

A multicellular developmental program in a close animal relative

<https://doi.org/10.1038/s41586-024-08115-3>

Marine Olivetta^{1,2}, Chandni Bhickta¹, Nicolas Chiaruttini³, John Burns⁴✉ & Omaya Dudin^{1,2}✉

Received: 29 April 2024

Accepted: 25 September 2024

Published online: 6 November 2024

 Check for updates

All animals develop from a single-celled zygote into a complex multicellular organism through a series of precisely orchestrated processes^{1,2}. Despite the remarkable conservation of early embryogenesis across animals, the evolutionary origins of how and when this process first emerged remain elusive. Here, by combining time-resolved imaging and transcriptomic profiling, we show that single cells of the ichthyosporean *Chromosphaera perkinsii*—a close relative that diverged from animals about 1 billion years ago^{3,4}—undergo symmetry breaking and develop through cleavage divisions to produce a prolonged multicellular colony with distinct co-existing cell types. Our findings about the autonomous and palintomic developmental program of *C. perkinsii* hint that such multicellular development either is much older than previously thought or evolved convergently in ichthyosporeans.

The evolution of multicellular organisms from their unicellular ancestors marks a major transition in the history of life on Earth^{5–7}. This transition was accompanied by fundamental developmental challenges such as generating diverse cell types, forming three-dimensional tissues and establishing overall coordination to drive body plan formation. Asymmetric cell division contributes to cellular diversity^{8,9}, and the formation of three-dimensional tissues relies on the precise coordination of cell divisions, adhesion and signalling^{10–12}. Compared to other eukaryotic lineages, animal multicellularity stands out by relying on an autonomous developmental program, predominantly driven by intrinsic signals, that drives the emergence of an extensive variety of cell types from a single-celled zygote^{13–15}. Although some processes of embryogenesis are found in plants, fungi and various algae, the sheer diversity of specialized cell types in animals highlights a unique aspect of their evolutionary trajectory. Throughout early animal embryogenesis, several ordered processes take place, including cleavage divisions, axis establishment, zygotic genome activation and spatial organization of germ layers. This sequence, although it displays remarkable conservation across species, exhibits a degree of developmental plasticity, highlighting the adaptability of development to distinct ecological pressures^{1,16}. Despite this plasticity, the fundamental aspects of this developmental sequence indicate that parts of the underlying program may have originated before the emergence of animals themselves¹⁷. One possibility is that embryogenesis represents an animal innovation; alternatively, it might have evolved from pre-existing developmental processes present in the last common unicellular ancestor. Animals are closely related to choanoflagellates, filastereans, pluriformeans and ichthyosporeans (Fig. 1a)⁴. These lineages not only partly share a genetic toolkit used by animals for development but also can form transient multicellular structures and display various temporary cell stages in distinct environments⁵. Current evidence indicates that the formation of clonal multicellular choanoflagellate colonies^{18–20} and the emergence of filastereans aggregates^{21–23}, both morphologically

distinct from any known embryonic stages from living animals and their fossil relatives²⁴, occur as a facultative response to external chemical cues^{18,19,25–28}. Moreover, although existing results in the choanoflagellate *Salpingoeca rosetta* indicate the presence of morphologically distinct cells in its facultative multicellular colonies^{29,30}, definitive evidence for a coordinated developmental program driving cell type differentiation in any close animal relatives remains lacking. These results argue that the animal intrinsic embryonic program orchestrating three-dimensional cell architecture and spatial cell type differentiation probably evolved simultaneously with the emergence of animal multicellularity. However, this assumption is largely based on studies of choanoflagellates and filastereans, which, with few exceptions^{18,19,21,31}, often overlook the life cycles, cellular physiology and development of other close relatives.

Among close animal relatives, the Ichthyosporea, comprising two primary lineages, Dermocystida and Ichthyophionida (Fig. 1a)^{32,33}, exhibit a variety of life cycles that combine fungal-like characteristics with transient multicellular stages reminiscent of early animal development³³. Most ichthyophionids, including the model *Sphaeroforma arctica*, undergo coenocytic development, characterized by synchronized nuclear divisions without cytokinesis^{34–37}. This process depends on an aceneteriolar microtubule organizing centre that drives a fungal-like closed mitosis³⁸. On reaching a specific nuclear-to-cytoplasmic ratio³⁹, *S. arctica* undergoes an actomyosin-dependent cellularization. During this process, a transient multicellular layer resembling an animal epithelium is formed before the release of new-born cells to repeat the cycle³⁶. In contrast, we recently showed that *C. perkinsii*, the only cultured free-living dermocystid³, undergoes centriole-mediated, animal-like open mitosis coupled with cleavage divisions³⁸. This discovery indicates that *C. perkinsii* proliferates through a palintomic developmental mode, highlighting the vast range of cellular and developmental diversity in the Ichthyosporea. However, nothing is known at present about the spatiotemporal dynamics of the ichthyosporean palintomic life cycle, nor is it clear whether it is orchestrated by an

¹Swiss Institute for Experimental Cancer Research, School of Life Sciences, Swiss Federal Institute of Technology (EPFL), Lausanne, Switzerland. ²Department of Biochemistry, University of Geneva, Geneva, Switzerland. ³Bioimaging and Optics Core Facility, Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland. ⁴Bigelow Laboratory for Ocean Sciences, East Boothbay, ME, USA. ✉e-mail: jburns@bigelow.org; omaya.dudin@unige.ch

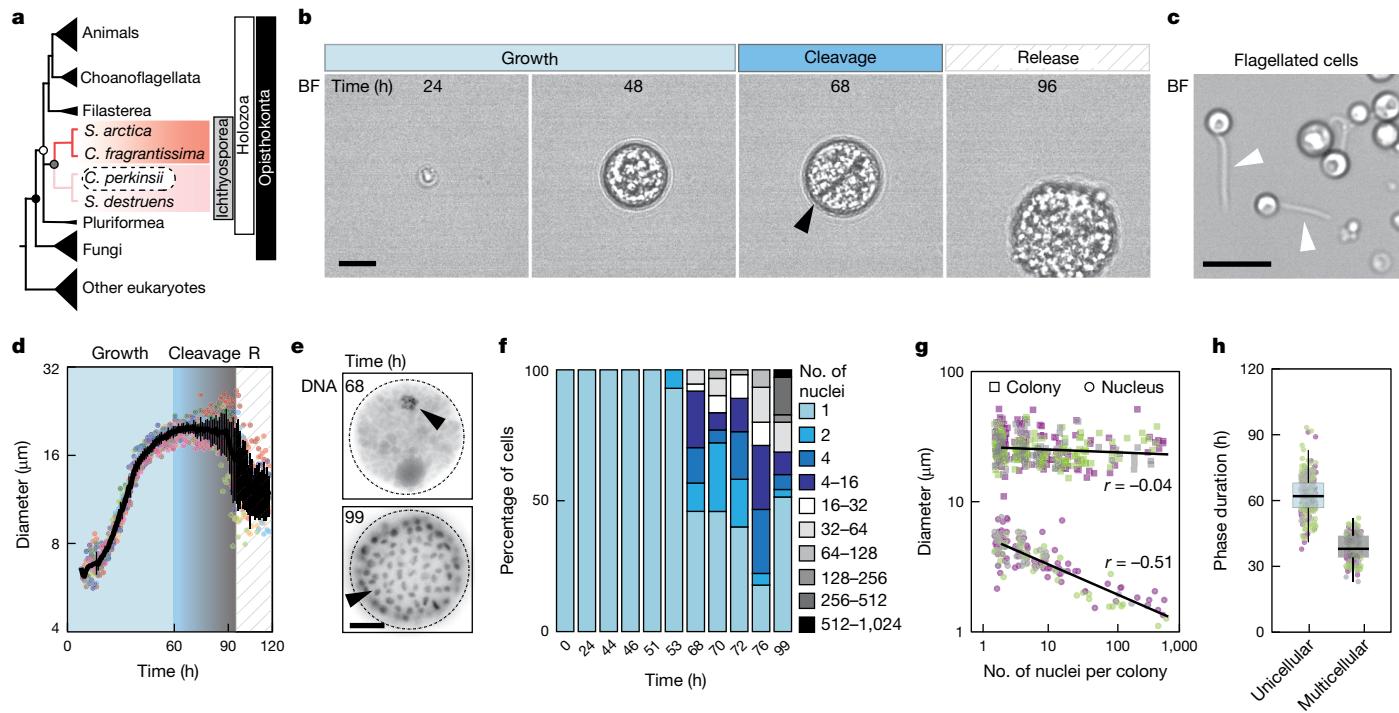


Fig. 1 | *C. perkinsii* undergoes a series of cleavage divisions at constant volume. **a**, Cladogram representing the position of ichthyosporeans, featuring ichthyophionids such as *S. arctica* (red), and dermocystids, including *C. perkinsii* (pink), in the eukaryotic tree. *C. fragrantissima*, *Creolimax fragrantissima*; *S. destruens*, *Sphaerothecum destruens*. **b**, Time-lapse bright-field (BF) images of the life cycle of *C. perkinsii* show cell-size increase, cleavage division at the one- to two-cell stage (arrow) and release of new-born cells (Supplementary Video 1). **c**, Microscopy of flagellated cells observed following cell release (arrows). **d**, Mean cell diameter over time of single-cell traces aligned to the first cell cleavage ($T = 60$ h) shows the duration of distinct cell stages: Growth, cleavage and cell release (R) ($n = 10$). Data are mean \pm s.d. **e**, Microscopy of DNA-stained cells at 68 and 99 h of growth, highlighting the

number and size of nuclei at each time point (arrows). **f**, Distribution of nuclear content of *C. perkinsii* colonies across the life cycle at 23 °C measured by microscopy of Hoechst-stained cells ($n > 82$ per time point). **g**, Distribution of colony (squares) and nucleus (circles) diameters relative to the number of nuclei per colony demonstrates that cellular size remains constant, whereas nuclear size decreases ($n = 1,571$ colonies). Lines represent correlation coefficients, and colours represent three independent replicates. **h**, Average phase duration for the unicellular and multicellular growth phases ($n = 244$ cell per colony). The colours represent three independent replicates. Box plots extend from the 25th to 75th percentile, including the median, and whiskers extend from the minimum to maximum value. Scale bars, 10 μ m.

autonomous developmental program that exhibits any similarities with early animal embryogenesis.

To better characterize the development of *C. perkinsii*, we conducted long-term brightfield live imaging of synchronized cell populations (Fig. 1b–d, Extended Data Fig. 1a–c and Supplementary Video 1). In spite of the light sensitivity of *C. perkinsii* (Methods), our results show that synchronized single cells undergo approximately 65 h of growth until the first cell cleavage (Fig. 1b,d and Extended Data Fig. 1b). Following this initial division, the now multicellular colonies maintain a constant size for a further 30 h period, which is followed by the release of hundreds of new-born cells (Fig. 1b–d, Extended Data Fig. 1e and Supplementary Video 1). Interestingly, released cells were of two types: first, proliferating cells without flagella and capable of division, termed ‘proliferative/mitotic’ (Supplementary Video 2); and second, small single cells with a flagellum that exhibited flagellar motility, named ‘flagellates’ (Fig. 1c and Supplementary Video 3). Notably, some flagellated cells remained immobile with dynamic membrane protrusions and were thus designated ‘amoeboflagellate’ (Supplementary Video 2). Counting flagellated cells shows an increase in their number around the time of cell release (more than 76 h), yet they represent less than 15% of all observed single cells (Extended Data Fig. 1d), indicating that only a fraction of cells differentiates into flagellated cells. These observations indicate that *C. perkinsii* differentiates into at least two distinct cell types, with flagellated cells potentially transitioning between flagellate and amoeboflagellate states (see also Fig. 4 and Extended Data Fig. 4).

To gain further insights into *C. perkinsii* development, we fixed and stained the nuclei of synchronized *C. perkinsii* cells throughout their

entire life cycle (Fig. 1e–g). We observe that cells exhibit similar size growth dynamics, although with increased variability, which can be attributed to heterogeneity in the cell population (Extended Data Fig. 1f). On counting the number of nuclei per cell, we find that after the first division (about 65 h), the nuclear number rapidly increases, surpassing 500 nuclei per colony before cell release (Fig. 1e–g and Extended Data Fig. 1g). As this significant and rapid increase in nuclear content occurs at constant colony size, we note it is accompanied by a gradual reduction in nuclear diameter (Fig. 1e,g). These results show that *C. perkinsii* develops through successive cleavage divisions while maintaining a constant overall colony volume (Fig. 1g). This multicellular colony stage is maintained for a substantial duration, accounting for about 33% of the life cycle in culture (Fig. 1h). In contrast, the epithelium-like multicellular stage in the ichthyosporean model *S. arctica* accounts for only 2% of its life cycle^{36,39}.

