibria even at low $\phi$; indeed, here, inter-compartmental adhesion is likely stronger than that of foam bubbles due to crosslinking between the walls. Cross sectional views (Fig. S13, SI Appendix) confirm that walls maintain strong radial contact throughout expansion.

More broadly, while structural transitions during development (including not only jamming (44, 57, 58) but also density-[$\phi$-] independent rigidity transitions (59)) have been widely studied in epithelia, we observe for the first time one in the extracellular matrix of *Volvox*. This perhaps paves way for yet another usage of *Volvox* as a model multicellular organism: for the development of epithelium-like structures and transitions in ECMs and bio-inspired materials. These are necessarily self-assembled structures, dependent on material properties like tension, adhesion, and permeability that are not under direct control by cells. This highlights more broadly the need to probe the rheology of the ECM—perhaps using microrheological techniques like those applied in cytoskeletal studies (60)—to understand how stochastic local interactions give way to robust global structures. A better understanding of the self-organization of crosslinking ECM components will shed light on the principles underlying the intricate geometry and stochasticity observed in multicellular ECMs.

**7. The shape of ECM compartment boundaries along the radial axis - an outlook.** In this study, we analyzed the surface features of the CZ3 in top view, which provides insight into the ECM packing geometry at the spheroid's expanding surface. We also provide preliminary data on cross sectional shapes (Fig. S13, SI Appendix), with the caveat that "floors" exhibit diffuse or poor YFP signal. From this, we estimate compartments' volumetric growth and find that they indeed increase from stages I-IV, though

DRAFT

nonlinearly, showing a sigmoidal growth (Table S4 and Fig. S14, SI Appendix). Moreover, the shape itself shows notable dynamics: while appearing roughly self-similar from stages I-III, it abruptly reverses its cross-sectional aspect ratio from III-IV from being slightly radially elongated to significantly tangentially elongated (Fig. S13, SI Appendix). This could indicate a transition in the growth dynamics at III-IV in which the boundary and deep zones are expanding faster than the cellular zone, with the latter building residual stress as a result. This picture of increasingly pressure-driven growth of the CZ (in which it is pulled apart laterally like the elastic surface of a balloon) is also weakly supported by observations of the spheroid's eccentricity (Fig. 10F) sharply decreasing towards sphericity from III-IV. The full 3D dynamic geometry of the ECM is the subject of future study.

## Materials and Methods

**Strains and culture conditions.** Female wild-type strains of *Volvox carteri* f. *nagariensis* were Eve10 and HK10. Eve10 is a descendant of HK10 and the male 69-1b, which originate from Japan. The strains have been described previously (61–64). Strain HK10 has been used as a donor for the genomic library. As a recipient strain for transformation experiments a non-revertible nitrate reductase-deficient (*nitA*$^-$) descendant of Eve10, strain TNit-1013 (65), was used. As the recipient strain is unable to use nitrate as a nitrogen source, it was grown in standard *Volvox* medium (66) supplemented with 1 mM ammonium chloride ($NH_4Cl$). Transformants with a complemented nitrate reductase gene were grown in standard *Volvox* medium without ammonium chloride. Cultures were grown at 28°C in a cycle of 8 h dark/16 h cool fluorescent white light (67) at an average of ∼100 $\mu$mol photons $m^{-2}$ $s^{-1}$ photosynthetically active radiation in glass tubes with caps that allow for gas exchange or Fernbach flasks aerated with ∼ 50 $cm^3$/min of sterile air.

**Vector construction.** The genomic library of *V. carteri* strain HK10 in the replacement lambda phage vector λEMBL3 (68) described by (27) has been used before to obtain a lambda phage, λ16/1, with a 22 kb genomic fragment containing three copies of the *phiI* gene (28). A subcloned 8.3 kb *Bam*HI-*Eco*RI fragment of this lambda phage contains the middle copy, the *phiI* gene B, used here. The 8.3 kb fragment also includes the *phiI* promoter region, 5'UTR and 3'UTR and is in the pUC18 vector. An artificial *Kpn*I side should be inserted directly upstream of the stop codon so that the cDNA of the *yfp* can be inserted there. This was done by cutting out a 0.5 kb subfragment from a unique *Mlu*I located 0.2 kb upstream of the stop codon to a unique *Cla*I located 0.3 kb downstream of the stop codon from the 8.3 kb fragment, inserting the artificial *Kpn*I with PCRs, and putting the *Mlu*I-*Cla*I subfragment back to the corresponding position. The primers 5'GTAACTAACGAATGTACGGC (upstream of *Mlu*I) and 5'*ATCGAT*TCAGGTACCTGGCCCCGTGCGGTAGATG were used for the first PCR and the primers 5'GGTACC**TGA**TTGCCGTAAGAGCAGTCATG and 5'TCTAGCCTCGTAACTGTTCG (downstream of *Cla*I) for the second PCR (The *Kpn*I side is underlined, the stop codon is shown in bold). One primer contains a *Cla*I (italics) at its 5'end to facilitate cloning. PCR was also utilized to add *Kpn*I sides to both ends of the *yfp* cDNA. In addition, a 15 bp linker sequence, which codes for a flexible pentaglycine interpeptide bridge, should be inserted before the *yfp* cDNA. The *yfp* sequence was previously codon-adapted to *C. reinhardtii* (69) but also works well in *V. carteri* (11). Since this *yfp* sequence was already provided with the linker sequence earlier (70), the primers 5'GGTACC*GGCGGAGGCGGTGGC*ATGAGC and 5'GGTACCCTTGTACAGCTCGTC and a corresponding template could be used (the *Kpn*I side is underlined, the 15 bp linker is shown in italics). The resulting 0.7 kb PCR fragment was digested with *Kpn*I and inserted into the artificially introduced *Kpn*I side of the above pUC18 vector with the 8.3 kb fragment. All PCRs were carried out as previously described (71–73) using a gradient PCR thermal cycler (Mastercycler Gradient; Eppendorf). The final vector pPHII-YFP (Fig. 2) was checked by sequencing.

**Nuclear transformation of *V. carteri* by particle bombardment.** Stable nuclear transformation of *V. carteri* strain TNit-1013 was performed as described earlier (74) using a Biolistic PDS-1000/He (Bio-Rad) particle gun (75). Gold microprojectiles (1.0 $\mu$m dia., Bio-Rad, Hercules, CA, USA) were coated according to earlier protocols (71, 72). Algae of the recipient strain where co-bombarded with the selection plasmid pVcNR15 (76), carrying the *V. carteri* nitrate reductase gene, and the non-selectable plasmid pPhII-YFP. Plasmid pVcNR15 is able to complement the nitrate reductase deficiency of the recipient strain and therefore allows for selection of transformants. For selection, the nitrogen source of the *Volvox* medium was switched from ammonium to nitrate and algae were then incubated under standard conditions in 9 cm. diameter petri dishes filled with ∼ 35 ml liquid medium. Untransformed algae of the recipient strain die under these conditions due to nitrogen starvation. After incubation for at least six days, the petri dishes were inspected for green and living transformants.

**Confocal laser scanning microscopy.** For live cell imaging of transformed algae, cultures were grown under standard conditions and induced with 10 $\mu$l medium of sexually induced algae in a 10 ml glass tube. Approximately 10 spheroids in culture medium were placed on a glass slide and covered with a glass coverslip. The coverslip was mounted on grease on all four edges in order to apply exactly the right force to hold the spheroids in place but minimize deformation. Intact spheroids with an average number of gonidia, embryos or daughter spheroids were selected for imaging. From all acquired images with stable spheroid positions, optimal focal planes were selected for segmentation and downstream analyses. An LSM780 confocal laser scanning microscope was used with a 63× LCI Plan-Neofluar objective and a 10× Plan-Apochromat (Carl Zeiss GmbH, Oberkochen, Germany). The pinhole diameter of the confocal was set to 1 Airy unit. Fluorescence of the PhII:YFP fusion protein was excited by an $Ar^+$ laser at 514 nm and detected at 520-550 nm. The fluorescence of chlorophyll was detected at 650-700 nm. Fluorescence intensity was recorded in bidirectional scan mode for YFP and chlorophyll in two channels simultaneously. Transmission images were obtained in a third channel by using a transmission-photomultiplier tube detector (trans-PMT). Images were captured at 12 bits per pixel and analyzed using ZEN black 2.1 digital imaging software (ZEN 2011, Carl Zeiss GmbH). Image processing and analysis used Fiji (ImageJ 1.51w) (77). To verify the signal as YFP fluorescence, spectral analysis was performed. The lambda scan function of ZEN was used in which a spectrum of the emitted light is generated by a gallium arsenide phosphide QUASAR photomultiplier detector that produces simultaneous 18-channel readouts. Emission spectra between 486 and 637 nm were recorded for each pixel with a spectral resolution of 9 nm using a 458/514 beam splitter and 488-nm laser light for excitation.

**Data, Materials, and Software Availability.** All data and code are available on Zenodo (DOI: 10.5281/zenodo.14066435) (78).

**ACKNOWLEDGMENTS.** We are grateful to Kordula Puls, Diana Thomas-McEwen and members of LiMiTec for technical assistance, and to Jane Chui and Kyriacos Leptos for inspiring conversations. Financial support for the work carried out in Bielefeld was provided by A.H.'s institutional funds. REG gratefully acknowledges the financial support of the John Templeton Foundation (#62220). This work was also supported in part by the Cambridge Trust (AS), and Wellcome Trust Investigator Grants 207510/Z/17/Z (SSMHH and REG) and 307079/Z/23/Z (SKB, SSMHH and REG).

1. M Abedin, N King, Diverse evolutionary paths to cell adhesion. *Trends Cell Biol.* **20**, 734–742 (2010).
2. L Stavolone, V Lionetti, Extracellular matrix in plants and animals: Hooks and locks for viruses. *Front. Microbiol.* **8**, 1760 (2017).
3. B Kloareg, Y Badis, JM Cock, G Michel, Role and evolution of the extracellular matrix in the acquisition of complex multicellularity in eukaryotes: a macroalgal perspective. *Genes* **12**, 1059 (2021).
4. DS Domozych, JG LoRicco, The extracellular matrix of green algae. *Plant Physiol.* **194**, 15–32 (2024).
5. A Hallmann, Evolution of reproductive development in the volvocine algae. *Sex Plant Reprod.* **24**, 97–112 (2011).

6. DL Kirk, *Volvox: molecular-genetic origins of multicellularity and cellular differentiation*, Developmental and Cell Biology series. (Cambridge University Press, Cambridge), 1 edition, (1998).
7. DL Kirk, A twelve-step program for evolving multicellularity and a division of labor. *BioEssays* **27**, 299–310 (2005).
8. I Nishii, SM Miller, *Volvox*: simple steps to developmental complexity? *Curr. Opin. Plant Biol.* **13**, 646–653 (2010).
9. MD Herron, Origins of multicellular complexity: *Volvox* and the volvocine algae. *Mol. Ecol.* **25**, 1213–23 (2016).
10. JG Umen, *Volvox* and volvocine green algae. *Evodevo* **11**, 13 (2020).
11. B von der Heyde, A Hallmann, Targeted migration of pherophorin-s indicates extensive extracellular matrix dynamics in *Volvox carteri*. *Plant J.* **103**, 2301–17 (2020).
12. EL von der Heyde, A Hallmann, Molecular and cellular dynamics of early embryonic cell divisions in *Volvox carteri*. *Plant Cell* **34**, 1326–1353 (2022).
13. B von der Heyde, A Hallmann, Cell type-specific pherophorins of *Volvox carteri* reveal interplay of both cell types in ecm biosynthesis. *Cells* **12**, 134 (2023).
14. A Hallmann, Extracellular matrix and sex-inducing pheromone in *Volvox*. *Int. Rev. Cytol.* **227**, 131–82 (2003).
15. A Hallmann, K Godl, S Wenzl, M Sumper, The highly efficient sex-inducing pheromone system of *Volvox*. *Trends Microbiol.* **6**, 185–9 (1998).
16. DL Kirk, MM Kirk, Heat shock elicits production of sexual inducer in *Volvox*. *Science* **231**, 51–4 (1986).
17. M Sumper, E Berg, S Wenzl, K Godl, How a sex pheromone might act at a concentration below 10(-16) m. *EMBO J.* **12**, 831–6 (1993).
18. DL Kirk, Germ-soma differentiation in *Volvox*. *Dev. Biol.* **238**, 213–23 (2001).
19. R Schmitt, Differentiation of germinal and somatic cells in *Volvox carteri*. *Curr. Opin. Microbiol.* **6**, 608–13 (2003).
20. DL Kirk, R Birchem, N King, The extracellular matrix of *Volvox*: a comparative study and proposed system of nomenclature. *J. Cell. Sci.* **80**, 207–31 (1986).
21. M Sumper, A Hallmann, Biochemistry of the extracellular matrix of *Volvox*. *Int. Rev. Cytol.* **180**, 51–85 (1998).
22. A Hallmann, The pherophorins: common, versatile building blocks in the evolution of extracellular matrix architecture in volvocales. *Plant J.* **45**, 292–307 (2006).
23. DH Miller, DTA Lamport, M Miller, Hydroxyproline heterooligosaccharides in *Chlamydomonas*. *Science* **176**, 918–20 (1972).
24. AM Showalter, D Basu, Extensin and arabinogalactan-protein biosynthesis: glycosyltransferases, research challenges, and biosensors. *Front. Plant Sci.* **7**, 814 (2016).
25. J Sommer-Knudsen, A Bacic, AE Clarke, Hydroxyproline-rich plant glycoproteins. *Phytochemistry* **47**, 483–497 (1998).
26. F Ender, K Godl, S Wenzl, M Sumper, Evidence for autocatalytic cross-linking of hydroxyproline-rich glycoproteins during extracellular matrix assembly in *Volvox*. *Plant Cell* **14**, 1147–60. (2002).
27. H Ertl, R Mengele, S Wenzl, J Engel, M Sumper, The extracellular matrix of *Volvox carteri*: molecular structure of the cellular compartment. *J. Cell. Biol.* **109**, 3493–501 (1989).
28. K Godl, A Hallmann, A Rappel, M Sumper, Pherophorins: a family of extracellular matrix glycoproteins from *Volvox* structurally related to the sex-inducing pheromone. *Planta* **196**, 781–7 (1995).
29. K Godl, A Hallmann, S Wenzl, M Sumper, Differential targeting of closely related ecm glycoproteins: the pherophorin family from *Volvox*. *EMBO J.* **16**, 25–34 (1997).
30. SE Prochnik, et al., Genomic analysis of organismal complexity in the multicellular green alga *Volvox carteri*. *Science* **329**, 223–6 (2010).
31. B Klein, D Wibberg, A Hallmann, Whole transcriptome rna-seq analysis reveals extensive cell type-specific compartmentalization in *Volvox carteri*. *BMC Biol.* **15**, 111 (2017).
32. S Wenzl, D Thym, M Sumper, Development-dependent modification of the extracellular matrix by a sulphated glycoprotein in *Volvox carteri*. *EMBO J.* **3**, 739–744 (1984).
33. S Wenzl, M Sumper, Sulfation of a cell surface glycoprotein correlates with the developmental program during embryogenesis of *Volvox carteri*. *Proc. Natl. Acad. Sci. U.S.A.* **78**, 3716–20 (1981).
34. S Wenzl, M Sumper, The occurrence of different sulphated cell surface glycoproteins correlates with defined developmental events in *Volvox*. *FEBS Lett.* **143**, 311–5 (1982).
35. S Wenzl, M Sumper, Early event of sexual induction in *Volvox*: chemical modification of the extracellular matrix. *Dev. Biol.* **115**, 119–128 (1986).
36. S Wenzl, M Sumper, *Pheromone-inducible glycoproteins of the extracellular matrix of Volvox and their possible role in sexual induction*, eds. W Wiessner, DG Robinson, RC Starr. (Springer-Verlag, Berlin), pp. 58–65 (1987).
37. M Sumper, J Nink, S Wenzl, Self-assembly and cross-linking of *Volvox* extracellular matrix glycoproteins are specifically inhibited by ellman's reagent. *Eur. J. Biochem.* **267**, 2334–2339 (2000).
38. M Dicko, et al., Geometry can provide long-range mechanical guidance for embryogenesis. *PLoS Comput. Biol.* **13**, e1005443 (2017).
39. DW Thompson, *On Growth and Form*, Canto ed. JT Bonner. (Cambridge University Press, Cambridge), (1992).
40. F Graner, D Riveline, 'The Forms of Tissues, or Cell-aggregates': D'Arcy Thompson's influence and its limits. *Development* **144**, 4226–4237 (2017).
41. T Day, et al., Cellular organizaion in lab-evolved and extant multicellular species obeys a maximum entropy law. *ELife* **11**, e72707 (2022).
42. T Aste, T Di Matteo, Emergence of gamma distributions in granular materials and packing models. *Phys. Rev. E* **77**, 021309 (2008).
43. M Tanemura, Statistical distributions of poisson voronoi cells in two and three dimensions. *FORMA-TOKYO-* **18**, 221–247 (2003).
44. L Atia, et al., Geometric constraints during epithelial jamming. *Nat. Phys.* **14**, 613–620 (2018).
45. A Srinivasan, S Höhn, RE Goldstein, Point processes and the statistics of cellular neighborhoods in simple multicellular organisms. *ArXiv* p. 2311.11939 (2023).
46. A Hallmann, DL Kirk, The developmentally regulated ecm glycoprotein isg plays an essential role in organizing the ecm and orienting the cells of *Volvox*. *J. Cell Sci.* **113**, 4605–4617 (2000).
47. P Gómez-Gálvez, et al., Scutoids are a geometrical solution to three-dimensional packing of epithelia. *Nat. Commun.* **9**, 2960 (2018) Publisher: Nature Publishing Group.
48. R Durrett, *Probability: Theory and Examples*, Cambridge Series in Statistical and Probabilistic Mathematics. (Cambridge University Press), 5 edition, (2019).
49. UW Goodenough, JE Heuser, Molecular organization of cell-wall crystals from *Chlamydomonas reinhardtii* and *Volvox carteri*. *J. Cell Sci.* **90**, 717–33 (1988).
50. UW Goodenough, B Gebhart, RP Mecham, JE Heuser, Crystals of the *Chlamydomonas reinhardtii* cell wall: polymerization, depolymerization, and purification of glycoprotein monomers. *J. Cell. Biol.* **103**, 405–17 (1986).
51. DL Kirk, Evolution of multicellularity in the volvocine algae. *Curr. Opin. Plant Biol.* **2**, 496–501 (1999).
52. K Esau, *Plant Anatomy*. (LWW) Vol. 75, (1953).
53. SB Lemke, CM Nelson, Dynamic changes in epithelial cell packing during tissue morphogenesis. *Curr. Biol.* **31**, R1098–R1110 (2021).
54. MP Miklius, S Hilgenfeldt, Analytical results for size-topology correlations in 2d disk and cellular packings. *Phys. Rev. Lett.* **108**, 015502 (2012).
55. A Dai, M Ben Amar, Minimizing the elastic energy of growing leaves by conformal mapping. *Phys. Rev. Lett.* **129**, 218101 (2022).
56. J Winkelmann, et al., 2D foams above the jamming transition: Deformation matters. *Colloids Surfaces A: Physicochem. Eng. Aspects* **534**, 52–57 (2017).
57. S Kim, M Pochitaloff, GA Stooke-Vaughan, O Campàs, Embryonic tissues as active foams. *Nat. Phys.* **17**, 859–866 (2021).
58. JA Park, et al., Unjamming and cell shape in the asthmatic airway epithelium. *Nat. Mater.* **14**, 1040–1048 (2015) Publisher: Nature Publishing Group.
59. D Bi, J Lopez, JM Schwarz, ML Manning, A density-independent rigidity transition in biological tissues. *Nat. Phys.* **11**, 1074–1079 (2015).
60. P Cicuta, A Donald, Microrheology: a review of the method and applications. *Soft Matter* **3**, 1449–1455 (2007).
61. A Kianianmomeni, G Nematollahi, A Hallmann, A gender-specific retinoblastoma-related protein in *Volvox carteri* implies a role for the retinoblastoma protein family in sexual development. *Plant Cell* **20**, 2399–419 (2008).
62. RC Starr, Structure, reproduction and differentiation in *Volvox carteri* f. *nagariensis* iyengar, strains hk 9 & 10. *Arch. Protistenkd.* **111**, 204–22 (1969).
63. RC Starr, Control of differentiation in *Volvox*. *Dev. Biol. Suppl.* **4**, 59–100 (1970).
64. CR Adams, et al., Patterns of organellar and nuclear inheritance among progeny of two geographically isolated strains of *Volvox carteri*. *Curr. Genet.* **18**, 141–53 (1990).
65. Y Tian, S Gao, EL von der Heyde, A Hallmann, G Nagel, Two-component cyclase opsins of green algae are atp-dependent and light-inhibited guanylyl cyclases. *BMC Biol.* **16**, 144 (2018).
66. L Provasoli, IJ Pintner, *Artificial media for fresh-water algae: problems and suggestions*, eds. CA Tryon, RT Hartman. (The Pymatuning Symposia in Ecology, Special Publication No. 2., University of Pittsburgh, Pittsburgh, PA), pp. 84–96 (1959).
67. RC Starr, L Jaenicke, Purification and characterization of the hormone initiating sexual morphogenesis in *Volvox carteri* f. *nagariensis* iyengar. *Proc. Natl. Acad. Sci. U.S.A.* **71**, 1050–4 (1974).
68. AM Frischauf, H Lehrach, A Poustka, N Murray, Lambda replacement vectors carrying polylinker sequences. *J. Mol. Biol.,* **170**, 827–42 (1983).
69. KJ Lauersen, O Kruse, JH Mussgnug, Targeted expression of nuclear transgenes in *Chlamydomonas reinhardtii* with a versatile, modular vector toolkit. *Appl. Microbiol. Biotechnol.* **99**, 3491–503 (2015).
70. EL von der Heyde, A Hallmann, Babo1, formerly vop1 and cop1/2, is no eyespot photoreceptor but a basal body protein illuminating cell division in *Volvox carteri*. *Plant. J.* **102**, 276–298 (2020).
71. K Lerche, A Hallmann, Stable nuclear transformation of *Gonium pectorale*. *BMC Biotechnol.* **9**, 64 (2009).
72. K Lerche, A Hallmann, Stable nuclear transformation of *Eudorina elegans*. *BMC Biotechnol.* **13**, 11 (2013).
73. K Lerche, A Hallmann, Stable nuclear transformation of *Pandorina morum*. *BMC Biotechnol.* **14**, 65 (2014).
74. B Schiedlmeier, et al., Nuclear transformation of *Volvox carteri*. *Proc. Natl. Acad. Sci. U.S.A.* **91**, 5080–4 (1994).
75. A Hallmann, S Wodniok, Swapped green algal promoters: aphviii-based gene constructs with *Chlamydomonas* flanking sequences work as dominant selectable markers in *Volvox* and vice versa. *Plant Cell Rep.* **25**, 582–91 (2006).
76. H Gruber, SH Kirzinger, R Schmitt, Expression of the *Volvox* gene encoding nitrate reductase: mutation-dependent activation of cryptic splice sites and intron-enhanced gene expression from a cdna. *Plant Mol. Biol.* **31**, 1–12 (1996).
77. J Schindelin, et al., Fiji: an open-source platform for biological-image analysis. *Nat. Methods* **9**, 676–82 (2012).
78. B von der Heyde, et al., Spatiotemporal distribution of the glycoprotein pherophorin II reveals stochastic geometry of the growing ecm of *Volvox carteri*. doi.org/10.5281/zenodo.14066435 (2024).

