



Genome-wide analysis of mRNA regionalization in a giant single cell

Ashley R. Albright^{a,b} , Connie Yan^{a,b} , David Angeles-Albores^c, Yina Hudnall^{a,b} , Tatyana Makushok^{a,b}, Jamarc Allen-Henderson^{a,b}, and Wallace F. Marshall^{a,b,1}

Edited by J. Matthew Taliaferro, University of Colorado Anschutz Medical Campus Alzheimer's and Cognition Center, Aurora, CO; received December 29, 2025; accepted May 12, 2026 by Editorial Board Member Susan Strome

The molecular mechanisms that produce cell structure are poorly understood. The giant ciliate *Stentor coeruleus* is a unicellular model organism whose large size, reproducible structure, and ability to heal wounds and regenerate allow the formation of structure in a single cell to be addressed using methods of experimental embryology. Such studies have shown that specific cellular structures always form in particular regions of the cell, which raises the question: what is the source of positional information within this organism? By analogy with embryonic development, in which regionalized mRNA is often used to mark position, we asked whether specific regionalized mRNAs might mark position along the anterior–posterior axis of *Stentor*. By physically bisecting cells and conducting bulk RNA sequencing, we were able to identify sets of messages enriched in either the anterior or posterior half and show that RNAi-mediated knockdown of posterior-enriched transcripts corresponding to MYB genes results in a cell's inability to regenerate the posterior portion of the cell body. We then conducted half-cell RNA-sequencing in paired anteriors and posteriors of cells in which the microtubule cytoskeleton was disrupted by RNAi of β -tubulin or dynein intermediate chains. We found that many messages either lost their regionalized distribution or switched to an opposite distribution, such that anterior-enriched messages in control became posterior-enriched in the RNAi cells, or vice versa. This study indicates that mRNA can be regionalized within a single giant cell and that microtubules may play a role, possibly by serving as tracks for the movement of the messages.

cytoskeleton | pattern formation | ciliates | mRNA localization | single cell

Understanding how complex patterns arise within individual cells is a challenging biological problem due to the vast difference in spatial scales from molecules to whole cells (1, 2). The questions that arise in understanding cellular morphogenesis are essentially the same as those that arise in understanding the development of embryos—how does the system know which structures to make and how to position them relative to an overall body plan? A classic model organism that facilitates the study of developmental biology in a single cell is the giant ciliate *Stentor coeruleus* (3, 4). *Stentor* cells are 1 mm long and covered with blue stripes of pigment (Fig. 1A), which alternate with longitudinal rows of cilia consisting of basal bodies associated with microtubule bundles (KM-fibers) that run the length of the cell (5). The cone-shaped cell has an array of cilia known as the membranellar band (MB) at its anterior end and a holdfast at the posterior. Orthogonal to the anterior–posterior (AP) axis defined by these structures, the pigment stripes show a graded distribution of width, defining a circumferential axis. Starting with the widest stripes, the stripe width gradually decreases as one moves counterclockwise around the circumference, until eventually, the narrowest stripes meet the widest stripes, creating a discontinuity known as the contrast zone.

Stentor was first used as a model system for its ability to regenerate after surgical manipulations. If any part of the *Stentor* cell is cut off, the missing piece regenerates in a matter of hours to yield a normal cell. If a cell is cut in half, each half regenerates its normal structure (7, 8). *Stentor* contains an elongated macronucleus that contains many copies of every gene, which likely contributes to the cell's ability to survive these surgical operations. *Stentor* can also employ a range of mechanical wound-closure mechanisms to keep its cytoplasm from leaking out while it patches over a wound (9). *Stentor*'s huge size, easily visible patterning, and amazing powers of regeneration attracted many developmental biologists during the turn of the last century, including Thomas Hunt Morgan (8). For almost a century, up until the 1970s, microsurgery was used to analyze morphogenetic processes in *Stentor*, resulting in a wealth of information about how cells respond to geometrical rearrangements (4, 10). But *Stentor* was never developed as a molecular model

Significance

Just as embryos develop into complex forms, single cells also form complicated structures, but much less is known about pattern formation in cells. We used a single-celled organism, *Stentor*, to ask how a cell forms different structures in different places. This model organism is uniquely powerful because its large size means that even a piece of the cell has enough material for genomic analysis, and because of its ability to regenerate after being cut in half, without the contents spilling out. By combining microsurgery with modern genomic methods, we find that mRNA produced by different genes accumulates in different regions of the cell, showing that single cells use similar mechanisms for developing patterns as seen in animal embryos.

Author affiliations: ^aDepartment of Biochemistry and Biophysics, University of California San Francisco, San Francisco, CA 94158; ^bCenter for Cellular Construction, University of California San Francisco, San Francisco, CA 94158; and ^cThe Organogenesis and Tissue Replacement Lab, San Francisco, CA 94107

Author contributions: A.R.A., T.M., and W.F.M. designed research; A.R.A., C.Y., Y.H., T.M., J.A.-H., and W.F.M. performed research; A.R.A. contributed new reagents/analytic tools; A.R.A., C.Y., and D.A.-A. analyzed data; and A.R.A. and W.F.M. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission. J.M.T. is a guest editor invited by the Editorial Board.

Copyright © 2026 the Author(s). Published by PNAS. This open access article is distributed under Creative Commons Attribution License 4.0 (CC BY).

¹To whom correspondence may be addressed. Email: wallace.marshall@ucsf.edu.

This article contains supporting information online at <https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2537760123/-/DCSupplemental>.

Published June 15, 2026.

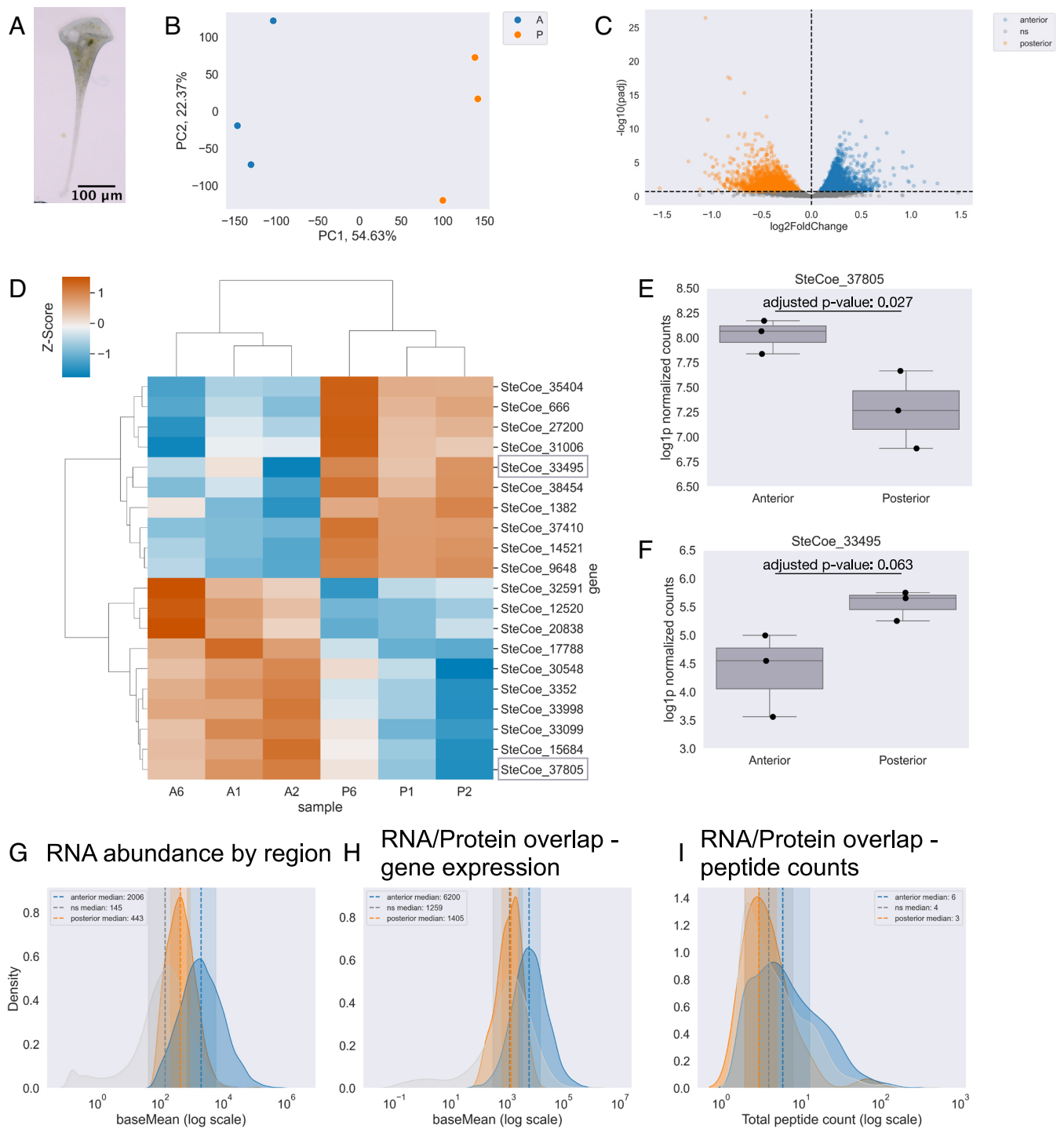


Fig. 1. Bulk RNA sequencing reveals RNAs are patterned along the AP axis in *Stentor coeruleus*. (A) *Stentor coeruleus*. (B) Principal Component Analysis (PCA) showing a clear separation of anterior (blue) and posterior (orange) samples based on their mRNA abundance profiles. (C) Volcano plot displaying results of differential enrichment analysis of anterior vs. posterior mRNA abundances. Points correspond to individual genes. Statistically significantly differentially enriched genes (q -value < 0.2) in the anterior are blue, posterior in orange, and not significant in gray. (D) Heatmap of relative abundance (z-scores) where rows correspond to individual genes and columns are samples for the top 10 enriched genes in anterior or posterior samples. Red indicates higher enrichment, and blue indicates relative depletion. Representative genes in (E and F) are circled in gray. (E) Representative example and boxplot of log-normalized counts of a posterior-enriched message, SteCoe_37805. (F) Representative example and boxplot of log-normalized counts of a posterior-enriched message, SteCoe_33495. (G) Kernel density estimate of mean normalized counts for all transcripts with baseMean > 0 (n : anterior = 3,440, posterior = 2,836, ns = 23,738). (H) Kernel density estimate of mean normalized counts for transcripts with corresponding proteins detected in ref. 6 (n : anterior = 859, posterior = 40, ns = 829). (I) Kernel density estimate of total peptide counts from ref. 6 with corresponding transcripts in panel H. Panels (G–I) are colored according to the identity as determined by our analysis: anterior-enriched (blue), posterior-enriched (orange), and nonsignificant (gray). Dashed vertical lines indicate group medians; shaded bands indicate interquartile ranges.

