Local cellular neighborhood controls proliferation in cell competition

Anna Bove[a,b,†], Daniel Gradeci[a,c,†], Yasuyuki Fujita[d], Shiladitya Banerjee[c,e,*], Guillaume Charras[a,b,e,*], and Alan R. Lowe[a,f,*]

[a]London Centre for Nanotechnology, University College London, London WC1H 0AH, United Kingdom; [b]Department of Cell and Developmental Biology, [c]Department of Physics and Astronomy, [e]Institute for the Physics of Living Systems, and [f]Institute for Structural and Molecular Biology, University College London, London WC1E 6BT, United Kingdom; [d]Division of Molecular Oncology, Institute for Genetic Medicine, Hokkaido University, Kita-ku, Sapporo, Hokkaido 060-0815, Japan

ABSTRACT Cell competition is a quality-control mechanism through which tissues eliminate unfit cells. Cell competition can result from short-range biochemical inductions or long-range mechanical cues. However, little is known about how cell-scale interactions give rise to population shifts in tissues, due to the lack of experimental and computational tools to efficiently characterize interactions at the single-cell level. Here, we address these challenges by combining long-term automated microscopy with deep-learning image analysis to decipher how single-cell behavior determines tissue makeup during competition. Using our high-throughput analysis pipeline, we show that competitive interactions between MDCK wild-type cells and cells depleted of the polarity protein scribble are governed by differential sensitivity to local density and the cell type of each cell's neighbors. We find that local density has a dramatic effect on the rate of division and apoptosis under competitive conditions. Strikingly, our analysis reveals that proliferation of the winner cells is up-regulated in neighborhoods mostly populated by loser cells. These data suggest that tissue-scale population shifts are strongly affected by cellular-scale tissue organization. We present a quantitative mathematical model that demonstrates the effect of neighbor cell–type dependence of apoptosis and division in determining the fitness of competing cell lines.

INTRODUCTION

Competition between cells is a phenomenon originally identified in development that results in the elimination of less fit cells (the loser cells) from a tissue (Levayer and Moreno, 2013; Vincent *et al.*, 2013; Merino *et al.*, 2016). The viability of loser cells depends strongly on context: when they are cultured alone, they thrive, but when they are cultured in a mixed population, they are eliminated by cells with greater fitness. Many of the mutations leading to competition give rise to a change in growth rate, with the faster-growing cells eventually eliminating the slower-growing ones (Morata and Ripoll, 1975; Simpson and Morata, 1981; Oliver *et al.*, 2004). The relationship between cell competition and cancer is complex. In some cases, competition can confer protection against tumorigenesis by eliminating cells with oncogenic mutations (Hogan *et al.*, 2009; Tamori *et al.*, 2010; Norman *et al.*, 2012; Martins *et al.*, 2014; Merino *et al.*, 2015; Porazinski *et al.*, 2016), but in other cases, cancer cells can turn competition to their advantage, leading to field cancerization (Rhiner and Moreno, 2009; Fernandez *et al.*, 2016). In cancer, many different lineages with distinct mutations are present in a tumor (Navin *et al.*, 2011; Sottoriva *et al.*, 2013), and their interaction can be viewed as a competition. Recent computational work has shown that a small competitive advantage, linked, for example, to better survival in the presence of a drug, can lead to one lineage taking over the whole tumor (Waclaw *et al.*, 2015). Thus, a detailed understanding of cell competition, how it depends on local context, and

how competition is modulated by the environment are of clear therapeutic interest.

Competition involves either the exchange of short-range biochemical information (biochemical competition) or longer-range mechanical cues (mechanical competition). In the former, competing cells must be in contact with one another and, as a result, cell death occurs only at the interface between cell types (Moreno *et al.*, 2002). In mechanical competition, cell death occurs because of differences in tolerance to cell density, and the role of contact interaction is less clear. To date, most studies have quantified competition at the tissue scale mainly concentrating on increases in apoptotic events and reporting these across tissues (Levayer *et al.*, 2015; Wagstaff *et al.*, 2016). Although these studies have yielded insight into the mechanisms of competition, many important aspects still remain unclear. Indeed, changes in population composition can arise not only from increases in apoptosis but also through changes in division rates. Furthermore, as competition is by nature context dependent, a tissue-scale description of outcome obscures key characteristics of how competition takes place at the single-cell level. For example, recent work has revealed that, when loser cells have a limited fraction of their surface in contact with winner cells, they survive (Levayer *et al.*, 2015).

These examples highlight the need to quantitatively characterize the apoptosis and division rates of each cell type and investigate how local cell density and neighborhood modulate these. Such characterization presents several challenges. First, cell trajectories must be accurately tracked (Hilsenbeck *et al.*, 2016) and the cell cycle states determined over long durations for several hundred cells. Second, new data analysis tools and metrics must be developed to characterize the spatiotemporal rules of competition. Here, we address these challenges by developing long-term automated microscopy together with deep-learning image analysis to decipher single-cell behaviors underlying population shifts during competition. As a test case for our analysis pipeline, we examine competition between MDCK cells depleted in the polarity protein scribble (scribble^kd) and wild-type cells (MDCK^WT).

Competition induced by loss of scribble is widely regarded as an example of mechanical competition (Norman *et al.*, 2012). Recent work has suggested that mechanical competition might stem from differences in the tolerances of each cell type to density. Indeed, when cultured in pure populations, the density in each cell population first increases before reaching a plateau (the homeostatic density). This homeostatic density is maintained over time, because the rate of cell death exactly compensates the rate of cell birth. Remarkably, scribble^kd cells reached lower homeostatic densities than MDCK^WT cells because of a threefold larger apoptosis rate (Wagstaff *et al.*, 2016). In competitive conditions, MDCK^WT cells compacted scribble^kd cells, causing an increase in their death rates. These results suggest changes in density may induce increases in apoptosis events that alone are sufficient to result in the elimination of the scribble^kd cells. However, a detailed characterization of the influence of density on apoptosis is lacking for a quantitative test of this hypothesis, and the sensitivity of cell division to density has never been characterized.

Here, we take a first step in characterizing apoptosis, cell division, and net growth as a function of local cell density. Next, we show that the local neighborhood influences apoptosis and division, suggesting a role for inductive phenomena in mechanical competition and emphasizing the importance of examining single-cell behaviors to characterize competition at the tissue scale. Finally, we implement a simple numerical model of competition to investigate how the density dependence of cell division and apoptosis determine the time evolution of cell count and the overall population fitness. Our analytical tools can be generally applied to determine the dependence of cell competition on local density and local neighborhood and to analyze cell interactions with relevance to cancer, developmental biology, and stem cell biology.

Experimental and computational strategy

To characterize how single-cell behaviors give rise to population shifts during cell competition, we developed a high-throughput imaging and analysis pipeline. We acquired time-lapse movies lasting several days, imaging cocultures of MDCK^WT cells and scribble^kd cells expressing H2b–green fluorescent protein (H2b-GFP) and H2b–red fluorescent protein (H2b-RFP), respectively. We first confirmed that exposure of scribble^kd cells to tetracycline for 70 h led to ~90% reduction in scribble expression, consistent with previous reports (Norman *et al.*, 2012) (Supplemental Figure S1E). We assembled a low-cost cell-imaging system for multiposition and multiwavelength acquisition inside a standard CO_2 incubator. Using fast-switching LEDs for illumination and an automated stage, we acquire brightfield, GFP, and RFP images for up to 12 regions of 530 × 400 µm² with a 20× objective with a temporal resolution of 4 min for more than 80 h (Figure 1, A and B). Regions can be sampled from multiple Petri dishes, allowing competitions and controls to run side by side in identical conditions. In these conditions, the fate of loser cells (scribble^kd) surrounded by winner cells (MDCK^WT) can be easily traced as they undergo apoptosis before extrusion from the monolayer (Figure 1B).

Movies were then automatically analyzed to track the position, state, and lineage of the cells using deep learning–based image classification in a first processing step. In this first step, cell nuclei are segmented (Figure 1, A–C) and classified using a deep neural network to assign them one of five states: interphase, prometaphase, metaphase, anaphase, or apoptosis (Figure 1, D–G). The positions and states of nuclei are then linked into trajectories over time (Figure 1H) using a Bayesian tracking method followed by hypothesis-based optimization to generate lineage trees (Figure 2). In a second processing step, we determined the rates of division, apoptosis, and net growth as a function of density and local neighborhood. These data were then used in numerical models of cell competition. Further details of the segmentation and tracking algorithms along with comparison to existing approaches can be found in the Supplemental Material.

Over the course of a single imaging experiment, we acquired 12 movies in parallel, following 800–1000 cells per field of view for 800–1200 frames. The analysis of such an experiment results in the determination of the fate of 9600–12,000 cells (800–1000 cells × 12) and returns 640,000–1,200,000 (800–1000 × 800–1200) discrete cellular-scale observations (cells/time) per field of view. In the following sections, we define an "observation" as the detection of an object (a cell) at a given frame, while an "event" refers to the detection of an apoptosis or a cell division.