To better understand whether *C. perkinsii*’s autonomous program is associated with a distinct transcriptional signature, we used time-resolved transcriptomic profiling. We isolated and sequenced mRNA from two independent synchronized cultures, covering key time points from the single-cell stage through the multicellular stage until cell release (Extended Data Fig. 2a and Extended Data Table 1). We identify 7,773 transcripts that are differentially expressed (false discovery rate less than 0.05) across these stages, representing 62% of *C. perkinsii*’s total genes. These genes clustered into five distinct expression patterns throughout development (Fig. 2a and Extended Data Fig. 2a). These patterns highlight the differences between the unicellular phases both before cleavage divisions (54 h) and after release

Article

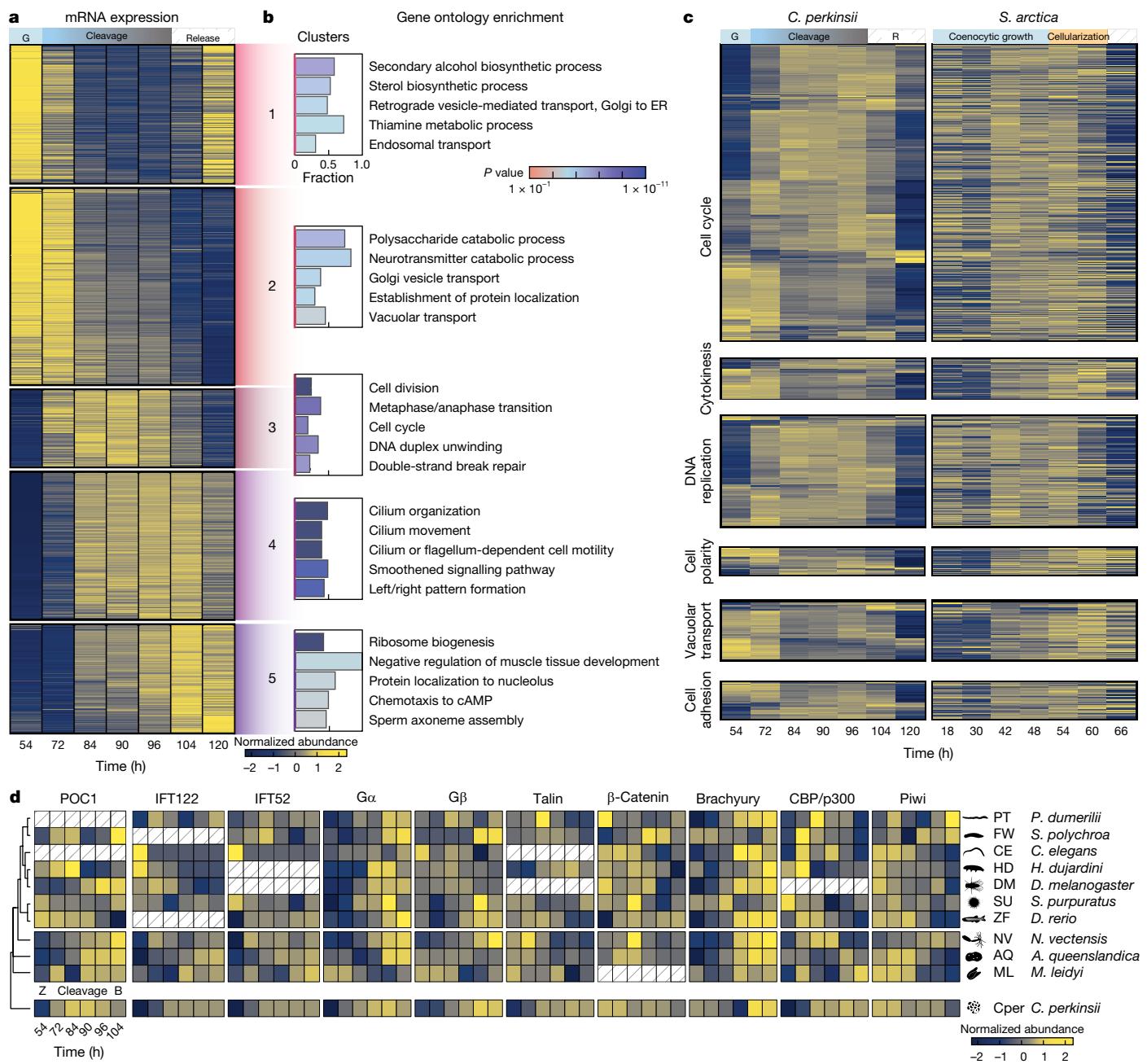


Fig. 2 | Transcriptional dynamics across the palintomic life cycle of *C. perkinsii*. **a**, A heatmap of 7,773 coding genes clustered into five dynamic clusters. **b**, Gene ontology enrichment (one-sided Fisher's exact test with 'weight' algorithm from topGO⁶⁷) of the five different dynamic clusters highlighting the patterned transcriptional program of *C. perkinsii* development linked to Supplementary Tables 1 and 2. ER, endoplasmic reticulum. **c**, A comparative heatmap of gene expression dynamics of orthologs between the palintomic *C. perkinsii* and the coenocytic *S. arctica*³⁵ life cycles for key gene ontology categories, including cell cycle (GO:0007049), cytokinesis (GO:0000910), DNA replication (GO:0006260), cell polarity (GO:0007163), vacuolar transport (GO:0007034) and cell adhesion (GO:0098609). It highlights the distinct transcriptional programs between both ichthyosporean species. The heatmaps for *C. perkinsii* and *S. arctica* have the same row order, allowing direct comparison of gene expression patterns of orthologs between the two species. **d**, A comparative heatmap of gene expression for key selected regulators involved in flagellar motility (POC1, IFT122, IFT52), cell signalling ($\text{G}\alpha$, $\text{G}\beta$), adhesion (Talin, β -Catenin) and transcriptional regulation of animal development (Brachyury, CBP/p300, Piwi) dynamics between the *C. perkinsii* (Cper) life cycle and early embryonic development stages of

early-branching animals. This comparison spans from the onset of zygote formation (Z) through various cleavage stages and up to the blastula (B) stage. G, growth. *D. melanogaster*, *Drosophila melanogaster*; *H. dujardini*, *Hypsibius dujardini*; *P. dumerilii*, *Platyneris dumerilii*. Credits: Silhouettes were obtained from PhyloPic (<https://www.phylopic.org/>). *Schmidtea polychroa*, created by M. A. Grohme under a CC0 1.0 Universal Public Domain licence; *Caenorhabditis elegans*, created by B. Goldstein, vectorization by J. Warner under a CC0 1.0 Universal Public Domain licence; *Isohypsibius dastychi*, created by B. Lang under a CC0 1.0 Universal Public Domain licence; *Sophophora melanogaster*, created by A. Wilson under a CC0 1.0 Universal Public Domain licence; *Strongylocentrotus purpuratus*, created by C. Schomburg under a CC0 1.0 Universal Public Domain licence; *Danio rerio*, created by Ian Quigley under a CC BY 3.0 licence; *Nematostella vectensis*, created by J. Warner under a CC0 1.0 Universal Public Domain licence; *Amphimedon queenslandica* obtained under a CC0 1.0 Universal Public Domain licence; *Mnemiopsis leidyi*, created by J. R. Winnikoff under a CC0 1.0 Universal Public Domain licence; *Creolimax*, created by Y. Wong using scanning electron microscopy images by A. Sebé-Pedrós (public domain agreed by I. Ruiz-Trillo) under a licence free of copyright, CC PD-M 1.0.

(104 and 120 h), compared to the multicellular phase (72–96 h) (Fig. 2a and Extended Data Fig. 2b). Gene ontology enrichment analysis shows that unicellular phase clusters (1, 2 and 5) are enriched in gene ontology terms related to nutrient and lipid biosynthesis and polysaccharide catabolism (Fig. 2b and Supplementary Tables 1 and 2), whereas multicellular phase clusters (3 and 4) show enrichment in cell division and flagellar differentiation (Fig. 2b). Therefore, before the first division, *C. perkinsii* produces and accumulates key cellular components such as proteins and lipids in a manner reminiscent of oocyte maturation⁴⁰. Following the first division, these biosynthetic and catabolic processes are turned off, indicating a gradual consumption of previously synthesized nutrients as the palintomic development progresses (Fig. 2b and Extended Data Fig. 2a), and only reappear around cell release. In contrast, the expression of genes driving catabolic processes remains low after release (cluster 2), probably due to nutrient depletion from the medium (Fig. 2a,b and Extended Data Fig. 2b). Towards the end of the multicellular phase, a distinct transcriptional pattern emerges (cluster 4), with enrichment in genes for cilia/flagella formation, motility and left/right patterning in animals (Supplementary Tables 1 and 2).

To determine if the developmental program of *C. perkinsii* is distinct among ichthyosporeans, we compared it with the transcriptional signature of the ichthyosporean *S. arctica*, which undergoes coenocytic growth followed by cellularization. We find that key gene ontology functional categories including the cell cycle, cytokinesis, DNA replication, polarity and differentiation do not overall correlate between *C. perkinsii* and *S. arctica*, indicating distinct transcriptional programs (Fig. 2c and Extended Data Fig. 2c,d). Despite these differences, certain cytoskeletal regulators previously highlighted and required in the *S. arctica* epithelium-like cell layer³⁶, such as formin 2, critical for actin nucleation, correlate between both species regardless of developmental type, indicating a housekeeping role (Extended Data Fig. 2e). In contrast, formin 5 exhibits an inverse correlation, implying a species-specific role in actin dynamics (Extended Data Fig. 2e). Next, we sought to examine whether the transcriptional program of *C. perkinsii* exhibits similarities with that of early embryonic development in early-branching animals. To accomplish this, we categorized genes from animals and ichthyosporeans into orthogroups: clusters of genes that originated from a single gene in their last common ancestor. We then analysed the expression patterns of these genes from the onset of zygote formation through the various cleavage divisions and up to the blastula stage, as previously described⁴¹ (Methods). By calculating the average correlation of gene expression among orthogroups across time points for *C. perkinsii* and the three earliest-branching animals—the ctenophore *Mnemiopsis leidyi*, the sponge *Amphimedon queenslandica* and the cnidarian *Nematostella vectensis*—we find an overrepresentation of positively correlated expression patterns indicating similarities in the transcriptional program between *C. perkinsii* and these three early-branching animals (Extended Data Fig. 2g). We observe a similar, although smaller, overrepresentation in positively correlated expression patterns between *C. perkinsii* and all animals (Extended Data Fig. 2h). These differences are statistically significant (P less than 2.2×10^{-16} by Kruskal–Wallis H -test for the overall effect), and each developmental cluster shows a significantly higher correlation between *C. perkinsii* and animals than non-developmentally regulated genes (Extended Data Fig. 2i–k; cluster 1: $P = 1.18 \times 10^{-15}$, cluster 2: $P = 4.54 \times 10^{-14}$, cluster 3: $P = 5.36 \times 10^{-49}$, cluster 4: $P = 1.13 \times 10^{-22}$, cluster 5: $P = 2.56 \times 10^{-102}$). By focusing on a select subset of genes important for flagellar formation (POC1, IFT122, IFT52)⁴², cellular signalling (Gα, Gβ, PKAc-α)^{43,44}, cell adhesion (Talin, β-Catenin, Vinculin, Grancalcin)^{45,46} and transcriptional regulation during animal development (Brachyury, CBP/p300, Piwi, Rbl1, P53, RunX)^{3,47,48}, we can observe comparable patterns of gene expression between *C. perkinsii* and early-diverging animals. In certain cases, this expression exhibits conserved temporal patterns across most lineages (Fig. 2d and Extended Data Fig. 2l).

A statistical analysis of gene expression patterns across various animals and *C. perkinsii* for individual genes shows that certain developmental genes, such as Brachyury ($P = 0.0065$, Kendall's coefficient of concordance), exhibit significant and deeply conserved developmental expression patterns (Supplementary Table 2). Notably, Brachyury, a transcription factor regulating gastrulation, and Piwi, a marker of germline formation, show such conserved expression patterns in many, although not all, animals (Fig. 2d). Although we cannot exclude the potential effect of noise from bulk culture asynchrony, orthogroup identification or developmental time averaging, our results show similar patterns of expression of genes critical for animal embryogenesis during the multicellular development of *C. perkinsii*.

To further explore the development of *C. perkinsii* at the cellular level, we tracked the plasma membrane in synchronized cells using long-term live imaging. Our results show that *C. perkinsii* undergoes a series of ordered cleavage divisions, starting from a single cell and progressing to a two-cell stage and then beyond (Extended Data Fig. 3a and Supplementary Video 4). Ultimately, this yields a spatially organized multicellular colony surrounded by an external cell wall and occasionally exhibiting a central cavity (Extended Data Fig. 3b). Owing to *C. perkinsii* photo-sensitivity and a need for increased resolution, we opted for imaging synchronized cultures at specific developmental times instead of long-term live imaging. We used three approaches: live imaging of membrane-stained cells to comprehend overall spatial organization (Fig. 3a and Supplementary Video 5), immunostaining of actin and nuclei to characterize nuclear positioning (Fig. 3b) and ultrastructural expansion microscopy (U-ExM) of microtubules, crucial for antibody accessibility³⁸, to show spindle orientation during early cleavage divisions (Fig. 3c). Initially, the nucleus of the one-cell stage (Cell A) is centrally located, before it migrates to the cortex, where it undergoes mitosis exhibiting striking polarity at this stage (Fig. 3b–d and Extended Data Fig. 3c–e). This yields two daughter cells (Aa and Ab) of distinct volumes, with Ab being on average about 1.3 times larger in volume (Fig. 3a,e and Extended Data Fig. 3d). The Aa/Ab volume ratio is highly variable, ranging from 1.05 to 2.6, indicating that the first division of *C. perkinsii* is often unequal and displays high stochasticity. Moreover, it is important to note that although we distinguish the two cells (Ab and Aa) solely by their volume difference, which may introduce biases, the observed difference remains statistically significant, with the volume ratio consistently greater than 1 (Wilcoxon signed-rank test, $P = 3.73 \times 10^{-9}$). Following the first cleavage division, we note that the Ab cell undergoes mitosis ahead of the Aa cell (Fig. 3b,c and Extended Data Fig. 3f), resulting in a transient three-cell stage (Fig. 3c, Extended Data Fig. 3g and Supplementary Video 6). This highlights an asynchrony in mitotic entry potentially mediated by the initial differences in volume between the Ab and Aa cells. Furthermore, we observe that the Ab cell undergoes mitosis perpendicular to the first division plane, whereas the Aa cell undergoes mitosis perpendicular to the Ab cell (Fig. 3c,f). Therefore, Aa and Ab orient their spindles and divide perpendicularly to each other, ultimately forming a tetrahedral four-cell stage (Fig. 3b,c,g). Tracking subsequent cleavage divisions proved more challenging; however, we find that the number of cells in the multicellular colonies seems to be variable (Fig. 1f,g and Extended Data Fig. 1g). Despite the need for further ultrastructural characterization of both the Aa and Ab cells to ensure the non-stochastic nature of the process, our results indicate that the development of multicellular *C. perkinsii* colonies is characterized by early symmetry breaking, initially in nuclear localization and subsequently during the first cleavage division, a phenomenon known to drive cell differentiation across different multicellular systems^{49,50}.

