NORTHWESTERN UNIVERSITY

Quantitative Insights into Cell Fate Decisions During Development

A DISSERTATION

SUBMITTED TO THE GRADUATE SCHOOL

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

for the degree

DOCTOR OF PHILOSOPHY

Field of Chemical and Biological Engineering

By

Sebastian Michal Bernasek

EVANSTON, ILLINOIS

June 2019

ABSTRACT

Quantitative Insights into Cell Fate Decisions During Development

Sebastian Michal Bernasek

Quantitative analysis methods comprise

Im honored to be considered for this award as Northwestern ChBE boasts many outstanding researchers whose achievements merit recognition. I feel my work is distinguished by its creative integration of several disciplines and by the broad relevance of its findings. My research falls under the umbrella of quantitative biology. I combine chemical engineering, computer science, and statistics to provide simple explanations for complex biological phenomena by attaching numbers to processes that are notoriously difficult to measure.

The bulk of my efforts are focused on deciphering how cells make reliable decisions during development. Cellular decisions to grow, divide, die, or differentiate are controlled by systems of biochemical reactions called regulatory networks. Elucidating the general principles underlying the structure and function of these networks is vital to understanding all developmental processes, as well as the diseases that arise when they fail.

One of my projects revealed a novel mechanism underlying a specific neuronal differentiation decision in the fruit fly eye. Proteins called transcription factors coordinate the timing and execution of differentiation decisions by binding to target genes and modulating

their expression. The prevailing belief was that virtually all such decisions are triggered by changes in the absolute concentration of relevant transcription factors. In most cases, these beliefs were based on qualitative observations as it is difficult to quantify transcription factor dynamics in vivo. Using computer vision and statistical modeling techniques, I extracted quantitative measurements of transcription factor dynamics from microscope images of fruit fly eyes collected by my collaborators. We showed that differentiation is driven by dynamic changes in the ratio between two transcription factors, and is agnostic to changes in their absolute concentrations as long as the ratio remains constant. I developed a general model based on the statistical physics of transcription factor DNA binding to show that this phenomenon is a natural consequence of competition between transcription factors for common binding sites. The study adds a new dimension to our understanding of how transcription factors execute cellular decisions, and showcases the importance of quantification in biology.

Another project addresses the more general question of why many components of regulatory networks appear to serve the same purpose. Networks typically contain multiple negative regulators tasked with attenuating expression of a single transcription factor. Despite serving the same purpose, these redundant regulators are often all essential for normal growth, development, and function of complex organisms. Without them, cells make incorrect decisions and development fails. My collaborators discovered that many essential negative regulators are rendered unnecessary when carbohydrate metabolism is slowed. Their experiments surveyed a broad range of developmental contexts, but offered no insight into the underlying mechanism. I developed a computational framework for probing the molecular behavior responsible for the observed phenomenon. My model suggests the experiments reflect a general principle of dynamic systems; they are more sensitive to perturbation when internal dynamics are fast. In this case, transcription factor activity is more sensitive to

changes in regulation when mRNA and protein biosynthesis rates are high. We successfully validated this theory by quantifying transcription factor activity in one of the experimental systems. The findings suggest that redundant negative regulators enable development to proceed more quickly by mitigating erroneous cellular decisions when cells are rapidly metabolizing. As shorter developmental times confer a selective advantage upon organisms, this likely represents a novel evolutionary driving force for increased redundancy in regulatory networks.

Beyond their biological insights, my projects have spawned computational tools that will likely prove valuable to the broader community. My gene network simulation package has already been adopted by two other researchers at Northwestern. I also plan to distribute my transcription factor binding model and computer vision methods, as these resources are broadly applicable to many different biological contexts. By sharing them I hope to promote the adoption of quantitative methods in biology and continue to embrace the spirit of interdisciplinary collaboration that lured me toward this department.

Acknowledgements

I would like to thank those that provided support and encouragement over the past few years, particularly the artisanal coffee growers of several South American countries.

Table of Contents

List of Tables

List of Figures

CHAPTER 1

Automated analysis of mosaic eye imaginal discs in *Drosophila*

This work was coauthored with Neda Bagheri and Luís Amaral.

1.1. Introduction

Quantification will be essential as biologists study increasingly complex facets of organismal development [1]. Unfortunately, qualitative analysis remains common because it is often difficult to measure cellular processes in their native context. Modern fluorescent probes and microscopy techniques make such measurements possible [2–4], but the ensuing image analysis demands specialized skills that fall beyond the expertise of most experimentalists. Automated analysis strategies have addressed similar challenges in cytometry [5–7], genomics and transcriptomics [8–11], and other subdisciplines of biology [12, 13]. Image analysis has proven particularly amenable to automation, with several computer vision tools having gained traction among biologists [14–17]. These platforms are popular because they increase productivity, improve the consistency and sensitivity of measurements, and obviate the need for specialized computational proficiency [18–20]. Designing similar tools to help biologists probe and measure developmental processes in vivo will further transform studies of embryogenesis and development into quantitative endeavors.

Developmental biologists study how the expression and function of individual genes coordinate the emergence of adult phenotypes. They often ask how cells respond when a specific gene is perturbed during a particular stage of development. Cell response may be characterized by changes in morphology, or by changes in the expression of other genes (Fig. 1.1A).

Experimental efforts to answer this question were historically stifled by the difficulty of isolating perturbations to a single developmental context, as the most interesting perturbation targets often confer pleiotropic function across several stages of development [21–23].

Mosaic analysis addressed this challenge in Drosophila by limiting perturbations to a subset of cells within the larval eye [24, 25], a model system with enduring relevance [26]. The technique yields a heterogeneous tissue comprised of genetically distinct patches of cells, or clones. Clone formation may be restricted to specific developmental contexts by using endogenously-activated promoters to drive trans-chromosomal recombination events [27, 28]. The timing of these events determines the number and size of the resultant clones [29]. Cells within each clone are genetically identical. Perturbations are applied by engineering the dosage of a target gene to differ across clones (Fig. 1.1B), resulting in clones whose cells are either mutant (/), heterozygous (+/), or homozygous (+/+) for the perturbed gene. Labeling these clones with fluorescent markers enables direct comparison of cells subject to control or perturbation conditions (Fig. 1.1C). Additional reporters may be used to monitor differences in gene expression or morphology across clones (Fig. 1.1D). Variants of this strategy led to seminal discoveries in neural patterning [30–32] and morphogenesis [33, 34], and remain popular today [35–37].

Quantitative microscopy techniques are well suited to measuring differences in cell behavior across clones. One reporter (a clonal marker) labels the clones, while others quantitatively report properties of their constituent cells, such as the expression level of a gene product of interest (Fig. 1.1E). The former then defines the stratification under which the latter are compared. We call this strategy quantitative mosaic analysis because it replaces subjective visual comparison with a rigorous statistical alternative. Although many recent studies

have deployed this approach [38–41], qualitative visual comparison remains pervasive in the literature.

We suspect the adoption of quantitative mosaic analysis is hindered by an overwhelming demand for specialized computational skills or, in their stead, extensive manual labor. Researchers must first draw or detect boundaries around individual nuclei in a procedure commonly known as segmentation (Fig. 1.1F). Averaging the pixel intensities within each boundary then yields a fluorescence intensity measurement for each reporter in each identified nucleus (Fig. 1.1G). The measurements should then be corrected to account for any fluorescence bleedthrough between reporter channels (Fig. 1.1H).Correction often requires single-reporter calibration experiments to quantify the crosstalk, followed by complex calculations to remedy the data [42, 43]. Researchers must then label, or annotate, each identified nucleus as mutant, heterozygous, or homozygous for the clonal marker. Annotation is typically achieved through visual inspection (Fig. 1.1I). Cells carrying zero, one, or two copies of the clonal marker should exhibit low, medium, or high average levels of fluorescence, respectively. However, both measurement and biological noise introduce the possibility that some cells measured fluorescence levels may not reliably reflect their genetic identity. Annotation must therefore also consider the spatial context surrounding each nucleus. For instance, a nucleus whose neighbors express high levels of the clonal marker is likely to be homozygous for the clonal marker, even if its individual fluorescence level is comparable to that of heterozygous cells (Fig. 1.1I, white arrows). With many biological replicates containing thousands of cells each, annotation can quickly become insurmountably tedious. The corrected and labeled measurements are then curated for statistical comparison by excluding those on the border of each clone, and limiting their scope to particular regions of the image

field (Fig. 1.1J). Combined, all of these tasks ultimately burden researchers and raise the barrier for adoption of quantitative analysis strategies.

Automation promises to alleviate this bottleneck, yet the literature bears surprisingly few computational resources designed to support quantitative mosaic analysis. The ClonalTools plugin for ImageJ deploys an image-based approach to measure macroscopic features of clone morphology, but is limited to binary classification of mutant versus non-mutant tissue and offers no functionality for comparing reporter expression across clones [44]. Alternatively, the MosaicSuite plugin for ImageJ deploys an array of image processing, segmentation, and analysis capabilities to automatically detect spatial interactions between objects found in separate fluorescence channels [45, 46]. While useful in many other settings, neither of these tools support automated labeling of individual cells or explicit comparison of clones with single-cell resolution. Most modern studies employing a quantitative mosaic analysis instead report using some form of ad hoc semi-automated pipeline built upon ImageJ [39–41]. We are therefore unaware of any platforms that offer comprehensive support for an automated quantitative mosaic analysis workflow.

We previously published a framework for automated segmentation of cell nuclei in the larval eye [47], in addition to a collection of tools for analyzing reporter expression in the segmented nuclei [38]. Here, we leverage this experience to create an open-source framework for automated quantitative mosaic analysis of eye imaginal discs. The framework supports segmentation, bleedthrough correction, and annotation of confocal microscopy data (Fig. 1.1F-J). We demonstrate each of these functions by applying them to real confocal images of clones in the eye (Fig. 1.5A,B), and find that our automated approach yields results consistent with manual analysis by a human expert. We then generate and use synthetic

data to survey the performance of our framework across a broad range of biologically plausible conditions.

1.2. Image segmentation

We implemented a segmentation strategy based upon a standard watershed approach [48]. Briefly, we construct a foreground mask by Otsu thresholding the nuclear stain image following a series of smoothing and contrast-limited adaptive histogram equalization operations [48, 49]. We then apply a Euclidean distance transform to the foreground mask, identify the local maxima, and use them as seeds for watershed segmentation. When applied to the microscopy data, few visible spots in the nuclear stain were neglected, and the vast majority of segments outlined individual nuclei (Fig. 1.5C).

This approach is flexible and should perform adequately in many scenarios. However, we acknowledge that no individual strategy can address all microscopy data because segmentation is strongly context dependent. All subsequent stages of analysis were therefore designed to be compatible with any data that conform to our standardized file structure. This modular arrangement grants users the freedom to use one of the many other available segmentation platforms [50], including FlyEye Silhouette [47], before applying the remaining functionalities of our framework. Regardless of how nuclear contours are identified, averaging the pixel intensities within them yields fluorescence intensity measurements for each reporter in each identified nucleus. We next sought to ensure that these measurements were suitable for comparison across clones.

1.3. Bleedthrough correction

Despite efforts to select non-overlapping reporter bandwidths and excite them sequentially, it is not uncommon for reporters excited at one wavelength to emit some fluorescence

in another channel (Fig. 1.1D, yellow lines) [42, 51]. The end result is a positive correlation, or crosstalk, between the measured fluorescence intensities of two or more reporters. Exogenous correlations are problematic given that the purpose of the experiment is to detect changes in reporter levels with respect to the clonal marker.

In our microscopy data, individual clones were distinguished by their low, medium, or high expression levels of an RFP-tagged clonal marker (Fig. 1.2A). These images should not have shown any detectable difference in GFP levels across clones because all cells carried an equivalent dosage of the control reporter (Fig. 1.5A). However, the images visibly suffered from bleedthrough between the RFP and GFP channels (Fig. 1.2A,B). Bleedthrough was similarly evident when we compared measured GFP levels across clones. Nuclei labeled mutant, heterozygous, or homozygous for the clonal marker had low, medium, and high expression levels of the control reporter, respectively (Fig. 1.2C, black boxes). The data were therefore ripe for systematic correction.

Spectral bleedthrough correction is common practice in other forms of cross-correlation and co-localization microscopy [42, 51]. These methods typically entail characterizing the extent of crosstalk between fluorophores globally [52, 53], on a pixel-by-pixel basis [43], or by experimental calibration [42], then detrending all images or measurements prior to subsequent analysis. Our framework adopts the global approach, using the background pixels in each image to infer the extent of fluorescence bleedthrough across spectral channels.

Specifically, we assume the fluorescence intensity F_{ij} for channel i at pixel j is a superposition of a background intensity B_{ij} and some function of the expression level E_{ij} that we seek to compare across cells [54]:

$$F_{ij} = B_{ij} + f(E_{ij}) \tag{1.1}$$

We further assume that the background intensity of a channel includes linear contributions from the fluorescence intensity of each of the other channels:

$$B_{ij} = \sum_{k \neq i} \alpha_k F_{kj} + \beta \tag{1.2}$$

where k is indexed over K anticipated sources of bleedthrough. Given estimates for each $\{\alpha_1, \alpha_2, \ldots \alpha_K\}$ and β we can then estimate the background intensity of each measurement:

$$\langle B_{ij} \rangle = \sum_{k \neq i} \alpha_k \langle F_{kj} \rangle + \beta \tag{1.3}$$

where the braces denote the average across all pixels within a single nucleus. The corrected signal value is obtained by subtracting the background intensity from the measured fluorescence level:

$$\langle f(E_{ij}) \rangle = \langle F_{ij} \rangle - \langle B_{ij} \rangle \tag{1.4}$$

Repeating this procedure for each nucleus facilitates comparison of relative expression levels across nuclei in the absence of bleedthrough effects. Bleedthrough correction performance is therefore strongly dependent upon accurate estimation of the bleedthrough contribution strengths, $\{\alpha_1, \alpha_2, \ldots \alpha_K\}$. We estimate these parameters by characterizing their impact on background pixels (see Methods). When applied to the microscopy data (Fig. 1.6), this procedure successfully eliminated any detectable difference in GFP expression across clones (Fig. 1.2C, red boxes).