RESULTS

Cell count, cell cycle length, and apoptosis in pure and mixed populations

Following tracking and identification of cell cycle state (Supplemental Movies 1–3), we could generate lineage trees for each cell identifying its progeny and potential termination by apoptosis (Figure 2, A–D). By measuring the temporal separation between the birth of a cell and that of its daughters on a lineage tree, we extract the duration of the cell cycle at single-cell resolution (Figure 2B) and plot its distribution for the entire population (Figure 2E). For MDCK^WT, this
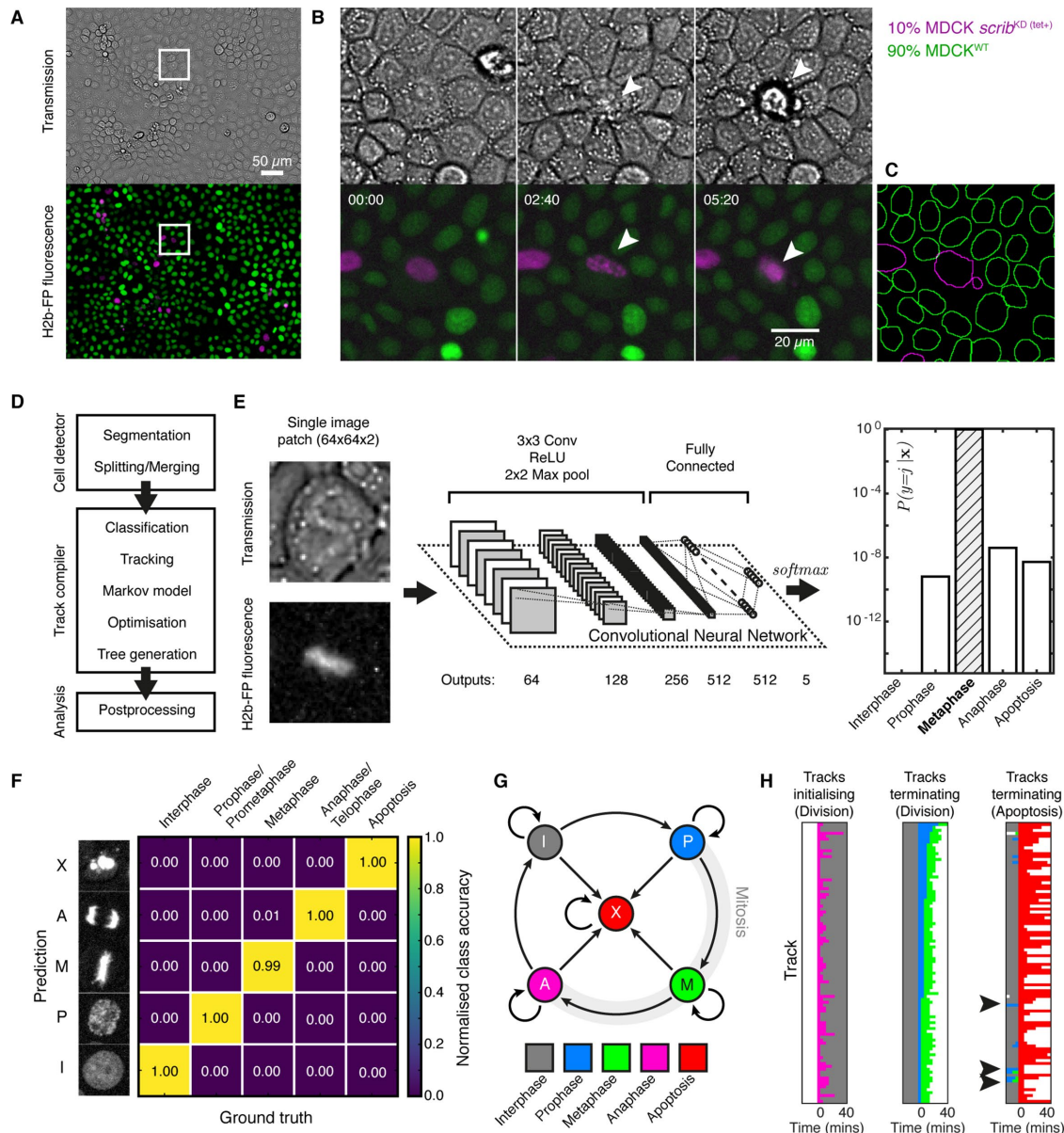
FIGURE 1: High-throughput image analysis pipeline for long-term analysis of single-cell dynamics. (A) Representative image containing H2b-GFP MDCK^WT (green) and H2b-RFP scribble^kd (magenta) cells mixed in an initial 90:10 ratio. The image corresponds to one field of view (1600 × 1200 pixels, 530 μm × 400 μm) acquired by wide-field epifluorescence using 20× magnification. The region of interest in the white rectangle is shown in B. (B) Time series from the competition assay in the region boxed in A. The white arrow indicates a scribble^kd cell surrounded by MDCK^WT neighbors that undergoes apoptosis. The acquisition of both transmission and fluorescence images enables detection of the apoptotic fragmentation of the loser's nucleus happening before extrusion of the cell body from the monolayer. Timings are indicated in the top left corner in hours and minutes. (C) Segmentation of the final image in B. MDCK^WT cells are outlined in green and scribble^kd cells in magenta. (D) Flowchart of the computational pipeline implemented for the study of competition dynamics. The strategy is based on segmentation of individual cells (cell detector), automatic annotation of morphological classes related to cell cycle state and apoptosis (track compiler), and postprocessing analysis of single-cell tracking data. (E) CNN for object classification. The CNN inputs are single-object patches, both in the transmission (BF) and fluorescence channels (left). The CNN stacks together four types of layers: convolutional/ReLU/max-pooling, and fully connected layers (middle). The CNN transforms the input image layer by layer from the original pixel values to the final class scores with the highest score reflecting the most probable classification of the image data (right). (F) Confusion matrix showing the matching of human annotations vs. the annotation of the CNN system. (G) HMM used for modeling progression through the cell cycle. The figure depicts the permitted directional transitions between five classes (interphase, prometaphase, metaphase, anaphase, and apoptosis). (H) Automated annotation of cell trajectories over time. A random selection of 100 trajectories (rows) is aligned and shown over a 40-min period. Colors refer to state labels as defined in F. Left, tracks following division start with anaphase before proceeding to interphase. Middle, tracks terminating in a division are preceded by interphase before going through prometaphase and metaphase. Right, tracks terminating with apoptosis are often preceded by interphase but can arise through failed division events (highlighted with arrows).
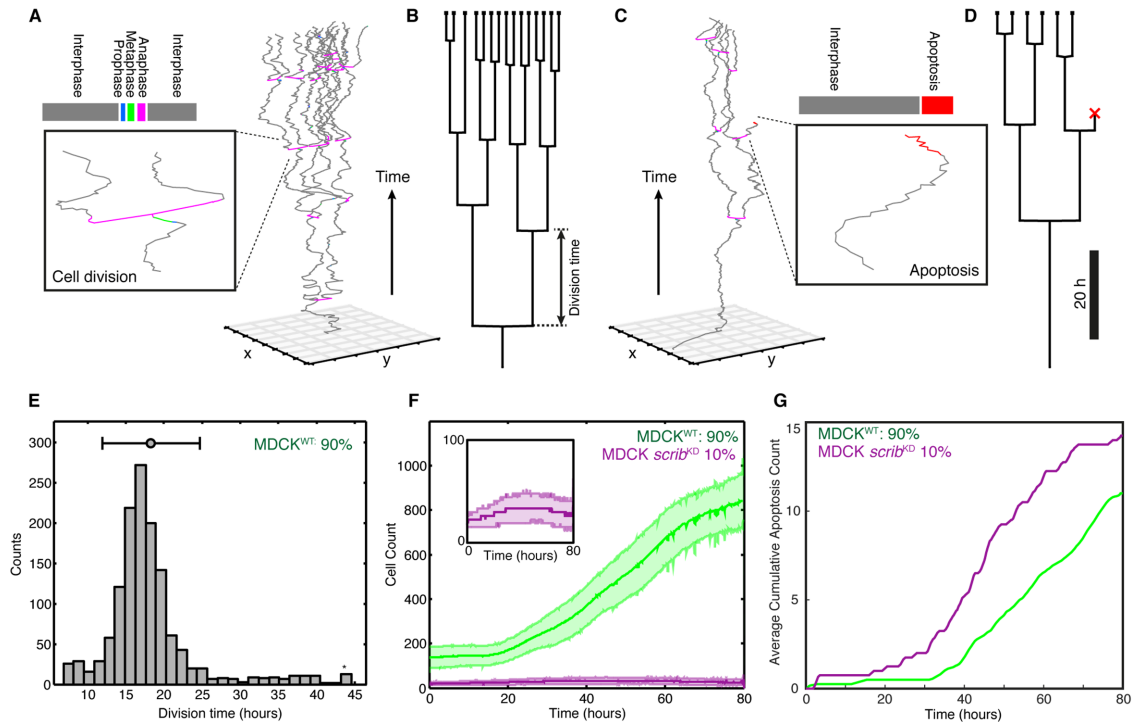
FIGURE 2: Cell lineage, cell cycle duration, and apoptosis detection during cell competition. (A) Representative cell tracks assembled into lineage trees as result of the global optimization. The original progenitor cell is at the bottom of the diagram, time is on the z-axis, and the position of all progeny is plotted in the x–y plane. Tracks in which the metaphase to anaphase transition is present are split and labeled as division events. In our classification, the metaphase state corresponds to a branching point, and new tracks start with the anaphase state. Each cell is assigned a unique cell ID (not shown in the figure). (B) Representative lineage tree for the cells shown in A. Cell cycle time can be measured for each individual cell as the time between consecutive division events on the lineage tree. (C) Representative cell tracks assembled into a lineage tree, showing a single apoptotic termination event, following multiple observations of the apoptotic state. The apoptotic states are labeled in red. (D) The lineage tree of the cells shown in C. (E) Histogram of cell cycle time measured for a population of MDCK[WT] showing a mean value of 18 ± 3.2 h. The distribution represents a population of at least 250 cells in the same field of view followed for 80 h. A total of 1326 division events were observed. (F) Proliferation profiles of MDCK[WT] (green) and scribble[kd] (magenta) during a competition. Data are pooled from three biological replicates imaging four fields of view for each replicate. The solid line indicates the mean of the experiments and the shaded area indicates the SD. The inset shows the evolution of cell count in the scribble[kd] cells on a smaller scale. (G) Quantification of apoptotic events for MDCK[WT] (green) and scribble[kd] (magenta) during competition. The number of apoptoses is detected and averaged across the four areas imaged during one competition assay.

yields a mean cell cycle time of 18 ± 3.2 h, consistent with other reports (Puliafito *et al.*, 2012; Gudipaty *et al.*, 2017) and validating our segmentation, classification, and tracking steps. It is important to note that, despite accurate estimation of cell cycle duration for the vast majority of the population, some errors subsist due to identity swaps at high cell densities or triple divisions that lead to underestimation of the cell cycle time (e.g., ~1.5% of cells have estimated cell cycle times of 5–10 h; Figure 2E).

As further validation of our algorithms, we analyzed the change in cell numbers and the number of apoptoses in a competition assay in which MDCK[WT] cells were mixed with scribble[kd] cells. We mixed cells in a 90:10 MDCK[WT]:scribble[kd] ratio with an initial seeding density of 10^{-3} cells/μm². Over the course of 80 h, scribble[kd] and MDCK[WT] proliferation differed markedly, with the scribble[kd] count peaking after 40 h before decreasing (Figure 2F, inset) and MDCK[WT] count showing a sevenfold increase (Figure 2F), consistent with previous work (Wagstaff *et al.*, 2016). Previous work has revealed increased cell death in the scribble[kd] cells when in the presence of MDCK[WT] cells (Wagstaff *et al.*, 2016). To confirm this, we plotted the cumulative count of apoptosis events for each cell type (Figure 2G). This revealed that scribble[kd] cells had significantly

higher counts of apoptosis than MDCK[WT] cells, despite being far scarcer, indicating a higher probability of apoptosis in scribble[kd] cells.