PNAS

# Supporting Information for

## Spatiotemporal distribution of the glycoprotein pherophorin II reveals stochastic geometry of the growing ECM of *Volvox carteri*

**B. von der Heyde, A. Srinivasan, S.K. Birwa, E.L. von der Heyde, S.S.M.H. Höhn, R.E. Goldstein, and A. Hallmann**

**Corresponding Authors: Raymond E. Goldstein and Armin Hallmann**
**E-mails: R.E.Goldstein@damtp.cam.ac.uk and armin.hallmann@uni-bielefeld.de**

**This PDF file includes:**

Supporting text
Figs. S1 to S17
Tables S1 to S5
SI References

## Supporting Information Text

### 1. Supplementary data: Pherophorin II overview and DNA sequences

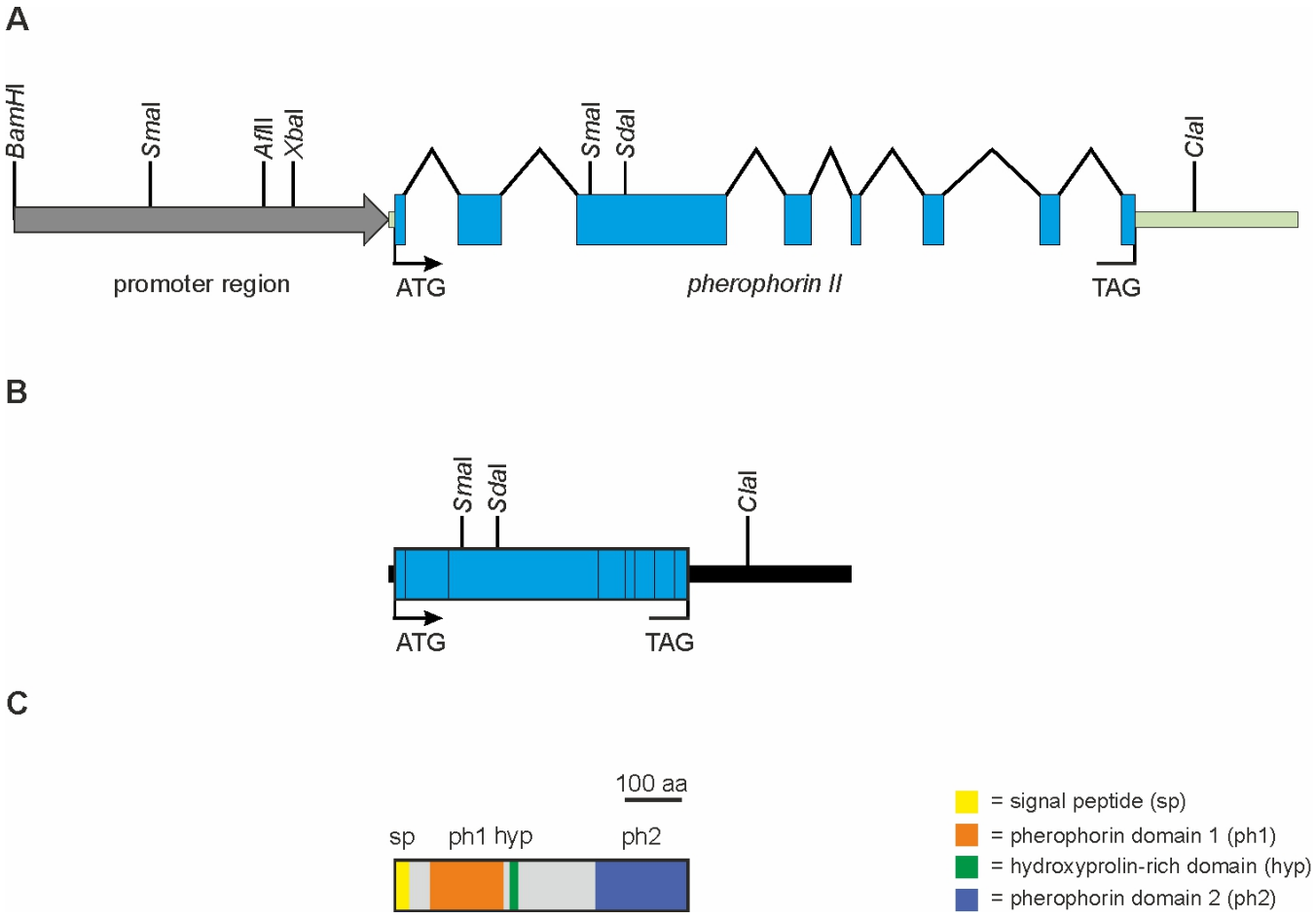


**Fig. S1. Schematic structure of the *phlI* gene, *phlI* mRNA and pherophorin-II protein.** (A) The genomic region schematized here corresponds to the 8329-bp genomic fragment utilized in plasmid pPhlI-YFP. The *phlI* gene (1) is located on scaffold 34 (nucleotides 980223 to 985045) of the *V. carteri* genome version 2.1 (2) in Phytozome v13 (3) on the reverse strand. The start codon is at nucleotide position 985025-985027 on the reverse strand. In the current *Volvox carteri* genome annotation available at Phytozome v13 (Volvox v2.1) pherophorin II is not annotated. Therefore the gene structure was established based on older annotations and confirmed with RNA-Sequencing data (4). The gene structure is indicated as follows: Coding sequences are represented by blue squares, intron sequences by carats, UTRs by green bars and the promoter region by a grey arrow. Start (ATG) and stop (TAG) codons are highlighted. The given restriction sites are also marked in SI Appendix, Fig. S2, which represents the genomic sequence of the *phlI* gene. (B) Structure of the *V. carteri phlI* mRNA. Sequence features are as indicated in A. The coding sequence (blue squares) totals 1557 nucleotides. The 5' UTR is 18 bp in length, while there is a quite long 3' UTR of 869 bp. The complete mRNA is 2,444 nucleotides in length. (C) Structure of the *V. carteri* pherophorin-II protein. The polypeptide comprises 518 amino acids and the calculated molecular weight amounts to 54.5 kDa. As pherophorin-II is an extracellular protein, it possesses a cleaved N-terminal signal peptide (sp) of 24 amino acids. In the mature protein three domains can be identified: an N-terminal pherophorin domain (ph1) with an E-value of 4.5e-27, a hydroxyproline-rich domain (hyp) in the middle (5) and a C-terminal pherophorin domain (ph2) with an E-value of 2.1e-32.The pherophorin domains where identified by blasting the Pfam database (6) using the hmmscan function (7). The short Hydroxyproline-rich domain consists of 78% (hydroxy) prolines (seven of nine amino acids).

GGATCCATGACTGGAAAACCCATCCATGACCAAGTCGTTCCCCGAGAGATAGCAGCTTTGGATCGACT *BamH*I
CGTCGCTTGTGAAAGCCTACCTGGTTACTTAGACGGATTTAGCGACTAGACCACTTATGAAGGTGTTG
ATAAACACCGGGGGTTTCCCTTGGGGTTTTATAGTCGCTTGCAGTCCGTCTGAGGCGCTCGGGATAAT
ATGCCGTACAGTAGTAGATATTCCTGCAGAAGAGAGTCGCACCTCCAGGGAAGAGTTACAATACAGGT
CTGGCCGTGAACGGGCAGACGTGTACAAGATGGTGTACGGAATCGAGTATGTGCACGGATCGTTACAG
CCATGGCCGTAGCCATGTACTGCTGAGTACGGCTGTAGTCGAAAGATGACGAAGATAAGATAGTTTTA
TGTACTGTGTGTACGATATTCGCCTCGACGATGGAATGATAAGAAGGAATGGAGTCCCTACTCGGAAG
TCCTGAGCCAGGCTCTTACACTGGGCACACATTGCCATCCAATGGATGGCCATGTGAATCACAGTGGT
GCCCGGCCTAAGGTAGCTAAGGACTTGGCCGTTATCATCGGTAGATTGCTCACCCACCTTCTCCGACC
AGGAAAACATCCTCGTCACTGGTGCACAGCTCGTCATAGCGATGAATGTTGCAGGGGTGCAAAGGCAA
GTTCCACCACACCATTATGAGCAACCTTAGGACCGCGTTCC*CCCGGG*TAGACACAGGAGGTTTCAGGG *Sma*I
GCTAGATATACTGGTCCAATAAGCAGGGTGCTATTGATGTAGGCAGTAATGTAGCACCCTCCCTCATC
CGACGTACCTCACTCCCTGCACATGTAGAGCCCCTGCGTAGAACGGCTATCTGTGATGAGGGGGAGTC
CCTGGTATAGAGGCCAATGGGAGCATGTGCCGTTGTTGGGCATTCGAGGGGTGCGTGACAGCGTTGTC
ATCGGTAGGGTCGTGCCGAGTTGTGAAGACGGGCTCCTGTAATGGGCCTTATCATGCAGTAGGTTTCT
TGTGGTTTCCGAGGCCGAGGTCATTTTGCCGGGGTGTACGGCATTGTTGCGGGTTGCGGGAAGTTGTA
CAGGTAAATGCATCGGCCATTTTTATCAGCCTTGTTTAAATATAAGTAGACAACCGAGCAAGAACCAA
AGTTTGTCGGAAAGAATTTGGGCAGAGGCCGAAGCTCGCCTGGCAAGCACTGCGCTGAAGGGATAGAG
AGAGACACGAATAAGAAGTGTACAACGGGGCCTCAATAGGCTTTGGGGTCCAGATCGCTGACGCCCTC
GCTCTCAGGCATCAAGTCCCGGACATGGGC*CTTAAG*ACCTGGAGTTGCTGGGCTATTCCCTCTCGCCC *Afl*II
TCCCCGCTCTGACCACTATCCCTCTCTCCAGTTCCCTTTCAGTTCGGGTGGTGCAACGCCAGGCCCCT
CGCCACACACACCGCACCGCACCCACGTTTGTTAAGCGCTAGCAGGCTGTTG*TCTAGA*GCTGTAGAAG *Xba*I
TTCCTAGGTCTATGTTGATGTAGTTTGGTCCCGTTGCTGATGCAACTGGCTTTCGCATCATTAACCGC
CGAGTAGCAGCCGTAGCTACCGGTGCGGAAAGGATGCATGCTCACCAACCGAAATGATGAAGTGCGTG
TGGCGAAGTTATATTCATTATGCCAGGATCAGAGATGTTGGTCAGTAAATAGCAAGTTATATAAAAGG
AATGGTGCTACAGATTACAACGACGGGTCGGATTCCGAAACGTACAACACGAGCAACATCCCAACGGA
TGAAATGCACATTTATATGTGTCCGCGACTAAC*GAATTC*TGCGCTGTGTTTCTAAGATTTATATTCTG *Eco*RI
CATAAAGATATTGATTCCAAAGGACGAAGCGCCTCTGTTGCGGTGGCCAGTCTCGGAAGCGGCTACCT
GCGCCCGCCGTTTACGGTAGTCGTACGGCCCCGCAGCAGCACAGCTGCCGCGGTCGGTCGGGTAGATA
AATAGCGCCACTATGATACAGAACATCATCAAAGTGACATAGCTCTGACAATATGGCATACTTTGCAA start
AGGTAGCTTTCATTTTGGCGACAGCGCTGGCGGTTGGCTGGTGCGTTTTTAGCATGGTGCATTCCCAT
ATGCCTGGGCTTACAATCTGAGGTTGGGTCTATATATTTAAGTATATTTAAGAGCCATATACTTGAAG
GCTTTCAGTCTTTGCCTTTGCGTCGCAAAAGACGTACATGTCCTGTCGCATGCCATAGTGATGAGCCT
TTTCTATAAAGCAATACATAAAAGATAAGACGTCTACAATCATAACTCGATCATGAAACAGCAAGCGT
CAACCATGAGCACTGACCTTGCTTGCTGCTTGTCGCCGCTGGTTCCAGGGCGCTCAGCGCTCATGCGC
AGGAATACAACGAGTACGATCCCTCAGCTGTCACGCCTGGCAGCATCCCCAACTTCCCATTCCGCGAC
TGCAACACCACAAACGGCGCGTACCGGCTGGCACCGGTGTGGCGGCCTTCGGGCAGCAACACGTACTG
CTTCAAGATCCAGGTCAGCCAGGATGCCCTGTCCTGCACTGGCGCCTGCTGCAGCGCCGACTTGCACA
AAATTGAGGTGAGCCCACGGCCGGCTGATGGTGTGGGGGATGCGACATGCGGACACAGTGCTACAGGT
CCTAGGGCCTCGGAAAGCACGGGAAGAGGCTCATTGGGCCCGCAACCAAATGTCAACGATTATTTTCC
AGTAGTTTCTCATGTGTGTGTGTGTGTGCATGTGCGCATGATGTCGGTGTTGAGGGCTCCGTTACCAT
CCTAGCTGCTCCTCCGGGGAAGTGGATTTCTAATCAGGGTTCGGAACGCGCCTCACAATATTTACATA
GCTCATTTACATGTGCTGGAGAGAATGAGAAGTTCGCACCCTCAAAAGTAGAGGGAACATCAAATTGT
GATGTTCCCTCCTGCCTCACCTTGCCCTTACTCCATCCTCCCACCACATGCCACCCCGCCATGTCCGC
AGTTCAACGTGTCCAGCAGCTGCCTGGTGGCCGGGGCCTCAGTTGTCGCCACCGTGAATGGAGTGCGG
A*CCCGGG*TTGGTGCGACCCTGGACAAGGCCCCGGCTGGGCCTGTTGGCTCCGCCGTTCTGCGGCTCAC *Sma*I
CCAGCTGGGGCTGGACACCACGACCGCCCAGGATGCGGAGGTGTGCCTCACCCTCAAGACCAACCGCG
GAGGCCAGGGCTGCACCACCCTGGAGCAACTGTGCTCGTCGCCAGGCTTCCCTGCAGGGACATGCACA
GCAGCTATGTTCGATGTCGCGTGTGATTGCTGCCCGGTTTCTCAAGCCGGTCAGGCCCTCCCGCCGCC
GCCACCAGTTATTCCGCCGGTCTTCCGGCCTTGTGACGTCTGCATCGCCGCGACCATTGTGCCCCCGG
CCAATGATGTGCGACCATACCGCTACGACAGCGCCACCTGCGCCGCCATTCAGCAGAACATCGCCGAT