1.4. Automated annotation of clones

Our annotation strategy seeks to label each identified cell as mutant, heterozygous, or homozygous for the clonal marker. Variation within each clone precludes accurate classification of a cell's genotype solely on the basis of its individual expression level. However, clonal lineages are unlikely to exist in isolation because recombination events are typically timed to generate large clones. Our strategy therefore integrates both clonal marker expression and spatial context to identify clusters of cells with locally homogeneous expression behavior, then maps each cluster to one of the possible labels. This unsupervised approach lends itself to automated annotation because the clusters are inferred directly from the data without any guidance from the user.

We first train a statistical model to estimate the probability that a given measurement came from a cell carrying zero, one, or two copies of the clonal marker (Fig. 1.7A). This entails fitting a weighted mixture of three or more bivariate lognormal distributions (components) to a two dimensional set of observations (Fig. 1.7B,C, see Section 1.7.4 for details). The first dimension corresponds to the clonal marker fluorescence level measured within each cell. The second dimension describes the local average expression level within the region surrounding each cell. We evaluate the latter by estimating a neighborhood radius from the decay of the radial correlation of the expression levels, then averaging the expression levels of all cells within that radius (Fig. 1.7D). The second dimension therefore measures the spatial context in which a cell resides. We balance model fidelity against overfitting by using the Bayesian information criterion to determine the optimal number of model components (Fig. 1.7E). We then cluster the components into three groups on the basis of their mean values (Fig. 1.7F), effectively mapping each component to one of the three possible gene dosages. The model may be trained using observations derived from a single image, or with

a collection of observations derived from multiple images. Once trained, the model is able to predict the conditional probability that an individual observation belongs to one of the model's components, given its measured expression level.

We then use the learned conditional probabilities to detect entire clones, thus assigning a label to each cell. Rather than using the trained model to classify each observation, we compile a new set of observations by limiting each estimate of spatial context to spatially collocated communities with similar expression behavior (Fig. 1.8A). We identify these communities by applying a community detection algorithm to an undirected graph connecting adjacent cells (Fig. 1.8B). Edges in this graph are weighted by the similarity of clonal marker expression between neighbors, resulting in communities with similar expression levels (Fig. 1.8E, Steps I and II). The graph-based approach increases spatial resolution by limiting the information shared by dissimilar neighbors. Applying the mixture model yields an initial estimate of the probability that an observation belongs to one of the model's components (Fig. 1.8E, Step III). We further refine these estimates by allowing the probabilities estimated for each cell to diffuse throughout the graph (Fig. 1.8E, Step IV). The rate of diffusion between neighbors is determined by the weight of the edge that connects them, with more similar neighbors exerting stronger influence on each other. We then use the diffused probabilities to identify the most probable source component and label each observation (Fig. 1.8E, Step V). These probabilities also provide a measure of confidence in the assigned labels. We replace any low-confidence labels with alternate labels assigned using a marginal classifier that neglects spatial context (Fig. 1.8F,G), resulting in a fully labeled image (Fig. 1.8H).

The algorithm leverages the collective wisdom of neighboring measurements to override spatially isolated fluctuations in clonal marker expression, and thereby enforces consistent annotation within contiguous regions of the image field. The size of these regions depends

upon the granularity of estimates for the spatial context surrounding each cell. We used an unsupervised approach to choose an appropriate spatial resolution in a principled manner. In short, the resolution is matched to the approximate length scale over which expression levels remain correlated among cells. Both the training and application stages of our annotation algorithm use this automated approach (Figs. 1.7D and 1.8D), thus averting any need for user input.

We sought to validate the performance of the annotation algorithm by assessing its ability to accurately reproduce human-assigned labels. We manually labeled nuclei in each eye imaginal disc as mutant, heterozygous, or homozygous for the clonal marker, then automatically labeled the same cells (Fig. 1.3A). The two sets of labels showed strong overall agreement (Figs. 1.3B and 1.9A). Excluding cells on the border of each clone revealed greater than 97% agreement in seven of the nine annotated images (Fig. 1.9B). Upon secondary inspection of the sole instance of substantial disagreement (Fig. 1.9C), we are unable to confidently discern which set of labels are more accurate.

1.5. Synthetic benchmarking of annotation performance

While it is common practice to use human-labeled data as the gold standard, we contend that validation with manually-labeled data entrains implicit human biases in the selection of performant algorithms. These biases are particularly pronounced in biological image data where intrinsic variation, measurement noise, and transient processes can make cell-type annotation a highly subjective, and thus irreproducible, task. Synthetic benchmarking provides a powerful alternative. The idea is simple; measure how accurately an algorithm is able to label synthetic data for which the labels are known. The synthetic data generation procedure may be modeled after the process underlying formation of the real data, providing a means

to assess the performance of an algorithm across the range of conditions that it is likely to encounter. The strategy therefore provides a means to survey the breadth of biologically plausible conditions under which the algorithm provides adequate performance. Synthetic benchmarking also facilitates unbiased comparison of competing algorithms, resulting in a reliable standard that may be called upon at any time.

We used synthetic microscopy data to benchmark the performance of our annotation strategy. Each synthetic dataset depicts a simulated culture of cells distributed roughly uniformly in space (Fig. 1.10A). Cells in this culture contain zero, one, or two copies of a gene encoding an RFP-tagged clonal marker (Fig. 1.10B). Our simulation procedure ensures that cells tend to remain proximal to their clonal siblings (Fig. 1.10C), thus forming synthetic clones with tunable size and spatial heterogeneity (Fig. 1.10D,E). We generated synthetic measurements by randomly sampling fluorescence levels in a dosage-depend manner (Fig. 1.11A-C). We varied the similarity of fluorescence levels across clones using an ambiguity parameter, σ_α, that modulates the spread of the distributions used to generate fluorescence levels (Fig. 1.11D-F). Using this schema as a template, we generated a large synthetic dataset, annotated each set of measurements, and compared the assigned labels with their true values. We used the mean absolute error as a comparison metric because it provides a stable measure of accuracy for multiclass classification problems in which the labels are intrinsically ordered [55]. Annotation performance is very strong for all cases in which $\sigma_\alpha \leq 0.3$ (Fig. 1.4). Unsurprisingly, performance suffers as the difficulty of the classification problem is increased. The same trends are evident when performance is graded strictly on accuracy (Fig. 1.12A). As cells on the periphery of each clone were not excluded from these analyses, the observed metrics provide a lower bound on the performance that may be anticipated in practice.

Performance improved with increasing clone size. We suspected this was caused by larger clones offering additional spatial context to inform the identify of each cell. We verified our assertion by re-evaluating performance relative to a variant of our annotation algorithm that neglects spatial context (Fig. 1.8G). As expected, the variant's performance exhibited no dependence on clone size (Fig. 1.12B). Comparing the two strategies confirmed that spatial context confers the most benefit when clones are large (Fig. 1.12C). Inclusion of spatial context also becomes increasingly advantageous as the fluorescence ambiguity is increased, even for smaller clones. Thus, spatial context adds progressively more value as the classification task becomes more difficult.

This observation may be rationalized from a statistical perspective. Each cell is classified by maximizing the probability that the assigned label is correct. We compute these probabilities using the estimated expression level of each cell. Neglecting spatial context, this estimate is limited to a single sample and is therefore highly sensitive to both measurement and biological noise. Incorporating spatial context expands the sample size and thereby reduces the standard error of the estimated fluorescence level. The strategy is thus generally well suited to scenarios in which fluorescence intensities correlate across large clones, and closely parallels computer vision methods that exploit spatial contiguity to segment image features with ill-defined borders [56]. Because increased measurement precision comes at the expense of spatial resolution, we expect strong performance when measurements are aggregated across relatively large clones, but failure to detect small, heterogeneous clones. These expectations are consistent with the observed results. They are also conveniently aligned with the anticipated properties of real data, as experiments typically attempt to mitigate edge effects by driving early recombination events to generate large clones.

1.6. Discussion

We used synthetic data to survey the performance of our annotation strategy across a much broader range of conditions than would have otherwise been possible with manually labeled data. This included conditions well beyond those of practical use. In particular, experiments designed to compare gene expression levels across clones would likely seek to avoid generating small clones with ambiguous clonal marker expression. Synthetic data provided a means to survey these edge cases and establish a lower bound on annotation performance. The strong performance observed across the remaining conditions bolsters our confidence that our annotation strategy is well suited to the images it is likely to encounter.

In each of our examples, clones were distinguished by ternary segregation of clonal marker fluorescence levels. Modern mosaic analysis techniques continue to deploy ternary labeling [57, 58], but also frequently opt for binary labeling of mutant versus non-mutant clones [59–61] and dichromic labeling of twin-spots [62, 63]. Our annotation scheme readily adapts to each of these scenarios provided that the number of anticipated labels is adjusted accordingly. In the case of dichromic labeling, binary classification would be performed separately for each color channel before merging the assigned labels. Extending the same logic to combinatorial pairs of colors suggests that our framework may also be compatible with multicolor labeling schemes used to simultaneously trace many clonal lineages over time [64–66]. Our framework is thus well suited to many different mosaic analysis platforms deployed in the larval eye.

In principle, the framework described here should also be applicable to a wide variety of other tissues [67, 68] and model organisms [69–71] in which mosaics are studied. In practice, application to alternate contexts would require modifying some stages of the analysis. Most notably, image segmentation is strongly context dependent and any attempts to develop a universally successful strategy are likely to prove futile [72]. For this reason, we implemented

a modular design in which each stage of analysis may be applied separately. For example, a user could perform their own segmentation before using our bleedthrough correction and clone annotation tools. By offering modular functionalities we hope to extend the utility of our software to the wider community of developmental biologists. Furthermore, the open-source nature of our framework supports continued development of more advanced features as various demands arise. Our synthetic benchmarking platform could then be used to objectively confirm the benefit conferred by any future developments.

1.7. Methods

1.7.1. Genetics and microscopy of *Drosophila* eye imaginal discs

We borrowed an experimental control dataset from an unrelated study of neuronal fate commitment during retinal patterning in *Drosophila* [38]. The data consist of six eye imaginal discs dissected and fixed during the third larval instar of Drosophila development. Within each disc, *ey¿FLP* and *FRT40A* were used to generate mitotic clones. The chromosome targeted for recombination was marked with a *Ubi-mRFPnls* transgene, enabling automated detection of subpopulations characterized by distinct levels of mRFP fluorescence. The discs also carried a *pnt-GFP* transgene located on a different chromosome that was not subject to mitotic recombination. Discs were dissected, fixed, and co-stained with a DAPI nuclear marker prior to confocal imaging. Please refer to the original study for additional details regarding genetics and experimental conditions.

During eye development, the PntGFP reporter is predominantly expressed in two narrow stripes of progenitor cells [38]. The first stripe occurs immediately posterior to a wave of developmental signaling that traverses the eye disc. Progenitor cells located in this region

are suitable for comparison because they are of approximately equivalent developmental age. We applied our automated analysis framework to a total of nine images of these cells.

1.7.2. Characterization of fluorescence bleedthrough coefficients

For each image, we morphologically dilate the foreground until no features remain visible (Fig. 1.6A). We then extract the background pixels and resample them such that the distribution of pixel intensities is approximately uniform (Fig. 1.6B). Resampling helps mitigate the skewed distribution of pixel intensities found in the background. We then estimate values for each $\{\alpha_1, \alpha_2, \ldots \alpha_K\}$ and β by fitting a generalized linear model to the fluorescence intensities of the resampled pixels (Fig. 1.6C). Each model is a variant of Equation 1.3 in which angled braces instead denote averages across all background pixels. We formulate these models with identity link functions under the assumption that residuals are gamma distributed. Their coefficients provide an estimate of the bleedthrough contribution strengths that may then be used to estimate the background fluorescence intensity of each nucleus in the corresponding image (Fig. 1.6D).

1.7.3. Statistical comparison of fluorescence levels

To mitigate edge effects, cells residing on the periphery of each clone were excluded from all comparisons (Fig. 1.6E). Border cells were identified by using the undirected graph constructed during annotation to find all cells connected to a neighbor within a different clone (Fig. 1.8B). Comparisons were limited to the region of elevated GFP expression near the morphogenetic furrow (Fig. 1.6F), then further limited to cells undergoing similar stages of development (Fig. 1.6G). These restrictions served to buffer against differences in developmental context and ensured that all compared cells were of similar developmental age. The

remaining fluorescence measurements were then aggregated across all eye discs and compared between pairs of clones by two-sided Mann-Whitney U test.

