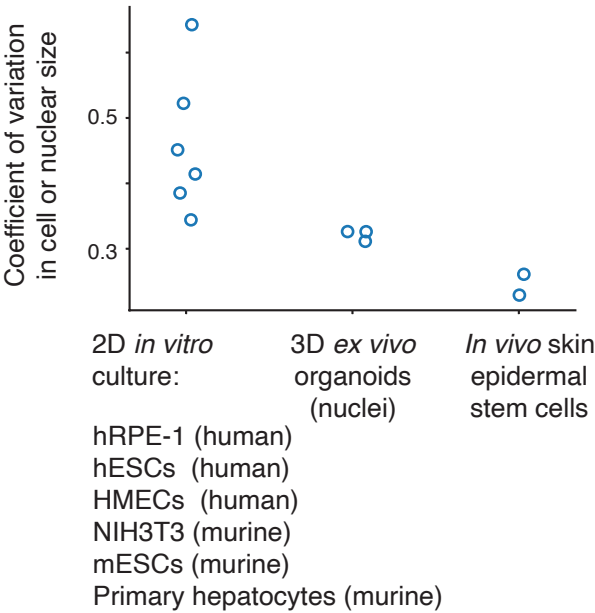


Supplemental Figure S1

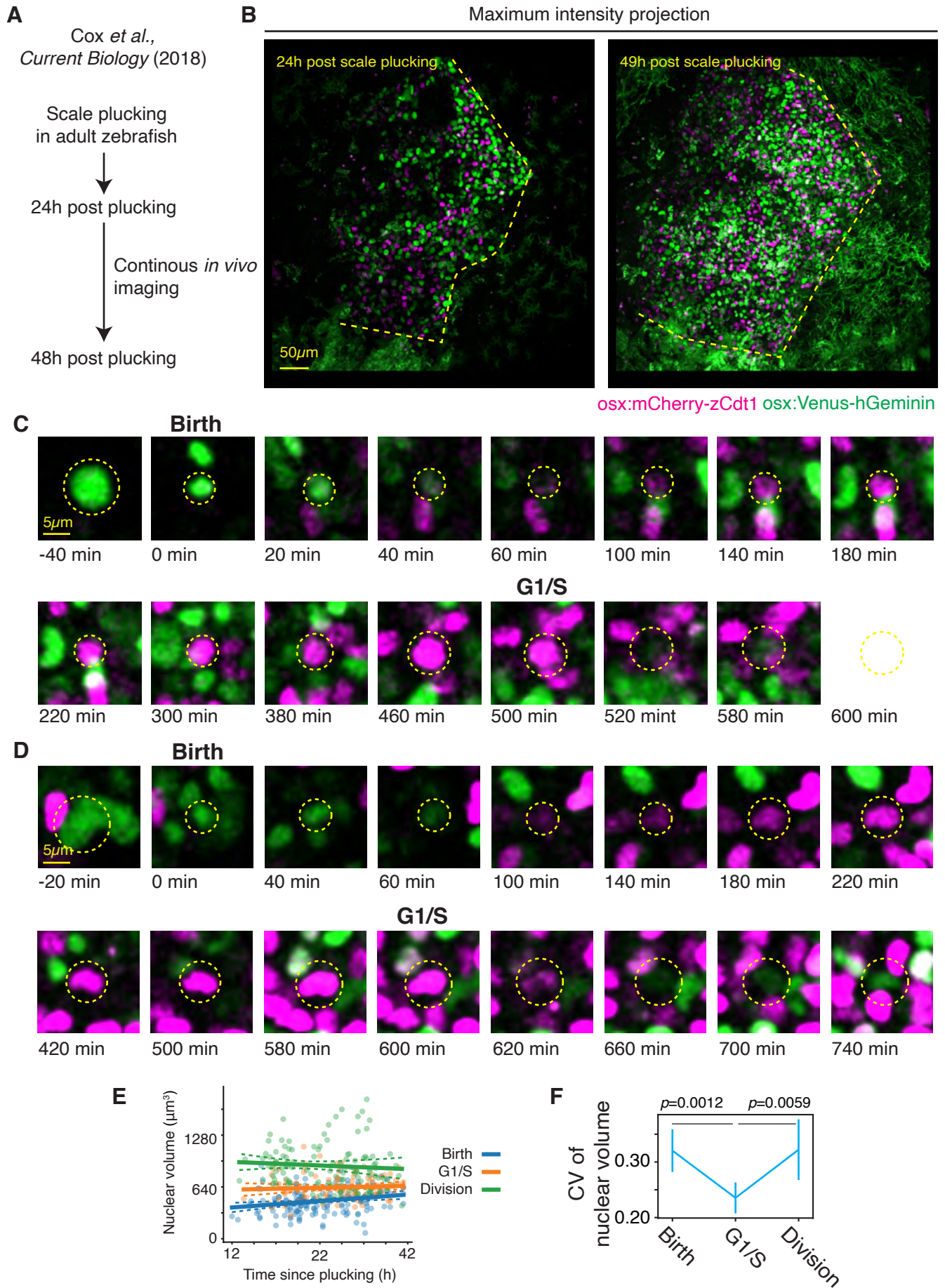


Supplementary Figure Legends

Supplementary Figure S1. Cell cycle variation in 2D cell culture, organoid cultures, and *in vivo*. The coefficient of variation (CV) of cell volume are shown for 6 mammalian cell lines and 2 different regions of *in vivo* mouse skin epidermal stem cells. The CV in nuclear volume is shown for 3 intestinal organoids during spheroid stage.

Source data are provided as a Source Data file.

Supplemental Figure S2



Supplementary Figure S2. Osteoblast growth during zebrafish scale regeneration.

A. Dataset acquired by Cox *et al.*, *Current Biology* (2018). Scales were plucked from an adult zebrafish, and confocal microscopy was used to monitor regenerating osteoblasts dynamics every 20 minutes from 24h to 48h post-injury.

B. Maximum intensity projections are shown of the regenerating scale in a zebrafish expressing the nuclear FUCCI cell cycle reporters *osx:mCherry-zCdt1* and *osx:Venus-hGeminin*, at 24h and 48h post-injury. Dotted yellow outlines the extent of the regenerating scale.

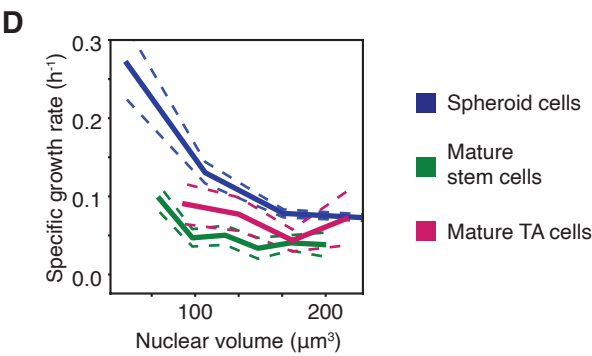
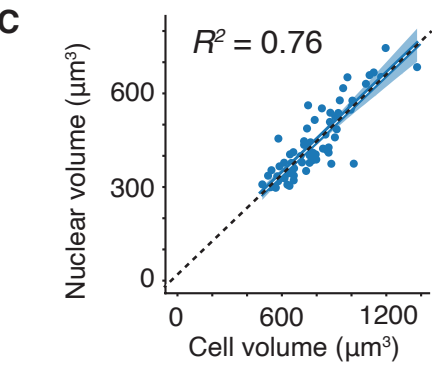
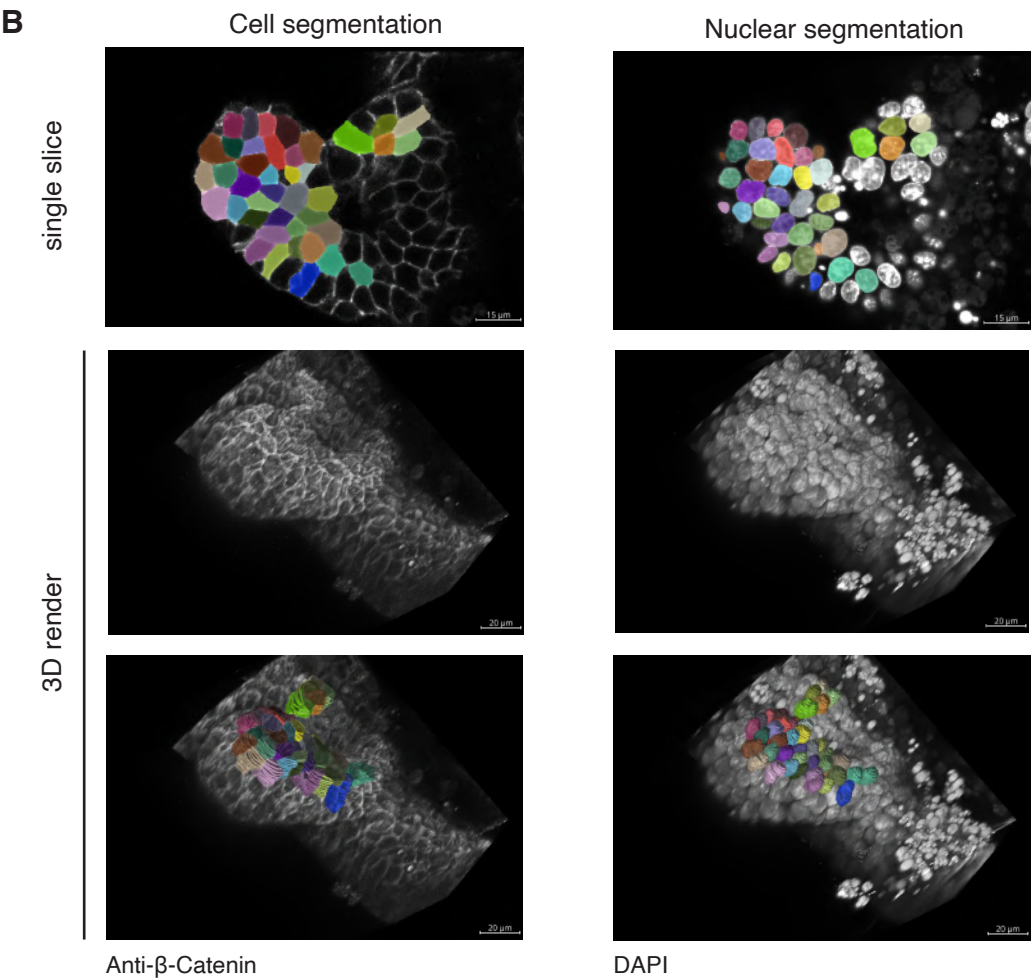
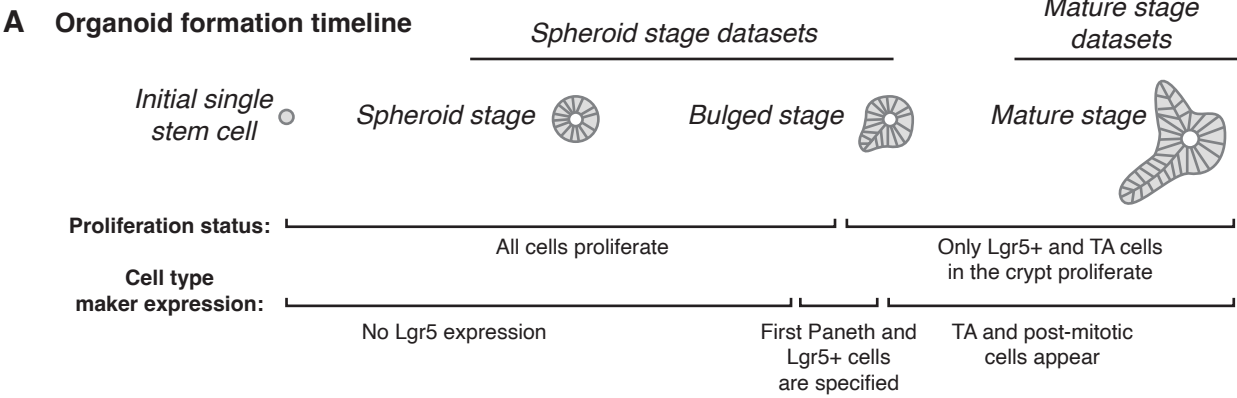
C-D. Two examples of a single osteoblast tracked from cell birth to the G1/S transition are shown. Note that *osx:Venus-hGeminin* expression remains for up to 1 hour after mitosis.

E. The nuclear volume is quantified for regenerating osteoblasts at birth, the G1/S transition, or division, as a function of time since injury. Solid line is linear regression. Dotted lines are 95% confidence intervals.

F. The CV of nuclear volume is shown for osteoblasts at birth, the G1/S transition, or division. Errorbar is mean $\pm$ std estimated by bootstrap. P-value of CV difference is estimated by bootstrap test.

Source data are provided as a Source Data file.

Supplemental Figure S3



Supplementary Figure S3. Intestinal organoid development.

A. A schematic of the stages of intestinal organoid development. When grown from single cells, organoids form initially as an undifferentiated spheroid where all cells actively proliferate. As the first Paneth and Lgr5⁺ cells are specified, the organoid transitions into a bulged state where cells arrest proliferation and active cycling becomes restricted to Lgr5⁺ stem cells and Lgr5⁻ TA cells. As the bulge domain buds out from the organoid body, it forms a mature crypt domain.