In contrast, control experiments mixing H2b-GFP MDCK[WT] with noninduced H2b-RFP scribble[kd] in a 90:10 ratio showed a twofold increase in cell count for H2b-RFP scribble[kd] and a threefold increase in MDCK[WT] cells after 50 h (Supplemental Figure S1A). These increases were comparable to those recorded in pure populations of each cell type over a similar duration (Supplemental Figure S1D). When we examined apoptosis in each cell population, we found that WT cells had significantly higher apoptosis counts, as would be expected from their 10-fold larger numbers (Supplemental Figure S1B) and in contrast to what is observed upon depletion (Figure 1G). Together, these data confirmed that, when scribble depletion is not induced, scribble[kd] and MDCK[WT] cells do not compete (Norman *et al.*, 2012).

Altogether, our large-scale quantification of growth rates and apoptosis in competitive and noncompetitive conditions are consistent with previous findings (Wagstaff *et al.*, 2016), confirming that our analysis pipeline can provide reliable high-throughput automated quantification of single-cell events during competition.

Probability of division and apoptosis depend on local cell density in competitive interactions

Previous experiments showed that scribble[kd] cells are less tolerant to density than MDCK[WT] cells, resulting in a higher rate of apoptosis for a given density (Wagstaff *et al.*, 2016). During competition, MDCK[WT] cells cause compaction of the scribble[kd] cells inducing them to apoptose. However, the exact relationship between cell density and probability of apoptosis was not determined. Furthermore, although a decrease in probability of division with density would also contribute to competition, the influence of cell density on cell division has not been examined.

Using our analysis pipeline, we investigated the impact of cellular density on the probability of division and apoptosis. For this, we implemented a measure of single-cell density based on a Delaunay triangulation of the centers of mass of cell nuclei in each image frame (Figure 3A). In computing density, we verify that cells are in contact with one another using a distance threshold (Supplemental Figures S2 and S3), and we exclude cells that are closer than 10 μm to the edge of the field of view, because not all of their neighbors can be identified, leading to an underestimation of density. We defined the local cellular density ρ as the sum of inverse areas of the triangles that share a common vertex with the cell of interest, as given by the equation

$$\rho = \sum_{i=1}^{n} \frac{1}{A(i)}$$

where $A(i)$ is the area of the triangle i, possessing a vertex in the center of the nucleus of the cell of interest (Figure 3A). Consistent with previous qualitative descriptions (Wagstaff *et al.*, 2016), the mean local density of noninduced scribble[kd] (tet–) and MDCK[WT] cells increases with time in pure populations, whereas the density of pure populations of induced scribble[kd] fluctuates around the initial density value (Figure 3B). After 80 h, the mean density of noninduced scribble[kd] and MDCK[WT] cells reaches a plateau (the homeostatic density) several fold higher (3.6- and 3-fold, respectively) than the density of induced scribble[kd] (Figure 3B). Thus, induced scribble[kd] cells possess a lower homeostatic density than MDCK[WT] or noninduced scribble[kd] cells. Strikingly, when induced scribble[kd] cells are placed in competition with MDCK[WT] cells, the temporal evolution of their density changes dramatically, augmenting fourfold over 80 h and reaching a final density 1.4-fold higher than the average density of surrounding MDCK[WT] cells (Figure 3C). In contrast, when MDCK[WT] cells and noninduced scribble[kd] cells are placed in competitive conditions, they follow behaviors similar to those observed in pure populations (Supplemental Figure S1A). Overall, our time-resolved mean local density measurements are consistent with the global density analysis previously performed by means of single–time-point quantification (Wagstaff *et al.*, 2016).

Having validated the efficacy of our approach to estimate density heterogeneity in a competing coculture, we experimentally determined the dependency of proliferation and apoptosis probability on local cellular density. To do this, we discretized the local densities for each cell ID at each time frame into 20 bins; the middle bin corresponds to the mean local density of all cells, and the first and last bins correspond to the minimum and maximum local density measured across the population, respectively. The probability of cells undergoing mitosis/apoptosis per cell per frame was then calculated for each bin as:

$$p^{\text{div/apo}}(\rho) = \frac{\sum_{\rho} f(\text{div/apo})}{\sum_{\rho} f}$$

with $f(\text{div/apo})$ being the number of observed events (division or apoptosis) in each bin and f being the total number of observations in that bin. The net growth per cell per frame $p^{\text{net growth}}$ is then defined as

$$p^{\text{net growth}}(\rho) = p^{\text{div}}(\rho) - p^{\text{apo}}(\rho)$$

In pure populations, MDCK[WT] cells showed a high baseline probability of division that decreased when density reached $\rho \sim 10^{-2}$ μm^{-2}, whereas this remained approximately constant at a fourfold smaller level for scribble[kd] cells (Figure 3D). Probability of apoptosis for both cell types was similar, ~10-fold smaller than the probability of division and increasing with density (Figure 3E). Overall, this resulted in net growths that were positive and relatively insensitive to density until $\rho \sim 10^{-2}$ μm^{-2}, after which they decreased (Figure 3F). These data suggest that homeostatic density is set differently for each cell type: in MDCK[WT] cells, it is controlled by a density-dependent decrease in proliferation; while in scribble[kd] cells, it results from an increase in apoptosis with density.

As hypothesized previously (Levayer *et al.*, 2015; Wagstaff *et al.*, 2016), we found that the probability of apoptosis p^{apo} of scribble[kd] cells increased sharply with density for densities $\rho > 4 \times 10^{-3}$ μm^{-2} when placed in contact with MDCK[WT] cells (Figure 3, E and H). For MDCK[WT] cells, p^{apo} was comparable to that observed in pure populations until $\rho \sim 10^{-2}$ μm^{-2}, after which it increased sharply. For high densities ($\rho \geq 10^{-2}$ μm^{-2}), the probability of apoptosis for scribble[kd] was ~2.5-fold higher than for MDCK[WT] cells. At the lowest densities ($\rho > 3 \times 10^{-3}$ μm^{-2}), the probability of apoptosis for scribble[kd] was similar in competitive and pure populations (compare Figure 3H with Figure 3E). Interestingly, the local density observed in competitions covered a larger range than in pure populations (Figure 3, C and H), perhaps because in pure populations, scribble[kd] cells strive to preserve the low homeostatic density that they prefer. Together, these data directly show that apoptosis is up-regulated with increasing density, as hypothesized previously (Wagstaff *et al.*, 2016).

To date, analysis of competition has mostly focused on increases in the probability of apoptosis. Interestingly, the analysis of division probabilities in competitive conditions revealed clear differences in behavior between scribble[kd] and MDCK[WT]. MDCK[WT] cells behaved as in pure populations, with p^{div} remaining largely insensitive to density before decreasing for $\rho \geq 10^{-2}$ μm^{-2} (Figure 3, D and G). In contrast, scribble[kd] cells showed clear sensitivity to local density with similar p^{div} than in pure populations at low density, but a larger p^{div} for densities $\rho \geq 10^{-2}$ μm^{-2} (Figure 3, D and G). In both pure populations and competitions, $p^{\text{div}}_{\text{WT}}$ was larger than $p^{\text{div}}_{\text{scribble}}$ until $\rho \sim 10^{-2}$ μm^{-2}. Together, the increase in p^{apo} with density and the decrease in p^{div} with density resulted in net growth that decreased sharply for densities $\rho \geq 10^{-2}$ μm^{-2} (Figure 3I). Overall, the MDCK[WT] cells dominate the competition at all densities, with a net growth of 0.2% that drops to 0.025% for the largest densities present in our experiments due to the combination of a decrease in p^{div} and an increase in p^{apo}. Thus, our single-cell analysis emphasizes the importance of considering division as well as apoptosis when examining cell competition.

Single-cell analysis reveals the presence of local neighborhood effect in net growth

