Deep learning-enabled integrated measurements of immune signaling in primary human macrophages

Thesis by
Emily Chiu Laubscher

In Partial Fulfillment of the Requirements for the
Degree of
Doctor of Philosophy

CALIFORNIA INSTITUTE OF TECHNOLOGY
Pasadena, California

2024
Defended May 7, 2024

Emily Chiu Laubscher
ORCID: 0009-0008-0242-0507

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my advisor, David Van Valen, for the privilege to work in his group. I have grown into the researcher that I always hoped to be under your guidance and am forever grateful for the numerous opportunities you have given me. I would also like to thank my committee, Lu Wei, Shasha Chong, and Ellen Rothenberg, for your helpful comments and enthusiasm about my projects.

Thank you to all of the Van Valen lab members that I had the pleasure of collaborating with: Edward Pao, Morgan Schwartz, Ellen Emerson, Changhua Yu, Julie Wang, Nitzan Razin, Will Graf, Ross Barnowski, Danielle Gallandt, Ekta Patel, John Soro, Erick Moen, Tom Dougherty, Uriah Israel, and Rohit Dilip. I would like to thank a few colleagues in particular. My success in graduate school would not have been possible without Ed and Morgan's daily support. Ed's assistance with the practical elements of research was what enabled all of the research in this thesis. Morgan's willingness to brainstorm and attention to detail have gotten my projects unstuck a number of times, but more importantly, her friendship has been a constant source of happiness through graduate school. Ellen's collaboration was invaluable to the completion of this thesis, expediting the analysis for the integrated measurements data in the eleventh hour. I had a great time collaborating with Changhua on the ML-driven reporter design project, and his friendship as my desk neighbor has meant a lot to me. I would like to thank Julie for being a kind, generous, and talented collaborator for Polaris' gene decoding method. I would like to thank Nitzan for initiating the work on the deep learning model for spot detection that grew into Polaris. Polaris would not be the quality analysis tool it is today without her insightful work on its algorithms and attention to detail for its design. Finally, I would like to thank Will and Ross who were great mentors for the software engineering aspects of my research.

Thank you to my collaborators Jeffrey Moffitt and Rosalind J. Xu for providing data and invaluable feedback during the development of Polaris. Thank you to Yisong Yue for providing insightful feedback during the optimization of Polaris' performance.

My journey as a researcher started when I was an undergraduate in the Groves lab at UC Berkeley. I would not have pursued this path without the mentorship

and encouragement of Shalini Low-Nam, Steven Alvarez, Jenny Lin, and Meredith Triplett. I would especially like to thank Shal, whose contagious enthusiasm for science and tenacity as a researcher continues to inspire me today. I would also like to thank my friends, Xavier, Howard, and James whose camaraderie during this time kept science fun.

Thank you to my parents, Jon and Mabel, who encouraged me from a young age to pursue my curiosity for science and created the perfect environment for me to flourish academically. My ability to pursue graduate school was entirely enabled by your continued commitment to supporting my education. Thank you to my brother, Alex. Your talent for all things quantitative and relentless pursuit of your passions are a continuous source of inspiration for me.

Finally, I would like to thank my partner, David, for providing stability through my tumultuous time in graduate school. My path through graduate school would not have been possible without your unwavering support of my scientific career.

ABSTRACT

Examination of biological systems at the single-cell level reveals heterogeneity in both time and space. Single-cell temporal and spatial heterogeneity allow communities of cells to process noisy stimuli and perform complex tasks. We leveraged state-of-the-art imaging technologies to characterize cell-to-cell heterogeneity in responses to environmental stimuli to reveal mechanisms of information transmission. Fluorescent live-cell reporters enable real-time visualization of the activity state of cell signaling proteins. Signaling dynamics allow cells to translate information about environmental stimuli into cellular behaviors. Chapter 2 explores the variety of live-cell reporters designed to characterize the dynamic patterns of activity of key signaling pathways, and covers the development of two live-cell reporters. Spatial transcriptomics assays, on the other hand, excel at capturing heterogeneity in spatial gene expression patterns, which is often required to enable a tissue to perform complex functions. Chapter 3 details the development of Polaris, a deep learning-enabled analysis method for spatial transcriptomics data. Polaris is an assay-agnostic, turnkey solution for analyzing images from spatial transcriptomics experiments, minimizing the time and expertise require to extract biological insights. In Chapter 4, we pair dynamic measurements of live-cell reporters with a spatial transcriptomics measurement in an integrated imaging assay in primary human macrophages. This imaging assay revealed transcriptional sub-populations of cells with differing distributions of dynamic immune signaling responses and morphological states.

This work contributes a number of methodological developments, including live-cell reporter expression in primary human macrophages and deep learning-enabled spatial transcriptomics image analysis. Expression of live-cell reporters in primary macrophages will enable the investigation of environmental cues shape macrophages' cell state, which is highly plastic and shaped by external stimuli. Polaris expedites the analysis of this multi-modal imaging data set, extracting single-cell gene expression values without manual parameter tuning. However, Polaris' impact extends beyond the scope of this work to the broader spatial biology field as its spot detection and gene decoding capabilities generalize to data sets from a variety of sample types and imaging modalities. Finally, our paired dynamics-spatial transcriptomics imaging assay can be generally applied to characterize information transmission from environmental stimuli through signaling dynamics to the expression of down-

stream genes for a wide variety of signaling pathways in primary and immortalized cell types.

PUBLISHED CONTENT AND CONTRIBUTIONS

1. Laubscher, E. *et al.* Accurate single-molecule spot detection for image-based spatial transcriptomics with weakly supervised deep learning. *Cell Systems* **15,** 475–482.e6. ISSN: 2405-4712. doi:10.1016/j.cels.2024.04.006. http://dx.doi.org/10.1016/j.cels.2024.04.006 (May 2024).

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

LIST OF TABLES

Number *Page*

Chapter 1

INTRODUCTION

1.1 Motivation

The viability of organisms requires dynamic homeostatic control of internal state in response to continuously changing environmental conditions.[1] Therefore, the success of living systems depends on their ability to effectively identify and represent information about their environment to perform a situationally-appropriate response.[2]

Examination of biological systems at the single-cell level reveals heterogeneity in time and space, whereas population-level measurements mask temporal and spatial heterogeneity, preventing the exploration of how single-cell variability enables the function of the system. (Fig. 1.1) Single-cell heterogeneity enables cells to act as individuals with a collective goal, increasing the complexity and robustness of achievable functions.

In this work, we aim to characterize heterogeneous cellular responses to environmental stimuli, leveraging breakthroughs in imaging technologies. We focus on macrophages, which are highly plastic and have cell states largely shaped by environmental cues.[3–5] We capture temporal heterogeneity of the macrophage signaling response, asking how the signaling state of macrophages evolves over time in response to immune stimuli and how the dynamic response of individual macrophages varies at the single cell level. We also interrogate their spatial heterogeneity, investigating how a macrophage's morphology may indicate its internal state and how neighborhoods of cells influence each other. Taken together in an integrated imaging assay, we explore the mechanisms by which macrophages transfer information about environmental stimuli into cellular behaviors.

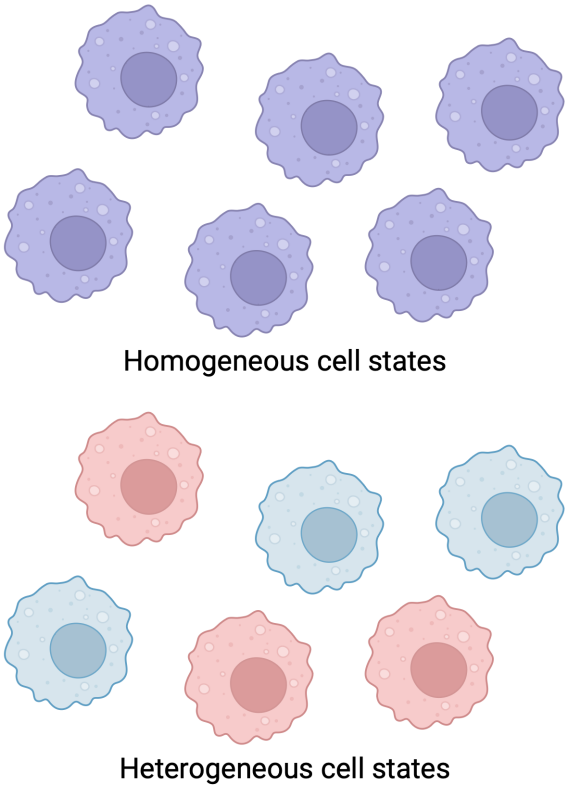


Figure 1.1: **Bulk measurements mask single-cell heterogeneity.** Cells in a homogeneous state at the single-cell level (upper) will be indistinguishable from cells in a heterogeneous state (lower) when observed at the population-level if the average state of both populations is the same.

1.2 Temporal heterogeneity stores and transmits information

Networks of interactions between signaling proteins allow cells to process the information from environmental stimuli into a biological response.[6] The architecture of these interactions allows signaling dynamics to encode environmental information and dictate the subsequent cellular behaviors. By studying signaling dynamics, we can explore how cells wade through noisy or overlapping stimuli to execute a stimulus-appropriate response.[7]

The nuclear factor-κB (NF-κB) transcription factor family is a critical signaling pathway that enables cells to respond to immune stimuli and is an illustrative example of this principle of cell signaling.[8,9] Complex molecular feedback mechanisms allow the NF-κB pathway to differentiate the identities and strengths of environmental stimuli.[7,10–14] For example, the NF-κB pathway encodes the identities of upstream stimuli, TNFα or LPS, with oscillatory activation or a single broad acti-

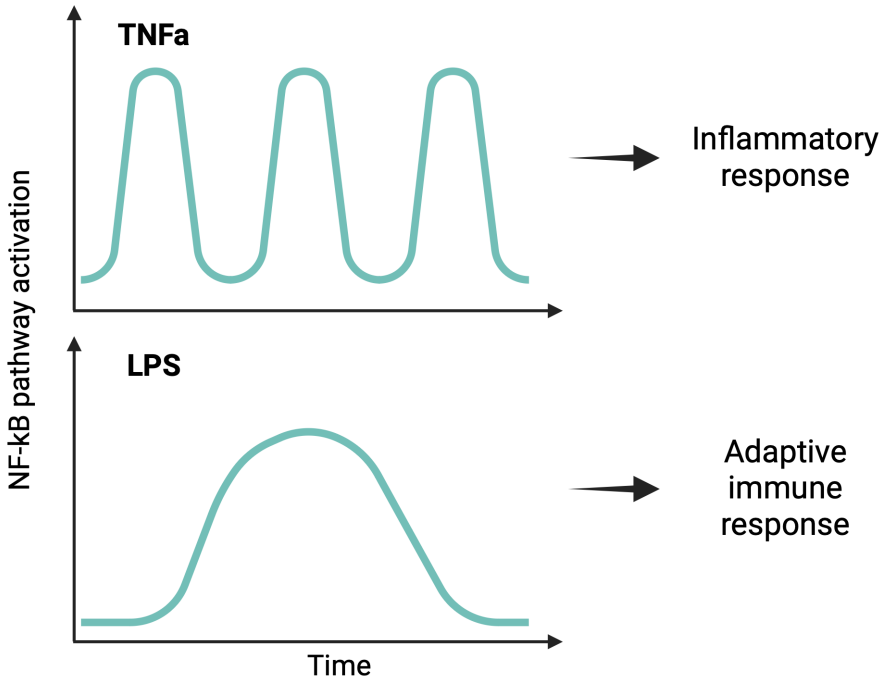


Figure 1.2: **Signaling dynamics dictate cellular behaviors.** NF-κB dynamics encode the identity of environmental stimuli and dictate the downstream cellular response. (Upper) In response to TNFα treatment, NF-κB activation will oscillate and trigger and inflammatory response. (Lower) In response to LPS treatment, NF-κB will undergo a single broad activation event and trigger an adaptive immune response. Figure adapted from Purvis et al. (2013).

vation event, respectively. (Fig. 1.2) These dynamic patterns of NF-κB activation elicit specific gene expression profiles, corresponding with stimulus-specific cellular responses.[2,15,16] Thus, cells are able to encode the identity and strength of immune stimuli in the dynamic pattern of NF-κB activation to elicit the stimulus-appropriate behavioral response.

Signaling dynamics are often heterogeneous in timing and strength at a single-cell level. Therefore, observation of dynamics at a population level can lead to misleading conclusions about the mechanisms of signal transduction occurring at the single-cell level.[7] For example, increasingly asynchronous oscillations can appear as a dampened oscillation when observed at the population-level. NF-κB dynamics are often oscillatory, but bulk measurements may obscure observation of this specific temporal pattern, apparent at the single-cell level. Alternatively, asynchronous digital activation events can appear as graded activation at the population-level. (Fig. 1.3)

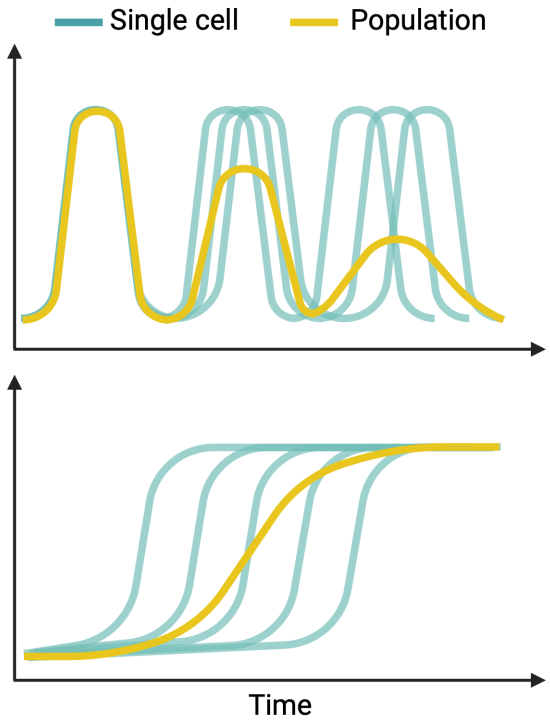


Figure 1.3: **Signaling dynamics vary at a single-cell level.** (Upper) Asynchronous single-cell oscillations in signaling pathway activation appear as dampened oscillations when observed at the population-level. (Lower) Asynchronous switch-like activation events appear as a graded activation when observed at the population-level. Figure adapted from Purvis et al. (2013).

Cell-to-cell heterogeneity in the NF-κB response is commonly reported, and in fact, has been shown to serve a functional role at the multicellular level. In response to TNF-α treatment, NF-κB activation has been shown to be digital at the single-cell level with stimulus strength encoded in the fraction of a population cells that respond.[17] Moreover, intrinsic and extrinsic noise in NF-κB signaling has been shown to drive the robustness of the cell population's response to a wider range of stimuli.[18]

For these reasons, observation of cell signaling at the single-cell level and dynamically through time is critical to the characterization of the mechanisms by which cells respond to environmental stimuli. Fluorescent live-cell reporters are a critical imaging tool for visualizing real-time signaling at the single-cell level. In Chapter 2, we explore a variety of fluorescent live-cell reporters currently available for visualization of real-time signaling pathway activation at the single-cell level. We also demonstrate two vignettes of live-cell reporter design and development.

1.3 Spatial heterogeneity enables complex functions

Technologies for profiling cells and tissues with spatial context are rapidly being developed in the emerging field of spatial biology. Spatial-"omics" methods aim to characterize the patterns of spatial organization in living systems that enable complex tissue function or drive pathological dysfunction.[19,20] These spatial-omics measurements can be targeted to characterize many categories of biological molecules, and the spatial transcriptome, or the collection of mRNA transcripts in a sample, is of particular interest to this work. Characterization of a cell's transcriptome yields a transient snapshot of its gene expression profile and offers a glimpse into the active processes shaping a cell's signaling state at a given point in time.[21] There are a number of emerging technologies for spatial transcriptomic characterization of cells and tissues.[22–24] These methods fall into two broad categories: (1) sequencing-based methods with a variety of capture strategies, such as spatial transcriptomics and Slide-seq,[25–28] and (2) image-based methods with a variety of multiplexed barcode strategies, such as in-situ sequencing, MERFISH, osmFISH, seqFISH, and splitFISH.[29–33]

Spatial transcriptomics methods excel at characterizing heterogeneous tissue microenvironments, consisting of functional structures of multiple cell types. Tumor microenvironments have been of particular research interest, because of their complex architectures and translational value.[34] A number of different spatial transcriptomics methods have offered critical insights into the organization of tumor microenvironments, including spatial methods for the whole-transcriptome and targeted in-situ methods.[35] The insights include interactions at the tumor boundary[36] or survival and therapy response predictions based on edge architecture characteristics,[37] amongst others. As another example of complex tissue architectures that have been explored with spatial transcriptomics, MERFISH has been applied to characterize the complex tissue-level organization of the mouse brain, elucidating the cellular structures that enable function.[38,39] These measurements identified thousands of transcriptionally distinct subpopulations of multiple cell types and quantified the cell-type composition of each specific brain region. This atlas of the spatial transcriptome of the mouse brain is an invaluable point of comparison for elucidating mechanisms of dysfunction in disease-state samples. Capturing the tissue-level organization of the tumor microenvironment and brain samples would have been impossible with RNA sequencing methods that require tissue dissociation, so spatial transcriptomics methods enable characterization of gene expression with multicellular spatial organization.

On the other hand, tissues often consist of structures of a single cell type in differing cell states, which allows the tissue to perform complex functions with a single cell type. Spatial transcriptomics methods are ideal for capturing this spatial heterogeneity in cell state. A clear example of structures of a single cell type in a spectrum of cell states to perform a collective function can be found in liver tissue. The liver consists of hexagonal units of hepatocytes, called lobules. The hepatocytes have a gradient of cell states from the periphery of the lobule to the center, varying in metabolic and immune signaling states. This complex architecture enables the liver tissue to perform complex functions.[40,41] Investigation of the link between the tissue-level organization of cell states in the liver and its function has been enabled by recent breakthroughs in spatial-omics technologies, revealing metabolic zonation, mechanisms of immunity, and morphological structures that could not have been revealed with sequencing technologies.[42,43]

The collaboration of spatially heterogeneous cell types and cell states requires cell-to-cell communication to enable function. Because spatial transcriptomics methods can often visualize the location of transcripts of interest with high spatial resolution, they are well-poised to investigate questions of intra-cellular signaling. For example, seqFISH+ has been used to quantify cell-to-cell communication between astrocytes, neurons, microglia, and endothelial cells in mouse olfactory bulb and visual cortex tissue samples. The high spatial resolution allowed the authors to identify sub-cellular spatial patterns of transcript localization and the large panel size of seqFISH+ allowed this study to identify subpopulations of each cell type with specific relative spatial distributions in the tissue.[32,44] These insights could not be learned from sequencing methods that require dissociation of the tissue sample and destroy sub-cellular and tissue-level organization information.

Because of the power of spatial transcriptomics methods to elucidate multi-cellular mechanisms of function, we sought to create an analysis pipeline to expedite the extraction of biological insights from spatial transcriptomics images. Because analysis of these large, complex datasets is often highly manual and time-intensive, we created a deep-learning enabled method, called Polaris, that is assay-agnostic and minimizes the time and expertise required for analysis. In Chapter 3, we detail the construction and validation of Polaris as a turnkey solution for the quantification of spatial transcriptomics images.

1.4 Integrated measurement of signaling dynamics and spatial transcriptomics

We sought to integrate both of these powerful imaging technologies, live-cell imaging of fluorescent reporters and spatial transcriptomics, to investigate the mechanisms of immune signaling pathways in primary human macrophages. Previous works have linked dynamic NF-κB activation to the expression of its target genes to understand how the pathway encodes environmental information to be decoded into cellular behaviors. The molecular feedback mechanisms of NF-κB signaling that allow for discrimination of concentrations of immune stimuli have been linked to population-level expression of target genes.[2,15] However, at the population level, these measurements cannot observe the information processing occurring in the NF-κB pathway at the single-cell level. Furthermore, it has been demonstrated that the fold-change of NF-κB activation corresponds with the expression of target genes at the single-cell level, but this work utilized a low-throughput spatial transcriptomics method, smFISH, to measure gene expression, sharply limiting the size of the panel of measured genes.[45] Dynamic observation of NF-κB activation has also previously been paired with single-cell RNA-sequencing. This measurement demonstrated subpopulations of macrophages have different dynamic NF-κB responses to immune stimuli, linked to specific patterns of target gene expression. Although this work demonstrates the link between NF-κB dynamics and transcriptome-level gene expression, it lacks the sub-cellular and cell-to-cell spatial sample organization that spatial transcriptomics measurements could capture.[16] More recently, increasingly nuanced mathematical models have been developed to characterize the mechanisms of information transmission via NF-κB signaling, but the accuracy of the outputs of this model are inherently limited by the abundance and quality of paired dynamics and target gene expression data.[46]

With this in mind, we aimed to develop an imaging assay to measure immune signaling in primary human macrophages at the single-cell level that addresses the limitations of previous characterizations of NF-κB signaling. We proposed to use characterize NF-κB signaling dynamics with time lapse imaging of live-cell reporters and measure the end-point spatial transcriptome with seqFISH in primary macrophage cells. Importantly, a paired measurement of this flavor has not been extended to primary human macrophages, which have much greater physiological relevance to macrophages *in vivo*. In Chapter 4, we demonstrate the application of integrated measurements of signaling dynamics and spatial transcriptomics to the characterization of immune signaling in primary human macrophages.

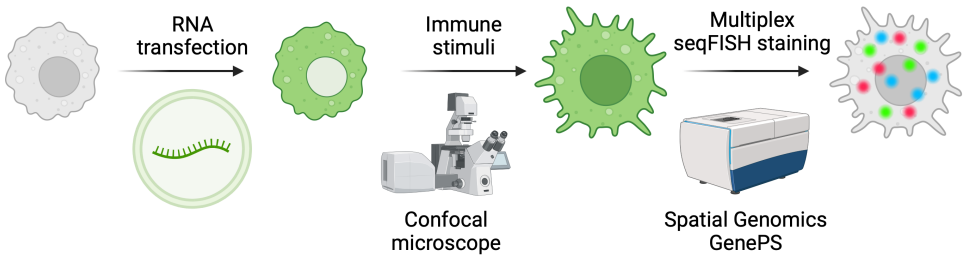


Figure 1.4: **Integrated measurements in primary human macrophages.** Fluorescent live-cell reporters are expressed in primary human macrophages by RNA transfection. The dynamic response of these cells to immune stimuli is captured with time-lapse imaging. Then, end-point seqFISH is performed to measure the spatial transcriptome of the cells.

Our proposed measurement can observe information transmission from environmental stimulus to gene expression profile through the NF-κB signaling pathway at the single-cell level. The single-cell resolution of our measurement is fundamental to its ability to investigate the capability of individual cells to process information about their environment. This experimental workflow can ultimately be applied as a general framework to investigate information processing at the single-cell level. This work demonstrates its applicability to classically hard-to-transfect cell types, increasing its theoretically scope. Moreover, it demonstrates the types of insights that can be yielded when pairing a dynamics measurement with a multiplexed spatial transcriptomics measurement.