Next, we sought to better understand the differentiation program that results in the formation of distinct cell types (Fig. 1c). Using brightfield microscopy, we find that following cell release, the fates of the cells seem distinct within the time frame of the experiment. Proliferative or mitotic cells exhibit a rapid increase in growth without

Article

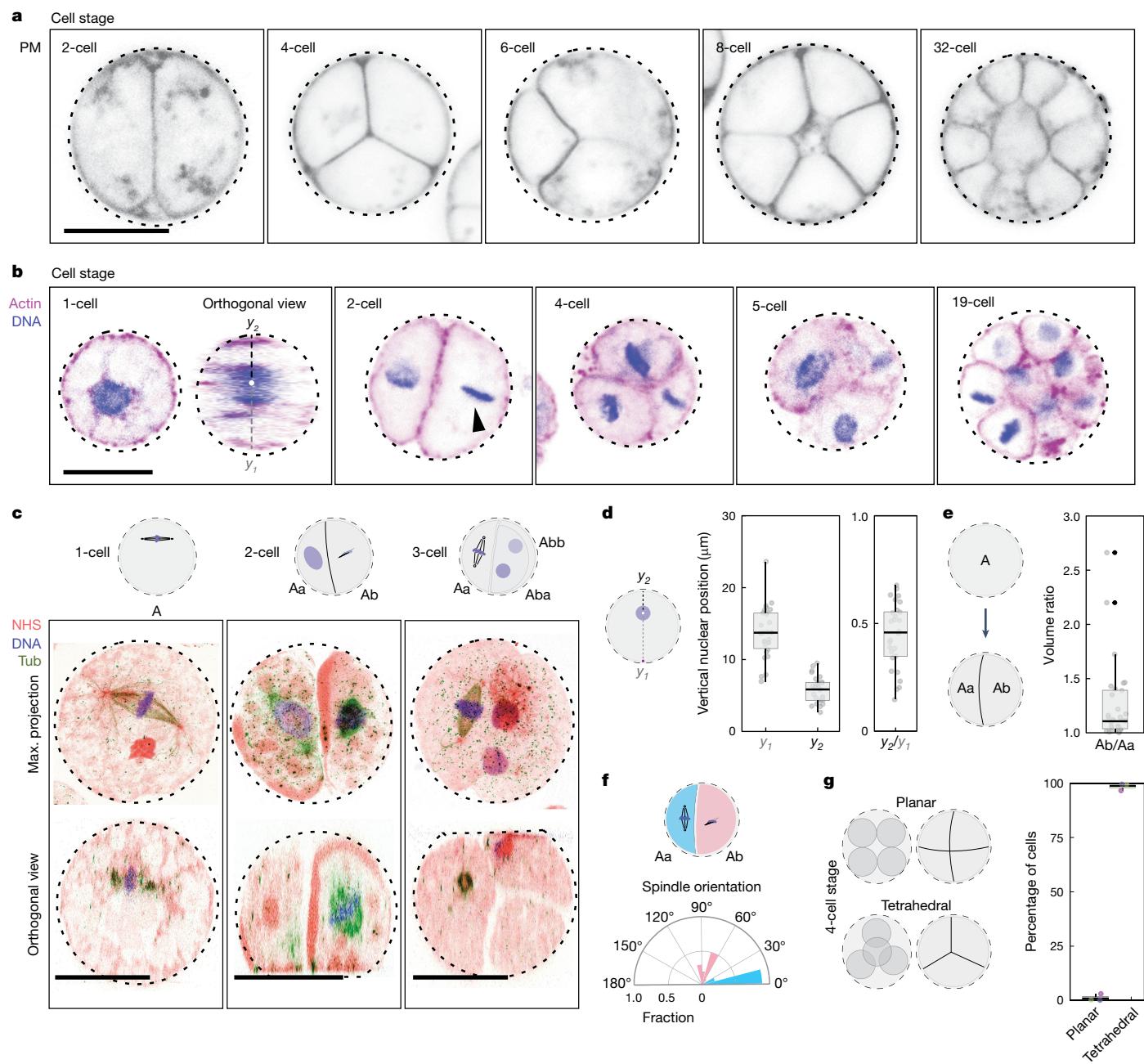


Fig. 3 | Early developmental patterns in *C. perkinsii*. **a**, Plasma membrane-stained (PM) live colonies at distinct cell stages, highlighting the patterned cleavage divisions, tetrahedral four-cell stage and formation of spatially organized multicellular colonies (Supplementary Video 5). **b**, Actin- (magenta) and DNA-stained (blue) colonies at distinct cell stages showcasing nuclear cortical positioning, asymmetrical cell division (in volume and in time) and the formation of a multicellular colony. This result has been reproduced at least three independent times. **c**, U-ExM stained colonies for pan-labeling with NHS-Ester (red), microtubules (green) and DNA (blue), highlighting the first mitotic division at the cortex, perpendicular spindles at the two-cell stage and a three-cell stage (Supplementary Video 6). **d**, Box plot showing the vertical localization of the nucleus at the one-cell stage before or at the first cleavage division highlighting the nuclear cortical localization ($n = 28$ cells). Box plots extend from the 25th to 75th percentile, including the median, and whiskers extend from the minimum to maximum value. **e**, Box plot illustrating the volumetric ratio following the first division resulting in Ab and Aa cells, highlighting the asymmetrical cell division ($n = 28$ two-cell-stage cells). The volume ratio is >1 (one-sided Wilcoxon signed-rank test, $P = 3.73 \times 10^{-9}$). Box plots extend from the 25th to 75th percentile, including the median, and whiskers extend from the minimum to maximum value. **f**, A polar plot representing the spindle angle for the Aa and Ab cells during mitosis demonstrates that the Ab cells divide perpendicular to the first cleavage division, whereas the Aa cells divide parallel to it ($n = 11$ colonies). **g**, Box plot showing the percent of the four-cell stage exhibiting a planar or tetrahedral spatial organization ($n = 108$ colonies). Box plots extend from the 25th to 75th percentile, including the median, and whiskers extend from the minimum to maximum value. Scale bars, $10 \mu\text{m}$. Max., maximum; tub, microtubules.

extend from the minimum to maximum value. **e**, Box plot illustrating the volumetric ratio following the first division resulting in Ab and Aa cells, highlighting the asymmetrical cell division ($n = 28$ two-cell-stage cells). The volume ratio is >1 (one-sided Wilcoxon signed-rank test, $P = 3.73 \times 10^{-9}$). Box plots extend from the 25th to 75th percentile, including the median, and whiskers extend from the minimum to maximum value. **f**, A polar plot representing the spindle angle for the Aa and Ab cells during mitosis demonstrates that the Ab cells divide perpendicular to the first cleavage division, whereas the Aa cells divide parallel to it ($n = 11$ colonies). **g**, Box plot showing the percent of the four-cell stage exhibiting a planar or tetrahedral spatial organization ($n = 108$ colonies). Box plots extend from the 25th to 75th percentile, including the median, and whiskers extend from the minimum to maximum value. Scale bars, $10 \mu\text{m}$. Max., maximum; tub, microtubules.

forming flagella or gaining motility, whereas the cell volume of flagellates remains constant, hinting at distinct developmental trajectories (Extended Data Fig. 4a–e and Supplementary Videos 7 and 8). We then focused on the flagellated cells, which are characterized by a small

nucleus (Extended Data Fig. 4f,g) and are easily identifiable using expansion microscopy with antibodies against both beta-tubulin and acetylated tubulin (Fig. 4a,b and Extended Data Fig. 4f). As flagellar gene expression starts much earlier than cell release (Fig. 2a–c), we

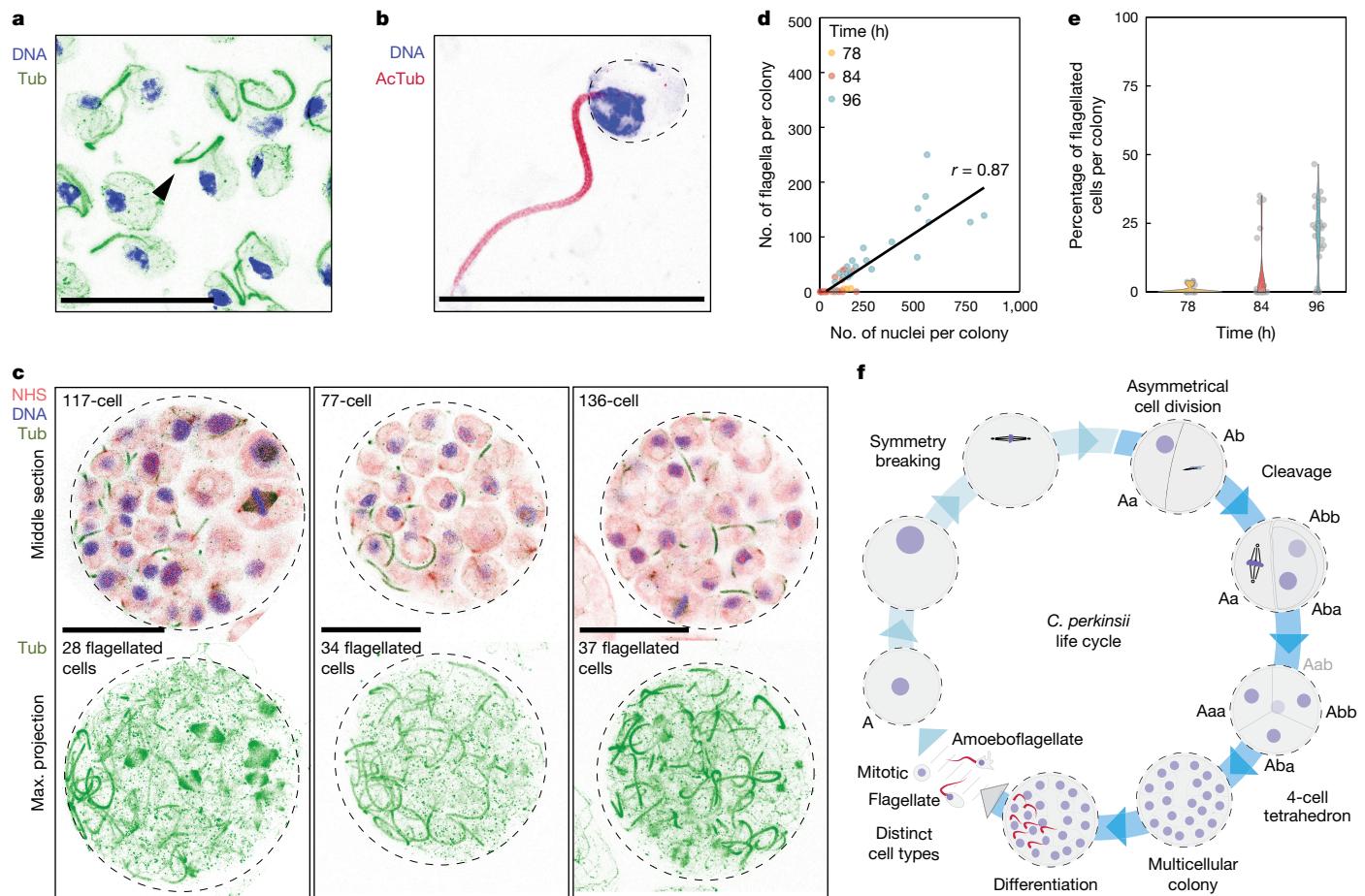


Fig. 4 | Cell differentiation in *C. perkinsii*'s colonies. **a**, U-ExM stained cells for microtubules (green) and DNA (blue), identifying flagellated cells (arrow). This result has been reproduced at least three independent times. **b**, U-ExM stained flagellated cell for acetylated tubulin (AcTub; dark red) and DNA (blue). This result has been reproduced at least three independent times. **c**, U-ExM stained late colonies for pan-labelling with NHS-Ester (red), microtubules (green) and DNA (blue), highlighting the co-existence of flagellated and non-flagellated cells in the multicellular colony (Supplementary Videos 9 and 10). **d**, A dot plot illustrating the number of flagellated cells per colony, which increases with the overall number of cells per colony. Notably, although there is

a positive correlation, the increase is not linear, showing that distinct cell types coexist in the multicellular colonies ($n_{78} = 33$, $n_{84} = 32$, $n_{96} = 27$ colonies per time point). **e**, Violin plot showing the percent of flagellated cells in a multicellular colony at distinct developmental times ($n_{78} = 33$, $n_{84} = 32$, $n_{96} = 27$ colonies per time point). **f**, A model representing the developmental program of *C. perkinsii*, beginning with the relocation of the nucleus to the cortex, symmetry breaking at the 1-to 2-cell stage, formation of a tetrahedral four-cell stage, development into a multicellular colony with intricate three-dimensional architecture, differentiation of a subset of cells into flagellated cells and release of distinct cell types. Scale bars, 10 μ m.

wondered whether the differentiation into flagellated cells occurs post-release, representing a case of temporal cell type, or whether it differentiates already in the multicellular colony. To this end, we used expansion microscopy to label flagella and nuclei in the late multicellular colonies (T78, T84, T96). Remarkably, we detect flagellated cells in the multicellular colonies, confirming that the differentiation process takes place before cell release (Fig. 4c and Supplementary Videos 9 and 10). As development progresses and the total number of cells per colony increases, the number of flagellated cells also rises (Fig. 4c,d). However, this number never exceeds 50% of all cells present in the colony and exhibits significant heterogeneity between colonies; moreover, we noted that flagellated cells often cluster spatially in the colony (Fig. 4c–e, Extended Data Fig. 4h and Supplementary Videos 9 and 10). These results indicate the co-existence of at least two distinct cell types in multicellular *C. perkinsii* colonies.

Discussion

Exploring the evolutionary history that led to the emergence of animal multicellularity has been challenging. This stems notably from the paucity of non-metazoan holozoans fossils and the limitations

stemming from studying only a few closely related protists, primarily focusing on their genomic content and often neglecting their cellular physiology^{5,6,51}. Here we characterize the life cycle of the ichthyosporean *C. perkinsii*, a close animal relative, providing evidence for an intrinsic and clonal multicellular developmental program. Our results show that *C. perkinsii* undergoes palintomic development (Fig. 4f), involving cleavage divisions coupled to open mitosis, diverging from the coenocytic development and cellularization in other studied ichthyosporeans³⁷. The observation that *C. perkinsii* positions the nucleus cortically in the one-cell stage, undergoes asymmetric cell division, forms organized multicellular colonies and spatially differentiates into distinct cell types demonstrates key morphological and transcriptional parallels with the early stages of animal embryogenesis¹. Notably, the development of *C. perkinsii* is stereotypical and occurs independently of apparent external triggers, leading to the formation of colonies harbouring hundreds of cells (Fig. 4f). Such capacity shows that an autonomous cell specification program with seemingly spatial cell differentiation has evolved in *C. perkinsii*, a developmental trait considered as a hallmark of 'complex' multicellular lineages such as animals, land plants, red and brown algae and mushroom-forming fungi⁷. Although this process is at present hypothesized to have emerged through the

Article

co-option of temporally regulated genes for spatially explicit differentiation, this transition occurred deep in evolutionary history, and, to our knowledge, no extant organisms have been identified as an intermediary form. Our findings on *C. perkinsii* potentially align with this model, as spatial differentiation seems evident; however, on the basis of our current understanding, the various cell types do not yet function together in a coordinated manner as known for ‘complex’ multicellular organisms. Future research will be essential to elucidate how spatial cell differentiation is established in *C. perkinsii*. Nevertheless, our study indicates that *C. perkinsii* represents a transitional form between temporal and spatial cell differentiation, providing insights into the evolutionary mechanisms that led to emergence of animal multicellularity⁵¹.