system, and we do not currently know the molecular basis for how it can develop or regenerate its structures. To exploit the unique opportunities presented by this organism, we have previously sequenced the *Stentor* genome (11), developed methods for RNAi

(12), and analyzed the transcriptional program of regeneration (13). Proteomic analysis was also reported for isolated MB (14), and transcriptomic analysis was reported for bisected cells of the related species *Stentor polymorphus* (15).

The source of positional information is perhaps the most fundamental question about pattern formation in any organism. That such information exists in *Stentor* can be inferred from the fact that cellular structures always form in defined and reproducible positions (4). For example, during regeneration or normal cell division, the new MB always forms halfway down the AP axis at the contrast zone between narrow and wide stripes (16). It is generally true in ciliates that specific surface structures form at positions that can be defined with respect to a cylindrical coordinate system (17–19). In this system, position is defined by two orthogonal axes—an AP axis that runs along the length of the cell, and a circumferential axis that runs around the equator, which in *Stentor* is manifest by the gradient of pigment stripe width. We have previously analyzed the distribution of the proteome relative to the AP axis by cutting individual *Stentor* cells into pieces and analyzing the fragments by mass spectrometry (6). This study estimated that approximately 25% of the proteome was polarized relative to this axis.

The role of regionalized mRNA as a source of positional information in embryonic development is well established, particularly in *Drosophila*. We anticipate that regionalized RNAs in different parts of the *Stentor* cell will be important sources of positional information, just like they are in embryos. Unlike the case of *Drosophila* development, where many important genes had already been identified in genetic experiments and served as candidates for localized messages, we do not currently have a list of candidate positional information molecules to map. Inspired by prior work in cryoslicing *Drosophila* embryos to obtain positional information (20, 21), we can use a similar unbiased approach to map the entire transcriptome to identify genes that might have a regionalized distribution.

The unbiased nature of a sequencing-based method for probing localization stands in contrast to microscopy-based methods. Many studies survey message localization using fluorescent in situ hybridization (FISH), single-molecule FISH (smFISH) (22), or multiplexed error-robust FISH (merFISH) (23) in cells that lack an obvious polarity axis, so collecting physical fragments of cells representing specific regions and mapping fragments back to any reference frame would not be possible. Consequently, in those cases, each message must be visualized sequentially by imaging, which is time-consuming and expensive. The giant size of *Stentor*, combined with the reproducible cell geometry and easily visualized body axis markers, makes it possible to use a sequencing-based method to map RNA regionalization relative to the body axis coordinates.

The first question is whether genes are regionalized in *Stentor*. If it turns out that RNAs are regionalized to different parts of the *Stentor* cell, the next question is: how might this be achieved? The surface of a *Stentor* is covered with parallel microtubule bundles. Studies in other systems, including yeast and human neurons, show that the cytoskeleton is involved in active RNA transport (24, 25); therefore, we hypothesize that transcripts in this giant cell may be polarized by trafficking specific messages along the cell's polarized cytoskeletal tracks.

Here, we investigate the regionalization of the *S. coeruleus* transcriptome using bulk RNA sequencing of anteriors and posteriors and show that roughly 15% of the transcriptome is polarized along the AP axis. Many of these polarized transcripts encode transcription factors, and we show that knockdown of two MYB transcription factor genes whose transcripts are enriched in the posterior half of intact cells prevents proper posterior regeneration. We also show that knockdown of both β -tubulin and cytoplasmic dynein intermediate chains (henceforth simply referred to as dynein) causes cells to display abnormal cell morphology. To investigate

whether disrupting the microtubule cytoskeleton disrupts RNA regionalization, we conducted half-cell RNA sequencing in paired anteriors and posteriors upon knockdown of β -tubulin or dynein. Overall, we find evidence for a global disruption in RNA regionalization upon perturbation of the microtubule cytoskeleton and identify many candidate transcripts with an anticorrelated shift in AP skew.

Methods

Stentor Culturing. *S. coeruleus* were originally obtained from Carolina Biological Supply and have been maintained in the Marshall lab in small Pyrex dishes at room temperature in filter-sterilized pasteurized spring water (Carolina Biological Supply—Burlington, NC), subsequently referred to as Carolina Spring Water (CSW). Once or twice a week, a lentil-sized pellet of *Chlamydomonas reinhardtii* grown on TAP plates (26) was resuspended in CSW and fed to *Stentor*. Before RNAi, *Stentor* were isolated from the main culture in 6-well tissue culture plates containing 2 mL CSW, starved for 3 d, and washed 3 times with CSW to eliminate residual *Chlamydomonas*.

RNAi. For sequencing experiments in cells with disrupted microtubules, β -tubulin RNAi was conducted by feeding as previously described (12) with minor changes. RNAi constructs were synthesized by Radiant Inc. using sequences from the *Stentor* genome. The sequence targeting the β -tubulin gene, SteCoe_5719, was amplified with the following oligos: forward—5' ATGGTGACTTAAATCACTGGTGTAGTC 3', reverse—5' TTAAGCAGCTTCCTCTTCCTCATCG 3'. Although *Stentor* has six β -tubulin genes, the high level of sequence conservation is such that this RNAi construct likely targets multiple genes in the family, based on the observation of multiple stretches greater than 20 base identity, which is also supported by our previous demonstration of loss of β -tubulin at the protein level by immunofluorescence (11). Transformed HT115 *Escherichia coli* were grown to log phase, and expression of dsRNA was induced with 1 mM IPTG overnight at room temperature. Overnight 50 mL cultures were spun down, washed with CSW, and resuspended in 30 mL of CSW. Aliquots of 1 mL were pelleted, with as much supernatant as possible removed, and frozen at -80°C for future use. For feeding, pellets were resuspended in 1 mL CSW. *Stentor* were fed 500 μL control (empty vector) or β -tubulin RNAi bacteria once a day for 7 d.

For sequencing experiments in cells with disrupted dynein motors, dynein intermediate chain RNAi was conducted as above. G04 targets SteCoe_24570 and SteCoe_13560 amplified with forward—5' AAGCAGGAACGTGAAGATGC 3', and 5' TGCCGATTATTGGGAAAGTG 3'. G05 targets SteCoe_14054 amplified with forward—5' AAAACGCTTGCTGAGCTTC 3' and reverse—5' TGGATGAGTATGGCTTCTGC 3'.

For experiments to obtain representative images (brightfield and fluorescence) as well as screening Myb DNA-binding domain genes, we used a recently developed method to deliver RNAi by microinjection (27). This method is more time-efficient and has a higher efficacy than RNAi by feeding, but was not yet established at the time the sequencing experiments were conducted. SteCoe_15300 and SteCoe_33495 constructs corresponding to the first half of each gene were synthesized using Twist Biosciences. Plasmids (2 μg) were linearized with PvuI-HF (NEB), gel-purified with the ZymoClean Gel DNA Recovery Kit (Zymo Research), and eluted in 6 μL of water. One μg of linear template was used for in vitro transcription with the TranscriptAid T7 High Yield Transcription Kit (Thermo). *Stentor* were incubated in 1.5% methylcellulose for 2 d before injection. dsRNA at a concentration between 100 and 1,000 ng/ μL was coinjected into cells with fluorescent dextran using a Nanoject II or Nanoject III (Drummond) and pulled capillary needles as described previously (28).

Dynein Phylogeny. We generated a fasta file containing protein sequences for SteCoe_24570, SteCoe_13560, and SteCoe_14054 as well as sequences from *Homo sapiens*, *Tetrahymena thermophila*, and *Chlamydomonas reinhardtii*. We performed multiple sequence alignment using a ClustalW (29) web server (<https://www.genome.jp/tools-bin/clustalw>). We then used the web server to run PhyML (30) with 100 bootstraps, followed by FigTree for visualization (<https://github.com/rambaut/figtree/releases/tag/v1.4.4>).

Bisection. Bisection is conducted as described previously using pulled capillary needles under a dissection scope for inducing regeneration (31); however, cells

were immediately processed for RNA extraction or sequencing rather than kept for observation.

Brightfield and Fluorescence Microscopy. For brightfield images, cells were placed in a 13.6 μ L droplet onto a slide with a 9 mm diameter $\times$ 0.12 mm depth SecureSeal spacer (Grace Bio-Labs) and covered with a coverslip. Images were collected on an AxioZoom V16 (Zeiss) with a 2.3 $\times$ objective at 32 $\times$ magnification and a Nikon Rebel T3i SLR Camera. MYB knockdown cells were imaged in low-light settings so as not to disturb the cells and allow them to fully extend. Brightness was adjusted during processing for visualization purposes.