The induction of behavior in one cell lineage by contact with another is a central concept in cell competition (Vincent *et al.*, 2013). To detect inductive behaviors during cell competition, we categorized each division and apoptosis as a function of the number of neighbors of each type. To do this, we implemented a
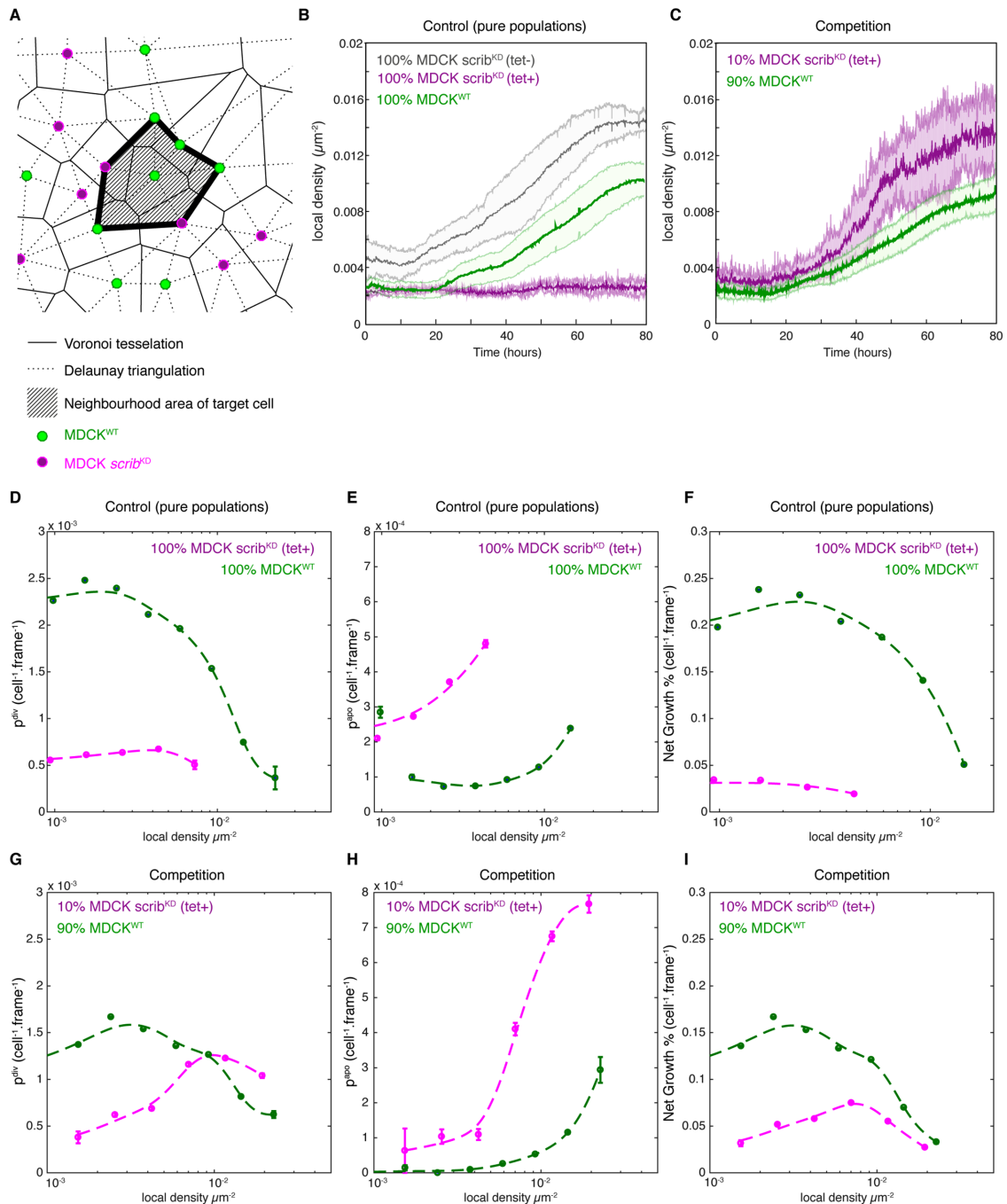
FIGURE 3: Probability of apoptosis and division are sensitive to local density in competitive conditions. (A) Definition of local density. The center of the nucleus of each cell is represented by a filled circle (green: MDCK^{WT}; magenta: scribble^{kd}). Nucleic centers are used to construct a Delaunay triangulation (dashed lines) and a Voronoi tessellation (solid lines). Local density was defined as the sum of inverse areas of the triangles sharing a common vertex at the nucleus of the cell of interest. The area taken into account in this example is shown in gray. (B, C) The local cellular density was computed for each cell ID and averaged among cells of the same lineage at each time point. The solid line indicates the mean of the experiments and the shaded area indicates the SD. Data are pooled from three biological replicates. (B) Temporal evolution of local density for pure populations of MDCK^{WT} (green), noninduced scribble^{kd} (tet−, black), and induced scribble^{kd} (tet+, magenta). (C) Temporal evolution of local density for MDCK^{WT} (green) and scribble^{kd} (magenta) seeded at a 90:10 ratio during a competition assay. (D–F) Probability of division (p^{div}), apoptosis (p^{apo}), and net growth ($p^{div} − p^{apo}$) per cell per frame as function of local density computed for induced scribble^{kd} (tet+, magenta) and MDCK^{WT} pure populations. (G–I) Probability of division, apoptosis, and net growth per cell per frame for MDCK^{WT} and induced scribble^{kd} during competition. Data are pooled from eight fields of view from two biological replicates. (D–I) Data points are indicated by solid circles. Each data point is computed from $N > 500$ observations. Trend lines computed using smoothing splines are plotted as dashed lines. In D–F, the mean local density used to define the bins is 2.6×10^{-3} μm^{-2} for scribble^{kd} cells and 6×10^{-3} μm^{-2} for MDCK^{WT}. In G–I, the mean local density value is 6×10^{-3} μm^{-2} for both cell types. For each data point, whiskers indicate the inverse of the number of observations as an estimate of accuracy in determining the probability.

neighborhood-based distance algorithm to retrieve the interaction network for each cell at each time point. First, we used the localization of centroids to infer the Voronoi diagram (Barber *et al.*, 1996) in each frame (Figure 4A and Supplemental Figures S2 and S3) and verified its accuracy in determining the number of neighbors of each cell (Supplemental Figure S8). For each cell ID, we computed the total number of neighbors and the number belonging to each cell lineage. Such calculations exclude cells closer than 10 µm to the edge of the field of view, because their entire neighborhood cannot be identified. This information enables the generation of "neighborhood plots," wherein the value of a parameter of interest is color-coded and placed in a grid as a function of the number of scribble^{kd} and MDCK^{WT} neighbors, respectively, on the *x*-axis and *y*-axis. The measurement at each grid position is typically computed from > 500 cells and often 10^4–10^5 cells (Supplemental Figure S4). In our diagrams, we annotated with an asterisk grid positions populated by more than 500 observations, but for which no event (e.g., a division or an apoptosis) of interest was detected. Thus, in these positions, we provide an upper bound ($1/N_{observations}$) for the probability of the event.

Next, we employed neighborhood plots to investigate how proliferation, apoptosis, and the resulting net growth depend on the local neighborhood (Figure 4, B–D). A diagram with uniform color indicates a behavior independent on the composition of the cell neighborhood, whereas a diagram showing asymmetry about the diagonal identifies a behavior dependent on neighborhood. To provide a metric for asymmetry, we computed

$$s = \left| \Sigma U - \Sigma L \right|$$

where *U* and *L* are the upper and lower triangular matrices of the neighborhood plot (see *Materials and Methods* and Figure 4E). With this metric, the higher the *s* value, the higher the inductive effect (Figure 4F). The probability of division (p^{div}) of MDCK^{WT} is strongly influenced by neighborhood, being higher in scribble^{kd}-dominated neighborhoods (Figure 4B, top, lower quadrant). In contrast, p^{div} for scribble^{kd} is insensitive to neighborhood (Figure 4B, bottom). Apoptosis diagrams displayed higher symmetry (Figure 4F), with p^{apo} higher in scribble^{kd} cells than MDCK^{WT} cells for most grid positions (Figure 4C). For both cell types, p^{apo} is lower than p^{div} by approximately an order of magnitude (Figure 4, B and C), consistent with Figure 3. As a result, the net-growth neighborhood plots reflect the prevalent contribution of division (Figure 4D). Net growth is positive and highest in the bottom right corner of the neighborhood plot for MDCK^{WT} cells (Figure 4D, top), while scribble^{kd} has either zero or negative net growth in the upper left corner (Figure 4D, bottom), the region dominated by MDCK^{WT}. Similar plots for mixed populations of noninduced scribble^{kd} (tet–) and MDCK^{WT} show less sensitivity to neighborhood, with a much lower degree of asymmetry in apoptosis, division, and net growth (Supplemental Figure S5). In particular, the noninduced scribble^{kd} (tet–) cells have a positive and quite uniform net growth across the entire grid, signifying that their behavior is independent on the composition of the cell neighborhood.

Rate-equation model of density-dependent growth and death quantitatively reproduces competition dynamics
Previous work has suggested that mechanical competition may be the result of cell-autonomous increases in apoptosis with density (Wagstaff *et al.*, 2016). Here, we test this hypothesis by developing a quantitative model based on our experimental findings. To this end, we implemented a coupled rate-equation model to investigate how the density dependence of cell division and apoptosis determine the time evolution of cell count and the overall population fitness. In this model, the density of the MDCK^{WT} (wt) and scribble^{kd} cells (kd) increases at a rate proportional to the density-dependent division rate (f_{div}^{wt}, f_{div}^{kd}), and decreases proportionally with the density-dependent death rate (f_{apo}^{wt}, f_{apo}^{kd}), as given by Eqs. 1 and 2. The rate equations for cell counts N_{wt} and N_{kd} are dependent on the cell density, as given by Eqs. 3 and 4.

$$\frac{\partial \rho_{wt}}{\partial t} = f_{div}^{wt}\left(\rho_{wt} + (1-a)\rho_{kd}\right)\cdot\left(\rho_{wt} + \rho_{kd}\right) - f_{apo}^{wt}(\rho_{wt} + \rho_{kd})\cdot\rho_{wt} \qquad (1)$$

$$\frac{\partial \rho_{kd}}{\partial t} = f_{div}^{kd}\left(\rho_{wt} + \rho_{kd}\right)\cdot\left(\rho_{wt} + \rho_{kd}\right) - f_{apo}^{kd}(\rho_{wt} + \rho_{kd})\cdot\rho_{kd} \qquad (2)$$

$$\frac{\partial N_{wt}}{\partial t} = [f_{div}^{wt}\left(\rho_{wt} + (1-a)\rho_{kd}\right) - f_{apo}^{wt}(\rho_{wt} + \rho_{kd})]\cdot N_{wt} \qquad (3)$$

$$\frac{\partial N_{kd}}{\partial t} = [f_{div}^{kd}\left(\rho_{wt} + \rho_{kd}\right) - f_{apo}^{kd}(\rho_{wt} + \rho_{kd})]\cdot N_{kd} \qquad (4)$$

Equations 1 and 2 describe the temporal evolution of local density of MDCK^{WT} and scribble^{kd}, respectively. Equations 3 and 4 describe the temporal evolution of cell count of MDCK^{WT} and scribble^{kd}, respectively. We solve Eqs. 1–4 simultaneously for the four unknowns (ρ_{wt}, ρ_{kd}, N_{wt}, N_{kd}), because the local density ρ and cell count *N* are not trivially related. To describe the density dependence of birth and death rates of MDCK^{WT} and scribble^{kd} cells, we fit logistic curves to the experimental data in Figure 3, G and H (Supplemental Figure S6, C, D, and F). The analytical form for the division rates of scribble^{kd} (f_{div}^{kd}) and MDCK^{WT} (f_{div}^{wt}) are determined by fitting a Gaussian and a logistic function, respectively, to the experimental data in Figure 3G (Supplemental Figure S6, A, B, and F). The resultant input functions to the model are shown in Figure 5A. Based on our experimental finding (Figure 4, B and C) that the MDCK^{WT} cells exhibit an asymmetric neighborhood dependence of division rate, we introduce a parameter, *a*, describing the degree of asymmetric dependence of MDCK^{WT} division rate on the densities of MDCK^{WT} and scribble^{kd} cells. The coupled Eqs. 1–4 are solved numerically to reproduce the temporal evolution of cell counts of MDCK^{WT} and scribble^{kd} cells (Figure 5, B and C, and Supplemental Figure S6E), subject to the experimental initial conditions for the cell count and density of the two cell types.

We hypothesize three different models for competitive interaction between the MDCK^{WT} and the scribble^{kd} cells, modulated by the asymmetry parameter, *a*. First, we consider an uncoupled model (*a* = 1), in which the growth rates of MDCK^{WT} and scribble^{kd} cells are independent of each other. Second, we examine a symmetric interaction model (*a* = 0), in which the evolution of cell density and count are equally affected by the densities of the MDCK^{WT} and scribble^{kd} cells. Finally, we investigate an asymmetric interaction model, with *a* > 0 treated as a fitting parameter, such that the division rate of MDCK^{WT} cells are enhanced in the presence of scribble^{kd} neighbors (*a* > 1).