Although the transcriptional signature of *C. perkinsii*’s development correlates positively with that of early-diverging animals, determining whether it is ancestral to all animals or whether it evolved independently requires still further investigation, including the spatial identification of cell type signatures at the single-cell level. If we hypothesize that developmental traits associated with palintomy are ancestral and shared with early animals, it indicates that the Holozoa have been engaging in complex developmental processes for much longer than previously thought. However, apart from partial observations in some coralochytrean⁵² and filasterean⁵³ species that may exhibit palintomic traits, no other animal relative seems to develop similarly. This discrepancy could be due to a lack of data on the complete developmental cycles of holozoans or the possibility that the ancestral developmental program was lost independently at least in choanoflagellates and ichthyophonids, if not more. Conversely, if we hypothesize that *C. perkinsii*’s developmental traits evolved independently, it indicates that the genetic toolkit present in the ancestor of the Holozoa was remarkably versatile⁶. This implies that the components necessary for multicellular development can be more easily repurposed than previously thought, partially challenging the idea that evolving an animal-like developmental program is inherently complex. To better understand the evolutionary history of this multicellular program and distinguish between these scenarios, we need to investigate further the cell physiology of non-animal holozoans across various life cycles and environments, as well as perform more comprehensive studies of the fossil record.

Indeed, the ability of free-living *C. perkinsii* to organize into a four-cell tetrahedral stage protected by a cell wall and form late spatially organized colonies with clustered flagellated cells (Fig. 4f) sets it apart from previously studied animal relatives including other ichthyosporeans, filastereans and choanoflagellates^{6,51,54,55}. These features exhibit morphological similarities with Ediacaran embryonic-like fossils of the early Ediacaran Weng'an Biota of the Doushantuo Formation (587.2 million years ago plus or minus 3.6 Myr)^{56,57} such as *Megasphaera*⁵⁸, *Helicoforamina*⁵⁹, *Spiralicella*⁶⁰ and *Caveasphaera*⁶¹, which are tentatively classified animal stem groups⁶². These fossils exhibit morphological features characteristic of some animal embryos such as cleavage divisions, Y-shaped junctions, tetrahedral four-cell stages and a proposed open mitotic strategy^{62–65}. However, their phylogenetic position, as stem animals or non-metazoan holozoans, is the subject of continuing debate^{24,58–60,65,66}, as it strongly relies on the cellular dissimilarities between these fossils and characterized close animal relatives until now, excluding *C. perkinsii*. Although these similarities do not fully elucidate the evolutionary path from Ichthyosporea to animals, our results on the multicellular development of *C. perkinsii* favour the interpretation of these fossils, particularly *Megasphaera*, as an ichthyosporean-like fossil rather than as a stem animal⁶⁶.

Altogether, our study highlights the multicellular developmental diversity in the Ichthyosporea, as illustrated by the palintomic program in *C. perkinsii* described here and the coenocytic program of *S. arctica*, reminiscent of *Drosophila*’s early embryos. With both species representing opposite ends of the developmental spectrum, the

Ichthyosporea provides, independently of the evolutionary question, a unique opportunity for a form of ‘comparative embryology’ that can show both shared and species-specific mechanisms driving multicellular development at the root of animals.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41586-024-08115-3>.

1. Kalinka, A. T. & Tomancak, P. The evolution of early animal embryos: conservation or divergence? *Trends Ecol. Evol.* **27**, 385–393 (2012).
2. Salazar-Ciudad, I. Morphological evolution and embryonic developmental diversity in metazoa. *Development* **137**, 531–539 (2010).
3. Grau-Bové, X. et al. Dynamics of genomic innovation in the unicellular ancestry of animals. *eLife* **6**, e26036 (2017).
4. Ocaña-Pallarès, E. et al. Divergent genomic trajectories predate the origin of animals and fungi. *Nature* **609**, 747–753 (2022).
5. Sebé-Pedrós, A., Degnan, B. M. & Ruiz-Trillo, I. The origin of Metazoa: a unicellular perspective. *Nat. Rev. Genet.* <https://doi.org/10.1038/nrg.2017.21> (2017).
6. Brunet, T. & King, N. in *The Evolution of Multicellularity* (eds Herron, M. D. et al.) Ch. 13 (CRC, 2022).
7. Knoll, A. H. The multiple origins of complex multicellularity. *Annu. Rev. Earth Planet Sci.* **39**, 217–239 (2011).
8. De Smet, I. & Beeckman, T. Asymmetric cell division in land plants and algae: the driving force for differentiation. *Nat. Rev. Mol. Cell Biol.* **12**, 177–188 (2011).
9. Gönczy, P. Mechanisms of asymmetric cell division: flies and worms pave the way. *Nat. Rev. Mol. Cell Biol.* **9**, 355–366 (2008).
10. Lechler, T. & Mapelli, M. Spindle positioning and its impact on vertebrate tissue architecture and cell fate. *Nat. Rev. Mol. Cell Biol.* **22**, 691–708 (2021).
11. Tsai, T. Y. C., Garner, R. M. & Megason, S. G. Adhesion-based self-organization in tissue patterning. *Annu. Rev. Cell Dev. Biol.* **38**, 349–374 (2022).
12. Wu, D., Yamada, K. M. & Wang, S. Tissue morphogenesis through dynamic cell and matrix interactions. *Annu. Rev. Cell Dev. Biol.* **39**, 123–144 (2023).
13. Arendt, D. et al. The origin and evolution of cell types. *Nat. Rev. Genet.* **17**, 744–757 (2016).
14. Bell, G. & Mooers, A. O. Size and complexity among multicellular organisms. *Biol. J. Linn. Soc.* **60**, 345–363 (1997).
15. Márquez-Zacarías, P. et al. Evolution of cellular differentiation: from hypotheses to models. *Trends Ecol. Evol.* **36**, 49–60 (2021).
16. Wolpert, L. Do we understand development? *Science* (1979) **266**, 571–572 (1994).
17. Ruiz-Trillo, I. & de Mendoza, A. Towards understanding the origin of animal development. *Development* <https://doi.org/10.1242/dev.192575> (2020).
18. Brunet, T. et al. A flagellate-to-amoeboid switch in the closest living relatives of animals. *eLife* **10**, 1–73 (2021).
19. Brunet, T. et al. Light-regulated collective contractility in a multicellular choanoflagellate. *Science* <https://doi.org/10.1126/science.aae2346> (2019).
20. Fairclough, S. R., Dayel, M. J. & King, N. Multicellular development in a choanoflagellate. *Curr. Biol.* <https://doi.org/10.1016/j.cub.2010.09.014> (2010).
21. Parra-Acero, H. et al. Integrin-mediated adhesion in the unicellular holozoan Capsaspora owczarzaki. *Curr. Biol.* <https://doi.org/10.1016/j.cub.2020.08.015> (2020).
22. Sebé-Pedrós, A. et al. Regulated aggregative multicellularity in a close unicellular relative of metazoa. *eLife* <https://doi.org/10.7554/eLife.01287> (2013).
23. Pérez-Posada, A., Dudin, O., Ocaña-Pallarès, E., Ruiz-Trillo, I. & Ondracka, A. Cell cycle transcriptomics of Capsaspora provides insights into the evolution of cyclin-CDK machinery. *PLoS Genet.* <https://doi.org/10.1371/journal.pgen.1008584> (2020).
24. Xiao, S. et al. The Weng'an biota and the Ediacaran radiation of multicellular eukaryotes. *Natl. Sci. Rev.* **1**, 498–520 (2014).
25. Ros-Rocher, N. et al. Chemical factors induce aggregative multicellularity in a close unicellular relative of animals. *Proc. Natl. Acad. Sci. USA* **120**, e221668120 (2023).
26. Wozniaka, A. et al. Bacterial lipids activate, synergize, and inhibit a developmental switch in choanoflagellates. *Proc. Natl. Acad. Sci. USA* <https://doi.org/10.1073/pnas.1605015113> (2016).
27. Levin, T. C., Greaney, A. J., Wetzel, L. & King, N. The Rosetteless gene controls development in the choanoflagellate *S. rosetta*. *eLife* <https://doi.org/10.7554/eLife.04070> (2014).
28. Levin, T. C. & King, N. Evidence for sex and recombination in the Choanoflagellate *Salpingoeca rosetta*. *Curr. Biol.* <https://doi.org/10.1016/j.cub.2013.08.061> (2013).
29. Dayel, M. J. et al. Cell differentiation and morphogenesis in the colony-forming choanoflagellate *Salpingoeca rosetta*. *Dev. Biol.* **357**, 73–82 (2011).
30. Laundon, D., Larson, B. T., McDonald, K., King, N. & Burkhardt, P. The architecture of cell differentiation in choanoflagellates and sponge choanocytes. *PLoS Biol.* **17**, e3000226 (2019).
31. Phillips, J. E., Santos, M., Konchwala, M., Xing, C. & Pan, D. Genome editing in the unicellular holozoan Capsaspora owczarzaki suggests a premetazoan role for the Hippo pathway in multicellular morphogenesis. *eLife* **11**, e77598 (2022).
32. Glockling, S. L., Marshall, W. L. & Gleeson, F. H. Phylogenetic interpretations and ecological potentials of the Mesomycetozoa (Ichthyosporea). *Fungal Ecol.* **6**, 237–247 (2013).
33. Mendoza, L., Taylor, J. W. & Ajello, L. The class Mesomycetozoea: a heterogeneous group of microorganisms at the animal-fungal boundary. *Annu. Rev. Microbiol.* **56**, 315–344 (2002).

34. Ondracka, A., Dudin, O. & Ruiz-Trillo, I. Decoupling of nuclear division cycles and cell size during the coenocytic growth of the Ichthyosporean Sphaeroforma arctica. *Curr. Biol.* **28**, 1964–1969.e2 (2018).
35. Dudin, O., Wielgoss, S., New, A. M. & Ruiz-Trillo, I. Regulation of sedimentation rate shapes the evolution of multicellularity in a close unicellular relative of animals. *PLoS Biol.* **20**, e3001551 (2022).
36. Dudin, O. et al. A unicellular relative of animals generates a layer of polarized cells by actomyosin-dependent cellularization. *eLife* **8**, e49801 (2019).
37. McCutney, B. & Dudin, O. Cellularization across eukaryotes: conserved mechanisms and novel strategies. *Curr. Opin. Cell Biol.* **80**, 102157 (2023).
38. Shah, H. et al. Life-cycle-coupled evolution of mitosis in close relatives of animals. *Nature* <https://doi.org/10.1038/s41586-024-07430-z> (2024).
39. Olivetta, M. & Dudin, O. The nuclear-to-cytoplasmic ratio drives cellularization in the close animal relative Sphaeroforma arctica. *Curr. Biol.* **33**, 1597–1605.e3 (2023).
40. Vasudevan, S., Selvi, E. & Steitz, J. A. Metazoan oocyte and early embryo development program: a progression through translation regulatory cascades. *Genes Dev.* **20**, 138–146 (2006).
41. Levin, M. et al. The mid-developmental transition and the evolution of animal body plans. *Nature* <https://doi.org/10.1038/nature16994> (2016).
42. Carvalho-Santos, Z., Azimzadeh, J., Pereira-Leal, J. B. & Bettencourt-Dias, M. Tracing the origins of centrioles, cilia, and flagella. *J. Cell Biol.* **194**, 165–175 (2011).
43. Brunet, T. & Booth, D. S. Cell polarity in the protist-to-animal transition. *Curr. Top. Dev. Biol.* **154**, 1–36 (2023).
44. Wright, B. A., Kvansakul, M., Schierwater, B. & Humbert, P. O. Cell polarity signalling at the birth of multicellularity: what can we learn from the first animals. *Front. Cell. Dev. Biol.* **10**, 1024489 (2022).
45. Klapholz, B. & Brown, N. H. Talin - the master of integrin adhesions. *J. Cell Sci.* **130**, 2435–2446 (2017).
46. Bays, J. L. & DeMali, K. A. Vinculin in cell-cell and cell–matrix adhesions. *Cell. Mol. Life Sci.* **74**, 2999–3009 (2017).
47. Sebé-Pedrós, A., De Mendoza, A., Lang, B. F., Degnan, B. M. & Ruiz-Trillo, I. Unexpected repertoire of metazoan transcription factors in the unicellular holozoan capsaspora owczarzaki. *Mol. Biol. Evol.* <https://doi.org/10.1093/molbev/msq309> (2011).
48. De Mendoza, A. et al. Transcription factor evolution in eukaryotes and the assembly of the regulatory toolkit in multicellular lineages. *Proc. Natl Acad. Sci. USA* <https://doi.org/10.1073/pnas.131181110> (2013).
49. Serra, D. et al. Self-organization and symmetry breaking in intestinal organoid development. *Nature* **569**, 66–72 (2019).
50. Palmer, A. R. Symmetry breaking and the evolution of development. *Science* **306**, 828–833 (2004).
51. Brunet, T. & King, N. The origin of animal multicellularity and cell differentiation. *Dev. Cell* <https://doi.org/10.1016/j.devcel.2017.09.016> (2017).
52. Raghu-kumar, S. Occurrence of the Thraustochytrid, Corallochytrium limacisporum gen. et sp. nov. in the coral reef lagoons of the Lakshadweep Islands in the Arabian Sea. *Botanica Marina* **30**, 83–90 (1987).
53. Hehenberger, E. et al. Novel predators reshape holozoan phylogeny and reveal the presence of a two-component signaling system in the ancestor of animals. *Curr. Biol.* **27**, 2043–2050.e6 (2017).
54. Ros-Rocher, N., Pérez-Posada, A., Leger, M. M. & Ruiz-Trillo, I. The origin of animals: an ancestral reconstruction of the unicellular-to-multicellular transition. *Open Biol.* **11**, 200359 (2021).
55. Larson, B. T. et al. Biophysical principles of choanoflagellate self-organization. *Proc. Natl Acad. Sci. USA* **117**, 1303–1311 (2020).
56. Yang, C. et al. The tempo of Ediacaran evolution. *Sci. Adv.* **7**, 9643 (2021).
57. Cunningham, J. A., Vargas, K., Yin, Z., Bengtson, S. & Donoghue, P. C. J. The Weng'an Biota (Doushantuo Formation): an Ediacaran window on soft-bodied and multicellular microorganisms. *J. Geol. Soc. London* **174**, 793–802 (2017).
58. Yin, Z. et al. Nuclei and nucleoli in embryo-like fossils from the Ediacaran Weng'an Biota. *Precambrian Res.* **301**, 145–151 (2017).
59. Yin, Z. et al. Developmental biology of Helicoforamina reveals holozoan affinity, cryptic diversity, and adaptation to heterogeneous environments in the early Ediacaran Weng'an biota (Doushantuo Formation, South China). *Sci. Adv.* **6**, 83–95 (2020).
60. Sun, W., Yin, Z., Liu, P., Zhu, M. & Donoghue, P. Developmental biology of Spiralicellula and the Ediacaran origin of crown metazoans. *Proc. R. Soc. Lond. B Biol. Sci.* **291**, 20240101 (2024).
61. Yin, Z. et al. The early Ediacaran Caveasphaera foreshadows the evolutionary origin of animal-like embryology. *Curr. Biol.* **29**, 4307–4314.e2 (2019).
62. Yin, Z. et al. Diverse and complex developmental mechanisms of early Ediacaran embryo-like fossils from the Weng'an Biota, southwest China. *Phil. Trans. R. Soc. Lond. B Biol. Sci.* **377**, 20210032 (2022).
63. Xiao, S. Mitotic topologies and mechanics of Neoproterozoic algae and animal embryos. *Paleobiology* **28**, 244–250 (2002).
64. Shuhai, X. et al. Fossil preservation in the Neoproterozoic Doushantuo phosphorite Lagerstätte, South China. *Lethaia* **32**, 219–238 (1999).
65. Chen, L., Xiao, S., Pang, K., Zhou, C. & Yuan, X. Cell differentiation and germ-soma separation in Ediacaran animal embryo-like fossils. *Nature* **516**, 238–241 (2014).
66. Huldtgren, T. et al. Fossilized nuclei and germination structures identify ediacaran ‘animal embryos’ as encysting protists. *Science* **334**, 1696–1699 (2011).
67. Alexa, A. & Rahnenfuhrer, J. topGO: enrichment analysis for gene ontology. R package version 2.42.0 <https://drdr.io/bioc/topGO/> (2020).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