For fluorescence images, MeOH fixed cells were stained with Hoechst 33342 (Thermo), mounted with DABCO in 70% glycerol in PBS, and pipetted onto a coverslip in a 13.6 μ L droplet. A slide with a spacer was placed onto the coverslip droplet upside down to ensure the cells remained close to the coverslip for imaging. Images were collected on a Nikon Ti-E inverted microscope with a 25 $\times$ silicone objective with a 405 nm or 561 nm laser, Hamamatsu Quest camera, and a VT-iSIM superresolution module. In the images shown, we imaged autofluorescence from the Stentorin pigment, which forms stripes on the cell surface, as a proxy for cortical organization.

Sequencing.

Bulk RNA sequencing. Within 20 min of bisection, 50 anteriors and 50 posteriors were placed in Trizol with as little CSW leftover as possible. RNA was extracted using a Direct-zol RNA Purification Kit (Zymo), and libraries were prepped using the NEBNext Ultra II RNA Library Prep Kit (NEB) following the manufacturer's instructions.

Bulk data processing and analysis. We used kallisto (32) to generate a reference index of predicted protein coding genes and quantify transcript abundance (SI Appendix, Table S1), sleuth to prepare the data (33), ComBat (34) to batch correct raw estimated counts (SI Appendix, Table S2), and pydeseq2 (35) to perform differential enrichment analysis with a permissive q-value cutoff for significance of 0.2 (SI Appendix, Table S3).

PCA. PCA was performed using the scikit-learn implementation (sklearn.decomposition.PCA) in Python. Prior to dimensionality reduction, data were standardized using StandardScaler (sklearn.preprocessing.StandardScaler). PCA was then applied to the transposed expression matrix (samples as observations, genes as features) and reduced to two components. For bulk transcriptomic analyses, this workflow was applied directly to the normalized expression matrix. For tubulin and dynein analyses, the same procedure was implemented via a custom utility function [get_pc, available in utils.py on Zenodo (36)].

Domain enrichment analysis. To identify protein domains overrepresented among anteriorly or posteriorly enriched transcripts, we performed a one-sided Fisher's exact test for each group. The background gene set consisted of all genes with nonzero expression and the test set consisted of transcripts significantly upregulated in the anterior or posterior. Genes were grouped by annotated protein domain according to ref. 11. Raw *P*-values were corrected for multiple testing with Benjamin-Hochberg correction. Anterior (SI Appendix, Table S3) and posterior (SI Appendix, Table S4) domain enrichment was tested separately.

Half-Cell RNA sequencing. Individual *Stentor* that displayed phenotypes were washed and transferred onto parafilm in a 2 μ L droplet of CSW and bisected laterally using a needle and processed within 20 min of bisection. Cells were lysed and prepped for sequencing with the NEBNext Single Cell/Low Input RNA Library Prep kit for Illumina (Cat. No. E6420) according to the manufacturer's protocol. Samples were pooled and submitted for sequencing on an Illumina NovaSeq6000 (SP200) or NovaSeqX (SP300). The experiments with β -tubulin and dynein RNAi were done separately, therefore each experiment has its own control samples.

Half-Cell data processing and analysis. Adapter trimming was performed using flexbar (37) as instructed by NEB for the NEBNext Single Cell Library Prep Kit: <https://github.com/nebiolabs/nebnext-single-cell-ma-seq>. We used kallisto (32) to generate a reference index and quantify transcript abundance, and sleuth (33) to prepare the data (tubulin—SI Appendix, Table S6 and dynein—SI Appendix, Table S8).

To reduce the effects of outlying transcript abundance on significance, we adjusted log-transformed transcripts per million (TPM) values by adding the sample mean log-transformed TPM. AP skew is then defined as the difference over the sum of adjusted log1TPM where log1TPM refers to the log base 2 of the TPM plus one.

$$Skew = \frac{A_{\log 1TPM(adj)} - P_{\log 1TPM(adj)}}{A_{\log 1RPM(adj)} + P_{\log 1TPM(adj)}}$$

Genes were selected for further analysis that passed a minimal cutoff of abundance, to ensure sufficiently robust statistical comparisons, and it was also required that genes had a coefficient of variation (variance divided by the mean) less than 2 to remove genes with such variable distributions that comparisons might be unreliable. For genes meeting these criteria, skews were compared for both control and β -tubulin RNAi samples, or control and dynein RNAi samples.

We were particularly interested in genes with anticorrelated skews, the most interesting candidates for disrupted regionalization. Such genes were identified with an unbiased feature selection by PCA, in which the genes were treated as factors and the samples as data points, to identify linear combinations of genes that give the highest variation among the samples. Feature selection ranks genes by their contributions (loadings) to the principal components that best separate samples along the AP axis. Feature selection was then followed by a *t* test for significance (tubulin—SI Appendix, Table S7 and dynein—SI Appendix, Table S9).

Results

Identification of Regionalized Messages in Stentor. To ask whether any genes have a regionalized distribution relative to the AP axis, we manually cut *Stentor* cells in half, using the MB as a marker to define the anterior end, and cutting approximately mid-way between the MB and the other end of the cell. For each sample, 50 anteriors and 50 posteriors were processed in bulk RNA-sequencing. The first principal component (PC1) of the PCA captures the primary variation in the dataset, with data points corresponding to anterior samples positioned distinctly from posterior samples, suggesting that PC1 effectively represents the AP axis (Fig. 1B). Using a permissive q-value threshold of 0.2 for statistical significance and future exploratory analysis of candidate genes, we identify 6,276 genes that are statistically significantly differentially enriched between the anterior and posterior half of the cell (Fig. 1C and SI Appendix, Table S3). This threshold allows for exploratory analysis of more candidate genes in the future, although we acknowledge that a q-value of 0.2 may result in a higher false discovery rate. Incorporating the 20% false discovery rate into our calculation, our result corresponds to approximately 15% of the total number of gene models.

Fig. 1D depicts a heatmap of the top 10 enriched genes in each direction based on log fold change. Log fold change in the positive direction corresponds to messages predominantly enriched in the anterior, with the greatest difference in abundance in SteCoe_37805, SteCoe_33099, and SteCoe_15684. SteCoe_37805 (Fig. 1E) and SteCoe_33099 are 91.2% identical and are predicted to encode membrane transporter proteins. The SteCoe_15684 gene product is a predicted cadherin domain-containing protein. Log fold change in the negative direction corresponds to messages predominantly enriched in the posterior, with the greatest difference in relative abundance in SteCoe_33495 (Fig. 1F), SteCoe_38454, and SteCoe_31006. The protein product of SteCoe_33495 contains Myb-like domains, SteCoe_38454 encodes an FCP1 homology domain-containing protein, and SteCoe_31006 encodes an RGS domain-containing protein. Transcripts identified as being differentially regionalized in the anterior have a higher median expression than the posterior (Fig. 1G). The full list of genes analyzed, along with significance data for regionalized genes, is given in SI Appendix, Table S3.

One hypothesis for why mRNA might be regionalized in different parts of the cell would be to support local translation of proteins needed to build structures in the corresponding region of the cell. We previously reported a proteomic analysis of dissected *Stentor* cells and found sets of proteins that were strongly enriched in either the MB, the anterior of the cell from which the

MB had been removed, or in the posterior of the cell (6). Given the huge size of a *Stentor* cell, local translation may be necessary to make protein localization more efficient, or else enrichment of transcripts in some region might be a mechanism for producing regional differences in protein composition. In either case, the local translation hypothesis would predict a strong correlation between these regionalized proteins and regionalized mRNAs. However, proteomic analyses are limited by a bias toward relatively abundant proteins, most likely resulting in fewer shared gene products with RNA-sequencing.

Considering the total set of proteins identified by mass spectrometry of dissected *Stentor* cells (6), 899 detected proteins were found to correspond to RNAs that we classified as either anterior or posterior in the analysis above. Out of these, 859 are anterior, and 40 are posterior. For each transcript, we calculated the average abundance across all samples. The median value of these averages was higher for the set of transcripts corresponding to proteins detected by mass spectrometry of dissected cells (6) than for all RNAs (Fig. 1H), indicating that proteins detected by proteomics tend to correspond to more abundant messages. There is also a bias in median total peptide counts for proteins encoded by anterior transcripts compared to proteins encoded by posterior transcripts (Fig. 1I).

The fact that proteomics is biased toward detecting more abundant proteins, combined with the tendency of anterior transcripts to encode abundant proteins, creates a tendency for anterior transcripts to encode anterior proteins, but makes it impossible to assess the opposite tendency for posterior transcripts. Increased proteomic coverage in future studies will be required to assess the degree of correlation between mRNA regionalization and protein localization.

One of the key structural components of the anterior *Stentor* cell is the MB. A proteomic analysis of isolated MBs (14) identified candidate proteins for components of this structure. When we compared their list of likely MB proteins with mRNA regionalization data (SI Appendix, Fig. S1), we found that 440 corresponded to anterior-enriched transcripts, 13 to posterior-enriched transcripts, and 335 to transcripts that did not show significant regionalization. This result could reflect a tendency for anterior transcripts to encode MB proteins, but it also may simply reflect the general trend, noted above, for anterior regionalized transcripts to correspond to more highly abundant proteins.