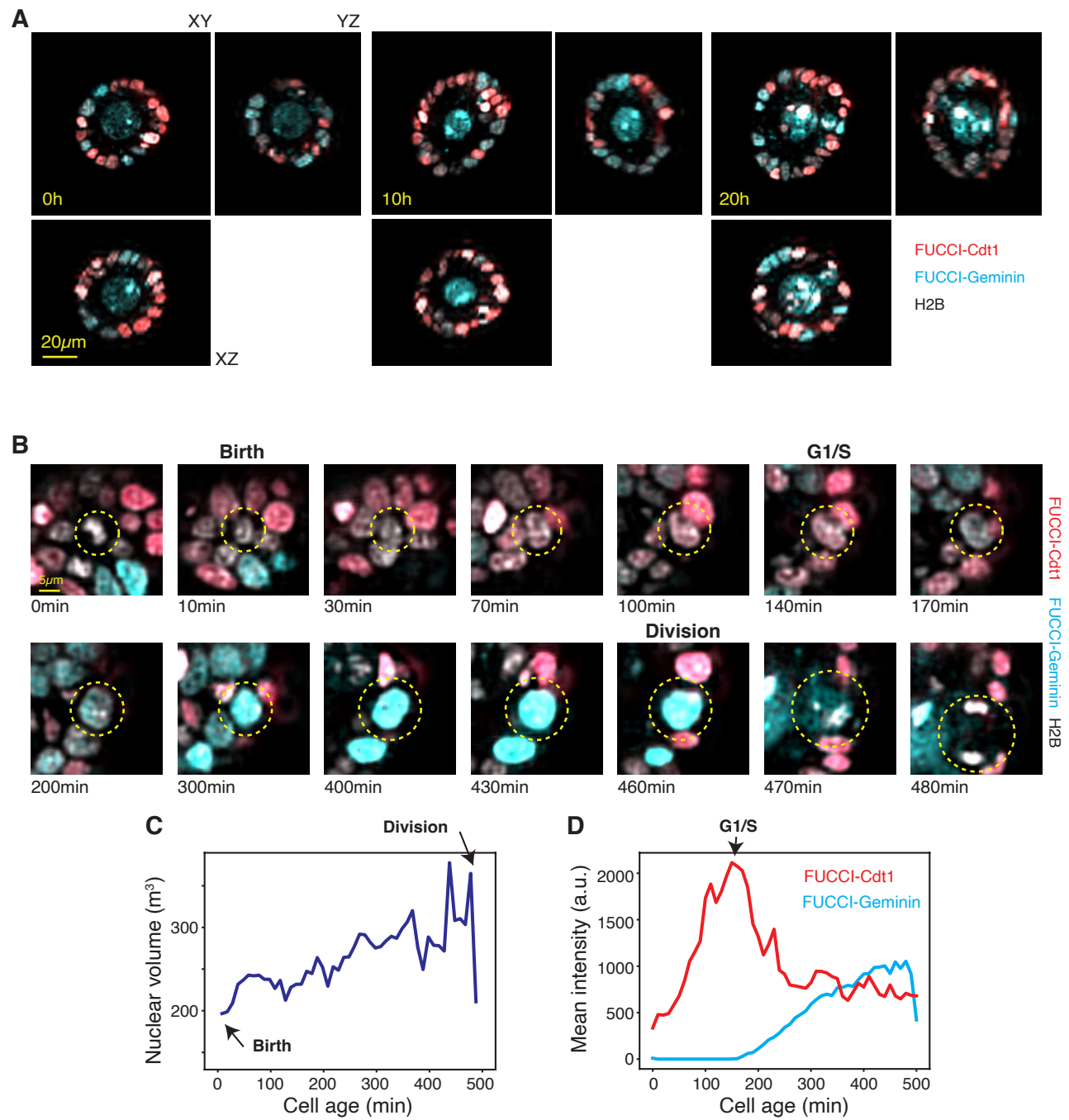
B. Example images of nuclear and cell volume segmentation of fixed intestinal organoids based on DAPI and β -Catenin signals, respectively. Top row is a slice. Middle row is a 3D render, while bottom row is a 3D render overlaid with the 3D cell or nuclear segmentation.

C. Quantification of the relationship between cell volume and nuclear volume in fixed intestinal organoids. The coefficient of determination is reported. N = 3 organoids, 79 cells. Dotted line is linear regression, and shaded area is 95% confidence intervals.

D. The size-binned specific nuclear growth rate is shown for spheroid, stem, and TA cells. N = 3 spheroid stage and N = 3 mature organoids, and >250 time points for each cell type. Solid line denotes the binned median. Dotted lines denote SEM of each bin.

Source data are provided as a Source Data file.

Supplemental Figure S4



Supplementary Figure S4. Analysis of spheroid stage intestinal organoids.

A. An example spheroid stage organoid expressing *H2B-miRFP670*, *FUCCI-Cdt1-mCherry*, and *FUCCI-Geminin-mVenus* is shown in 3 orthogonal views.

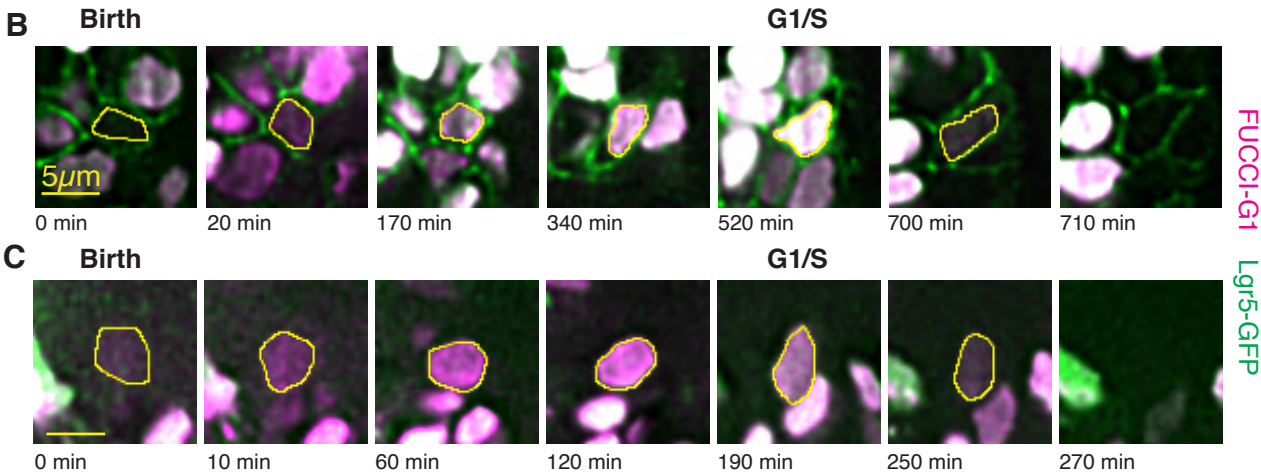
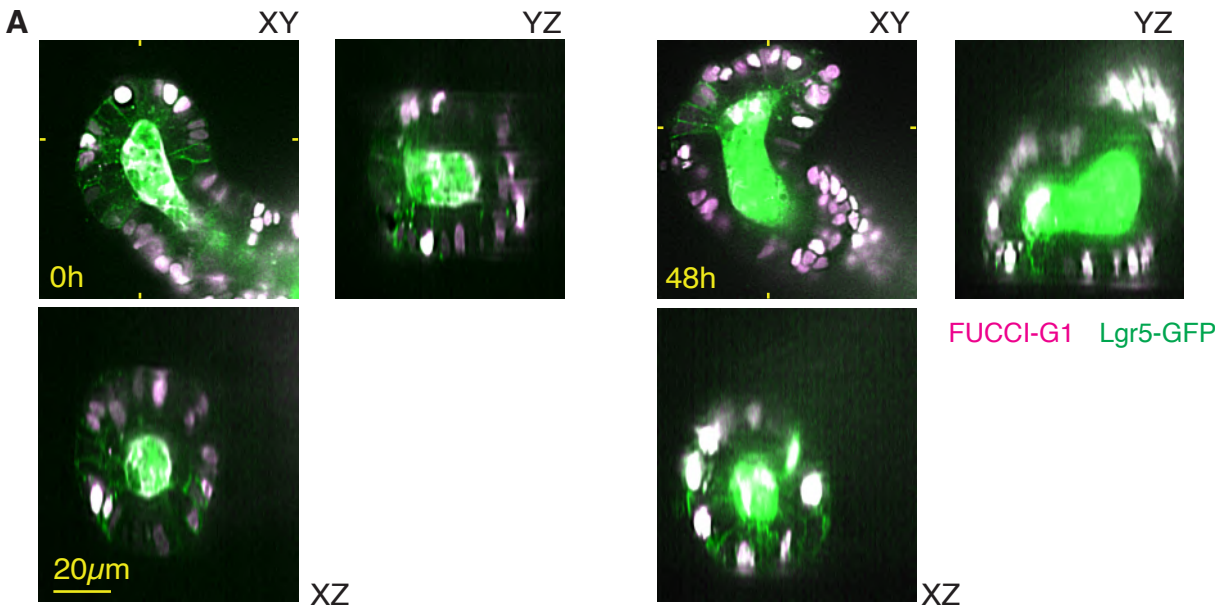
B. An example spheroid stage cell is tracked from cell growth to cell division. Dotted yellow outlines the tracked cell.

C. The time-series of nuclear volume is shown for the example cell in (B).

D. The time-series of *FUCCI-Cdt1* and *FUCCI-Geminin* are shown for the example cell in (B).

Source data are provided as a Source Data file.

Supplemental Figure S5



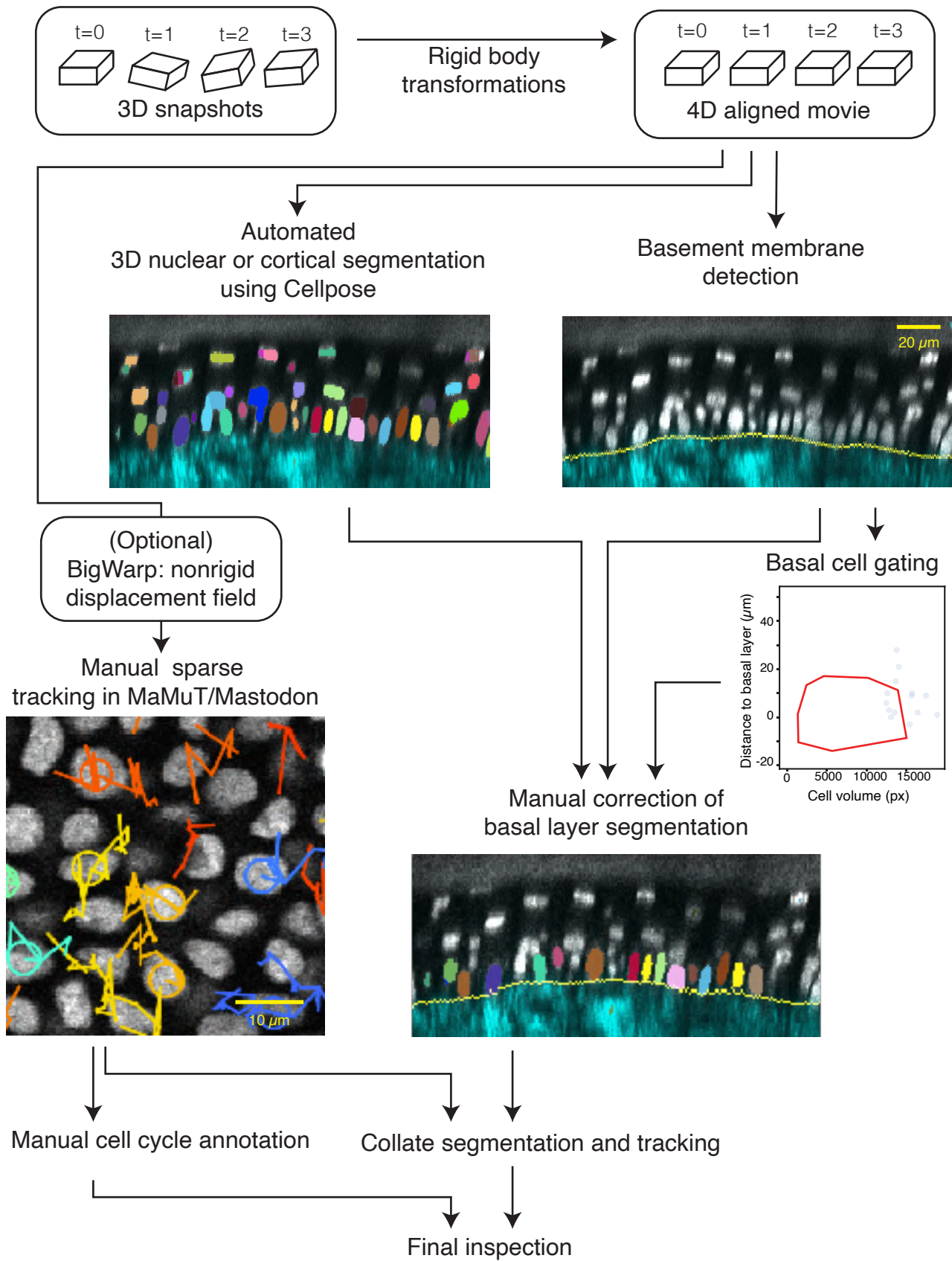
Supplementary Figure S5. Analysis of mature stage intestinal organoids.

A. An example mature stage organoid movie expressing *Lgr5-GFP* and *FUCCI-Cdt1-mKO2* is shown in 3 orthogonal views.

B. Example montage of an *Lgr5*⁺ intestinal stem cell tracked from birth to G1 exit. Scale bar is 5µm. G1/S transition is defined as the point at which the *FUCCI-G1* signal decreases.