Both the symmetric model (Figure 5B) and the uncoupled model (Supplemental Figure S6E) fail to quantitatively reproduce the experimental cell count, predicting a lower count of MDCK^{WT} cells. By contrast, the asymmetric model (*a* = 2.6) quantitatively reproduces the experimentally observed cell counts (Figure 5C) by accounting for the enhancement of division rate of MDCK^{WT} when in neighborhoods with high proportions of scribble^{kd} cells (Figure 4, B and F). This result reinforces our hypothesis that the MDCK^{WT} cells exhibit division induction dependent on local neighborhood. The model further predicts how the net growth of the competing cell lines varies with the densities of MDCK^{WT} and scribble^{kd} cells (Figure 5, D and E). First, the net growth of the MDCK^{WT} is larger than that of the scribble^{kd} cells, as seen experimentally (Figure 4D). Second, in
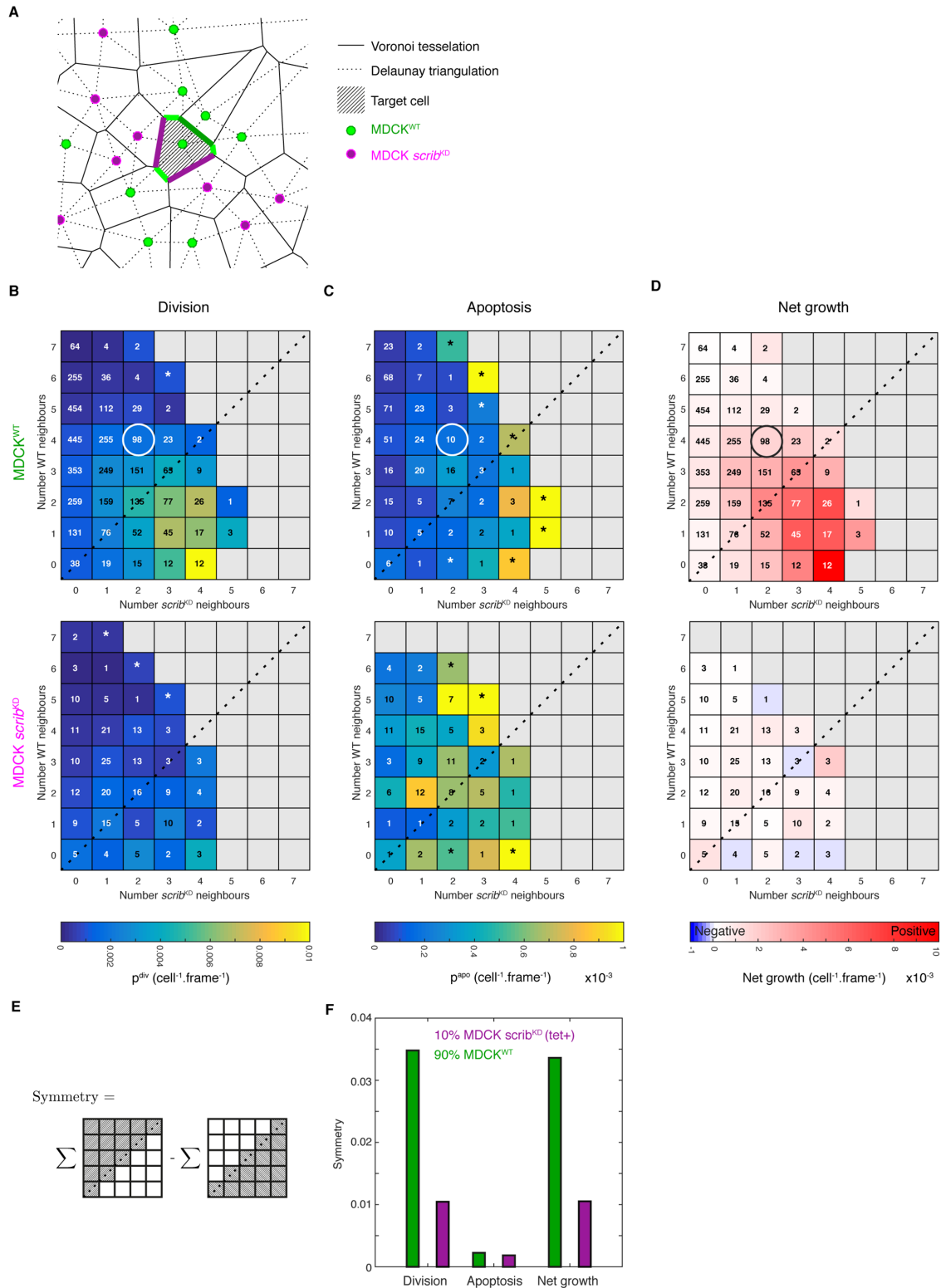
FIGURE 4: Probability of apoptosis, division, and net growth are sensitive to the composition of local neighborhood. (A) Definition of local neighborhood. The center of the nucleus of each cell is represented by a filled circle (green: MDCK^{WT}; magenta: scribble^{kd}). Nucleic centers are used to construct a Voronoi tessellation (solid lines). To determine the composition of the local neighborhood, we determine how many intercellular junctions the cell of interest (gray) has in common with the MDCK^{WT} cells (junctions shown in green) or scribble^{kd} cells (junction shown in magenta). The position of the cell of interest depicted here is shown by the grid position circled with the white frame in B in the neighborhood plot for MDCK^{WT} cells. (B–D) Neighborhood plots showing the color-coded probability per cell per frame of division, apoptosis, and net growth for MDCK^{WT} (top) and scribble^{kd} (bottom) during competition. The number of scribble^{kd} neighbors is shown on the x-axis and the number of MDCK^{WT} cells is shown on the y-axis. The diagonal (black dashed line) indicates grid positions with equal numbers of MDCK^{WT} and scribble^{kd} neighbors. Numbers in each grid
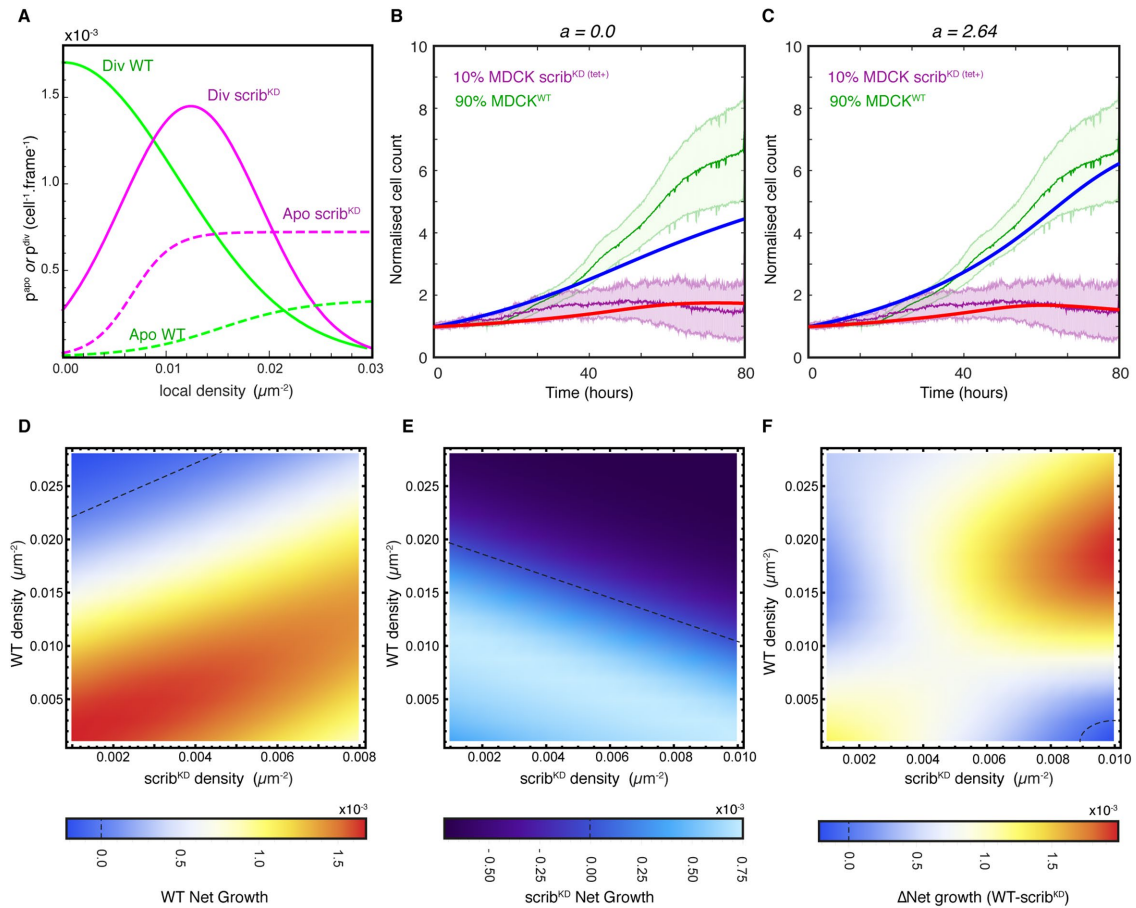
FIGURE 5: A density-dependent rate-equation model for cell competition. (A) Probability (per unit time per cell) of apoptosis (dashed lines) and probability of division (solid lines) for MDCK^{WT} (green lines) and scribble^{kd} (magenta lines) cells determine the rates f_{apo} and f_{div} for our model. These functions were fit to the experimental data in Figure 3, G–I, and Supplemental Figure S5, A–D. (B) Temporal evolution of cell count predicted by the symmetric interaction model ($a = 0.0$) initialized with the mean experimental cell count at $t = 0$ for MDCK^{WT} (solid blue line) and scribble^{kd} (solid red line). The model curves are overlaid with the experimental cell count from Figure 2D for MDCK^{WT} (green) and scribble^{kd} cells (magenta). (C) Temporal evolution of cell count predicted by the asymmetric interaction model ($a = 2.64$) initialized with the mean experimental cell count at $t = 0$ for MDCK^{WT} (solid blue line) and scribble^{kd} (solid red line). (D, E) Heat maps of net growth as a function of local density of MDCK^{WT} cells on the *y*-axis and local density of scribble^{kd} cells on the *x*-axis for the asymmetric interaction model for MDCK^{WT} cells ($a = 2.64$, D) and scribble^{kd} cells ($a = 2.64$, E). Warm colors indicate high net growths, while cold colors indicate low net growths. Dashed line indicates the contour of zero net growth, or density homeostasis. (F) Relative fitness landscape for the asymmetric density-dependent model ($a = 2.64$), defined as the net growth of the MDCK^{WT} cells minus the net growth of the scribble^{kd} cells. MDCK^{WT} cells have a higher fitness than scribble^{kd} cells, except within the region delineated by the dashed line.