The demonstration that a set of transcripts are differentially regionalized raises the question of whether or not regionalized transcripts play any functional role in generating regionally different patterns within the cell. A domain enrichment analysis was performed as described in *Methods*, with results given in SI Appendix, Tables S4 and S5, which describe domains enriched in proteins encoded by the anterior and posterior transcript sets, respectively. Based on this domain analysis, there is a clear tendency for anterior transcripts to encode highly abundant structural or metabolic proteins like Cpn60/TCP1, proteasome, ribosomes, and mitochondrial ATP synthase. In contrast, posterior transcripts have a clear tendency to encode proteins with signaling or regulatory functions based on their domain structure, including kinase and MYB family members.

Regionalized Transcription Factors Are Necessary for Proper Cell Morphology. Of the many differentially enriched transcripts, 50 correspond to Myb DNA-binding domain proteins, 10 of which are anterior-enriched and 40 of which are posterior-enriched (SI Appendix, Table S3). Indeed, one of the most posterior-enriched transcripts discussed above, SteCoe_33495, corresponds to a transcription factor with a Myb DNA-binding domain. Genes

in this expanded family of transcription factors are known to control morphogenesis and patterning in plant cells (38) and are well-known proto-oncogenes in humans (39, 40). To ask whether Myb DNA-binding domain transcription factors are important for cellular patterning in *Stentor*, we conducted a targeted RNAi screen of MYB genes found in our study and asked whether these cells regenerated normally.

In each case, following injection of dsRNA, cells were bisected into anterior and posterior half-cells. *Stentor* typically regenerates 8 h after bisection; however, here we imaged cells at 24 h postbisection to ensure they had sufficient time to completely regenerate. For controls we used LF4, a gene that affects ciliary length but not cell morphology, as well as MOB1, which has characteristic effects on cell morphology, to confirm the efficacy of the RNAi procedure. Cells were only counted when the image provided a clear view of the cell in which the MB and posterior were visible. We found that a considerable number of MYB knockdown cells were unable to regenerate a posterior, unlike either of our controls (LF4 and MOB1). As shown in Fig. 2, RNAi targeting either Myb-encoding gene causes a fraction of anterior cells to lack the characteristic tapered posterior tail region that gives *Stentor* its trumpet-like shape. Additional images showing how the shapes are classified in anterior half-cells are present in SI Appendix, Fig. S2. This phenotype was seen in 26% of SteCoe-15300 (6/23) and 40% of SteCoe_22495 (8/20) RNAi anterior half-cells, compared to 10% (2/20) of LF4 and 5% (1/19) of Mob1 anterior halves. In posterior half-cells, this apparent phenotype occurred at a frequency matching what was seen in the controls.

We can imagine three possible functions for localized mRNA in *Stentor* regeneration. First, the mRNA might be localized to the region where it is functionally required, for example to promote local translation of proteins. Second, the mRNA might be localized to a region of the cell because it is being stored there, and will be used after bisection to quickly form structures needed to regenerate the other half of the cell. Third, localization might be functionally irrelevant.

When a cell is bisected, the anterior half-cell has to regenerate its posterior structures, while the posterior half-cell has to regenerate anterior structures. If messages are localized in the region where they are required to make structures specific for that region, then knockdown of posterior localized mRNA would be expected to affect formation of posterior structures. This takes place in the anterior half-cell. So in that case, the prediction is that anterior half cells would be impaired in regenerating posterior structures, while the posterior half-cells which already have posterior structures, would not be impaired in regenerating anterior structures. If, on the other hand, messages are stored in one half of the cell to allow it to regenerate the other half, then knockdown of posterior localized mRNA would be expected to affect formation of anterior structures by posterior half-cells. So in that case, the prediction is that posterior half-cells would be impaired in regenerating anterior structures, while the anterior half-cells which already have their anterior structures, would not be impaired in regenerating posterior structures.

Our results clearly support the first model. Knockdown of SteCoe_15300 and SteCoe_33495 results in a strong defect in regeneration of the posterior structures in the anterior half-cell following bisection, indicating a requirement for these posterior-enriched genes in regeneration of posterior structures. While these transcripts are enriched in the posterior region under steady-state conditions, their depletion leads to impaired regenerative capacity specifically in anterior half-cells, which have to form new posterior structures. In contrast, the posterior half-cell, which only has to regenerate anterior structures, is unaffected by knockdown of posterior-localized mRNA.

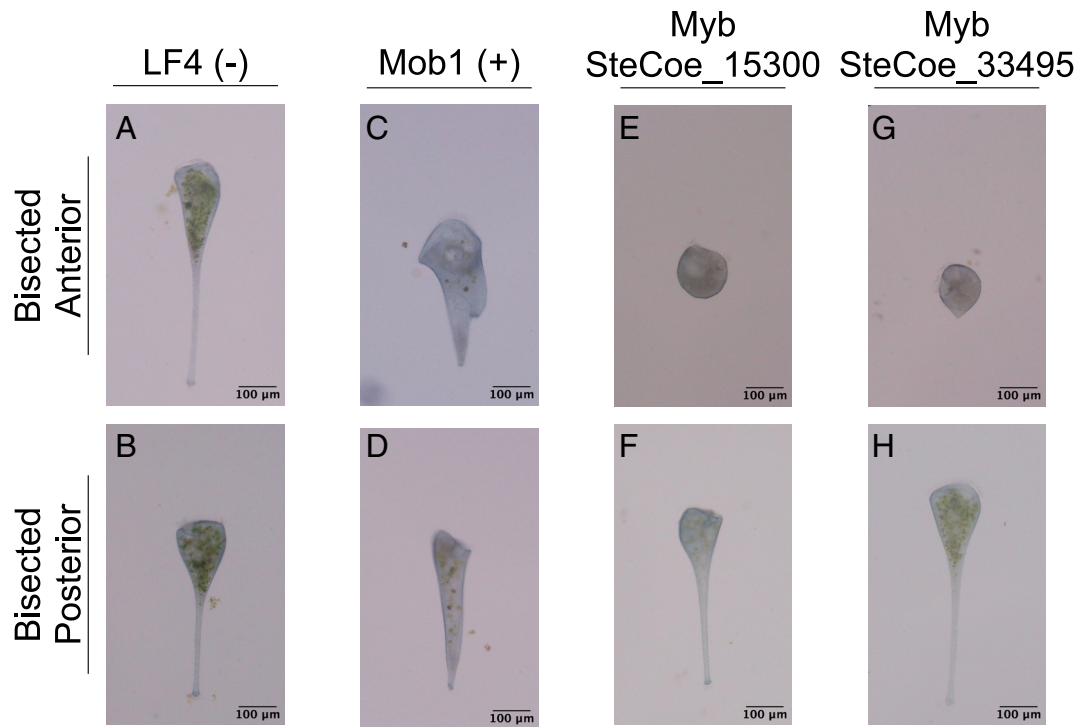


Fig. 2. Myb DNA-binding transcription factor knockdown impairs posterior regeneration. Brightfield images of anterior and posterior knockdown cells 24 h postbisection. The negative control, LF4 (A–anterior and B–posterior) corresponds to a gene uninvolved in regeneration. The positive RNAi control, Mob1 (C–anterior and D–posterior) corresponds to a gene important for posterior cell morphology but with a distinct phenotype (extra tails and a less tapered body) from what was observed with the Myb constructs SteCoe_15300 (E–anterior and F–posterior) and SteCoe_33495 (G–anterior and H–posterior) correspond to MYB genes.

We therefore favor a model in which posterior-enriched transcripts contribute to establishing or maintaining posterior identity and/or provide components required for posterior regeneration. In this framework, the defect observed in bisected anteriors reflects the inability to properly reestablish posterior structures and polarity rather than a requirement for these gene products in anterior cellular function.

Role of Microtubules in Cell Morphology. In many cell types, mRNA can be moved along microtubule tracks. If regionalized transcripts are important for cell morphogenesis, we would expect disruption of such tracks to affect cell shape. Prior work has shown that RNAi knockdown of β -tubulin causes cells to lose their typical structure and orderly microtubule bundles (12). Here, we replicate the findings that knocking down β -tubulin causes cells to lose their characteristic shape and resemble a tulip rather than a trumpet (Fig. 3 A and B). In addition, we show that knocking down either a pair of genes encoding a cytoplasmic dynein 1 intermediate chain 2 (SteCoe_24570 and SteCoe_13560) or a gene encoding cytoplasmic dynein 1 intermediate chain 1 (SteCoe_14054) also caused cells to lose their characteristic shape, as well as their ability to extend their holdfast (Fig. 3 C and D). SteCoe_24570, SteCoe_13560, and SteCoe_14054 were selected for functional characterization based on their classification as cytoplasmic dynein intermediate chains. Phylogenetic analysis of dynein intermediate chains places SteCoe_24570, SteCoe_13560, and SteCoe_14054 within the cytoplasmic dynein intermediate chain clade, distinct from axonemal dynein intermediate chains (SI Appendix, Fig. S3). This classification is further supported by PANTHER database annotation (41). Using autofluorescence of Stentorin, the pigment that gives *S. coeruleus* their blue-green color, we also observe aberrations in structure upon β -tubulin and dynein intermediate chain knockdown. Compared to normal *Stentor* (Fig. 3E), we see

an underdeveloped MB upon β -tubulin knockdown (Fig. 3F), although the orderly cortical rows appear normal. In Fig. 3 G and H, we see more extreme defects in the cortical structure upon dynein knockdown.

Despite the clear defects in β -tubulin and dynein intermediate chain knockdown cells, the cells remain alive for a period of time and retain a MB at one end and a holdfast at the other. It is thus still possible to recognize the AP axis, allowing us to repeat our analysis of mRNA regionalization by bisecting such cells to produce anterior and posterior half-cells and thereby ask whether disruption of the microtubule tracks or cytoplasmic dynein motors, which we hypothesize to carry messages, has any effect on mRNA regionalization.