C. Example montage of an *Lgr5*⁻ TA cell tracked from birth to G1 exit.

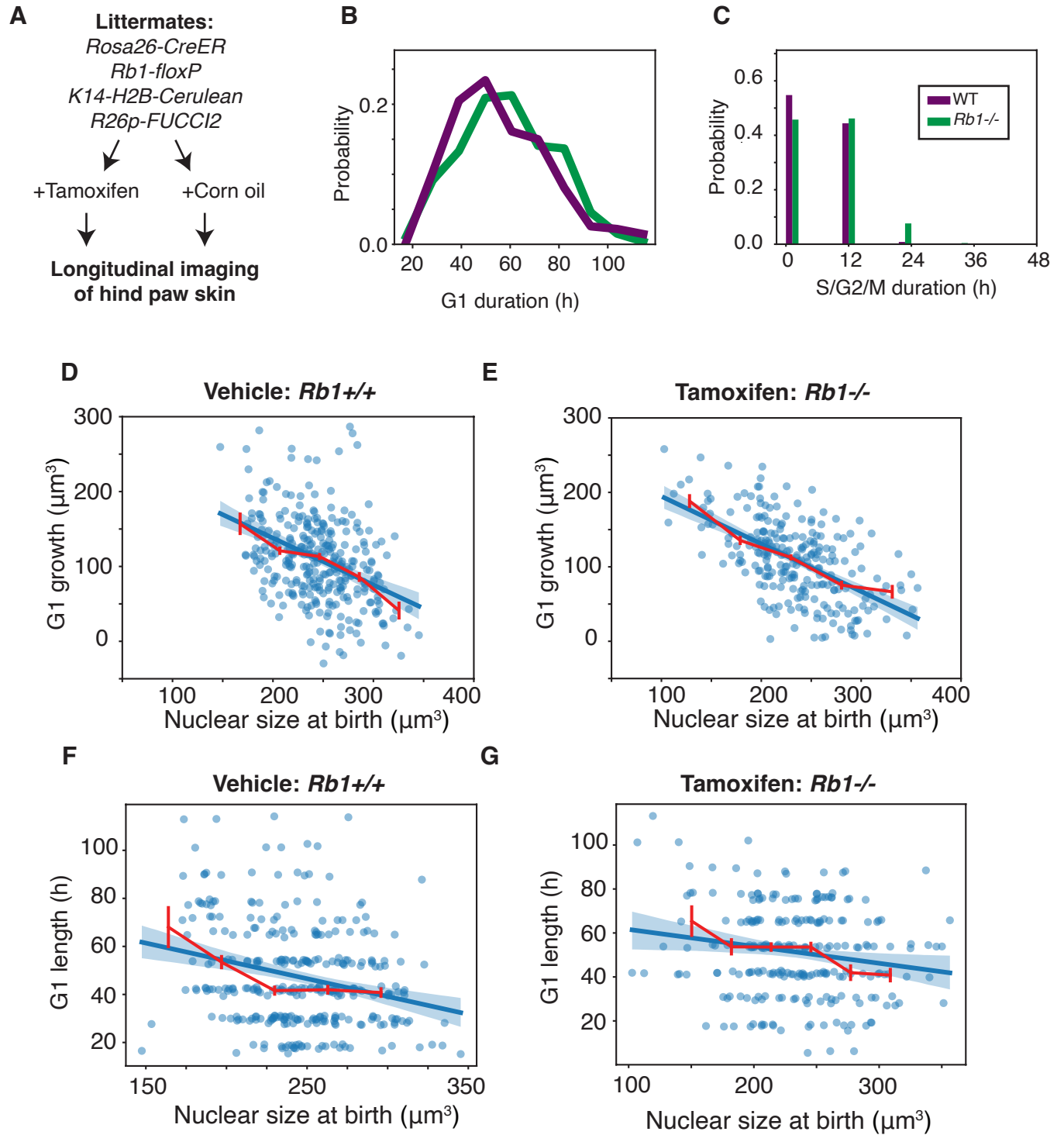
Supplemental Figure S6



Supplementary Figure S6. Image analysis pipeline for tracking single cells from longitudinal intravital imaging experiments. A flow chart showing the major steps of the image analysis process used to track single epidermal basal cells through complete cell divisions. Details can be found in Methods.

Source data are provided as a Source Data file.

Supplemental Figure S7



Supplementary Figure S7. *Rb1*^{-/-} single mutation does not affect cell size control.

A. Schematic of the experimental setup. *Rb1-flox*; *Rosa26-CreERT2*; *K14-H2B-Cerulean*; *R26p-FUCCI-G1-mCherry* litter mates were either injected with tamoxifen or corn oil vehicle control, and their paw skin epidermis were longitudinally imaged.

B-C. The durations of G1 phase (B) or S/G2/M phases combined (C) are shown for WT and *Rb1*^{-/-} cells. Note that since the sampling rate is every 12h, many cells with S/G2/M phases that were shorter than 12h were annotated to have S/G2/M durations of 0h.

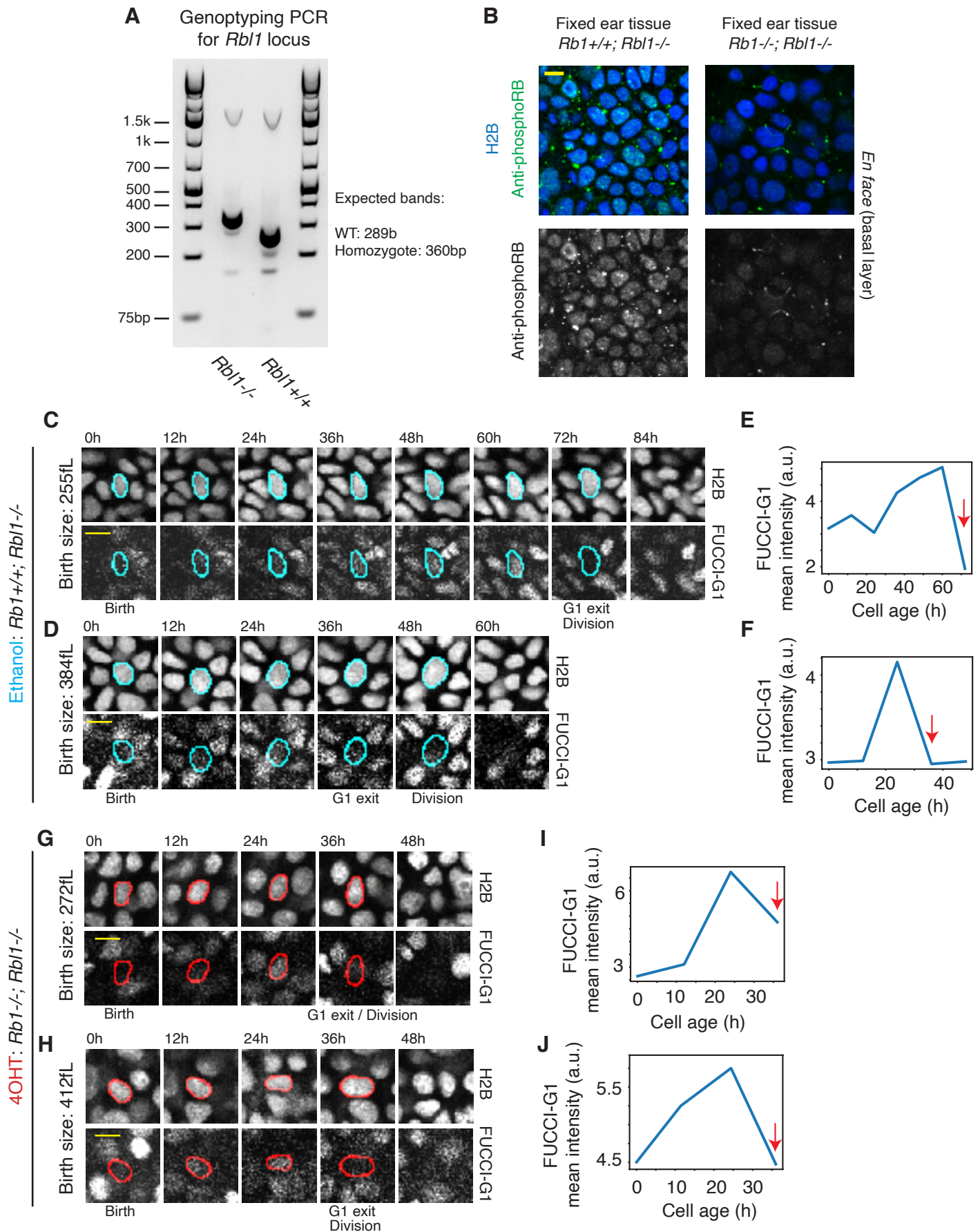
D-E. The relationship between nuclear size at birth and amount of nuclear growth during G1 is shown for WT (D) and *Rb1*^{-/-} (E) cells. WT: slope = -0.63; N = 2 mice, 317 cells; *Rb1*^{-/-}: slope = -0.63, N = 2 mice, 244 cells.

F-G. The relationship between nuclear size at birth and the duration of G1 phase is shown for WT (F) and *Rb1*^{-/-} (G) cells. WT: Pearson's R = -0.25; *Rb1*^{-/-}: Pearson's R = -0.18. For visual clarity, a small random jitter was added to the G1 length.

Solid line is the linear regression, and the shaded area is the 95% confidence interval. Red line is the binned data showing the median and SEM.

Source data are provided as a Source Data file.

Supplemental Figure S8



Supplementary Figure S8. Characterization of *Rb1*^{-/-}; *Rbl1*^{-/-} DKO tissues.

A. Genotyping PCR results are shown for the *Rbl1*^{-/-} mouse, compared to a B6 wildtype control.

B. Immunofluorescence staining using antibody recognizing phosphorylated Rb1 in fixed *Rb1*^{+/+}; *Rbl1*^{-/-} (SKO) and *Rb1*^{-/-}; *Rbl1*^{-/-} (DKO) tissues. An *en face* view of the basal layer is shown. The topical application of 4OHT to the ear skin led to the complete loss of phosphorylated Rb1 signal in DKO tissues. Scale bar is 10µm. n = 1 mouse.

C-D. Example cells are shown for in smaller-born (C) and larger-born (D) cell in SKO tissue. Blue outline shows the segmentation of the nuclear volume.

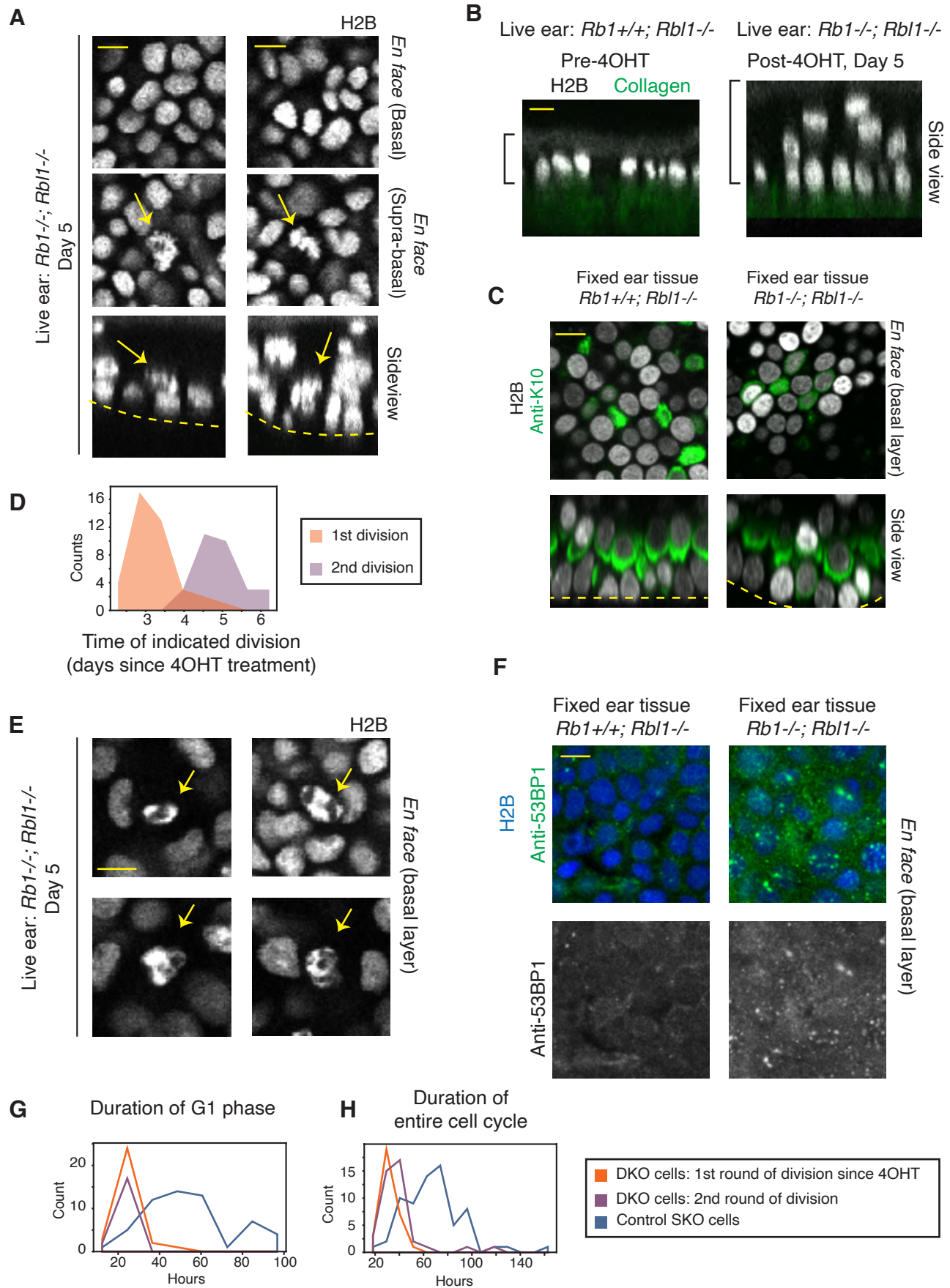
E-F. The *FUCCI-G1* mean fluorescence intensity is quantified for the example SKO cells shown in (C) and (D), respectively. Red arrow denotes the G1 exit frame.