agreement with the experimental data (Figure 4D, bottom), the net growth of scribble^{kd} cells exhibit a symmetric dependence on ρ_{wt} and ρ_{kd} (Figure 5E). Finally, the relative fitness (net growth of the MDCK^{WT} – net growth of the scribble^{kd}) heat map (Figure 5F) predicts that the net growth of the MDCK^{WT} cells is always larger than

scribble^{kd} cells, except in a small region with high density of scribble^{kd} cells and low density of MDCK^{WT} cells (Figure 5F, dashed line). This prediction quantitatively delineates the influence of asymmetric density dependence of MDCK^{WT} division rate on the overall fitness landscape.

position indicate the number of detected events (division/apoptosis). The total number of observations for each grid position is shown in Supplemental Figure S4. Measurements for each grid position are typically computed from observations of >500 cells. The probability for each grid position is defined as $p = N_{events}/N_{observations}$. Grid positions for which many observations were made but no event detected are marked by an asterisk and colored as $1/N_{observations}$ to provide an upper bound for the probability in that position. Data are pooled from 12 time-lapse movies from three biological replicates. (E) Definition of *s* as parameter used for calculating the symmetry of neighborhood plots around the diagonal. For each neighborhood plot, *s* is computed as difference between the sum of the lower quadrant and the sum of the upper quadrant for grid positions ranging from [0,0] to [4,4] inclusive. (F) Symmetry calculation performed on checkerboard plot shown in B–D. The symmetry defines whether the behavior is cell autonomous (low *s* values) or dependent on the cell type of neighbors (high *s* values).

DISCUSSION

Competition is a process during which two (or more) cell types interact and whose outcome is the elimination of the less fit cells. Previous work has shown that competition can arise through biochemical induction via intercellular contact or through different tolerances to cell density (Levayer and Moreno, 2013; Vincent et al., 2013). Furthermore, as competition typically takes place over several days, population shifts may result from changes in division rate as well as apoptosis rate. Thus, quantitatively characterizing cell competition necessitates high-throughput automated analysis strategies to mine long-duration time lapses of cell interactions.

Here, we describe a new high-throughput analysis pipeline for characterizing the single-cell behaviors giving rise to population shifts during cell competition. We developed a low-cost time-lapse acquisition system for imaging cells over long durations (up to 80 h) that we coupled with an image-analysis pipeline that tracks cells, automatically annotates cell cycle state, and generates lineage trees for each cell at each time point. Next, we designed tools to investigate how key parameters in competition, such as the probabilities of apoptosis and division, are affected by the local cellular density and the composition of the local cellular neighborhood. To benchmark our analysis, we examined the interaction between MDCK[WT] cells and scribble[kd] cells, as previous work has highlighted it as an example of mechanical competition resulting from differential sensitivity of the cell lines to cell density (Wagstaff et al., 2016).

Previous work showed that scribble[kd] cells are more sensitive to cell density leading to apoptosis and suggested that compaction caused by the MDCK[WT] cells leads to their eventual elimination (Wagstaff et al., 2016). We found that, when scribble[kd] cells are placed in the presence of MDCK[WT] cells, their local density increases threefold compared with when they are in pure populations. Using our automated analysis pipeline, we quantitatively characterized the dependency of the probability of apoptosis and division as a function of density. Strikingly, interaction with MDCK[WT] cells causes the probability of apoptosis of scribble[kd] cells to increase sharply at higher densities. As the probability of division is approximately one order of magnitude larger than the probability of apoptosis, any effect on division will be the dominant effect on competition. For instance, at high density ($\sim 10^{-2}$ µm^{-2}), the division rate for scribble[kd] cells is higher than the division rate of MDCK[WT] cells and higher than in pure populations of scribble[kd] cells. Despite this, the net growth of scribble[kd] cells at high density remains lower than that of MDCK[WT] because of a concomitant increase in the probability of apoptosis of scribble[kd] cells. Therefore, in the range of density explored in our experiments, there is no regime where scribble[kd] cells have higher net growth than MDCK[WT] cells. These competition-specific changes were intriguing, because they did not fit in a framework in which cell density is the main predictor of competition outcome and suggested that other factors may participate.

We addressed this question by using our single cell–analysis approach to investigate the impact of the local neighborhood makeup on population dynamics. For this, we generated neighborhood plots, which display how the probability of apoptosis or division depends on the number and cell type of neighbors. Our neighborhood plots suggest that apoptosis is increased in scribble[kd] cells possessing many neighbors, consistent with the notion that apoptosis increases at high density. Our metric for asymmetry revealed that apoptosis in scribble[kd] cells was more sensitive to neighborhood than apoptosis in MDCK[WT] cells. However, the most striking neighborhood dependence was revealed in neighborhood plots of division in MDCK[WT] cells. Interestingly, we found that the probability of division is significantly higher for MDCK[WT] cells in a neighborhood mostly populated by scribble[kd] cells. Thus, proliferation also seems to be strongly affected by local cellular neighborhood and, surprisingly, this is the case in the winner cell type. These data suggest that some inductive cell behavior may be at play in this competition, something that is a well-known marker of cell competition in its traditional definition (Vincent et al., 2013). In addition, our neighborhood analysis can be extended to include a temporal aspect, so that changes in competition at high or low cell density can be assessed, for example (Supplemental Figure S7). Time-resolved neighborhood plots may also enable comparison of competition before and after drug treatment.

To explore the dependence of apoptosis and division on neighbor cell type evident from our experimental data, we introduced a coupled rate-equation model for the evolution of cell count, wherein rates of division and apoptosis depended on cell density. We found that the scenario that best simulated the experimental cell counts assumed an asymmetric dependency of the division rate of MDCK[WT] on the density of scribble[kd] cells. There is a clear difference when we compare this scenario with the symmetric interaction model, which underestimates the temporal evolution of MDCK[WT] cell count.

Interestingly, our numerical simulation shows that scribble[kd] cells may outcompete MDCK[WT] cells in a region of high scribble[kd] density and low MDCK[WT] density (bottom right hand corner of Figure 5F). Such a regime was never observed in our experimental conditions and would require external manipulation to be applied.

Initial seeding density is a key parameter in the competition phenomenon described here. All of our experiments were performed with an initial density of 10^{-3} cells/µm^2 and a 90:10 ratio of MDCK[WT] and scribble[kd] cells. Future experiments will be necessary to assess how the competition outcome depends on initial seeding density and seeding ratios of the competing lines.

One intriguing question arising from our analysis is whether increased division of MDCK[WT] cells in majority scribble[kd] neighborhoods is cause or consequence of increased apoptosis in scribble[kd] cells. Further experimental work will be needed to understand the molecular mechanisms underlying the sensitivity of mitotic behavior to density and neighborhood and provide a dynamic characterization of the molecular changes occurring at the interface between cell types.

Our quantitative analysis of competition has suggested original hypotheses underlying the eventual elimination of loser cells from the tissue and emphasizes the need to examine how the probability of division changes with density and neighborhood. In addition to competition, our pipeline and characterization tools are broadly applicable to any interaction between cell types leading to outcomes such as division, death, or differentiation in processes such as cancer, stem cell biology, and development.

MATERIALS AND METHODS
Cell culture

The MDCK cell lines used for this study (MDCK[WT] and pTR scribble short hairpin RNA [shRNA], scribble[kd]) were generated as described by Norman et al. (2012). All cell lines used in this publication have been tested for mycoplasmal infection and were found to be negative (MycoAlert Plus Detection Kit, Lonza, LT07-710).

MDCK cells were grown in DMEM (Thermo-Fisher) supplemented with 10% fetal bovine serum (Sigma-Aldrich), HEPES buffer (Sigma-Aldrich), and 1% penicillin/streptomycin in a humidified incubator at 37°C with 5% CO$_2$. The scribble[kd] cells were cultured as WT cells, except that we used tetracycline-free bovine serum (Clontech, 631106) to supplement the culture medium. For inducing expression of scribble shRNA, doxycycline (Sigma-Aldrich, D9891-1G) was added to the medium at a final concentration of 1 µg/ml.

To enable visualization of nucleic acid organization during the cell cycle, we established cell lines stably expressing fluorescently tagged histone markers. Use of different fluorescent proteins enabled us to distinguish the two competing cell types and allowed for accurate segmentation. To do this, we transduced MDCK^{WT} cells with lentiviruses encoding H2b-GFP (Addgene; plasmid #25999) and the scribble^{kd} cells with lentiviruses encoding H2B-RFP (Addgene; plasmid #26001). After transduction, cells were sorted using fluorescence-activated flow cytometry based on GFP or RFP fluorescence to yield populations with homogeneous levels of fluorescence.

Western blotting

We performed Western blotting on MDCK H2b-GFP cells, noninduced MDCK scribble^{kd} H2b-RFP (tet–), and induced MDCK scribble^{kd} H2b-RFP cells. Induction of scribble shRNA was carried out as previously described (Norman et al., 2012). Briefly, cells were induced with 1 µg/ml doxycycline for 70 h before lysis. For preparation of protein extracts, cells were placed on ice and washed with cold phosphate-buffered saline (PBS). After removal of PBS, the cells were lysed using RIPA lysis buffer (Santa Cruz Biotechnology) to which protease and phosphatase inhibitors were added at appropriate concentrations. The lysates were clarified by centrifugation at $8000 \times g$ for 4 min at 4°C, diluted 1:1 with 2× Laemmli buffer (Sigma-Aldrich), denatured for 5 min at 95°C, and loaded onto NuPage 4–20% gradient gels (Bio-Rad). For immunoblotting, we used goat anti-Scribble primary antibody (1:500; Santa Cruz; sc-11048) and mouse anti-GAPDH (1:1000; Novus Biologicals; NB300-221) as loading control. For secondary antibodies, we used horseradish peroxidase (HRP)-coupled anti-mouse (GE Healthcare; NXA931) and HRP-coupled anti-goat (Abcam; ab 97110). All HRP-coupled secondary antibodies were used at 1:10,000 dilution. Protein bands were visualized using the ECL Detection kit (GE Healthcare).