GCCATGAACTCTCTCCTGGGAGGCGCCAACATCGGCGCCTTCTCACCTTTCGCGCCAAATGCAAGCCA
GTGCTTTGACACACAAATCATTACATGTGGCAGGTTCAACGGTTCGGACAGTGATGCGTTGGCCAATC
TGACGGAAGCAGTGCAGCAGCAGCTTTCCGCCTTCATTGGAGTCGCGTCTGGCGGCAACGTCTGCAAT
CCAAAGTTGGAGAAATACACGGTGACGGTCTCCACCATCAACAATGTAGCAGATCAGTGCCTGGACCT
CTCGCAGTCGGCTAGTTGCTTCCTGCCTGGCGTTCCATTCCCCAACTGCACGTGAGTCACCCTACTTC
ATGGCGTCGTCTCTTTTGCGGCGCGGCCGCCCTCCGTAAAGATTTGGGGGCGAGGATGTAGGGCCTAA
TAACCAAAGCTTCCCCCCGCAATGGATGGACCCATGTTTGGGATCTGGCTAGCGTCTTTGTCGTGTAT
GCAGAGCTTTGGATGAATGGCTGAAGGCAAAGACACCGCGCTTGCATGGATACCTTGCTCACATGGCT
ATTTGAAAAAATCAGGAACGTTGGGTGAATATGTGATAGTACTAGTTACCGACCCTGCTGCCGCCTCC
GCTCGCCCTCTGACGCGCAGGTGCAACACAACCCAGGGAGTCATGCCGTTCACAGTCTCCCCCACCTG
GTATGCACAGCCGGCCAATGTGCGCTGGGGCCGCAACGTCACTGAGTACTGCTTCACCGTCAACACGC
TACAGCCAAGTCAGGTCGTGCCGGTGAGGCCAGATAGGTGCATGGAGGACCCGTGGGGTGGCGCTTTG
TTTTCTTCCACGCACCCATGCATGCTTGGCTGCATGCCTGATGAGCTCGGTTAGTTGTATCGGGTCTG
GACAAGGCACACGGGCCTGTCTGCTGCGCGTAGCTCACCCAGACAAAGTTAAGAAATTACATGACTGA
GGATGCTGCCTCGTTCGCTGTCCCCCTGCAGAGCACCTGCTACAACGCAAACGACGCGCTCGCCAAGA
TCGAGTGGTATGCAAGTAAGTTATCATTGCTGTTGCTTGCTTGCGTAGTCTTCTATCGTTTCAACGTT
TCTTAGATAAGACCCGTCCTTTCAACCTGGGCAGACATGCATGGATATGTGGGAGTACGGCATGCGAA
TGGATATGGATATGTGCAAACTTCTACTGTCGCAGCATGATGTGAACTCTGGCCCTGCCGGCTGCACA
CATACGCATACATGACTTCCGTAATGCGAGCAGACTGCAGCATAGCACCCAGTACGGCAGCATCAGTG
GTTAACCTTTTTTGGTCCTCGCTACAGCCTAGGGTTTCACCTGCCGTGCATTTTGGCTTGTGTTTCCA
TGACTGCAGGTGATGCGTTCAGGTCGGCGGTCAAGGGTTTCACGGTGTACCCCGCCGGGGGATCAAAC
AAGACCATTGCCGACAGCTGGGGCGCCACTGGAACCGATACCCTGAAAGGTGAGGCCGGCAAGCTAGA
TAGCAGCTCAGGTAGATGCACCTCGTGCGGGGCTCATTCGCTGCTTCGATTGTGGATTCTGTGCGGCA
CTTACCTGCCGTGTGTCACAAGGGTAACTGTCGTGCGAGACCTTCACTTACAATACACTCGAGGAAAT
CTCGACGGACCTGCATCTCGGAAGGCCTCTTCGATGACCCTATCAAAAGATCAGCAATAAGCTCCAAC
AGTGTATGAAGATTGACGGAGCTGTCAAGCTCAATAACATAATCCCGAAAATGGCAAGCACTACGGCT
ACTGTATAAAACCCAGCGCAGAGGCTACACTCCCTTCGCTCCTGGCGCATCTAAGCTGCAATTACAAT
TACAGTAACTCTGTTCAGTAGCCAGCCATGCGCCGGTTTGTTTGTACCGTACCTGTCCAATTGCCATG
TGCTGCACCATCACTGCTAATGCCTGCATGTCAAATAAGATCACATCGTTCCCTGCCTGCGCATCCTT
TTGTTCGTGTGTGCAGTCAATCTCAACTGGAACCTGCTGCAAGCCAACGGCGGCAAGGTTTGCGTGGC
GATTCAGAACCCGTTCACCATGGGCGACATCTGCAAGGGAGCCCTTGGCCAGTGGTAAGTCCTGAGGC
AGGGGTTTGCGAGGCCTATCCTTCTCAACTCAACTAAGATGACCTCCCGCGTACCCGCATCAATACCA
ATACAACGTCCATGGCCCCCCCGGCTTTTACGGACACACCTGCACTTGTGTGCGGAACTTGTAACTTG
TAACTAACGAATGTACGGCCGCGCCACGTGCAAGAA**ACGCGT**TTTCCTTCCCGCACTTTTAGACCACA ***Mlu*I***
AGTTGCACACGTGTCAATTAGCAGCAGTGGTAGTAGGGATGAGTGTGCCTGATTGCTATGTCGCTAAT
AATACTCCTTCATGCCCCTGCTGTATCTTGCCTGCGCCCGTGCAGCTACGCCAGCATCTTCAACAGGG
ACAACTCGGACTACTGCTGCCCCATCTACCGCACGGGGCCA**TGA**TTGCCGTAAGAGCAGTCATGGCGC **stopp**
TGGAGAGCGCAGCAGCCGGAAGTAGGGGTAGAAGTTGTTGTTGCTACAAGTGCAGCAGCAGGCAGAAG
ATAAGGAGCACAATATGGCATCTTAGTACTGCATGAATGAACTGACATACAGACATGTTTACAGATCA
TCAGTCGGCGGGTTTTTCACTTTATCTCTTCGTTCGGTTGCAGCAGTCGATAGATAGATAGACAGACG
AGCTAGACAGTATGATTGATCTGTTGGACGGAATCTTTTCCTTTCCCCGACACAACTGCTTGTAGTAA
CATTTTTTGTAATGATTT**ATCGAT**GTACCACAATTCCTCTCGAACAGTTACGAGGCTAGATACTCGGA ***Cla*I***
CAGGCCTGGTAGGCAGGTGCTGCTGTGCTGCATATATCTATTAGGATGGTGGCGCAGAGGTTTCTTTG
GTGGGTTTGCATGTTGCCATAGACACCGCGGCCGGATGTGCACGGTTTATTAAATGATAATGTGGCTT
GTGAAATTAACGCAAGTAGCTGGCGAGCAGTGGCTACTTGGTCAAGTGGACTGAGTCTGAGCATGACG
TGGTGGGACGCGGTGCGGTCTATTTGGAGCACTAAAGTACATCTCGTCATCAGCAGGCTCTGCTGGGT
CCTGCGGGTGGCGCAAGTTTGAAACGAACGCGCGGTGGCACATTCTCGTGGGTGTCTTCACCAAGGAG
TTCCAAATCCTAAGAACAAGGTTAGCCCAATGGAACTGCAGTGGCGTCTGAGGTAGCTGACGGTAGAG
ACAGATCACGGTGATACAGTAACCTTTTACCGGGGCACATTGTGGCGACTGGAGCCGCATGACCGTCC
CCGTCACTTTGTAACCGCTCAAACTGCACGATAACTCCAACCCACACCGTCTTCAATGCGGCAGCTTG
CGACGGTTTCGC**AGATCT**TCCGAACTGCTCCAATAGCAGCTTGGGATCGACCCGTCATTTCTGCACCA ***Bgl*II**
AGTCGCAAAATAATGACAGATCCTAGGTGGCCGTGGTTATATGCTTTGCCACGTTGTAGAGTCCGAAC
TATATTGTCTTCCCTATGGGCACCCATGGCCATGCTGGATGGCGAAGCACAGTGGTGAGCGGCCTAAG

GTAGCTCTCCGACCAGGAAAACATCACCAGTGCACAGCTCGTCATAGCGATGAATGTTGCAGGGGTGC
AAAGGCATGTTCCACCACACCATTACGAGCAACCTTAGGACCGCGTTCCCCCAGGTAAACAGGGGAGG
TTTCAGGGGCTAGATATACTGGTCCAATAAGCAGGGTGCTATTGATGCAGGCGGTAATGTAGCACCCT
CCCTCATCCGACGTACCTCACTCCCTGCACATGTAGAGCCCCTGCGTAGAACGGCTATCTGTGATGAG
GGGGAGTCCCTGGTATAGAGGCCAATCATGCAAACTCGCTGTGCAGTGCTTCATCAGTTAGGGAGCAT
GTGCCGTTGTTGGACATTCGGGGACTGCGGGACAGCATTGTCAACATCGTTGCAGGCCCTGTAATGGG
CCACATCATGCAGTAGGTTTCTAGCGGTTTCCGAGGCCGAGGTCATTTTGCCGGGGTGTACGGCATTG
TTGCGGGTTGCGGGAAGTTGTACAGGTAAATGCACCGGCCACATTTACCAACATTGCTTTAAATATAA
GTAGACAACCGAGAAAGAACCAAAAGATGTCGGAAAGAATTTGGGCAGAGGCCGAAGCTCGCCTGGCA
AGCACTGCGCTGAACGGATAGAGAGAGACACGAATAAGAAGTGTACAACGGGGCCTCAATAGGCTTTG
GGGTCCATAGATACCGGGCGCCCTCGCTCTCAGGCATTAAGTCTCGGACGCGGGCCACAAGACCTGGA
GCTGCTTGGCTGTTCCCTCTCGTCCCCCCCCTTCAGAAGAAAAACTCTGACCACTATCCCTCTCTCCA
GTTCCCTCTCAGTCTGGGTGGTGCAATACCAGACCCCTCGCCACACACACCGCCACCTCAGCACCAAA
CCGTCACATTACGCTGCAAACATTGAGTCACGTTTGTTAGGCGCTAGCACGCTGTTGTATGAATTAGT
TTCTGAGTCTATGTTGATGTGGTTTGGTCCCGTTGCTGATGCAACTGGCTTTCGCATCATTAACCGCC
GAGTAGCAGCCGTAGCTACCGGTGCGGAAAACGATGCATGCTCACCAACCTAAATGATGAAGGGTGTG
TGGCGAAGTTATATTCATTATGCCAGGATCAGAGCTGTTGGTCAGTAAATATCAAGTTATATAAAAGA
AATGGTGCTACAGATTACAACGACGGGTCGGATTCCGACACGTACAACACGAGCAACATCCCAACGGA
TGAAATGCACATTTATATGTGTCTGCGATTAAC**GAATTC** ***Eco*RI**

**Fig. S2. Genomic sequence of the *V. carteri phll* gene.** The sequence depicted here corresponds to the 8329-bp genomic fragment utilized in plasmid pPhII-YFP. The *phll* gene (1) is located on scaffold 34 (nucleotides 980223 to 985045) of the *V. carteri* genome version 2.1 (2) in Phytozome v13 (3) on the reverse strand. The start codon is at nucleotide position 985025-985027 on the reverse strand. In the current *Volvox carteri* genome annotation available at Phytozome v13 (Volvox v2.1) pherophorin II is not annotated. Therefore, the gene structure was established based on older annotations and confirmed with RNA-Sequencing data (4). The gene structure is indicated as follows: Coding sequences are shown with blue background, UTRs with green background and the promoter region with grey background. Start and stop codons are highlighted (violet font). The 5' UTR is just 18 bp in length, while there is a quite long 3' UTR of 869 bp. The coding sequence totals 1557 nucleotides. The restriction sites that are shown in Figure 2 (main text) and SI Appendix, Fig. S1 A and B are marked (bold, underlined). Restriction sides that were used for inserting the *yfp* coding sequence are marked with asterisks.

**GGATCC**ATGACTGGAAAACCCATCCATGACCAAGTCGTTCCCCGAGAGATAGCAGCTTTGGATCGACT *BamHI*
CGTCGCTTGTGAAAGCCTACCTGGTTACTTAGACGGATTTAGCGACTAGACCACTTATGAAGGTGTTG
ATAAACACCGGGGGTTTCCCTTGGGGTTTTATAGTCGCTTGCAGTCCGTCTGAGGCGCTCGGGATAAT
ATGCCGTACAGTAGTAGATATTCCTGCAGAAGAGAGTCGCACCTCCAGGGAAGAGTTACAATACAGGT
CTGGCCGTGAACGGGCAGACGTGTACAAGATGGTGTACGGAATCGAGTATGTGCACGGATCGTTACAG
CCATGGCCGTAGCCATGTACTGCTGAGTACGGCTGTAGTCGAAAGATGACGAAGATAAGATAGTTTTA
TGTACTGTGTGTACGATATTCGCCTCGACGATGGAATGATAAGAAGGAATGGAGTCCCTACTCGGAAG
TCCTGAGCCAGGCTCTTACACTGGGCACACATTGCCATCCAATGGATGGCCATGTGAATCACAGTGGT
GCCCGGCCTAAGGTAGCTAAGGACTTGGCCGTTATCATCGGTAGATTGCTCACCCACCTTCTCCGACC
AGGAAAACATCCTCGTCACTGGTGCACAGCTCGTCATAGCGATGAATGTTGCAGGGGTGCAAAGGCAA
GTTCCACCACACCATTATGAGCAACCTTAGGACCGCGTTCC**CCCGGG**TAGACACAGGAGGTTTCAGGG *SmaI*
GCTAGATATACTGGTCCAATAAGCAGGGTGCTATTGATGTAGGCAGTAATGTAGCACCCTCCCTCATC
CGACGTACCTCACTCCCTGCACATGTAGAGCCCCTGCGTAGAACGGCTATCTGTGATGAGGGGGAGTC
CCTGGTATAGAGGCCAATGGGAGCATGTGCCGTTGTTGGGCATTCGAGGGGTGCGTGACAGCGTTGTC
ATCGGTAGGGTCGTGCCGAGTTGTGAAGACGGGCTCCTGTAATGGGCCTTATCATGCAGTAGGTTTCT
TGTGGTTTCCGAGGCCGAGGTCATTTTGCCGGGGTGTACGGCATTGTTGCGGGTTGCGGGAAGTTGTA
CAGGTAAATGCATCGGCCATTTTTATCAGCCTTGTTTAAATATAAGTAGACAACCGAGCAAGAACCAA
AGTTTGTCGGAAAGAATTTGGGCAGAGGCCGAAGCTCGCCTGGCAAGCACTGCGCTGAAGGGATAGAG
AGAGACACGAATAAGAAGTGTACAACGGGGCCTCAATAGGCTTTGGGGTCCAGATCGCTGACGCCCTC
GCTCTCAGGCATCAAGTCCCGGACATGGGC**CTTAAG**ACCTGGAGTTGCTGGGCTATTCCCTCTCGCCC *AflII*
TCCCCGCTCTGACCACTATCCCTCTCTCCAGTTCCCTTTCAGTTCGGGTGGTGCAACGCCAGGCCCCT
CGCCACACACACCGCACCGCACCCACGTTTGTTAAGCGCTAGCAGGCTGTTG**TCTAGA**GCTGTAGAAG *XbaI*
TTCCTAGGTCTATGTTGATGTAGTTTGGTCCCGTTGCTGATGCAACTGGCTTTCGCATCATTAACCGC
CGAGTAGCAGCCGTAGCTACCGGTGCGGAAAGGATGCATGCTCACCAACCGAAATGATGAAGTGCGTG
TGGCGAAGTTATATTCATTATGCCAGGATCAGAGATGTTGGTCAGTAAATAGCAAGTTATATAAAAGG
AATGGTGCTACAGATTACAACGACGGGTCGGATTCCGAAACGTACAACACGAGCAACATCCCAACGGA
TGAAATGCACATTTATATGTGTCCGCGACTAAC**GAATTC**TGCGCTGTGTTTCTAAGATTTATATTCTG *EcoRI*
CATAAAGATATTGATTCCAAAGGACGAAGCGCCTCTGTTGCGGTGGCCAGTCTCGGAAGCGGCTACCT
GCGCCCGCCGTTTACGGTAGTCGTACGGCCCCGCAGCAGCACAGCTGCCGCGGTCGGTCGGGTAGATA
AATAGCGCCACTATGATACAGAACATCATCAAAGTGACATAGCTCTGACAAT**ATG**GCATACTTTGCAA start
AGGTAGCTTTCATTTTGGCGACAGCGCTGGCGGTTGGCTGGTGCGTTTTTAGCATGGTGCATTCCCAT
ATGCCTGGGCTTACAATCTGAGGTTGGGTCTATATATTTAAGTATATTTAAGAGCCATATACTTGAAG
GCTTTCAGTCTTTGCCTTTGCGTCGCAAAAGACGTACATGTCCTGTCGCATGCCATAGTGATGAGCCT
TTTCTATAAAGCAATACATAAAAGATAAGACGTCTACAATCATAACTCGATCATGAAACAGCAAGCGT
CAACCATGAGCACTGACCTTGCTTGCTGCTTGTCGCCGCTGGTTCCAGGGCGCTCAGCGCTCATGCGC
AGGAATACAACGAGTACGATCCCTCAGCTGTCACGCCTGGCAGCATCCCCAACTTCCCATTCCGCGAC
TGCAACACCACAAACGGCGCGTACCGGCTGGCACCGGTGTGGCGGCCTTCGGGCAGCAACACGTACTG
CTTCAAGATCCAGGTCAGCCAGGATGCCCTGTCCTGCACTGGCGCCTGCTGCAGCGCCGACTTGCACA
AAATTGAGGTGAGCCCACGGCCGGCTGATGGTGTGGGGGATGCGACATGCGGACACAGTGCTACAGGT
CCTAGGGCCTCGGAAAGCACGGGAAGAGGCTCATTGGGCCCGCAACCAAATGTCAACGATTATTTTCC
AGTAGTTTCTCATGTGTGTGTGTGTGTGCATGTGCGCATGATGTCGGTGTTGAGGGCTCCGTTACCAT
CCTAGCTGCTCCTCCGGGGAAGTGGATTTCTAATCAGGGTTCGGAACGCGCCTCACAATATTTACATA
GCTCATTTACATGTGCTGGAGAGAATGAGAAGTTCGCACCCTCAAAAGTAGAGGGAACATCAAATTGT
GATGTTCCCTCCTGCCTCACCTTGCCCTTACTCCATCCTCCCACCACATGCCACCCCGCCATGTCCGC
AGTTCAACGTGTCCAGCAGCTGCCTGGTGGCCGGGGCCTCAGTTGTCGCCACCGTGAATGGAGTGCGG
A**CCCGGG**TTGGTGCGACCCTGGACAAGGCCCCGGCTGGGCCTGTTGGCTCCGCCGTTCTGCGGCTCAC *SmaI*
CCAGCTGGGGCTGGACACCACGACCGCCCAGGATGCGGAGGTGTGCCTCACCCTCAAGACCAACCGCG
GAGGCCAGGGCTGCACCACCCTGGAGCAACTGTGCTCGTCGCCAGGCTTCCCTGCAGGGACATGCACA
GCAGCTATGTTCGATGTCGCGTGTGATTGCTGCCCGGTTTCTCAAGCCGGTCAGGCCCTCCCGCCGCC
GCCACCAGTTATTCCGCCGGTCTTCCGGCCTTGTGACGTCTGCATCGCCGCGACCATTGTGCCCCCGG
CCAATGATGTGCGACCATACCGCTACGACAGCGCCACCTGCGCCGCCATTCAGCAGAACATCGCCGAT