Chapter 2

CHARACTERIZATION OF METABOLIC AND IMMUNE SIGNALING DYNAMICS WITH FLUORESCENT LIVE-CELL REPORTERS

2.1 Overview of select fluorescent live-cell reporters

Information is transduced through signaling pathways by a wide variety of signaling events. Commonly, a protein substrate is phosphorylated by a kinase in response to an upstream stimulation event. (Fig. 2.1a) This phosphorylation event confers a negative charge to the target protein that can trigger a number of molecular events. Phosphorylation can trigger the target protein to interact with other macromolecules, such as genomic DNA or proteins. (Fig. 2.1b,e) It can also alter the localization pattern of a target protein, leading to trafficking to sub-cellular compartments such as the plasma membrane or nucleus. (Fig. 2.1d) Alternatively, phosphorylation can mark a protein molecule for degradation by the ubiquitin-proteasome pathway. (Fig. 2.1c) Finally, phosphorylation can trigger a conformational change in a target protein, which can, for example, modulate its enzymatic activity. (Fig. 2.1e) A wide variety of live-cell reporters have been engineered to capitalize on all of these molecular events, amongst others, to read out the signaling activity of proteins of interest in real time.[47,48] These reporter methods have the ability to quantify information storage and transmission in signaling pathways at the single-cell level, making them invaluable to the characterization of cellular strategies for responding to environmental stimuli.

The development, selection, and implementation of live-cell reporters should be rigorously considered for accurate quantification of signaling pathway activation.[49] Live-cell reporter development requires demonstration of the specificity and accuracy of the read out of the reporter. Often this process involves demonstration of activity and comparison to existing quantification methods in response to known upstream stimuli. Selection of the appropriate existing live-cell reporter for a research question should involve consideration of dynamic range, spectral bandwidth, and mechanism of action, amongst others. Such considerations will determine the veracity of the reporter's read out of signaling pathway activity and capability to be multiplexed with other measurements. Finally, implementation of an imaging

assay using live-cell reporters requires attention to the ideal mode of expression and imaging conditions. Development of an imaging assay that implements the live-cell reporter in its ideal conditions will offer the best quantification of signaling pathway activity. In this chapter, we discuss the advantages and limitations of a variety of live-cell reporters to demonstrate their ranging applications. We also detail the ongoing development of two live-cell reporters to illustrate the process of developing reporters to address specific research questions.

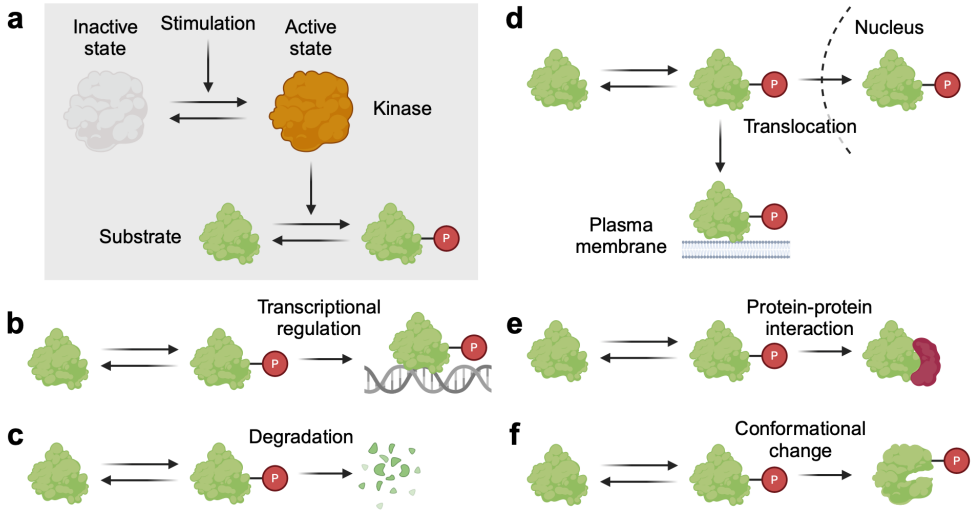


Figure 2.1: **Signaling events for reporters.** See section 2.5 for caption.

Fluorescent fusion proteins

The discovery and continued development of fluorescent proteins (FPs) has revolutionized our ability to characterize spatial and temporal heterogeneity in living systems.[50–53] FPs can be fused to proteins of interest to visualize changes in their abundance and localization over time, making them an ideal tool for the quantification of a number of signaling events that involve translocation or expression level changes. However, there are a number of important considerations when expressing fluorescent fusion proteins, including chromophore maturation rate, aggregation, photobleaching rate, and brightness of individual fluorescent proteins, and for pairs of fluorescent proteins, spectral overlap.[49,54]

Fluorescent fusion proteins are an ideal imaging method for visualize proteins that undergo a significant shift in sub-cellular localization upon signaling pathway activation. A key category of proteins that undergo such a shift upon activation are transcription factors, which are commonly sequestered in the cytoplasm in their

inactive state and translocate to the nucleus upon upstream signaling activity. An important consideration when designing a fluorescent fusion protein is the tendency of some fluorescent proteins to aggregate when over-expressed. Aggregation can modulate the local concentration levels of signaling proteins in sub-cellular regions, which can affect the propensity of certain chemical reaction steps in signaling networks. Selection of monomeric fluorescent proteins can minimize this effect. Additionally, the endogenous expression levels of some signaling proteins can be relatively low, requiring bright fluorescent proteins to minimize the effect of imaging noise during visualization. When expressing fluorescent fusion proteins, the potential for protein over-expression to perturb the endogenous balance of a cell's signaling state should be carefully considered. Previous works have introduced loss-of-function mutations to transcription factors to allow safe over-expression. For example, the F3aN400-Venus reporter for FoxO3 contains all phospho-residues for interaction with upstream signaling proteins, but contains a loss-of-function mutation that prevents DNA binding and thus modulation of FoxO3 target gene expression, minimizing the effect of reporter expression on the cell's signaling state.[55]

Fluorescent fusion proteins also excel at visualizing dynamic changes in protein abundance. The fluorescent fusion construct can be knocked in to a target genome as an entirely exogenous locus, or the fluorescent protein can be introduced to the native locus, offering flexibility in the method of expression for different applications. For example, a fluorescent protein sequence can be inserted into the reading frame for a target gene of a transcription factor of interest. Such a strategy would visualize the induction of gene expression in response to pathway activation, but it also allows for visualization of the degradation of a specific protein. (Fig. 2.1c) The rate of chromophore maturation is an important consideration when using a fluorescent fusion protein to visualize the expression of genes of interest. Selection of a fluorescent protein with a relatively fast maturation rate will minimize the delay between protein translation and ability to be visualized, offering more a faithful read out of the concentration of the expressed protein.

Live-cell reporters for signaling network activity

In recent years, the toolbox of available live-cell reporters has grown significantly, expanding the range of cell signaling questions that can be investigated.[56,57] These live-cell reporters are engineered constructs that are exogenously expressed in a living system to visualize its single-cell signaling state in real time. A wide variety of reporter architectures have been developed with different advantages for their specific application.

Förster resonance energy transfer (FRET) involves the transfer of energy from a donor fluorophore to an acceptor fluorophore, the efficiency of which is highly dependent on the proximity of the donor and acceptor.[58] Therefore, FRET is often used in live-cell reporters designed to measure the proximity of two peptide constructs.[59–62] The proximity of the peptides can be driven by a protein-protein interaction event following a post-translational modification like phosphorylation or acetylation.[63,64] (Fig. 2.1e) FRET reporters can also be designed to read out the conformational changes of a protein of interest, strategically fusing two FPs to the protein such that their proximity is determined by the conformational state of the protein.[65,66] (Fig. 2.1f)

Because the read out of FRET reporters depends solely on intra-molecular processes, they excel at visualizing signaling activity with sub-cellular resolution. FRET-based A-Kinase Activity Reporter (AKAR) targeted to sub-cellular compartments reveal differential basal and agonist-induced activity of cAMP-dependent protein kinase (PKA).[67,68] FRET reporters are also uniquely suited to investigating structural dynamics, such as protein conformational dynamics or complex formation. For example, single-molecule FRET has been used recently to study agonist-induced organizational dynamics of G protein-coupled receptors in live cells.[69]

Kinase translocation reporters (KTRs) are a single color, translocation-based live-cell reporter architecture that addresses some of the limitations of FRET reporters. These reporters consist of optimized NES and NLS sequences that regulate cytoplasmic to nuclear shuttling upon phosphorylation. A kinase docking sequence recruits the kinase of interest, which then processively phosphorylates sites on the NES and NLS sequences.[70] Therefore, these reporters exploit native charge-based nuclear translocation machinery to read out the activity of a target kinase (Fig. 2.1d). Therefore, they are best applied to study kinases that trigger translocation events, because they do not report kinase activity with sub-cellular spatial resolution. KTRs have been rigorously characterized in mathematical models corresponding their read

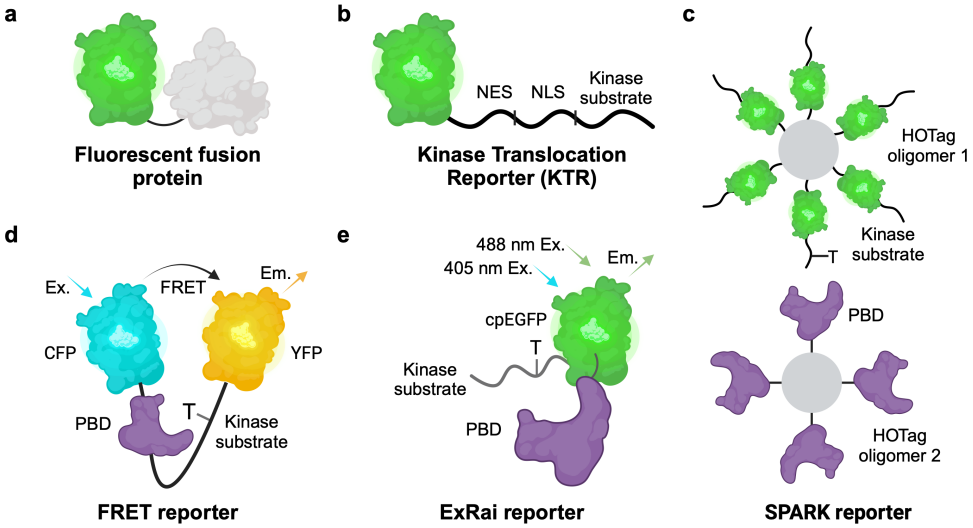


Figure 2.2: **Survey of live-cell reporter architectures.** See section 2.5 for caption.

out to kinase activity and are readily expressed in cell lines of interest.[71] As a single color reporter, KTRs occupy a small range of the visible spectrum and are very amenable to multiplexing.[70]

The fidelity and multiplexing capabilities of KTRs have led to a number of impactful discoveries of properties of signaling networks. As an example, JNK KTR-mCerulean3 was expressed in the same cells as p65-dsRed to reveal coordination between the JNK and NF-κB signaling pathways in response to a variety of immune stimuli. In another example, co-expression of ERK KTR, JNK KTR, and p38 KTR revealed mechanisms of cell cycle regulation by interactions between the three kinases.[70]

Optimization for spatial resolution with a single fluorophore

However, FRET reporters and KTRs have some key short-comings that limit the breadth of their applications. Although FRET reporters excel at reading out the proximity of pairs of peptides, they are often difficult to multiplex with other live-cell reporters because their pairs of fluorescent proteins occupy a large portion of the visible spectrum. Furthermore, their reliance on the sensitive FRET process as a read out can limit their dynamic range. KTRs, on the other hand, are limited by their inability to read out kinase activity with sub-cellular resolution as they use translocation to visualize kinase activity. These limitations of FRET reporters and KTRs have been addressed by other reporter architectures.

Live-cell reporters for intracellular calcium address these limitations with circularly permuted fluorescent proteins (cpFPs). These reporters consist of two peptide fragments fused to the cpFP. Upon Ca^{2+} binding, the two fragments form a stable complex, altering the conformation of the cpFP. In the bound form, the cpGFP has higher quantum yield than the unbound form, leading to an increase in fluorescence intensity from the reporter dependent on the intracellular concentration of Ca^{2+}.[72,73] This reporter architecture establishes a pattern that has been iteratively improved by subsequent live-cell reporters: an intracellular signaling event triggers a molecular binding event that can be read out by changes in patterns of fluorescent intensity.

Excitation ratiometric (ExRai) reporters extend previous work with cpGFP reporters to read out kinase activity. As will be discussed further in section 2.3, these reporters rely on protein-protein interaction in response to kinase activity to alter the excitation spectrum of cpGFP in a phosphorylated reporter molecule. (Fig. 2.1e)[74] The most recent iterations of this reporter demonstrate much greater dynamic range than FRET reporters, increasing their applicability to a wide range of imaging assays. Furthermore, with a single fluorescent protein, they are more amenable to multiplexing than FRET reporters.[75]

The ExRai reporter architecture has been generalized to a variety of color variations and applied to a number of target kinases.[74] Mutations to the cpGFP in ExRai reporters increased its dynamic range, allowing it to be used for *in vivo* imaging of protein kinase A (PKA) activity in the brains of awake mice.[75] Furthermore, its compact design allows it to be targeted to organelles, such as mitochondria and lysosomes, for specific, a sub-cellular read out of target kinase activity. Observation of ExRai AMPKAR targeted to individual organelles has revealed unique kinetic profiles of activation in mitochondria and lysosomes and identified cytoplasmic to nuclear shuttling of AMPK.[76]

Separation of Phases-based Activity Reporter of Kinase (SPARK) reporters utilize a phase separation mechanism to reporter kinase activity with sub-cellular resolution with a single fluorescent protein.[77] This reporter architecture utilizes the protein-protein interaction motif to drive phase separation of phosphorylated reporter molecules (Fig. 2.1e). As will be highlighted in section 2.2, SPARK reporters have a number of key advantages. They have been reported to read out kinase activity with many-fold greater dynamic range than FRET reporters, which makes them feasible to image in tissue samples and in high-throughput screens. They occupy a smaller range of the visible light spectrum than ratiometric reporters like FRET and ExRai reporters, increasing their capacity for multiplexing.

SPARK reporters for a number of target kinases have been created and applied to novel observation of signaling pathway dynamics. PKA SPARK revealed heterogeneous PKA dynamics in response to β-adrenergic receptor and adenosine receptor activation. ERK SPARK has been used to read out spatiotemporal ERK activity in Drosophila tissues *in vivo*. Challenging *in vivo* imaging of SPARK reporters is enabled by its excellent dynamic range in a single fluorescence channel.[77] Finally, ataxia telangiectasia mutated (ATM) SPARK has been used to observe the spatiotemporal DNA damage response. ATM SPARK was employed in a high-throughput screen that revealed enzyme and small-molecule inhibitors of ATM. The wide dynamic range of SPARK reporters allowed for the low magnification imaging required for the high-throughput screen.[78]

We considered the characteristics of these currently available reporter architectures when designing reporters relevant to the scope of this work. The following vignettes detail the design and validation of two live-cell reporters targeted to visualize the activity of metabolic and immune signaling pathway for the ongoing goal to characterize immuno-metabolic signaling in primary macrophages.

2.2 Vignette 1: Development and validation of AMPK SPARK

Construction of a novel live-cell reporter for AMPK

AMP-activated protein kinase (AMPK) is a central metabolic regulator, controlling cellular processes such as glucose metabolism, autophagy, and mitochondrial function in response to cellular ATP levels. AMPK is critical to a cell's ability to balance catabolic and anabolic processes in response to metabolic stress. Its activity is modulated by two independent upstream signaling pathways: the CAMKK2-AMPK pathway and the LKB1-AMPK pathway.[79–81] The CAMKK2-AMPK pathway mod-

ulates the cell's metabolic state in response to spikes in intracellular calcium levels, and has been implicated in fatty acid metabolism and adipocyte differentiation.[82] The LKB1-AMPK pathway is a central energetic checkpoint, regulating ATP production in response to the availability of environmental nutrients.[83] Its role as a fundamental metabolic regulator makes AMPK a target of basic and translational research. We are particularly interested in the mechanisms by which AMPK regulates cellular behaviors in response to dynamic changes in a cell's environment. We aimed to improve upon the leading live-cell reporter for AMPK at the time of conception of this project, AMPKAR.[84] As a FRET-based reporter, AMPKAR had two main limitations that could be alleviated with an alternative reporter architecture: (1) limited dynamic range, and (2) limited multiplexing capabilities, because FRET reporters rely on a sensitive photon transfer step between two fluorescent proteins.[60] We sought a live-cell reporter architecture that utilizes a single fluorescent protein and achieves a greater dynamic range than FRET-based reporters.

Overview of SPARK reporters

SPARK reporters are a reporter architecture that reads out kinase activity at subcellular spatial resolution with a single fluorescent protein. The reporter molecules freely diffuse when the kinase of interest is not active. However, upon phosphorylation with the kinase of interest, the reporter molecules phase separate into clusters (Fig. 2.3). The sharp increase in local reporter concentration allows the clusters to be detected above the background fluorescence intensity of the cell. The dynamic activity of the kinase of interest is quantified by measuring fraction of the total fluorescence intensity of the clusters to the total fluorescence intensity of the cell.

To achieve this, SPARK reporters consist of two constructs whose interaction drives clustering in response to kinase activity. The first construct contains a peptide substrate for the kinase of interest fused to a fluorescent protein, EGFP and a homo-oligomerization domain (HOTag). The HOTag is a computationally-designed de novo alpha helical domain that forms oligomers spontaneously.[85,86] In the SPARK construct, the HOTag drives homo-hexamer formation of the kinase substrate-EGFP-HOTag construct. The second construct contains a phospho-binding domain (PBD), which binds the phosphorylated kinase substrate, and an orthogonal HOTag, which drives homo-tetramer formation of the second SPARK construct. The two HOTags have been shown to oligomerize specifically and orthogonally, such that mixed oligomers of the two different HOTags will not form.[77]

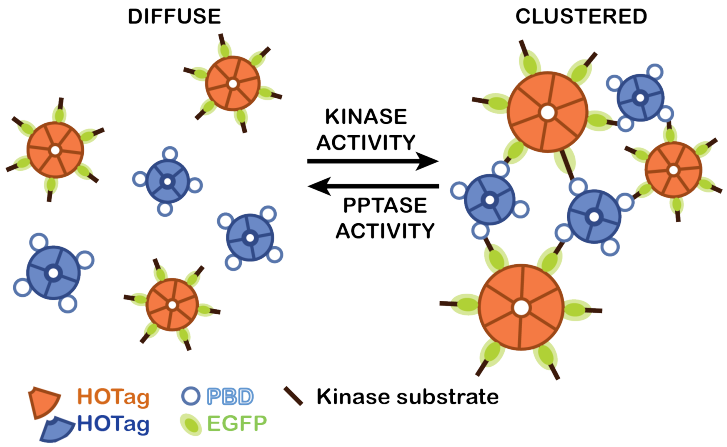


Figure 2.3: **SPARK reporters cluster in response to kinase activity.** HOTag1 (orange) forms homo-hexamers containing the kinase substrate and the fluorescent protein. HOTag2 (blue) forms separate homo-tetramers containing the phospho-binding domain (PBD). Upon phosphorylation by the kinase of interest, multivalent interactions form between the two homo-oligomers, leading to clustering of the reporter.

Therefore, upon phosphorylation of the kinase substrate, the PBDs will bind to the kinase substrates, which will drive the multivalent interactions between the two types of homo-oligomer, resulting in phase separation of the reporter. It has been demonstrated that these interactions can be reversed when the kinase substrate is de-phosphorylated by native phosphatases, causing dissociation of the PBD from the kinase substrate. Previous work demonstrated that this SPARK reporter architecture could achieve many-fold changes in reporter read out range in kinase activity, which is much greater than the $\sim$ 30% dynamic range expected from FRET-based reporters.[77]

For these reasons, we designed a novel SPARK reporter to read out the dynamic activity of AMPK, using the previously published PKA SPARK reporter as a template.[77] The construction of this reporter is benefited by a number of design choices. The kinase substrate for AMPK (RRVA**T**LVDL) was derived from its native substrate, acetyl-CoA carboxylase (ACC), and has been shown previously to be specifically phosphorylated by AMPK.[84] FHA1 domain binds specifically to phospho-serine and phospho-threonine residues, making it an ideal candidate as a PBD for this reporter.[87] The AMPK substrate was selected to be optimal for FHA1 binding, with the FHA1-preferred LVD sequence following the phospho-threonine residue.[88] The two HOTag constructs were selected out of a number of candidate combinations,

because previous work demonstrated their ability to drive phase separation with the greatest sensitivity of all tested pairs. Finally, the T2A sequence encodes a self-cleaving peptide, allowing two separate protein construct to be expressed in a single open reading frame, ensuring that the two constructs are expressed in the desired 1:1 stoichiometry.[89]

Figure 2.4: **AMPK SPARK construct.** AMPK substrate is a verified substrate of AMPK activity derived from ACC. EGFP is enhanced green fluorescent protein. HOTag1/2 are homo-oligomerization domains. T2A is a self-cleaving peptide sequence. FHA1 is a phospho-binding domain.

Validation of AMPK SPARK

To assess the viability of the reporter, we measured the response of AMPK SPARK to known AMPK agonists in HeLa cells. As previously mentioned, there are two major signaling pathways that activate AMPK, the CAMKK2-AMPK pathway and the LKB1-AMPK pathway, so we will test agonists of both pathways. Under homeostatic culture conditions, we expect low intra-cellular calcium levels and low metabolic stress, so therefore, we expect minimal basal AMPK activity from both upstream pathways.[90]

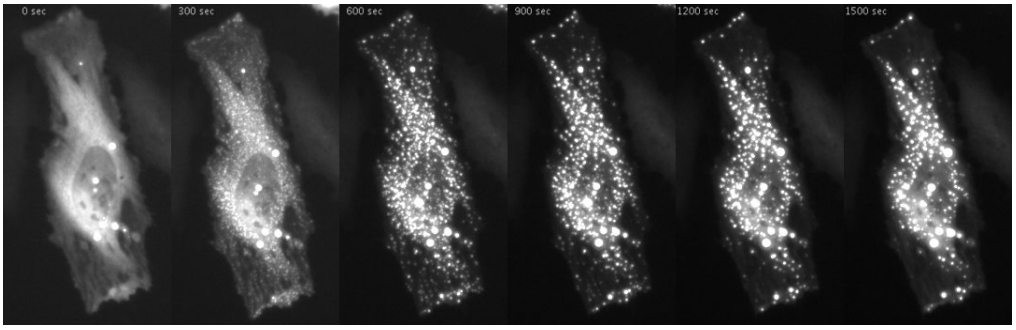


Figure 2.5: **Montage of AMPK SPARK response to ionomycin treatment.** Example HeLa cell transiently expressing AMPK SPARK was dosed with 2μM ionomycin between the first and second frames. The reporter clusters in response to AMPK activity.

Ionomycin is a membrane-permeable calcium ionophore that causes extracellular calcium ions to flux into the cytoplasm.[91] This calcium flux activates CAMKK2, a calcium-dependent protein kinase, to phosphorylate and thus activate AMPK.[92,93] As expected, we observed low basal AMPK activity as read out by AMPK SPARK

and a sharp increase in AMPK activity in response to 2μM ionomycin treatment, demonstrating AMPK activation in response to a flux in intracellular calcium (Fig. 2.5).