G-H. Example cells are shown for in smaller-born (G) and larger-born (H) cell in DKO tissue. Red outline shows the segmentation of the nuclear volume.

I-J. The *FUCCI-G1* mean fluorescence intensity is quantified for the example DKO cells shown in (G) and (H), respectively. Red arrow demarcates the G1 exit frame.

Unless noted otherwise, scale bar is 5µm. Source data are provided as a Source Data file.

Supplemental Figure S9



Supplementary Figure S9. Cell cycle defects in DKO tissues.

A. Representative images of two regions in the DKO tissue where supra-basal mitotic figures could be seen. The *H2B-Cerulean* reporter was imaged in a live mouse 5 days after 4OHT treatment. *En face* views of the basal and supra-basal layers and side views are shown. Arrows indicate the location of ectopic mitotic figures. Dotted line indicates the location of the basement membrane.

B. Representative side views of an ear epidermis showing the thickening of the epidermis before (left) and after (right) induction of DKO. Bracket denotes the thickness of the epidermis. In the post-4OHT tissue, the indicated epidermal thickness is likely an underestimate. n = 2 animals.

C. Keratin10 (K10) staining in fixed SKO and DKO tissues. Dotted line indicates the location of the basement membrane.

D. The timing of the first or second round of division is shown for DKO cells, with respect to the 4OHT treatment time. SKO; n = 57 cells; DKO: n = 49 cells.

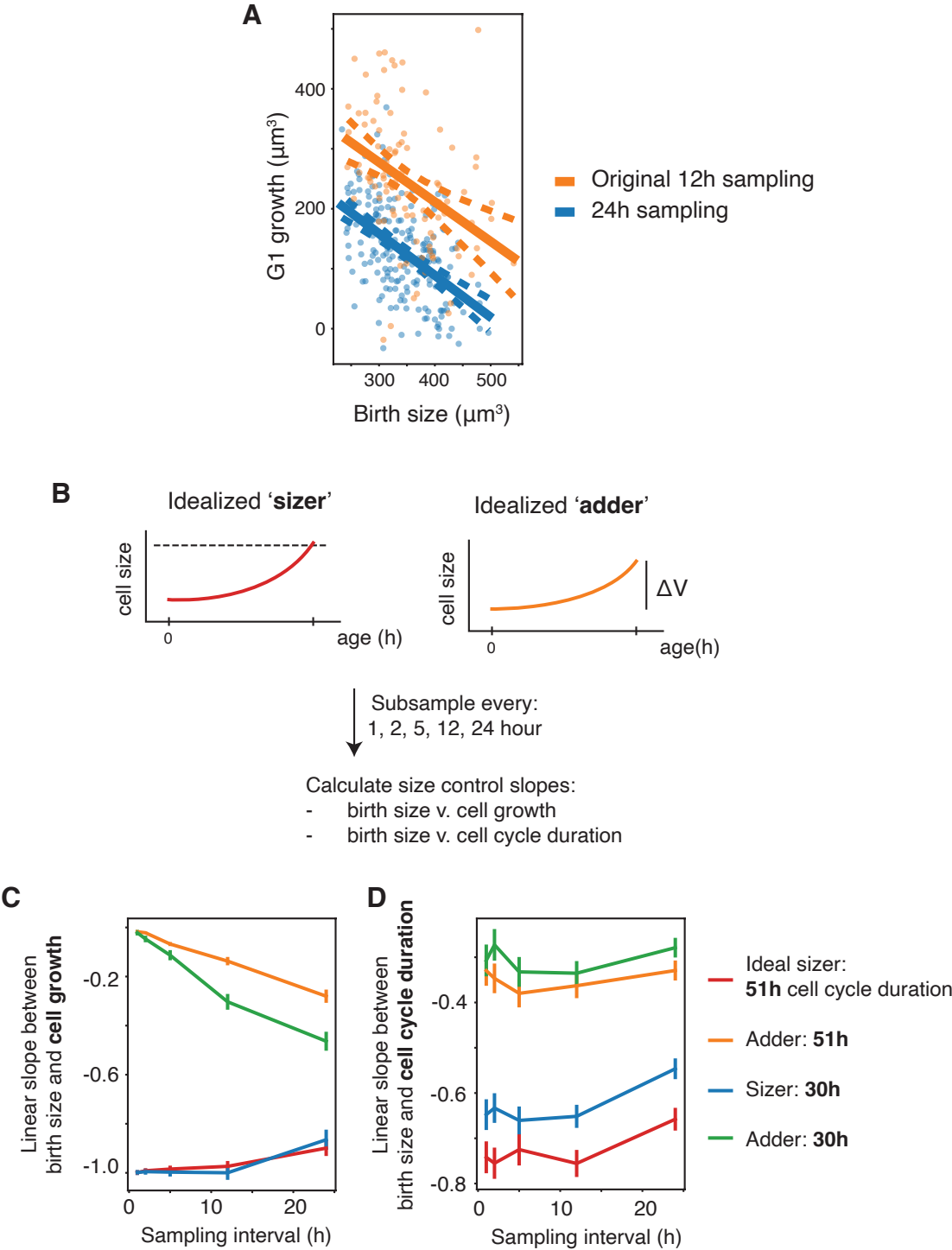
E. Representative images of SKO or DKO fixed tissues stained with a 53BP1 antibody. N = 1 animal.

F. Representative images of the *H2B-Cerulean* reporter in DKO live tissues with dead nuclear bodies indicated with arrows. N = 1 animal.

G-H. Histograms of the durations of G1 phase (G) or the entire cell cycle (H) for SKO cells, or the first and second round of division in DKO cells.

Scalebar is 10µm. Source data are provided as a Source Data file.

Supplemental Figure S10



Supplementary Figure S10: Analysis of the effect of temporal sampling rate on cell size control measurements.

A. The correlation between cell volume at birth is anti-correlated with total growth during G1 phase in epidermal stem cells that were sampled every 12h. A 2x random down-sampling to a 24h-sampling rate did not change the characteristic negative relationship between cell birth size and G1 growth, although it leads to a systematic underestimate of the total G1 growth amount overall. $n = 173$ cells.

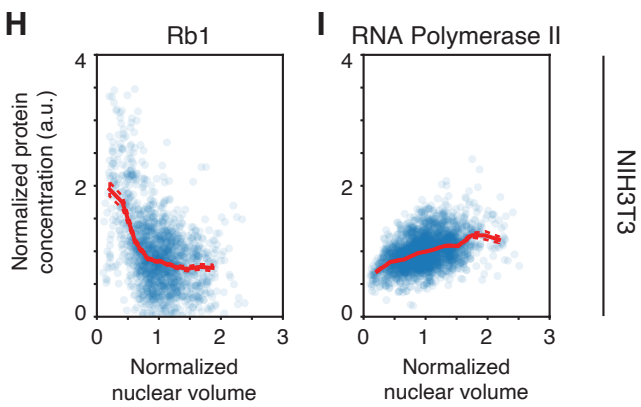
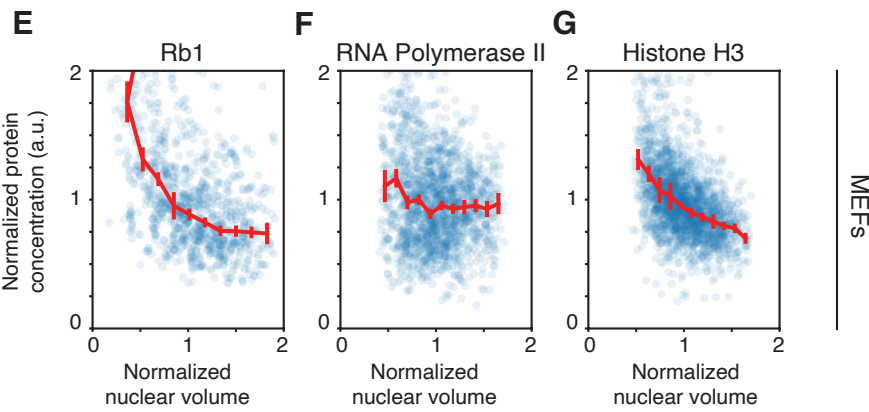
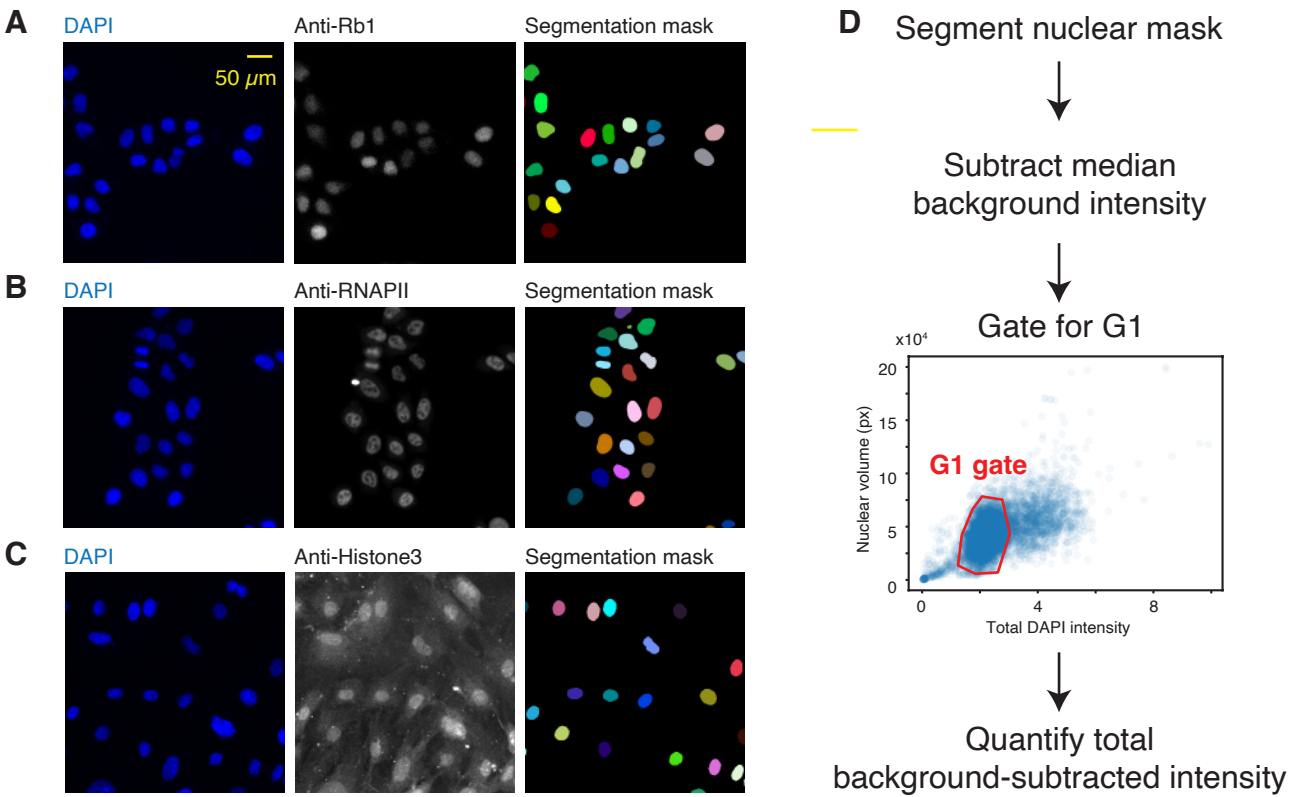
B. Schematic of the effect of sampling rate on cell size analysis with different total cell cycle durations. Idealized 'adder' and 'sizer' growing populations were generated with either 40h or 72h mean total cell cycle durations. Their growth curves were then subsampled to 1, 2, 5, 12, or 24h sampling rates. Then, the relationship between birth size and cell growth or cell cycle duration was calculated.

C. Linear slopes between birth size and cell growth are calculated for 'sizer' and 'adder' behaviors with mean cell cycle duration time similar to wildtype (72h) or DKO (40h) cells. The slope is shown with respect to the sampling interval time, showing that for 40h-cell cycle adders, the slope tends towards spurious negative relationship due to bias in the estimator. Errorbar is the mean slope $\pm$ mean 95% confidence interval of 1000 iterations.

D. Linear slopes between birth size and cell cycle duration are calculated for 'sizer' and 'adder' behaviors with mean cell cycle duration time similar to wildtype (72h) or DKO (40h) cells. The slope is shown with respect to the sampling interval time, showing that for 'adders', the slope is less sensitive to sampling rate. Errorbar is the mean slope $\pm$ mean 95% confidence interval of 1000 iterations.

P-values are two-sided T-tests. Source data are provided as a Source Data file.

Supplemental Figure S11



Supplementary Figure S11. Quantification of Rb1 protein subscaling in murine cell lines.