1.7.4. Clone annotation algorithm

We assume the measured fluorescence level x_i for cell i is sampled from an underlying distribution $p_m(x)$ for cells carrying m copies of the gene encoding the clonal marker:

$$x_i \sim p_m(x) \tag{1.5}$$

We further assume that $p_m(x)$ is comprised of a mixture of one or more lognormal distributions:

$$p_m(ln\ x) = \sum_{n=1}^{N} \lambda_n \mathcal{N}(ln\ x | \theta_n) \tag{1.6}$$

$$\sum_{n=1}^{N} \lambda_n = 1 \tag{1.7}$$

where $0 \leq \lambda \leq 1$ are the mixing proportions, $\theta_n = (\mu_n, \sigma_n^2)$ are the mean and variance of the nth distribution. This assumption is supported by both empirical observations and theoretical insights [73, 74]. By superposition, the global distribution of measured fluorescence levels $p(lnx)$ for all values of m are also sampled from a mixture of K components:

$$p(ln\ x) = \sum_{m=0}^{2} \alpha_m p_m(ln\ x) = \sum_{m=0}^{2} \alpha_m \sum_{n=1}^{N} \lambda_n \mathcal{N}(ln\ x | \theta_n) = \sum_{k=1}^{K} \lambda_k \mathcal{N}(ln\ x | \theta_k) \tag{1.8}$$

$$\sum_{k=1}^{K} \lambda_k = 1 \tag{1.9}$$

where α_m denotes the overall fraction of cells with m copies of the gene encoding the clonal marker. For brevity, we substitute $X = lnx$ yielding:

$$p(X) = \sum_{k=1}^{K} \lambda_k \mathcal{N}(X|\theta_k) \tag{1.10}$$

Given a collection of sampled fluorescence levels, $\{X_i\}_{i=1...N}$, we use expectation maximization to find values of θ_k and λ_k for each of the model's K components that maximize the log-likelihood of the observed sample. We repeat this procedure for a range of sequential values of K, resulting in multiple models of increasing size. We then balance model resolution against overfitting by selecting the model that yields the smallest value of the Bayesian Information Criterion (BIC):

$$BIC(K) = ln(N)q_K - 2ln(\hat{L}_K) \tag{1.11}$$

$$q_K = K - 1 + 2^K \tag{1.12}$$

where N is the sample size, $ln(\hat{L})_K$ is the maximum value of the log-likelihood, the subscript K denotes the number of mixture components in the model, and q_K is the total number of parameters (i.e. $K-1$ values of λ_k and 2^K values of μ_k and σ_k^2).

Applying Bayes' rule to the selected model infers the posterior probabilities that each sample X_i belongs to the kth component:

$$p(k|X_i) = \frac{p(X_i|k)p(k)}{p(X_i)} = \frac{p(X_i|k)\lambda_k}{p(X_i)} \tag{1.13}$$

where $p(X_i \mid k)$ is evaluated using the model's likelihood function and $p(X_i)$ is evaluated by marginalizing across each of the model's K components. The end result is a mixture

model that allows us to predict the probability that a given measurement of clonal marker expression belongs to a particular one of its component distributions.

We then define a many-to-one mapping, f, from each of the K components of the mixture to each of the three possible values of m:

$$f : \{0, 1, ...K\} \rightarrow \{0, 1, 2\} \tag{1.14}$$

We determine the mapping by k-means clustering the K component distributions into three groups on the basis of their mean values, e^{μ_k}. We may then assign a genotype label m to each measurement X_i by predicting the component k from which it was sampled. The accuracy of these labels depends upon how closely the fitted mixture model reflects the true partitioning of gene copies among clones. While finite mixtures are always identifiable given a sufficiently large sample [75], the algorithm used to fit the mixture tends toward local maxima of the likelihood function when the true components are similar (Wu, 1983). An approach based on a univariate mixture is thus inherently prone to failure when expression levels extensively overlap across clones, as variation within each clone precludes accurate classification of a cell's genotype solely on the basis of its individual expression level. However, clonal lineages are unlikely to exist in isolation because recombination events are usually timed to generate large clones. Our strategy therefore integrates both clonal marker expression and spatial context to identify clusters of cells with locally homogeneous expression behavior.

We incorporate spatial context by introducing a second jointly-distributed variable Y_i:

$$Y_i = \frac{1}{M_i} \sum_{j=0}^{M_i} X_j \tag{1.15}$$

where the subscript j indexes all M_i neighbors of cell i. The new variable reflects the average expression level among the neighbors surrounding each cell. We define neighbors as pairs of cells located within a critical distance of each other. This distance, or sampling radius, is derived from the approximate length scale over which cells retain approximately similar clonal marker expression levels. Specifically, we determine the exponential decay constant of the spatial correlation function, $\psi(\delta)$:

$$\psi(\delta) = \frac{< (X_i - \mu_X)(X_j - \mu_X) >_{i,j \in \delta}}{\sigma_X^2} \tag{1.16}$$

where μ_X and σ_X^2 are the global mean and standard deviation, and angled brackets denote the mean across all pairs of cells separated by distance δ. We efficiently implement this procedure by fitting an exponential decay function to the down-sampled moving average of $\psi(\delta)$ as a function of increasing separation distance.

Following the introduction of spatial context, the mixture model becomes:

$$p(X, Y) = \sum_{k=1}^{K} \lambda_k \mathcal{N}(X, Y | \theta_k) \tag{1.17}$$

where $\theta_k = (\vec{\mu}_k, \vec{\sigma}_k^2)$ contains the mean and variance of each component given by vectors of length two. This formulation constrains each component's covariance matrix to be diagonal. The posterior is now:

$$p(k|X_i, Y_i) = \frac{p(X_i, Y_i | k)\lambda_k}{p(X_i, Y_i)} \tag{1.18}$$

We can recover the univariate model by marginalizing the posterior over all values of Y:

$$p(k|X_i) = \sum_{j} p(k|X_i, Y_j) \tag{1.19}$$

When neglecting spatial context, we use this expression to classify each sample by applying the mapping f to the value of k that maximizes $p(k \mid X_i)$:

$$f(\underset{k}{\mathrm{argmax}}\, p(k|X_i)) \tag{1.20}$$

In all other cases, we deploy a graph-based approach to refine the estimate of $p(k \mid X_i, Y_i)$. This first entails constructing an undirected graph connecting adjacent cells within each image. We obtain the graph's edges through Delaunay triangulation of the measured cell positions, then exclude distant neighbors by thresholding the edge lengths. Each edge is assigned a weight w_{ij} reflecting the similarity of clonal marker expression between adjacent cells i and j:

$$w_{ij} = exp\left(\frac{-E_{ij}}{\langle E \rangle}\right) \tag{1.21}$$

$$E_{ij} = |X_i - X_j| \tag{1.22}$$

where E_{ij} is the absolute log fold-change in measured expression level and angled brackets denote the mean across all edges. We chose an exponential formulation because it yields an approximately uniform distribution of edge weights. We then detect communities within the graph using the Infomap algorithm [76]. The algorithm provides a hierarchical partitioning of nodes into non-overlapping clusters. We aggregate all clusters below a critical level that is again chosen by estimating the spatial correlation decay constant. We then enumerate $p(k \mid X_i, Y_i^c)$ where Y_i^c is the spatial context obtained by averaging expression levels among all neighbors in the same community as cell i.

We further incorporate spatial context by allowing the posterior probabilities $p(k \mid X_i, Y_i^c)$ to diffuse among adjacent cells. We define the modified posterior probability $\hat{p}(k \mid X_i, Y_i^c)$

through a recursive relation analogous to the Katz centrality [?], initialized by $p(k \mid X_i, Y_i^c)$:

$$\hat{p}(k \mid X_i, Y_i^c) = \alpha \sum_j w_{ij} \hat{p}(k \mid X_i, Y_i^c) + \beta \tag{1.23}$$

$$\beta = (1 - \alpha) p(k \mid X_i, Y_i^c) \tag{1.24}$$

where α is the attenuation factor and w_{ij} are the edge weights. Expressed in matrix form, the solution for $\hat{p}(k \mid X, Y^c)$ is given by:

$$\hat{p}(k \mid X, Y^c) = (I - \alpha W)^{-1} (1 - \alpha) p(k \mid X, Y^c) \tag{1.25}$$

where I denotes the identity matrix and W is the matrix of edge weights w_{ij}. We then assign a label to each measurement X_i by applying f to the value of k that maximizes $\hat{p}(k \mid X_i, Y_i^c)$:

$$f(\operatorname*{argmax}_k \hat{p}(k \mid X_i, Y_i^c)) \tag{1.26}$$

Finally, we assess the total posterior probability of each assigned label, $\hat{P}(m_i)$:

$$\hat{P}(m_i) = \sum_{\{k \mid f(k) = m_i\}} \hat{p}(k \mid X_i, Y_i^c) \tag{1.27}$$

This measure reflects the overall confidence that m_i is the appropriate label. Labels whose confidence falls below 80% are replaced by their counterparts estimated using the marginal classifier. This substitution helps preserve classification accuracy in situations where spatial context is not informative, and is particularly useful when the annotated clones are relatively small.

1.7.5. Measurement curation

Our framework includes a simple graphical user interface that permits manual curation of which regions of the image field are included in subsequent analyses. This tool allows for the exclusion of certain aspects of the segmented data on a context-dependent basis. For example, when studying *Drosophila* eye imaginal discs it is common to restrict analysis to a narrow column of cells in which the cellular processes of interest are taking place (Fig. 1.6F). Alternatively, users may exclude regions of the image field marked by extensive tissue deformation or cell death. This functionality is convenient but is not a necessary component of a quantitative clonal analysis workflow.

1.7.6. Cell growth simulations

We simulated the two dimensional growth of a cell culture seeded with a single cell. In these simulations, growth proceeds through sequential division of cells (Fig. 1.10A). Not all cells divide at each time-step because cell division is a stochastic process. Instead, each cell divides stochastically with a rate controlled by a global growth rate parameter.

Cells are simultaneously subject to mitotic recombination (Fig. 1.10C). Each cell contains either zero, one, or two copies of a gene encoding an RFP-tagged clonal marker. Each time a cell divides, its genes are duplicated and equally partitioned between the two daughter cells. However, in some instances a heterozygous parent may instead partition its two duplicate genes unequally, with one daughter receiving both and the other receiving none. These mitotic recombination events occur stochastically with a frequency defined by a global recombination rate parameter.

The timing and duration of recombination events affects the number and size of the resultant clones. In real experiments, recombination events are restricted to a particular stage

of the developmental program through localized exogenous expression of the recombination machinery. We incorporated this feature into our cell growth simulations via two adjustable parameters. The first determines the minimum population size at which recombination may begin, while the second determines the number of generations over which recombination may continue to occur. These two parameters provide a means to tune the average number and size of clonal subpopulations in the synthetic data (Fig. 1.10D). Early recombination events generally entail larger clones, while shorter recombination periods limit the extent of clone formation (Fig. 1.10E).

After each round of cell division, all cells are repositioned in order to preserve approximately uniform spatial density (Fig. 1.10C). Repositioning is achieved by equilibrating a network of springs connecting each cell with its neighbors. This undirected network is constructed through Delaunay triangulation of all cells spatial positions. Edges on the periphery of the culture are systematically excluded by establishing a maximum polar angle between neighbors. This filtration removes spurious edges between distant pairs of cells. Edges connecting pairs of cells with the same clonal marker dosage are assigned a 10% higher spring constant than edges that connect dissimilar cells. This modest bias ensures that clonal lineages remain spatially collocated after repositioning. Cell positions are then updated using a force-directed graph drawing algorithm [77]. Alternating cell division and repositioning steps are then repeated until a predefined population size is reached.

We used this platform to generate all synthetic data presented in the manuscript. All simulations were terminated when the total population exceeded 2048 cells. We assigned each cell a 20% probability of division upon each iteration, and each cell division event was accompanied by a 20% chance of mitotic recombination. Parent cells containing zero or two copies of the recombined genes were ineligible for recombination, effectively sealing

the genetic fates of their respective lineages. Except where stated otherwise, all simulations limited recombination to the first sixteen cell division events.

1.7.7. Generation of synthetic microscopy data

Cell growth simulations yield a list of spatial coordinates and gene dosages for each nucleus (Fig. 1.10B). The corresponding fluorescence levels $\{x_1, x_2, \ldots x_{i=N}\}$ were sampled from a lognormal distribution conditioned upon the corresponding gene dosage (Fig. 1.11A-C):

$$ln\ x \sim \mathcal{N}_n(\theta_n) \tag{1.28}$$

where the subscript n denotes the gene copy number and $\theta_n = (\mu_n, \sigma_\alpha^2)$ are the mean and variance of the corresponding distribution. We define μ_n such that the mean fluorescence level doubles for each additional copy of the gene:

$$\mu_n = ln(2^{n-1}) \tag{1.29}$$

We refer to σ_α as the fluorescence ambiguity coefficient because it modulates the similarity of fluorescence levels across gene dosages. Increasing σ_α increases the overlap among $\mathcal{N}_0$, $\mathcal{N}_1$, and $\mathcal{N}_2$ (Fig. 1.11D,E), and consequently increases the difficulty of the annotation task (Fig. 1.11F).

1.7.8. Synthetic benchmarking of automated annotation performance

We created a synthetic microscopy dataset by varying the timing of recombination events to generate cell cultures spanning a range of sixteen average clone sizes (Fig. 1.10D, only half are shown). We performed 50 replicate simulations for each condition. The mixture model was independently trained and applied to each replicate. Training a single model on all

replicates yields stronger performance on average (not shown), but also yields more variable results across the parameter space because all labels are dependent upon the outcome of a single expectation maximization routine.

We then labeled each cell and quantified annotation performance by evaluating the mean absolute difference between the predicted labels and their respective true values. This metric preserves ordinality, meaning egregious misclassifications are penalized more severely than mild ones [55]. The entire procedure was repeated for sixteen different fluorescence ambiguity coefficient values.