Wide-field microscopy

A custom-built automated epifluorescence microscope was built inside a standard CO$_2$ incubator (Thermo Scientific Heraeus BL20) that maintained the temperature at 37°C and in a 5% CO$_2$ atmosphere. The microscope comprised a high-performance motorized stage (Prior Proscan III; H117E2IX), with a motorized focus controller (Prior FB201 and PS3H122R) and a 9.1MP CCD camera (Point Grey; GS3-U3-91S6M). Bright-field illumination was provided using a green LED (Thorlabs M520L3; 520 nm). Fluorescence illumination in two channels, GFP and mCherry/RFP, was via a blue (Thorlabs; M470L3, 470 nm) or yellow (Thorlabs; M565L3, 565 nm) LED, respectively. These were combined using a dichroic beamsplitter (Semrock) and focused onto the back focal plane of a 20× air objective (Olympus 20×/0.4 NA) in an epifluorescence configuration. The camera and the LEDs were synchronized using TTL pulses from an external D/A converter (Data Translation; DT9834). A custom-built humidified chamber maintained the humidity within the sample chamber and was fitted with a thermocouple and humidity sensor to continuously monitor the environment. The microscope setup was controlled via custom-written software in Python and C++.

Long-term live imaging and competition assay

Cell competition assays were carried out in 35-mm glass-bottom Petri dishes (WillCo). At the start of each experiment, cells were seeded at an initial density of 1×10^{-3} cells µm^{-2}. MDCK^{WT} cells expressing H2b-GFP were mixed with scribble^{kd} H2b-RFP cells at a ratio of 90:10. In some experiments, the expression of scribble shRNA had been induced in scribble^{kd} cells by exposure to doxycycline for 70 h

before seeding. In other experiments, the cells were maintained in tetracycline-free medium to prevent scribble shRNA induction.

Imaging was started 2–3 h after seeding. Imaging medium used during the assay was phenol red–free DMEM (Thermo Fisher Scientific; 31053) supplemented with tetracycline-free bovine serum, HEPES, antibiotics, and, for experiments involving induction, doxycycline at the dose indicated earlier. Multilocation imaging was performed inside the incubator scope for duration of 50–80 h. Bright-field, GFP, and RFP fluorescence images were acquired with a frequency of 1 frame every 4 min for each position.

Image processing and cell tracking

After having acquired time-lapse movies of cells using the incubator microscope, we segmented the images into foreground (cells) and background. Several preprocessing steps were performed to restore the images. Flat-field illumination correction was performed and CCD "hot pixels" were removed.

Following image restoration, segmentation of the fluorescence images was performed using a Gaussian mixture model (GMM). Briefly, the combined intensity histogram of three images taken from the beginning, middle, and end of the movie were fitted to a GMM using the expectation maximization algorithm to learn the appropriate parameters (Xu and Jordan, 1996). The intensity distribution was described as a weighted sum of n normal distributions:

$$P(x|\theta) = \sum_{k=1}^{n} \lambda_k N_k \left(\mu_k, \sigma_k^2 \right)$$

where θ represents the learned parameters for the N models: λ_k is the normalized weight, μ_k the mean intensity, and σ_k^2 the variance for each normal distribution in the mixture model (Prince, 2012). We typically used $n = 3$, and separate parameters were learned for the GFP and RFP fluorescence movies. In general, when ordered by increasing μ_k, the three normal distributions reflect the intensity distributions of background, interphase, and mitotic/apoptotic cells. The output of the segmentation method is a binary classification of the image into background and cells. Dense regions of cells were separated using either a marker-controlled watershed transform, a custom-written object-splitting algorithm based on calculating regions of concavity in convex objects (Wienert et al., 2012), or a hybrid of both methods.

Next, we used an additional merging step to recombine fragments arising from oversegmentation of nuclei with a weak fluorescence signal. The algorithm attempts to find the best possible hypothesis for merging the objects, based on separation distance and image features. This works in several phases. First, a Delaunay graph is calculated to make putative clusters of fragments. Second, hypotheses for combinations of fragments constituting a single object are constructed. Each hypothesis has an equal prior probability. Third, successive Bayesian updates are performed using the separation distance and image feature information. Finally, the algorithm selects the merging hypotheses with the highest posterior probabilities. This algorithm has the advantage of not merging apoptotic fragments with nonapoptotic nuclei.

Once the objects are segmented and split/merged, each nuclear marker in the original image is classified according to its position in the cell cycle. We used the following classes for simplicity: interphase, pro(meta)phase, metaphase, anaphase/telophase, and apoptotic. We trained a deep convolutional neural network (CNN) to perform object classification.

Our CNN architecture was broadly based on the LeNet-5 architecture (Lecun et al., 1998) and consisted of several layers of 3×3 convolution, rectified linear units (ReLU) (Nair and Hinton, 2010) and 2×2 max-pooling units (Scherer et al., 2010), which decrease spatial

dimensionality and increase the number of filters. These layers are followed by several fully connected layers, which reduce the output to a one-dimensional tensor representing the five mutually exclusive cell cycle classes. A final Softmax layer (normalized exponential function) returns the output probabilities for each class.

To train the deep CNN, we generated three data sets: 1) training, 2) test, and 3) validation. Each set has an identical number of training examples that are shuffled, class balanced, and augmented (rotations, noise, translations) to yield a large number of training examples. CNNs were implemented in Caffe (Jia *et al.*, 2014) or TensorFlow (Abadi *et al.*, 2016). Training was performed using a momentum optimizer with an exponentially decaying learning rate until convergence. We measured the accuracy of the CNN classification by calculating a confusion matrix that compares a ground truth based on human operator classification and the CNN prediction using the validation data set. Following the training steps, the CNN achieved an overall accuracy of >99% (Figure 1F).

We also trained a nonlinear support vector machine (SVM) with a radial basis function using image features such as fluorescence intensity, intensity gradient, histogram of oriented gradients (HoG) features (Dalal and Triggs, 2005), orientation, eccentricity, and texture. Although the SVM performed well at cell cycle state classification, it did not match the performance of the CNN, particularly with apoptosis detection, with a maximum accuracy of ~80% (unpublished data). We used the CNN for all further data analysis.

Next, classified and segmented objects were assembled into tracks. The tracking algorithm assembles reliable sections of track that do not contain cell division events (tracklets). Each new tracklet initiates a probabilistic model in the form of a Kalman filter (Kalman, 1960) and uses this to predict future states (and error in states) of each of the objects in the field of view. We assigned new observations to the growing tracklets (linking) by evaluating the posterior probability of each potential linkage from a Bayesian belief matrix for all possible linkages (Narayana and Haverkamp, 2007). The best linkages are those with the highest posterior probability. Despite the high instantaneous accuracy of the CNN classification, occasional errors occur. We corrected errors using a temporal model of the cell cycle (Held *et al.*, 2010) implemented as a hidden Markov model (HMM) comprising interphase, the three states of mitosis, and a dead-end state of apoptosis. Any tracklets containing a metaphase to anaphase transition are split into separate tracks so that they can be labeled as division events in later steps of the algorithm.

The tracklets are then assembled into lineage trees by using multiple hypothesis testing and integer programming (Al-Kofahi *et al.*, 2006; Bise *et al.*, 2011) to identify a globally optimal solution. We built upon this previous work to incorporate hypotheses specific to apoptosis/extrusion and used additional geometric features and CNN classifications in the hypothesis generation. The following hypotheses were generated: 1) true positive track, 2) false positive track, 3) initializing at the beginning of the movie or near the edge of the FOV, 4) termination at the end of the movie or near the edge of the FOV, 5) a merge between two tracklets, 6) a division event, or 7) an apoptotic event. The likelihood of each hypothesis was calculated for some or all of the tracklets based on heuristics. The global solution identified a sequence of high-likelihood hypotheses that accounted for all observations. The global solution having been identified, the fates of individual cells were updated, tracks were merged and lineage trees were generated using a breadth-first search to traverse the trees.

At the end of the image-processing and tracking steps, the xyt-position, cell cycle state, and lineage of each cell in the field of view has been determined.

All code was implemented in Python and C/C++ using CVXOPT, GLPK, Numpy, Scipy, TensorFlow, and Caffe libraries. All image processing was performed on a Dell Precision workstation running Ubuntu 16.04LTS with 32 Gb RAM and a NVIDIA GTX1080 GPU. Computational time was on the order of minutes to hours depending on the complexity of the data.

Postprocessing analysis

The output of the tracking software is a table containing a time-resolved list of unique cell IDs; for each cell ID, the software saves the centroid coordinates, the assigned cell cycle state, and lineage information (mother ID).

To investigate the role of local neighborhood in cell competition, we implemented a neighborhood-based distance algorithm to retrieve the cellular interaction network (Figure 3A). A custom-written MATLAB (MathWorks) script was created to calculate the Voronoi diagram (Barber *et al.*, 1996) using the known localization of cell centroids in each frame. The distance between Voronoi cells was computed (Supplemental Figure S2) and compared with a threshold value (D_{thresh}). We calculated the mean value of internuclear separation over time, determining D_{thresh} to be 30 μm for MDCK^{WT} and 60 μm for induced MDCK *scribble*^{kd}. Among the neighboring Voronoi cells centered at a distance below D_{thresh}, we defined true neighbors as those cells sharing one common vertex with the target cell. This definition is important to detect and remove neighbors situated too far from the cell of interest to truly interact with it. Thus, for each cell ID, we could compute the number of neighbors and the fraction of neighbors belonging to each cell lineage. This allowed investigation of the dependency of proliferation and apoptosis on local neighborhood (Figure 3, B and C).

To investigate how apoptosis and division depended on local cell density, we implemented a custom-written MATLAB script to compute a local cellular density measurement based on the Delaunay triangulation of nucleic centers of mass in each image frame (Figure 3A). We defined the local cellular density as the sum of inverse areas of the triangles that share a common vertex with the cell of interest, and as given by the equation

$$\rho = \sum_{i=1}^{n} \frac{1}{A(i)}$$

where $A(i)$ is the area of the triangle i sharing a vertex with the target cell. Local cell density (ρ) was computed for each cell ID and averaged among cells of the same lineage at each time point. This average was then plotted as a function of time for each cell type separately in mixed populations.