A-C. Representative images are shown for mouse embryonic fibroblasts (MEFs) stained with DAPI and anti-Rb1 (A), anti-RNAPII (B), or anti-H3 (C) antibodies. The segmentation prediction masks are also shown.

D. The quantification pipeline is shown. After segmentation, the background pixel median intensity was calculated and subtracted from the images. Then, total DAPI intensity and nuclear volume was used to gate for G1 cells.

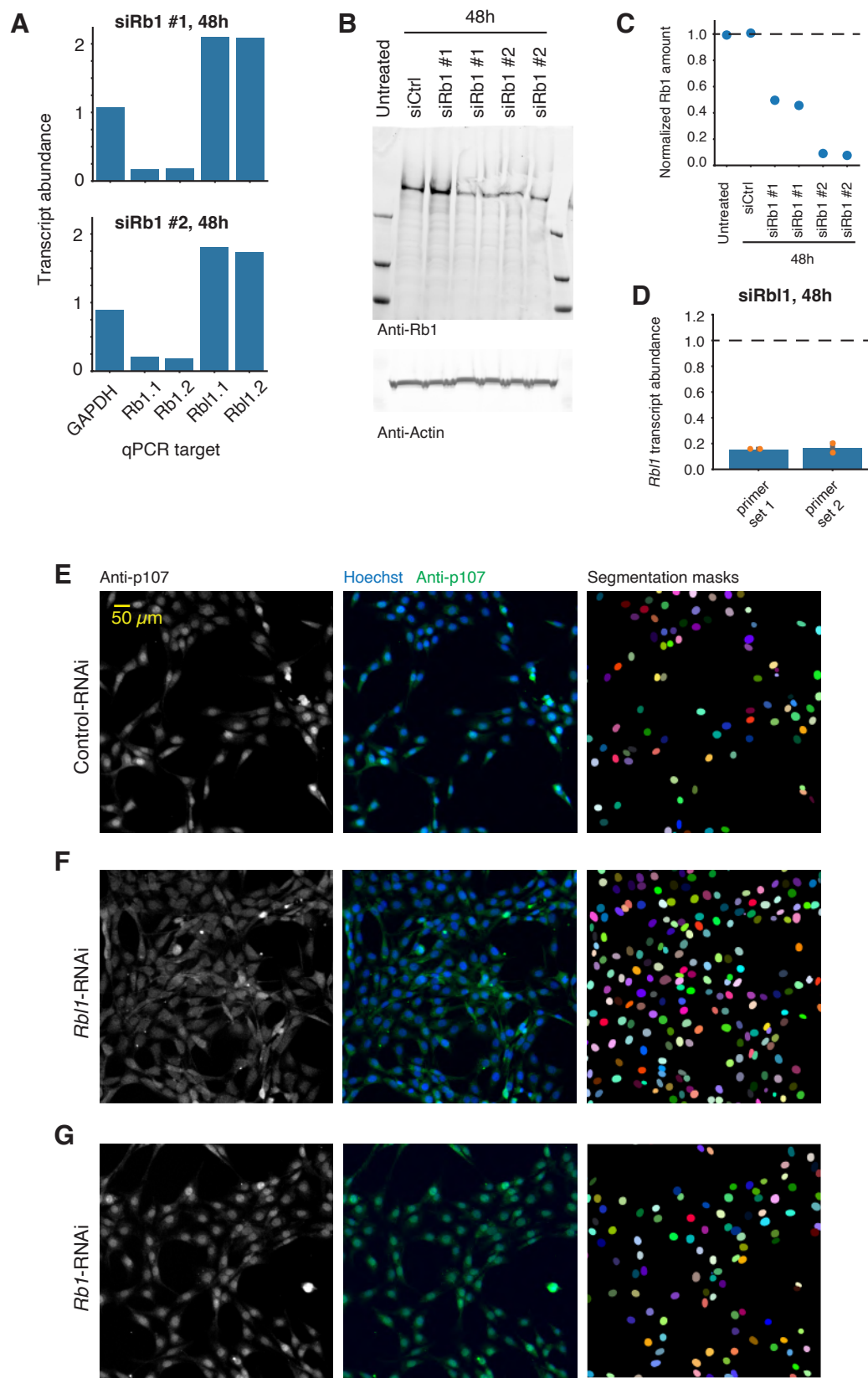
E-G. For MEFs, the mean-normalized protein concentration is quantified for Rb1 (E), RNAPII (F), or H3 (G) against the mean-normalized nuclear volume. Rb1: n = 1,098 cells; RNAPII: n = 1,751 cells; H3: n = 2,115 cells.

H-I. For NIH3T3s, the mean-normalized protein concentration is quantified for Rb1 (E), RNAPII (F), or H3 (G) against the mean-normalized nuclear volume. Rb1: n = 1617 cells; RNAPII: n = 2,464 cells.

Solid red line is the median. Error bars are SEMs.

Source data are provided as a Source Data file.

Supplemental Figure S12



Supplementary Figure S12. Quantification of Rbl1 subscaling in NIH3T3s.

A. qPCR was used to quantify *Gapdh*, *Rb1*, or *Rbl1* transcripts in NIH3T3s when *Rb1* was knocked down for 48h using two different siRNA constructs. mRNA abundance was normalized to *Actb*.

B. Immunoblot was used to probe against Rb1 in NIH3T3s treated with siRb1 or siControl for 48h. Anti-Actin was used as the loading control. Uncropped blot is shown in **Supplementary Figure S22**.

C. Quantification of the immunoblot in (B) shows that for siRb1 #2, the knockdown efficiency is ~90% after 48h of treatment.

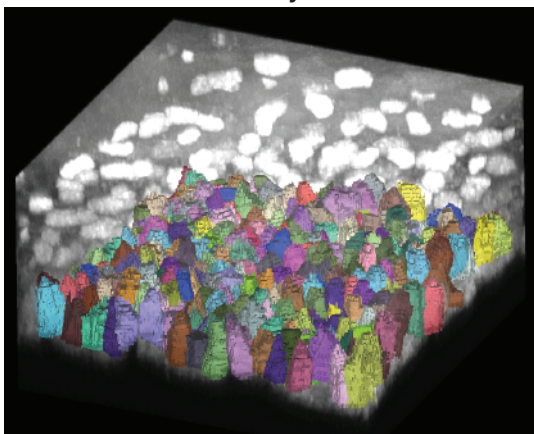
D. qPCR was used to quantify *Rbl1* transcript abundance in NIH3T3s when *Rbl1* was knocked down for 48h. mRNA abundance was normalized to *Actb*.

E-G. NIH3T3s treated with control (E), *Rbl1* (F), or *Rb1* (G) RNAi constructs for 96h are shown. The cells were stained with Hoechst and an anti-Rbl1 antibody. The segmentation masks are shown.

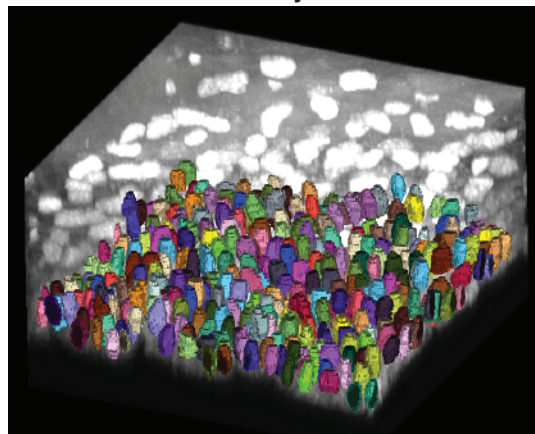
Source data are provided as a Source Data file.

Supplemental Figure S13

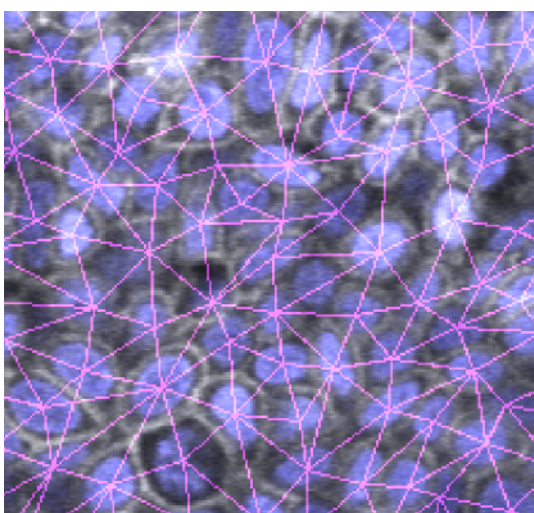
A Cortical segmentation of all basal layer cells



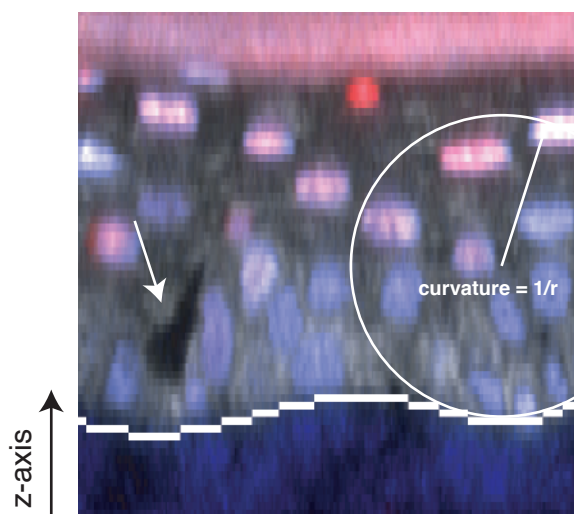
B Nuclear segmentation of all basal layer cells



C Adjacency annotation of all basal layer cells

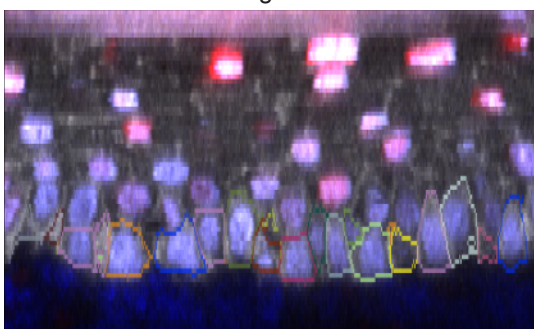


D Non-keratinocytes annotation and BM curvature estimation

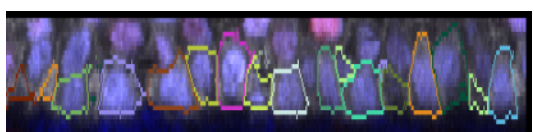


E Flatten tissue to measure cell geometry

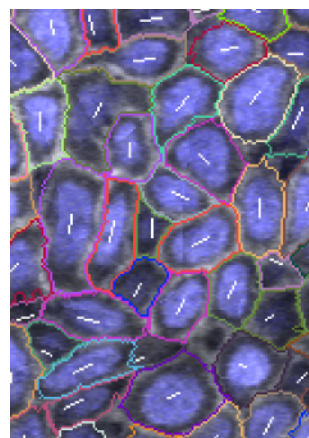
Original



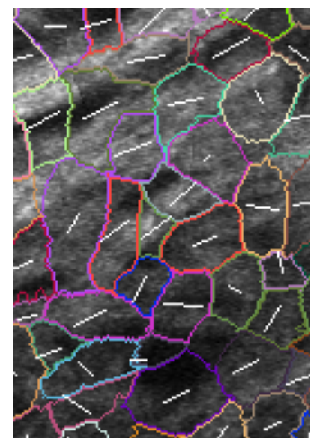
Flattened



F Basal shape orientation



G Dermal collagen orientation



Supplementary Figure S13. Measurement of cell and microenvironment morphometrics

A. Example image of the exhaustive 3D cell segmentation of all basal layer cells within a $\sim 120\mu\text{m} \times 120\mu\text{m} \times 100\mu\text{m}$ tissue volume.