The second tested agonist of AMPK was a glucose analog, 2-deoxy-D-glucose (2-DG), which inhibits glucose metabolism. 2-DG has a hydrogen at the 2' carbon where glucose has a hydroxyl group, preventing its conversion to fructose and leading to the accumulation of glucose metabolism intermediates.[94] Exposure to an excess of 2-DG at concentrations greater than glucose concentration in the culture media prevents efficient production of ATP, triggering AMPK signaling by energy-sensing kinase LKB1 and direct interaction of AMP/ADP with AMPK.[83] AMPK SPARK clusters in response to the 2-DG treatment, demonstrating AMPK activation in response to glycolysis inhibition. We have therefore demonstrated that AMPK SPARK reads out AMPK activity triggered by both major upstream signaling pathways, CAMKK2-AMPK and LKB1-AMPK (Fig. 2.6).

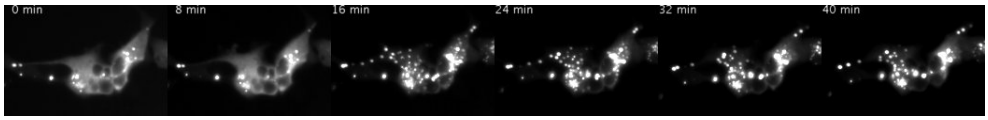


Figure 2.6: **Montage of AMPK SPARK response to 2-DG treatment.** Example HEK293T cell transiently expressing AMPK SPARK was dosed with 20mM 2-DG between the first and second frames. The reporter clusters in response to AMPK activity.

To demonstrate that the mechanism of reporter clustering relies specifically on binding of the PBD to the phosphorylated AMPK substrate, the phospho-site of the AMPK substrate (T12) was mutated to an alanine residue, using site-directed mutagenesis. This construct was also transiently transfected into HeLa cells, using the same protocol as was used for the wild type AMPK SPARK construct. The AMPK SPARK T12A mutant did not demonstrate a clustering response after treatment with 2μM ionomycin, demonstrating that the reporter clustering is driven by phosphorylation of the threonine residue in the AMPK substrate. (Fig. 2.7)

Quantification of SPARK reporter clustering

To quantify the degree of reporter clustering, we initially followed the absolute intensity thresholding method of the original work.[77] Because cluster formation increases the local concentration of the reporter, the locations of reporter clusters can be identified by setting an absolute intensity threshold, above which a pixel is con-

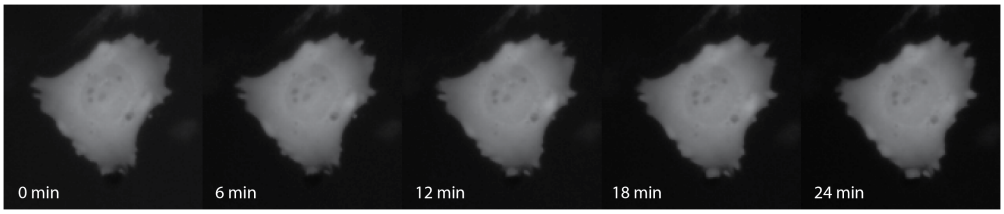


Figure 2.7: **Montage of AMPK SPARK T12A mutant response to ionomycin treatment.** Example HeLa cell transiently expressing AMPK SPARK T12A was dosed with 2μM ionomycin between the first and second frame. The reporter does not cluster in response to AMPK activity.

sidered to be a part of a cluster. This threshold was set manually for each individual cell because of cell-to-cell variation in reporter expression level. The whole-cell area was similarly segmented by absolute intensity thresholding (Fig. 2.8a-c). The intensity of the pixels inside the masked cluster and cell areas was integrated to calculate the total amount of reporter in each masked area. The integrated cluster intensity was normalized by the total whole-cell intensity to calculate the fraction of the reporter that is clustered, which we take as a read out of kinase activity. This method was applied to time lapse images to quantify AMPK SPARK clustering over time (Fig. 2.8d).

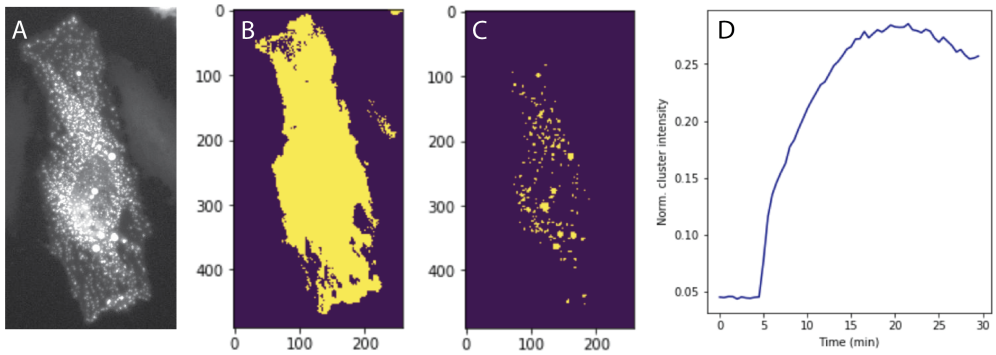


Figure 2.8: **Image analysis of SPARK reporter with intensity thresholding.** (A) Image of clustered AMPK SPARK in a HeLa cell in response to ionomycin. (B) Cell area segmented with manually-tuned absolute intensity thresholding. (C) SPARK cluster area segmented with manually-tuned absolute intensity thresholding. (D) Fraction of clustered reporter, reading out AMPK activity, through time.

This intensity thresholding method was used to analyze the response of a population of AMPK SPARK-expressing cells to ionomycin stimulation, yielding time traces of normalized SPARK cluster intensity (Fig. 2.9). The traces reveal heterogeneity at the single-cell level in the dynamic AMPK activities. The cells have heterogeneous

initial cluster intensities before ionomycin stimulation, demonstrating the variation in basal metabolic state at the single-cell level. Furthermore, the cells react heterogeneously to ionomycin stimulation, demonstrating either external noise in dosing efficiency or internal noise in calcium signaling states. These results illustrate the need to measure AMPK signaling dynamics at the single-cell level.

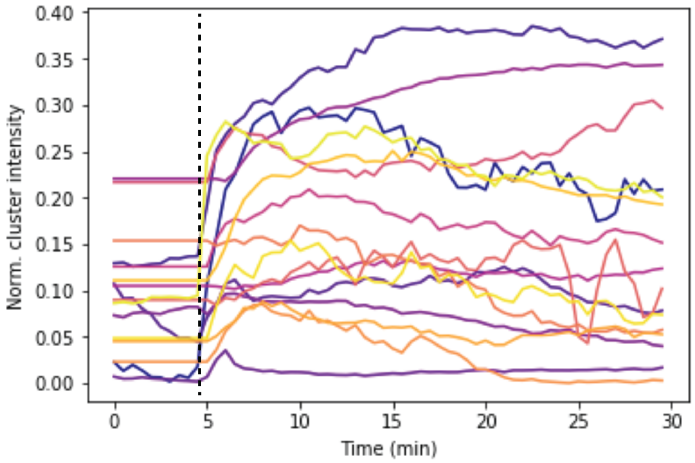


Figure 2.9: **Time series of dynamic AMPK activation of live-cells in response to ionomycin stimulation.** Ionomycin (2μM) was added to HeLa cells at the 5 min. time point, as indicated by the black dotted line. The fluorescence signal from clustered reporter was quantified by intensity thresholding and is normalized by the total fluorescence signal in the cell volume, representing the dynamically changing fraction of reporter in the clustered state in each individual cell.

The images of AMPK SPARK clustering in response to 2-DG were also quantified with this method, yielding time traces representing the dynamics AMPK activation of individual cells (Fig. 2.10). This population of cells also demonstrates heterogeneity in basal metabolic state before stimulation and heterogeneity in dynamic response to AMPK stimulation. The population-level patterns in AMPK response to ionomycin and 2-DG differ, corresponding to our intuition of the biological mechanisms driving both agonists. AMPK activation spikes immediately after exposure to ionomycin, reflecting the sudden onset of the calcium flux and activation of the CAMKK2-AMPK pathway. On the other hand, AMPK activation ramps more slowly after exposure to 2-DG, reflecting a slower onset of metabolic stress as the cells consume their remaining ATP resources.

However, this method of cluster quantification has its limitations, because segmentation by an absolute intensity threshold can not accurately differentiate all pixels corresponding to clusters from those corresponding to diffuse reporter. For example,

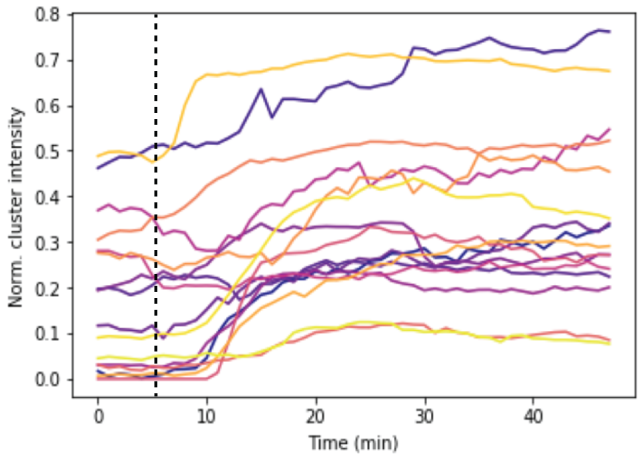


Figure 2.10: **Time series of dynamic AMPK activation of live-cells in response to 2-DG stimulation.** 2-DG (20mM) was added to HeLa cells at the 5 min. time point, as indicated by the black dotted line. The fluorescence signal from clustered reporter was quantified by intensity thresholding and is normalized by the total fluorescence signal in the cell volume, representing the dynamically changing fraction of reporter in the clustered state in each individual cell.

uneven background fluorescence intensity may make areas of the diffuse reporter brighter than clusters in the dimmer parts of the cell (Fig. 2.8c). Moreover, because the thresholds were set manually, they are time-intensive to generate and are inherently subjective.

For these reasons, accurate quantification of SPARK reporter clustering necessitates a more sophisticated image analysis method. Instead of using thresholding, we segmented the SPARK reporter clusters with Polaris, our deep learning-based spot detection method.[95] We find that Polaris more accurately detects clusters in areas of the cell with ranging background intensities than intensity thresholding (Fig. 2.11). Moreover, it does not require manual parameter tuning to detect spots, increasing the reproducibility of its quantification and enabling measurements at much greater scale. The construction and performance of Polaris is detailed in Section 3.1.

Advantages of AMPK SPARK

This work constructing and validating AMPK SPARK as a novel live-cell reporter AMPK activity demonstrates that we have improved upon the limitations of AMPKAR. As is inherent to its design architecture, AMPK SPARK utilizes a single fluorescent protein and occupies a smaller region in the visible light spectrum than FRET-based reporters, making it more amenable to multiplexing with other live-

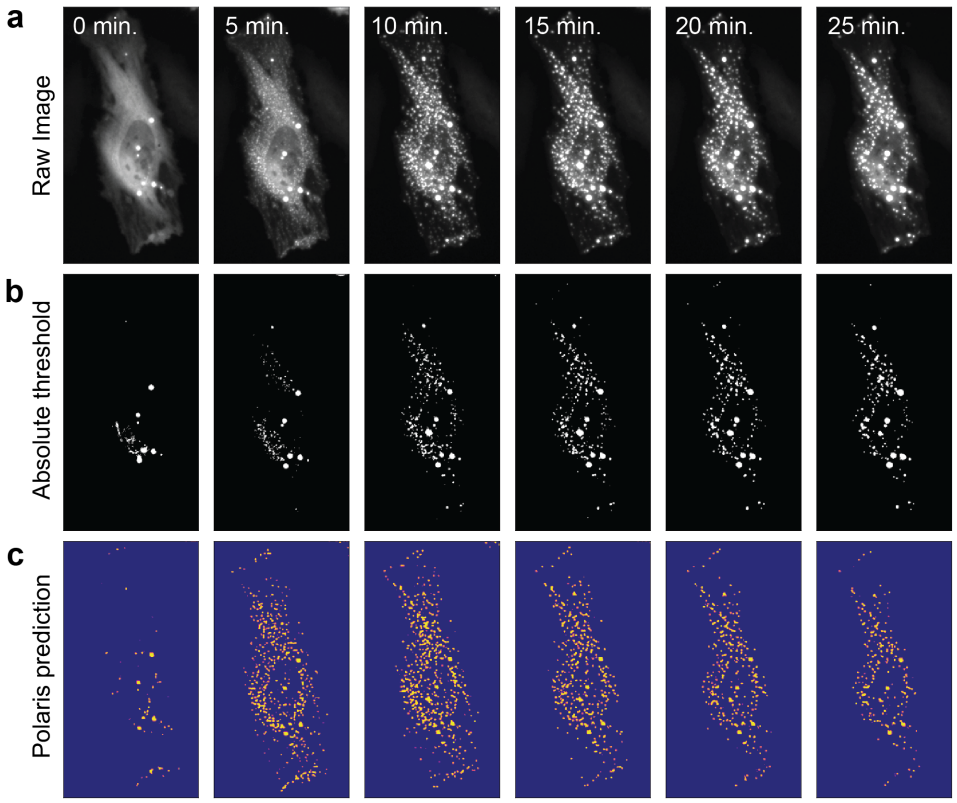


Figure 2.11: **Comparison of SPARK cluster segmentation with absolute thresholding and Polaris.** (a) Montage of AMPK SPARK response to 2 μM ionomycin treatment. (b) Absolute intensity thresholding of montage in (a). (c) Polaris prediction of spot locations in montage in (a).

cell reporters. AMPK SPARK could even theoretically be multiplexed with other SPARK reporters, if both reporters use orthogonal HOTag sequences and different PBDs.[77]

As with the previously published SPARK reporters, AMPK SPARK demonstrated much greater dynamic range than the leading FRET-based reporter, AMPKAR. In response to both ionomycin and 2-DG, AMPK SPARK demonstrated a many-fold increase in reporter clustering whereas AMPKAR demonstrated an $\sim$ 30% increase in intensity of FRET signal.[84] This improvement in dynamic range allows AMPK SPARK to detect more subtle differences in AMPK activity states at the single-cell level. Moreover, it makes AMPK SPARK signal less susceptible than AMPKAR signal to the external noise inherent to live-cell imaging assays.

An additional advantage of AMPK SPARK that potentially warrants further investigation is its ability to visualize sub-cellular AMPK activity. In HeLa cells expressing AMPK SPARK, we observed oscillatory AMPK activity that was occasionally localized to a sub-cellular compartments of the cell in response to treatment with LPS (Fig. 2.12). Both gradual accumulation and rapid spiking of fluorescent reporter can be observed in these images. These phenotypes align with plausible biological mechanisms for AMPK signaling in response to LPS stimulation. Mounting an immune response to perceived pathogens is an energetically expensive process, which would lead to gradual AMPK activation as energy resources are consumed, explaining the observation of persistently accumulating clusters. Furthermore, LPS treatment has been linked to calcium signaling activity, including fluxes from the extracellular space and stores in the endoplasmic reticulum (ER), which could explain the variety of observed whole-cell and sub-cellular pulses in AMPK SPARK clustering.[96] For these reasons, we believe that dynamic AMPK activity in response to LPS treatment warrants further investigation. Immune cells, unlike HeLa cells, express all of the signaling pathway components necessary to respond to LPS. For this reason, a measurement of the AMPK SPARK response to LPS in immune cells would be of great interest to the field.

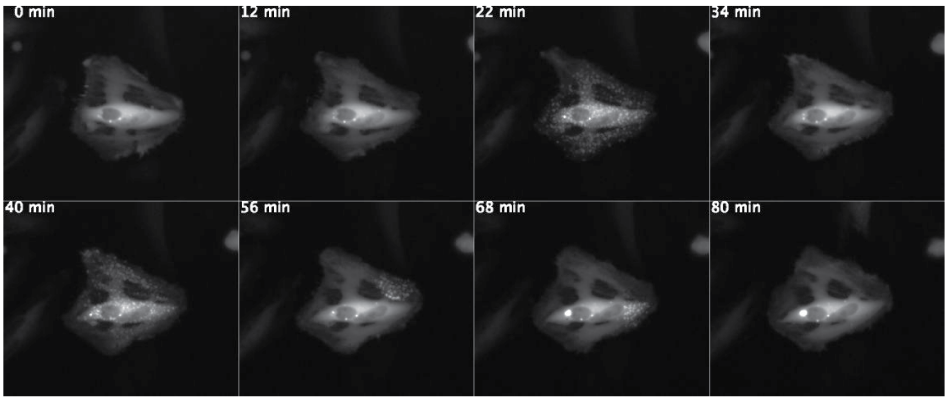


Figure 2.12: **Montage of AMPK SPARK response to LPS treatment.** Example HeLa cell transiently expressing AMPK SPARK was dosed with 1μg/mL LPS between the first and second frames. The reporter clusters in response to AMPK activity.

These observations were enabled by AMPK SPARK's improved dynamic range, which sensitively reports rapid oscillations of sub-cellular kinase activity. These oscillations occur at the low end of AMPK SPARK's dynamic range, with only a small proportion of the reporter clustering at any given time point. Therefore, a

reporter with lower dynamic range like AMPKAR would likely struggle to read out these oscillations above the degree of external noise inherent to FRET-based reporters.

Limitations of the SPARK reporter architecture

The SPARK reporter architecture is potentially limited by the kinetics and the varied mechanisms of the phase separation that it relies on to read out dynamic kinase activity. The mechanisms of reporter phase separation are similar to those of endogenous proteins.[97] A rapid spike in the concentration of phosphorylated reporter would elicit a sudden phase separation event whereas slower reporter phosphorylation may lead to gradual coalescence (Fig. 2.13a). Depending on the size of cluster that can be reliably visualized with a given imaging setup, there may be a considerable population of phosphorylated, clustered reporter that is unreliably detected, especially for the coalescence mechanisms of reporter clustering. The varied kinetics by which the phosphorylated reporter becomes visualizable limit the reliability of the SPARK clustering read out of kinase activity.

Additionally, SPARK reporter clustering is a lagging indicator of kinase activity (Fig. 2.13b). For a theoretical spike in kinase activity, the SPARK reporter would only exhibit clusters when the local concentration of phosphorylated reporter exceed the critical concentration required for phase separation. It follows that the SPARK reporter would not become diffuse again until the local concentration of phosphorylated reporter dips below the critical concentration.[98,99] Therefore, theoretically, the read out of kinase activity fails to report the earliest changes in kinase activity, but the degree to which this limits kinase kinetics reported by SPARK reporters is unknown, because the critical concentration for SPARK reporter clustering has not been rigorously characterized.

Because of the demonstrated value of the SPARK reporter architecture, we believe that the degree to which the kinetics of phase separation limits SPARK reporters' ability to read out real-time kinase activity warrants further investigation. Possible experiments to investigate this question include characterization of the critical concentration for phase separation and characterization of SPARK clustering kinetics with single molecule fluorescence imaging. Furthermore, an alternative live-cell reporter for AMPK that was published after the conception of this project, called ExRai AMPKAR, could be used as an alternative read out for dynamic AMPK activity in live cells. ExRai AMPKAR has a similar dynamic range to AMPK

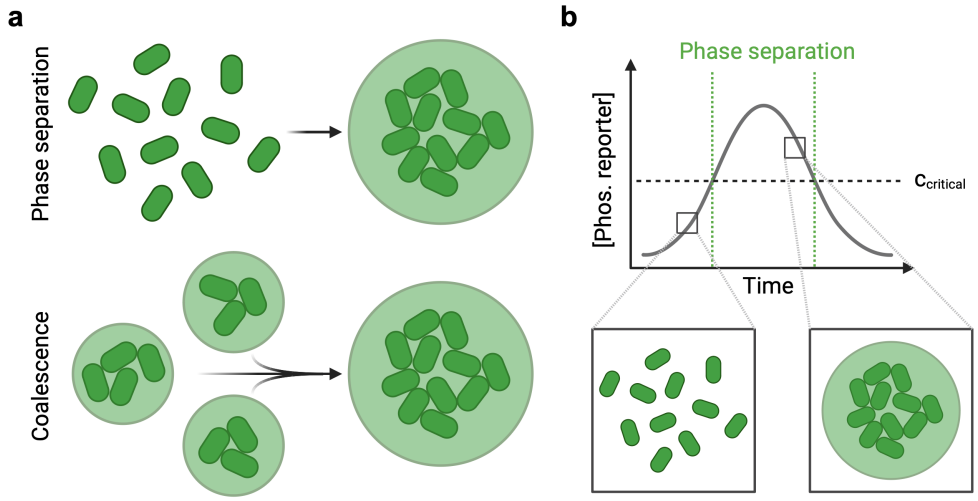


Figure 2.13: **Kinetics of phase separation limit the responsiveness of the SPARK reporter read out of kinase activity.** (a) Schematic diagram of two possible mechanisms of reporter clustering: sudden phase separation and gradual coalescence. Figure adapted from Yoo et al. (2019). (b) Example time trace of phosphorylated reporter concentration, illustrating that SPARK reporter clustering is a lagging indicator of kinase activity. The critical concentration of phosphorylated reporter for phase separation is noted as $c_{critical}$. Figure adapted from Alberti et al. (2017).

SPARK, making it a valuable resource for characterizing AMPK activity dynamics. However, its read out of kinase activity relies on measuring the ratio of two fluorescence channels, potentially limiting its multiplexing capabilities and highlighting the field's need for a single-channel live-cell reporter like AMPK SPARK.[76]

2.3 Vignette 2: ExRai reporter development with large language models

Kinase cascades in immune signaling pathways

IFN regulatory factor 3 (IRF3) is a key transcription factor involved in the innate immune response, modulating the expression of type I interferons (IFNs) particularly in response to viral infection. IRF3 is constitutively expressed in its monomeric form in the cytoplasm. Upon infection, IRF3 is phosphorylated by an upstream kinase, TANK-binding kinase 1 (TBK1), which triggers it to undergo dimerization and translocation to the nucleus where it activates the expression of its target genes. Its activation is tightly regulated by a cascade of signaling events following detection of infection (Fig. 2.14).[100–102]

Another fundamental immune signaling pathway is nuclear factor κ-light-chain-enhancer of activated B cells (NF-κB), which is a family of transcription factors activated by a number of immune stimuli, such as lipopolysaccharide, viral nucleic acid material, or cytokines. The family consists of five structurally similar members that can create a variety of homo- and hetero-dimers with different transcriptional activities. The canonical NF-κB signaling pathways involves a hetero-dimer of two subunits, p65 (also named RelA) and p50, which is sequestered in the cytoplasm by an inhibitory protein, inhibitor of nuclear factor κ-B α (IκBα). Upon detection of immune stimuli, IκBα is phosphorylated by an upstream kinase, primarily inhibitor of nuclear factor kappa-B kinase β (IKKβ), triggering its degradation and allowing the p65-p50 dimer to translocate to the nucleus and modulate the expression of its target genes (Fig. 2.14). NF-κB translocation is stereotypically oscillatory, a behavior manifested by complex regulatory feedback within a network of signaling proteins.[9,103]

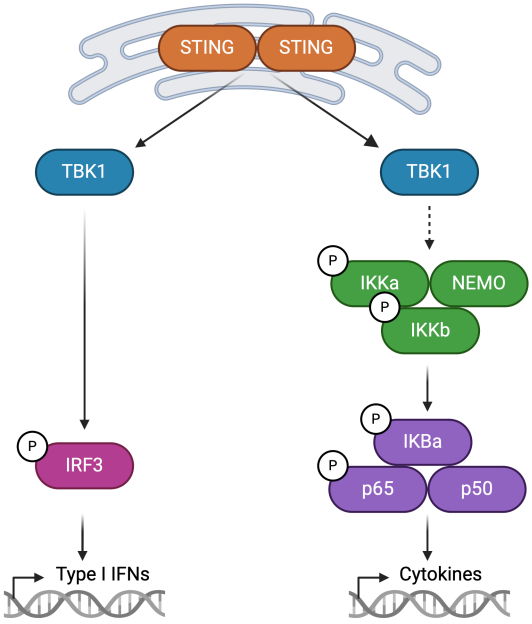


Figure 2.14: **Parallel immune signaling pathways: NF-κB and IRF3.** STING, which is present in the endoplasmic reticulum, detects viral DNA and triggers a cascade of signaling kinase activity, which culminates in the translocation of IRF3 and p65 into the nucleus, activating the expression of their target genes. Dashed arrows indicate the omission of known molecular steps. Figure adapted from Balka et al. (2020).