GCCATGAACTCTCTCCTGGGAGGCGCCAACATCGGCGCCTTCTCACCTTTCGCGCCAAATGCAAGCCA
GTGCTTTGACACACAAATCATTACATGTGGCAGGTTCAACGGTTCGGACAGTGATGCGTTGGCCAATC
TGACGGAAGCAGTGCAGCAGCAGCTTTCCGCCTTCATTGGAGTCGCGTCTGGCGGCAACGTCTGCAAT
CCAAAGTTGGAGAAATACACGGTGACGGTCTCCACCATCAACAATGTAGCAGATCAGTGCCTGGACCT
CTCGCAGTCGGCTAGTTGCTTCCTGCCTGGCGTTCCATTCCCCAACTGCACGTGAGTCACCCTACTTC
ATGGCGTCGTCTCTTTTGCGGCGCGGCCGCCCTCCGTAAAGATTTGGGGGCGAGGATGTAGGGCCTAA
TAACCAAAGCTTCCCCCCCGCAATGGATGGACCCATGTTTGGGATCTGGCTAGCGTCTTTGTCGTGTAT
GCAGAGCTTTGGATGAATGGCTGAAGGCAAAGACACCGCGCTTGCATGGATACCTTGCTCACATGGCT
ATTTGAAAAAATCAGGAACGTTGGGTGAATATGTGATAGTACTAGTTACCGACCCTGCTGCCGCCTCC
GCTCGCCCTCTGACGCGCAGGTGCAACACAACCCAGGGAGTCATGCCGTTCACAGTCTCCCCCACCTG
GTATGCACAGCCGGCCAATGTGCGCTGGGGCCGCAACGTCACTGAGTACTGCTTCACCGTCAACACGC
TACAGCCAAGTCAGGTCGTGCCGGTGAGGCCAGATAGGTGCATGGAGGACCCGTGGGGTGGCGCTTTG
TTTTCTTCCACGCACCCATGCATGCTTGGCTGCATGCCTGATGAGCTCGGTTAGTTGTATCGGGTCTG
GACAAGGCACACGGGCCTGTCTGCTGCGCGTAGCTCACCCAGACAAAGTTAAGAAATTACATGACTGA
GGATGCTGCCTCGTTCGCTGTCCCCCTGCAGAGCACCTGCTACAACGCAAACGACGCGCTCGCCAAGA
TCGAGTGGTATGCAAGTAAGTTATCATTGCTGTTGCTTGCTTGCGTAGTCTTCTATCGTTTCAACGTT
TCTTAGATAAGACCCGTCCTTTCAACCTGGGCAGACATGCATGGATATGTGGGAGTACGGCATGCGAA
TGGATATGGATATGTGCAAACTTCTACTGTCGCAGCATGATGTGAACTCTGGCCCTGCCGGCTGCACA
CATACGCATACATGACTTCCGTAATGCGAGCAGACTGCAGCATAGCACCCAGTACGGCAGCATCAGTG
GTTAACCTTTTTTGGTCCTCGCTACAGCCTAGGGTTTCACCTGCCGTGCATTTTGGCTTGTGTTTCCA
TGACTGCAGGTGATGCGTTCAGGTCGGCGGTCAAGGGTTTCACGGTGTACCCCGCCGGGGGATCAAAC
AAGACCATTGCCGACAGCTGGGGCGCCACTGGAACCGATACCCTGAAAGGTGAGGCCGGCAAGCTAGA
TAGCAGCTCAGGTAGATGCACCTCGTGCGGGGCTCATTCGCTGCTTCGATTGTGGATTCTGTGCGGCA
CTTACCTGCCGTGTGTCACAAGGGTAACTGTCGTGCGAGACCTTCACTTACAATACACTCGAGGAAAT
CTCGACGGACCTGCATCTCGGAAGGCCTCTTCGATGACCCTATCAAAAGATCAGCAATAAGCTCCAAC
AGTGTATGAAGATTGACGGAGCTGTCAAGCTCAATAACATAATCCCGAAAATGGCAAGCACTACGGCT
ACTGTATAAAACCCAGCGCAGAGGCTACACTCCCTTCGCTCCTGGCGCATCTAAGCTGCAATTACAAT
TACAGTAACTCTGTTCAGTAGCCAGCCATGCGCCGGTTTGTTTGTACCGTACCTGTCCAATTGCCATG
TGCTGCACCATCACTGCTAATGCCTGCATGTCAAATAAGATCACATCGTTCCCTGCCTGCGCATCCTT
TTGTTCGTGTGTGCAGTCAATCTCAACTGGAACCTGCTGCAAGCCAACGGCGGCAAGGTTTGCGTGGC
GATTCAGAACCCGTTCACCATGGGCGACATCTGCAAGGGAGCCCTTGGCCAGTGGTAAGTCCTGAGGC
AGGGGTTTGCGAGGCCTATCCTTCTCAACTCAACTAAGATGACCTCCCGCGTACCCGCATCAATACCA
ATACAACGTCCATGGCCCCCCCGGCTTTTACGGACACACCTGCACTTGTGTGCGGAACTTGTAACTTG
TAACTAACGAATGTACGGCCGCGCCACGTGCAAGAA**ACGCGT**TTTCCTTCCCGCACTTTTAGACCACA *MluI**
AGTTGCACACGTGTCAATTAGCAGCAGTGGTAGTAGGGATGAGTGTGCCTGATTGCTATGTCGCTAAT
AATACTCCTTCATGCCCCTGCTGTATCTTGCCTGCGCCCGTGCAGCTACGCCAGCATCTTCAACAGGG
ACAACTCGGACTACTGCTGCCCCATCTACCGCACGGGGCCA**GGTACC**GGCGGAGGCGGTGGCATGAGC *KpnI*
AAGGGCGAGGAGCTGTTCACCGGCGTGGTGCCCATCCTGGTGGAGCTGGACGGCGACGTGAACGGCCA 5xG spacer
CAAGTTCAGCGTGAGCGGCGAGGGCGAGGGCGACGCCACCTACGGCAAGCTGACCCTGAAGCTGATCT
GCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTGGTGACCACCCTGGGCTACGGCCTGCAGTGC
TTCGCCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGAGCGCCATGCCCGAGGGCTACGT mVenus coding sequence
GCAGGAGCGCACCATCTTCTTCAAGGACGACGGTAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGG
GCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGC
CACAAGCTGGAGTACAACTACAACAGCCACAACGTGTACATCACCGCCGACAAGCAGAAGAACGGCAT
CAAGGCCAACTTCAAGATCCGCCACAACATCGAGGACGGCGGCGTGCAGCTGGCCGACCACTACCAGC
AGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCTACCAGAGCAAG
CTGAGCAAGGACCCCAACGAGAAGCGCGACCACATGGTGCTGCTGGAGTTCGTGACCGCCGCCGGCAT
CACCCTGGGCATGGACGAGCTGTACAAG**GGTACC**TGATTGCCGTAAGAGCAGTCATGGCGCTGGAGAG *KpnI*, stopp
CGCAGCAGCCGGAAGTAGGGGTAGAAGTTGTTGTTGCTACAAGTGCAGCAGCAGGCAGAAGATAAGGA
GCACAATATGGCATCTTAGTACTGCATGAATGAACTGACATACAGACATGTTTACAGATCATCAGTCG
GCGGGTTTTTCACTTTATCTCTTCGTTCGGTTGCAGCAGTCGATAGATAGATAGACAGACGAGCTAGA
CAGTATGATTGATCTGTTGGACGGAATCTTTTCCTTTCCCCGACACAACTGCTTGTAGTAACATTTTT
TGTAATGATTT**ATCGAT**GTACCACAATTCCTCTCGAACAGTTACGAGGCTAGATACTCGGACAGGCCT *ClaI**

GGTAGGCAGGTGCTGCTGTGCTGCATATATCTATTAGGATGGTGGCGCAGAGGTTTCTTTGGTGGGTT
TGCATGTTGCCATAGACACCGCGGCCGGATGTGCACGGTTTATTAAATGATAATGTGGCTTGTGAAAT
TAACGCAAGTAGCTGGCGAGCAGTGGCTACTTGGTCAAGTGGACTGAGTCTGAGCATGACGTGGTGGG
ACGCGGTGCGGTCTATTTGGAGCACTAAAGTACATCTCGTCATCAGCAGGCTCTGCTGGGTCCTGCGG
GTGGCGCAAGTTTGAAACGAACGCGCGGTGGCACATTCTCGTGGGTGTCTTCACCAAGGAGTTCCAAA
TCCTAAGAACAAGGTTAGCCCAATGGAACTGCAGTGGCGTCTGAGGTAGCTGACGGTAGAGACAGATC
ACGGTGATACAGTAACCTTTTACCGGGGCACATTGTGGCGACTGGAGCCGCATGACCGTCCCCGTCAC
TTTGTAACCGCTCAAACTGCACGATAACTCCAACCCACACCGTCTTCAATGCGGCAGCTTGCGACGGT
TTCGC**AGATCT**TCCGAACTGCTCCAATAGCAGCTTGGGATCGACCCGTCATTTCTGCACCAAGTCGCA ***Bgl*II**
AAATAATGACAGATCCTAGGTGGCCGTGGTTATATGCTTTGCCACGTTGTAGAGTCCGAACTATATTG
TCTTCCCTATGGGCACCCATGGCCATGCTGGATGGCGAAGCACAGTGGTGAGCGGCCTAAGGTAGCTC
TCCGACCAGGAAAACATCACCAGTGCACAGCTCGTCATAGCGATGAATGTTGCAGGGGTGCAAAGGCA
TGTTCCACCACACCATTACGAGCAACCTTAGGACCGCGTTCCCCCAGGTAAACAGGGGAGGTTTCAGG
GGCTAGATATACTGGTCCAATAAGCAGGGTGCTATTGATGCAGGCGGTAATGTAGCACCCTCCCTCAT
CCGACGTACCTCACTCCCTGCACATGTAGAGCCCCTGCGTAGAACGGCTATCTGTGATGAGGGGGAGT
CCCTGGTATAGAGGCCAATCATGCAAACTCGCTGTGCAGTGCTTCATCAGTTAGGGAGCATGTGCCGT
TGTTGGACATTCGGGGACTGCGGGACAGCATTGTCAACATCGTTGCAGGCCCTGTAATGGGCCACATC
ATGCAGTAGGTTTCTAGCGGTTTCCGAGGCCGAGGTCATTTTGCCGGGGTGTACGGCATTGTTGCGGG
TTGCGGGAAGTTGTACAGGTAAATGCACCGGCCACATTTACCAACATTGCTTTAAATATAAGTAGACA
ACCGAGAAAGAACCAAAAGATGTCGGAAAGAATTTGGGCAGAGGCCGAAGCTCGCCTGGCAAGCACTG
CGCTGAACGGATAGAGAGAGACACGAATAAGAAGTGTACAACGGGGCCTCAATAGGCTTTGGGGTCCA
TAGATACCGGGCGCCCTCGCTCTCAGGCATTAAGTCTCGGACGCGGGCCACAAGACCTGGAGCTGCTT
GGCTGTTCCCTCTCGTCCCCCCCCTTCAGAAGAAAAACTCTGACCACTATCCCTCTCTCCAGTTCCCT
CTCAGTCTGGGTGGTGCAATACCAGACCCCTCGCCACACACACCGCCACCTCAGCACCAAACCGTCAC
ATTACGCTGCAAACATTGAGTCACGTTTGTTAGGCGCTAGCACGCTGTTGTATGAATTAGTTTCTGAG
TCTATGTTGATGTGGTTTGGTCCCGTTGCTGATGCAACTGGCTTTCGCATCATTAACCGCCGAGTAGC
AGCCGTAGCTACCGGTGCGGAAAACGATGCATGCTCACCAACCTAAATGATGAAGGGTGTGTGGCGAA
GTTATATTCATTATGCCAGGATCAGAGCTGTTGGTCAGTAAATATCAAGTTATATAAAAGAAATGGTG
CTACAGATTACAACGACGGGTCGGATTCCGACACGTACAACACGAGCAACATCCCAACGGATGAAATG
CACATTTATATGTGTCTGCGATTAAC**GAATTC** ***Eco*RI**

**Fig. S3. Genomic sequence of the *V. carteri phlI* gene with *yfp* coding sequence added.** The sequence depicted here is present in plasmid pPhII-YFP. The chimeric gene consists of 8.3 kb genomic DNA including 5' and 3' regulatory sequences and the complete transcribed region of pherophorin II including all introns as present on scaffold 34. The *yfp* (mVenus) (8) coding sequence was fused in frame with the coding sequence of the last *phlI* exon and a 15 bp sequence coding for a pentaglycine flexible spacer. For the insertion of mVenus, artificial *Kpn*I sides were introduced by recombinant PCR which also allow for a later exchange of the fluorescent marker. The gene structure is indicated as follows: Coding sequences are shown with blue background, UTRs with green background and the promoter region with grey background. Start and stop codons are highlighted (violet font). The 5' UTR is just 14 bp in length, while there is a quite long 3' UTR of 1,168 bp. The coding sequence of mVenus is highlighted in yellow. The pentaglycine spacer is highlighted in orange. The restriction sites that are shown in SI Appendix, Fig. S1 A and B are marked (bold, underlined). Restriction sides that were used for inserting the *yfp* coding sequence are marked with asterisks.

## 2. Supplementary methods: semi-automated image segmentation and geometric analysis

**A. Overview.** We employ a semi-automated image analysis pipeline which uses Cellpose (9) as a key step.

1. Contrast stretching of the image is performed by predetermined cutoffs, e.g. 2nd and 98th percentile intensity.

2. A J-invariant filtration is performed using (depending on the channel, fluorescence or trans-PMT) either (i) total-variation (TV) denoising by minimizing the Rudin-Osher-Fatemi functional

$$\min_{u\in \mathrm{BV}(\Omega)} \int_\Omega \left[ \|\nabla u\| + \frac{\lambda}{2}(f-u)^2 \right] \qquad [1]$$

where $f$ is the intensity profile of an image supported on the image domain $\Omega$ to be denoised, or (ii) wavelet denoising with adaptive thresholding. A J-invariant filter is defined as one whose output value at every pixel is independent of the value of the source pixel (i.e. is only a function of source pixels at other locations). We find TV denoising in particular to be effective at filtering Poisson noise and significantly enhance the performance of Cellpose on low-SNR fluorescence images. We did not use the trained denoising model available in Cellpose3 (10).

3. A user-prompted input polygon $P$ is used to estimate the diameter $d = \max\{\|\boldsymbol{v}_i - \boldsymbol{v}_j\| \mid \boldsymbol{v}_i, \boldsymbol{v}_j \in P\}$ of typical instances to be identified in the image, passed as the *diameter* input to the Cellpose *cyto3* model (10).

4. Objects identified as pixel-space masks by Cellpose are converted to polygons in the plane by either (i) taking the convex hull, for convex objects such as somatic cells, or (ii) identifying outlines in the mask. Degenerate and invalid polygons are suppressed by taking a single binary erosion-dilation step. Further conversion to ellipses for approximately elliptical objects (such as parent and offspring spheroids) is performed by computing the ellipse with the same $n$th-order moments of area as the polygon (see §B) up to $n = 2$.

5. False-positives are manually rejected where identified. False-negatives, where identified, are re-prompted to Cellpose by restricting to a user-specified region of interest around the object, and re-iterating from step 1.

6. For identification of the somatic CZ3 geometry in particular, only the somatic cell-CZ3 compartment pairs (as seen in Fig. 5, main text) which have jointly been successfully identified are retained.

7. Downstream analysis of the resulting polygons and/or ellipses is performed as described in Table 1 (main text) and further detailed below in §B.

**B. Geometric moments of area.** The geometric moments of bounded planar domains, analogous to the moments of bivariate uniform random variables, quantify shape properties such as size, center of mass, eccentricity, skew, and so on. In the case of polygonal domains, the geometric moments up to sufficiently high order completely determine the vertices (11).

**Definition 2.1** (Planar $n$th area moment tensor)**.** Let $D \subset \mathbb{R}^2$ be an open measurable set with boundary given by a simple closed curve $\partial D = C$. The $n$th moment tensor for the domain $D$ of uniform mass density is

$$\mu_I^{(n)}(D) = \iint_D \boldsymbol{x}^I d^2\boldsymbol{x} \qquad [2]$$

where $I = (i_1, \ldots, i_n)$ is a multi-index such that $i_j \in \{1, 2\}$ and

$$\boldsymbol{x}^I = \prod_{j=1}^{n} x_{i_j}. \qquad [3]$$

The 0th moment $\mu^{(0)}$ is the area of $D$. Accordingly, one may define the radius $R_D$ of an equivalent (same-area) circle as $R_D = \sqrt{\mu^{(0)}/\pi}$. The $n$th-order moments can be calculated for polygons by (i) applying the divergence theorem to write Eq. (2) as a boundary integral on $C$, and (ii) computing the integral as a finite sum using the piecewise-linearity of the sides $C$.

#### *B.1. First moment and centrality.*

**Definition 2.2** (Centroid)**.** The center of mass of $D$ is

$$\mu_j = \frac{\mu_j^{(1)}}{\mu^{(0)}}. \qquad [4]$$

The notation $\boldsymbol{\mu} = [\mu_1, \mu_2]$ evokes the probabilistic interpretation as the expected value of a uniform distribution supported on $D$. The basis in which Eq. (2) is computed will unless otherwise specified be taken, for moments of order $\geq 2$, to be one in which $\boldsymbol{\mu}$ is at the origin.

**Definition 2.3** (Dimensionless centrality of a test point)**.** For a test point $\boldsymbol{y} \in \mathbb{R}^2$, we define the centrality metric

$$d_{\boldsymbol{\mu}}(\boldsymbol{y}) = \|\mathsf{W}(\boldsymbol{y}-\boldsymbol{\mu})\| = \sqrt{(\boldsymbol{y}-\boldsymbol{\mu})\cdot\Sigma^{-1}(\boldsymbol{y}-\boldsymbol{\mu})} \quad [5]$$

with $\mathsf{W}$ a matrix defined as $\mathsf{W}^\intercal\mathsf{W} = \Sigma^{-1}$, and $\Sigma$ the covariance matrix of $D$, defined in Eq. (7).

This *whitening* procedure (again evoking the probabilistic interpretation of $\Sigma$) enables comparison across domains $D$ of varying second moment, resulting in a quantity which is dimensionless. In probability terms, Eq. (5) is the Mahalanobis distance of $\boldsymbol{y}$ to the uniform distribution supported on $D$.

In particular, $d$ is scale-invariant; for dilations of space $\mathbb{R}^2 \mapsto \rho\mathbb{R}^2$, $\rho > 0$, we have $\boldsymbol{\mu} \mapsto \rho\boldsymbol{\mu}, \boldsymbol{y} \mapsto \rho\boldsymbol{y}$, and $\Sigma \mapsto \rho^2\Sigma$ by Eq. (7), hence $d_{\boldsymbol{\mu}}(\boldsymbol{y}) \mapsto d_{\boldsymbol{\mu}}(\boldsymbol{y})$ by Eq. (5). Moreover, dilations preserve aspect ratios, since

$$\Sigma \overset{\mathbb{R}^2\mapsto\rho\mathbb{R}^2}{\mapsto} \frac{1}{\rho^2\,|D|}\iint_D \rho(\boldsymbol{x}-\boldsymbol{x'})\otimes\rho(\boldsymbol{x}-\boldsymbol{x'})\rho^2 d^2\boldsymbol{x'} = \rho^2\Sigma, \quad [6]$$

hence its eigenvalues map as $\lambda_j \mapsto \rho^2\lambda_j$ and the ratio $\lambda_{\max}/\lambda_{\min}$ is preserved.

### *B.2. Second moment and isotropy.*

**Definition 2.4** (Covariance matrix of a domain)**.** The normalized second central moment, or covariance matrix, is

$$\Sigma = \frac{\mu^{(2)}}{\mu^{(0)}} \quad [7]$$

As before, *central* indicates that $\mu^{(2)}$ is computed in a basis in which $\boldsymbol{\mu}$ is at the origin.

In probability terms, $\Sigma$ is the covariance matrix of a uniform distribution supported on the domain $D$. If $D$ has radial symmetry about $\boldsymbol{\mu}$ (either continuous, as a circle does, or discrete, as regular $n$-gons do), then $\Sigma = c\mathsf{I}$ for some $c > 0$. This is verified by the existence of an eigenspace of dimension 2. $\Sigma^{-1} = \mathsf{M}$ is the matrix defining an ellipse with the same aspect ratio as $D$.

**Definition 2.5** (Principal axes of a domain)**.** Using Eq. (7), we may define the principal axes and stretches of a domain $D$ via Hermitian eigendecomposition

$$\Sigma = \mathsf{P}\Lambda\mathsf{P}^{-1} \quad [8]$$

with $\boldsymbol{v}_i$ the columns of $\mathsf{P}$ being the principal axes and and $\lambda_i > 0$ in ascending order (as $\Sigma$ is symmetric positive-definite for non-degenerate curves).