B. Example image of the exhaustive 3D nuclear segmentation of all basal layer cells.

C. Example annotation of basal layer cell adjacencies, with each cell-cell interface denoted in magenta.

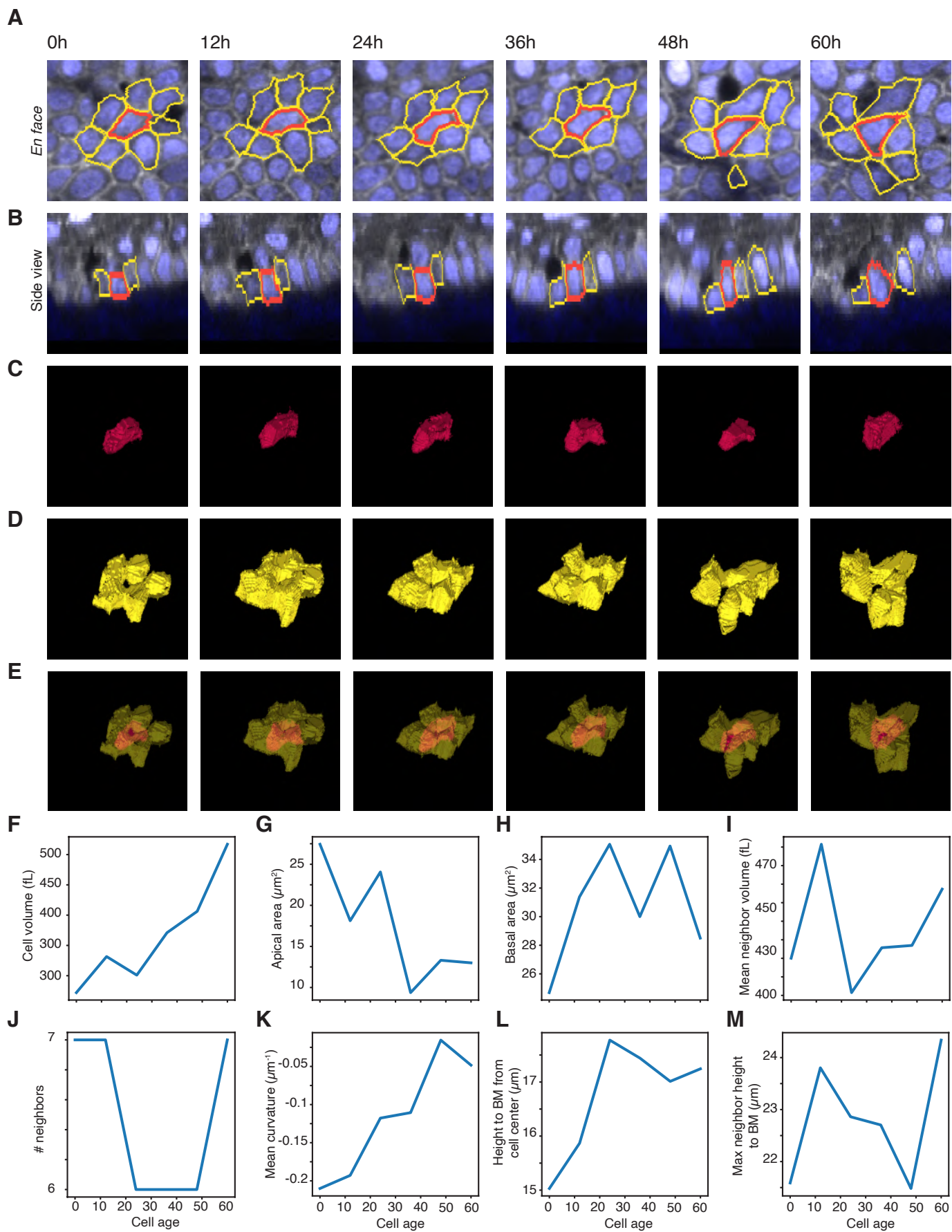
D. Example annotation of basement membrane (BM) location (solid line), calculation of BM curvature (white circle), and annotation of the location of a non-keratinocyte cell shadow in the epidermis (arrow).

E. Example of a curved epidermal tissue image being re-sliced to generate a flattened version of the tissue.

F. Example measurement of basal shape derived from flattened versions of basal layer cells. Colored outlines show the basal shape segmentation on top of a flattened basal image data. The angle of the white line shows the orientation of the basal shape. The length of the white line shows the eccentricity of the shape.

G. Example measurement of the orientation of dermal collagen fibrils. The same basal cell shapes are shown as in (F). The angle of the white line shows the mean orientation of the fibrils within the basal shape. The length of the white line shows the coherence of the fibril orientations within the shape.

Supplemental Figure S14



Supplementary Figure S14. Example cell and microenvironment morphometrics.

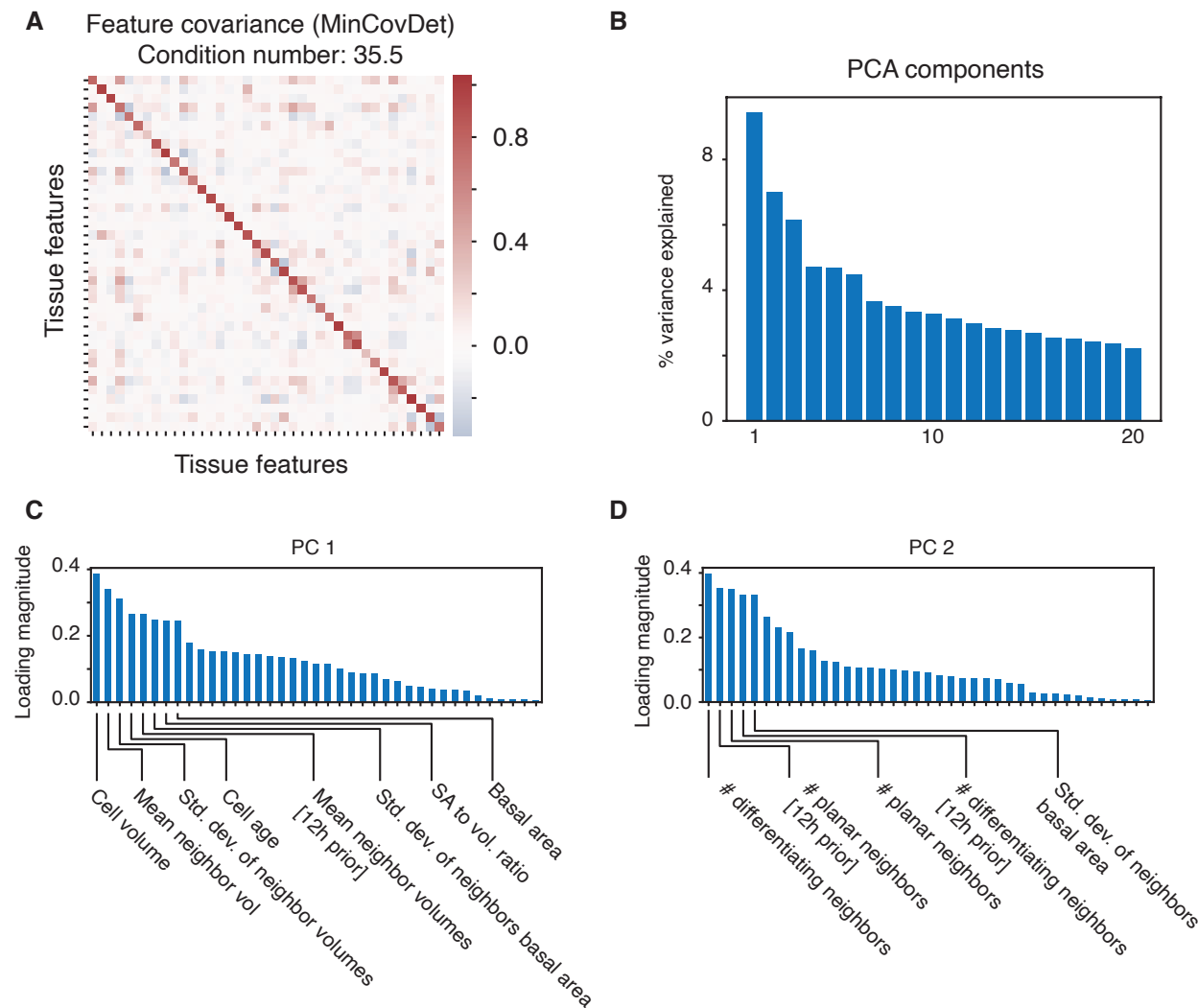
A-B. The 3D cell segmentations of a central cell (red) and its neighboring cells (yellow) are shown in *en face* (A) and as a side view (B). The cluster of cells are tracked from cell birth to cell division.

C-E. A montage of the 3D rendering of the same central cell (C) and its neighboring cells (D) are shown. The merged montage is shown in (E).

F-M. For the same central cell, the time-series of its cell volume (F), apical area (G), basal area (H), mean neighbor volume (I), number of neighbors (J), local tissue curvature (K), cell height from basement membrane (L), and maximum neighbor height to basement membrane (M) are shown.

Source data are provided as a Source Data file.

Supplemental Figure S15



Supplementary Figure S15. Cell and microenvironment morphometric feature-set.

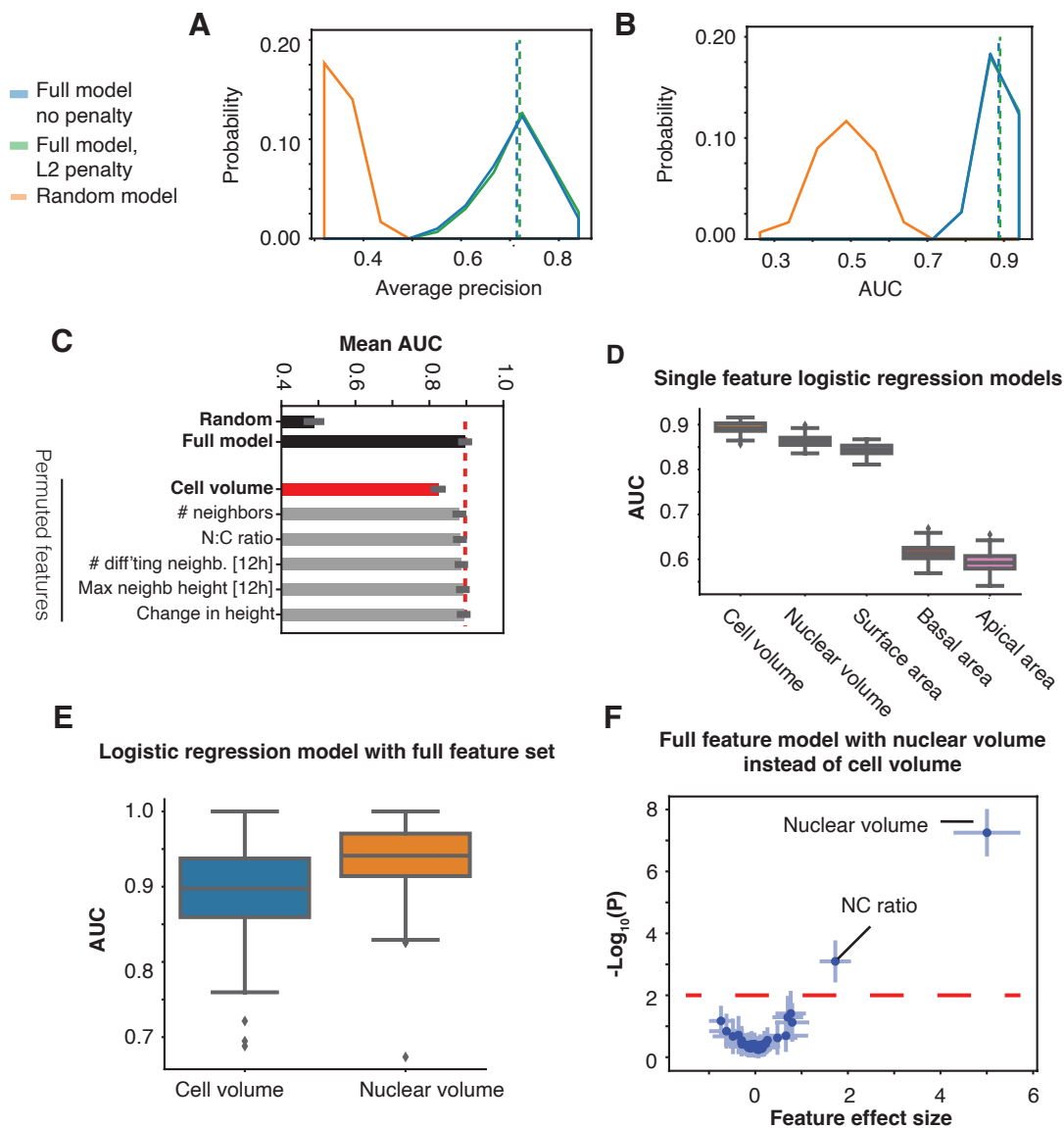
A. The covariance matrix of the feature set used for statistical modeling estimated using the minimum covariance determinant (MinCovDet) is shown ($n = 707$ samples). The condition number of invertibility, which is the ratio of the largest and smallest eigenvalue, is presented.

B. The percent of total variability explained is shown for each component of the data resulting from the principal component decomposition with 20 components.

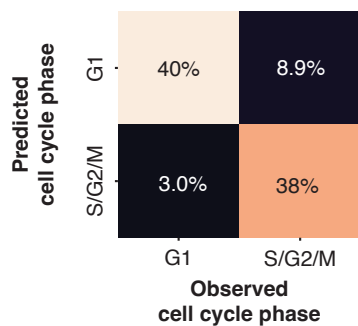
C-D. The top contributing features to the top principal components PC1 (C) and PC2 (D) are shown.

Source data are provided as a Source Data file.

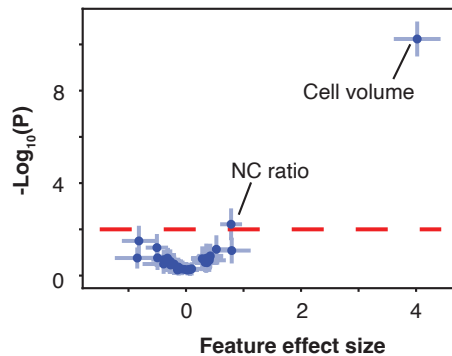
Supplemental Figure S16



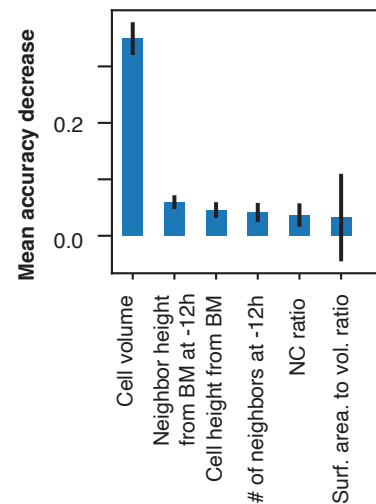
G



H Model with all S/G2/M time points included



I



Supplementary Figure S16. Multiple logistic regression model of cell cycle variability in basal layer epidermal cells.