1.7.9. Data and software availability

We have distributed the automated mosaic analysis framework as an open-source python package available at `https://github.com/sebastianbernasek/clones`. We also intend to incorporate its core features into future version of *FlyEye Silhouette*, our open-source platform for quantitative analysis of the larval eye. The code used to generate synthetic microscopy data is also freely available at `https://github.com/sebastianbernasek/growth`. All segmented and annotated eye discs are accessible via our data repository. These include a series of Jupyter notebooks that enable complete reproduction of all figures and analysis presented in this manuscript.
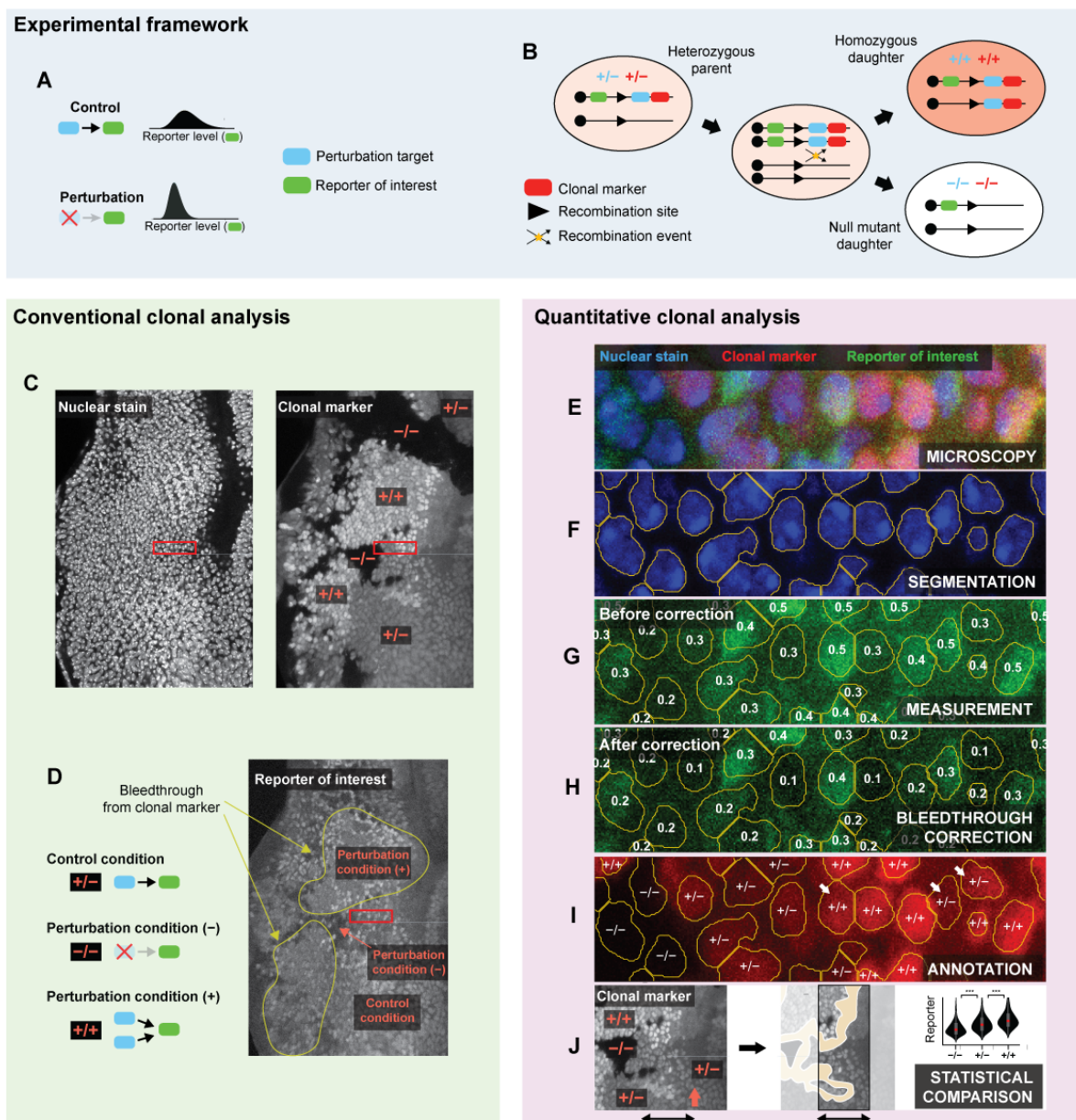
1.8. Supplementary figures

Figure 1.1. (Continued on next page.)

Figure 1.1. **A framework for conducting quantitative clonal analysis.** (A,B) Experimental framework using mitotic clones to test whether or not regulatory interactions occur between a perturbation target and reporter of interest. Blue and green ovals represent the respective genes encoding the perturbation target and the reporter. (A) A perturbation-induced decrease in reporter levels would confirm that regulation occurs. (B) Mitotic recombination generates clonal subpopulations carrying zero, one, or two copies of the gene encoding a perturbation target. Black lines depict a genetic locus. Only genes downstream of the recombination site are subject to recombination. Red ovals represent a gene encoding a clonal marker used to identify the resultant clones.(C,D) Conventional clonal analysis in the larval fruit fly eye. (C) Clones are identified by visual comparison of clonal marker fluorescence among nuclei. (D) Regions labeled mutant (/) or homozygous (+/+) for the clonal marker are compared with those labeled heterozygous (+/) to assess whether reporter expression differs across clones. Fluorescence bleed-through is arbitrarily diagnosed. (E-J) Quantitative clonal analysis. Panels depict a magnified view of the region enclosed by red rectangles in panels C and D. (E) Raw confocal image of the nuclear stain, clonal marker, and reporter of interest. (F) Segmentation identifies distinct nuclei. (G) Reporter expression is quantified by averaging the pixel intensities within each segment. Numbers reflect measured values. (H) Measurements may be corrected to mitigate fluorescence bleedthrough. (I) Individual nuclei are labeled mutant, heterozygous, or homozygous for the clonal marker. White arrows mark nuclei with ambiguous fluorescence levels. (J) Reporter levels are compared across clones to determine whether the perturbation affects reporter expression. Yellow region marks excluded clone borders. Comparison may exclude clone borders (yellow regions) and focus on a particular region of the image field (black arrows). In the larval eye, comparison is often limited to a narrow window near the MF (orange arrow).
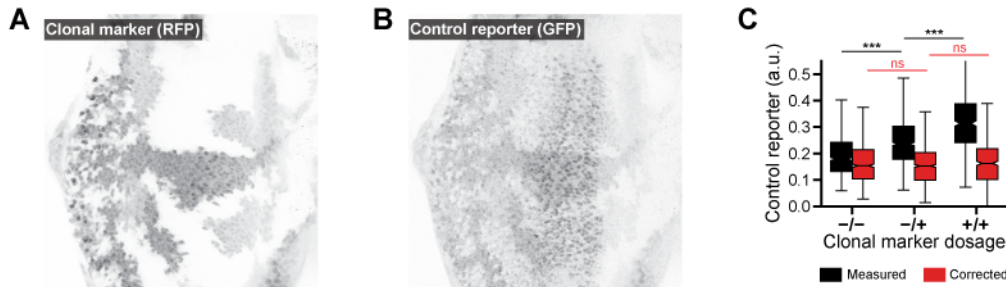
Figure 1.2. **Automated correction of fluorescence bleedthrough in the larval eye.** (A) Low, medium, and high expression levels of the RFP-tagged clonal marker. (B) GFP-tagged control reporter expression. RFP fluorescence bleedthrough is visually apparent upon comparison with A. (C) Comparison of control reporter expression between clones. Includes data aggregated across nine images taken from six separate eye discs. Data were limited to cells within the region of elevated GFP expression that were of approximately comparable developmental age (see Fig. 1.12E). Measurements are stratified by their assigned labels. Before correction, expression differs between clones (black boxes, $p < 10^{-5}$). No difference is detected after correction (red boxes, $p > 0.05$).
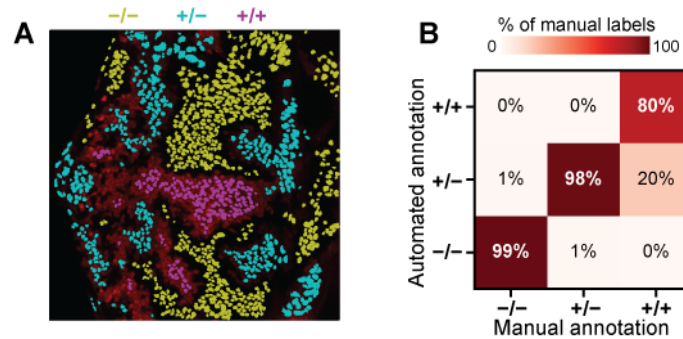

Figure 1.3. **Automated unsupervised annotation of clones in the larval eye.** (A) Labels assigned by automated annotation. Yellow, cyan, and magenta denote the label assigned to each contour. Labels are overlaid on the RFP channel of the image shown in 1.1B. Cells on the periphery of each clone are excluded. (B) Comparison of automated annotation with manually-assigned labels. Confusion matrix includes data aggregated across nine images taken from six separate eye discs. Cells on the periphery of each clone are included. Columns sum to one.
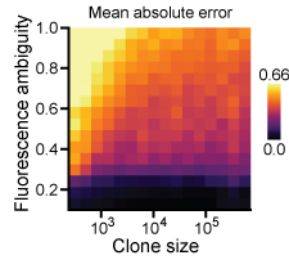
Figure 1.4. **Synthetic benchmarking of automated annotation performance.** Grid shows the mean absolute error (MAE) between assigned labels and the corresponding ground truth as a function of fluorescence ambiguity and clone size. Values shown reflect the average across 50 replicate simulations. Performance improves with increasing clone size and worsens with increasing fluorescence ambiguity. Clone size reflects the mean number of cells per clone.
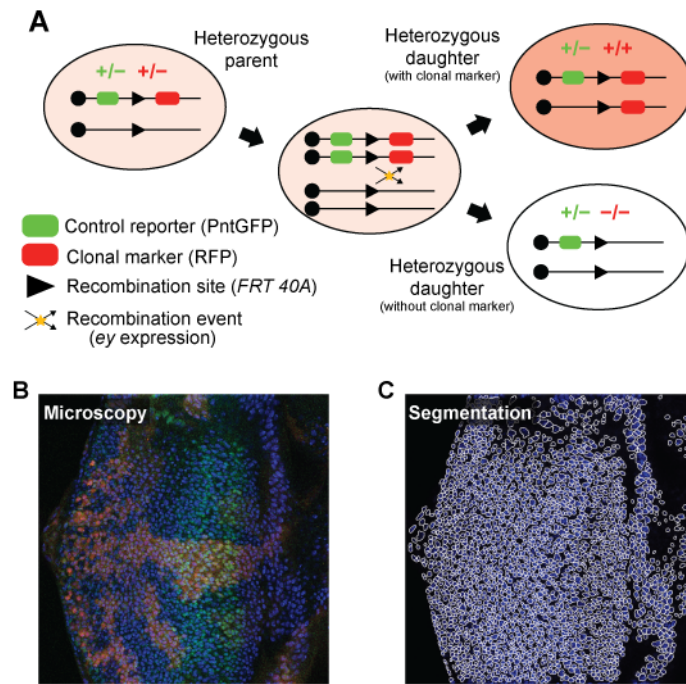