**Definition 2.6** (Aspect ratio of a domain)**.** Let $\lambda_1, \lambda_2$ and $v_1, v_2$ be the principal stretches and axes (in ascending order) of $\Sigma$ as in Eq. (8). The aspect ratio is

$$\alpha = \sqrt{\frac{\lambda_2}{\lambda_1}}. \quad [9]$$

Ellipses (objects defined uniquely by their moments to order $n = 2$) with the same orientation, aspect ratio, and area ($\pi ab$) as $D$ have the minor and major axes

$$a = \sqrt{\frac{\mu^{(0)}}{\pi\alpha}}, \quad b = \alpha a. \quad [10]$$

Eq. (10) is an exact elliptical representation of a polygon (or arbitrary domain), in contrast to elliptical approximations, e.g. (i) weighted $\ell^2$-minimization of vertex distance or (ii) convex programs for minimum-area bounding ellipses.

***B.3. Moments under affine transforms.*** Let $\mathsf{F} > 0$ be a symmetric positive-definite matrix (e.g. a strain tensor) and define an affine transform of a domain $D$ about its center of mass by

$$T(\boldsymbol{x}) = \mathsf{F}(\boldsymbol{x}-\boldsymbol{\mu}) + \boldsymbol{\mu}. \quad [11]$$

Let $T(D)$ be the transformed region. By change of coordinates for integrals, the relevant moments of $T(D)$ are

$$\mu^{(0)}(T(D)) = \iint_{T(D)} d^2\boldsymbol{y} = \iint_D \det\mathsf{F}\, d^2\boldsymbol{x} = \mu^{(0)}(D)\det\mathsf{F} \quad [12]$$

$$\mu^{(1)}(T(D)) = \iint_{T(D)} y d^2\boldsymbol{y} = \iint_D (\mathsf{F}(\boldsymbol{x}-\boldsymbol{\mu})+\boldsymbol{\mu})\det\mathsf{F}\, d^2\boldsymbol{x} = \mu^{(1)}(D)\det\mathsf{F} \quad [13]$$

$$\boldsymbol{\mu}(T(D)) = \mu^{(1)}(T(D))/\mu^{(0)}(T(D)) = \boldsymbol{\mu}(D) \quad [14]$$

$$\mu^{(2)}(T(D)) = \iint_{T(D)} (\boldsymbol{y}-\boldsymbol{\mu})\otimes(\boldsymbol{y}-\boldsymbol{\mu}) d^2\boldsymbol{y} = \iint_D \mathsf{F}(\boldsymbol{x}-\boldsymbol{\mu})\otimes\mathsf{F}(\boldsymbol{x}-\boldsymbol{\mu})\det\mathsf{F}\, d^2\boldsymbol{x} = \mathsf{F}\mu^{(2)}(D)\mathsf{F}^\intercal\det\mathsf{F} \quad [15]$$

$$\Sigma_{T(D)} = \mu^{(2)}(T(D))/\mu^{(0)}(T(D)) = \mathsf{F}\Sigma_D\mathsf{F}^\intercal \quad [16]$$

hence the centroid is preserved and the covariance matrix maps as $\Sigma \mapsto \mathsf{F}\Sigma\mathsf{F}^\intercal$.

***B.4. Whitening a domain.*** Let a domain $D$ have covariance matrix $\Sigma$ Eq. (7). Define the *whitened* domain $D_W$ by the affine transform $T$ as defined in Eq. (11),

$$D_W = \{T(\boldsymbol{x}) \mid \boldsymbol{x} \in D\}, \quad \mathsf{F} = \Sigma^{-1/2}, \tag{17}$$

with the matrix square root $\mathsf{F}$ typically approximated by singular value decomposition (SVD) as

$$\mathsf{F} \overset{\varepsilon \to 0^+}{=} \mathsf{U}(\mathsf{S} + \varepsilon)^{-1/2}\mathsf{U}^*, \quad \Sigma = \mathsf{USV}^\intercal, \tag{18}$$

and $\mathsf{F}$ the *ZCA whitening matrix* (12) with regularization constant $\varepsilon \ll 1$. Then $D_W$ has aspect ratio 1 as defined in Eq. (9).

**C. Isoperimetric problems.** We recall here several quantities which can be used as measures of the deviation of a domain $D$ from a disk.

***C.1. Classical isoperimetric inequalities.*** One has

$$4\pi A \leq L^2, \tag{19}$$

where $A$ is the area of $D$ and $L$ the total arclength of $C$ (which we may now require to be a rectifiable Jordan curve) and is an equality only for circles. Accordingly, one defines an *isoperimetric quotient*, which we term the *circularity* in the main text,

$$q = \frac{\sqrt{4\pi A}}{L} \in [0, 1], \tag{20}$$

maximized for disks. For regular $n$-gons, one has

$$q_n = \sqrt{\frac{\pi}{n}\cot\frac{\pi}{n}}. \tag{21}$$

A natural question is whether Eq. (19) can be used to define a set distance, in the sense of Hausdorff metric, of $D$ to the "best" disk. This turns out (13) to be related to the problem of making Eq. (19) *quantitative*, in the sense of a nonnegative quantity $\nu(D)$ such that

$$4\pi A + \nu(D) \leq L^2 \tag{22}$$

for all $D$, with $\nu(D) = 0$ iff $D$ is a disk. Without proof, we cite (13) the result that the *isoperimetric deficit*, defined as the dimensionless quantity

$$\mathrm{ID}(D) = \frac{1-q}{q} \tag{23}$$

upper-bounds any such quantitative inequality $\nu(D)$ via

$$\nu(D) \leq C_2\sqrt{\mathrm{ID}(D)} \tag{24}$$

for some dimension-dependent constant $C_2$. For convex $D$ (applicable e.g. to the cells of a Voronoi tessellation), Eq. (23) upper-bounds the Hausdorff distance $d_H$ to the best-fit equal-volume ball $B$ as

$$\inf_{\boldsymbol{x}\in\mathbb{R}^2} d_H(D, B + \boldsymbol{x}) \leq C_2\mathrm{ID}(D)^{\alpha_2} \tag{25}$$

for dimension-dependent constants $C_2, \alpha_2$ (13). Equality is achieved for $D$ which is a ball.

***C.2. Weighted isoperimetric inequalities.*** Recalling the second moment $M_2$ as defined in the main text, eq. 1, written here with explicit arguments $M_2(\boldsymbol{x}, D)$, we recall by classical results that its minimization is also an isoperimetric problem.

**Lemma 1** (Disks minimize $\mathrm{Tr}(M_2)$)**.** Let $D \subset \mathbb{R}^2$ be as in Definition 2.1, $|D|$ its Lebesgue measure, and $\boldsymbol{x} \in \mathbb{R}^2$. We have $\mathrm{Tr}(M_2(\boldsymbol{x}, D)) \geq \frac{|D|^2}{2\pi}$ with equality iff $D$ is a disk and $\boldsymbol{x}$ is its centroid.

*Proof.* Write $\mathrm{Tr}(M_2)$ as

$$U(\boldsymbol{x}, D) = \mathrm{Tr}(M_2(\boldsymbol{x}, D)) = \iint_D \mathrm{Tr}[(\boldsymbol{x}' - \boldsymbol{x}) \otimes (\boldsymbol{x}' - \boldsymbol{x})]d^2\boldsymbol{x}' = \iint_D \left\|\boldsymbol{x}' - \boldsymbol{x}\right\|^2 d^2\boldsymbol{x}'. \tag{26}$$

Then optimality in the first argument implies

$$0 = \frac{\partial U}{\partial \boldsymbol{x}} = 2\iint_D (\boldsymbol{x}' - \boldsymbol{x})d^2\boldsymbol{x}', \tag{27}$$

which holds iff $\boldsymbol{x} = \boldsymbol{\mu}$ is the centroid of $D$. Locating $\boldsymbol{\mu}$ at the origin without loss of generality, we have

$$U(\boldsymbol{x}, D) \geq U(\boldsymbol{\mu}, D) = \iint_D \left|\boldsymbol{x}'\right|^2 d^2\boldsymbol{x}'. \tag{28}$$

Recognizing the last expression as the polar moment of inertia and applying a *weighted* isoperimetric inequality (7.2, (14)),

$$\geq \iint_{B_{R_D}(0)} \left|\boldsymbol{x}'\right|^2 d^2\boldsymbol{x}', \tag{29}$$

where $B_{R_D}(0)$ is a disk at the origin of the same area as $D$ (i.e. $R_D$ is $D$'s circular radius as in §B). Thus

$$= \frac{\pi R_D^4}{2} = \frac{|D|^2}{2\pi}. \tag{30}$$

□

Recall also the (standardized) sum of second moments, Eq. 3, main text. Using the previous lemma we may establish the following optimality for it in terms of honeycombs.

**Lemma 2** (Second moment bound for tessellations)**.** For any space packing $\{\boldsymbol{x}_i\}_{i=1}^n, \{D_i\}_{i=1}^n$ with standardized sum of second moments $m_2$ defined as

$$m_2 = n \sum_{i=1}^n \mathrm{Tr}(\mathsf{M}_2^{(i)}(\boldsymbol{x}_i, D))/\left(\sum_{i=1}^n |D_i|\right)^2, \tag{31}$$

we have $m_2 \geq \frac{1}{2\pi}$.

*Proof.* Let $A = \sum_{i=1}^n |D_i|$ and $\mathsf{M}_2(\boldsymbol{x}_i, D_i) = \mathsf{M}_2^{(i)}$. Then

$$m_2 = \frac{n}{A^2} \sum_{i=1}^n \mathrm{Tr}\, \mathsf{M}_2^{(i)} \tag{32}$$

$$\overset{\text{CS}}{\geq} \frac{1}{A^2}\left(\sum_{i=1}^n (\mathrm{Tr}\, \mathsf{M}_2^{(i)})^{1/2}\right)^2 \tag{33}$$

$$\overset{\text{L1}}{\geq} \frac{1}{A^2}\left(\sum_{i=1}^n \frac{|D_i|}{\sqrt{2\pi}}\right)^2 \tag{34}$$

$$= \frac{1}{2\pi}, \tag{35}$$

where the first inequality is Cauchy-Schwarz and the second is Lemma 1. □

Note that the first inequality is sharp only for configurations consisting of congruent cells $D_i$. The second inequality cannot be sharp for space *partitions*, only *packings* consisting of congruent disks. Space *partitions* optimize the second only if they are asymptotically (in $n$) regular hexagonal lattices (15, 16).

**D. Average number of neighbors in a spherical Voronoi tessellation.** Let $\{\boldsymbol{x}_i\}_{i=1}^n$ be a collection of points on the surface of a sphere and $\{V_i\}_{i=1}^n$ the corresponding set of spherical polygons given by their Voronoi tessellation. The number of neighbors of each $V_i$ is precisely the degree (number of incident edges) $d(\boldsymbol{x}_i)$ of each node in the topological dual, the Delaunay triangulation $\{T_j\}_{j=1}^f$ of $\{\boldsymbol{x}_i\}$. Let $m$ be the number of edges; since this is a triangulation of a compact (boundaryless) surface, we have the relation $3f = 2m$. By the Euler theorem,

$$n - m + f = \chi = 2 \tag{36}$$

where $\chi$ is the Euler characteristic of the sphere. Substituting the triangulation property into Eq. (36) yields $m = 3n - 6$, hence the average degree $\overline{d}$ is

$$\overline{d} = \frac{1}{n}\sum_{i=1}^n d(\boldsymbol{x}_i) = \frac{2m}{n} = \frac{6n-12}{n} \overset{n\to\infty}{\longrightarrow} 6, \tag{37}$$

thus the average number of neighbors of a Voronoi polygon is asymptotically 6.

**E. Distortion correction of compartment features due to elastic deformation against a glass slide.** Segmented compartment features as described in §2.A occur in a single or tight range of focal planes due to the physical planarity of the surface, which arises from the mounting procedure described in Materials and Methods in the main text. This can induce some elastic distortion of the geometry of the compartments, as the surface of the organism itself is generally ellipsoidal and cannot be embedded isometrically in the plane. This distortion varies from spheroid to spheroid and also in principle depends on the constitutive (stress-strain) relations of the ECM material, making the physical inversion of this elastic distortion an ill-posed problem. However, we find using the following 3D analysis a generic shape for the organism's section close to and in contact with the mounting slide, and propose a simple, general distortion correction mechanism which we use to verify that the conclusions drawn from 2D features are accurate reflections of the 3D-corrected shapes.

***E.1. Extraction of volumetric compartment hull from voxel data.*** The PhII:yfp-stained CZ3 compartment boundaries appear in fluorescence z-stacks as local maxima in intensity, providing a basis from which to extract the bounding shape (convex hull) of the overall compartment geometry by thresholding. This procedure is implemented as follows.

1. Apply masking to z-stacks using the outer boundaries identified in Fig. S5 to isolate the spheroid of interest in 3D space.

2. Equalize intensity across successive z-planes (whose mean generally drops off further in $z$) by multiplying intensity data of yfp emission at plane $z$ by the ratio $\mu_z/\mu$, where $\mu_z$ is the mean intensity at plane $z$ and $\mu$ is that of the entire volume.

3. Apply a top-hat box transform to subtract background features from the voxel data with radius large compared to the scale of compartent wall thicknesses (on the order of $\sim 2\mu$m as estimated in the main text, p. 4).

4. Obtain an intensity mask by Otsu thresholding (17).

5. Obtain an approximate triangulated surface using the marching cubes algorithm (18).

6. Compute the convex hull of this surface to obtain a convex triangulated surface representing the outer boundary of the CZ3 geometry. Extract the vertices of this hull for use in the following section.

We find consistently that the resulting shape is well-approximated by an ellipsoidal section bounded by a planar slice as shown in Fig. S17; this plane is where the majority of surface compartment features appear and the section in which segmentation (Fig S5, blue dots) is feasible.

Not all spheroids analyzed are sufficiently stably mounted in order to facilitate consistent extraction of the 3D shape from volumetric fluorescence data (nevertheless, surface CZ3 analysis as in the main text is possible from a single or a tight range of focal planes); the ones for which it is feasible are shown in Fig. S5.

***E.2. Fitting the ellipsoidal section.*** We now assume that the organism body is mirror-symmetric about some plane $z > 0$; this $z$ is unknown as it depends on the exact spacing between the two surfaces mounting the spheroid. However, we present here a method for estimating it.

1. For a plane at some height $z > 0$ parallel to the XY plane, make a mirror-image of the vertices extracted in §2.E.1 about it.

2. Compute the minimum-volume bounding ellipsoid containing this point cloud $\{x_i\} \in \mathbb{R}^{n\times 3}$ by solving the convex program

$$\max \log\det A \qquad [38]$$

$$\text{subject to } \|Ax_i + b\| \leq 1 \ \forall i, \ A \geq 0 \qquad [39]$$

   where the resulting ellipsoid $(M, v)$ defined by the equation $(x - v)^\intercal M(x - v) = 1$ is given by the solution as $M = A^\intercal A$, $v = -A^{-1}b$. $A \geq 0$ means that $A$ is positive semidefinite.

3. Compute the cost of the original points not passing through the surface as

$$c = \sum_i ((x_i - v)^\intercal M(x_i - v) - 1)^2 \qquad [40]$$

4. Repeat steps 1-3 with perturbations in the height $z$ to find a fit with locally minimum cost $c$.

As shown in Fig. S17, this ellipsoidal section is a good approximation of the vertices obtained from the CZ3 convex hull obtained in §2.E.1

***E.3. Projection of the compartment geometry.*** We first note that cells do not experience the same geometric distortion in top view as compartments due to compression because (1) they are approximately spherical objects with similar cross-section from any viewing angle and (2) they are not (significantly) physically deformed by the mounting procedure—rather the outer structures which contain them, i.e. the cellular zone (CZ) and boundary zone (BZ), are. Therefore, the following projections are performed only for the compartment geometry.

1. Given the ellipsoid $(M, v)$ computed in §2.E.2, translate $v$ in $z$ so $v_z > r_{\max}$, where $r_{\max}$ is the maximum of the elliptic radii $1/\sqrt{\lambda_i}$, where $\lambda_i > 0$ are the eigenvalues of $M$. This ensures $(M, v)$ lies in the upper-half $z$ space.

2. For each polygon in the $XY$ plane defined as a set of points $\{(x_i, y_i)\}_i$, identify the coordinate $z_i$ projected onto the ellipsoid by solving the quadratic equations $w_i^\intercal M w_i = 1$, where $w_i = (x_i, y_i, z_i) - v$. This is solved using Vieta's formula to avoid floating-point cancellation. The roots $z_i$ with smallest value are chosen.

3. The resulting surface $\{(x_i, y_i, z_i)\}_i$ is no longer planar, so find the best-fit plane using the standard SVD solution (12) and take the orthogonal projection of these points onto the plane. This results in a set of CZ3 compartment polygons each lying in their own plane approximately tangential to the ellipsoidal surface computed in §2.E.2. All downstream analyses can then be computed within a basis defined by this plane.

This projection is visualized in Fig. S17.