A-B. Histograms of the average precision (A) or area under the curve (AUC) (B) of 1000 iterations of cross-validation (10% withheld test data) are shown for the full model or a model of the same size but generated from random data obtained from gaussian distributions. The mean performance of the full model is shown in dotted line.

C. The mean AUC is shown for cross-validation models with the entire feature-set, completely random features, or feature-sets where one of the indicated features was randomly permuted. Only cell volume shows significant decrease in model performance when permuted. Error bars indicate 95% confidence interval from 1000 cross-validations.

D. The AUC is shown for single-feature logistic regression models predicting G1 or post-G1 phases of the cell cycle.

E. The AUC is shown for full-feature multiple-logistic regression models that include either cell volume or nuclear volume alongside the rest of the cell and microenvironment features.

F. The effect size and statistical significance of each feature is quantified for a multiple logistic regression model that includes nuclear volume instead of cell volume.

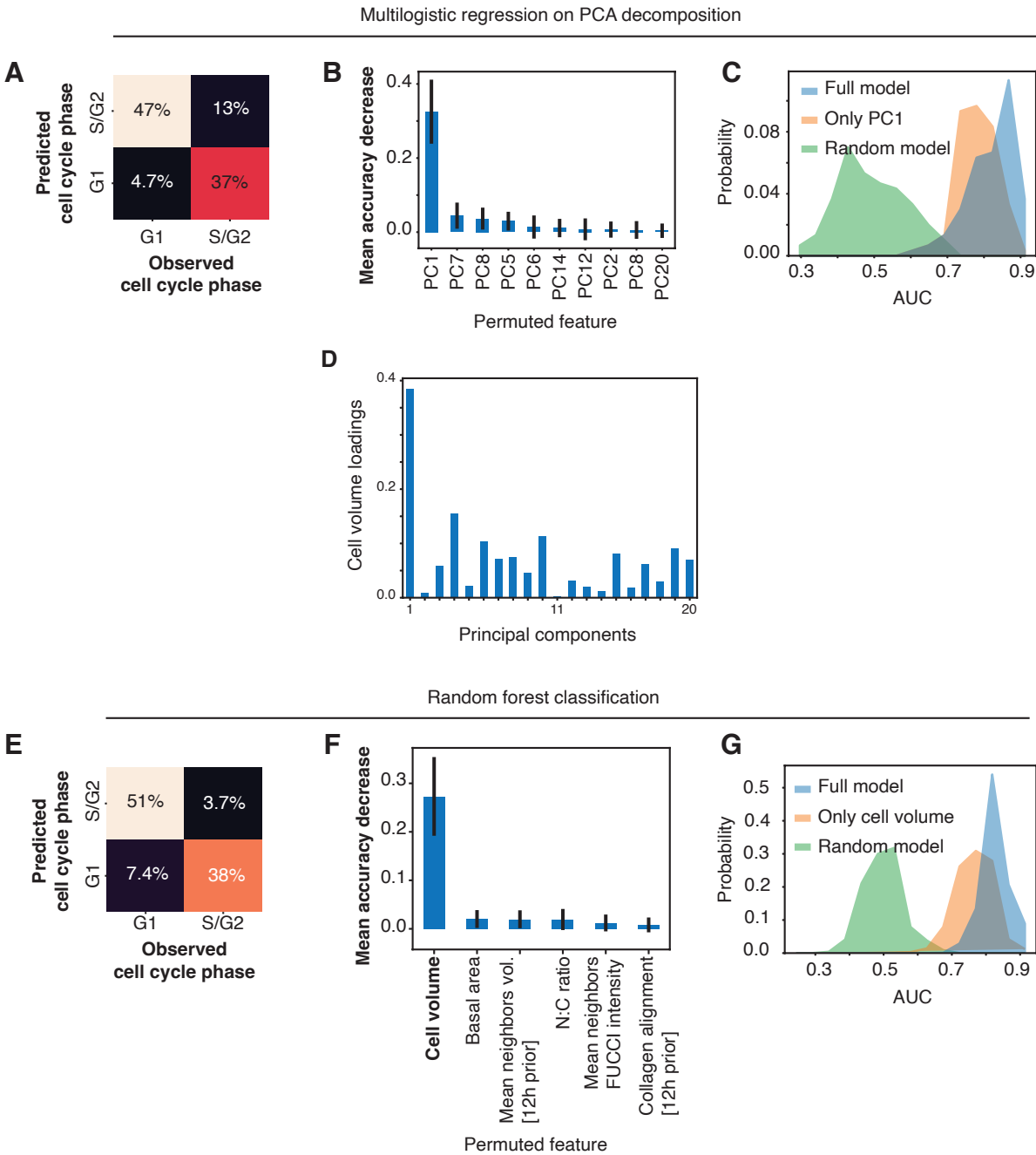
G. The confusion matrix of a multiple-logistic regression model is shown when trained with the full feature-set and *all* G1 and S/G2/M datapoints, not just the first post-G1 S/G2 time point as shown in Figure 4.

H. The effect size and p-value for individual features are shown for a multiple-logistic regression model trained with full feature-set and all S/G2/M datapoints.

I. The permutation importance is shown for individual features for a multiple-logistic regression model trained with full feature-set and all S/G2/M datapoints. Error bars are SEM from 1000 iterations.

Source data are provided as a Source Data file.

Supplemental Figure S17



Supplementary Figure S17. PCA-diagonalized models and random forest models.

A. The error matrix of a logistic regression model trained on PCA-diagonalized data is shown.

B. The relative model accuracy decrease is shown for logistic regression models trained on PCA-diagonalized data, where each indicated principal component was randomly permuted. Error bars indicate std.

C. The AUC metric is shown for logistic regression models trained on the full features, random features, or only PC1.

D. The loadings of cell volume as a feature onto each PC is shown.

E. The error matrix of a random forest model trained on the original data is shown.

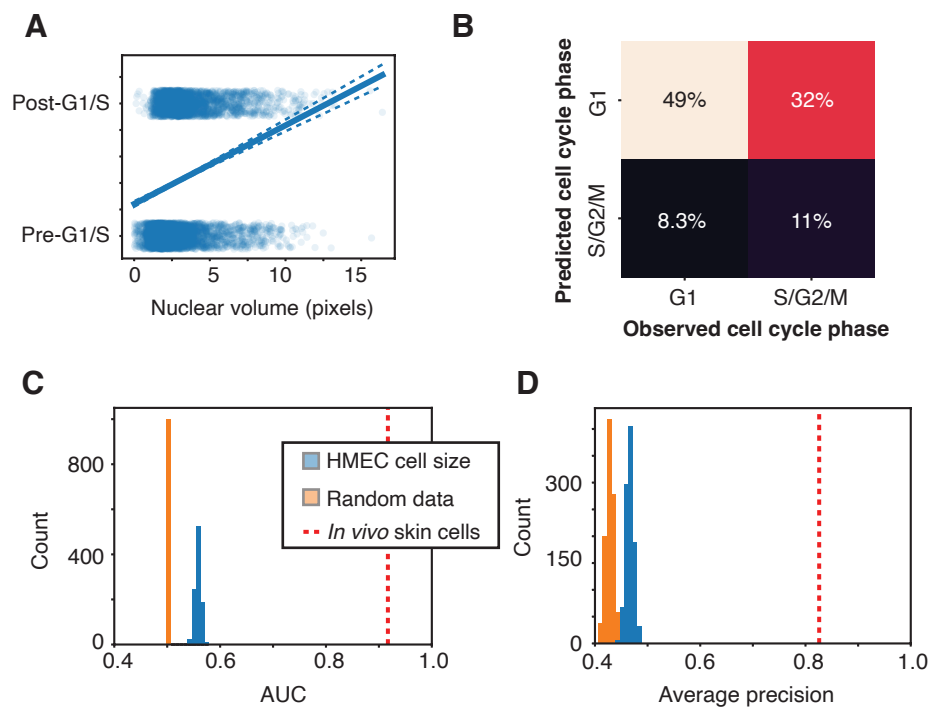
F. The relative model accuracy decrease is shown for random forest regression models trained on the original data, but where each indicated feature was randomly permuted. Error bars indicate std.

G. The AUC metric is shown for logistic regression models trained on the full features, random features, or only on cell volume.

Source data are provided as a Source Data file.

Supplemental Figure S18

HMECs culture *in vitro*



Supplementary Figure S18. G1 phase prediction model for *in vitro* HMEC cells.

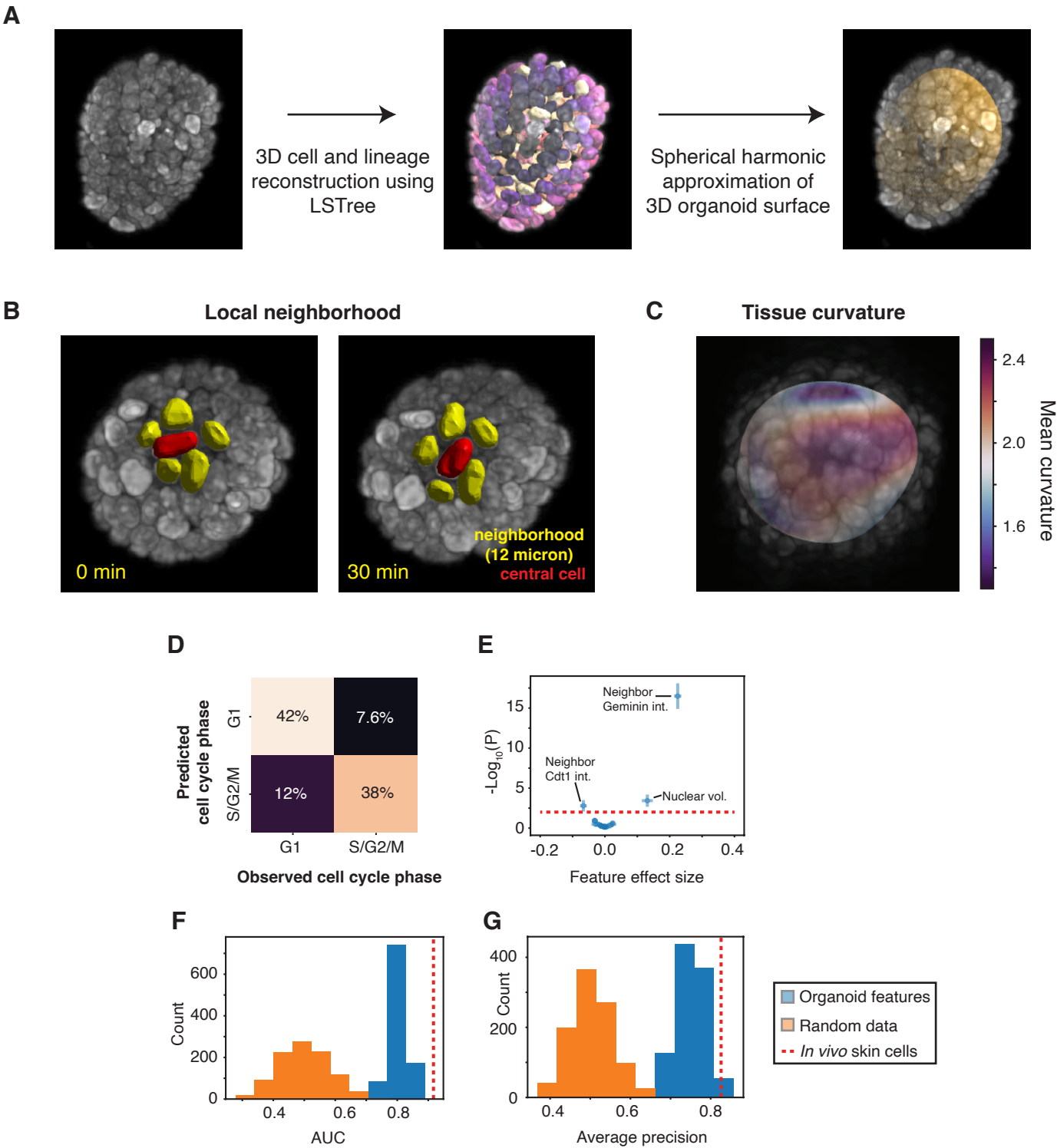
A. Single-variable logistic regression prediction for pre- or post-G1/S cell cycle phase is shown trained on live-cell imaging of HMEC cells grown *in vitro*. Nuclear volume was used as the endogenous variable. Solid line is logistic regression curve. Dotted line is 95% confidence interval.

B. The error matrix of the logistic regression model in (A) is shown.

C-D. The AUC (C) or average precision (D) metrics are shown for the logistic regression model in (A), compared with the performance of either randomized data, or the performance of the *in vivo* skin dataset.

Source data are provided as a Source Data file.

Supplemental Figure S19



Supplementary Figure S19. Microenvironmental feature extraction and cell cycle prediction for cells in intestinal organoids.

A. The 3D geometric feature extraction pipeline based on LSTree is shown.³⁸ Cells and their lineages were reconstructed using LSTree, and subsequently, a spherical harmonic representation was used to approximate the overall 3D geometry of the entire organoid.