Error bars correspond to the inverse of the number of observations for each data point. For generating the neighborhood diagrams shown in Figure 4, we implemented a custom-written MATLAB script to compute the interaction network for each cell at each time point, based on a Voronoi tessellation. For each cell ID, we computed the total number of neighbors and the number belonging to each cell lineage, MDCK^{WT}, or scribble^{kd}. We categorized each division and apoptosis event as a function of the number of neighbors of each type. We color-coded the probability of division, apoptosis, and net growth and display these parameters as function of neighborhood by placing them in a grid, where the *x*-axis and *y*-axis represent the number of scribble^{kd} and MDCK^{WT} neighbors, respectively. The measurement within each grid position was computed from $N_{observations} > 500$ cells. We annotated with an asterisk grid positions populated by more than 500 observations, but for which no event (e.g., a division or an apoptosis) of interest

was detected. In these positions, we provide an upper bound ($1/N_{observations}$) for the probability of the event.

Numerical simulations

The experimental data for the density dependence of the probability of division and apoptosis were fit with analytic functions. These functions correspond to the density-dependent division and apoptotic rate for both respective cell types used in the model (f_{div}^{wt}, f_{apo}^{wt}, f_{div}^{kd}, f_{apo}^{kd}). The mathematical form of the fitting functions was chosen to minimize the number of parameters (Figures 5A and Supplemental Figure S5, A–D) while providing an accurate fit that converges at extremities. For the density-dependent apoptotic rate of MDCK^{WT} and density-dependent apoptotic rate of scribble^{kd}, logistic functions satisfied these criteria. For the density-dependent division rate of wild-type and scribble^{kd}, Gaussian functions best described the experimental trends (Supplemental Figure S5, A and B). We numerically solved the four coupled rate equations (Eqs. 1–4) for the density of MDCK^{WT}, density of scribble^{kd}, cell count of MDCK^{WT}, and the cell count of scribble^{kd} using Mathematica (Wolfram Research), and plotted the normalized cell counts in Figure 5C. We studied two interacting limits of the rate-equation models, one describing a symmetric interaction of local densities of MDCK^{WT} and scribble^{kd} ($a = 0$) and the other describing asymmetric interaction ($a > 0$ in Eqs. 1–4). We found that the asymmetric model best replicated the experimental findings. For Figure 5, D and E, the heat maps of net growth against density of MDCK^{WT} and density of scribble^{kd} were then plotted by using the built-in Density Plot function in Mathematica, in which the net growth of a cell type is calculated as the density-dependent division rate minus the density-dependent apoptotic rate. The relative fitness is calculated as the net growth of MDCK^{WT} minus the net growth of scribble^{kd}.

Software availability

MATLAB scripts for analysis of cell trajectories are available at https://github.com/quantumjot/CellTracking. The Bayesian tracking library is available at https://github.com/quantumjot/BayesianTracker.

REFERENCES

Abadi M, Agarwal A, Barham P, Brevdo E (2016). Tensorflow: large-scale machine learning on heterogeneous distributed systems. arXiv:1603.04467v2.

Al-Kofahi O, Radke RJ, Goderie SK, Shen Q, Temple S, Roysam B (2006). Automated cell lineage construction—a rapid method to analyze clonal development established with murine neural progenitor cells. Cell Cycle 5, 327–335.

Barber CB, Dobkin DP, Huhdanpaa H (1996). The Quickhull algorithm for convex hulls. ACM T Math Software 22, 469–483.

Bise R, Yin ZZ, Kanade T (2011). Reliable cell tracking by global data association. IEEE International Symposium on Biomedical Imaging: From Nano to Macro, held March 30–April 2, 2011, Chicago, IL, 1004–1010.

Dalal N, Triggs B (2005). Histograms of oriented gradients for human detection. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, held June 20–25, 2005, San Diego, CA.

Fernandez LC, Torres M, Real FX (2016). Somatic mosaicism: on the road to cancer. Nat Rev Cancer 16, 43–55.

Gudipaty SA, Lindblom J, Loftus PD, Redd MJ, Edes K, Davey CF, Krishnegowda V, Rosenblatt J (2017). Mechanical stretch triggers rapid epithelial cell division through Piezo1. Nature 543, 118–121.

Held M, Schmitz MH, Fischer B, Walter T, Neumann B, Olma MH, Peter M, Ellenberg J, Gerlich DW (2010). CellCognition: time-resolved phenotype annotation in high-throughput live cell imaging. Nat Methods 7, 747–754.

Hilsenbeck O, Schwarzfischer M, Skylaki S, Schauberger B, Hoppe PS, Loeffler D, Kokkaliaris KD, Hastreiter S, Skylaki E, Filipczyk A, et al. (2016). Software tools for single-cell tracking and quantification of cellular and molecular properties. Nat Biotechnol 34, 703.

Hogan C, Dupre-Crochet S, Norman M, Kajita M, Zimmermann C, Pelling AE, Piddini E, Baena-Lopez LA, Vincent JP, Itoh Y, et al. (2009). Characterization of the interface between normal and transformed epithelial cells. Nat Cell Biol 11, 460–467.

Jia Y, Shelhamer E, Donahue J, Karayev S, Long J, Girshick R, Guadarrama S, Darrell T (2014). Caffe: convolutional architecture for fast feature embedding. arXiv:1408.5093v1.

Kalman R (1960). A new approach to linear filtering and prediction problems. Trans ASME J Basic Engineering 82(Series D), 35–45.

Lecun Y, Bottou L, Bengio Y, Haffner P (1998). Gradient-based learning applied to document recognition. Proc IEEE 86, 2278–2324.

Levayer R, Hauert B, Moreno E (2015). Cell mixing induced by myc is required for competitive tissue invasion and destruction. Nature 524, 476–480.

Levayer R, Moreno E (2013). Mechanisms of cell competition: themes and variations. J Cell Biol 200, 689–698.

Martins VC, Busch K, Juraeva D, Blum C, Ludwig C, Rasche V, Lasitschka F, Mastitsky SE, Brors B, Hielscher T, et al. (2014). Cell competition is a tumour suppressor mechanism in the thymus. Nature 509, 465–470.

Merino MM, Levayer R, Moreno E (2016). Survival of the fittest: essential roles of cell competition in development, aging, and cancer. Trends Cell Biol 26, 776–788.

Merino MM, Rhiner C, Lopez-Gay JM, Buechel D, Hauert B, Moreno E (2015). Elimination of unfit cells maintains tissue health and prolongs lifespan. Cell 160, 461–476.

Morata G, Ripoll P (1975). Minutes: mutants of Drosophila autonomously affecting cell division rate. Dev Biol 42, 211–221.

Moreno E, Basler K, Morata G (2002). Cells compete for decapentaplegic survival factor to prevent apoptosis in Drosophila wing development. Nature 416, 755–759.

Nair V, Hinton GE (2010). Rectified linear units improve restricted Boltzmann machines. Proceedings of the 27th International Conference on Machine Learning, Haifa, Israel.

Narayana M, Haverkamp D (2007). A Bayesian algorithm for tracking multiple moving objects in outdoor surveillance video. IEEE Conference on Computer Vision and Pattern Recognition, held June 17–22, 2007, Minneapolis, MN.

Navin N, Kendall J, Troge J, Andrews P, Rodgers L, McIndee J, Cook K, Stepansky A, Levy D, Esposito D, et al. (2011). Tumour evolution inferred by single-cell sequencing. Nature 472, 90–94.

Norman M, Wisniewska KA, Lawrenson K, Garcia-Miranda P, Tada M, Kajita M, Mano H, Ishikawa S, Ikegawa M, Shimada T, Fujita Y (2012). Loss of Scribble causes cell competition in mammalian cells. J Cell Sci 125, 59–66.

Oliver ER, Saunders TL, Tarle SA, Glaser T (2004). Ribosomal protein L24 defect in belly spot and tail (Bst), a mouse Minute. Development 131, 3907–3920.

Porazinski S, de Navascues J, Yako Y, Hill W, Jones MR, Maddison R, Fujita Y, Hogan C (2016). EphA2 drives the segregation of Ras-transformed epithelial cells from normal neighbors. Curr Biol 26, 3220–3229.

Prince SJD (2012). Computer Vision: Models, Learning, and Inference, New York: Cambridge University Press.

Puliafito A, Hufnagel L, Neveu P, Streichan S, Sigal A, Kuchnir Fygenson D, Shraiman BI (2012). Collective and single cell behavior in epithelial contact inhibition. Prac Natl Acad Sci USA 109, 739–744.

Rhiner C, Moreno E (2009). Super competition as a possible mechanism to pioneer precancerous fields. Carcinogenesis 30, 723–728.

Scherer D, Muller A, Behnke S (2010). Evaluation of pooling operations in convolutional architectures for object recognition. Lect Notes Comput Sc 6354, 92–101.

Simpson P, Morata G (1981). Differential mitotic rates and patterns of growth in compartments in the *Drosophila* wing. Dev Biol 85, 299–308.

Sottoriva A, Spiteri I, Piccirillo S, Touloumis A, Collins V, Marioni J, Curtis C, Watts C, Tavaré S (2013). Intratumor heterogeneity in human glioblastoma reflects cancer evolutionary dynamics. Proc Natl Acad Sci USA 110, 4009–4014.

Tamori Y, Bialucha CU, Tian AG, Kajita M, Huang YC, Norman M, Harrison N, Poulton J, Ivanovitch K, Disch L, *et al.* (2010). Involvement of Lgl and Mahjong/VprBP in cell competition. PLoS Biol 8, e1000422.

Vincent JP, Fletcher AG, Baena-Lopez LA (2013). Mechanisms and mechanics of cell competition in epithelia. Nat Rev Mol Cell Biol 14, 581–591.

Waclaw B, Bozic I, Pittman M, Hruban R, Vogelstein B, Nowak M (2015). A spatial model predicts that dispersal and cell turnover limit intratumour heterogeneity. Nature 525, 261–264.

Wagstaff L, Goschorska M, Kozyrska K, Duclos G, Kucinski I, Chessel A, Hampton-O'Neil L, Bradshaw CR, Allen GE, Rawlins EL, *et al.* (2016). Mechanical cell competition kills cells via induction of lethal p53 levels. Nat Commun 7, 11373.

Wienert S, Heim D, Saeger K, Stenzinger A, Beil M, Hufnagl P, Dietel M, Denkert C, Klauschen F (2012). Detection and segmentation of cell nuclei in virtual microscopy images: a minimum-model approach. Sci Rep 2, 503.

Xu L, Jordan MI (1996). On convergence properties of the EM algorithm for Gaussian mixtures. Neural Comput 8, 129–151.