## 3. Supplementary analyses

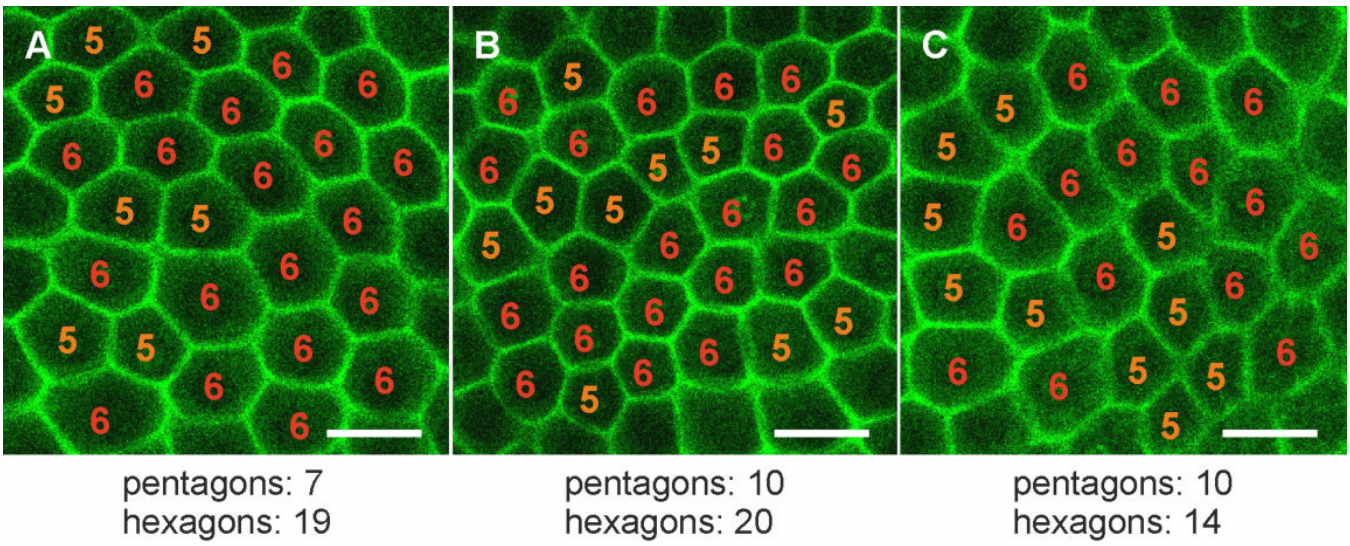


**Fig. S4. Share of pentagonal and hexagonal somatic CZ3 compartments in middle aged adults (early stage II).** Sexually induced transformants expressing the *phII:yfp* gene under the control of the endogenous *phII* promoter were analyzed in vivo for the localization of the PHII:YFP fusion protein. Magnified view of the PhII:YFP-stained compartments surrounding the somatic cells corresponding to CZ3. In areas where no gonidia lie below the somatic cell sheet, the compartments form a pattern of hexagons and pentagons. In the exemplary regions shown here, 27 pentagonal and 53 hexagonal compartments were counted, representing a ratio of roughly 1:2. Scale bars are 20 $\mu$m.

| | **Spheroid** | | | | | | |
|---|---|---|---|---|---|---|---|
| **Stage I** | I.1 | I.2 | I.3 | I.4 | I.5 | I.6 | I.7 |
| Circular radius ($\mu$m) | 96 | 107 | 104 | 109 | 103 | 117 | 108 |
| Segmented cell-compartment pairs | 152 | 317 | 267 | 260 | 265 | 249 | 248 |
| **Stage II** | II.1 | II.2 | II.3 | II.4 | II.5 | | |
| Circular radius ($\mu$m) | 188 | 230 | 221 | 247 | 219 | | |
| Segmented cell-compartment pairs | 397 | 494 | 548 | 535 | 495 | | |
| **Stage III** | III.1 | III.2 | III.3 | III.4 | III.5 | | |
| Circular radius ($\mu$m) | 245 | 243 | 266 | 241 | 223 | | |
| Segmented cell-compartment pairs | 447 | 404 | 520 | 439 | 437 | | |
| **Stage IV** | IV.1 | IV.2 | IV.3 | IV.4 | IV.5 | | |
| Circular radius ($\mu$m) | 416 | 419 | 419 | 428 | 430 | | |
| Segmented cell-compartment pairs | 519 | 446 | 430 | 550 | 529 | | |
| **Stage S** | S.1 | S.2 | S.3 | S.4 | S.5 | S.6 | S.7 |
| Circular radius ($\mu$m) | 254 | 281 | 269 | 278 | 205 | 287 | 280 |
| Segmented cell-compartment pairs | 415 | 466 | 497 | 531 | 396 | 580 | 551 |

**Table S1. Summary of individual spheroid data: equivalent circular radius and number of analyzed cell-compartment pairs.**

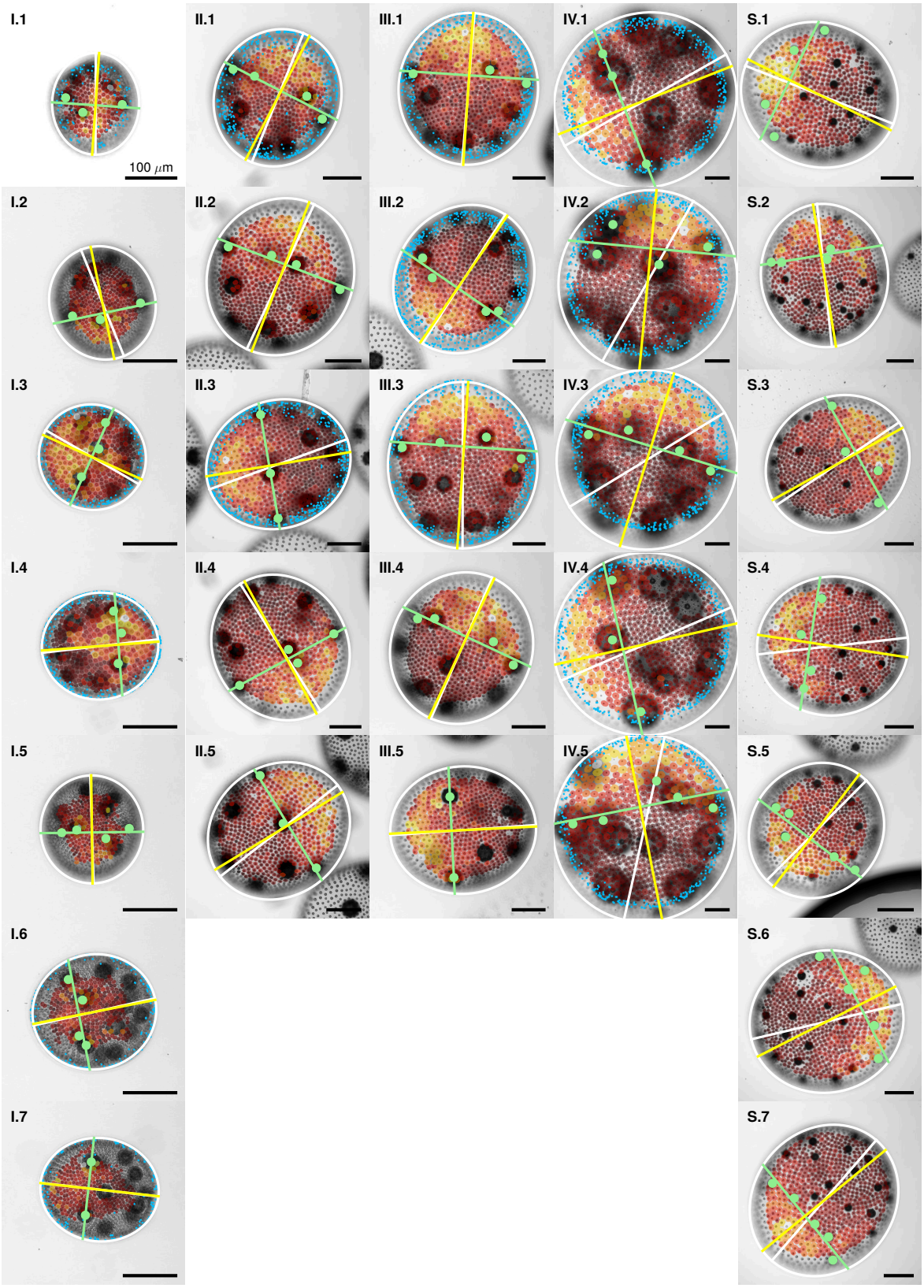


**Fig. S5. Segmented PhII:YFP signal and anterior-posterior axis identification.** Trans-PMT images of spheroid in stage I-IV and S (defined in Fig. 3 in the main document) with posterior-anterior axis (yellow line) estimated as the line passing through the elliptical (white outline) center which is normal to the best-fit line (green) through manually identified offspring (green dots) located in the anterior section. The white line represents the major axis of the ellipsoid. Overlaid are segmentations of the CZ3 compartments (PhII:YFP), colored by area (dark to light by size). Blue dots are the $(x, y)$ coordinates of the CZ3 convex hull coordinates as defined in §2.E.2 (in the spheroids from which extraction is possible) and whose full 3D coordinates are also rendered in Fig. S17. The approximately elliptical middle regions where points are absent are the flat contact patches between organisms and cover slips (Materials and Methods, main text), defining the feasible segmentation region.

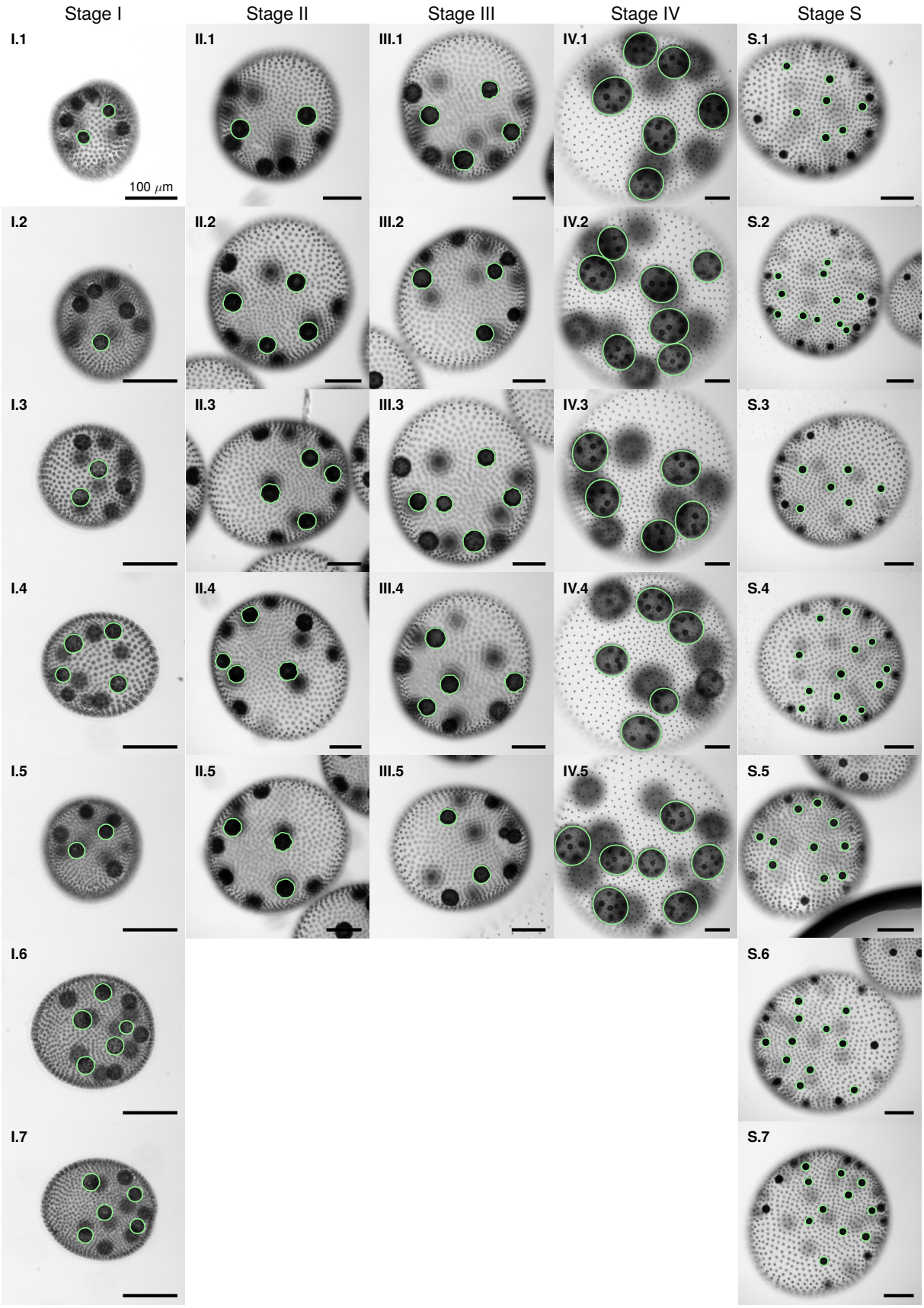


**Fig. S6. Offspring.** Offspring are identified using the same semi-automated procedure as that used for identification of somatic cells and CZ3 compartment geometry. Being far larger than either of those, however, offspring appear in focus at different planes. We identify offspring from planes where clearly in focus, displaying here particular planes for each spheroid. The boundaries are used to estimate the volumetric growth rates displayed in Table 2, main text.

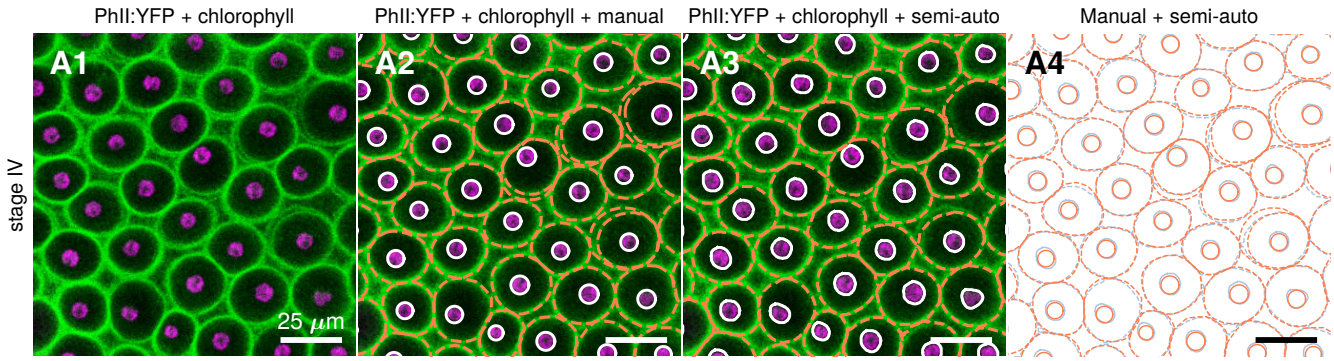


**Fig. S7. Comparison of semi-automated segmentation against fully manual segmentation.** 1. Overlay of YFP fluorescence of PhII:YFP protein (green) and chlorophyll fluorescence (magenta), detected at 650-700 nm. 2. Same as 1 with *manually* segmented CZ3 (orange) and cell boundaries (white), identical to panel A2, Fig. 5 in the main text. 3. Same as 2 with *semi-automatically* segmented (using the procedure described in §2) CZ3 (orange) and cell boundaries (white). 4. Comparison of semi-automated segmentation (blue, CZ3 compartments and cells) with fully manual segmentation (orange, CZ3 compartments and cells).

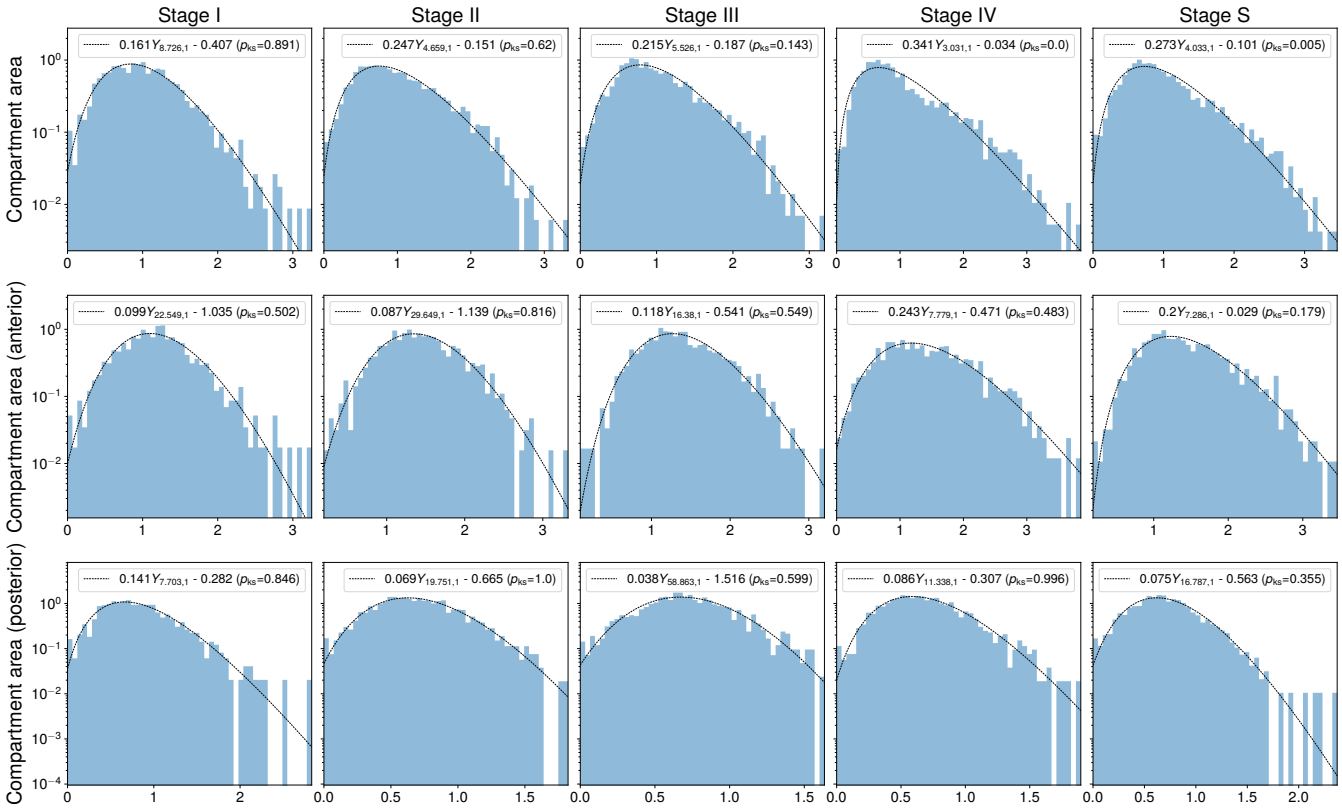


**Fig. S8. Areas are gamma-distributed throughout growth with changing shape parameters reflecting anterior-posterior differentiation.** Areas of CZ3 compartments robustly follow gamma-distributions, as highlighted in Fig. 8B2 and §E, main text. Plotted on y-axes (shared by row) are normalized counts per bin ($n = 50$ bins), and on x-axes are the reduced CZ3 areas $\tilde{a}_{\mathrm{cz3}} = (a_{\mathrm{cz3}} - a_{\mathrm{min}})/(a_{\mathrm{avg}} - a_{\mathrm{min}})$, where $a_{\mathrm{min}}$ and $a_{\mathrm{avg}}$ are minimum and average values of $a_{\mathrm{cz3}}$ for the respective spheroid. Such normalization enables distributional fit across different organisms. The empirical mean is therefore fixed at $1$ in these distributions, and the maximum-likelihood fit parameters are $k$ and offset of the support (with $\lambda$ determined by the relation between $k$ and the fixed mean). A Kolmogorov-Smirnov (KS) goodness-of-fit test is performed and its $p$-value is recorded as $p_{\mathrm{ks}}$ in each plot. One key observation we make, evident in the first row, is that the extreme degree of anterior-posterior differentiation (shown in Fig. 8B1, main text) suggests that a single fit from this distribution family may not be valid in later stages of the life cycle. Indeed, $p_{\mathrm{ks}} \approx 0.9$ in stage I, indicating no strong evidence that $\tilde{a}_{\mathrm{cz3}}$ do not arise from a gamma distribution, yet drops below $0.001$ by stage IV, supporting rejection of the fit hypothesis in this case. Qualitatively, one observes the formation of "shoulders" in the distribution (exhibiting a loss of log-concavity which does not hold in the gamma distribution for shape parameter $k \geq 1$). Separating the data by anterior/posterior hemispheres, however, almost completely resolves the issue and reveals that the respective $\tilde{a}_{\mathrm{cz3}}$ values follow gamma distributions (of very different shape parameters). Lastly, fitted parameters of the sexual stage S place it distributionally closer to Stage IV than to III, which contrasts with the grouping inferred from the spheroid radii (Fig. 10E, main text).