B. An example of the 12 μ m local neighborhood (yellow) is shown surrounding a central cell (red).

C. The local tissue curvature quantification is shown for the example organoid.

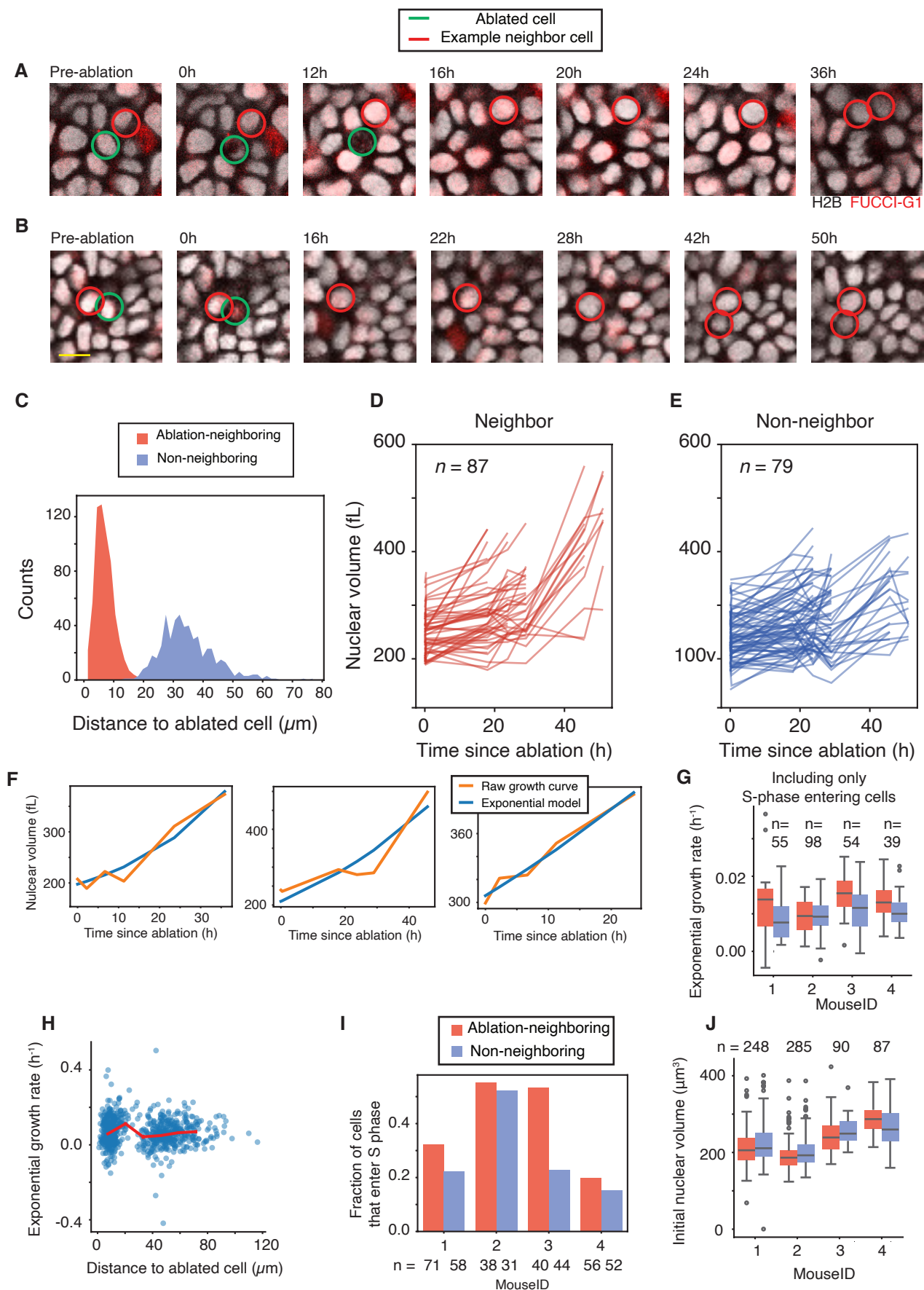
D. The error matrix of the multiple logistic regression model that incorporate cell-intrinsic and cell-extrinsic organoid features is shown (see Methods for details).

E. The feature effect size and significance are shown for significant predictors of G1/S phase in the model shown in (D). Neighbor cell FUCI intensity as well as nuclear volume have significant predictive power, indicating that spheroid cells are more likely to divide synchronously.

F-G. The AUC (F) or average precision (G) metrics are shown for the multiple-logistic regression model in (D), compared with the performance of either randomized data, or the performance of the *in vivo* skin dataset.

Source data are provided as a Source Data file.

Supplemental Figure S20



Supplementary Figure S20. Single cell ablation.

A-B. Example montages of cell ablation time-series, showing central ablated cells (green) and single neighbor cells (red) tracked through time. Only a single tracked neighbor cell is highlighted in each montage for clarity. Scale bar is 10 μ m. Cells were ablated using a two-photon laser as described in the Methods.

C. The distributions of distances to the ablation site from the centroids of ablation-neighboring and non-neighboring cells are shown. N = 4 mice, n = 749 cells.

D-E. The nuclear volume growth curves are shown in a representative ablation experiment, showing cells that neighbored the ablation site (D) and those that were farther away (E). N = 4 ablation experiments.

F. Example nuclear growth curves and their fitted exponential growth models.

G. The exponential growth rate estimated from only the subset of ablation-neighboring or non-neighboring cells that enter S phase within the span of the respective datasets.

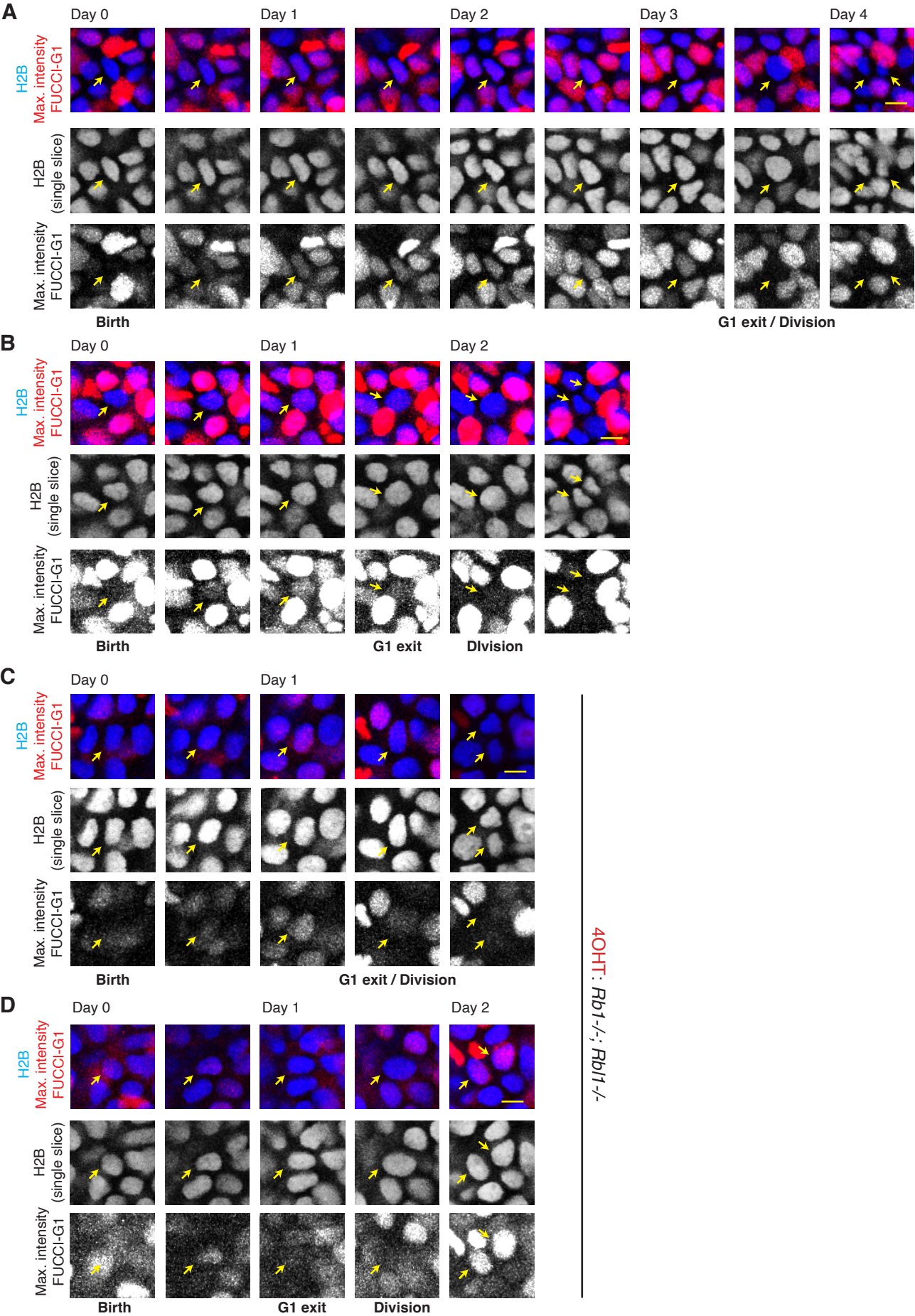
H. The exponential growth rate of all measured cells are quantified against the distance to the nearest ablation site. N = 4 mice, n = 749 cells. Red line is median $\pm$ SEM.

I. The fraction of cells that enter S phase within the span of their respective datasets are quantified for ablation-neighboring or non-neighboring regions.

J. The initial nuclear volumes of ablation-neighboring and non-neighboring cells are quantified right after ablation for each animal.

For boxplots, center is the median, box range is the quartile range, and whiskers show non-outlier data range. Source data are provided as a Source Data file.

Supplemental Figure S21



Supplementary Figure S21. Example *FUCCI-G1* signal in SKO and DKO tissues.

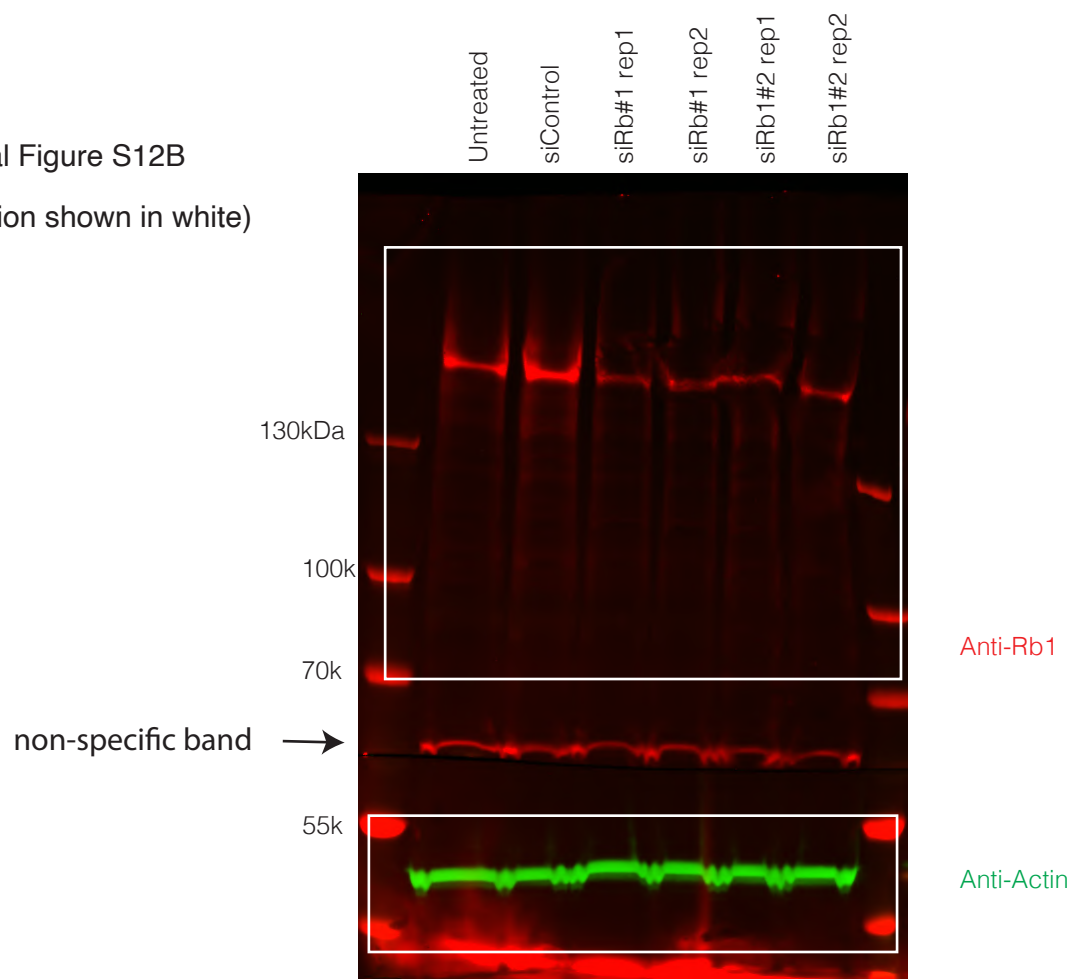
A-B. Example *Rb1*^{-/-} cells. Middle row shows a single middle z-plane of the cell tracked from birth to division. Bottom row shows the maximum intensity projection of the *FUCCI-G1* signal through the nuclear volume of the cell. Top row shows the merge.

C-D. Example *Rb1*^{-/-}; *Rb1*^{-/-} cells. Middle row shows a single middle z-plane of the cell tracked from birth to division. Bottom row shows the maximum intensity projection of the *FUCCI-G1* signal through the nuclear volume of the cell. Top row shows the merge.

Supplemental Figure S22

Related to
Supplemental Figure S12B

(cropped region shown in white)



Western blot of whole cell lysates of NIH3T3s treated with the indicated siRNA

Supplementary Figure S22. Uncropped gel images.

The uncropped Western blot is shown of the blot in Figure S12B.