© The Author(s), under exclusive licence to Springer Nature Limited 2024

Methods

Culture conditions

C. perkinsii cultures, previously described in refs. 3,38, were maintained at 23 °C in *C. perkinsii* medium (CpM) medium (yeast extract 3 g l⁻¹, malt extract 3 g l⁻¹, peptone 5 g l⁻¹, glucose 10 g l⁻¹, NaCl 20 g l⁻¹) protected from light. Cultures were either streaked onto CpM agar plates (2% agar) or grown in liquid CpM medium in a rectangular canted neck cell culture flask with vented cap (Falcon; 353108). For maintenance, liquid cultures are propagated monthly from single colonies growing on CpM agar and always protected from light. Liquid cultures were refreshed every 2 weeks (1:1,000 dilution) and restarted from a cryopreserved stock every 6 months. For synchronization, 10 ml of a 6-day-old culture is filtered using a 5 µm filter to obtain a homogenous small-celled population that is then diluted 1:100 in CpM medium and tracked for an entire life cycle (120 h).

Microscopy

Imaging of live and fixed *C. perkinsii* cells was conducted using a fully motorized Nikon Ti2-E epifluorescence inverted microscope, which was equipped with a hardware autofocus PFS4 system, a Lumencor SOLA SMII illumination system and a Hamamatsu ORCA-spark digital CMOS camera. Imaging was performed using a CFI Plan Fluor ×20 objective with a numerical aperture of 0.50, a CFI Plan Fluor ×40 air objective and a CFI Plan Fluor ×60 oil objective with a numerical aperture ranging from 0.5 to 1.25.

To maintain the temperature at a consistent 23 °C, a cooling/heating P Lab-Tek S1 insert (Pecon GmbH) connected to a Lauda Loop 100 circulating water bath was used. To minimize light-induced phototoxicity, the light source beam was manually adjusted to its lowest setting, and a 715 nm LongPass Color Filter (Thorlabs; FGL715S) was used to selectively allow far-red wavelengths to pass through. Moreover, we meticulously controlled the imaging frequency and the number of Z-stacks during long-term imaging to prevent cell lysis. Our observations showed that an imaging interval of 15 min, combined with seven Z-slices over 24 h, represented the upper limit for optimal imaging conditions, beyond which most cells exhibited abnormal behaviour and lysis. For confocal imaging of membrane-stained, phalloidin-stained and expanded samples, an upright Leica SP8 confocal microscope with an HC PL APO ×40/1.25 glycerol objective was used.

Cell fixation and staining

Cell fixation was carried out either using 4% formaldehyde and 250 mM sorbitol for 20 min or using prechilled methanol (100%) at -20 °C for 7 min, followed by two phosphate-buffered saline (PBS) washes. To stain the nuclei, cultures were allowed to settle for 15 min at room temperature before fixation, and Hoechst 33342 nuclear stain (Thermo Fisher; 62249) was added at a concentration of 20 mM. For actin staining, fixed cells underwent a single PBS wash before adding the F-actin stain Alexa Fluor 488-Phalloidin (Thermo Fisher; A12379) at a final concentration of 0.165 mM. Before imaging, the fixed and stained cells were concentrated and placed between a slide and coverslip. For live-cell imaging, saturated cultures were diluted 500× in CpM medium in an eight-well Ibidi chamber (Ibidi; 80826) or a 35 mm dish (Mattek; P35G-1.5-14-C). The plastic cover was removed to ensure oxygenation during the whole experiment period. To reduce evaporation, silicon oil 100 cSt was added on top (Sigma; 378364). For plasma membrane live staining, FM4-64 (Thermo Fisher; T13320) at a final concentration of 10 mM from 100× dimethyl sulfoxide diluted stock solution was directly added to the medium before imaging.

U-ExM

U-ExM was adapted from refs. 38,68. Initially, cells were fixed in a 4% formaldehyde solution with 250 mM sorbitol for 10 min. Following fixation, they underwent two washes with 1× PBS and were

subsequently resuspended in 1 ml PBS. Coverslips were precoated with poly-L-lysine for an hour, and the fixed cells were then added and allowed to adhere for 1 h. Anchoring in acrylamide/formaldehyde (1% acrylamide / 0.7% formaldehyde) was performed for 12 h at 37 °C. Gelation relied on a monomer solution (19% (wt/wt) sodium acrylate (Chem Cruz, AKSci 10% (wt/wt) acrylamide (Sigma-Aldrich; A4058), 0.1% (wt/wt) N, N'-methylenbisacrylamide (Sigma-Aldrich; M1533) in PBS and was performed at 37 °C for 1 h in a moist chamber. For denaturation, gels were transferred to the denaturation buffer for 15 min at room temperature and then shifted to 95 °C for 1.5 h. Following denaturation, expansion was performed with water exchanges as previously described^{38,68}. Post expansion, gel diameter was measured and used to determine the expansion factor. For microtubule staining of U-ExM gels, mouse primary antibodies targeting beta-tubulin (DSHB; E7) (1/300 dilution) or acetylated tubulin (Sigma; T6793) and goat antimouse IgG (H + L) cross-adsorbed secondary antibodies coupled to Alexa Fluor 488 (Thermo Fisher; A-11001) (1/500 dilution) were used. For protein pan-labelling, gels were incubated with Alexa Fluor 568 NHS-Ester (Thermo Fisher; A20003) in NaHCO₃ (pH = 8.28) for 1.5 h. For nuclear staining, Hoechst 33352 was incubated in PBS for 10 min before re-expansion in water. For gel mounting, gels were cut to appropriate sizes and attached to precoated poly-D-lysine coverslips and sealed using i-Spacers (Sunjin Lab; IS013).

Image analysis

Image analysis was done using ImageJ software (v.1.52) and Imaris (v.10.0.1) (Bitplane). For cell and nuclear diameter measurements (Fig. 1d,g and Extended Data Fig. 1e,f), brightfield (live imaging) and immunostained images were transformed into a binary before using the particle analysis function in ImageJ with a circularity parameter set to 0.15 to 1 to measure the perimeter. As cells are spherical, we computed the cell diameter as $D = 2(C/2\pi)$, where C is circumference. For nuclear content distribution (Fig. 1f and Extended Data Fig. 1g), fixed and Hoechst-stained cells and colonies were imaged, and the number of nuclei per cell were counted either using the ObjectJ plugin in ImageJ or by segmenting the nuclei using the Imaris 'surface' tool when the number of nuclei were difficult to count manually. For nuclear position measurements (Fig. 3d), nuclear-stained and U-ExM cells were first oriented with the nucleus positioned at the top. Using the line tool in ImageJ, we measured the distance between the cortex, either from the bottom (y_1) or from the top (y_2), and the centre of the metaphase plate. For volume ratio measurements of the two-cell stage (Fig. 3e), we used LimeSeg⁶⁹ in ImageJ/Fiji⁷⁰ for semi-automated segmentation of nuclei and phalloidin-stained cells. Briefly, after manual visual inspection, a spherical seed was manually added to the region of interest (ROI) manager for each detected nucleus. These spherical seeds were then inflated with LimeSeg using two main parameters: 0.8 microns between each surface element and applying a pressure of 0.009 arbitrary units (a.u.) on the phalloidin channel. Typically, segmentation convergence to the three-dimensional cell shape was achieved in a matter of seconds (Extended Data Fig. 3d). For spindle orientation measurements (Fig. 3f), U-ExM cells stained for microtubules were used where the entirety of the spindle, including both spindle poles, was visible. The angles were then measured between the plane of the first cleavage division and the plane defined by the positions of both spindle poles. To quantify the number of flagellated cells released in the environment (Extended Data Fig. 1d), we used brightfield or U-ExM images and counted the single cells harbouring a flagellum that were not localized in any multicellular colonies. For the analysis of flagellated cells in multicellular colonies (Fig. 4c–e), U-ExM cells stained for microtubules or flagella (acetylated tubulin) and DNA (Hoechst) were examined. Initially, we measured nuclear content using the surface tool in Imaris. Subsequently, flagellated cells were counted using ObjectJ in ImageJ or the surface tool in Imaris,

depending on the number of flagellated cells. In instances where there were more than 150 flagellated cells per colony, a manual approach combining the skeletonize tool and ObjectJ in ImageJ was used. For quantification of spatial clustering of flagellated cells in multicellular colonies (Extended Data Fig. 4h), two maximum projections were obtained for each cell: one for the upper half and one for the bottom half. Each projection was then split into two quarters, resulting in four quarters per cell (Q1, Q2, Q3, Q4). Flagellated cells were counted in each quarter and projected onto a radial plot. To allow for comparison between different cells, the quarter with the highest number of flagellated cells was aligned to Q1. Means and standard deviations were calculated for five cells per time point and represented on the final radial plot. All figures were assembled with Illustrator CC 2020 (Adobe). Several figures were generated using ggplot2 in R v.0.5.53.

RNA isolation, library preparation and sequencing

Synchronized cultures of *C. perkinsii* at 23 °C protected from light were sampled at critical time points throughout their entire life cycle (120 h). Total RNA was extracted by Trizol and purified using an RNA Clean & Concentrator-5 kit (Zymo; R1013) from 300–600 ml of culture at each time point. Libraries were prepared using 400 ng of RNA using the Illumina stranded mRNA ligation prep (ISML). Libraries were sequenced on NovaS6000 in a PE60 at the EPFL gene expression core facility. We obtained between 51.2 and 104.5 M reads per sample. Reads were trimmed for ISML (Nextera) adaptors.

Gene expression analysis in *C. perkinsii*

RNA-sequencing (RNA-seq) reads were mapped to the *C. perkinsii* transcriptome⁷¹ using Salmon (v.1.1.0)⁷² in mapping-based mode and the full genome³ to generate decoys for accurate read mapping. Salmon quant.sf files were read into the R programming environment (v.4.3.1) (www.r-project.org/) using the tximport (v.1.30.0)⁷³ tool. Differential expression analyses in *C. perkinsii* were conducted with edgeR (v.4.0.3)^{74–76}. RNA was collected and sequenced for three replicate cultures; however, on examination of the data, one replicate was discarded because of anomalous expression of all time points from that culture as observed in a multiple-dimensional scaling plot (Extended Data Fig. 2a) and examination of expression patterns from that culture indicative of a possible infection or contamination of the culture (Extended Data Table 1). The remaining two replicates (A and C, Extended Data Fig. 2a) were examined for differential gene expression across time by considering the differential expression of each time point relative to expression in the first time point using genewise negative binomial generalized linear models with quasi-likelihood tests (glmQLFTest)⁷⁶ in edgeR. Differentially expressed genes across all time points (false discovery rate adjusted P value less than 0.05) were collected for further analyses of expression patterns. Expression patterns were extracted from the differentially expressed genes by fuzzy c-means clustering in R from package e1071 (v.1.7.14)⁷⁷. Considering in-cluster sum of squares for different values of k-clusters, five clusters were chosen as the optimal cluster number. Gene expression patterns across time were plotted using the viridis (v.0.6.4)⁷⁸ cividis palette and superheat (v.1.0.0) plotting functions⁷⁹.

Gene ontology functional enrichment analyses

Functional annotation of the *C. perkinsii* transcriptome was conducted by BLASTP (v.2.10.0+)⁸⁰ analysis of *C. perkinsii* predicted peptides against the UniProtKB/Swiss-Prot expertly curated protein database (downloaded 16 June 2021, containing 565,254 entries)⁸¹. BLASTP hits were filtered by best ‘homology-derived structure of proteins’ distance score⁸² as in ref. 83. The gene ontology annotations for best Swiss-Prot hits to each *C. perkinsii* peptide were converted to a.goa format file and used for functional enrichment analyses of expression clusters with topGO (v.2.54.0)⁶⁷.