Figure 1.5. **Example clones in the larval fly eye.** (A) Genetic schema for a bleedthrough control experiment. Red and green ovals represent genes encoding a RFP-tagged clonal marker and a GFP-tagged control reporter, respectively. Black lines depict a genomic locus. Recombination does not affect gene dosage of the control reporter, so GFP variation across clones is attributed to fluorescence bleedthrough. (B) Confocal image of an eye imaginal disc. Red, green, and blue reflect clonal marker, control reporter, and nuclear stain fluorescence, respectively. (C) Segmentation of the DAPI nuclear stain. White lines show individual segments.
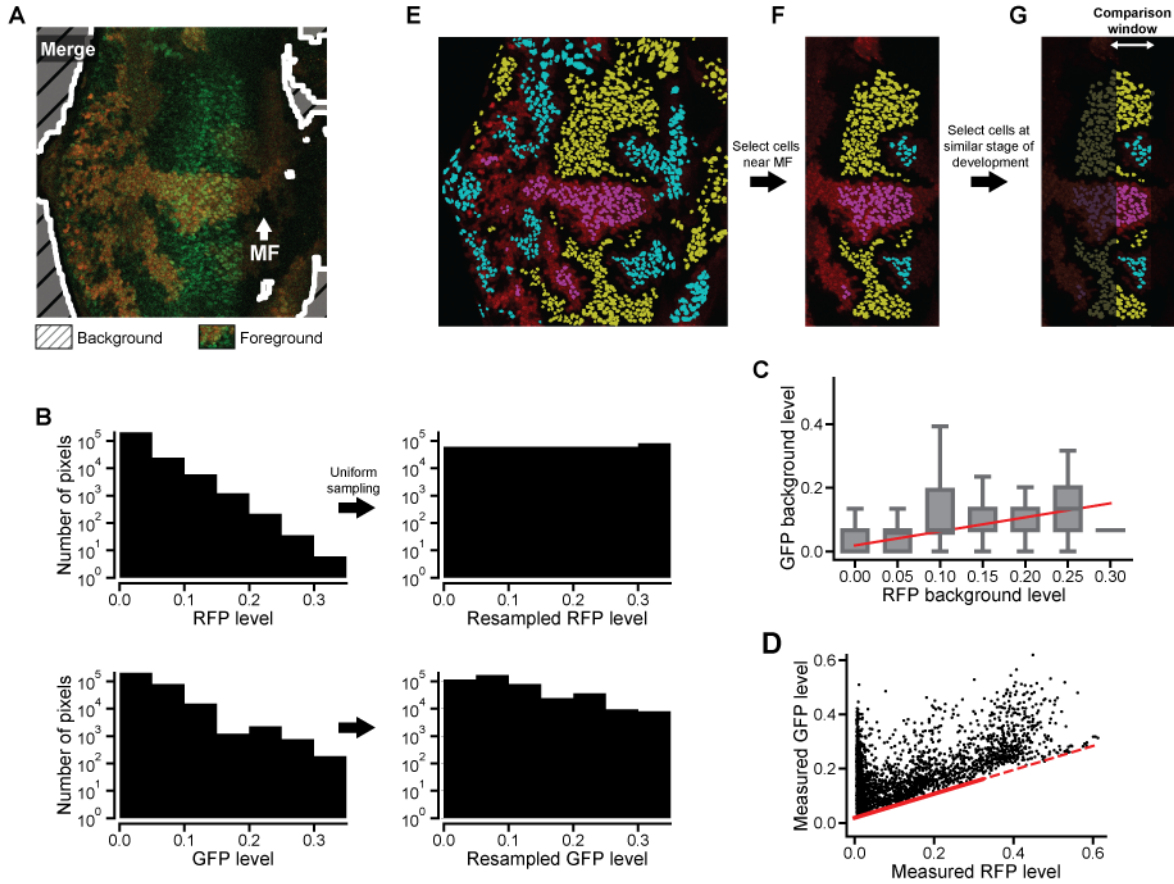
Figure 1.6. **Intermediate stages of bleedthrough correction.** (A) Extraction of background pixels (striped region). Foreground includes the merged RFP and GFP images, surrounded by a white line. White arrow marks the morphogenetic furrow (MF). (B) Background pixel values are resampled such that RFP intensities are uniformly distributed. (C) A generalized linear model characterizes the contribution of RFP bleedthrough to GFP fluorescence. Boxes reflect windowed distributions of resampled background pixel intensities. Red line shows the model fit. (D) Measured GFP levels before bleedthrough correction. Markers represent individual nuclei. Red line shows the inferred contributions of RFP fluorescence bleedthrough. Dashed portion is extrapolated. (E-G) Data curation prior to statistical comparison of GFP levels. (E) Cells on the periphery of each clone are excluded. (F) The selection is limited to the region of elevated GFP expression near the MF. (G) It is further limited to cells of the same developmental age, defined by their relative positions along the x-axis.
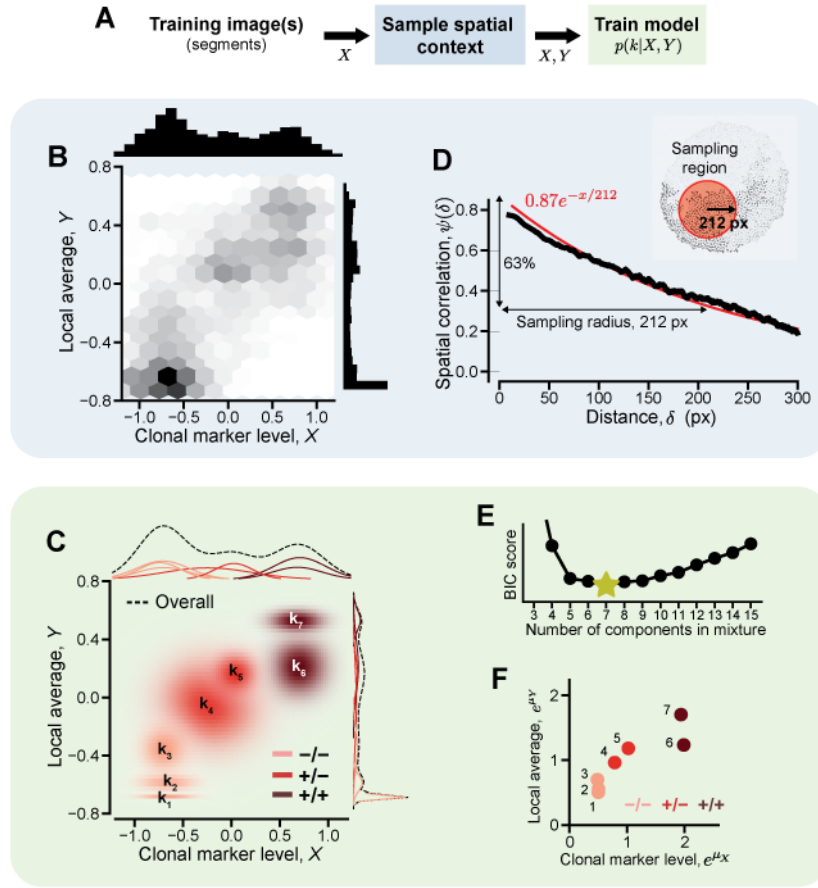
Figure 1.7. **Training a clone annotation model.** (A) One or more images are segmented, yielding a set of fluorescence measurements X. These are used to sample the spatial context Y of the neighborhood surrounding each cell. Both sets of values are used to train a mixture model. Subsequent panels demonstrate these procedures using the example shown in Figure 1.7C. (B) Expression levels are jointly distributed with the local average among neighboring cells. Center panel shows the joint distribution. Top and right bar plots show marginal distributions. (C) Mixture model identifies seven distinct components k_i. Center panel shows position and spread of each component. Top and right panels show marginal components scaled by their respective weights. Red shading denotes the label m_i assigned to each component. The model predicts the posterior probabilities that a given sample (X, Y) belongs to each component. (D) Neighborhood size is estimated by computing the decay constant of the spatial correlation function, $\psi(\delta)$. Black line shows the moving average of $\psi(\delta)$, red line shows an exponential fit. Inset shows the resultant sampling region. (E) The optimal number of mixture components is determined by minimizing BIC score. (F) Mixture components are labeled by k-means clustering their mean values. Markers reflect the component means, colors denote the assigned label.
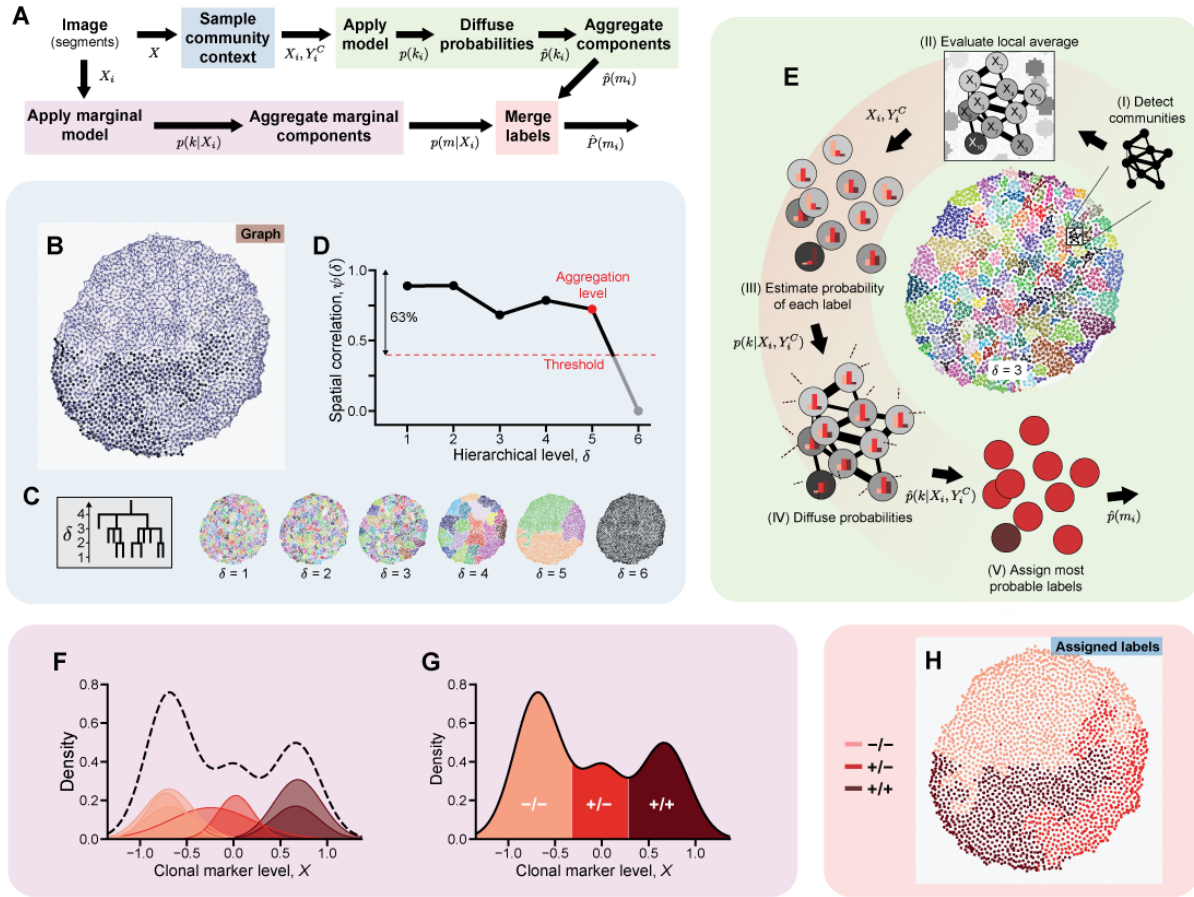
Figure 1.8. **Label assignment using a trained clone annotation model.** (Continued on next page.)

Figure 1.8. **Label assignment using a trained clone annotation model.** (A) The measurements X from a segmented image are used to sample the spatial context Y^C of the community surrounding each cell before the mixture model is applied (blue and green path). They are simultaneously labeled using a marginal projection of the trained model (magenta path). The two sets of labels are then merged (red path). Subsequent panels demonstrate these procedures using the example shown in Figure 1.7C. (B-D) Spatial context sampling. (B) Weighted undirected graph connecting adjacent cells. Line thickness reflects the expression similarity between neighbors. (C) Community resolution is defined by aggregating clusters that fall below a cut level δ in the hierarchy. Images show potential levels of aggregation. Colors denote distinct communities. (D) Cut level is chosen by finding the maximum level (red dot) that remains lower than the decay constant of the spatial correlation function, $\psi(\delta)$ (black line). In this example, clusters are aggregated below the fifth level. Panel E instead depicts aggregation below the third level for ease of visualization. (E) Application of the mixture model. *(I)* The graph connecting adjacent cells contains distinct communities of locally similar expression. *(II)* Mean expression level within each community serves as the local average for each cell. *(III)* Mixture model estimates the probability that each cell belongs to each of its component. Bar plots within each cell illustrate the cumulative probability of each label. *(IV)* Posterior probabilities are diffused across the entire graph. *(V)* Each cell is assigned its most probable label. (F,G) Application of a marginal mixture model that neglects spatial context. (F) Marginal mixture model components obtained by summing across the spatial context dimension of the full mixture model. Red shading denotes the label assigned to each component. Dashed black line is the overall marginal density. (G) Marginal classifier that labels cells strictly on the basis of their individual fluorescence level. Red shading denotes the most probable label for each expression level. (H) Annotated measurements. Red shading denotes the assigned label. Labels with low confidence $\hat{P}(m_i) < 0.8$ are replaced by their marginal counterparts.

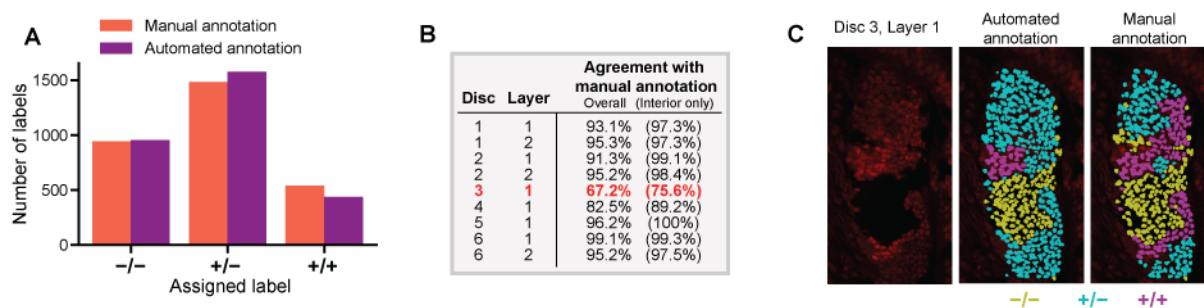| Disc	Layer	Agreement with manual annotation	
		Overall	(Interior only)
1	1	93.1%	(97.3%)
1	2	95.3%	(97.3%)
2	1	91.3%	(99.1%)
2	2	95.2%	(98.4%)
3	1	67.2%	(75.6%)
4	1	82.5%	(89.2%)
5	1	96.2%	(100%)
6	1	99.1%	(99.3%)
6	2	95.2%	(97.5%)