Role of Microtubules in mRNA Regionalization. To determine whether the microtubule cytoskeleton plays a role in mRNA regionalization, we bisected cells and conducted half-cell RNA sequencing in paired anterior and posterior halves. Although these knockdown cells lose the elongated, tapering cell shape of wild-type cells, it was still possible to recognize the equator of the cell based on the longitudinal striping, MB, and the convergence of stripes at the posterior pole, enabling us to cut cells into anterior and posterior halves. Moreover, even in cells with normal morphology, as soon as we touch them with the needle, they immediately round up into a spherical shape, which is very similar in the tubulin and dynein knockdowns, so in all cases the cell morphology ends up being similar to the moment we cut the cells.

First, we performed a PCA on control samples to determine if anterior or posterior identity stratifies the samples as shown in Fig. 1B for bulk RNA sequencing. In this case, PC2 represents the AP axis (Fig. 4A). Next, we projected the β -tubulin knockdown data onto the control PC space (Fig. 4B) and found that the observable distinction between anterior and posterior samples

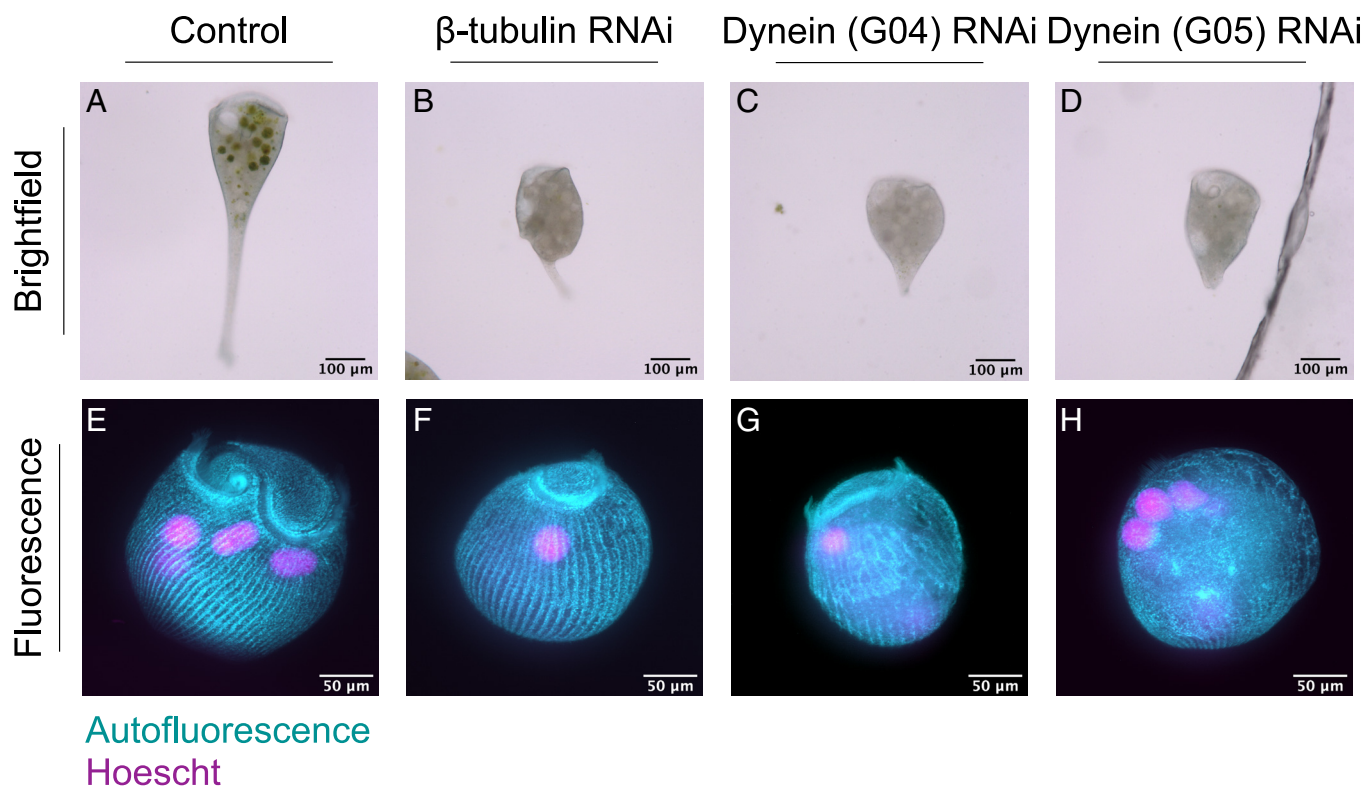


Fig. 3. β -tubulin and dynein knockdown result in loss of typical cell morphology. Brightfield images of (A) control, (B) β -tubulin, and (C and D) dynein intermediate chain knockdown cells. Maximum intensity projections for pigment autofluorescence (cyan) and Hoechst staining (magenta) of (E) control, (F) β -tubulin, and (G and H) dynein intermediate chain knockdown cells. Image brightness is adjusted for visualization purposes. G04 and G05 refer to two different RNAi constructs that target distinct dynein intermediate chains encoded by genes *SteCoe_24570* and *SteCoe_13560* (for G04) and *SteCoe_14054* (for G05).

on PC2 is eliminated upon β -tubulin knockdown (Fig. 4C). These differences at the whole-sample level suggest that the microtubule cytoskeleton is broadly important for proper RNA regionalization, such that anterior and posterior samples become more similar to each other in the absence of the microtubule tracks.

We then define skew in terms of the difference in abundance of transcripts between the two half-cells in each sample. To quantify polarization along the AP axis, we log (1 + x) transformed and adjusted the normalized TPM by adding the sample mean log-transformed TPM. Then, we calculated the difference over the sum of AP values for each gene within each sample to obtain a skew value (Methods). We removed genes with highly variable abundance between samples or abundances at the extreme end of high or low to ensure that these characteristics do not affect our overall calculation of skew. We expect that skews will be variable to some extent even in normally polarized cells, given that the precision of bisection is subject to human error. With this analysis, each gene is assigned an average skew, taken over all samples, with positive values reflecting anterior enrichment and negative values reflecting posterior enrichment. The distribution of skews in β -tubulin knockdown shows a taller, sharper peak, further suggesting that mRNA regionalization is disrupted (Fig. 4D).

If microtubule tracks worked to separate a subset of transcripts along the AP axis, one would expect a loss of transcript skew in the RNAi cells, which is certainly seen for many genes. For other genes, in which the skews remain correlated, it suggests that some degree of regionalization is still possible without intact microtubule tracks. The track model can thus account for the loss of correlation or retention of positive correlation. A more surprising outcome would be anticorrelation, such that a gene switches from anterior-enriched in the control to posterior-enriched in the knockdown, or vice versa.

To look for such genes, we performed feature selection by PCA to detect genes with anticorrelated skews (Methods). Feature selection uses PCA to identify the genes that contribute the most to the overall weightings of the PCA components, which are then for statistical significance. Although changes in skew that remain correlated may be significant (i.e., cases where an anterior-enriched gene still has an anterior skew, but perhaps more extreme), candidates with anticorrelated skews upon β -tubulin knockdown are the most interesting for further analysis, as these account for the largest differences in skew between control and β -tubulin cells. Following feature selection, we found 156 candidates out of 500 tested with a significant anticorrelated skew (SI Appendix, Table S7). *SteCoe_33099*, a putative membrane transporter and one of the most anterior-enriched mRNAs in bulk sequencing, is skewed more posteriorly upon β -tubulin knockdown (Fig. 4E). In the other direction, *SteCoe_7537* is a PP2C domain-containing protein that skews more anteriorly upon β -tubulin knockdown. We have previously shown that a PP2A subunit is important for the MB, an array of motile cilia at the anterior of the cell, morphogenesis (6), thus *SteCoe_7537* may be an interesting gene to examine for its role in posterior patterning.

In addition to *SteCoe_33099*, we wanted to know if other regionalized mRNAs, as determined by bulk RNA sequencing, showed an anticorrelated skew upon β -tubulin knockdown. Of the 156 differentially skewed genes upon β -tubulin knockdown, 50 were also differentially enriched in one half of the cell in bulk RNA sequencing (Fig. 4G). For most of these genes, including *SteCoe_33099* and *SteCoe_7537*, their skew in control half-cells from the RNAi experiments matches that of the bulk RNA sequencing, but they are oppositely skewed upon β -tubulin knockdown (Fig. 4H).