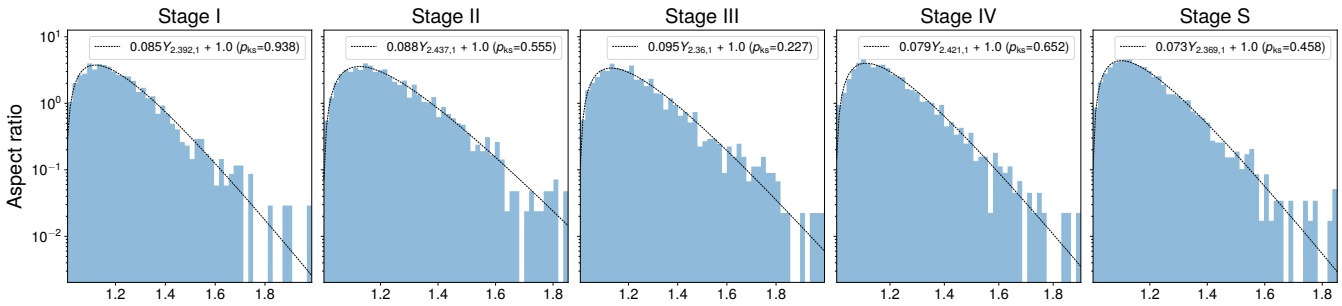


**Fig. S9. Aspect ratios remain stably gamma-distributed throughout growth.** As seen, the $k$ parameter remains in a remarkably tight range between $2.35 - 2.45$. Maximum-likelihood estimated (MLE) parameters include only rate $\lambda$ and shape $k$, with location fixed at $1$. Both this distribution family and the value of the shape parameter for aspect ratios are robustly observed in a variety of living organisms and inert jammed systems (19). Kolmogorov-Smirnov tests performed in each case indicate no strong evidence that the data do not follow gamma distributions.

| Stage | Spheroid volume change (est., $mm^3$) | Total offspring volume change (est., $mm^3$) | Total cell volume change (est., $\mu m^3$) | % spheroid volume change to offspring (est., %) | % spheroid volume change to cells (est., %) | Elapsed time (h) | Spheroid growth rate (est., $mm^3$/h) | Total offspring volume growth rate (est., $mm^3$/h) | Total cell volume growth rate (est., $\mu m^3$/h) |
|---|---|---|---|---|---|---|---|---|---|
| I | ↓ | ↓ | ↓ | ↓ | ↓ | ↓ | ↓ | ↓ | ↓ |
| II | 0.0403 | 0.0011 | $3.1 \times 10^5$ | 2.8 | 0.76 | 15 | 0.0027 | $7.5 \times 10^{-5}$ | 10.2 |
| III | 0.0153 | 0.0002 | $4.2 \times 10^4$ | 1.2 | 0.28 | 6 | 0.0025 | $3.1 \times 10^{-5}$ | 3.5 |
| IV | 0.2553 | 0.0258 | $7.9 \times 10^4$ | 10.1 | 0.03 | 16 | 0.0160 | $1.6 \times 10^{-3}$ | 2.5 |

**Table S2. Estimated volumetric growth changes by life cycle stage (I-IV), supplementary to Table 2, main text. Number of juveniles per spheroid were manually estimated from the trans-PMT image (as in Fig. S6), with the average value over all $n = 29$ spheroids being exactly $13$. The estimated total offspring volume change (column 3) is defined as this number, times the average estimated juvenile volume per stage (as given in Table 2, main text, computed from segmentation of the trans-PMT image, Fig. S6). Further, assuming a fixed count of $2^{11} \approx 2000$ somatic cells per spheroid and multiplying this times the average somatic cell volume per stage (again Table 2, main text) yields the estimated total volume change due to somatic cells (column 4). Assuming $2^{12}$ somatic cells changes the contribution to the overall spheroid volume negligibly, as evident from column 6. Taken together, columns 2, 3, and 4 yield an estimation of the contribution to overall spheroid volume change by offspring, somatic cells, and parental ECM (Table 2, column 5, main text) by subtracting columns 3 and 4 from 2. Lastly, knowledge of the approximate elapsed time between life cycle stages (column 7) yield the corresponding estimated growth rates (columns 8-10 here and Table 2, columns 5-6, main text).**

| | Cell area | Compartment area | Aspect ratio | Circularity | Offset | Offset (whitened) | Voronoi error |
|---|---|---|---|---|---|---|---|
| % change PA in I | 13.72 | 44.25 | 9.85 | -2.49 | 46.90 | 22.59 | 1.66 |
| % change PA in II | 13.44 | 93.90 | 3.06 | -0.69 | 117.00 | 65.15 | 15.61 |
| % change PA in III | 13.28 | 121.88 | 1.01 | -0.67 | 82.04 | 36.93 | -6.55 |
| % change PA in IV | 8.12 | 129.56 | 14.39 | -0.53 | 49.86 | 8.72 | -10.31 |
| % change PA in S | 9.90 | 128.90 | -1.76 | -1.94 | 195.36 | 105.17 | -3.06 |
| Mean value in stage I | 26.37 | 57.85 | 1.20 | 0.88 | 0.54 | 0.24 | 0.21 |
| % change mean I to II | 81.47 | 201.09 | 0.93 | 1.95 | 141.60 | 43.29 | 21.60 |
| Mean value in stage II | 47.85 | 174.18 | 1.21 | 0.90 | 1.29 | 0.35 | 0.25 |
| % change mean II to III | 2.38 | 30.02 | 0.86 | -0.78 | 27.36 | 12.05 | -1.41 |
| Mean value in stage III | 48.99 | 226.48 | 1.22 | 0.89 | 1.65 | 0.39 | 0.25 |
| % change mean III to IV | 12.21 | 165.02 | -2.74 | 1.82 | 52.87 | -5.58 | 15.60 |
| Mean value in stage IV | 54.97 | 600.21 | 1.19 | 0.91 | 2.52 | 0.37 | 0.29 |

**Table S3. Summary of changes in mean value along posterior-anterior axis (PA) and by life cycle stage (I-IV, S). Rows 1-5 show the changes in empirical mean value from posterior to anterior end (corresponding to dashed mean lines, column 1, Fig. 8, main text) within each life cycle stage . Rows 6-12 show the changes in mean value (corresponding to vertical black bars overlaying the histograms, column 2, Fig. 8, main text) across successive life cycle stages. While the underlying distributions are highly skewed and mean-based comparisons should be interpreted carefully as noted in §E1 (main text), they provide one quantification of the underlying trends visible in the empirical distributions, Fig. 8, main text.**

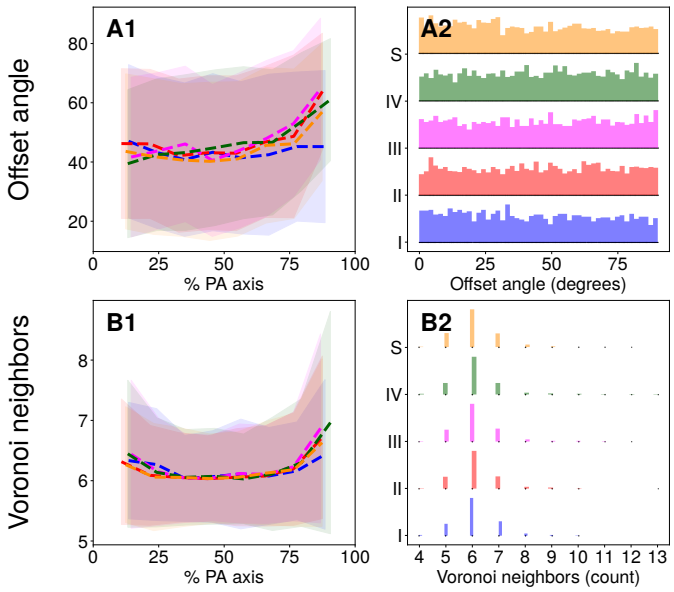


**Fig. S10. Additional invariances in the stochastic geometry of the CZ3 compartments across the life cycle.** Rows A and B follow the same formatting as Fig. 8, main text. Row A shows variation in the offset angle $\theta_{\text{cell}} = \arccos(|\Delta \boldsymbol{x} \cdot \boldsymbol{v}| \,/ (\|\Delta \boldsymbol{x}\| \, \|\boldsymbol{v}\|))$ between the somatic cell offset vector $\Delta \boldsymbol{x}$ (Table 1, main text) and the principal stretch axis $\boldsymbol{v}$ of the CZ3 compartment (given by the eigenvector corresponding to $\lambda_{\max}$). Since the latter has no polarity, $\theta_{\text{cell}} \in [0, 90]$ degrees. Panel A1 shows that there is little variation along the PA axis or by life cycle, with the exception of upward tails toward the anterior pole in stages II-IV and S. Panel A2 reveals furthermore that $\theta_{\text{cell}}$ effectively follows a uniform distribution over its support throughout the lifcycle, indicating that there is no correlation between the somatic cell offset vector $\Delta \boldsymbol{x}$ and the principal stretch axis. This is perhaps surprising given that one might naively expect translation of the cell along the deformation axis of its compartment. Row B shows, in the same format, the number of neighbors (edges) of each Voronoi partition. Panel B1 again shows that there is almost no variation along the PA axis or by life cycle stage, providing quantitative evidence that the topology of the CZ3 space partition does not change during growth. Panel B2 displays the distribution of Voronoi neighbors by life cycle stage, whose mean is around 6, consistent with Euler's theorem for Voronoi tessellations of the sphere (§D).

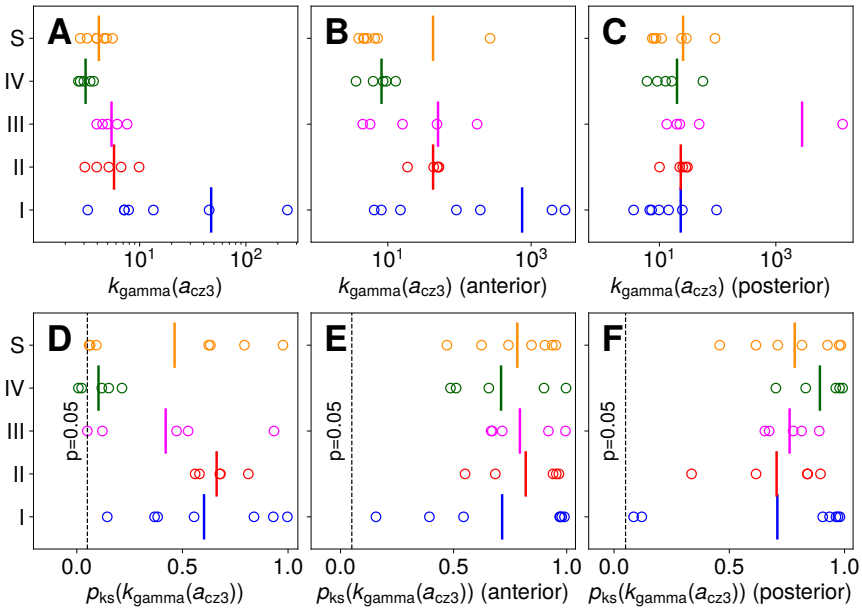


**Fig. S11. Increasing area polydispersity during the life cycle is primarily due to changes in the anterior hemisphere.** Panel A is the same as Fig. 10B (main text) while panels B and C show the same metrics separated by anterior and posterior regions respectively. As seen, B shows a decrease in $k$ (in increase in disorderedness as explained in main text) of the anterior configuration by life cycle stage, while C shows relatively stable values in the posterior configuration. The former therefore likely underlies the increasing disorderedness observed in the whole spheroid (panel A). Panels D-F display the respective $p$-values ($p_{\text{ks}}$) of a Kolmogorov-Smirnov goodness-of-fit test performed in each instance of the $k_{\text{gamma}}$ value reported in panels A-C. All panels D-F show that stage I can exhibit varying degrees of goodness-of-fit, which is likely reflected by the large spread of $k$-values in stage I (panels A-C). However, in the progression from stages II-IV and S, $p_{\text{ks}}$ drops significantly, sometimes below the threshold (plotted as dashed vertical lines), as expected from Fig. S8, despite the qualitatively good match observed in row 1 of that figure. As in Fig. S8 rows 2-3, this is remedied by separating into anterior/posterior hemispheres, at which point the data collapses well, indicated by the sharp increases in $p_{\text{ks}}$ in panels E-F. Remarkably, this increase in goodness-of-fit accompanies a clear observation in panels B-C that the anterior hemisphere is primarily responsible for increasing disorder in the somatic CZ3 configuration.

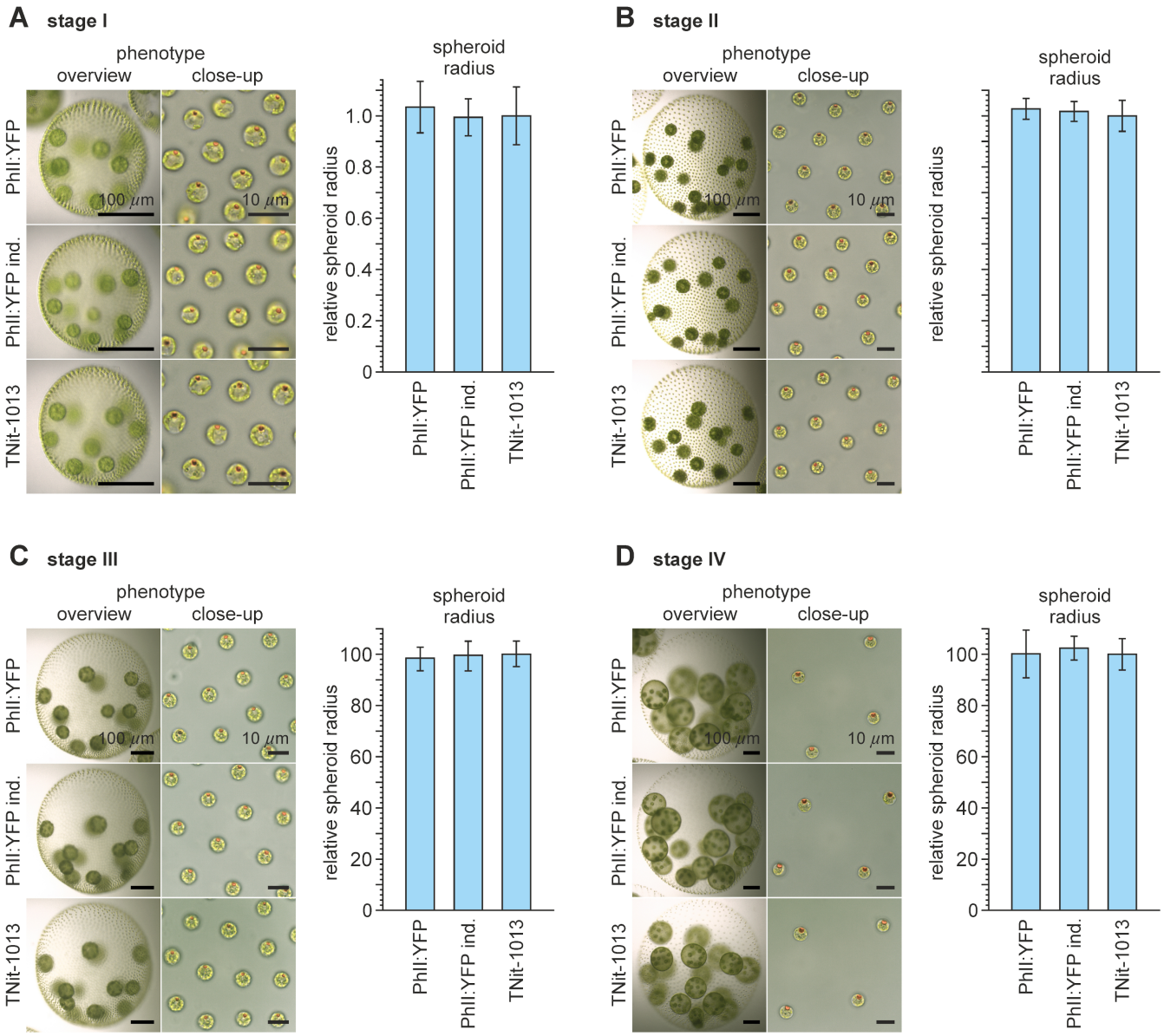


**Fig. S12. Phenotypes of PhlI:YFP-expressing transformant and recipient (control) *V. carteri* strains at developmental stages I to IV.** The phenotypes of transformant PhlI:YFP and the recipient TNit-1013 (control) strains were compared to check whether the expression of the PhlI:YFP fusion protein alters the phenotype. It was also investigated, whether the addition of the sex-inducer protein for only 24 h (inducer added, ind.), i.e. for the period used in the experiments, already leads to a change in the vegetative phenotype of PhlI:YFP transformants. (A-D) Developmental stages I to IV. Representative light microscopic images on the left show overviews and close-ups. The overviews reveal the spheroid morphology, spheroid size, expansion kinetics, the arrangement of the progeny and the regularity of the arrangement of the somatic cells. The close-ups show the regularity of the arrangement of the somatic cells in detail, the size of the somatic cells and the orientation of the eyespots. Bar charts on the right show the analysis of the spheroid radii. The average spheroid radius is shown in relation to the recipient (control) strain TNit-1013 (=1.0). The standard deviation is indicated (n=20). There is no evidence of changes in phenotype due to the expression of PhlI-YFP or the addition of the sex-inducer protein for 24 h.

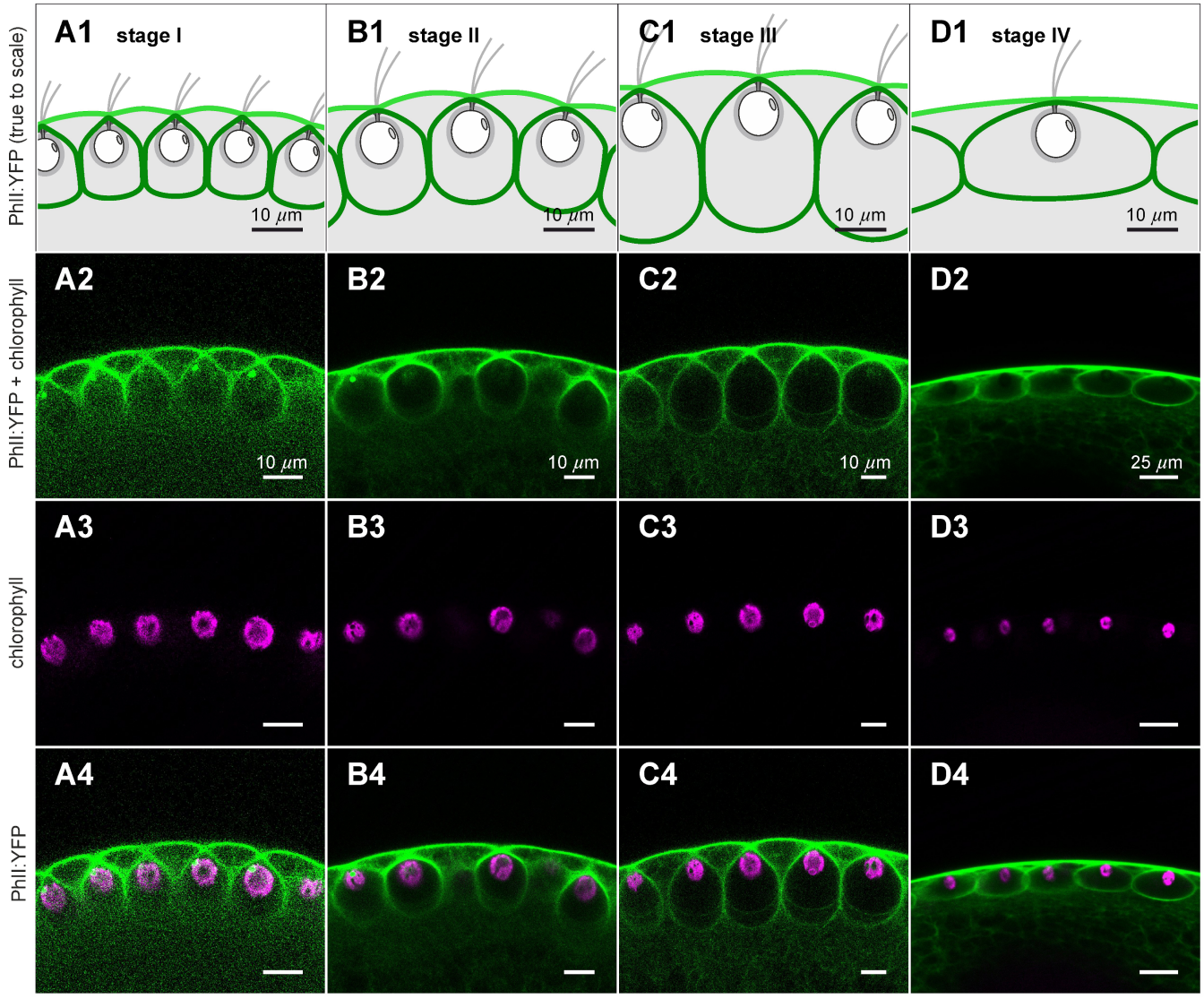


**Fig. S13. Cross sections through PhII:YFP-stained somatic CZ3 compartments at developmental stages I to IV.** Sexually induced transformants expressing the *phII:yfp* gene under the control of the endogenous *phII* promoter were analyzed in vivo in developmental stages I-IV (columns 1-4) for the localization of the PHII:YFP fusion protein. Side view orthogonal to the geometrically analyzed top view showing the shapes of the CZ3 compartments along the Z-axis of the top view. PhII:YFP is located in the CZ3 of somatic cells and in the BZ. It should be noted that when looking at all cLSM images taken as a whole, a considerable proportion of the "floors" of the chambers are poorly recognizable or appear as a diffuse, unclear boundary. Row 1 (A1-D1): True-to-scale schematic cross section, showing PhII:YFP localization in CZ3 of somatic cells (dark green) and BZ (light green). The somatic cells and further ECM material are shown in white or gray. Row 2 (A2-D2): YFP fluorescence of the PhII:YFP protein (green). Row 3 (A3-D3): Chlorophyll fluorescence (magenta). Row 4 (A4-D4): Overlay of YFP and chlorophyll fluorescence.