Controlling the phosphorylation of both IRF3 and NF-κB are intersecting cascades of signaling kinases, leading to crosstalk between the activation of both signaling pathways in response to immune stimuli. For example, TBK1 activity downstream of stimulator of interferon genes (STING) detection of the presence of viral DNA activates both IRF3 and NF-κB translocation (Fig. 2.14). This pattern of redundancy is also observed for IκB kinase ϵ, which also activates both IRF3 and NF-κB translocation in response to pro-inflammatory stimuli.[104–106]

Dysregulation of both the NF-κB and IRF3 signaling pathways is implicated in a number of disease states, including cancer, inflammatory disorders, and autoimmune disorders.[107,108] Therefore, characterization of these signaling pathways is of particular translational interest, as mechanistic insight into the regulation of their activation has previously revealed druggable signaling proteins.[109–112]

Beyond translational applications, modeling these immune signaling pathways as toy systems for motifs of regulatory feedback is of significant interest to the systems biology field, and mathematical models of increasing nuance and complexity continue to be developed.[113–118] These models require biochemical parameters of kinase activity, such as catalytic rates and affinity for binding partners. These parameters are often characterized with population-level measurements, which can miss the heterogeneity in signaling responses that would be evident at the single-cell level. Fluorescent live-cell reporters, on the other hand, have the capability to measure dynamic kinase activity at the single-cell level, which may reveal heterogeneity in dynamic behaviors that was previously impossible to measure. Characterization of the information flow through these pathways at the single-cell level could lead to the development of models for these signaling pathways with greater complexity and nuance.

We aim to create a set of live-cell reporters to characterize the kinase cascades in the NF-κB and IRF3 signaling pathways. Such a toolbox of reporters could quantify the information flow between various immune stimuli and cellular responses at a single-cell level. We first created reporters for IKKβ and TBK1, which are the kinases directly upstream of the p65 and IRF3 translocation steps, respectively. Additionally, in experiments where direct visualization of p65 or IRF3 is not feasible, these reporters can offer an alternative read out for signaling pathway activity because of the strong correspondence between the activity of IKKβ and TBK1 to the translocation of p65 and IRF3, respectively.[115,119,120]

Designing live-cell reporters with large language models

Selection a substrate for a kinase of interest is fundamental to designing a live-cell reporter for the activity of such kinase. Candidate substrates could be identified from curated databases of known kinase-substrate interactions (KSIs)[121–123], but screening candidate live-cell reporter sequences in an arrayed fashion is time and resource intensive. To identify the most promising kinase substrates, we use our large language model for predicting KSIs, called KINBERT. KINBERT was trained with the most comprehensive data set (131,081 KSIs) and its predictions have been shown to recapture many times more known KSIs when compared to other previously published language models for KSI prediction.

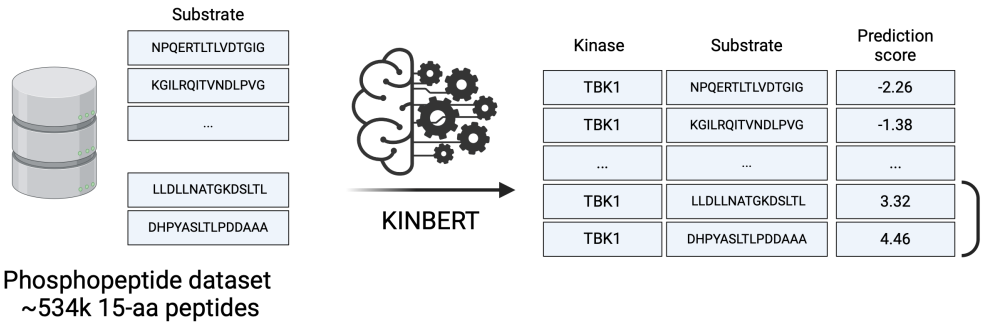


Figure 2.15: **Designing live-cell reporters with large language models.** Large language models can be used to predict kinase-substrate interactions (KSIs). Figure courtesy of Changhua Yu.

In addition to being an amenable target sequence for the kinase of interest, the designed kinase substrate must meet a few more requirements to be a good candidate for a live-cell reporter. As described for AMPK SPARK in Vignette 1, the kinase substrate must be compatible with FHA1 binding, which is the typical PBD used in ExRai reporters. Candidate IKKβ and TBK1 substrates were filtered based on their compatibility with the selective binding preferences of FHA1.[87,88] Based on these criteria, a set of candidate substrates were selected for IKKβ and TBK1 from the substrates identified by KINBERT for further screening in mammalian cells (Table 2.1).

Overview of ExRai reporters

We selected the ExRai architecture to construct live-cell reporters for IKKβ and TBK1. ExRai reporters contain cpGFP, which has two conformations with different fluorescent excitation spectra. Phosphorylation of the kinase substrate in ExRai

Table 2.1: **Candidate kinase substrates for IKKβ and TBK1.** Target phosphosites in the kinase substrates are shown in bold. The score returned by KINBERT is a logit score corresponding to the probability of kinase-substrate interaction.

Kinase	ID No.	Original protein	Amino acid sequence	KINBERT score
IKKβ	1	HS90B	NPQERTL**T**LVDTGIG	4.620
IKKβ	2	PRDX1	KGILRQI**T**VNDLPVG	4.460
IKKβ	3	CSN5	EGPSEYQ**T**IPLNKIE	3.944
TBK1	1	DDX3X	LLDLLNA**T**GKDSLTL	4.476
TBK1	2	DDAH2	DHPYASL**T**LPDDAAA	4.561

reporters drives the conformation change in the cpGFP (Fig. 2.16). In the unphosphorylated conformation, the cpGFP has a higher excitation peak at 405 nm, and in its phosphorylated conformation, the cpGFP has a higher excitation peak at 488 nm. Therefore, the read out of kinase activity for ExRai reporters is the ratio of the fluorescence intensity in the 488 nm and 405 nm excitation images.[74]

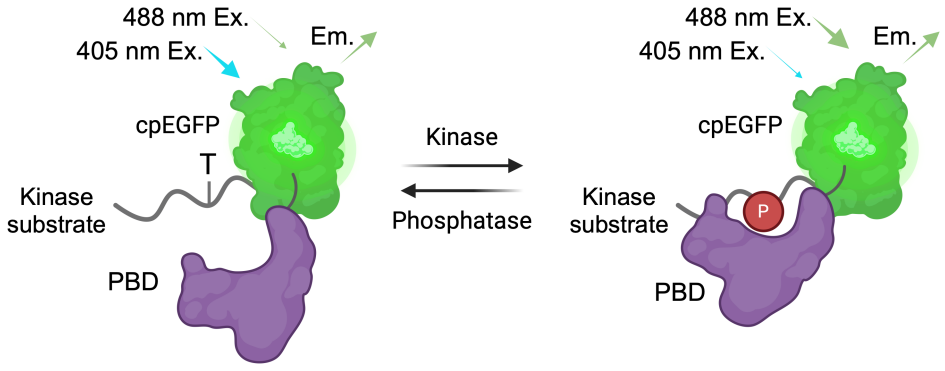


Figure 2.16: **ExRai reporters undergo a conformation change in cpGFP in response to kinase activity.** ExRai reporters contain a cpGFP molecule fused to a substrate for the kinase of interest and a phospho-binding domain (PBD). Upon phosphorylation, the reporter undergoes a conformation change that shifts the primary peak in the excitation spectrum of cpGFP from 405 nm to 488 nm.

Similar to SPARK reporters, ExRai reporters have been shown to have a wide dynamic range and the ability to read out kinase activity with sub-cellular spatial resolution.[74–76] Unlike SPARK reporters, ExRai reporters are not limited by a

complex design that leads to nuances in read out dynamics or limits multiplexing capacity. For these reasons, we selected the ExRai reporters architecture to construct novel live-cell reporters to IKKβ and TBK1. We incorporated the substrates identified by the large language model method into ExRai reporters, which we call ExRai IKKAR and ExRai TBKAR henceforth.

ExRai IKKAR and ExRai TBKAR: Novel ExRai reporters for IKKβ and TBK1

To screen the performance of the set of candidate ExRai IKKAR and ExRai TBKAR reporters, we measured the response of the reporters to poly (I:C) treatment in HT-1080 cells. Poly (I:C) is an dsRNA analog and is a known agonist of Toll-like receptor 3 (TLR3) signaling, activating immune signaling proteins including IKKβ and TBK1.[124] We expect low basal activity from both kinases in homeostatic conditions before poly (I:C) treatment and significant kinase activation upon treatment.

The screened substrates for IKKβ and TBK1 yielded ExRai reporters with varying sensitivity to kinase activity. The substrates that demonstrated the greatest sensitivity to target kinase activity were IKKβ substrate 1, derived from HS90B, and TBK1 substrate 2, derived from DDAH2. These substrates were also assigned the highest probability of kinase-substrate interaction by KINBERT (Table 2.1), demonstrating the efficacy of identifying candidate kinase substrates for live-cell reporter design with KINBERT.

In response to poly (I:C) treatment and read out by ExRai IKKAR and ExRai TBKAR, we observe dynamic responses of IKKβ and TBK1, respectively, that are heterogeneous at a single cell level over time (Fig. 2.17). Population-level methods such as Western blotting would not be able to quantify the activity of these kinases at a single-cell level, which may obscure the sub-populations of cells that may respond divergently to the same stimulus. End-point measurements like immunofluorescence staining cannot capture the changes in kinase activity over time, which prevents interrogation of the information processing mechanisms of the dynamic activation of these kinases. ExRai IKKAR is of particular value for characterization of the relationship between the activity of IKKβ and p65, whose stereotypical oscillatory nuclear localization is obscured at the population level.

The ExRai reporter architecture has the distinct advantage over other established live-cell reporter architectures, such as KTRs, of reporting kinase activity with sub-cellular spatial resolution. Both ExRai IKKAR and ExRai TBKAR demonstrate lower activity in the cell nuclei than cytoplasm volumes, a pattern that is

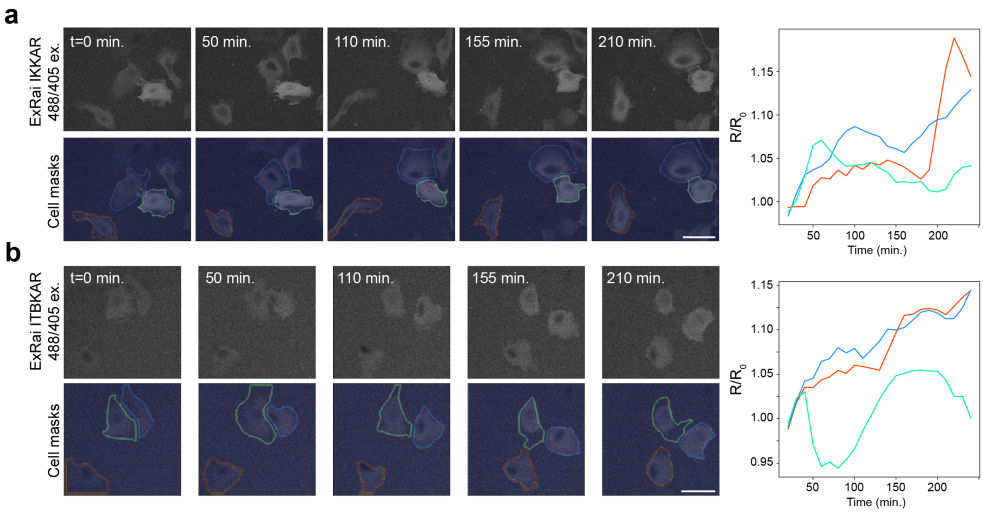


Figure 2.17: **ExRai IKKAR and ExRai TBKAR demonstrate heterogeneous single-cell responses to poly(I:C) treatment.** (a,b) HT-1080 cells expressing ExRai IKKAR and ExRai TBKAR, respectively, respond heterogeneously to 2 μg/mL poly(I:C) treatment at a single-cell level. The cells were dosed 15 minutes after the initiation of the time series. The colors of cell masks correspond with the colors of the plotted time series. Scale bar is 20μm. The time series are normalized to the average reporter activity before poly(I:C) treatment.

exacerbated by treatment with an immune agonist (Fig. 2.18). This observation follows established models of the sub-cellular activity of these kinases as primarily cytoplasmic.[115,119,120]

We demonstrate that this pattern is does not simply reflect differences in reporter abundances with line profile analysis. If we take a line profile across the nucleus of a cell expressing ExRai IKKAR, we observe that the fluorescence intensities at the termini of the line, distant from the nucleus have fluorescence intensities in both channels that are quite correlated. However, in the perinuclear region, the intensity of the 405 nm ex. image sharply increases whereas the intensity of the 488 nm ex. image modestly increases. This trend is further exacerbated in the nucleus of the cell, reading out lower kinase activity in the nucleus.

We compare the nuclear line intensity profile to one taken in a region of the cell's cytoplasm with ranging reporter abundance. The intensity of both the 405 nm ex. image and the 488 nm ex. images are highly correlated. Therefore, the ratio of intensities from the 488 nm ex. image to the 405 nm ex. image are not altered by changes in reporter abundance. This control analysis illustrates the robustness of the ExRai reporter read out of kinase activity to changes in abundance.

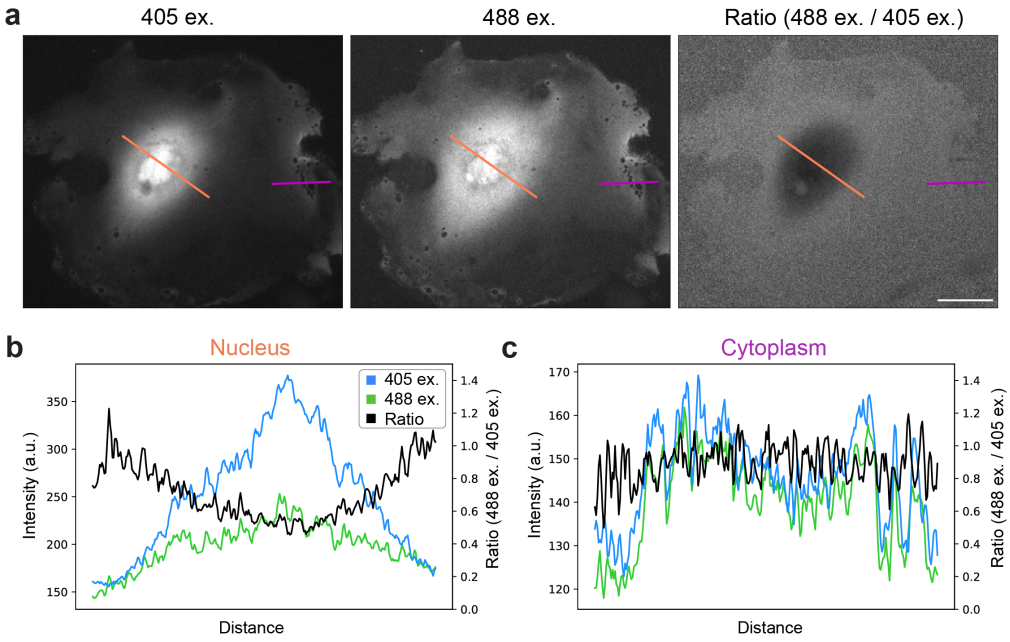


Figure 2.18: **ExRai IKKAR reveals sub-cellular patterns of IKKβ activation.** (a) Images of ExRai IKKAR expressed in an individual HT-1080 cell, treated with 2 μg/mL poly(I:C) and visualized in 405 nm excitation (ex.) and 488 nm ex. channels and a ratio of the two channels. The orange and magenta line profiles correspond to the plots in (b-c). Scale bar is 20μm. (b) Line intensity profile of 405 nm ex., 488 nm ex., and ratio images across the nucleus. Profile corresponds to the orange line in (a). (c) Line intensity profile of 405 ex., 488 ex., and ratio images across a region of the cytoplasm with varying reporter concentration. Profile corresponds to the magenta line in (a).

The observation of varying kinase activities in different sub-cellular compartments for IKKβ and TBK1 warrants further explicit characterization. Additional experiments could include fusing the ExRai reporters to NES and NLS sequences to target the reporter to specific sub-cellular compartments. Such experiments can yield biophsyical parameters of the activity of these kinases that would be of great value to mathematical modeling of immune signaling pathways.

Discussion

This work illustrates the construction of ExRai IKKAR and ExRai TBKAR with substrates identified with a large language model for prediction of KSIs, KINBERT. Live-cell reporters for IKKβ and TBK1 would be of particular translational interest, as both kinases are druggable targets of active research interest.[125–130] However, the greatest value of this work is the demonstration of KINBERT as a computational

method for candidate kinase substrate sequence generation and ExRai reporters as an experimental platform for screening such sequences. In future works, KINBERT can be used to streamline candidate substrate identification for the design ExRai reporters for other kinases of interest.

A number of control experiments are still needed to verify the validity of ExRai IKKAR and ExRai TBKAR's read out of the activity of their respective kinases. The specificity of the reporters for their target kinase needs to be demonstrated by activating the reporters in the background of inhibitors specific to the respective kinases. Additionally, the specificity of the read out to phosphorylation of the expected phospho-site in the reporter needs to be demonstrated by mutating the phospho-site to an amino acid that cannot undergo phosphorylation, such as alanine. Finally, it would be interesting to demonstrate that the dynamic range of the reporters is not limited by its expression level. In the regime where the expression level of the reporter far exceeds the capacity of the kinase of interest to phosphorylate it, the large pool of unphosphorylated reporter could limit the dynamic range of the reporter. To test this consideration, stable cell lines with varying reporter expression could be generated by Lentiviral transduction at different multiplicities of infection, and their dynamic range in response to immune stimuli could be tested.

Once validated, the performance of ExRai IKKAR and ExRai TBKAR should be further optimized, because the dynamic range of both reporters is approximately 15%, which is much lower than other published ExRai reporters.[75,76] Mutagenesis to the kinase substrate and cpGFP sequences, as was done previously for ExRai AKAR2[75], could yield an improved version of ExRai IKKAR and ExRai TBKAR by modulating the rate at which the substrate is phosphorylated and the efficiency of the read out of its phosphorylation state.

2.4 Methods

Plasmid construction

The plasmids for AMPK SPARK, ExRai IKKAR, and ExRai TBKAR were cloned using Gibson assembly of an oligo containing the kinase substrate and linearized backbone plasmid containing the reporter sequence (Takara In-Fusion HD). The reporter sequence that was used as a template for AMPK SPARK was derived from PKA SPARK (Addgene Plasmid #106920). The reporter sequence that was used as a template for the ExRai reporters was ExRai AMPKAR (gift from the Zhang lab). The plasmid backbones contain a mammalian promoter, EF-1α, and a T7

Table 2.2: **Reporter peptide sequences.** Target phospho-sites in the kinase substrates are shown in bold.

Reporter	Construct	Sequence
AMPK SPARK	Kinase substrate	RRVA**T**LVDL
AMPK SPARK	HOTag1	GEIAKSLKEIAKSLKE IAWSLKEIAKSLKG
AMPK SPARK	HOTag2	TLREIEELLRKIIEDS VRSVAELEDIEKW LKKI
ExRai IKKAR	Kinase substrate	QERTL**T**LVDTG
ExRai TBKAR	Kinase substrate	PYASL**T**LPDDA

promoter for expression by transient DNA and RNA transfections, respectively. For each plasmid, the cloning scheme was verified by sequencing of the junctions of the assembly.

Cell culture

HeLa (CCL-2) cells were received from the American Type Culture Collection. The cells were cultured in Eagle's Minimum Essential Medium (Cytiva #SH30024LS) supplemented with 2 mM L-glutamine (Gibco), 100 U/mL penicillin, 100 μg/mL streptomycin (Gibco or Caisson), and 10% fetal bovine serum (Omega Scientific or Thermo Fisher). Cells were incubated at 37°C in a humidified 5% CO_2 atmosphere and were passaged when they reached 70%-80% confluence.

HEK 293T (CRL-3216) and HT-1080 (CCL-121) cells were received from the American Type Culture Collection. Both cell lines were cultured in Dulbecco's Modified Eagle's Medium (Cytiva #SH3024302) supplemented with 2 mM L-glutamine (Gibco), 100 U/mL penicillin, 100 μg/mL streptomycin (Gibco or Caisson), and 10% fetal bovine serum (Omega Scientific or Thermo Fisher). Both cell lines were incubated at 37°C in a humidified 5% CO_2 atmosphere and were passaged when they reached 70%-80% confluence.

THP-1 (TIB-202) cells were received from the American Type Culture Collection. The cells were cultured in Roswell Park Memorial Institute (RPMI) 1640 Medium (Gibco) supplemented with 2 mM L-glutamine (Gibco), 100 U/mL penicillin, 100 μg/mL streptomycin (Gibco or Caisson), and 10% fetal bovine serum (Omega Scientific or Thermo Fisher). To make the complete medium, 2-mercaptoethanol (BME) (Sigma-Aldrich #M6250) was added to a concentration of 0.05mM before

every use. Cells were incubated at 37°C in a humidified 5% CO_2 atmosphere and were passaged to maintain a concentration of 0.3-1 x 10^6 cells/mL.

Transient transfections

For all imaging experiments, the cells were seeded in a fibronectin-functionalized (Fisher Scientific #33010018) glass-bottom 96-well plate (Cellvis P96-1.5H-N) at 80%-90% confluence. For plasmid DNA transfections, 100 ng of plasmid DNA was transfected per well with TransIT-LT1 Transfection Reagent (Mirus #MIR 2304). For both types of transfections, cells were allowed to incubate at 37°C in a humidified 5% CO_2 atmosphere overnight before imaging.

Imaging experiments

Ionomycin treatment: Ionomycin (Sigma #I3909) was received reconstituted at a stock concentration of 1 mM in DMSO. Cells were treated with 2 μM ionomycin. Cells were imaged at a frame rate of 2 minutes to capture the rapid AMPK activation following this treatment.