Figure 1.9. **Comparison of automated annotation with manually assigned labels.** (A) Distribution of labels among each possible value. (B) Agreement between automated and manual annotation. Values in parentheses exclude cells on the periphery of each clone. The sole instance of low agreement is marked in red. (C) Visual comparison of the instance in which automated and manual annotation differ. Image shows clonal marker fluorescence, overlayed colors denote the assigned label.
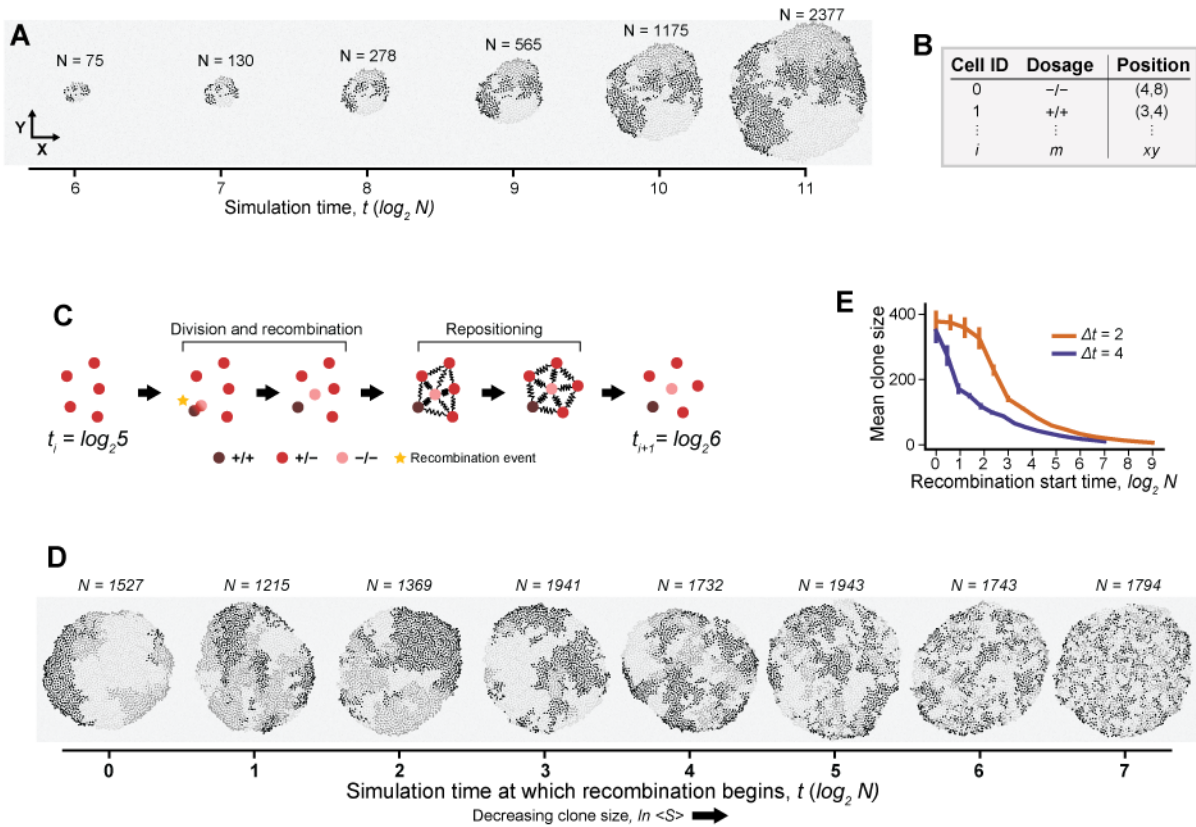
Figure 1.10. **Simulated growth of a synthetic cell culture.** (A) Partial simulation time course. Each marker depicts a cell. Greyscale intensity reflects clonal marker gene dosage. Simulation time reflects the approximate number of cell divisions since the initial seed. (B) Simulations yield gene dosages and spatial coordinates for each cell. (C) Single iteration of an example simulation. Circles represent individual cells, red shading denotes clonal marker dosage. Cycles of cell division, recombination, and repositioning are repeated until the simulation reaches a specified end time ($t > 11$ in panel A). (D) Cultures simulated with varying recombination start times. All cultures were subject to four generations of recombination ($\delta t = 4$). Recombination start time increases from left to right. Later recombination events generally yield smaller clones. (E) Mean clone size (cells per clone) as a function of the recombination start time. Colors denote recombination period duration. Error bars reflect standard error of the mean across 50 replicates. Clone size generally decreases as recombination is limited to later times.
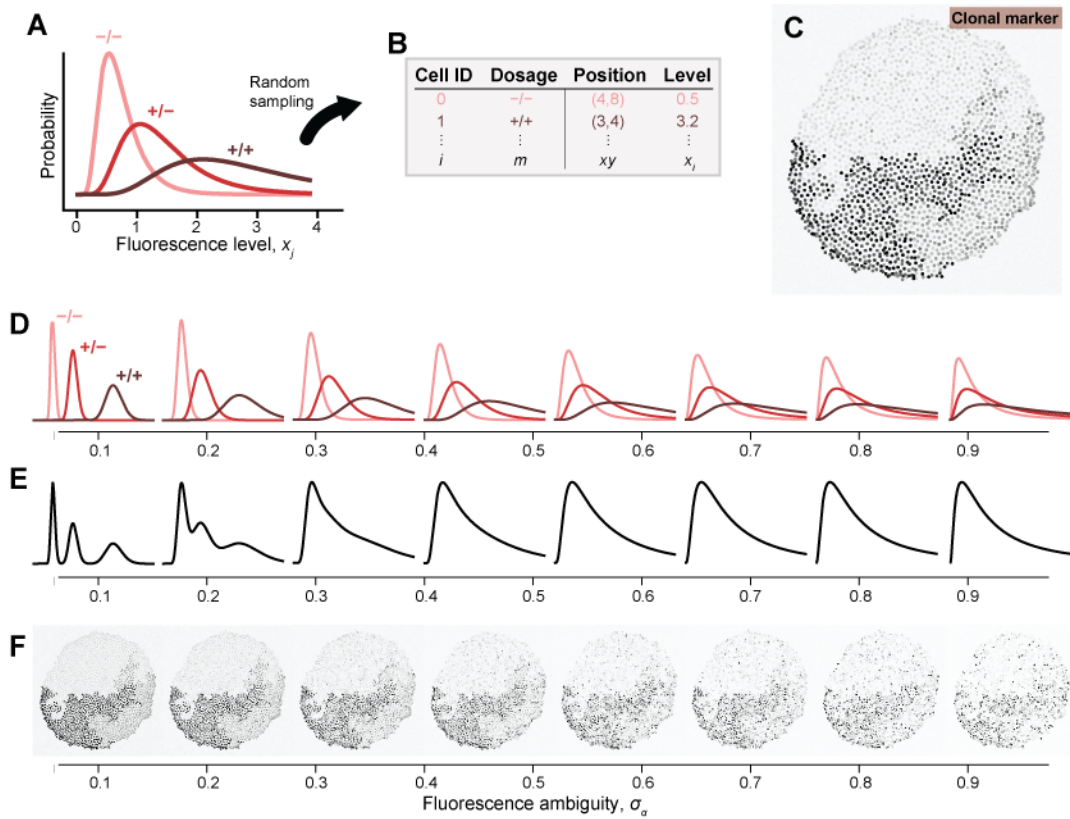
Figure 1.11. **Tunable generation of synthetic microscopy data.** (A) Fluorescence levels are sampled from lognormal distributions conditioned upon gene dosage. (B) Synthetic data include a measured fluorescence level for each reporter in each cell. Text color reflects the generative distribution in A. (C) Synthetic image of clonal marker fluorescence when $\sigma_\alpha = 0.25$. Each nucleus is shaded in accordance with its sampled fluorescence intensity. (D-F) Left to right, increasing the fluorescence ambiguity parameter broadens the overlap in fluorescence levels across gene dosages. (D) Distributions used to generate clonal marker fluorescence levels. Red shading denotes gene dosage. (E) Evenly weighted sum of the generative distributions. (F) Example images of clonal marker fluorescence.
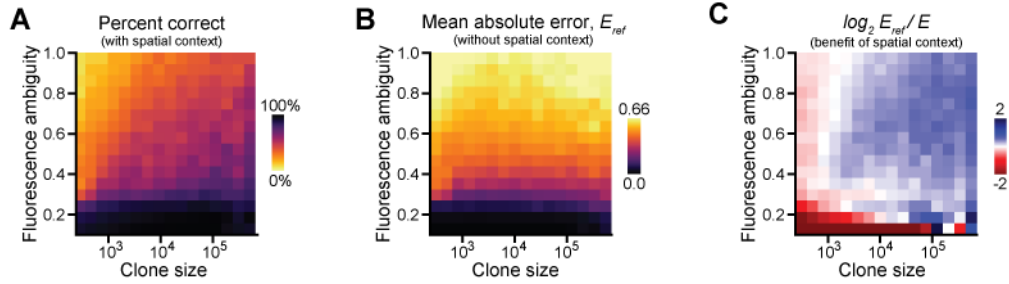
Figure 1.12. **Additional results for synthetic benchmarking of annotation algorithm.** (A) Annotation accuracy as a function of fluorescence ambiguity and clone size. Accuracy is defined as the fraction of cells that were correctly labeled. Performance improves with increasing clone size and worsens with increasing fluorescence ambiguity. (B) Annotation performance of a marginal classifier that neglects spatial context. Each point in the grid reflects the mean absolute error (MAE) between assigned labels and the corresponding ground truth. Performance worsens with increasing fluorescence ambiguity but does not depend upon clone size. (C) Annotation performance relative to the marginal classifier. Color scale reflects the log_2 fold-change in MAE when spatial context is ignored. Blue indicates that spatial context improves performance. Spatial context is informative for larger clones, particularly when fluorescence levels are ambiguous.

References

[1] Andrew C. Oates, Nicole Gorfinkiel, Marcos González-Gaitán, and Carl Philipp Heisenberg. Quantitative approaches in developmental biology, aug 2009. ISSN 14710056. URL http://www.nature.com/articles/nrg2548.

[2] Dale Muzzey and Alexander van Oudenaarden. Quantitative Time-Lapse Fluorescence Microscopy in Single Cells. *Annual Review of Cell and Developmental Biology*, 25(1): 301–327, nov 2009. ISSN 1081-0706. doi: 10.1146/annurev.cellbio.042308.113408. URL http://www.annualreviews.org/doi/10.1146/annurev.cellbio.042308.113408.

[3] Ernst H K Stelzer. Light-sheet fluorescence microscopy for quantitative biology. *Nature Methods*, 12(1):23–26, dec 2014. ISSN 1548-7091. doi: 10.1038/nmeth.3219.

[4] Thai V. Truong and Willy Supatto. Toward high-content/high-throughput imaging and analysis of embryonic morphogenesis. *Genesis*, 49(7):555–569, jul 2011. ISSN 1526954X. doi: 10.1002/dvg.20760. URL http://doi.wiley.com/10.1002/dvg.20760.

[5] Nima Aghaeepour, Greg Finak, Holger Hoos, Tim R Mosmann, Ryan Brinkman, Raphael Gottardo, Richard H Scheuermann, David Dougall, Alireza Hadj Khodabakhshi, Phillip Mah, Gerlinde Obermoser, Josef Spidlen, Ian Taylor, Sherry A. Wuensch, Jonathan Bramson, Connie Eaves, Andrew P. Weng, Edgardo S. Fortuno, Kevin Ho, Tobias R. Kollmann, Wade Rogers, Stephen De Rosa, Bakul Dalai, Ariful Azad, Alex Pothen, Aaron Brandes, Hannes Bretschneider, Robert Bruggner, Rachel Finck, Robin Jia, Noah Zimmerman, Michael Linderman, David Dill, Gary Nolan, Cliburn Chan, Faysal El Khettabi, Kieran O'Neill, Maria Chikina, Yongchao Ge, Stuart Sealfon, István Sugár, Arvind Gupta, Parisa Shooshtari, Habil Zare, Philip L. De Jager, Mike Jiang, Jens Keilwagen, Jose M. Maisog, George Luta, Andrea A. Barbo, Peter Májek, Jozef Vilček, Tapio Manninen, Heikki Huttunen, Pekka Ruusuvuori, Matti Nykter, Geoffrey J. McLachlan, Kui Wang, Iftekhar Naim, Gaurav Sharma, Radina Nikolic, Saumyadipta Pyne, Yu Qian, Peng Qiu, John Quinn, Andrew Roth, Pablo Meyer, Gustavo Stolovitzky, Julio Saez-Rodriguez, Raquel Norel, Madhuchhanda Bhattacharjee, Michael Biehl, Philipp Bucher, Kerstin Bunte, Barbara Di Camillo, Francesco Sambo, Tiziana Sanavia, Emanuele Trifoglio, Gianna Toffolo,

S. D. Slavica Dimitrieva, Rene Dreos, Giovanna Ambrosini, Jan Grau, Ivo Grosse, Stefan Posch, Nicolas Guex, Miron Kursa, Witold Rudnicki, Bo Liu, Mark Maienschein-Cline, Petra Schneider, Michael Seifert, Marc Strickert, and Jose M.G. Vilar. Critical assessment of automated flow cytometry data analysis techniques. *Nature Methods*, 10(3):228–238, mar 2013. ISSN 15487091. doi: 10.1038/NMETH.2365. URL `http://www.nature.com/articles/nmeth.2365`.

[6] Xiaoyi Chen, Milena Hasan, Valentina Libri, Alejandra Urrutia, Benoît Beitz, Vincent Rouilly, Darragh Duffy, Étienne Patin, Bernard Chalmond, Lars Rogge, Lluis Quintana-Murci, Matthew L. Albert, and Benno Schwikowski. Automated flow cytometric analysis across large numbers of samples and cell types. *Clinical Immunology*, 157(2):249–260, apr 2015. ISSN 15217035. doi: 10.1016/j.clim.2014.12.009. URL `https://www.sciencedirect.com/science/article/pii/S1521661614002873`.