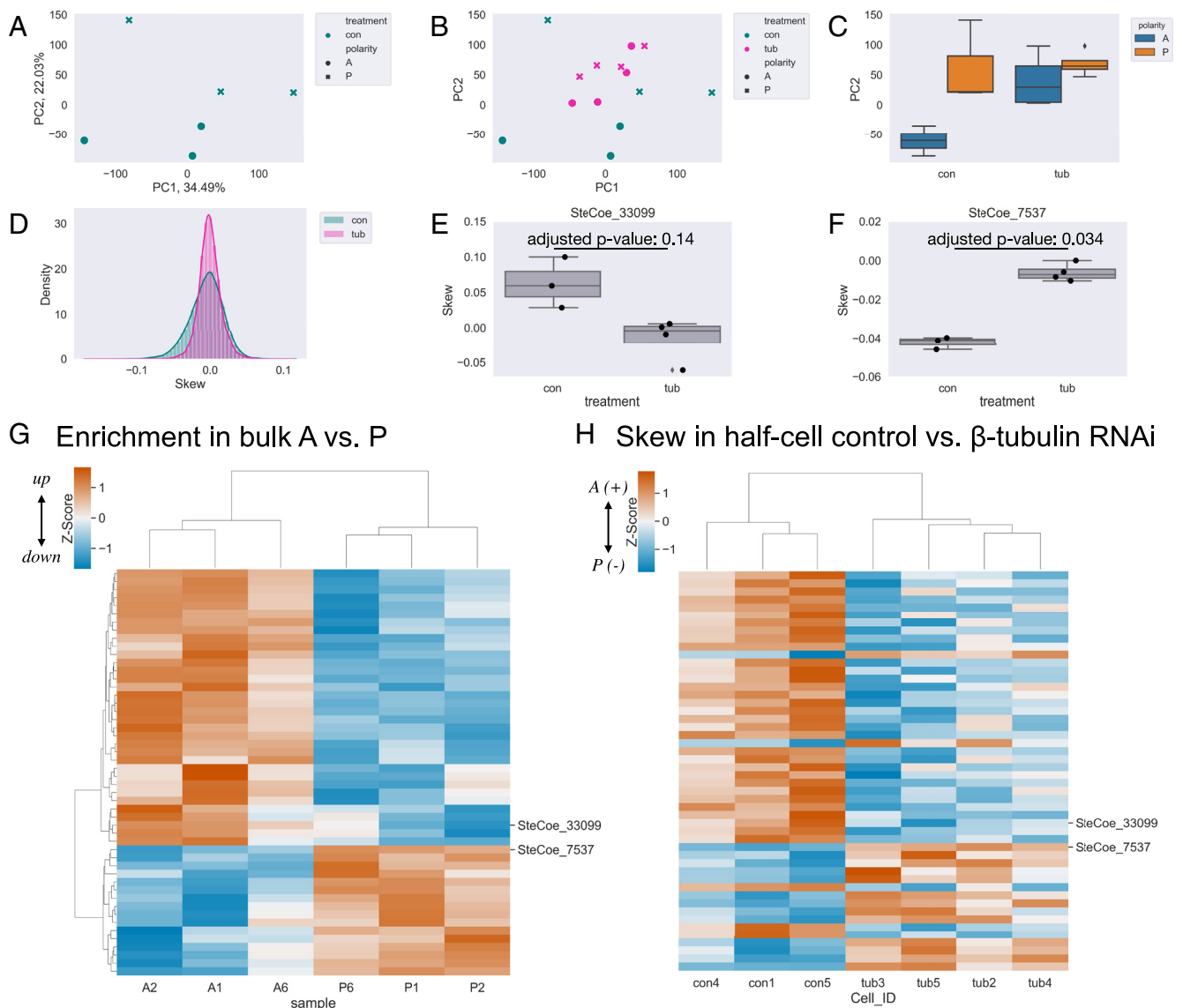


Fig. 4. β -tubulin knockdown results in loss of typical cell morphology and mRNA regionalization. (A) PCA on control half-cell RNA data showing a clear separation of anterior (circles) and posterior (x's) on PC2. (B) Projection of β -tubulin (pink) knockdown half-cells onto the control (green) PC space. (C) Boxplot of PC2 scores for control (Left) and β -tubulin (Right) data group by anterior (blue) or posterior (orange). (D) Distribution of individual gene skew values in control (green) and β -tubulin RNAi (pink). (E) Representative example, SteCoe_33099, with anterior skew that becomes more posterior upon β -tubulin knockdown. (F) Representative example, SteCoe_7537, with posterior skew that becomes more anterior upon β -tubulin knockdown. (G) Heatmap of relative abundance (z-scores) where rows correspond to individual genes and columns are both differentially enriched in bulk sequencing and differentially skewed upon β -tubulin knockdown. Red indicates higher enrichment, and blue indicates lower enrichment. (H) Heatmap of skew (z-scores) in control and β -tubulin RNAi, maintaining the same gene order as (G). Red indicates positive skew corresponding to relative enrichment in the anterior, and blue indicates negative skew corresponding to relative enrichment in the posterior. The four columns on the right cluster the four β -tubulin RNAi cell datasets, and it can be seen that for many genes the skew is now opposite in sign from that seen in controls.

Role of Cytoplasmic Dynein in mRNA Regionalization. Given that the surface of *Stentor* is covered with parallel, unidirectional longitudinal cortical microtubule bundles (5), among other cortical microtubule populations, an obvious hypothesis is that microtubule motors move transcripts to one end or the other of the cell. There is an extensive literature on cytoplasmic dynein functioning to transport mRNA in other organisms using adaptor proteins like BicD that bind specific messages and couple them to dynein-mediated trafficking (42). To understand the role of cytoplasmic dynein in cellular patterning, we conducted half-cell RNA-sequencing of paired anteriors and posteriors in control and dynein knockdown *Stentor* using two different dynein RNAi constructs. Construct G04 targets two genes, SteCoe_24570 and SteCoe_13560. G05 targets one gene, SteCoe_14054. In figures and in text, we refer to these as G04 and G05. Similar to β -tubulin

knockdown cells, we see that dynein knockdown cells display altered morphology and disorder in the cortical rows (Fig. 3 C, D, G, and H).

Here again, we performed a PCA on control-only samples (Fig. 5A), and in this case, PC1 corresponds to the AP axis. While there is partial overlap between anterior and posterior points along PC1, the samples show a consistent directional shift, where each anterior sample lies more toward the negative (left/anterior) side of PC1 relative to its matched posterior sample.

We then projected the dynein knockdown data onto the control PC space (Fig. 5B), and found that observable polarization on PC1 is eliminated upon dynein knockdown (Fig. 5C). These results suggest that microtubule motors are also important for proper RNA regionalization on a broad scale. The distribution of skews, specifically with the G04 but not the G05 construct, shows

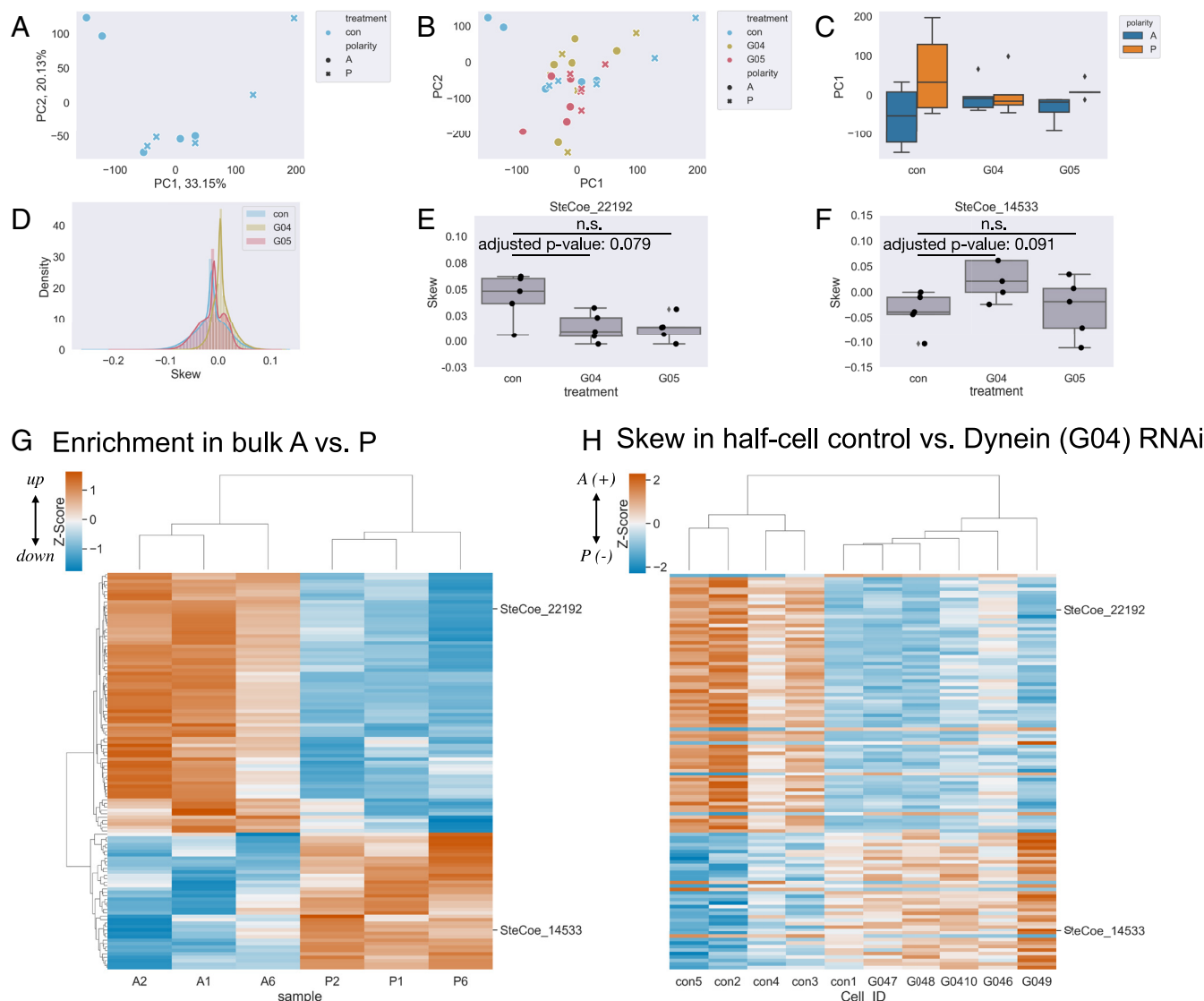


Fig. 5. Dynein knockdown results in loss of typical cell morphology and mRNA regionalization. (A) PCA on control half-cell RNA data showing a clear separation of anterior (circles) and posterior (x's) on PC1. (B) Projection of G04 (yellow) and G05 (red) knockdown half-cells onto the control (blue) PC space. (C) Boxplot of PC1 scores for control (Left), G04 (Middle), and G05 (Right) data grouped by anterior (blue) and posterior (orange). (D) Distribution of individual gene skew values in control (blue), G04 (yellow), and G05 (red). (E) Representative example, SteCoe_22192, with anterior skew that becomes more posterior upon dynein knockdown. (F) Representative example, SteCoe_14533, with posterior skew that becomes more anterior upon dynein knockdown. (G) Heatmap of relative abundance (z-scores) where rows correspond to individual genes and columns are both differentially enriched in bulk sequencing and differentially skewed upon dynein knockdown. Red indicates higher enrichment and blue indicates lower enrichment. (H) Heatmap of skew (z-scores) in control and dynein RNAi, maintaining the same gene order as (G). Red indicates positive skew corresponding to relative enrichment in the anterior, blue indicates negative skew corresponding to relative enrichment in the posterior. The six columns on the right cluster the 5 dynein RNAi datasets and 1 control cell. Note in this figure that the controls for the dynein RNAi experiment are different from the controls for the β -tubulin RNAi experiment in Fig. 3 because controls were performed separately for each RNAi experiment.