Comparison of *C. perkinsii* gene expression to *S. arctica* and ten animal developmental time courses

To compare gene expression patterns across evolutionary divides, we used OrthoFinder (v.2.5.4)⁸⁴ to cluster genes into orthogroups across ten animals found in ref. 41—*Caenorhabditis elegans*, *Platynereis dumerili*, *Drosophila melanogaster*, *Strongylocentrotus purpuratus*, *Danio rerio*, *Mnemiopsis leidyi*, *Nematostella vectensis*, *Schmidtea polychroa*, *Hypsibius dujardini*, *Amphimedon queenslandica*—as well as *C. perkinsii* and *S. arctica*³⁶; four more ichthyosporeans: *Amoebiidium parasiticum*, *Creolimax fragrantissima*, *Ichthyophonus hoferi*, *Sphaerothecum destruens*; one Corallochytrean: *Corallochytrium limacisporum*; two fungi: *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe*; and one amoebozoan: *Dictyostelium discoideum*. Transcriptomes and read count data for the ten animals from ref. 41 were downloaded from the associated National Center for Biotechnology Information GEO database (accession number GSE70185). Transcriptomes were translated using TransDecoder (v.5.5.0)⁸⁵. Peptide files for further ichthyosporeans and the corallochytrean were downloaded from the Ruiz-Trillo Multicellgenome lab resources on FigShare⁷¹. The yeast and amoebozoan predicted peptide data (outgroups) were downloaded from their National Center for Biotechnology Information genome pages^{86–88}. Orthogroups from the predicted peptides of the 20 organisms were predicted using the tool OrthoFinder⁸⁴.

To compare developmental transcriptional patterns among animals and *C. perkinsii*, we selected time points from animal development for each animal that went from the one- or two-cell stage through the blastula stage, as detailed in Extended Data Table 2, on the basis of the developmental time courses detailed in ref. 41. We used sliding windows to create an equivalent number of time points for each animal’s development, reducing the number of time points to match those of *C. perkinsii*. We used the first six time points of *C. perkinsii* development for comparison with animals (54, 72, 84, 90, 96, 104 h). The final time point (120 h), aligning with cell release, was considered to be unique to *C. perkinsii* development and probably not shared with multicellular animals whose multicellular developmental programs continue beyond the blastula stage.

Orthogroups were used for comparing expression patterns across animals. Given that an orthogroup may encompass several genes, each exhibiting unique expression patterns, we used network analysis to segregate these genes on the basis of their expression profiles before conducting cross-species comparisons (Extended Data Fig. 2f). Specifically, we calculated the Euclidean distance between all gene pairs in an orthogroup to quantify expression dissimilarities. This analysis facilitated the construction of a graph where genes, represented as nodes, were interconnected by edges if their expression distance fell below a predefined threshold (threshold of 1.61), thereby indicating similarity. Conversely, genes with distances surpassing this threshold remained unlinked. Subsequently, isolated nodes were eliminated from the graph, which was then subjected to a connected component analysis using the igraph package (v.1.6.0) in R. This step identified clusters of three or more interconnected genes from within or across organisms as distinct expression patterns in the orthogroup, which were then extracted for further analyses. In instances where an organism lacked a gene corresponding to a specific expression pattern in a connected component yet possessed other genes from the same orthogroup, the expression values of these genes were averaged to represent that organism’s contribution to the orthogroup’s overall expression profile. The network procedure allowed analyses of all distinct expression patterns per orthogroup across animals and is illustrated in Extended Data Fig. 2f.

Functional annotation of orthogroups was achieved by (1) aligning all proteins in an orthogroup using mafft (v.7.487)⁸⁹; (2) converting that alignment to a hidden Markov model using hmmer (v.3.3.2) (<http://hmmer.org/>) and emitting a strict consensus sequence from that

Article

hidden Markov model; and then (3) using BLASTP of each consensus sequence against the UniProtKB/Swiss-Prot database (downloaded 16 June 2021, containing 565,254 entries)⁸¹ and collecting the best hit by homology-derived structure of proteins score as above for *C. perkinsii* genes. The UniProtKB/Swiss-Prot gene names and annotations were transferred to the orthogroups in this way and used for functional inference and gene ontology analyses.

Statistical analyses were conducted to determine the correspondence between gene expression in *C. perkinsii* and animals. This was achieved by comparing the distribution of Pearson correlation values between *C. perkinsii* gene expression and gene expression in each animal. The analysis included all orthogroups, focusing on genes important for development in *C. perkinsii*. Specifically, we examined the five identified clusters of differentially expressed genes in *C. perkinsii* and assessed whether these clusters, on average, showed a higher correlation with gene expression in animals compared to genes that are not differentially expressed during *C. perkinsii* development. We used the Kruskal–Wallis *H*-test to ask whether there are significant differences in the distribution of correlations between non-developmental and differentially expressed gene clusters and then completed pairwise one-sided Wilcoxon rank-sum tests to determine whether clusters of differentially expressed genes were more correlated than non-differentially expressed (non-developmental) genes. The resultant distributions are plotted in Extended Data Fig. 2i with the following *P* values from the Wilcoxon test (with a Bonferroni multiple testing correction): cluster 1 $P = 1.18 \times 10^{-15}$, cluster 2 $P = 4.54 \times 10^{-174}$, cluster 3 $P = 5.36 \times 10^{-49}$, cluster 4 $P = 1.13 \times 10^{-22}$, cluster 5 $P = 2.56 \times 10^{-102}$. Pairwise correlation measurements for each cluster among animals and the resultant plots (Extended Data Fig. 2k) were conducted using the corrplot package (v.0.92) in R⁹⁰. We implemented further per-orthogroup statistical tests of correspondence using Kendall's coefficient of concordance to test whether correspondence in gene expression was significant across animals for select orthogroups (Supplementary Table 1). Empirical statistics were measured for a set of developmental genes. This was done by calculating the average correlation of 10,000 randomly selected gene subsets across animals to determine how likely it was that the correlation value associated with the developmental genes could have been selected at random (Extended Data Fig. 2j). The approach was validated by computing the correlation value for a set of control genes not involved in developmental progression—specifically, glycolysis genes—which showed a correlation value similar to that of a random selection of genes (Extended Data Fig. 2j). Silhouettes were obtained from PhyloPic (<https://www.phylopic.org>), which also now linked to the R package rphylopic⁹¹.

Reporting summary

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

Data availability

All data used for quantifications and Supplementary Videos 1–10 are available at Figshare (<https://figshare.com/s/f20f6d471c719990471c>)⁹². For all other raw images required, they are available on request due to their large data sizes. Raw RNA-seq data are publicly available on <https://www.ncbi.nlm.nih.gov/bioproject/PRJNA1091032>.

Code availability

All the generated code for transcriptomic analysis is available at Zenodo (<https://doi.org/10.5281/zenodo.13352464>)⁹³.

68. Gambarotto, D. et al. Imaging cellular ultrastructures using expansion microscopy (U-ExM). *Nat. Methods* **16**, 71–74 (2018).
69. Machado, S., Mercier, V. & Chiaruttini, N. LimeSeg: a coarse-grained lipid membrane simulation for 3D image segmentation. *BMC Bioinf.* **20**, 1–12 (2019).
70. Schindelin, J. et al. Fiji: an open-source platform for biological-image analysis. *Nat. Methods* **9**, 676–682 (2012).
71. Multicellgenome Lab. Genome - *Chromosphaera perkinsii*. Figshare https://figshare.com/articles/dataset/Genome_-_Chromosphaera_perkinsii/5426494 (2024).
72. Patro, R., Duggal, G., Love, M. I., Irizarry, R. A. & Kingsford, C. Salmon provides fast and bias-aware quantification of transcript expression. *Nat. Methods* **14**, 417–419 (2017).
73. Soneson, C., Love, M. I. & Robinson, M. D. Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences. *F1000Res.* **4**, 1521 (2016).
74. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2010).
75. McCarthy, D. J., Chen, Y. & Smyth, G. K. Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. *Nucleic Acids Res.* **40**, 4288 (2012).
76. Chen, Y., Lun, A. T. L. & Smyth, G. K. From reads to genes to pathways: differential expression analysis of RNA-Seq experiments using Rsubread and the edgeR quasi-likelihood pipeline. *F1000Res.* **5**, 1438 (2016).
77. Meyer, D. et al. e1071: misc functions of the Department of Statistics, Probability Theory Group (formerly: E1071), TU Wien. R package version 1.7-15 <https://rdrr.io/rforge/e1071/> (2024).
78. Garnier, S. et al. viridis: colorblind-friendly color maps for R. R package version 0.6.4 <https://rdrr.io/cran/viridis/> (2024).
79. Barter, R. & Yu, B. superheat: a graphical tool for exploring complex datasets using heatmaps. R package version 0.1.0 <https://rdrr.io/cran/superheat/> (2019).
80. Camacho, C. et al. BLAST+: architecture and applications. *BMC Bioinf.* **10**, 1–9 (2009).
81. Boutet, E. et al. UniProtKB/Swiss-Prot, the manually annotated section of the UniProt KnowledgeBase: how to use the entry view. *Methods Mol. Biol.* **1374**, 23–54 (2016).
82. Rost, B. Enzyme function less conserved than anticipated. *J. Mol. Biol.* **318**, 595–608 (2002).
83. Burns, J. A., Paasch, A., Narechania, A. & Kim, E. Comparative genomics of a bacterivorous green alga reveals evolutionary causalities and consequences of phago-mixotrophic mode of nutrition. *Genome Biol. Evol.* **7**, 3047–3061 (2015).
84. Emms, D. M. & Kelly, S. OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biol.* **20**, 1–14 (2019).
85. Haas, B. J. et al. TransDecoder source. GitHub <https://github.com/TransDecoder/TransDecoder> (2024).
86. Engel, S. R. et al. New data and collaborations at the Saccharomyces Genome Database: updated reference genome, alleles, and the Alliance of Genome Resources. *Genetics* **220**, iyab224 (2022).
87. Wood, V. et al. The genome sequence of *Schizosaccharomyces pombe*. *Nature* **415**, 871–880 (2002).
88. Eichinger, I. et al. The genome of the social amoeba *Dictyostelium discoideum*. *Nature* **435**, 43–57 (2005).
89. Katoh, K. & Standley, D. M. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* **30**, 772–780 (2013).
90. Wei, T. & Simko, V. corrplot: visualization of a correlation matrix. R package version 0.92. <https://rdrr.io/packages/corrplot/versions/0.94> (2024).
91. Gearty, W. & Jones, L. A. rphylopic: an R package for fetching, transforming, and visualising PhyloPic silhouettes. *Methods Ecol. Evol.* **14**, 2700–2708 (2023).
92. Olivetta, M. et al. Source data for ‘A multicellular developmental program in a close animal relative’. Figshare <https://figshare.com/s/f20f6d471c719990471c> (2024).
93. Olivetta, M. et al. Code for ‘A multicellular developmental program in a close animal relative’. Zenodo <https://doi.org/10.5281/zenodo.13352464> (2024).

Acknowledgements We thank P. Gönczy, G. Dey, A. de Mendoza, C. Martín-Durán, H. Shah and A. Woglart for comments on the manuscript and general feedback; H. Suga for *C. perkinsii* cultures; and both the EPFL Bio Imaging and Optics Platform and the EPFL Gene Expression Core Facility for their support. We acknowledge funding from US National Science Foundation grant no. OIA-1826734 (J.B.), Swiss National Science Foundation Ambizione grant no. PZOOOP3_185859 (O.D. and M.O.) and Swiss National Science Foundation Starting grant no. TMSGI3_218007 (O.D. and M.O.).

Author contributions Conceptualization: O.D. Methodology: M.O., C.B., N.C., J.B., O.D. Investigation: M.O., C.B., N.C., J.B. and O.D. Visualization: M.O., N.C., J.B. and O.D. Funding acquisition: J.B. and O.D. Project administration: O.D. Supervision: O.D. Writing—original draft: O.D. Writing—review and editing: M.O., N.C., J.B. and O.D.

Competing interests The authors declare no competing interests.

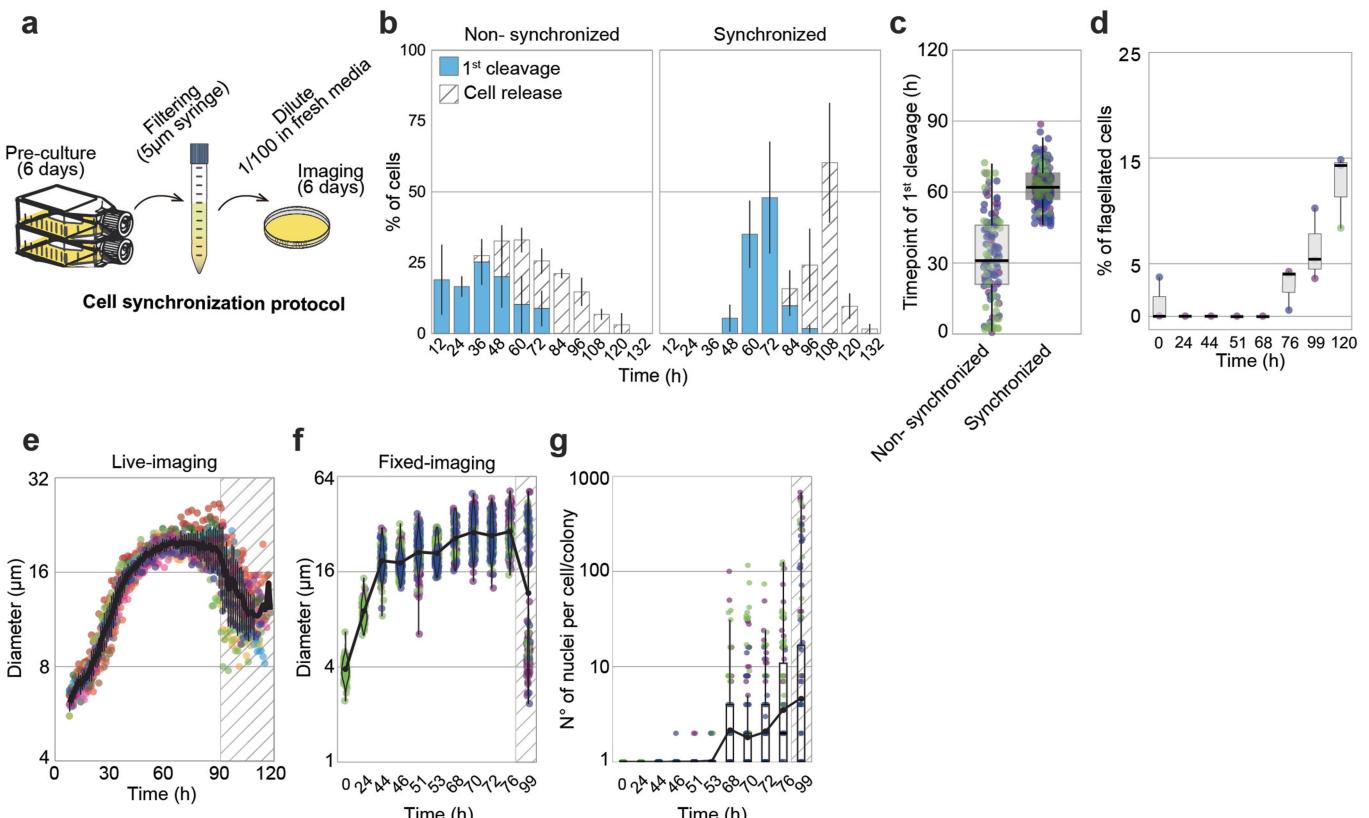
Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41586-024-08115-3>.