2-DG treatment: 2-DG (Sigma #D6134) was received in powdered form and reconstituted at a stock concentration 200 mM in water. Cells were treated with 20 mM 2-DG, which exceeds the typical 5.5 mM glucose concentration in most cell culture media formulations. Cells were imaged at a frame rate of 5 minutes.

LPS treatment: LPS O55:B5 (Sigma #L4524) was received in powdered form and reconstituted at a stock concentration of 5 mg/mL in water. Cells were treated with 100 ng/mL or 1 μg/mL. Cells were imaged at a frame rate of 5 minutes.

Poly (I:C) treatment: Poly (I:C) (Tocris #4287) was receive in powdered form and reconstituted at a stock concentration of 3 mg/mL in water. Cells were transfected with poly (I:C) using TransIT-LT1 Transfection Reagent (Mirus #MIR 2304), following the manufacturer's specifications. Cells were imaged at a frame rate of 5 minutes.

Imaging conditions

The AMPK SPARK samples were imaged with a Nikon Ti2-E fluorescence microscope controlled by Nikon Elements. Images were acquired with a Nikon SOLA SE II light source, a 40X air objective, and a Photometrics Prime 95B CMOS camera. The samples were incubated at 37°C in a humidifed 5% CO_2 atmosphere during imaging.

The ExRai reporter samples were imaged with a Nikon Ti2-E with Crest X-Light V3 spinning disk confocal microscope controlled by Nikon Elements. Images were acquired with a Lumencor Celesta light source, a Photometrics Prime 95B CMOS camera, and a 40X air objective. The cpGFP was visualized with two imaging configuations: (1) 405 nm laser excitation with a 488 nm emission filter, and (2) 488 nm laser excitation with a 488 nm emission filter. The samples were incubated at 37°C in a humidifed 5% CO_2 atmosphere during imaging.

Image analysis

SPARK reporter images: Whole-cell segmentation was performed by absolute intensity thresholding and watershed segmentation. Cells were tracked through time with an intersection over union (IOU)-based cell tracking algorithm. Clustered reporter was identified by absolute intensity thresholding or with Polaris'[95] spot detection model. Images were background-subtracted, using the mean intensity values of assigned to the background during segmentation. The total amount of reporter in each masked area was taken as the sum of the pixel values in the masked pixels. The reporter activity was quantified from the ratio of the total fluorescence intensity of the clustered reporter and the total fluorescence intensity of the cell.

ExRai reporter images: Whole-cell segmentation was performed with Deepcell's deep learning models for cell segmentation.[131] Cells were tracked through time with an IOU-based cell tracking algorithm and minor manual corrections.[132] Images in both fluorescence channels were background-subtracted, using the 5th percentile of intensity values of pixels assigned to the background during segmentation. The reporter activity was visualized by taking the ratio of the images acquired in the 488 nm and 405 nm excitation conditions. ExRai reporter activity was quantified from the ratio images using randomly selected ROIs inside the cell area as previously published.[76]

2.5 Figure captions

Figure 2.1: Signaling events for reporters. (a) Kinase activity is regulated by internal factors of cell state or external stimuli. A kinase in its active state will phosphorylate its substrate proteins. (b) Phosphorylation of a protein can lead to interaction with genomic DNA, regulating transcription of target genes. (c) Phosphorylation can signal for a protein to be degraded. (d) Phosphorylation can lead to translocation of the substrate protein to sub-cellular compartments. Destinations can include the nucleus of a cell or the plasma membrane. (e) Phosphorylation can

drive interaction or complex formation between proteins. (f) Phosphorylation can lead to a conformation change in the substrate protein.

Figure 2.2: Survey of live-cell reporter architectures. (a) Fluorescent fusion proteins consist of a fluorescent protein connected to a protein of interest with a flexible peptide linker. (b) KTRs consist of a fluorescent protein fused to a nuclear export signal (NES), nuclear localization signal (NLS), and a substrate for a kinase of interest. (c) SPARK reporters consist of two types of oligomers: (1) one oligomer consisting of fluorescent proteins and substrates for a kinase of interest, and (2) one oligomer consisting of phospho-binding domains (PBDs). (d) FRET reporters consist of a pair of fluorescent proteins that are amenable to FRET, such as CFP and YFP, that are connected by a flexible peptide linker containing a kinase substrate and a PBD. (e) ExRai reporters consist of a cpGFP molecule fused to a kinase substrate and a PBD.

Chapter 3

POLARIS: ACCURATE SINGLE-MOLECULE SPOT DETECTION FOR IMAGE-BASED SPATIAL TRANSCRIPTOMICS WITH WEAKLY SUPERVISED DEEP LEARNING

1. Laubscher, E. *et al.* Accurate single-molecule spot detection for image-based spatial transcriptomics with weakly supervised deep learning. *Cell Systems* **15,** 475–482.e6. ISSN: 2405-4712. doi:10.1016/j.cels.2024.04.006. http://dx.doi.org/10.1016/j.cels.2024.04.006 (May 2024).

3.1 Introduction

Advances in spatial transcriptomics have enabled system-level gene expression measurement while preserving spatial information, enabling new studies into the connections between gene expression, tissue organization, and disease states[75,133]. Spatial transcriptomics methods fall broadly into two categories. Sequencing-based methods leverage arrays of spatially barcoded RNA capture beads to integrate spatial information and transcriptomes[25–28]. Image-based methods, including multiplexed RNA fluorescent in situ hybridization (FISH) and in situ RNA sequencing (ISS), perform sequential rounds of fluorescent staining to label transcripts to measure the expression of thousands of genes in the same sample[29–33]. Because these methods rely on imaging, the data that they generate naturally contain the sample's spatial organization. While image-based spatial transcriptomics enables measurements with high transcript recall and subcellular resolution[75,133], rendering the raw imaging data interpretable remains challenging. Specifically, the computer vision pipelines for image-based spatial transcriptomics must reliably perform cell segmentation, spot detection, and gene assignment across diverse imaging data. Prior methods that sought an integrated solution to this problem rely on manually-tuned algorithms to optimize performance for a particular sample or spatial transcriptomics assay[134,135]. Thus, there remains a need for an integrated, open-source pipeline that can perform these steps reliably across the diverse images generated by spatial transcriptomics assays with minimal human intervention.

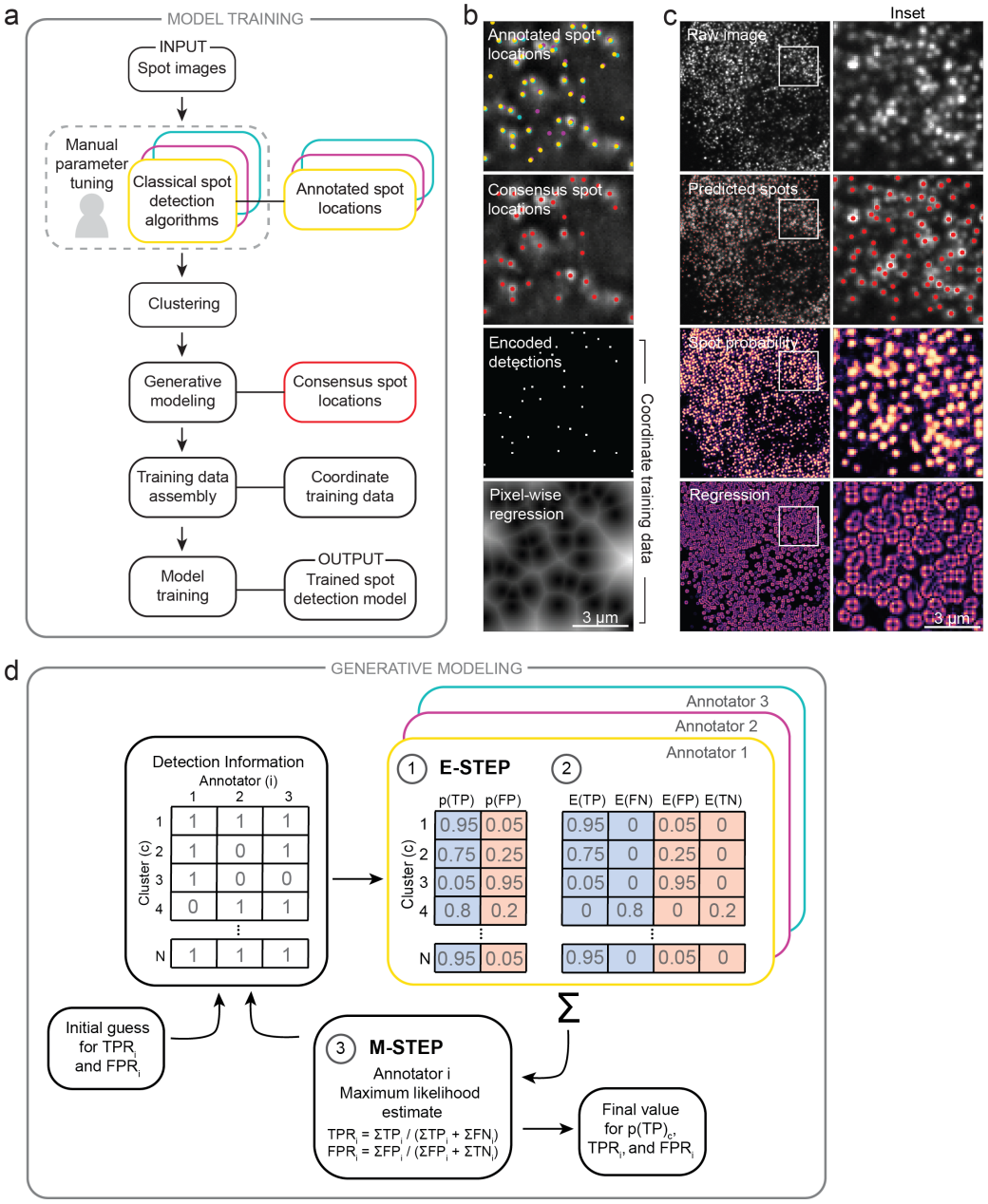


Figure 3.1: **A weakly supervised deep learning framework for accurate fluorescent spot detection of spatial transcriptomics imaging data.** See section 3.6 for caption.

Deep learning methods are a natural fit for this problem. Prior work by ourselves and others has shown that deep learning methods can accurately perform cell segmentation with minimal user intervention[136–139], providing a key computational primitive for cellular image analysis. Here, we focus on the problem of spot detection for image-based spatial transcriptomics data. Existing spot detection methods fall into

two categories: "classical" and "supervised"[140]. Classical methods are widely used but require manual parameter fine-tuning to optimize performance[141,142]. The optimal parameter values are often different within regions of the same image, making implementation of classical methods time-intensive and fundamentally limiting their scalability. Supervised methods[143–145], which often rely on deep learning methodologies, learn how to detect spots from labeled training data. These methods eliminate the need for manual parameter tuning to optimize spot detection performance. However, the requirement for labeled training data presents a major challenge, as experimentally-generated data contain too many spots for manual annotation to be feasible. Training data derived from classical algorithms are limited by the characteristics of those algorithms, imposing a ceiling on model performance. Further, simulated training data lack the artifacts present in experimentally-generated data which can limit their performance on real data.

In this work, we combine deep learning a with weakly supervised training data construction scheme to create a universal spot detector for image-based spatial transcriptomics data. We demonstrate the performance of our spot detection model on simulated and experimentally-generated images. Given that training deep learning models with weak supervision can yield a computational primitive for spot detection, we then constructed Polaris, an integrated deep learning pipeline for image-based spatial transcriptomics. Constructed in this fashion, Polaris offers a turnkey analysis solution for data from various image-based spatial transcriptomics methods while removing the need for manual parameter tuning or extensive user expertise.

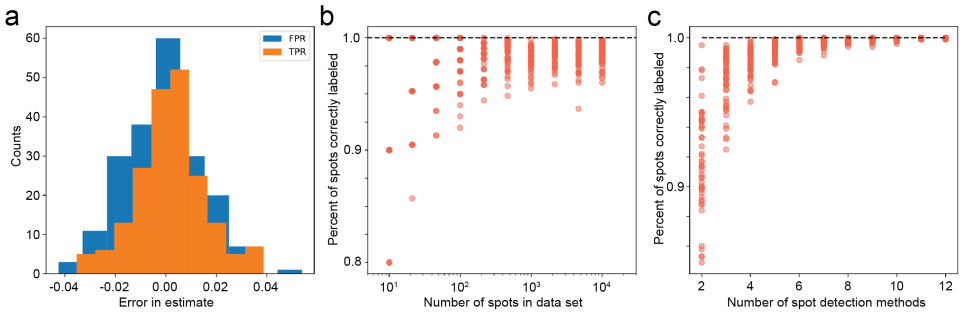


Figure 3.2: **Benchmarking consensus annotation output of the generative model.** (a) Error distribution for EM estimates of TPR and FPR values for 100 trials with three simulated classical methods. (b) Fraction of simulated detections correctly classified with increasing dataset size (number of spots in the dataset). (c) Fraction of simulated detections correctly classified as a true or false detection by EM for an increasing number of classical spot detection methods used in the EM method.

3.2 Training spot detection models with weak supervision

Here, we describe two key aspects of Polaris' spot detection model — constructing consensus training data with annotations from multiple classical spot detection algorithms, and deep learning model design and training.

Constructing consensus training data. Accurate training data is an essential component of every deep learning method. In this work, we have sought to create training data for spot detection models by finding consensus among several commonly used classical spot detection algorithms (Fig. 3.1a). In our approach, we first create annotations for representative fluorescent spot images by manually fine-tuning a collection of classical algorithms on each image. We refer to these algorithms as "annotators." This process generates conflicting sets of coordinate spot locations, as each annotator detects or misses different sets of spots. To determine which annotators detected each spot in an image, the detections from all annotators are clustered based on their proximity. Inspired by prior work on programmatic labeling[146], we then de-noise conflicting spot annotations with a generative model. The generative model characterizes annotators with two parameters: (1) true positive rate (TPR), which is an annotator's probability of detecting a ground-truth true spot, and (2) false positive rate (FPR), which is an annotator's probability of detecting a ground-truth false spot. The model characterizes clustered spots by their probability of corresponding to a ground-truth true spot (p(TP)). The generative model is given an initial guess for the TPR and FPR of each classical algorithm and a matrix of annotation data, which we name the "detection information matrix." This matrix of annotation data, $x = \{x_{ic_{i=1,\ldots,n}}\}$ consists of binary variables, x_{ic}, which are equal to 1 if annotator i detected cluster c, and 0 if not. The model is then fit with expectation maximization (EM)[147] by iteratively calculating the p(TP) of each cluster and estimating the TPRs and FPRs of each annotator until convergence (Fig. 3.1d). The remainder of this section provides more detail on the mathematical execution of these steps.

Our generative model assumes that each annotator i produces Bernoulli-distributed annotations. We define z_c to be a binary variable indicating whether cluster c is a true spot or not (1 if true, 0 if not). $p(z_c)$ — conditioned on the data and annotator characteristics — corresponds to p(TP), which we wish to compute. Let us also define $\theta_i(z_c)$ to be a variable that represents an annotator i's probability of detecting a cluster, conditioned on whether the cluster is a true spot or not. This notation is a more compact way of representing the annotator characteristics, as $\theta_i(1) = \text{TPR}_i$ and

$\theta_i(0) = \text{FPR}_i$. For every cluster c and annotator i, the distribution of x_{ic} given the cluster assignment z_c and annotator characteristics $\theta_i(z_c)$ is a Bernoulli distribution:

$$P(x_{ic}|z_c, \theta_i(z_c)) = \theta_i(z_c)^{x_{ic}}(1 - \theta_i(z_c))^{1-x_{ic}}. \tag{3.1}$$

We assume the variables x_{ic} are independent; the probability to observe the data $x = \{x_{ic}\}$, given $\{\theta_i\}$ and $\{z_c\}$ is then

$$P(\{x_{ic}\}|\{z_c\}, \{\theta_i\}) = \prod_{i,c} \theta_i(z_c)^{x_{ic}}(1 - \theta_i(z_c))^{1-x_{ic}}. \tag{3.2}$$

To offer a concrete example of this formula in action, consider the following situations for a hypothetical set of three annotators. For a "true detection" (e.g., $z_c = 1$), the probability that all three annotators detect the spot (e.g., $x_{ic} = 1$ for all i) given by the above formula reduces to

$$p(x_{i,c}|z_c, \theta) = \prod_{i=1}^{3} \text{TPR}_i, \tag{3.3}$$

which is simply the product of the TPRs of each annotator. Alternatively, the probability that the first two annotators detect a ground-truth true spot while the third annotator (incorrectly) does not is given by:

$$p(x_{i,c}|z_c, \theta) = \text{TPR}_1 \cdot \text{TPR}_2 \cdot (1 - \text{TPR}_3). \tag{3.4}$$

We utilize formula 3.2 with the EM algorithm to infer the annotator and cluster characteristics. The EM algorithm consists of two computation steps: an expectation step and a maximization step. To perform the expectation step, we define the probability of a cluster corresponding to a true or false detection with Bayes' theorem:

$$p(z_c|x_{ic}, \theta) = \frac{p(x_{ic}|z_c, \theta)p(z_c)}{p(x_{ic}|\theta)}. \tag{3.5}$$

The term $p(z_c)$ is the prior probability of a cluster corresponding to a true or false detection; we use the least informative value for the prior by setting $p(z_c) = 1/2$, indicating an equal probability that a spot is a true or false detection. The term $p(x_{ic}|\theta)$ can be expressed as:

$$p(x_{ic}|\theta) = \sum_{Z} p(x_{ic}|z_c, \theta)p(z_c). \tag{3.6}$$

Therefore, the probability $p(z_c|x_{ic}, \theta)$ can be expressed as the likelihood of each possible label (e.g., true or false) normalized by the sum of the likelihood of both labels:

$$p(z_c|x_{ic}, \theta) = \frac{p(x_{ic}|z_c, \theta)p(z_c)}{\sum_z p(x_{ic}|z_c, \theta)p(z_c)}, \tag{3.7}$$

$$= \frac{p(x_{ic}|z_c, \theta) \cdot \frac{1}{2}}{\sum_z p(x_{ic}|z_c, \theta) \cdot \frac{1}{2}}, \tag{3.8}$$

$$= \frac{p(x_{ic}|z_c, \theta)}{\sum_z p(x_{ic}|z_c, \theta)}. \tag{3.9}$$

Using this method to calculate $p(z_c|x_{ic}, \theta)$, we can then calculate $\mathbb{E}(\text{TP})$, $\mathbb{E}(\text{FN})$, $\mathbb{E}(\text{FP})$, and $\mathbb{E}(\text{TN})$ for each annotator. Two scenarios can arise when calculating these values.

1. If an annotator detects a spot in a particular cluster, i.e., $x_{ic} = 1$, $\mathbb{E}(\text{TP})$ for that method is equal to $p(\text{TP}) = p(z_c|x_{ic}, \theta)$ for that cluster, and $\mathbb{E}(\text{FP})$ for that annotator is equal to $p(\text{FP})$ for that cluster. $\mathbb{E}(\text{TN})$ and $\mathbb{E}(\text{FN})$ are set to zero.

2. If an annotator does not detect a spot in a particular cluster, i.e., $x_{ic} = 0$, $\mathbb{E}(\text{TN})$ for that annotator is equal to $p(\text{FP}) = 1 - p(\text{TP})$ for that cluster, and $\mathbb{E}(\text{FN})$ for that annotator is equal to $p(\text{TP})$ for that cluster. $\mathbb{E}(\text{TP})$ and $\mathbb{E}(\text{FP})$ are set to zero.

To perform the maximization step, we sum $\mathbb{E}(\text{TP})$, $\mathbb{E}(\text{FN})$, $\mathbb{E}(\text{FP})$, and $\mathbb{E}(\text{TN})$ across all clusters to calculate an updated maximum likelihood estimate for TPR_i and FPR_i for method i with equations of the following form:

$$\text{TPR}_i = \frac{\sum_i \mathbb{E}(\text{TP}_i)}{\sum_i \mathbb{E}(\text{TP}_i) + \sum_i \mathbb{E}(\text{FN}_i)} \tag{3.10}$$

$$\text{FPR}_i = \frac{\sum_i \mathbb{E}(\text{FP}_i)}{\sum_i \mathbb{E}(\text{FP}_i) + \sum_i \mathbb{E}(\text{TN}_i)} \tag{3.11}$$

The expectation and maximization steps are performed iteratively until the values for $p(\text{TP})_c$, TPR_i, and FPR_i converge. The consensus spot locations are taken at the centroid of the clusters with a $p(\text{TP})_c$ value that exceeds a defined probability threshold. We demonstrated that this method yields accurate estimation of TPR and FPR and that the resulting spot labels approach 100% correct as data set size and numbers of annotators increase (Fig. 3.2).

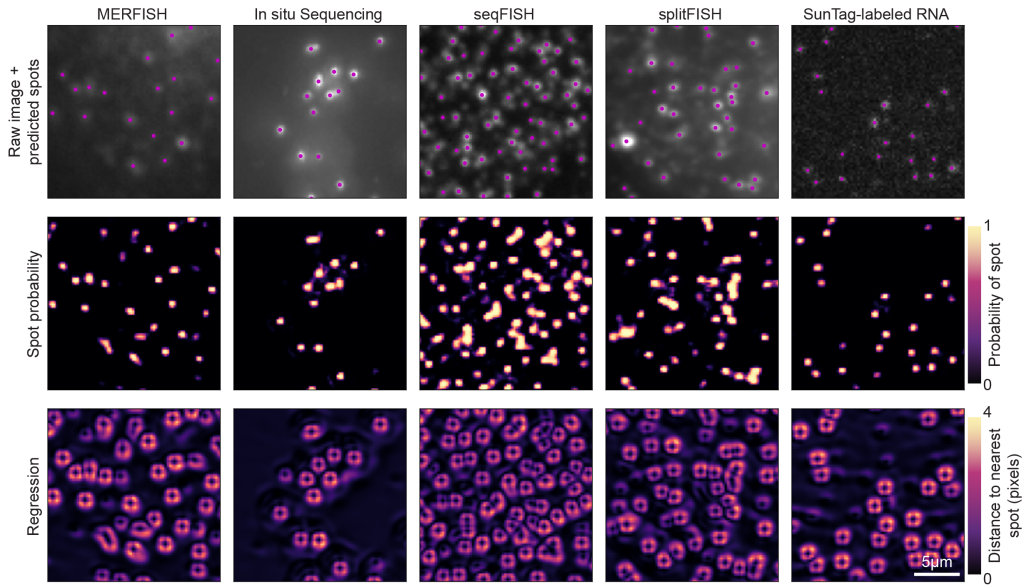


Figure 3.3: **Polaris' spot detection model generalizes to spot images generated with a variety of single-molecule assays.** The spot probability prediction images encode the pixel-wise spot probability. The regression image is the sum of the square of the subpixel distances to the nearest spot in the x- and y-dimensions. Pixels beyond a threshold value are set to zero. These outputs are used together to generate a set of predicted spot locations with subpixel resolution, plotted over the raw image.