a taller, sharper peak, suggesting that perhaps G04 has a greater impact on mRNA regionalization (Fig. 5D). Interestingly, we detect 403 anticorrelated skews in G04 (SI Appendix, Table S9) following the same method as described above for β -tubulin and none in G05. Whether a lack of significance in G05 is biologically relevant or a result of technical effects is unclear.

SteCoe_22192, one of many genes that encode for actin, mRNA is enriched in the anterior and skewed anterior in control half-cells, but is skewed more posterior upon dynein intermediate chain (significant in G04, not significant in G05) knockdown (Fig. 5E). SteCoe_14533 is annotated as a protein kinase. Of the 403 differentially skewed genes upon dynein intermediate chain (G04) knockdown, 115 were also differentially enriched in one half of the cell in bulk RNA sequencing (Fig. 5G). For most of

these genes, including SteCoe_22192 and SteCoe_14533, their skew in control half-cells matches that of the bulk RNA sequencing, and they are oppositely skewed upon dynein knockdown (Fig. 5H).

Discussion

Here, we provide evidence that mRNAs are regionalized and the microtubule cytoskeleton is broadly important for proper mRNA regionalization in the giant ciliate, *S. coeruleus*. Among all of the regionalized transcripts, we found a large number encoding transcription factors, including 50 Myb DNA-binding domain proteins (SI Appendix, Table S3). Myb DNA-binding domains are also overrepresented in the posterior (SI Appendix, Table S5).

In addition to its role in patterning plant cells and as a proto-oncogene in humans, MYB is also a key regulator of ciliogenesis and centriole amplification during the formation of multiciliated epithelia (43, 44), both of which take place during morphogenesis and regeneration in *Stentor*. Upon knocking down two Myb genes, *SteCoe_15300* and *SteCoe_33495*, we found that a considerable number of cells are unable to regenerate their posterior. It may seem surprising to find such messages that encode transcription factors destined to enter the nucleus, regionalized within a cell that has a single nucleus. However, we note that although *S. coeruleus* has a single macronucleus, this macronucleus is highly polyploid and spatially extended, running the length of the whole cell. Moreover, the macronucleus takes the shape of a set of nodes joined by small junctions, like beads on a string, such that different nodes might be able to sustain different programs of gene expression. We speculate that transcription factors encoded by regionalized messages might function to locally drive patterns of gene expression in different parts of the macronucleus.

Knockdown of β -tubulin, a major component of the microtubule cytoskeleton, disrupts normal cell morphology and causes cells to lose their orderly cortical rows [Fig. 3, (12)]. The mRNA encoded by the specific tubulin gene that we targeted in this analysis (*SteCoe_5719*) is significantly enriched in the anterior (*SI Appendix, Table S3*). Knockdown of cytoplasmic dynein intermediate chains, a major component of molecular motors that carry cargo along the microtubule cytoskeleton, also disrupts normal cell morphology and causes cells to lose their orderly cortical rows (Fig. 3). Half-cell RNA-sequencing in paired anteriors and posteriors of the cell also reveals global disruptions in regionalization by PCA upon knockdown of β -tubulin or dynein intermediate chains (Figs. 4 A–C and 5 A–C). Overall, we find many candidates with significant and anticorrelated skews (β -tubulin—*SI Appendix, Table S7*, dynein (G04)—*SI Appendix, Table S9*). However, despite an apparent loss of regionalization in one of our dynein intermediate chain knockdowns globally (G05—Fig. 5C), we find no statistically significant different transcript skews (*SI Appendix, Table S10*). The two cytoplasmic dynein transcripts targeted by the G04 construct (*SteCoe_24570* and *SteCoe_13560*) were themselves anteriorly enriched, while the transcript targeted by G05 (*SteCoe_14054*) was found to be

dispersed in both halves of the cell (*SI Appendix, Table S3*). It is perhaps informative that RNAi targeting the first two genes resulted in inversion of skew, while the third did not. Given the potential role of dynein motors in organizing microtubules, such as is seen in the mitotic spindle, it could be that some of the defects we observe extend beyond simple cargo transport.

We speculate that transcripts with significant and anticorrelated skews upon β -tubulin or dynein knockdown (G04) may be important for cellular patterning, given the morphological defects and change in transcript skew upon knockdown, and the connection between patterned mRNAs and cell patterning in other contexts. We do note, however, that for a small number of genes, the skew in the control half-cells was opposite to that seen in bulk RNA sequencing, but then became the same as in bulk RNA sequencing in the RNAi cells.

This work shows that mRNA regionalization is disrupted upon the loss of β -tubulin or dynein intermediate chains, but raises several additional questions: 1) Are the messages we found directly involved in cellular patterning or the regeneration of new patterned structures? 2) Is mRNA regionalization by the microtubule cytoskeleton essential for regeneration, given its role in patterning the cell? 3) Does protein enrichment in one half over the other arise from differential transcript regionalization or possibly translation?

RNA regionalization could be a key to understanding pattern formation and regeneration in *Stentor*. Based on our RNAi phenotypes, which show that depletion of posterior-localized transcripts causes anterior half-cells to become impaired in regenerating posterior structures, we favor a model in which mRNA transcripts are localized to the point at which their translation products will be required. Our results suggest a potential model for how this regionalization is achieved. Given that RNAi knockdown of posterior-regionalized transcripts corresponding to transcription factors inhibits proper posterior regeneration and that disruption of the cortical microtubule bundles led to a loss or reversal of polarized regionalization for many messages, we propose that messages may both be transcribed near where they are needed in the cell and associate with microtubule motors via RNA-binding adaptors, such as BicD, and then move to the ends of the bundles (Fig. 6). Future work will be needed to test this model by interfering with kinesin as well as potential adaptor proteins.

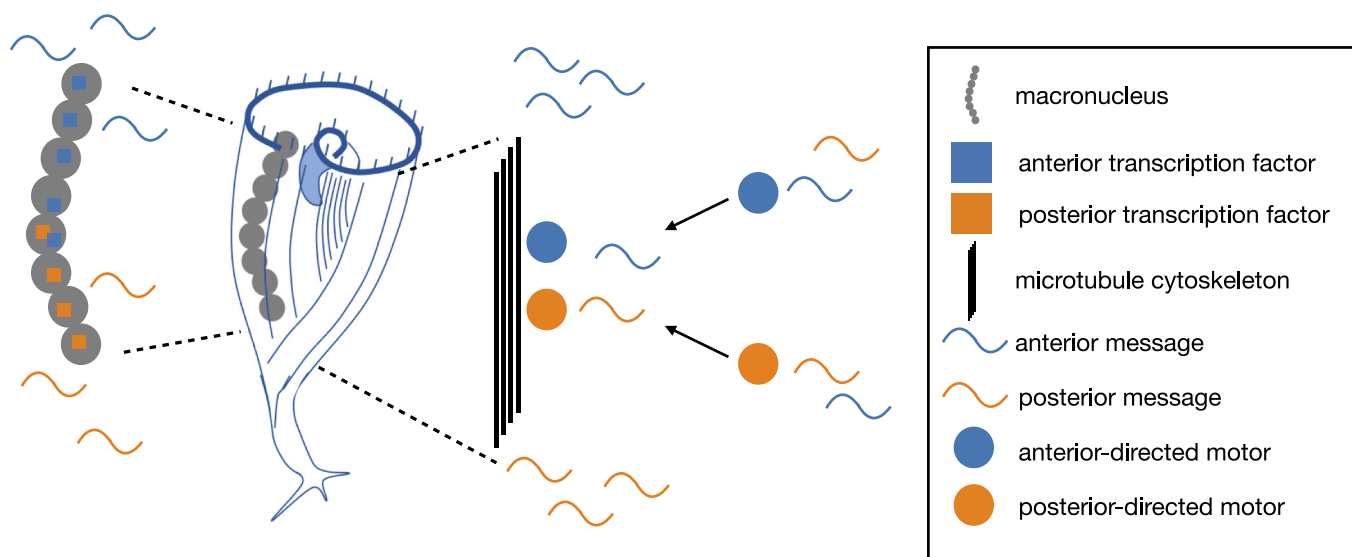


Fig. 6. Model for RNA regionalization in *Stentor* via regionalized transcription factors and/or motor-directed transport.

Data, Materials, and Software Availability. Data and code have been deposited in Dryad (45) and Zenodo (36). Other data are included in the manuscript and/or supporting information.