| Stage | Time (h) | Cross-sectional major axis length ($\mu$m) | Cross-sectional minor axis length ($\mu$m) | Cross-sectional area ($\mu$m$^2$) | Cross-sectional aspect ratio | Cross-sectional elongation | Ellipsoidal est. volume ($\mu$m$^3$) |
|---|---|---|---|---|---|---|---|
| I | 0 | 14.06 | 11.82 | 130.90 | 1.20 | Radial | 1045.19 |
| II | 15 | 19.45 | 17.57 | 268.30 | 1.11 | Radial | 3145.50 |
| III | 21 | 30.20 | 23.50 | 557.51 | 1.29 | Radial | 8750.71 |
| IV | 37 | 38.91 | 18.75 | 572.54 | 0.48 | Tangential | 14892.58 |

**Table S4. Estimates of cross-sectional shape from developmental stages I to IV. As seen from Fig. S13 columns B-C, compartment cross sections can form double-walls. We use the outer wall in these estimates, which also appears of generally higher fluorescent intensity. Compartment cross-sections are segmented (where feasible in the sections shown in Fig. S13) into polygons whose major and minor axis lengths are computed from the second area moment as in §2.B.2, reported in columns 3 and 4. Columns 3-8 are averages across the polygons within each image. Column 6 indicates the aspect ratio defined as the ratio of radial and tangential axes. Column 7 indicates whether the compartment's major axis is oriented radially or tangentially with respect to the spheroid. Column 8 is an estimate of the compartment's 3D volume as an ellipsoid with cross-sectional axes as in columns 3 and 4, revolved around the radial axis.**

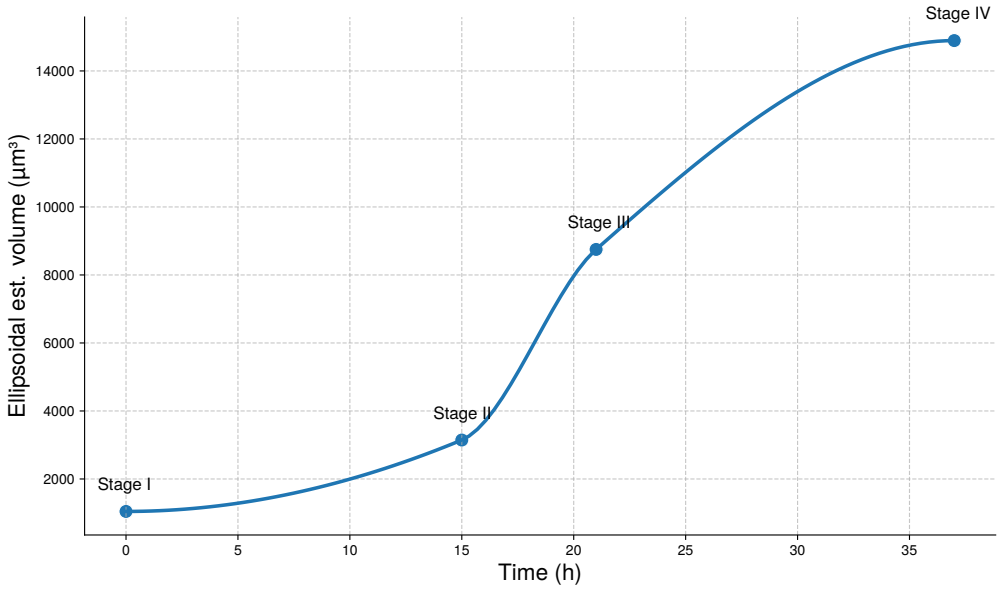


**Fig. S14. Interpolated volumetric growth curve estimated from cross-section.** Columns 2 (time) and 8 (estimated compartment volume) from Table S4 are shown with a monotone interpolant (piecewise cubic Hermite polynomial). As seen, estimated volumetric growth accelerates between Stages II and III, displaying an overall sigmoid-like shape.

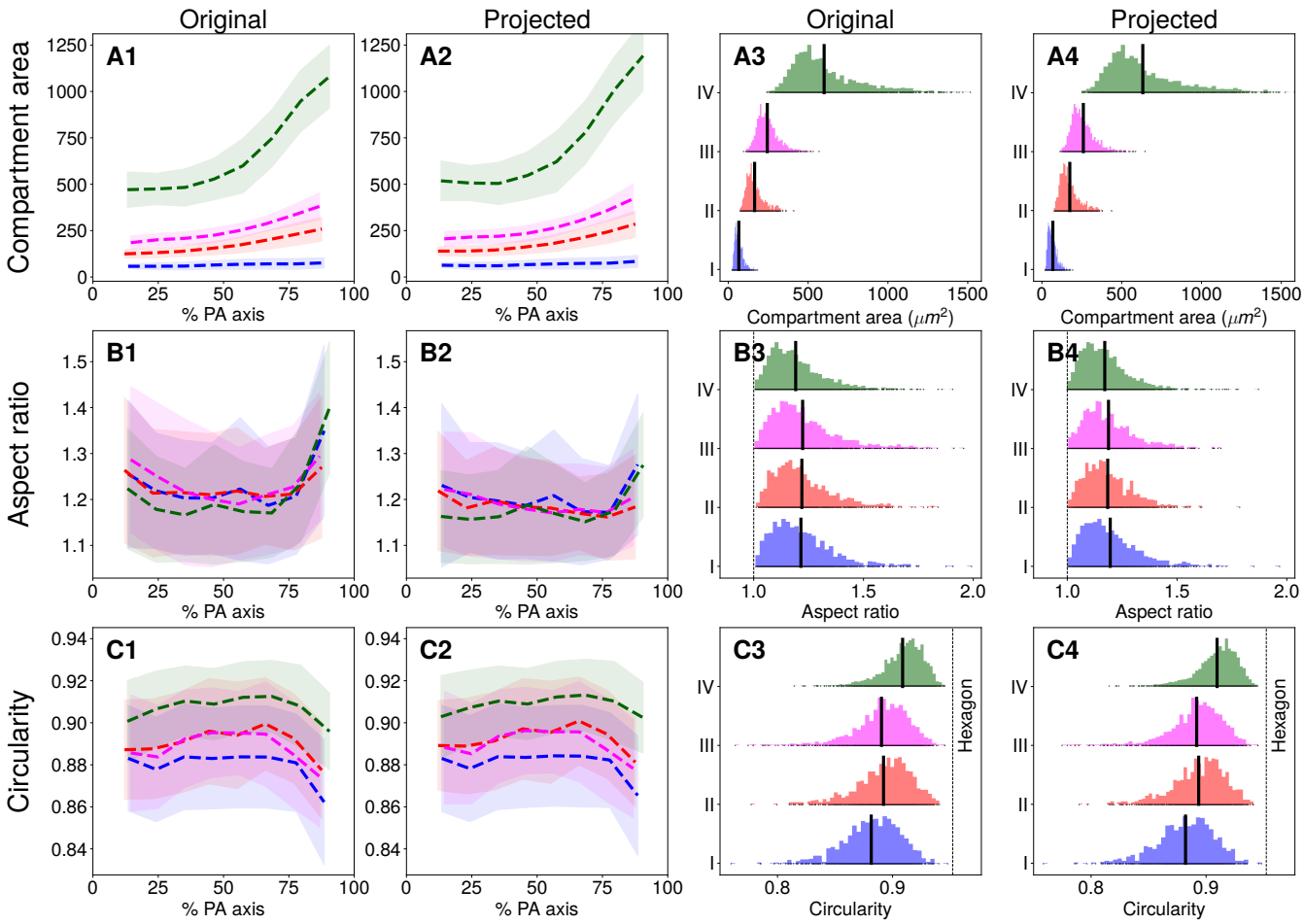


**Fig. S15. Comparison of original and projected compartment metrics.** From the subset of spheroids in which accurate extraction of the 3D geometry is possible as described in §2.E.1 (in particular none from stage S), the distortion correction procedure described in §2.E.3 is applied to the compartment polygons and plots identical to those in Fig. 8, main text, are shown for these polygons. X- and Y-axes are matched between "Original" and "Projected" columns. As seen, the only noticeable variation between the original and projected compartment features are in the extreme anterior region, particularly in stage IV, with respect to area and aspect ratio. Area (A1-2) slightly increases in the projected compartments compared to the original. Aspect ratio (B1-2), which showed upward spikes at the anterior and posterior extremes in the original geometry (B1) now shows a more flat response throughout the AP axis in the projected geometry (B2), as expected. Other than a loss of right outliers in aspect ratio (B3-4) and left outliers in circularity (C3-4) the distributions of these features (columns 3-4) is essentially unchanged between the original and projected geometries. In particular, the aspect ratio distributions are even more stable in mean in the projected geometry relative to the original geometry, from which we first made this observation (main text).

| Stage | Compartment area (original) ($\mu m^2$) | Compartment area (projected) ($\mu m^2$) | Aspect ratio (original) | Aspect ratio (projected) | Circularity (original) | Circularity (projected) |
|---|---|---|---|---|---|---|
| I | 65.69 | 68.53 | 1.22 | 1.20 | 0.88 | 0.88 |
| II | 164.56 | 174.71 | 1.22 | 1.18 | 0.89 | 0.89 |
| III | 244.12 | 259.56 | 1.22 | 1.19 | 0.89 | 0.89 |
| IV | 600.21 | 632.07 | 1.19 | 1.17 | 0.91 | 0.91 |

**Table S5. Quantitative comparison of original and projected compartment metrics by mean. Reflecting the black bars showing empirical means in Fig. S15, we find minimal change in aspect ratio and circularity between original and projected geometries. Compartment area shows consistently around a 5% increase from the original to the projected geometry.**

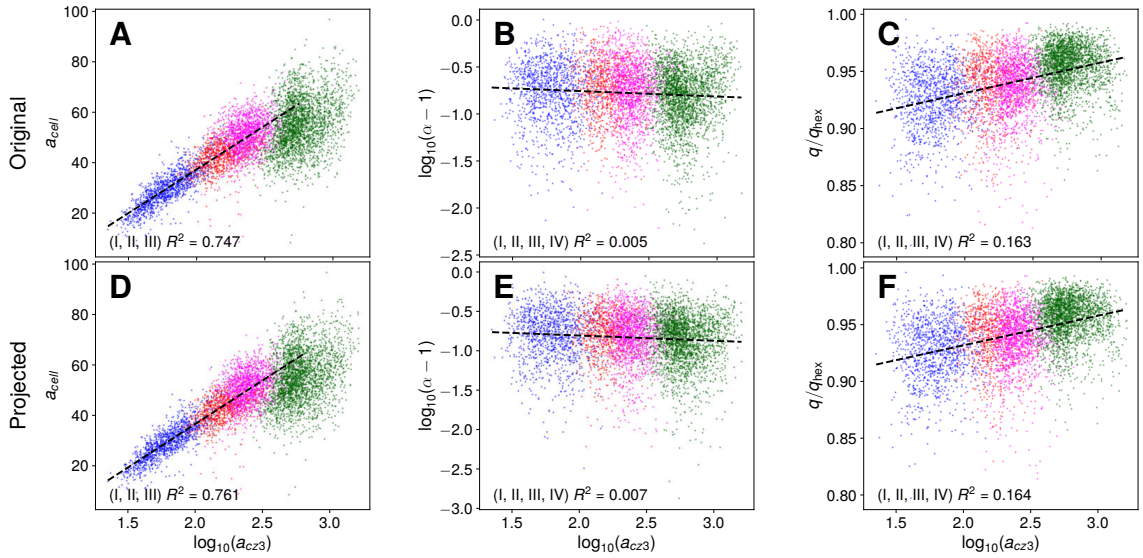


**Fig. S16. Comparison of feature correlations in original and projected compartment geometries.** Correlations between compartment area ($a_{cz3}$) and various metrics (cell area ($a_{cell}$), aspect ratio ($\alpha$), and circularity ($q$) in column order) are shown, as in the main text Fig. 9. As shown, minimal change is present from the original to the projected compartment geometries, with a slight increase in correlation between cell and compartment area in the projected geometry (D), which is consistent with the slight increase in compartment area as seen in Fig. S15 and Table S5. As discussed in §2.E.3, projection is performed only for compartment geometry, as cellular cross-sections are not deformed by the mounting procedure; $a_{\mathrm{cell}}$ is identical in panels A and D.

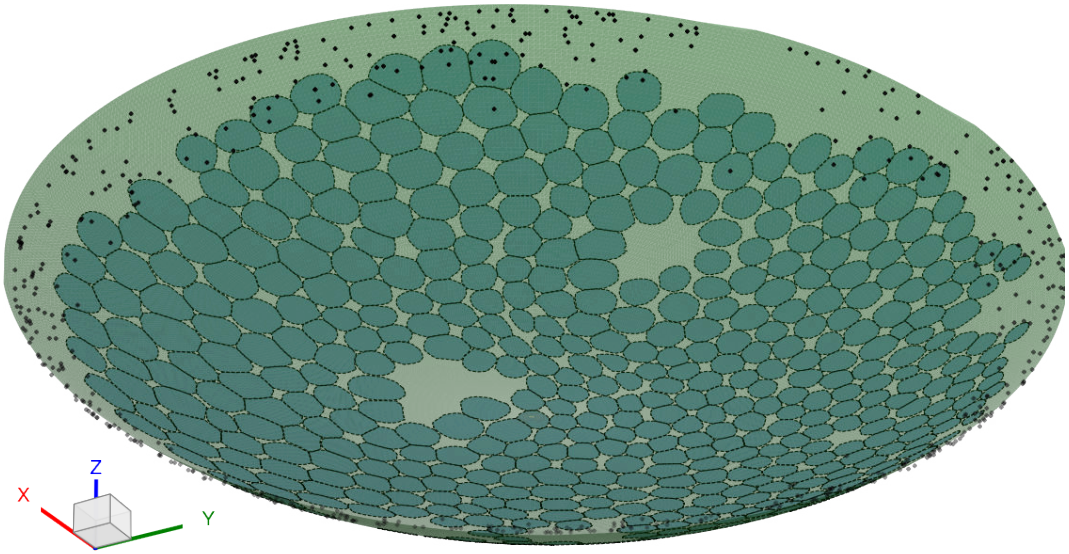


**Fig. S17. Elastic distortion correction by 3D projection of surface features onto ellipsoidal cap**. Black points are the vertices of the yfp 3D convex hull defined in §E.1. The green shell is the ellipsoidal cap defined in §E.2. Blue regions are the compartment geometry projected onto this surface as defined in §E.3. As seen, the inferred ellipsoidal sections are in good agreement with hull points extracted from 3D fluorescence intensity data. The region in the middle in which few hull points appear is the planar section of compression of the organism against the glass slide, as seen from top view in Fig. S5. The ellipsoidal cap (green) "re-inflates" this region. The shown data is from a spheroid in stage IV, in which we expect the greatest amount of distortion due to the organism's size.

## References


1. B von der Heyde, A Hallmann, Targeted migration of pherophorin-S indicates extensive extracellular matrix dynamics in *Volvox carteri*. *The Plant J.* **103**, 2301–2317 (2020).
2. SE Prochnik, et al., Genomic analysis of organismal complexity in the multicellular green alga *Volvox carteri*. *Science* **329**, 223–226 (2010).
3. DM Goodstein, et al., Phytozome: a comparative platform for green plant genomics. *Nucleic acids research* **40**, D1178–D1186 (2012).

4. B Klein, D Wibberg, A Hallmann, Whole transcriptome RNA-Seq analysis reveals extensive cell type-specific compartmentalization in *Volvox carteri*. *BMC Biol.* **15**, 111 (2017).
5. A Hallmann, Extracellular matrix and sex-inducing pheromone in *Volvox*. *Int. Rev. Cytol.* **227**, 131–182 (2003).
6. S El-Gebali, et al., The pfam protein families database in 2019. *Nucleic acids research* **47**, D427–D432 (2019).
7. SC Potter, et al., Hmmer web server: 2018 update. *Nucleic acids research* **46**, W200–W204 (2018).
8. KJ Lauersen, O Kruse, JH Mussgnug, Targeted expression of nuclear transgenes in *Chlamydomonas reinhardtii* with a versatile, modular vector toolkit. *Appl. Microbiol. Biotechnol.* **99**, 3491–3503 (2015).
9. C Stringer, T Wang, M Michaelos, M Pachitariu, Cellpose: a generalist algorithm for cellular segmentation. *Nat. Methods* **18**, 100–106 (2021).
10. C Stringer, M Pachitariu, Cellpose3: one-click image restoration for improved cellular segmentation. *bioRxiv* (2024) Preprint.
11. P Milanfar, GC Verghese, WC Karl, AS Willsky, Reconstructing polygons from moments with connections to array processing. *IEEE Transactions on Signal Process.* **43**, 432–443 (1995).
12. G Strang, *Introduction to Linear Algebra, Sixth Edition.* (Wellesley-Cambridge Press, Philadelphia, PA), (2022).
13. N Fusco, F Maggi, A Pratelli, The sharp quantitative isoperimetric inequality. *Annals Math.* pp. 941–980 (2008).
14. G Polya, G Szegö, *Isoperimetric Inequalities in Mathematical Physics.* (Princeton University Press, Princeton), (1951).
15. L Fejes Toth, Sur la représentation d'une population infinie par un nombre fini d'éléments. *Acta Math. Acad. Sci. Hungarica* **10**, 299–304 (1959).
16. D Newman, The hexagon theorem. *IEEE Transactions on Inf. Theory* **28**, 137–139 (1982).
17. N Otsu, , et al., A threshold selection method from gray-level histograms. *Automatica* **11**, 23–27 (1975).
18. WE Lorensen, HE Cline, Marching cubes: a high resolution 3D surface construction algorithm in *Seminal graphics: pioneering efforts that shaped the field, Volume 1.* (Association for Computing Machinery, New York, NY, USA) Vol. Volume 1, pp. 347–353 (1998).
19. L Atia, et al., Geometric constraints during epithelial jamming. *Nat. Phys.* **14**, 613–620 (2018).