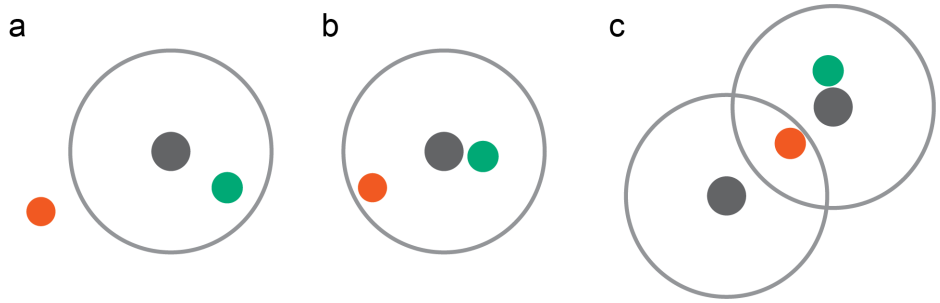


Figure 3.4: **Example cases handled by a mutual nearest-neighbor matching algorithm.** (a) Example with spots inside and outside the threshold distance to a ground-truth spot. Ground truth spots and their threshold distance are shown in grey. True positive detections are shown in green and false positive detections are shown in orange. (b) Example with two spots inside the threshold distance to a ground-truth spot. (c) Example with two spots within the threshold distance of two ground-truth spots.

To construct a training dataset for Polaris' spot detection model, we applied this consensus training data construction method to images from various spatial tran-

scriptomics assays. We assembled a set of representative images from sequential fluorescence in situ hybridization (seqFISH), multiplexed error-robust FISH (MERFISH)[148,149], and SunTag-labeled imaging data[150]. We use these consensus annotations to train a deep learning model for spot detection.

Model design and training. To train our spot detection model, we frame the problems as a classification and regression task (Fig. 3.1b). For each pixel we seek to predict whether that pixel contains a spot and compute the distance from the pixel to the nearest spot centroid. To train our model with our consensus spot labels, the coordinate spot locations are converted into two image types: (1) an image containing one-hot encoded spot locations, and (2) regression images encoding the sub-pixel distance to the nearest spot in the x- and y-directions (Fig. 3.1b). The deep learning model returns two output images from a given input image: (1) the pixel-wise probability of a spot, and (2) x- and y-regression images (Fig. 3.1c). Our model is trained with a weighted cross entropy and a custom mean squared error loss (e.g., computed only in a neighborhood around each spot) for the two output images.

We trained the network using a custom loss function composed of a classification loss and a regression loss, which considers the outputs of both of the model's prediction heads. The loss function has the following form:

$$L(y, \hat{y}) = L_{\text{cla}}(C, \hat{C}) + L_{\text{reg}}(R, \hat{R}) \tag{3.12}$$

where C is the classification head output, R is the regression head output, and $y = (C, R)$. The classification loss is the weighted cross-entropy with inverse class frequency-based weights. The regression loss is given by:

$$L_{\text{reg}}(R, \hat{R}) = \frac{1}{|G_d|} \sum_{i \in G_d} \ell((dy_i, dx_i), (\hat{dy}_i, \hat{dx}_i)) \tag{3.13}$$

where $R_i = (dy_i, dx_i)$ (i denotes a single pixel), $G_d = \{\text{pixels } i = (i_y, i_x) |$ a spot-containing pixel j exists with $L_\infty(i, j) = \max_{k \in x,y} |i_k - j_k| \leq d\}$, and ℓ is the smooth L_1 function. dx_i is the x-coordinate of the position difference between the nearest spot to pixel i, and pixel i's center. Similarly, dy_i is the y-coordinate of this position difference. d is a configurable parameter that determines the threshold distance from the nearest spot under which the estimated nearest spot's position for that pixel is taken into account in the loss function.

For our model architecture, we utilize FeatureNets, a family of models that are parameterized by their receptive field[136]. We perform hyperparameter optimization experiments to find the optimal receptive field size of 13 (Fig. 3.6). To return the location of the spots in an image, we use maximum intensity filtering to detect the local maxima in the spot probability image and assign sub-pixel location of each maximum with the regression images.

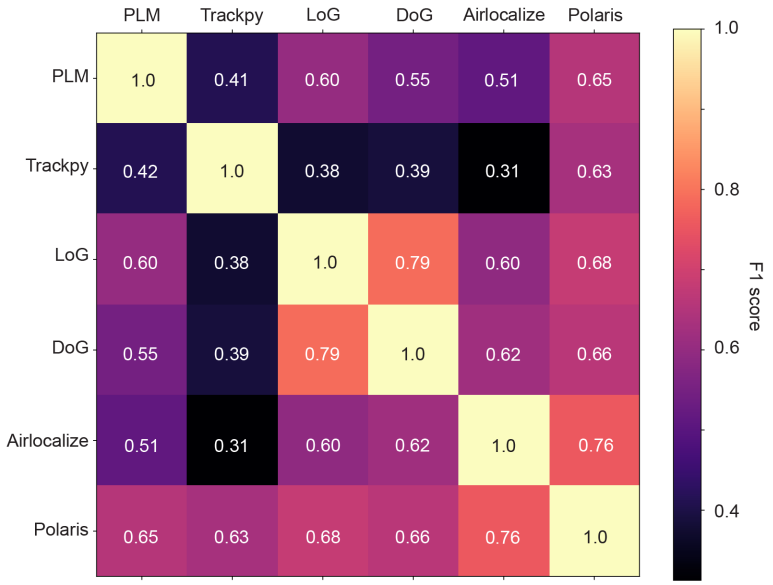


Figure 3.5: **Quantification of agreement between Polaris' deep learning model and different classical spot detection methods.** The benchmarked methods include maximum-intensity filtering (PLM), the Crocker-Grier centroid-finding algorithm (Trackpy), Laplacian of Gaussian (LoG), difference of Gaussians (DoG), Airlocalize, and Polaris.

3.3 Results

We demonstrate Polaris' spot detection capabilities on held-out experimentally-generated images. Visual inspection showed that our model generalized to out-of-distribution, spot-like data generated by various spatial-transcriptomics assays, such as ISS[29] and splitFISH images[33] (Fig. 3.3). Additionally, we used held-out images to quantify the agreement between Polaris and the classical methods used to create our consensus training data. Agreement between sets of detected spots was determined with a mutual nearest neighbors matching method. (Fig. 3.4).

We observed higher agreement between Polaris and the classical methods than exists among the classical methods themselves. This analysis demonstrates Polaris' learning of consensus labels generalizes to images held-out from the training data set. (Fig. 3.5)

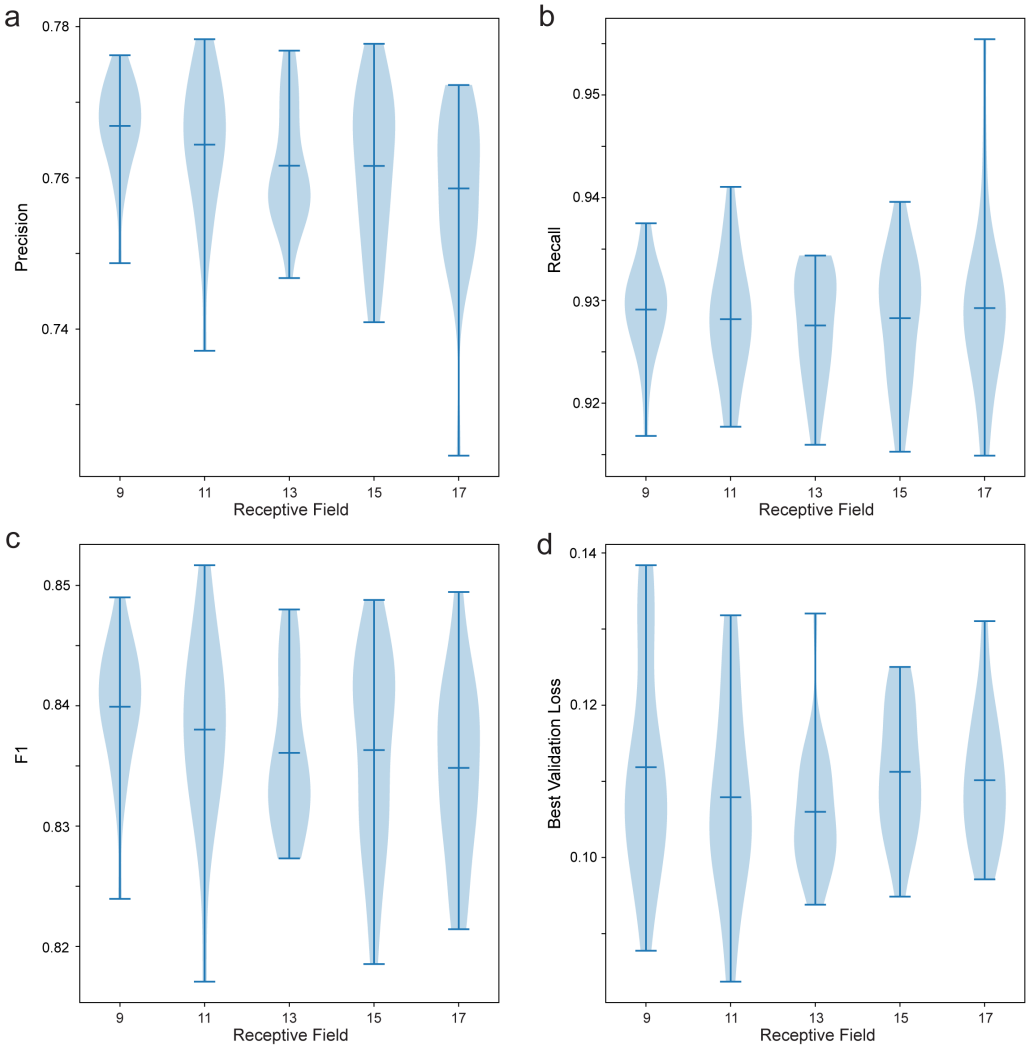


Figure 3.6: **Benchmarking the receptive field parameter of Polaris' spot detection model.** (a-d) Violin plot quantifying the performance metrics ((a) precision, (b) recall, (c) F1, (d) best validation loss during training) for models trained with different values for receptive field of Polaris' spot detection model. n=24 trained models per receptive field condition.

The ambiguity of ground-truth annotations for experimental data presents challenges for quantitatively benchmarking spot detection methods. To evaluate the accuracy of our approach, we followed prior work by simulating spot images, which have unambiguous ground truth spot locations.[151,152] When accurate simulation of

experimental data is possible, simulations remove the need for unambiguous ground-truth annotations for benchmarking. We note that our spot simulations add signal on top of autofluorescence images. Because we control the image generation, we can explore model performance as a function of image difficulty by tuning parameters such as the spot density and signal-to-noise ratio. Benchmarking on simulated data demonstrated that our method outperforms models trained with either simulated data or data labeled with a single classical algorithm. We found that this performance gap held across the tested range of spot intensity and density (Fig. 3.7a-b). We concluded that the consensus annotations more accurately capture the ground truth locations of spots in training images than any single classical algorithm and that there is significant value to training with experimentally generated images. We also found that Polaris' spot detection model outperforms other recently published spot detection methods when evaluated on these simulated data, demonstrating greater robustness to ranging spot density and signal-to-noise ratios (Fig. 3.7c-d). The combination of benchmarking on simulated data, visual inspection, and analysis of inter-algorithm agreement led us to conclude that Polaris can accurately perform spot detection on a diverse array of challenging single-molecule images.

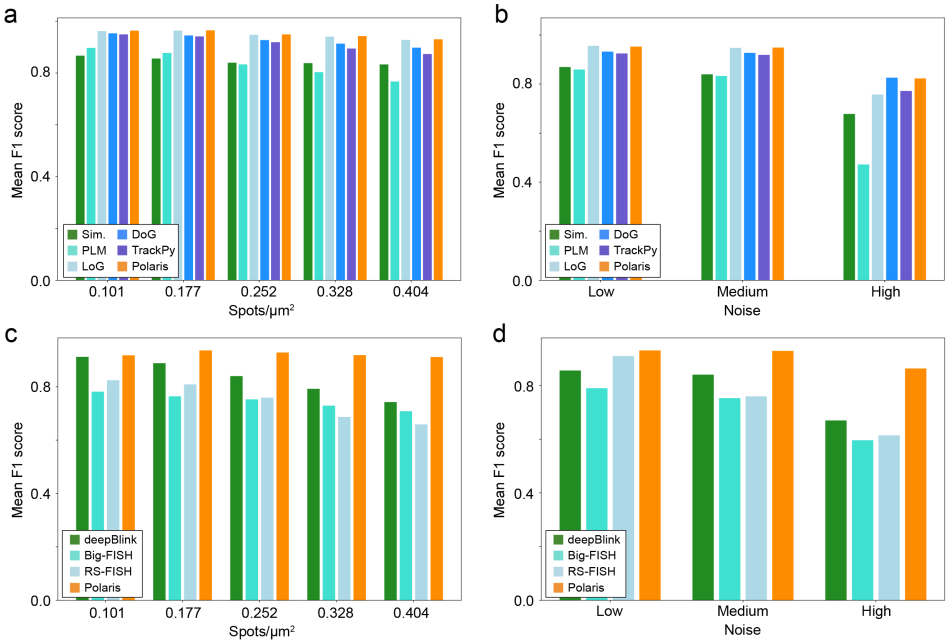


Figure 3.7: **Benchmarking model performance on simulated spot images with a range of spot intensities and densities.** See section 3.6 for caption.

Polaris packages this model into an analysis pipeline implemented in Python for multiplexed spatial transcriptomics data. Polaris integrates multiple analysis steps to yield coordinate spot locations with assigned gene identities. (Fig. 3.8a) First, it utilizes classical computer vision methods for image alignment and performs spot detection on images across all staining rounds. Then, cell segmentation is performed with models from DeepCell software library[138]. Polaris' spot detection model predicts the pixel-wise spot probability for each imaging round. For multiplexed spatial transcriptomics images, Polaris considers a codebook of up to thousands of barcodes that define the rounds and colors of fluorescent staining for each gene. To assign gene identities for barcoded spatial transcriptomics images, we fit a graphical model of a mixture of relaxed Bernoulli distributions to the pixel-wise probability values with variational inference[153,154] (Fig. 3.8b-c). This model estimates the characteristic relaxed Bernoulli distributions of pixel values for "spots" and "background" in each imaging round, which may vary due to factors such as fluorophore identity, staining efficiency, or image normalization. These distributions are used in combination with the experimental codebook to estimate the probability of each barcode identity and ultimately assign spots to a gene identity or "background."

As with our benchmarking of spot detection methods, we used simulated data to benchmark the performance of our barcode assignment method quantitatively. Here, simulated data allowed us to explore our method's dependency on spot dropout. This event can occur due to labeling failure, image quality, or failure in spot detection. Regardless of origin, the presence of dropout imposes a robustness constraint on the gene decoding methodology, as decoding schemes robust to dropout would better tolerate labeling and spot detection model failures. Our benchmarking of spot decoding with simulated data demonstrates that decoding with a generative model based on the relaxed Bernoulli distribution was more robust to dropout than other benchmarked methods (Fig. 3.9).

We demonstrated Polaris' performance on a variety of previously published data: a MERFISH experiment in a mouse ileum tissue sample[149] (Fig. 3.8d-f), a MERFISH experiment in a mouse kidney tissue sample[155] (Fig. 3.10a), a seqFISH experiment in cultured macrophages (Fig. 3.10b), and an ISS experiment of a pooled CRISPR library in HeLa cells[156] (Fig. 3.12). We found that Polaris detected marker genes from expected cell types - even in areas with high cell density and heterogeneous cell morphologies in tissue samples. These results highlight the power of spatial transcriptomics methods to quantify gene expression while retaining multicellular

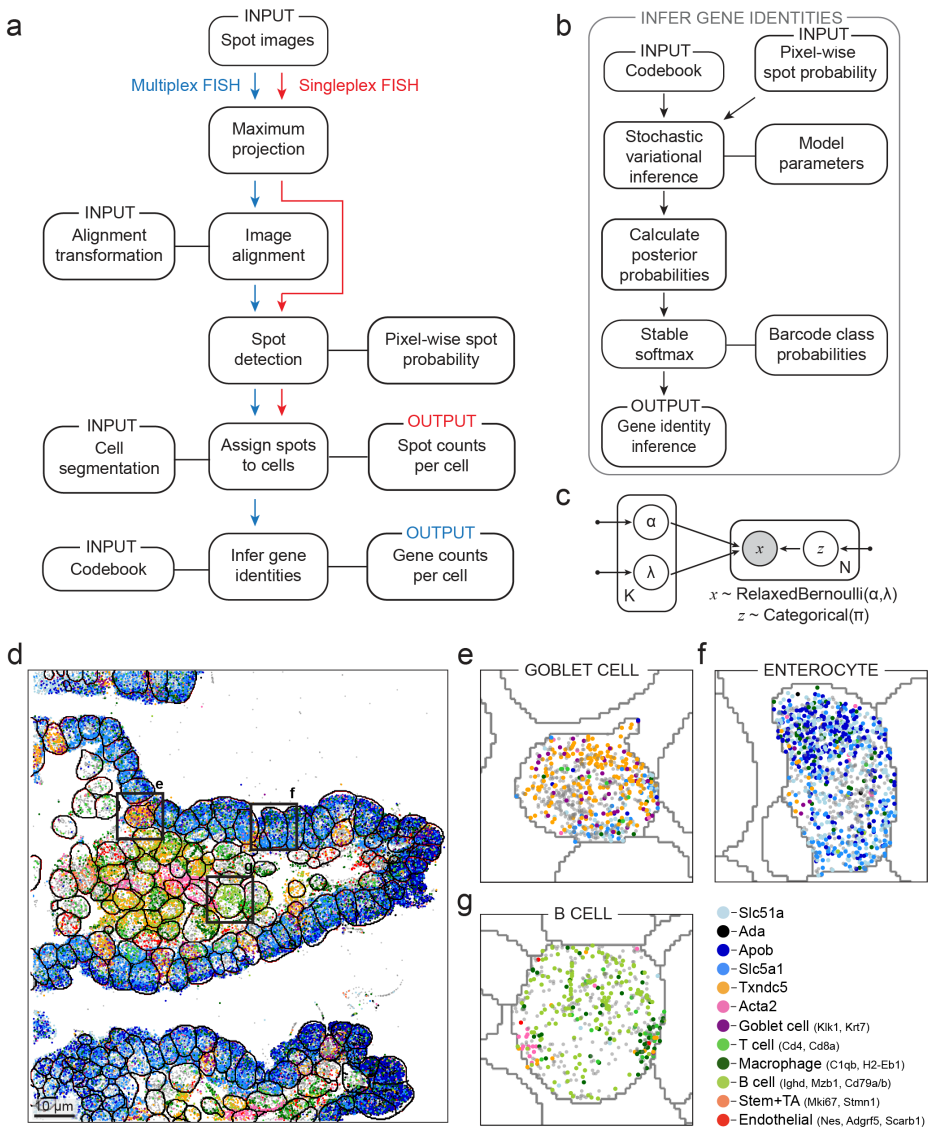


Figure 3.8: **Polaris produces single-cell, spatial gene expression maps for multiplex spatial transcriptomics images.** See section 3.6 for caption.

and sub-cellular spatial organization.

For both tissue MERFISH datasets, we found that Polaris' output gene expression counts have similar correlations with bulk sequencing data as the original analyses (r=0.796 and r=0.683; r=0.537 and r=0.565), and for both datasets, the two analysis outputs were highly correlated (r=0.936; r=0.911) (Fig. 3.11a-d). For the cell culture seqFISH dataset, Polaris' output has a similar correlation with bulk sequencing data as the output of the original analysis tool (r=0.809 and r=0.694) and the two outputs are highly correlated (r=0.910) (Fig. 3.11e-f). For the ISS dataset, the bar-

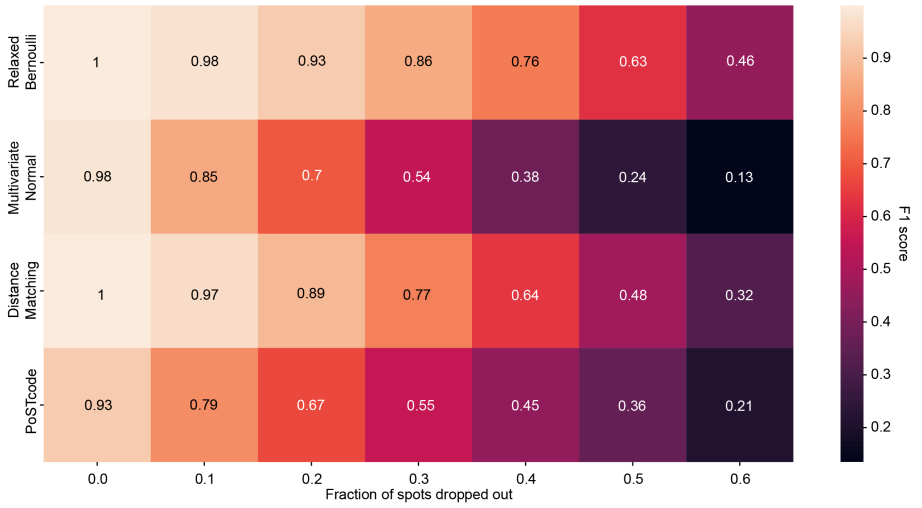


Figure 3.9: **Benchmarking of the robustness of gene decoding methods to dropout.** Quantification of F1 score for four barcode decoding methods (a graphical model of relaxed Bernoulli distributions, a graphical model of multivariate normal distributions, Hamming distance matching, and PoSTcode) for simulated barcode pixel values with a range of dropout rates.

code counts quantified with Polaris and the original analysis[156] are highly correlated (r=0.946) with Polaris consistently yielding higher counts (Fig. 3.12). Spatial transcriptomics methods often encounter overdispersion in measuring gene expression counts, potentially limiting the efficacy of comparing these counts using a linear regression model.[155,157]. Despite this limitation, these results demonstrate that Polaris can generalize across sample types, imaging platforms, and spatial transcriptomics assays without manual parameter tuning.