ACKNOWLEDGMENTS. We would like to acknowledge Mark Slabodnick, Dept. Biology Knox College for his pioneering work on *Stentor* RNAi by both feeding and microinjection. We thank Dyche Mullins, Tanner Fadero, and Sam Lord for microscopy assistance and use of the iSIM. We would also like to thank members of the Marshall Lab, Mark Slabodnick, and Xin Xiang for helpful discussion and

feedback. Sequencing was performed at the University of California San Francisco (UCSF) Center for Advanced Technology (CAT), supported by UCSF Program in Breakthrough Biomedical Research (PBBR), UCSF Research Resilience Program (RRP) Institutional Matching Instrumentation Award, and NIH 1S100D028511-01 Grants. A.R.A. was supported by NIH 5K12GM081266, NIH K99GM154059, and UCSF Research Resilience Initiative. C.Y. was supported by an NSF-GRFP (Graduate Research Fellowship Program) fellowship. This work was supported by NIH R35 GM130327, on which W.F.M. is the principal investigator.

1. M. Kirschner, J. Gerhart, T. Mitchison, Molecular, "vitalism." *Cell* **100**, 79–88 (2000).
2. J. M. Shulman, D. St. Johnston, Pattern formation in single cells. *Trends Cell Biol.* **9**, M60–M64 (1999).
3. W. F. Marshall, Regeneration in *Stentor coeruleus*. *Front. Cell Dev. Biol.* **9**, 753625 (2021).
4. V. Tartar, *The Biology of Stentor* (Pergamon Press, New York, 1961).
5. B. Huang, D. R. Pitelka, The contractile process in the ciliate, *Stentor coeruleus*. I. The role of microtubules and filaments. *J. Cell Biol.* **57**, 704–728 (1973).
6. A. Lin *et al.*, Determining protein polarization proteome-wide using physical dissection of individual *Stentor coeruleus* cells. *Curr. Biol.* **32**, 2300–2308.e4 (2022).
7. F. R. Lillie, On the smallest parts of stentor capable of regeneration; a contribution on the limits of divisibility of living matter. *J. Morphol.* **12**, 239–249 (1896).
8. T. H. Morgan, Regeneration of proportionate structures in *Stentor*. *Biol. Bull.* **2**, 311–328 (1901).
9. K. S. Zhang, L. R. Blauch, W. Huang, W. F. Marshall, S. K. Y. Tang, Microfluidic guillotine reveals multiple timescales and mechanical modes of wound response in *Stentor coeruleus*. *BMC Biol.* **19**, 63 (2021).
10. N. De Terra, Evidence for cortical control of macronuclear behavior in *Stentor*. *J. Cell Physiol.* **78**, 377–386 (1971).
11. M. M. Slabodnick *et al.*, The macronuclear genome of *Stentor coeruleus* reveals tiny introns in a giant cell. *Curr. Biol.* **27**, 569–575 (2017).
12. M. M. Slabodnick *et al.*, The kinase regulator mob1 acts as a patterning protein for *Stentor* morphogenesis. *PLoS Biol.* **12**, e1001861 (2014).
13. P. Sood *et al.*, Modular, cascade-like transcriptional program of regeneration in *Stentor*. *Elife* **11**, e80778 (2022), 10.7554/eLife.80778.
14. W. Wei, C. Jiang, W. Yang, W. Miao, J. Xiong, Proteomic identification and expression of oral apparatus constituents in cell regeneration of giant ciliate *Stentor coeruleus* (strain WHL). *Gene* **743**, 144624 (2020).
15. H. Onsrbring, M. Jamy, T. J. G. Ettema, RNA sequencing of *Stentor* cell fragments reveals transcriptional changes during cellular regeneration. *Curr. Biol.* **28**, 1281–1288.e3 (2018).
16. V. Tartar, Grafting experiments concerning primordium formation in *Stentor coeruleus*. *J. Exp. Zool.* **131**, 75–121 (1956).
17. J. Frankel, *Pattern Formation* (Oxford University Press, New York, NY, 1990).
18. D. Wloga, J. Frankel, From molecules to morphology: Cellular organization of *Tetrahymena thermophila*. *Methods Cell Biol.* **109**, 83–140 (2012).
19. E. Cole, J. Gaertig, Anterior-posterior pattern formation in ciliates. *J. Eukaryot. Microbiol.* **69**, e12890 (2022).
20. P. A. Combs, M. B. Eisen, Genome-wide measurement of spatial expression in patterning mutants of *Drosophila melanogaster*. *F1000Res.* **6**, 41 (2017).
21. J. E. Haines, M. B. Eisen, Patterns of chromatin accessibility along the anterior-posterior axis in the early *Drosophila* embryo. *PLoS Genet.* **14**, e1007367 (2018).
22. A. Raj, P. van den Bogaard, S. A. Rifkin, A. van Oudenaarden, S. Tyagi, Imaging individual mRNA molecules using multiple singly labeled probes. *Nat. Methods* **5**, 877–879 (2008).
23. C. Xia, J. Fan, G. Emanuel, J. Hao, X. Zhuang, Spatial transcriptome profiling by MERFISH reveals subcellular RNA compartmentalization and cell cycle-dependent gene expression. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 19490–19499 (2019).
24. P. A. Takizawa, A. Sil, J. R. Swedlow, I. Herskowitz, R. D. Vale, Actin-dependent localization of an RNA encoding a cell-fate determinant in yeast. *Nature* **389**, 90–93 (1997).
25. P. Litman, J. Barg, I. Ginzburg, Microtubules are involved in the localization of tau mRNA in primary neuronal cell cultures. *Neuron* **13**, 1463–1474 (1994).
26. E. H. Harris, *The Chlamydomonas Sourcebook: A Comprehensive Guide to Biology and Laboratory Use* (Academic Press, San Diego, 1989).
27. A. Kuecks, P. Arra, S. Hoffmann-Weitsmann, M. F. Craig, M. M. Slabodnick, Microinjected dsRNA triggers a robust RNAi response in *Stentor coeruleus*. bioRxiv [Preprint] (2025). <https://doi.org/10.1101/2025.08.28.672911> (Accessed 28 August 2025).
28. U. Diaz, W. Marshall, B. Riggs, *Drosophila* embryo preparation and microinjection for live cell microscopy performed using an automated high content analyzer. *J. Vis. Exp.* **167**, e61589 (2021), 10.3791/61589.
29. J. D. Thompson, D. G. Higgins, T. J. Gibson, CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**, 4673–4680 (1994), 10.1093/nar/22.22.4673.
30. S. Guindon *et al.*, New algorithms and methods to estimate maximum-likelihood phylogenies: Assessing the performance of PhyML 3.0. *Syst. Biol.* **59**, 307–321 (2010), 10.1093/sysbio/syq010.
31. A. Lin, T. Makushok, U. Diaz, W. F. Marshall, Methods for the study of regeneration in *Stentor*. *J. Vis. Exp.* **136**, e57759 (2018), 10.3791/57759.
32. N. L. Bray, H. Pimentel, P. Melsted, L. Pachter, Near-optimal probabilistic RNA-seq quantification. *Nat. Biotechnol.* **34**, 525–527 (2016).
33. H. Pimentel, N. L. Bray, S. Puente, P. Melsted, L. Pachter, Differential analysis of RNA-seq incorporating quantification uncertainty. *Nat. Methods* **14**, 687–690 (2017).
34. W. E. Johnson, C. Li, A. Rabinovic, Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics* **8**, 118–127 (2007).
35. B. Muzellec, M. Teleńczuk, V. Cabeli, M. Andreux, PyDESeq2: A python package for bulk RNA-seq differential expression analysis. *Bioinformatics* **39**, btad547 (2023).
36. A. Albright, D. Angeles-Albores, Data from "Genome-wide analysis of mRNA regionalization in a giant single cell." Zenodo. <https://doi.org/10.5281/zenodo.20452282>. Deposited 29 May 2026.
37. M. Dodt, J. T. Roehr, R. Ahmed, C. Dieterich, FLEXBAR-flexible barcode and adapter processing for next-generation sequencing platforms. *Biology (Basel)* **1**, 895–905 (2012).
38. K. Baumann *et al.*, Control of cell and petal morphogenesis by R2R3 MYB transcription factors. *Development* **134**, 1691–1701 (2007).
39. R. G. Ramsay, T. J. Gonda, MYB function in normal and cancer cells. *Nat. Rev. Cancer* **8**, 523–534 (2008).
40. Y. Weinstein, J. N. Ihle, S. Lavu, E. P. Reddy, Truncation of the c-myc gene by a retroviral integration in an interleukin 3-dependent myeloid leukemia cell line. *Proc. Natl. Acad. Sci. U.S.A.* **83**, 5010–5014 (1986).
41. P. D. Thomas *et al.*, PANTHER: Making genome-scale phylogenetics accessible to all. *Protein Sci.* **31**, 8–22 (2022).
42. S. L. Bullock, D. Ish-Horowitz, Conserved signals and machinery for RNA transport in *Drosophila* oogenesis and embryogenesis. *Nature* **414**, 611–616 (2001).
43. S. P. Choksi, G. Lauter, P. Swoboda, S. Roy, Switching on cilia: Transcriptional networks regulating ciliogenesis. *Development* **141**, 1427–1441 (2014).
44. F. E. Tan *et al.*, Myb promotes centriole amplification and later steps of the multiciliogenesis program. *Development* **140**, 4277–4286 (2013).
45. A. Albright, W. Marshall, Data from "Parent dataset: Single and half-cell RNA-sequencing in *Stentor coeruleus* control, beta-tubulin, and dynein knockdown cells" [with Child Datasets 1–4]. Dryad. <https://datadryad.org/dataset/doi:10.6078/D1WT6W>. Deposited 1 December 2023.