Correspondence and requests for materials should be addressed to John Burns or Omaya Dudin.

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

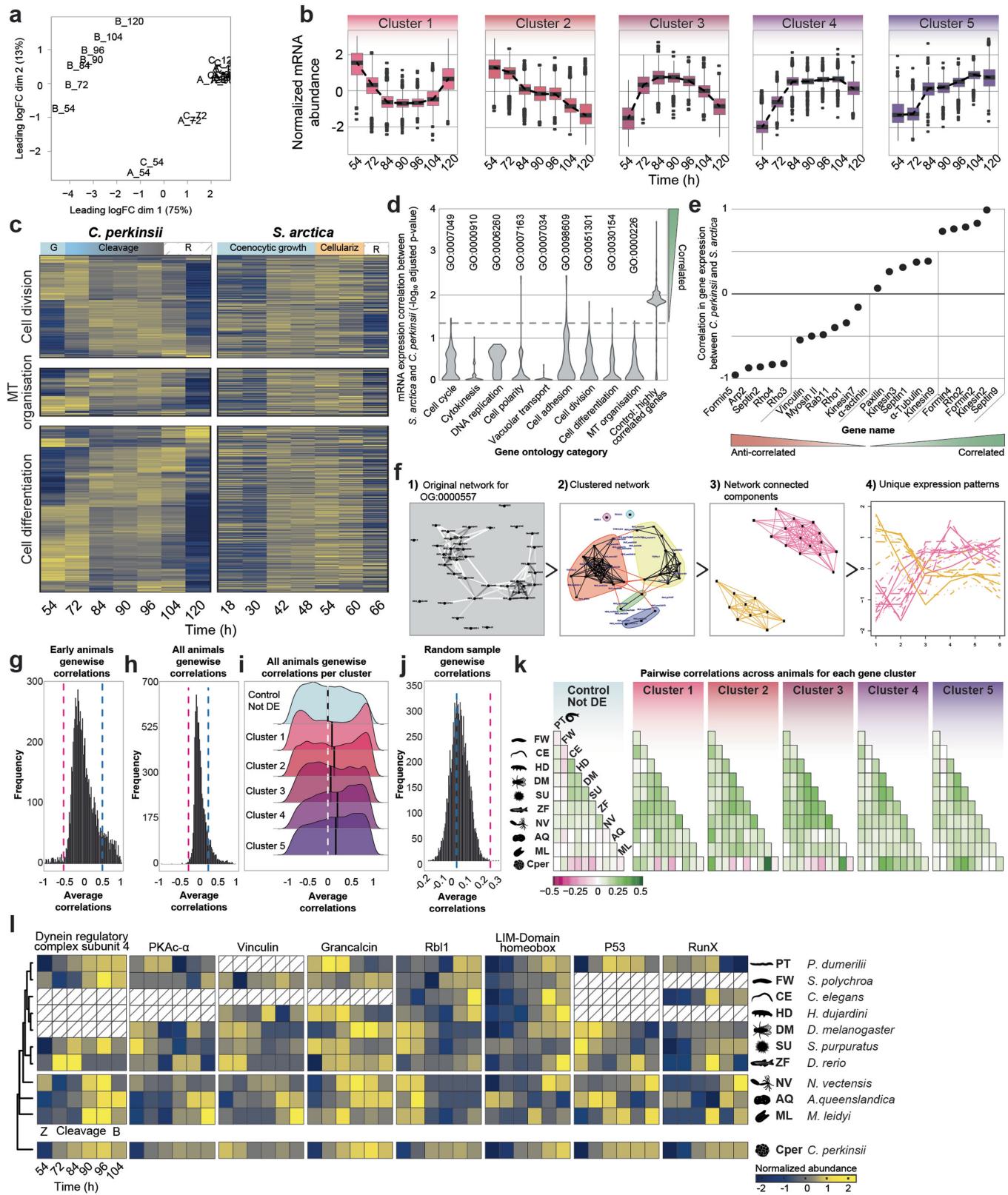
Reprints and permissions information is available at <http://www.nature.com/reprints>.



Extended Data Fig. 1 | Culture synchronization and growth dynamics of *C. perkinsii*. (a) Schematic of the synchronization protocol of *C. perkinsii*. (b) Distribution of cells undergoing their first cleavage division or cell release throughout the life cycle at 23 °C, following either synchronization or not ($n_{\text{non-synchronized}}=129$ cell/colony, $n_{\text{synchronized}}=233$ cell/colony). Data are mean \pm s.d. (c) The average time at which non-synchronized and synchronized cells undergo their 1st cleavage in a bulk culture ($n_{\text{non-synchronized}}=129$ cell/colony, $n_{\text{synchronized}}=233$ cell/colony). Box plots extend from the 25th to 75th percentile, including the median, and whiskers extend from the minimum to maximum value. (d) Percentage of free-swimming flagellated cells over time ($n_0=112$ cells, $n_{24}=84$ cells, $n_{44}=110$ cells, $n_{46}=70$ cells, $n_{51}=99$ cells, $n_{53}=73$

cells, $n_{68}=116$ cells, $n_{70}=105$ cells, $n_{72}=68$ cells, $n_{76}=77$ cells, $n_{99}=104$ cells, $n_{120}=67$ cells). Box plots extend from the 25th to 75th percentile, including the median, and whiskers extend from the minimum to maximum value. (e) Cell diameter over time of 10 single cell traces aligned to time. Variability increases from 75 hr onwards due to asynchronous cell release. Data are mean \pm s.d. (f) Cell diameter over time of fixed cells ($n=1571$ cell/colony). The colours represent 3 independent replicates. (g) Distribution of number of nuclei per cell/colony across the life-cycle ($n=1571$ cell/colony). The colours represent 3 independent replicates. Box plots extend from the 25th to 75th percentile, including the median, and whiskers extend from the minimum to maximum value.

Article

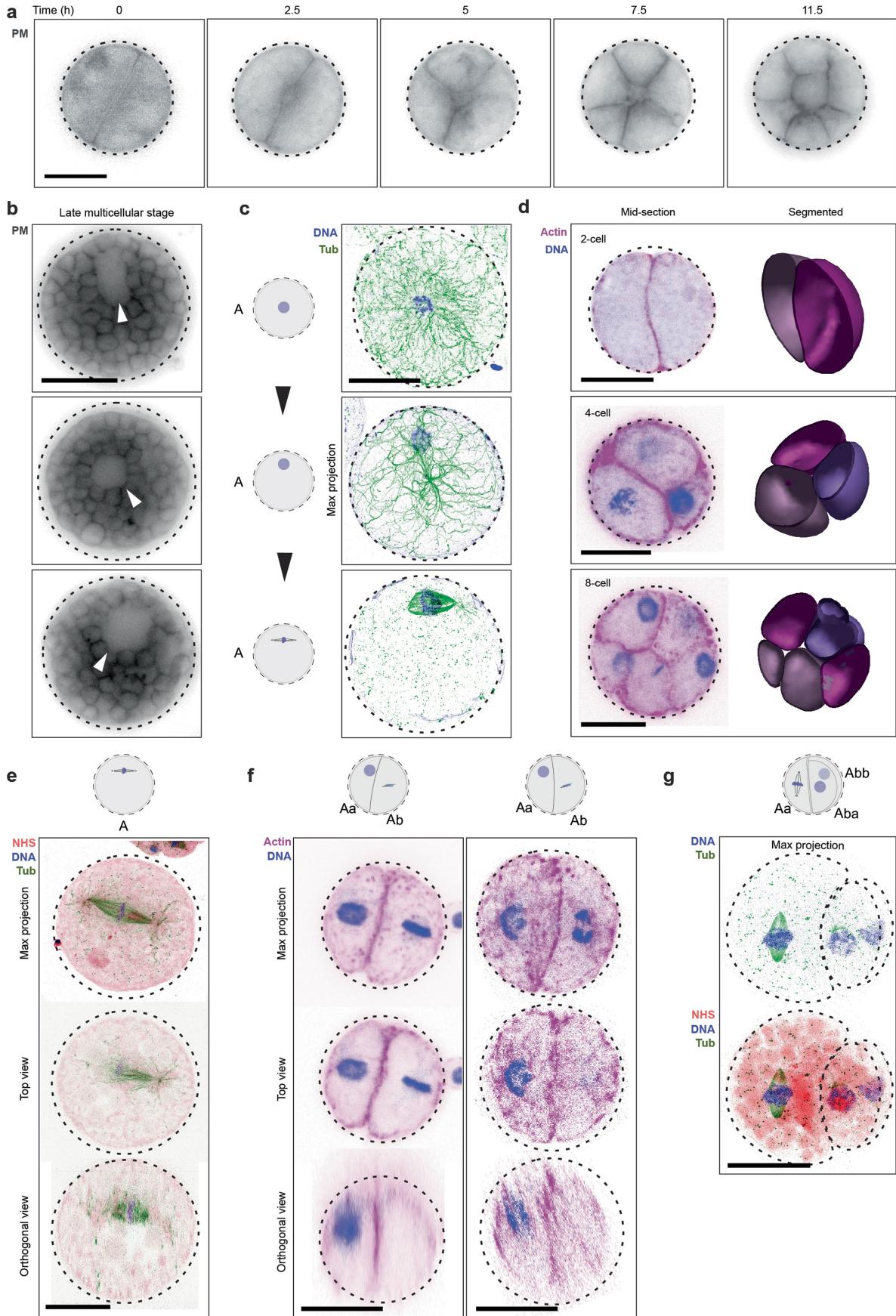


Extended Data Fig. 2 | See next page for caption.

Extended Data Fig. 2 | Transcriptional dynamics in *C. perkinsii*'s palintomic lifecycle and comparative analysis with *S. arctica* and early-branching animals. (a) Multiple dimensional scaling (MDS) plot of RNAseq counts of 3 independent developmental time courses of *C. perkinsii*. "B" replicates were excluded from further analyses due to evidence of infection (Extended Data Table 1). (b) Expression profile distributions for the 5 distinct gene expression clusters for both replicates. Box plots extend from the 25th to 75th percentile, including the median, and whiskers extend from the minimum to maximum value. (c) Heatmaps of gene expression dynamics between *C. perkinsii* and *S. arctica* highlight distinct transcriptional programs. (d) Statistical evaluation of gene expression correlation between *C. perkinsii* and *S. arctica* in select GO categories. (e) Correlation values comparing *C. perkinsii* to *S. arctica* expression patterns for genes important during coenocytic development of *S. arctica*³⁵. (f) Network strategy used to decouple distinct expression patterns occurring within an orthogroup (methods). (g) Average correlation of OG expression values between early diverging animals and *C. perkinsii*. Orthogroups with an average correlation value greater than 0.5 (blue line) were subjected to GO enrichment analyses (Supplementary Tables 1, 2). (h) Average correlation of OG expression values between all animals and *C. perkinsii*. OGs with an average correlation value greater than 0.25 (blue line) were subjected to GO enrichment analyses (Supplementary Tables 1, 2). (i) Distribution of gene expression correlation values between *C. perkinsii* and each animal for non-developmental genes (Control Not DE) and differentially expressed gene clusters. The white

dashed line indicates the median correlation value of the control distribution. (j) Average gene expression correlation among all animals plus *C. perkinsii* for a random sample of 34 OGs (black bars), developmental genes (pink line), and glycolysis genes (not developmentally regulated; blue line). (k) Pairwise correlation coefficient plots of gene clusters showing pairwise correlations among all animals and *C. perkinsii*. (l) Heatmap of gene expression for select OGs between *C. perkinsii* and developmental stages of early-branching animals spanning from zygote formation (Z) to the blastula stage (B). Credits: Silhouettes were obtained from PhyloPic (<https://www.phylopic.org/>). *Schmidtea polychroa*, created by M. A. Grohme under a CC0 1.0 Universal Public Domain licence; *Caenorhabditis elegans*, created by B. Goldstein, vectorization by J. Warner under a CC0 1.0 Universal Public Domain licence; *Isohypsibius dastychi*, created by B. Lang under a CC0 1.0 Universal Public Domain licence; *Sophophora melanogaster*, created by A. Wilson under a CC0 1.0 Universal Public Domain licence; *Strongylocentrotus purpuratus*, created by C. Schomburg under a CC0 1.0 Universal Public Domain licence; *Danio rerio*, created by Ian Quigley under a CC BY 3.0 licence; *Nematostella vectensis*, created by J. Warner under a CC0 1.0 Universal Public Domain licence; *Amphimedon queenslandica* obtained under a CC0 1.0 Universal Public Domain licence; *Mnemiopsis leidyi*, created by J. R. Winnikoff under a CC0 1.0 Universal Public Domain licence; *Creolimax*, created by Y. Wong using scanning electron microscopy images by A. Sebé-Pedrós (public domain agreed by I. Ruiz-Trillo) under a licence free of copyright, CC PDM 1.0.