3.4 Discussion

We sought to create a key computational primitive for spot detection and an integrated, open-source pipeline for image-based spatial transcriptomics. Our weakly supervised deep learning model for spot detection provides a universal spot detection method for image-based spatial transcriptomics data. Our training data generation methodology effectively tackles a fundamental data engineering challenge in generating annotations for supervised spot detection methods, surpassing the performance achieved by using simulated data or a single classical method. Polaris packages this model and others into a unified pipeline that takes users from raw data to interpretable spatial gene expression maps with single-cell resolution. We believe Polaris will help standardize the computational aspect of image-based spatial transcriptomics, reduce the amount of time required to go from raw

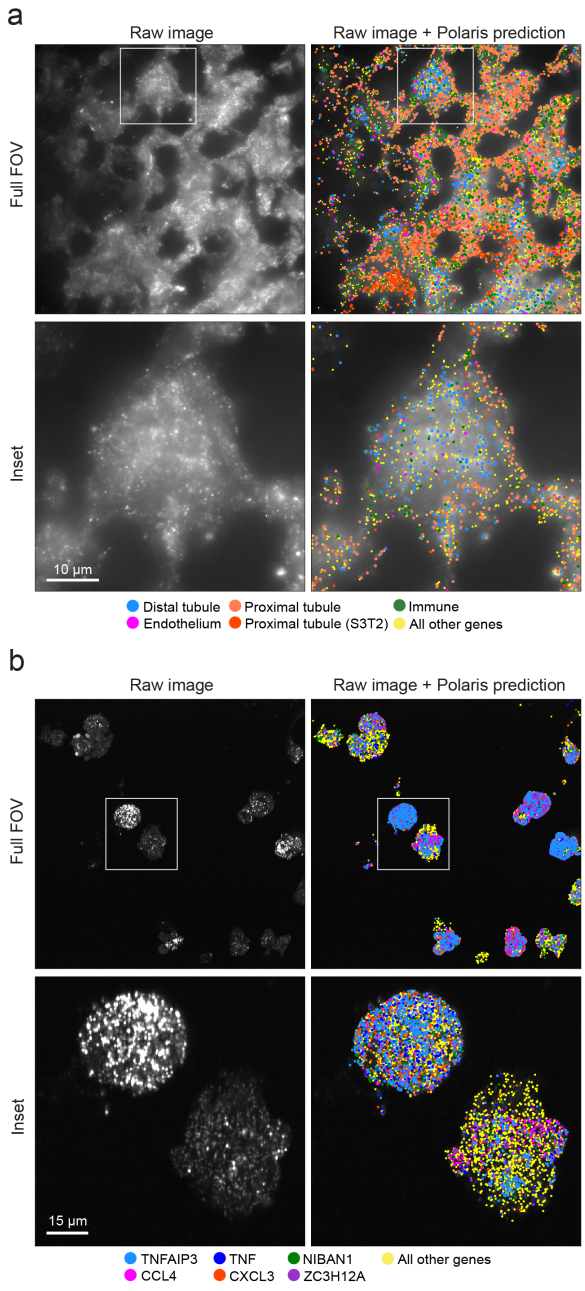


Figure 3.10: **Demonstration of Polaris' performance on a MERFISH and seqFISH data.** See section 3.6 for caption.

data to insights, and facilitate scaling analyses to larger datasets. Polaris' outputs are compatible with downstream bioinformatics tools, such as squidpy[19] and Seurat[158]. Polaris is available for academic use through the DeepCell software library `https://github.com/vanvalenlab/deepcell-spots` and as a Python package distributed on PyPI `https://pypi.org/project/DeepCell-Spots/`. A

singleplex deployment of the pipeline is available through the DeepCell web portal `https://deepcell.org`.

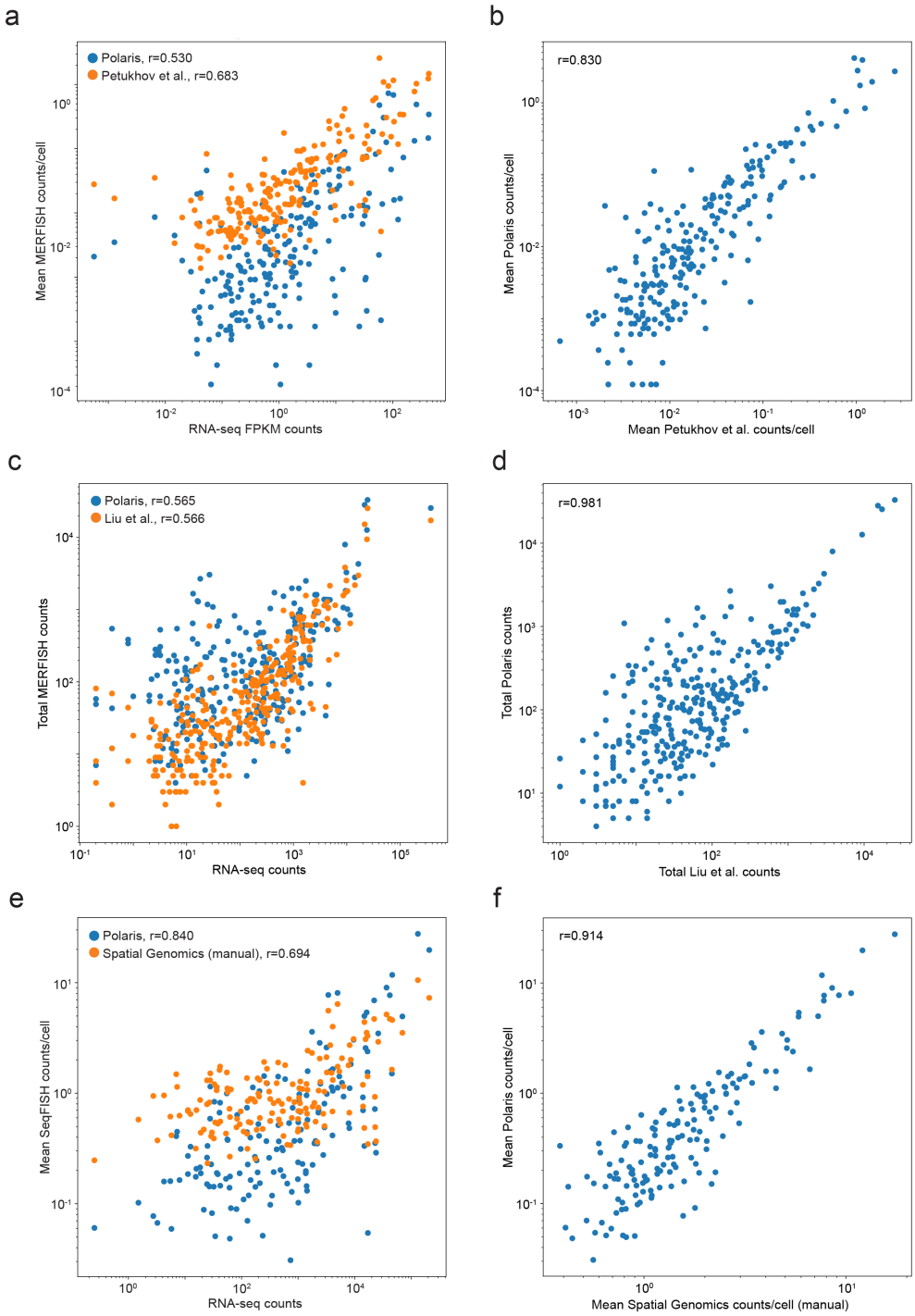


Figure 3.11: **Correlation of Polaris' quantification of MERFISH data with other quantification methods.** See section 3.6 for caption.

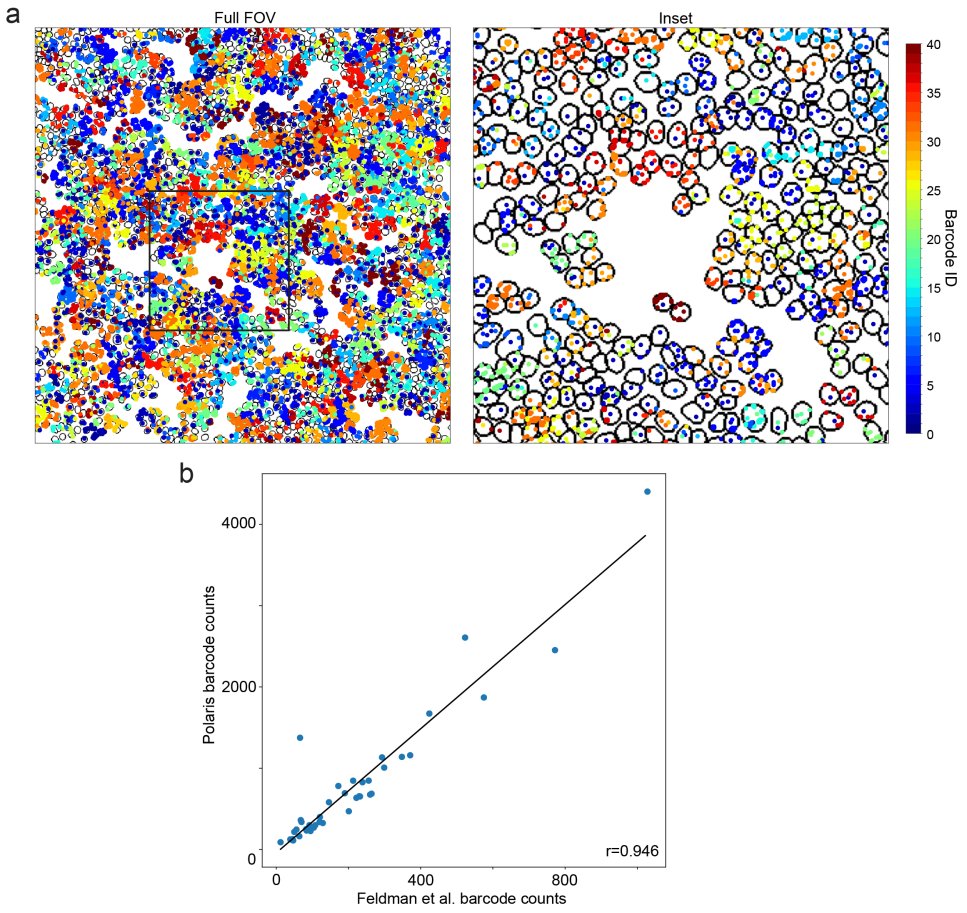


Figure 3.12: **Demonstration of Polaris' performance on an ISS dataset in HeLa cells.** (a) Example Polaris prediction for the ISS sample. The spot colors correspond with barcode identities. The inset location is defined by the black box in the full field of view (FOV). (b) Scatter plot correlating total counts for each barcode decoded by the original published analysis with counts quantified by Polaris (r=0.946).

3.5 Methods

Generation of sequential fluorescent in situ hybridization (seqFISH) images for spot training data

Probe design: mRNA transcripts were targeted with single-stranded DNA probes, as previously described[32]. Primary probes were designed to target a panel of 10 genes with OligoMiner using balanced coverage settings[159]. The primary probes were designed to have secondary probe-binding sites flanking both ends of the sequence that binds to the mRNA transcript. The secondary probes were 15 bases long and consisted of nucleotide combinations that were optimized to have 40%-60% GC content and minimal genomic off-target binding.

Probe construction: Single-stranded DNA primary probes were obtained from Integrated DNA Technologies (IDT) as an oPools Oligo Pool. Oligos were received lyophilized and dissolved in ultrapure water at a stock concentration of 1 μM per probe. Single-stranded DNA secondary probes were also obtained from IDT and were 5'-functionalized with Alexa Fluor 647, Alexa Fluor 546, or Alexa Fluor 488. Secondary probes were received lyophilized and dissolved in ultrapure water at a concentration of 100 nM.

Cell culture: HeLa (CCL-2) cells were received from the American Type Culture Collection. The cells were cultured in Eagle's minimum essential medium (Cytiva #SH30024LS) supplemented with 2 mM L-glutamine (Gibco), 100 U/mL penicillin, 100 μg/mL streptomycin (Gibco or Caisson), and 10% fetal bovine serum (Omega Scientific or Thermo Fisher). Cells were incubated at 37°C in a humidified 5% CO_2 atmosphere and were passaged when they reached 70%-80% confluence.

Buffer preparation: The primary probe hybridization buffer consisted of 133 mg/mL high-molecular-weight dextran sulfate (Calbiochem #3710-50GM), 2X saline-sodium citrate (SSC) (IBI Scientific #IB72010), and 66% formamide (Bio Basic #FB0211-500) in ultrapure water. The 55% wash buffer was comprised of 2X SSC, 55% formamide, and 1% Triton-X (Sigma-Aldrich #10789704001) in ultrapure water. The secondary probe hybridization buffer consisted of 2X SSC, 16% ethylene carbonate (Sigma-Aldrich #E26258-500G), and 167 mg/mL of high-molecular-weight dextran sulfate in ultrapure water. The ethylene carbonate was first melted at 50°C for 30-60 minutes. The 10% wash buffer was comprised of 2X SSC, 10% formamide, and 1% Triton-X in ultrapure water. The imaging buffer base consisted of 0.072M Tris HCl (pH 8) (RPI #T60050-1000), 0.43 M NaCl (Fisher #MK-7581-500), and 3 mM Trolox (Sigma-Aldrich #238813) in ultrapure water. The anti-bleaching buffer was comprised of 70% imaging buffer base, 2X SSC, 1% catalase (10X dilution of stock) (Sigma #C3155), 0.005 mg/mL glucose oxidase (Sigma-Aldrich #G2133-10KU), and 0.08% D-glucose (Sigma #G7528) in ultrapure water.

seqFISH sample preparation: The cells were seeded in a fibronectin-functionalized (Fisher Scientific #33010018) glass-bottom 96-well plate (Cellvis P96-1.5H-N) at 80%-90% confluence. Cells were rinsed with warm 1X phosphate-buffered saline (PBS) (Gibco), fixed with fresh 4% formaldehyde (Thermo #28908) in 1X PBS for 10 minutes at room temperature, and then permeabilized with 70% ethanol overnight at -20°C. Prior to probe hybridization, the cells were rinsed with 2X SSC.

The primary probes were diluted to 10 nM in the primary probe hybridization buffer and 100 μL of this solution was added to each well. Cells were incubated with the primary probe solution for 24 hours at 37°C. The primary probes were rinsed out twice with 55% wash buffer. Cells were incubated in 55% wash buffer for 30 minutes in the dark at room temperature and then rinsed three times with 2X SSC buffer. The secondary probes were diluted to 50 nM in the secondary probe hybridization buffer and 100 μL was added to each of the wells. The probes were then incubated for 15 minutes in the dark at room temperature. The secondary probes were then washed twice in 10% wash buffer. The cells were then incubated in the 10% wash buffer for 5 minutes in the dark at room temperature. Finally, the cells were washed once with 2X SSC buffer and once with imaging buffer. Before imaging, the buffer was changed to anti-bleaching buffer (100 μL).

Images of cell autofluorescence were acquired with non-specific secondary probe staining to generate simulated spot images. These samples were prepared with the same seqFISH method described above, without the addition of primary probes.

Imaging conditions: The seqFISH samples were imaged with a Nikon Ti2 fluorescence microscope controlled by Nikon Elements. Images were acquired with a Nikon SOLA SE II light source, a 100X oil objective, and a Photometrics Prime 95B CMOS camera.

Creation of spot training data

Image annotation: Our training dataset consisted of 1000 128x128 pixel images: 400 images generated as described above by performing seqFISH on cell culture samples, 400 previously published images generated with multiplexed error-robust FISH on tissue samples[148,149], and 200 previously published images generated with SunTag labeling of nascent proteins in cell culture samples[150]. All data were scaled so that the pixels had the same physical dimension of 110 nm prior to training. These images contained up to approximately 200 spots per image and were min-max normalized to a range of [0,1] prior to annotation. We annotated each image with a collection of five classical spot detection algorithms: maximum intensity filtering (`skimage.feature.peak_local_max`), difference of Gaussians (`skimage.feature.blob_dog`), Laplacian of Gaussian (`skimage.feature.blob_log`), the Crocker-Grier centroid-finding algorithm (`trackpy.locate`), and Airlocalize. To accelerate image labeling with all methods except Airlocalize, we created a Tkinter graphical user interface to tune the algo-

rithm parameters (intensity threshold, minimum distance between spots, etc.) on a per-image basis. To create annotations with Airlocalize, we utilized its previously published Matlab GUI to fine tune the algorithm parameters on a per-image basis.[160]

For the spot detection algorithms that return spot locations with pixel-level resolution (`peak_local_max`, `blob_dog`, and `blob_log` in `skimage.feature`), the subpixel localization was determined by fitting a 2D isotropic Gaussian to the spot intensity[161]. A 10x10 pixel portion of the image surrounding the detected spot was cropped out and used for the Gaussian fitting. A nonlinear least squares regression was performed, with initial parameters of Gaussian mean at the pixel center, an amplitude of 1 (the image's maximum value after min-max normalization) and a standard deviation of 0.5 pixels. The spot location was constrained to be in the middle 20% of the image, and the spot standard deviation was constrained to be between 0 and 3.

Clustering of spot annotations: Spots detected by the four classical algorithms were clustered into groups by proximity. For clarity, we refer to these classical algorithms as "annotators". Each group of detections was presumed to be derived from the same spot in the image. To perform this clustering, we first constructed a graph with each detected spot being a node. Two detections were connected by an edge if they were within 1.5 pixels of each other. We then considered the connected components of this graph to be clusters. We screened the clusters to ensure that they contained at most one detection from each algorithm. If a cluster contained more than one detection from the same algorithm, the detection closest to the cluster centroid was retained, and all other detections were separated into new clusters.

From this graph, we derived the "detection information matrix," which identifies the annotators that contribute a detection to each cluster. This matrix has dimensions of $C \cdot A$, where C is the number of connected components or clusters in the graph and A is the number of annotators. The matrix has a value of 1 or 0 when a particular annotator does or does not have a detection in a particular cluster, respectively.

Creation of consensus annotations with expectation maximization: A generative model fit with the expectation-maximization (EM) algorithm[147] was used to estimate the probability that a cluster of detections corresponds to a true spot in the image. The detection information matrix, described above, was used as the input into the generative model along with an initial guess for the true positive rate (TPR) and false positive rate (FPR) of each algorithm and a prior probability of a spot being a true detection. The initial guesses for the TPR and FPR were 0.9 and 0.1, respectively.

The prior probability of a spot being a true detection was defined as 0.5. Briefly, the EM algorithm consists of two steps - an expectation step and a maximization step. The expectation step yields an estimate for the probability that each detection cluster corresponds to a true detection. The maximization step yields an updated estimate for the TPR and FPR of each annotator. The expectation and maximization steps were performed iteratively 20 times, sufficient iterations for convergence to a local maximum of the likelihood. The resulting values were used as an estimate for the Bayesian probability of each cluster corresponding to a "true" spot. We provide further details of these steps in the main text. Clusters with a probability above 0.9 were used as spots in the training dataset, and the spot location was taken to be the centroid of the detection cluster.

Model benchmarking with simulated data

Creation of simulated spot images: Simulated spot images were used to benchmark Polaris' spot detection model. We created simulated images by adding Gaussian spots at random locations to cellular autofluorescence images. The location and number of spots in an image were sampled from uniform distributions. The intensity and width of the simulated Gaussian were sampled from normal distributions reflecting a spot distribution that is characteristic of experimental images.

Creation of simulated detection information: Simulated detection information was used to benchmark the EM algorithm for generating consensus spot annotations. First, we simulated a set of spot identities, as either true detections or false detections. The ratio of ground-truth true and false spots was determined by a pre-defined prior probability of true spots. We then used the defined TPR and FPR values to simulate detections. The probability that true spots are detected by the simulated spot detection methods is defined by the simulated TPR, and the probability that false spots are detected is defined by the simulated FPR. As with detections from experimental images, the simulated detections are stored in a detection information matrix.

Creation of simulated barcode pixel values: Simulated barcode pixel values were used to benchmark Polaris' barcode assignment method's robustness to dropout. We generated simulated barcode pixel values by sampling from distributions of spot and background pixel values from experimental images. The benchmarked methods include a graphical model of relaxed Bernoulli distributions, a graphical model of multivariate normal distributions, barcode matching by Hamming distance,

and PoSTcode[154]. The relaxed Bernoulli graphical model, the multivariate normal graphical model, and the distance matching method take input pixel values sampled from a distribution simulating spot probability values output by Polaris. Alternatively, PoSTcode takes input pixel values from a distribution simulating pixel values of the raw imaging data, consistent with its original methodology. Our method for simulating barcode values does not consider correlations in pixel values between images acquired in the same fluorescence channel or imaging round. PoSTcode relies on this nuanced characteristic of experimental data; thus, its performance on the simulated data used in this benchmarking analysis may have been negatively impacted.

Mutual nearest neighbors point matching method: To quantitatively benchmark the performance of our deep learning models, we need a method for comparing sets of coordinate spot locations. Our method finds sets of mutual nearest neighbors to compare sets of ground-truth and predicted spot locations. To be considered a true detection, a detection must be within some threshold distance of a ground-truth detection. All spots outside this threshold distance are considered false detections (Supplementary Fig. 3.4a). If more than one detection is within the threshold distance of a ground-truth spot, the detection that is the closest to the ground-truth spot is considered a true detection, and all others are considered false detections (Supplementary Fig. 3.4b). If more than one detection is within the threshold distance for more than one ground-truth spot, edge cases may arise. For a detection to be considered a true detection, that detection and the corresponding ground-truth spot must be mutual nearest neighbors. Therefore, if a detection is within the threshold distance for two ground-truth spots, it is only paired with a ground-truth spot if they are each others' mutual nearest neighbors. Otherwise, the detection is considered a false detection, even if the detection is within the threshold distance of a ground-truth spot (Supplementary Fig. 3.4c).

Spot detection deep learning model architecture

Preparation of coordinate annotations for training data: The coordinate spot locations were converted into two different types of images before being used for deep learning model training. The first image type is a classification image array in which pixels corresponding to spots and background are one-hot encoded. The second image type is a regression image array, in which pixel values correspond to the distance to the nearest spot in the x- and y-direction.

Image preprocessing: We performed preprocessing of images prior to model training. Pixel intensities were clipped at the 0.1 and 99.9th percentiles and then min-max normalized so that all pixel values were scaled between 0 and 1.

Model architecture: Our deep learning model architecture is based on FeatureNets, a previously published backbone implemented in TensorFlow[162] where the receptive field is an explicit hyperparameter[136]. We attach two prediction heads to this backbone: a classification head (to predict the probability a given pixel contains a spot) and a regression head (to predict the distance to the nearest spot with sub-pixel resolution). The receptive field of the network is an explicit hyperparameter that was set to a default value of 13 pixels.

All models were trained with stochastic gradient descent with Nesterov momentum. We used a learning rate of 0.01 and momentum of 0.9. We performed image augmentation during training to increase data diversity; augmentation operations included rotating (0°-180°), flipping, and scaling (0.8X-1.2X) input images. The labeled data were split into training and validation sets, with the training set consisting of 90% and the validation set consisting of 10% of the data. The test set for benchmarking model performance consisted of simulated spot images and held-out experimentally generated spot images.

Image postprocessing: We processed the classification and regression predictions to produce a list of spots. The local maxima of the classification prediction output were determined using maximum intensity filtering, with a default intensity threshold of 0.95. For most datasets, we found that the spot detection results do not vary widely with changes to this threshold value. In the pixels determined to correspond to local maxima, the sub-pixel localization is determined by adding the value of the regression prediction in the x- and y-directions. These sub-pixel locations are returned as the output of the model.

Generation of multiplexed seqFISH dataset in cultured macrophages

Cell culture: THP-1 (TIB-202) cells were received from the American Type Culture Collection. The cells were cultured in Roswell Park Memorial Institute (RPMI) 1640 Medium (Gibco) supplemented with 2 mM L-glutamine (Gibco), 100 U/mL penicillin, 100 μg/mL streptomycin (Gibco or Caisson), and 10% fetal bovine serum (Omega Scientific or Thermo Fisher). To make the complete medium, 2-mercaptoethanol (BME) (Sigma-Aldrich #M6250) was added to a concentration of 0.05mM before every use. Cells were incubated at 37°C in a humidified 5% CO_2 atmosphere and were passaged to maintain a concentration of 0.3-1 x 10^6 cells/mL.

seqFISH sample preparation and imaging: The THP-1 monocyte cells were seeded on a fibronectin-functionalized glass slide (Corning #2980-246) at 80%-90% confluence contained by a rubber gasket (Grace Bio-Labs #JTR8R-2.5). To differentiate the THP-1 monocytes into macrophages, the cells were incubated with 10ng/mL phorbol 12-myristate 13-acetate (PMA) (Sigma-Aldrich #P8139) in RPMI with BME for 24 hours. The media was then replaced with fresh RPMI with BME and incubated for an additional 24 hours. Differentiation was confirmed visually based on changes in adherence and morphology.