[7] Saumyadipta Pyne, Lisa M Maier, T.-I Lin, Kui Wang, Elizabeth Rossin, Xinli Hu, Pablo Tamayo, David A Hafler, Philip L De Jager, Geoffrey J McLachlan, Clare Baecher-Allan, and Jill P Mesirov. Automated high-dimensional flow cytometric data analysis. *Proceedings of the National Academy of Sciences*, 106(21):8519–8524, may 2009. ISSN 0027-8424. doi: 10.1073/pnas. 0903028106. URL `http://www.ncbi.nlm.nih.gov/pubmed/19443687http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2682540`.

[8] Bradley E Bernstein, Myles Brown, David S Johnson, X Shirley Liu, Chad Nussbaum, Richard M Myers, Tao Liu, Wei Li, Clifford A Meyer, Jérôme Eeckhoute, and Yong Zhang. Model-based Analysis of ChIP-Seq (MACS). *Genome Biology*, 9(9):R137, sep 2008. ISSN 1465-6906. doi: 10.1186/gb-2008-9-9-r137. URL `http://genomebiology.biomedcentral.com/articles/10.1186/gb-2008-9-9-r137`.

[9] Jan Hellemans, Geert Mortier, Anne De Paepe, Frank Speleman, and Jo Vandesompele. qBase relative quantification framework and software for management and automated analysis of real-time quantitative PCR data. *Genome Biology*, 8(2), feb 2007. ISSN 1465-6914. doi: 10.1186/gb-2007-8-2-r19. URL `http://genomebiology.biomedcentral.com/articles/10.1186/gb-2007-8-2-r19`.

[10] Ben Langmead and Steven L Salzberg. Fast gapped-read alignment with Bowtie 2. *Nature methods*, 9(4):357–9, apr 2012. ISSN 1548-7105. doi: 10.1038/nmeth.1923. URL `http://www.nature.com/articles/nmeth.1923http://www.ncbi.nlm.nih.gov/pubmed/22388286{%}0Ahttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3322381`.

[11] Cole Trapnell, Lior Pachter, and Steven L. Salzberg. TopHat: Discovering splice junctions with RNA-Seq. *Bioinformatics*, 25(9):1105–1111, may 2009. ISSN 13674803. doi: 10.1093/bioinformatics/btp120. URL `https://academic.oup.com/bioinformatics/article-lookup/doi/10.1093/bioinformatics/btp120`.

[12] Sylvain V. Costes, Dirk Daelemans, Edward H. Cho, Zachary Dobbin, George Pavlakis, and Stephen Lockett. Automatic and quantitative measurement of protein-protein colocalization in live cells. *Biophysical Journal*, 86(6):3993–4003, jun 2004. ISSN 00063495. doi: 10.1529/biophysj.103.038422. URL `https://www.sciencedirect.com/science/article/pii/S0006349504744392`.

[13] Lawrence A Kelley, Stefans Mezulis, Christopher M Yates, Mark N Wass, and Michael J.E. Sternberg. The Phyre2 web portal for protein modeling, prediction and analysis. *Nature Protocols*, 10(6):845–858, jun 2015. ISSN 17502799. doi: 10.1038/nprot.2015.053. URL `http://www.nature.com/articles/nprot.2015.053`.

[14] Anne E Carpenter, Thouis R Jones, Michael R Lamprecht, Colin Clarke, In Han Kang, Ola Friman, David A Guertin, Joo Han Chang, Robert A Lindquist, Jason Moffat, Polina Golland, and David M Sabatini. CellProfiler: Image analysis software for identifying and quantifying cell phenotypes. *Genome Biology*, 7(10):R100, oct 2006. ISSN 14747596. doi: 10.1186/gb-2006-7-10-r100. URL `http://genomebiology.biomedcentral.com/articles/10.1186/gb-2006-7-10-r100`.

[15] Ahmad Paintdakhi, Bradley Parry, Manuel Campos, Irnov Irnov, Johan Elf, Ivan Surovtsev, and Christine Jacobs-Wagner. Oufti: An integrated software package for high-accuracy, high-throughput quantitative microscopy analysis. *Molecular Microbiology*, 99(4):767–777, feb 2016. ISSN 13652958. doi: 10.1111/mmi.13264. URL `http://doi.wiley.com/10.1111/mmi.13264`.

[16] Johannes Schindelin, Ignacio Arganda-Carreras, Erwin Frise, Verena Kaynig, Mark Longair, Tobias Pietzsch, Stephan Preibisch, Curtis Rueden, Stephan Saalfeld, Benjamin Schmid, Jean Yves Tinevez, Daniel James White, Volker Hartenstein, Kevin Eliceiri, Pavel Tomancak, and Albert Cardona. Fiji: An open-source platform for biological-image analysis, jul 2012. ISSN 15487091. URL `http://www.nature.com/articles/nmeth.2019`.

[17] Christoph Sommer, Christoph Straehle, Ullrich Kothe, and Fred A. Hamprecht. Ilastik: Interactive learning and segmentation toolkit. In *Proceedings - International Symposium on Biomedical Imaging*, pages 230–233. IEEE, mar 2011. ISBN 9781424441280. doi: 10.1109/ISBI.2011.5872394. URL `http://ieeexplore.ieee.org/document/5872394/`.

[18] Florian Jug, Tobias Pietzsch, Stephan Preibisch, and Pavel Tomancak. Bioimage Informatics in the context of Drosophila research. *Methods*, 68(1):60–73, jun 2014. ISSN 10959130. doi: 10.1016/j.ymeth.2014.04.004. URL `https://www.sciencedirect.com/science/article/pii/S1046202314001480`.

[19] Ivo F. Sbalzarini. Seeing is believing: Quantifying is convincing: Computational image analysis in biology. *Advances in Anatomy, Embryology and Cell Biology*, 219:1–39, 2016. ISSN 03015556. doi: 10.1007/978-3-319-28549-8_1. URL `http://link.springer.com/`

10.1007/978-3-319-28549-8{_}1.

[20] Johannes Schindelin, Curtis T. Rueden, Mark C. Hiner, and Kevin W. Eliceiri. The ImageJ ecosystem: An open platform for biomedical image analysis, jul 2015. ISSN 10982795. URL http://doi.wiley.com/10.1002/mrd.22489.

[21] Ian T. Simpson and David J. Price. Pax6; a pleiotropic player in development, nov 2002. ISSN 02659247. URL http://doi.wiley.com/10.1002/bies.10174.

[22] T R Parody and M. A T Muskavitch. The pleiotropic function of Delta during postembryonic development of Drosophila melanogaster. *Genetics*, 135(2):527–539, 1993. ISSN 00166731.

[23] Ben Zion Shilo and Erez Raz. Developmental control by the Drosophila EGF receptor homolog DER, nov 1991. ISSN 01689525. URL https://www.sciencedirect.com/science/article/pii/016895259190261N.

[24] T Xu and G M Rubin. Analysis of genetic mosaics in developing and adult Drosophila tissues. *Development (Cambridge, England)*, 117(4):1223–37, apr 1993. ISSN 0950-1991. doi: 8404527. URL http://www.ncbi.nlm.nih.gov/pubmed/8404527.

[25] Tian Xu and Gerald M Rubin. The effort to make mosaic analysis a household tool. *Development*, 139(24):4501–4503, dec 2012. ISSN 0950-1991. doi: 10.1242/dev.085183. URL http://www.ncbi.nlm.nih.gov/pubmed/23172911http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3509720.

[26] Jorge V. Beira and Renato Paro. The legacy of Drosophila imaginal discs, sep 2016. ISSN 14320886. URL http://link.springer.com/10.1007/s00412-016-0595-4.

[27] T P Newsome, B Asling, and B J Dickson. Analysis of Drosophila photoreceptor axon guidance in eye-specific mosaics. *Development (Cambridge, England)*, 127(4):851–60, feb 2000. ISSN 0950-1991. URL http://www.ncbi.nlm.nih.gov/pubmed/10648243.

[28] Nicole A. Theodosiou and Tian Xu. Use of FLP/FRT system to study Drosophila development. *Methods: A Companion to Methods in Enzymology*, 14(4):355–365, apr 1998. ISSN 10462023. doi: 10.1006/meth.1998.0591. URL https://www.sciencedirect.com/science/article/pii/S1046202398905916?via{%}3Dihub.

[29] Gary Struhl and Konrad Basler. Organizing activity of wingless protein in Drosophila. *Cell*, 1993. ISSN 00928674. doi: 10.1016/0092-8674(93)90072-X.

[30] K Halfar, C Rommel, H Stocker, and E Hafen. Ras controls growth, survival and differentiation in the Drosophila eye by different thresholds of MAP kinase activity. *Development (Cambridge, England)*, 128(9):1687–96, may 2001. ISSN 0950-1991. URL

`http://www.ncbi.nlm.nih.gov/pubmed/11290305`.

[31] A Tomlinson and G Struhl. Delta/Notch and Boss/Sevenless signals act combinatorially to specify the Drosophila R7 photoreceptor. *Molecular cell*, 7(3):487–95, mar 2001. ISSN 1097-2765. URL `http://www.ncbi.nlm.nih.gov/pubmed/11463374`.

[32] L Yang and N E Baker. Role of the EGFR/Ras/Raf pathway in specification of photoreceptor cells in the Drosophila retina. *Development (Cambridge, England)*, 128(7): 1183–91, apr 2001. ISSN 0950-1991. URL `http://www.ncbi.nlm.nih.gov/pubmed/11245584`.

[33] Jianbin Huang, Shian Wu, Jose Barrera, Krista Matthews, and Duojia Pan. The Hippo signaling pathway coordinately regulates cell proliferation and apoptosis by inactivating Yorkie, the Drosophila homolog of YAP. *Cell*, 122(3):421–434, aug 2005. ISSN 00928674. doi: 10.1016/j.cell.2005.06.007. URL `https://www.sciencedirect.com/science/article/pii/S0092867405005520`.

[34] Barry J. Thompson and Stephen M. Cohen. The Hippo Pathway Regulates the bantam microRNA to Control Cell Proliferation and Apoptosis in Drosophila. *Cell*, 126(4): 767–774, aug 2006. ISSN 00928674. doi: 10.1016/j.cell.2006.07.013. URL `https://www.sciencedirect.com/science/article/pii/S0092867406009536{#}bib4`.

[35] Mardelle Atkins. Drosophila Genetics: The Power of Genetic Mosaic Approaches. In *Methods in Molecular Biology*, volume 1893, pages 27–42. Humana Press, New York, NY, 2019. doi: 10.1007/978-1-4939-8910-2_2. URL `http://link.springer.com/10.1007/978-1-4939-8910-2{_}2`.

[36] Masato Enomoto, Carmen Siow, and Tatsushi Igaki. Drosophila as a cancer model. In *Advances in Experimental Medicine and Biology*, volume 1076, pages 173–194. Springer, Singapore, 2018. ISBN 9789811305290. doi: 10.1007/978-981-13-0529-0_10. URL `http://link.springer.com/10.1007/978-981-13-0529-0{_}10`.

[37] Federico Germani, Cora Bergantinos, and Laura A Johnston. Mosaic analysis in Drosophila. *Genetics*, 208(2):473–490, feb 2018. ISSN 19432631. doi: 10.1534/genetics.117.300256. URL `http://www.ncbi.nlm.nih.gov/pubmed/29378809http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC5788516`.

[38] Sebastian M. Bernasek, Jean-François Boisclair Lachance, Nicolás Peláez, Rachael Bakker, Heliodoro Tejedor Navarro, Luis A. N. Amaral, Neda Bagheri, Ilaria Rebay, and Richard W. Carthew. Ratio-based sensing of two transcription factors regulates the transit to differentiation. *bioRxiv*, page 430744, sep 2018. doi: http://dx.doi.org/10.1101/430744. URL `https://www.biorxiv.org/content/10.1101/430744v1`.

[39] Haley Burrous, Thomas Kenney, Wei Dai, Denise J. Montell, and Amy Peterson. Quantitative microscopy of the Drosophila ovary shows multiple niche signals specify progenitor cell fate. *Nature Communications*, 8(1):1244, dec 2017. ISSN 2041-1723. doi: 10.1038/s41467-017-01322-9. URL `http://www.nature.com/articles/s41467-017-01322-9`.

[40] Christian Ghiglione, Patrick Jouandin, Delphine Cérézo, and Stéphane Noselli. The Drosophila insulin pathway controls Profilin expression and dynamic actin-rich protrusions during collective cell migration. *Development*, 145(14):dev161117, jul 2018. ISSN 0950-1991. doi: 10.1242/dev.161117. URL `http://dev.biologists.org/content/145/14/dev161117.abstract`.

[41] Ke Li and Nicholas E Baker. Regulation of the Drosophila ID protein Extra macrochaetae by proneural dimerization partners. *eLife*, 7, 2018. ISSN 2050-084X. doi: 10.7554/elife.33967. URL `http://www.ncbi.nlm.nih.gov/pubmed/29687780http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC5915177`.

[42] Kirsten Bacia, Zdeněk Petrášek, and Petra Schwille. Correcting for spectral cross-talk in dual-color fluorescence cross-correlation spectroscopy. *ChemPhysChem*, 13(5):1221–1231, apr 2012. ISSN 14394235. doi: 10.1002/cphc.201100801. URL `http://doi.wiley.com/10.1002/cphc.201100801`.

[43] Masilamani Elangovan, Horst Wallrabe, Ye Chen, Richard N. Day, Margarida Barroso, and Ammasi Periasamy. Characterization of one- and two-photon excitation fluorescence resonance energy transfer microscopy. *Methods*, 29(1):58–73, jan 2003. ISSN 10462023. doi: 10.1016/S1046-2023(02)00283-9. URL `https://www.sciencedirect.com/science/article/pii/S1046202302002839`.