Article



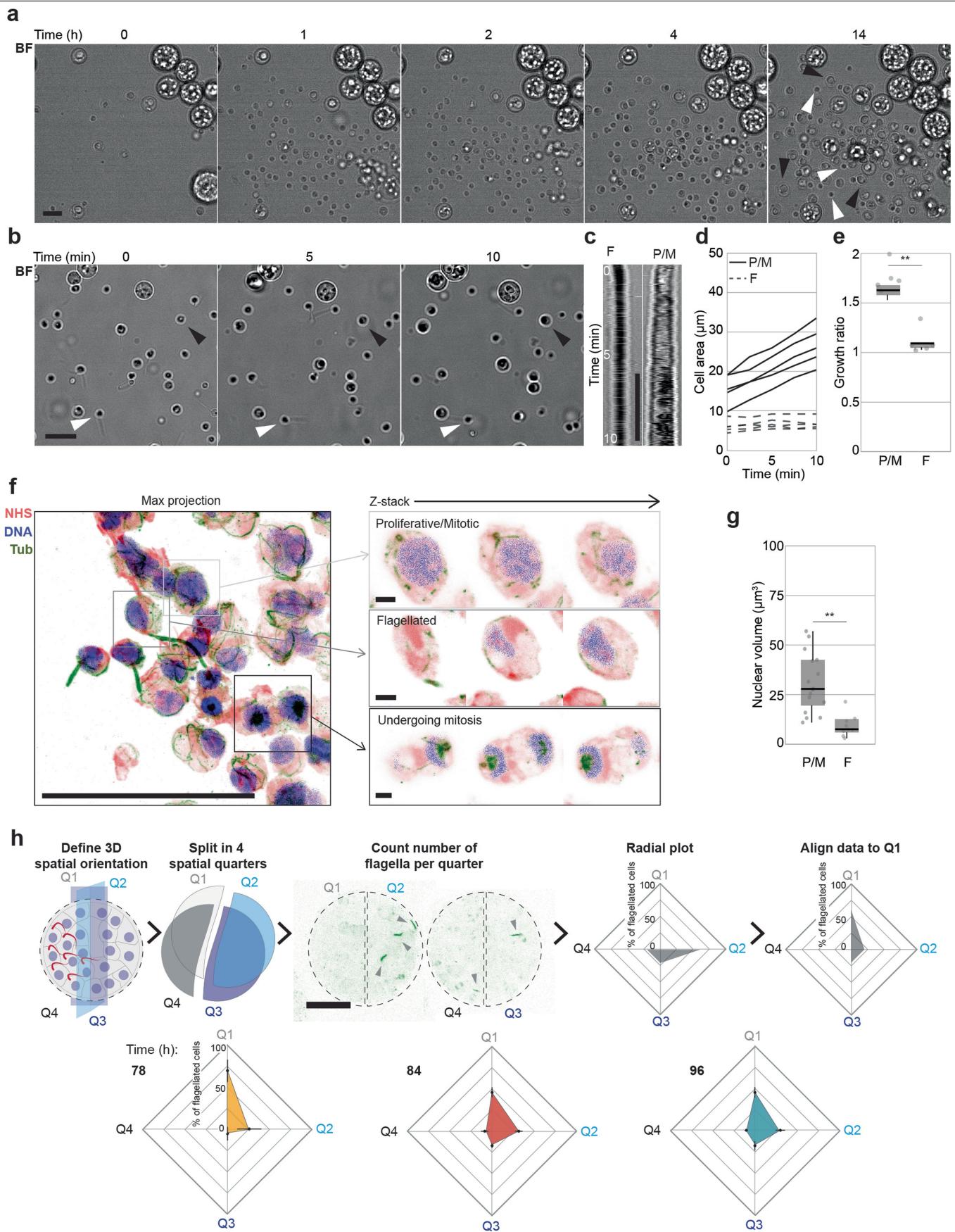
Extended Data Fig. 3 | See next page for caption.

Extended Data Fig. 3 | Symmetry breaking during early development of *C. perkinsii*.

(a) Cleavage division dynamics at constant volume over time, visualized using a plasma membrane (PM) marker FM4-64 (Supplementary Video 4). Scale bar, 10 μm . **(b)** PM-stained live late colonies exhibiting internal cavities (arrows). This result has been reproduced at least 2 independent times. Scale bar, 10 μm . **(c)** U-ExM stained cells for microtubules (green), and DNA (blue), highlighting nuclear migration to the cortex prior to 1st mitotic division. This result has been reproduced at least 3 independent times. Scale bar, 10 μm . **(d)** Actin (magenta) and DNA-stained (blue) colonies at various cell stages, accompanied by volumetric segmentation of cells, highlighting the asymmetry in volume following the first cleavage division. This result has been reproduced

at least 3 independent times. Scale bar, 10 μm . **(e)** U-ExM stained cell for pan-labelling with NHS-Ester (red), microtubules (green), and DNA (blue), illustrating the first mitotic division at the cortex. This result has been reproduced at least 3 independent times. Scale bar, 10 μm . **(f)** Actin (magenta) and DNA-stained (blue) colonies at distinct cell stages showing the asymmetrical cell division in volume and time between the Aa and Ab cells. This result has been reproduced at least 3 independent times. Scale bar, 10 μm . **(g)** U-ExM stained colonies for pan-labelling with NHS-Ester (red), microtubules (green), and DNA (blue), highlighting a three-cell stage. This result has been reproduced at least 3 independent times. Scale bar, 10 μm .

Article



Extended Data Fig. 4 | See next page for caption.

Extended Data Fig. 4 | Distinct cell fates following the release of *C. perkinsii*'s colonies. (a) Time-lapse images over 14 h of new born *C. perkinsii* cells post-release show a cell-size increase in proliferative/mitotic cells (black arrows) and no increase in flagellated cells (white arrows) (Supplementary Video 7). Scale bar: 10 μm . (b) A short-duration high-resolution time-lapse of new born *C. perkinsii* cells post-release ($T = 104$ h) shows a cell-size increase in proliferative/mitotic cells (black arrows) and no increase in flagellated cells (white arrows) (Supplementary Video 8). Bar, 10 μm . (c) Kymographs of both flagellated (F) and proliferative/mitotic (P/M) cells over 10 min show the increase in cell size for the P/M cell. Bar, 10 μm . (d) Cell area over time of single cell traces reveals the increase in size for the proliferative/mitotic (P/M) cells in contrast to the flagellated cells (F) ($n_{\text{each}} = 5$ cells). (e) Growth ratio between $t = 0$ and $t = 10$ min for both proliferative/mitotic (P/M) and flagellated cells (F) ($n_F = 5$, $n_{P/M} = 5$ cells). Statistical analysis using a two-tailed Student's t-test, ** $P = 0.001$. Box plots extend from the 25th to 75th percentile, including the median, and whiskers

extend from the minimum to maximum value. (f) U-ExM stained cells for pan-labelling with NHS-Ester (red), microtubules (green), and DNA (blue), identifying proliferative/mitotic and flagellated cells, as well as cells undergoing mitosis. Scale bar: 10 μm . Zoomed-in insets represent Z-stacks of each cell. Scale bar: 1 μm . (g) Nuclear volume measured from U-ExM images for both proliferative/mitotic (P/M) and flagellated cells (F) ($n_F = 10$, $n_{P/M} = 16$ cells). Statistical analysis using a two-tailed Student's t-test, ** $P = 0.00016$. Box plots extend from the 25th to 75th percentile, including the median, and whiskers extend from the minimum to maximum value. (h) Image analysis workflow and quantifications showing the spatial clustering of flagellated cells within multicellular colonies (see methods). Two maximum projections per cell (upper and bottom halves) were split into four quarters (Q1, Q2, Q3, Q4). Flagellated cells were counted in each quarter, and the quarter with the highest number was aligned to Q1. Mean and standard deviation were calculated and represented ($n_{\text{each}} = 5$ cells).

Article

Extended Data Table 1 | Enriched GO terms in differentially expressed genes in *C. perkinsii* replicate B compared to A and C replicates

GO.ID	Term	Annotated	Significant	Expected	weight
GO:0015074	DNA integration	174	60	19.61	1.70E-16
GO:0075713	establishment of integrated proviral latency	40	18	4.51	7.20E-08
GO:0044826	viral genome integration into host DNA	40	18	4.51	7.20E-08
GO:0032196	transposition	40	17	4.51	4.60E-07
GO:0046718	viral entry into host cell	57	21	6.42	2.10E-06
GO:0042133	neurotransmitter metabolic process	15	9	1.69	7.50E-06
GO:0009083	branched-chain amino acid catabolic process	21	10	2.37	3.40E-05
GO:0006641	triglyceride metabolic process	54	17	6.08	4.00E-05
GO:0075732	viral penetration into host nucleus	27	11	3.04	8.20E-05
GO:0010438	cellular response to sulfur starvation	6	5	0.68	9.80E-05

GO term enrichment is indicative of a viral infection, which explains the distinct RNAseq profile and exclusion of B replicate compared to that of A and C.

Extended Data Table 2 | Developmental time points and time averaging windows for animal gene expression comparisons from the one or 2-cell stage through the blastula stage

TwoLetterCode	Species	Blastula time point cutoff	Time point windows for averaging
PT	<i>P. dumerilii</i>	14	1-4, 4-6, 6-8, 8-10, 10-12, 12-14
FW	<i>S. polychroa</i>	9	1-3, 3-5, 5-6, 6-7, 7-8, 8-9
CE	<i>C. elegans</i>	20	1-5, 5-8, 8-11, 11-14, 14-17, 17-20
HD	<i>H. dujardini</i>	15	1-4, 4-7, 7-9, 9-11, 11-13, 13-15
DM	<i>D. melanogaster</i>	9	1-3, 3-5, 5-6, 6-7, 7-8, 8-9
SU	<i>S. purpuratus</i>	22	1-5, 5-9, 9-13, 13-16, 16-19, 19-22
ZF	<i>D. rerio</i>	6	1, 2, 3, 4, 5, 6
NV	<i>N. vectensis</i>	10	1-3, 3-5, 5-7, 7-8, 8-9, 9-10
AQ	<i>A. queenslandica</i>	11	1-3, 3-5, 5-7, 7-9, 9-10, 10-11
ML	<i>M. leidyi</i>	11	1-3, 3-5, 5-7, 7-9, 9-10, 10-11

The number of time points was selected based on developmental information in Levin et al.⁴¹.

Reporting Summary

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

Statistics

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

n/a Confirmed

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

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection All data was collected on commercial microscopes running the following acquisition software: NIS-Elements (Nikon), LAS X (Leica).

Data analysis Image analysis was performed in Fiji running ImageJ and Imaris (v992). Transcriptome analyses were carried out using Salmon (v1.1.0), R programming environment (v4.3.1), tximport (v1.30.0), edgeR (v4.0.3), BLASTP (v2.10.0+), topGO (v2.54.0), OrthoFinder (v 2.5.4), mafft (v7.487) and hmmer (v3.3.2).

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

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

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

All data used for quantifications, movies and the generated code

For transcriptomic analysis is accessible on figshare (<https://figshare.com/s/f20f6d471c719990471c>) with the raw images available upon request. Raw

RNAseq data are on <https://tinyurl.com/PRJNA1091032>. Note that the data and code archives are embargoed while the manuscript is under review.

Research involving human participants, their data, or biological material

Policy information about studies with [human participants or human data](#). See also policy information about [sex, gender \(identity/presentation\), and sexual orientation](#) and [race, ethnicity and racism](#).

Reporting on sex and gender

NA

Reporting on race, ethnicity, or other socially relevant groupings

NA

Population characteristics

NA

Recruitment

NA

Ethics oversight

NA

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

Field-specific reporting

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

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

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

Life sciences study design

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

Sample size

All data in this study represented either qualitative or quantitative analyses extracted from unmodified microscopy images. For fluorescence microscopy images minimum of 5-10 cells per condition/timepoint per species were obtained. Following initial analysis, additional analysis was collected for any condition with notably lower counts. Sample sizes were determined to enable rigorous testing of hypotheses raised by preliminary experimental data. Data acquisition was continued until further acquisition was causing no measurable difference in results.

Data exclusions

No data were excluded from the analyses.

Replication

All experiments were repeated with two or three independent biological replicates tested on different days with different source cultures, as specified in legends and methods section.

Randomization

This is not applicable for our study. All microscopy images were obtained randomly from synchronized cell populations at different days.

Blinding

Data were analysed where possible by automated or semi-automated methods. In cases where manual filtering or correction was incorporated, same criteria were applied to all strains or treatments under investigation.

Reporting for specific materials, systems and methods

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

Materials & experimental systems

n/a	Involved in the study
<input type="checkbox"/>	<input checked="" type="checkbox"/> Antibodies
<input checked="" type="checkbox"/>	<input type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology and archaeology
<input type="checkbox"/>	<input checked="" type="checkbox"/> Animals and other organisms
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data
<input checked="" type="checkbox"/>	<input type="checkbox"/> Dual use research of concern
<input checked="" type="checkbox"/>	<input type="checkbox"/> Plants

Methods

n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input checked="" type="checkbox"/>	<input type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging

Antibodies

Antibodies used

Anti-Tub E7 antibody (DSHB), Anti- Acetylated Tubulin (Sigma; T6793)
 Goat anti-Mouse Secondary Antibody, Alexa Fluor 488(Thermo A-11001)

Validation

All primary antibodies used in the study are well established in the field and frequently used for immunofluorescence of the targets.
 All antibodies were supported by manufacturer's validation statements for purposes utilised in the study.

Anti- Tub E7 was previously validated in various single-celled eukaryotes :
 Choanoflagellates : <https://www.sciencedirect.com/science/article/pii/S0960982213011275?via%3Dihub> , <https://journals.plos.org/plosbiology/article?id=10.1371/journal.pbio.3000226>
 Leishmania chagasi: <https://www.sciencedirect.com/science/article/pii/S0021925820890899?via%3Dihub>
 Ichthyosporea: <https://www.biorxiv.org/content/10.1101/2023.05.10.540163v1>

Anti- Acetylated Tubulin (Sigma; T6793) was previously validated in various single-celled eukaryotes :
 Dinoflagellates : <https://www.biorxiv.org/content/10.1101/2023.10.17.562701v1.full.pdf>
 Choanoflagellates: <https://evodevojournal.biomedcentral.com/articles/10.1186/s13227-016-0042-x>
 Filasterea: <https://www.sciencedirect.com/science/article/pii/S0960982215008878>

Animals and other research organisms

Policy information about [studies involving animals](#); [ARRIVE guidelines](#) recommended for reporting animal research, and [Sex and Gender in Research](#)

Laboratory animals

The study did not involve any laboratory animals- only lab strains of the unicellular models Ichthyosporean Chromosphaera perkinsii

Wild animals

This study did not involve wild animals

Reporting on sex

For non-animal model organisms, reporting on sex is not relevant.

Field-collected samples

This study did not involve field-collected samples

Ethics oversight

Ethics approval is not required for research using unicellular holozoa.

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