[44] Richard L. Mort. Quantitative analysis of patch patterns in mosaic tissues with ClonalTools software. *Journal of Anatomy*, 215(6):698–704, dec 2009. ISSN 00218782. doi: 10.1111/j.1469-7580.2009.01150.x. URL `http://doi.wiley.com/10.1111/j.1469-7580.2009.01150.x`.

[45] Jo A Helmuth, Grégory Paul, and Ivo F Sbalzarini. Beyond co-localization: Inferring spatial interactions between sub-cellular structures from microscopy images. *BMC Bioinformatics*, 11(1):372, jul 2010. ISSN 14712105. doi: 10.1186/1471-2105-11-372. URL `http://bmcbioinformatics.biomedcentral.com/articles/10.1186/1471-2105-11-372`.

[46] Arun Shivanandan, Aleksandra Radenovic, and Ivo F Sbalzarini. MosaicIA: An ImageJ/Fiji plugin for spatial pattern and interaction analysis. *BMC Bioinformatics*, 14(1):349, dec 2013. ISSN 14712105. doi: 10.1186/1471-2105-14-349. URL `http://bmcbioinformatics.biomedcentral.com/articles/10.1186/1471-2105-14-349`.

[47] Nicolás Peláez, Arnau Gavalda-Miralles, Bao Wang, Heliodoro Tejedor Navarro, Herman Gudjonson, Ilaria Rebay, Aaron R. Dinner, Aggelos K. Katsaggelos, Luís A. Nunes Amaral, and Richard W. Carthew. Dynamics and heterogeneity of a fate determinant during transition towards cell differentiation. *eLife*, 4, nov 2015. ISSN 2050-084X. doi: 10.7554/eLife.08924. URL `http://www.ncbi.nlm.nih.gov/pubmed/26583752http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC4720516`.

[48] Stéfan van der Walt, Johannes L. Schönberger, Juan Nunez-Iglesias, François Boulogne, Joshua D. Warner, Neil Yager, Emmanuelle Gouillart, and Tony Yu. scikit-image: image processing in Python. *PeerJ*, 2014. ISSN 2167-8359. doi: 10.7717/peerj.453.

[49] Nobuyuki Otsu. A Threshold Selection Method from Gray-Level Histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, 1979.

[50] Milica Bugarski, Maysam Mansouri, Axel Niemann, Aurélien Rizk, Philipp Berger, Urs Ziegler, Ivo F Sbalzarini, Pietro Incardona, and Grégory Paul. Segmentation and quantification of subcellular structures in fluorescence microscopy images using Squassh. *Nature Protocols*, 9(3):586–596, mar 2014. ISSN 1754-2189. doi: 10.1038/nprot.2014.037. URL `http://www.nature.com/articles/nprot.2014.037`.

[51] Vadim Zinchuk, Olga Zinchuk, and Teruhiko Okada. Quantitative Colocalization Analysis of Multicolor Confocal Immunofluorescence Microscopy Images: Pushing Pixels to Explore Biological Phenomena. *Acta Histochem. Cytochem.*, 40(4):101–111, 2007. ISSN 0044-5991. doi: 10.1267/ahc.07002. URL `http://joi.jlc.jst.go.jp/JST.JSTAGE/ahc/07002?from=CrossRef`.

[52] Paul T. Arsenovic, Carl R. Mayer, and Daniel E. Conway. SensorFRET: A Standardless Approach to Measuring Pixel-based Spectral Bleed-through and FRET Efficiency using Spectral Imaging. *Scientific Reports*, 7(1), dec 2017. ISSN 20452322. doi: 10.1038/s41598-017-15411-8. URL `http://www.nature.com/articles/s41598-017-15411-8`.

[53] Dahan Kim, Nikki M Curthoys, Matthew T Parent, and Samuel T Hess. Bleed-through correction for rendering and correlation analysis in multi-colour localization microscopy. *Journal of Optics*, 15(9), sep 2013. ISSN 20408978. doi: 10.1088/2040-8978/15/9/094011. URL `http://www.ncbi.nlm.nih.gov/pubmed/26185614http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC4501387`.

[54] Patrick D. McMullen, Richard I. Morimoto, and Luís A. Nunes Amaral. Physically grounded approach for estimating gene expression from microarray data. *Proceedings of the National Academy of Sciences*, 107(31):13690–13695, aug 2010. ISSN 1091-6490. doi: 10.1073/pnas.1000938107. URL `http://www.ncbi.nlm.nih.gov/pubmed/20643961http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2922271http://dx.doi.org/10.1073/pnas.1000938107{%}5Cnhttp://www.pnas.org/content/107/31/13690.abstract{%}5Cnhttp://www.pnas.org/`

`content/107/31/13690.ful`.

[55] Lisa Gaudette and Nathalie Japkowicz. Evaluation methods for ordinal classification. In *Lecture Notes in Computer Science*, volume 5549 LNAI, pages 207–210. Springer, Berlin, Heidelberg, 2009. ISBN 3642018173. doi: 10.1007/978-3-642-01818-3_25. URL `http://link.springer.com/10.1007/978-3-642-01818-3{_}25`.

[56] Thanh Minh Nguyen and Q. M.Jonathan Wu. Gaussian mixture-model-based spatial neighborhood relationships for pixel labeling problems. *IEEE Transactions on Systems, Man, and Cybernetics*, 42(1):193–202, feb 2012. ISSN 10834419. doi: 10.1109/TSMCB. 2011.2161284. URL `http://ieeexplore.ieee.org/document/5983453/`.

[57] Alexis Gambis, Pierre Dourlen, Hermann Steller, and Bertrand Mollereau. Two-color in vivo imaging of photoreceptor apoptosis and development in Drosophila. *Developmental Biology*, 351(1):128–134, mar 2011. ISSN 1095564X. doi: 10.1016/j.ydbio.2010.12.040. URL `https://www.sciencedirect.com/science/article/pii/S0012160610012960`.

[58] Pierre Dourlen, Clemence Levet, Alexandre Mejat, Alexis Gambis, and Bertrand Mollereau. The Tomato/GFP-FLP/FRT Method for Live Imaging of Mosaic Adult Drosophila Photoreceptor Cells. *Journal of Visualized Experiments*, (79):e50610, sep 2013. ISSN 1940-087X. doi: 10.3791/50610. URL `http://www.ncbi.nlm.nih.gov/pubmed/24084155http://www.pubmedcentral. nih.gov/articlerender.fcgi?artid=PMC3923918`.

[59] Yvette E Fisher, Helen H Yang, Jesse Isaacman-Beck, Marjorie Xie, Daryl M Gohl, and Thomas R Clandinin. FlpStop, a tool for conditional gene control in Drosophila. *eLife*, 6, 2017. ISSN 2050084X. doi: 10.7554/ eLife.22279. URL `http://www.ncbi.nlm.nih.gov/pubmed/28211790http://www. pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC5342825`.

[60] Joy S Wu and Liqun Luo. A protocol for mosaic analysis with a repressible cell marker (MARCM) in Drosophila. *Nature Protocols*, 1(6):2583–2589, jan 2007. ISSN 1754-2189. doi: 10.1038/nprot.2006.320. URL `http://www.ncbi.nlm.nih.gov/pubmed/ 17406512http://www.nature.com/doifinder/10.1038/nprot.2006.320`.

[61] Qingxiang Zhou, Scott J. Neal, and Francesca Pignoni. Mutant analysis by rescue gene excision: New tools for mosaic studies in Drosophila. *Genesis*, 54(11):589–592, nov 2016. ISSN 1526968X. doi: 10.1002/dvg.22984. URL `http://doi.wiley.com/10.1002/dvg. 22984`.

[62] Elleard Heffern, Norbert Perrimon, Amber M Hohl, Alberto del Valle Rodriguez, Chris Bakal, Marianne Bonvin, Anne Sustar, Didier Grunwald, Ruth Griffin, Christians Villalta, C-ting Wu, Jack R Bateman, Claude Desplan, Richard Binari, and

Gerold Schubiger. The twin spot generator for differential Drosophila lineage analysis. *Nature Methods*, 6(8):600–602, aug 2009. ISSN 1548-7091. doi: 10.1038/ nmeth.1349. URL `http://www.ncbi.nlm.nih.gov/pubmed/19633664http://www. pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2720837`.

[63] Hung Hsiang Yu, Chih Fei Kao, Yisheng He, Peng Ding, Jui Chun Kao, and Tzumin Lee. A complete developmental sequence of a Drosophila neuronal lineage as revealed by twin-spot MARCM. *PLoS Biology*, 8(8):39–40, aug 2010. ISSN 15449173. doi: 10.1371/ journal.pbio.1000461. URL `https://dx.plos.org/10.1371/journal.pbio.1000461`.

[64] Alexandru S. Denes, Emmanuel Caussinus, Markus Affolter, Oguz Kanca, and Anthony Percival-Smith. Raeppli: a whole-tissue labeling tool for live imaging of Drosophila development. *Development*, 141(2):472–480, jan 2013. ISSN 0950-1991. doi: 10.1242/ dev.102913. URL `http://dev.biologists.org/content/141/2/472.short`.

[65] Dafni Hadjieconomou, Shay Rotkopf, Cyrille Alexandre, Donald M Bell, Barry J Dickson, and Iris Salecker. Flybow: Genetic multicolor cell labeling for neural circuit analysis in Drosophila melanogaster. *Nature Methods*, 8(3):260–266, mar 2011. ISSN 15487091. doi: 10.1038/nmeth.1567. URL `http://www.nature.com/articles/nmeth.1567`.

[66] Stefanie Hampel, Phuong Chung, Claire E McKellar, Donald Hall, Loren L Looger, and Julie H Simpson. Drosophila Brainbow: a recombinase-based fluorescence labeling technique to subdivide neural expression patterns. *Nature Methods*, 8(3):253–259, mar 2011. ISSN 1548-7091. doi: 10.1038/nmeth.1566. URL `http://www.nature.com/ articles/nmeth.1566`.

[67] Thomas P Neufeld, Aida Flor A. De La Cruz, Laura A Johnston, and Bruce A Edgar. Coordination of growth and cell division in the Drosophila wing. *Cell*, 93(7):1183– 1193, jun 1998. ISSN 00928674. doi: 10.1016/S0092-8674(00)81462-2. URL `https: //www.sciencedirect.com/science/article/pii/S0092867400814622`.

[68] Michael Tworoger, Michele Keller Larkin, Zev Bryant, and Hannele Ruohola-Baker. Mosaic analysis in the Drosophila ovary reveals a common Hedgehog- inducible precursor stage for stalk and polar cells. *Genetics*, 1999. ISSN 00166731.

[69] Russell T Collins, Claudia Linker, and Julian Lewis. MAZe: A tool for mosaic analysis of gene function in zebrafish. *Nature Methods*, 7(3):219–223, mar 2010. ISSN 15487091. doi: 10.1038/nmeth.1423. URL `http://www.nature.com/doifinder/10.1038/nmeth. 1423`.

[70] Celia Muñoz-Jiménez, Cristina Ayuso, Agnieszka Dobrzynska, Antonio Torres-Mendéz, Patricia de la Cruz Ruiz, and Peter Askjaer. An efficient FLP-based toolkit for spatiotemporal control of gene expression in Caenorhabditis elegans. *Genetics*, 206(4):1763–1778, aug 2017. ISSN 19432631. doi: 10.1534/genetics.

117.201012. URL http://www.ncbi.nlm.nih.gov/pubmed/28646043http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC5560786.

[71] Wei Wang, Madhuri Warren, and Allan Bradley. Induced mitotic recombination of p53 in vivo. *Proceedings of the National Academy of Sciences*, 104(11):4501–4505, mar 2007. ISSN 0027-8424. doi: 10.1073/pnas.0607953104. URL http://www.ncbi.nlm.nih.gov/pubmed/17360553http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC1838630.

[72] Erik Meijering. Cell Segmentation: 50 Years Down the Road. *IEEE Signal Processing Magazine*, 29(5):140–145, sep 2012. ISSN 1053-5888. doi: 10.1109/msp.2012.2204190. URL http://ieeexplore.ieee.org/document/6279591/.

[73] Chikara Furusawa, Takao Suzuki, Akiko Kashiwagi, Tetsuya Yomo, and Kunihiko Kaneko. Ubiquity of Log-normal Distributions in Intra-cellular Reaction Dynamics. *Biophysics*, 1:25–31, 2005. ISSN 1349-2942. doi: 10.2142/biophysics.1.25. URL http://www.ncbi.nlm.nih.gov/pubmed/27857550http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC5036630http://arxiv.org/abs/q-bio/0503040.

[74] Jacob Beal. Biochemical complexity drives log-normal variation in genetic expression. *Engineering Biology*, 1(1):55–60, jun 2017. ISSN 2398-6182. doi: 10.1049/enb.2017.0004. URL https://digital-library.theiet.org/content/journals/10.1049/enb.2017.0004.

[75] Henry Teicher. Identifiability of Finite Mixtures. *The Annals of Mathematical Statistics*, 34(4):1265–1269, dec 1963. ISSN 0003-4851. doi: 10.1214/aoms/1177703862. URL http://projecteuclid.org/euclid.aoms/1177703862.

[76] M. Rosvall, D. Axelsson, and C. T. Bergstrom. The map equation. *European Physical Journal: Special Topics*, 2009. ISSN 19516355. doi: 10.1140/epjst/e2010-01179-1.

[77] Tomihisa Kamada and Satoru Kawai. An algorithm for drawing general undirected graphs. *Information Processing Letters*, 31(1):7–15, apr 1989. ISSN 00200190. doi: 10.1016/0020-0190(89)90102-6. URL https://www.sciencedirect.com/science/article/pii/0020019089901026?via{%}3Dihub.