The macrophages were dosed with 1ug/mL LPS (Sigma Aldrich #L4524) in RPMI with BME for 3 hours. Cells were rinsed with warm 1X PBS, fixed with fresh 4% formaldehyde (Thermo #28908) in 1X PBS for 10 minutes at room temperature, and then permeabilized with 70% ethanol overnight at -20°C. The primary probe library (Spatial Genomics) was added to the sample in a flow chamber provided by Spatial Genomics and incubated overnight at 37°C. The sample was washed several times with primary wash buffer (Spatial Genomics). The nuclei of the sample were stained with staining solution (Spatial Genomics).

The macrophage sample was imaged with the Spatial Genomics Gene Positioning System (GenePS). Image tiling and secondary probe staining were performed programmatically by this instrument.

Spatial Genomics image analysis: To generate a point of comparison for Polaris' output, we analyzed the seqFISH dataset with the Spatial Genomics software. The spot detection for this analysis was performed via manual parameter tuning. For each imaging round and channel, the threshold intensity used to detect spots was defined visually. The validity of this threshold value was confirmed across several randomly selected fields of view. The DAPI channel was used as the input for their supervised nuclear segmentation method; nuclear masks were dilated to create whole-cell masks.

Multiplex FISH analysis pipeline

Cell segmentation: For cell culture samples, cell segmentation was performed with nuclear and whole-cell segmentation applications from the Deepcell software library. For tissue samples, cell segmentation was performed with Mesmer[138]. The source code for these models is available at `https://github.com/vanvalenlab/deepcell-tf`; a persistent deployment is available at `https://deepcell.org`.

Gene identity assignment: Existing spot decoding methods for image-based spatial transcriptomics fall into two main categories: (1) pixel-wise decoding, which attempts to decode every pixel in the input image, and (2) spot-wise decoding, which attempts to detect spots before decoding them. Polaris' spot decoding method uses elements of both methods. For spot decoding, Polaris' pixel-wise spot probability output was used to determine which pixels to decode. The maximum intensity projection of the spot probability image was performed across all rounds and channels and the set of pixels to be decoded was determined by an intensity threshold with a default value of 0.01. For each thresholded pixel, the array of spot probability values at its coordinate location through the rounds and channels was used as the input for gene assignment.

Gene assignment was performed by fitting a generative model to the probability intensities at identified spot locations, a similar method to previously published work.[154] Our model consists of a mixture of $2 \cdot R \cdot C$ relaxed Bernoulli distributions, where R is the number of imaging rounds in the experiment and C is the number of fluorescent channels in each round. The model for pixel intensities was based on a mixture of relaxed Bernoulli distributions by default, but Polaris also offers two alternative distributions for modeling pixel intensities: Bernoulli and multivariate Gaussian. Therefore, the model consists of a "spot" distribution and a "background" distribution for each imaging round and channel. This model requires two inputs:

1. Spot probabilities at the pixel location across all imaging rounds and channels of the detected spots, as predicted by Polaris' deep learning model

2. A codebook defining the imaging rounds and channels in which each gene in the sample is labeled, referred to as the gene's barcode. An empty barcode is added to the input codebook, corresponding to a "background" — or false positive — assignment.

The distributions are fit to the pixel values of the detected spots with stochastic variational inference[153]. The codebook is used to constrain the logit function of the relaxed Bernoulli distribution, and the temperature is learned. We assume independence across channels and imaging rounds, but the distribution parameters are shared across all genes. For each thresholded pixel, the probability of each barcode assignment is calculated, based on the probability that the pixel value was out of the "spot" or "background" distribution for each round and channel. These probability values are compared to the codebook to calculate the probability of each gene assignment for the pixel. The gene whose barcode has the highest probability is assigned to the pixel. If the highest probability assignment does not exceed a threshold probability value, set to 0.95 by default, the pixel is instead given an "unknown" assignment.

To find the coordinate locations of decoded genes, we create a mask with the pixels successfully decoded to a gene in the codebook and apply this mask to the maximum intensity-projected spot intensity image. We perform peak finding with maximum intensity filtering on the projected image to yield the coordinate location of each decoded gene.

Gene assignment rescue methods: After prediction, "background" and "unknown" assignments can be rescued to be assigned gene identities through two methods. The first method, which we refer to as "error rescue", compares the pixel values of the spot to each barcode in the codebook. If the Hamming distance of the pixel values to a barcode is less than or equal to 1, the assignment is updated. This method catches the rare cases in which the probabilistic decoder misses an assignment. The second method, which we refer to as "mixed rescue", catches the spots that contain two mixed barcodes. This situation occurs when two RNA molecules are in close proximity in the sample and the signal from their barcode labels is mixed. Mixed barcodes often lead to a low probability assignment, so all spots below a threshold probability, set to 0.95 by default, are checked for this case. For this method, the

barcode values for the original assignment are subtracted and the update pixel values are compared to each barcode in the codebook. If the Hamming distance of the pixel values to a barcode is less than or equal to 1, the assignment is updated.

Background masking: Bright objects in the background of smFISH images can interfere with barcode assignment, so detected spots in these regions can be masked out. This step requires a background image without FISH staining to assess the initial fluorescence intensity in the sample. The background image is min-max normalized so that pixels are scaled between 0 and 1, and a mask is created with a threshold intensity, which defaults to 0.5. Any detected spots in the masked regions are excluded from downstream analysis. Polaris also allows a user to manually define a mask for image regions to exclude from analysis.

3.6 Figure captions

Figure 3.1: A weakly supervised deep learning framework for accurate fluorescent spot detection for spatial transcriptomics imaging data. (a) Training data generation for spot detection. Spot labels were generated by finding consensus among a panel of commonly used classical spot detection algorithms through generative modeling. These consensus labels were then used to train Polaris' spot detection model. Sequential steps are linked with an arrow; associated methods and data types are linked with a solid line. (b) Demonstration of the training data generation for an example spot image. Spot locations are converted into encoded detections and distance maps which guide the classification and regression tasks performed during model training. Spot colors correspond to the annotation colors in (a). (c) Output of Polaris' spot detection model for an example seqFISH image. Regression values above a default threshold are set to zero. The regression images in (b-c) are the sum of the squared pixel-wise regression in the x- and y-directions. (d) Schematic diagram of the EM method to fit the generative model for consensus spot annotation creation.

Figure 3.7: Benchmarking model performance on simulated spot images with a range of spot intensities and densities. (a) Performance quantification for models with Polaris' deep learning architecture trained with various datasets predicting on images with a range of spot density. (b) Performance quantification for models with Polaris' deep learning architecture trained with various datasets predicting on images with a range of levels of simulated noise. The low noise condition corresponds to a signal-to-noise ratio of greater than approximately 16. The medium noise condition

corresponds to a SNR of approximately 8-15. The high noise condition corresponds to a SNR of approximately 3-7. (c) Performance quantification for various spot detection methods on images with a range of spot densities. (d) Performance quantification for various spot detection methods on images with a range of levels of simulated noise. The SNR ranges for each noise condition are the same as those in (b).

Figure 3.8: Polaris produces single-cell, spatial gene expression maps for multiplex spatial transcriptomics images. (a, b) Analysis steps for Polaris for singleplex (red) and multiplex (blue) spatial transcriptomics imaging data. Sequential steps are linked with an arrow, and associated methods and data types are linked with a solid line. Deep learning models perform spot detection and cell segmentation, while a probabilistic graphical model infers gene identities. (c) A probabilistic graphical model for inferring gene identities from spot detections. This model consists of a mixture of K relaxed Bernoulli distributions, parameterized by their probability, α, and their temperature, λ, for generating observed data, x, of size N spots. Shaded vertices represent observed random variables; empty vertices represent latent random variables; edges signify conditional dependency; rectangles ("plates") represent independent replication; and small solid dots represent deterministic parameters. (d) Spatial organization of marker gene locations in a mouse ileum tissue sample. Each spot corresponds to a decoded transcript for a cell type marker gene. Whole-cell segmentation was performed with Mesmer[138]. (e-g) Locations of decoded genes in an example Goblet cell, enterocyte, and B cell, respectively.

Figure 3.10: Demonstration of Polaris' performance on a MERFISH and seqFISH data. (a) Example Polaris prediction for a MERFISH experiment in a mouse kidney tissue sample (Liu et al. 2022). (b) Example Polaris prediction for a seqFISH experiment in a macrophage cell culture sample. The spot colors of the Polaris prediction in (a,b) denote the predicted gene identities. The inset image location is defined by the white box in the full field of view (FOV).

Figure 3.11: Correlation of Polaris' quantification of MERFISH data with other quantification methods. (a,c) Scatter plot plotted in logspace, comparing gene expression counts quantified with MERFISH with counts measured with RNA-seq. (b,d) Scatter plot plotted in logspace, comparing previously published MERFISH gene expression counts with counts quantified with Polaris. (e) Scatter plot plotted in logspace, comparing mean gene expression counts per cell quantified with seqFISH with counts measured with RNA-seq. (f) Scatter plot plotted

in logspace, comparing mean gene counts per cell obtained by manual analysis of seqFISH data with gene counts quantified with Polaris.

Chapter 4

PAIRED LIVE-CELL IMAGING AND SEQFISH MEASUREMENTS OF NF-κB SIGNALING IN PRIMARY HUMAN MACROPHAGES

4.1 Introduction

Macrophages are innate immune cells deposited in tissues during development that perform a diverse array of tissue-specific roles.[163,164] They have an important immunological role, acting as the first line of defense against pathogens and clearing foreign materials. They are also thought of as "gatekeepers of tissue homeostasis," responsible for the resolution of inflammation and tissue self-renewal.[165] Macrophages are able to perform diverse roles because of their highly plastic cell state that informs their inflammatory and metabolic functions.[3–5]

Macrophages are highly dynamic cells that can be readily polarized to a wide variety of cell states depending on environmental stimuli.[165–167] They adeptly encode the identity and strength of perceived stimuli as dynamic patterns of immune signaling pathway activation, which have been linked previously to unique transcriptional profiles in macrophages.[16,168] Therefore, to characterize the mechanisms by which macrophages perceive and encode environmental information, their signaling response must be measured dynamically and at the single-cell level. Moreover, because of the influence of their environment on their cell state, measurements of macrophage signaling with spatial context can reveal the cell-to-cell communication shaping their responses.

In this work, we integrated dynamic observation of the signaling response of primary human macrophages to lipopolysaccharide (LPS) treatment with an end-point measurement of the expression of a panel of inflammatory genes (Fig. 4.1a). We utilized deep learning-based image analysis methods to generate paired signaling pathway activation dynamics and spatial transcriptome data at the single-cell level (Fig. 4.1b-d). We demonstrated live-cell reporter expression in primary human macrophages, adapting a toolbox of imaging methods that were previously limited to immortalized cell systems to primary cells. We leveraged live-cell reporters in primary macrophages to illustrate their capability to perceive and encode different LPS concentrations at the single-cell level. We then utilized seqFISH to measure

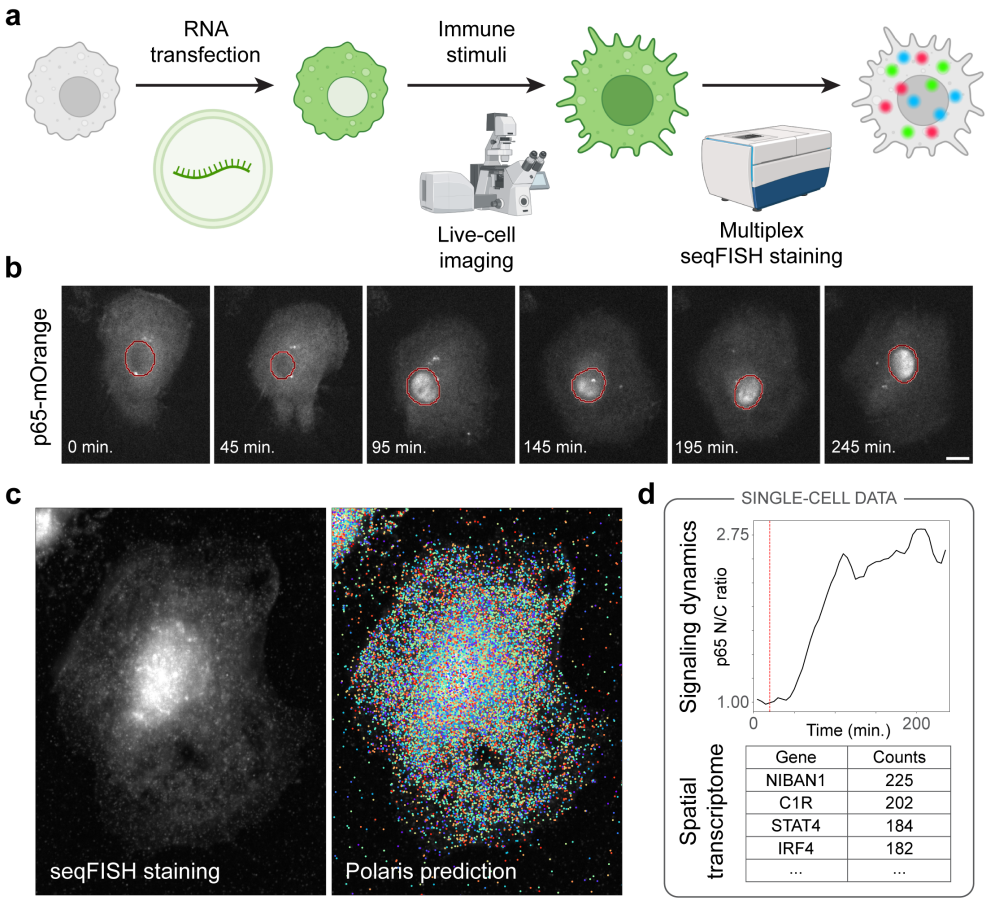


Figure 4.1: **Overview of integrated measurements method in primary human macrophages.** See section 4.5 for caption.

spatial gene expression patterns to reveal distinct morphological and transcriptional subpopulations of cells. Finally, in an integrated measurement of paired live-cell imaging and seqFISH, we demonstrated that the transcriptional subpopulations also have distinct dynamic patterns of immune signaling pathway activation in response to LPS. With this work, we aimed to demonstrate how cell state informs an individual macrophage's response to LPS treatment and how this response is informed by intra-cellular communication from its neighbors.

4.2 Results

Single-cell morphological heterogeneity in cultured macrophages

Macrophages utilize cell-to-cell heterogeneity to perform diverse, specialized roles in tissues. It has been well characterized that macrophages spontaneously form two distinct morphological populations, linked to their inflammatory states. The

"round" morphotype is characterized by a smaller, more circular morphology that has been associated with a pro-inflammatory signaling state. The "spindle" morphotype is characterized by an elongated appearance and is associated with an anti-inflammatory cell state.[169–173] These cell states are highly plastic and allowing macrophages to take on roles determined by environmental stimuli.[165–167] Considering this relationship between a macrophage's morphology and its inflammatory state, we sought to develop an imaging-based assay that could capture a macrophage's immune signaling state with spatial context.

We quantified the morphologies of primary human macrophages using a set of previously published morphology metrics.[174] (Fig. 4.2) A considerable amount of heterogeneity in macrophage morphology is found on a circularity-eccentricity axis. Circular cells are round and have relatively similar values for their major and minor axis lengths. On the other hand, eccentric cells come to sharp points, meaning the distance between their major foci is nearly the distance of their major axis. The macrophages also vary in size, which can be quantified with two related metrics, area and perimeter. A cell's aspect ratio is an alternative metric to quantify its elongation, calculated as the ratio of the lengths of its major and minor axes. Finally, the solidity of a cell is calculated as the ratio of a cell's area and the area of its convex hull.[174]

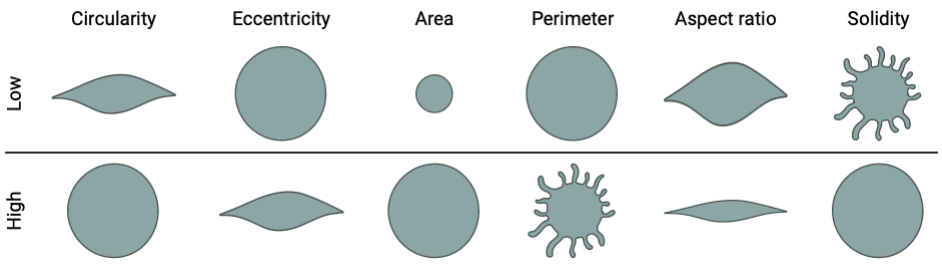


Figure 4.2: **Metrics for morphology quantification.** Example cell morphologies demonstrate the axis of variation the metric captures.

We utilized these metrics to identify subpopulations of macrophages with distinct morphologies. In naive macrophages, we observed a mixture of round and spindle macrophages, in alignment with previous work characterizing macrophage morphologies.[169–173]. (Fig. 4.3a) We utilized the morphology metrics[174] and find that the two populations are particularly distinct in the distributions of cell circularity and eccentricity. (Fig. 4.3d)

We further explored morphology as a function of inflammatory cell state by treating naive macrophages with polarizing cytokines. Once treated with a pro-inflammatory cytokine, IFNγ, the circularity of the macrophages increased and their eccentricity decreased, reflecting polarization towards the round morphological state (Fig. 4.3b,e). Once treated with an anti-inflammatory cytokine, IL4, the circularity and solidity of the macrophages sharply decreased and their eccentricity and aspect ratio sharply increased, indicating strong polarization towards the spindle morphological state (Fig. 4.3c,f). These results confirm that the observed morphotypes align with the field's current understanding of the morphological characteristics of macrophages in different inflammatory states. (Fig. 4.3d)

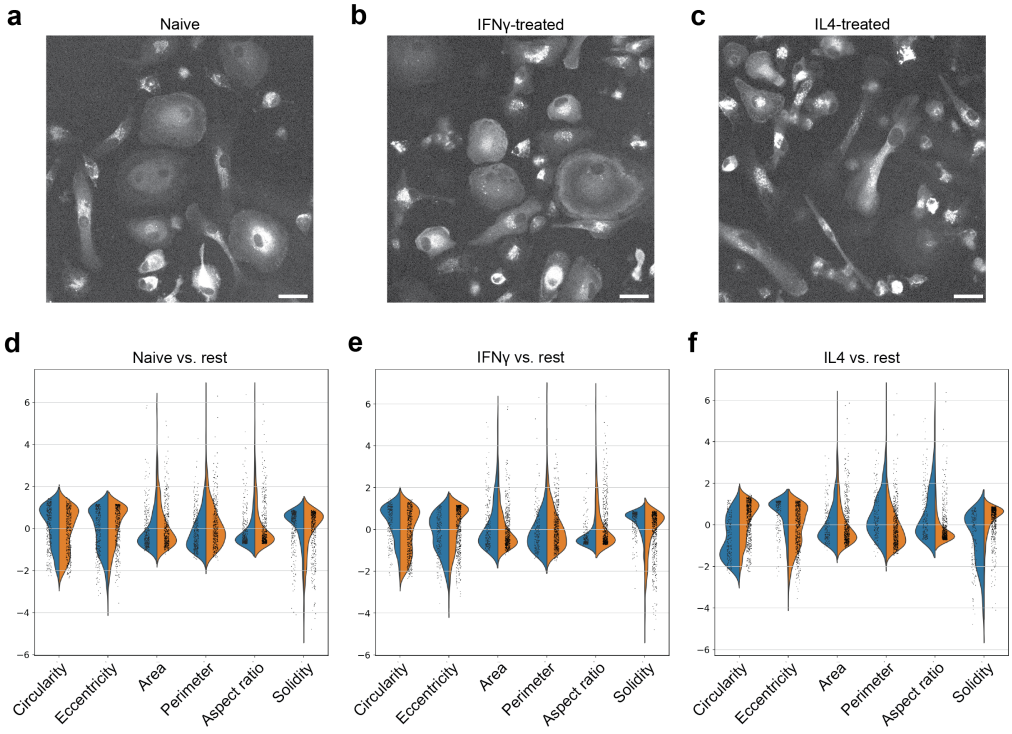


Figure 4.3: **Treatment with pro-inflammatory and anti-inflammatory cytokines alters the morphology of primary macrophages.** (a,c,e) Images of primary macrophages in naive, IFNγ-treated, and IL4-treated states, respectively. Scale bar corresponds to 30 μm. (b,d,f) Violin plots comparing the morphology metrics of primary macrophages in naive (n=348), IFNγ-treated (n=226), and IL4-treated (n=189) states (blue), respectively, to the rest of the cells from other conditions (orange). All metrics were normalized by removing the mean and scaling to unit variance.

This characterized heterogeneity in macrophage morphology and its relationship to a cell's inflammatory state was a key motivation for our image-based investi-

gation of their information processing capabilities at the single-cell level. In this work, we leveraged live-cell imaging and spatial transcriptomics to reveal the relationship between a macrophage's morphology and its response to environmental stimuli.[169–173]

Modified RNA transfections in primary macrophages

We exogenously expressed fluorescent live-cell reporters to observe dynamic signaling pathway activity at the single-cell level by mRNA transfection. Plasmid DNA transfection is the most common method for transient expression of exogenous constructs. However, monocytes and macrophages are classically difficult to transiently transfect for a number of reasons.[175] First, in order to be transcribed, transfected plasmid DNA needs to be delivered to the nucleus of cells, which typically occurs during nuclear envelope breakdown in cell division. As non-proliferative cells, nuclear delivery of plasmid DNA remains challenging for macrophages.[176,177] To address this consideration, we utilized a protocol that transfects macrophages with *in vitro* synthesized messenger RNA (IVT-mRNA), which can be expressed from the cell's cytoplasm.[178] Second, macrophages are programmed to recognize and mount an immune response to exogenous DNA or RNA, which would typically be an indicator of viral infection.[179] It has been demonstrated that chemical modifications to the nucleotides used to synthesize the IVT-mRNA can allow it to evade immune detection.[180] Therefore, we synthesize IVT-mRNA with pseudouridine and 5'-methyl cytidine to minimize the immunogenicity of its transfection (Fig. 4.4a).[178]

We extended this protocol to the expression of a variety of constructs, such as transcription factor fluorescent fusion proteins, kinase translocation reporters (KTRs), and ratiometric ExRai reporters, demonstrating the general applicability of this method to a range of live-cell measurements in primary human macrophages. (Fig. 4.5) For more information on the sampled live-cell reporter architectures, see section 2.1. The efficiency of IVT and transfection as a function of construct length could have limited the broad applicability of this live-cell reporter expression protocol. However, this panel demonstrates the robustness of the modified RNA transfection method for reporters ranges in construct size from about 800 to over 3,000 base pairs, which covers a considerable amount of the size range applicable to live-cell reporters.[49]

We utilized this IVT-mRNA transfection protocol to achieve efficient simultaneous transfection of two fluorescent constructs. (Fig. 4.6a) Transfection with lipofec-

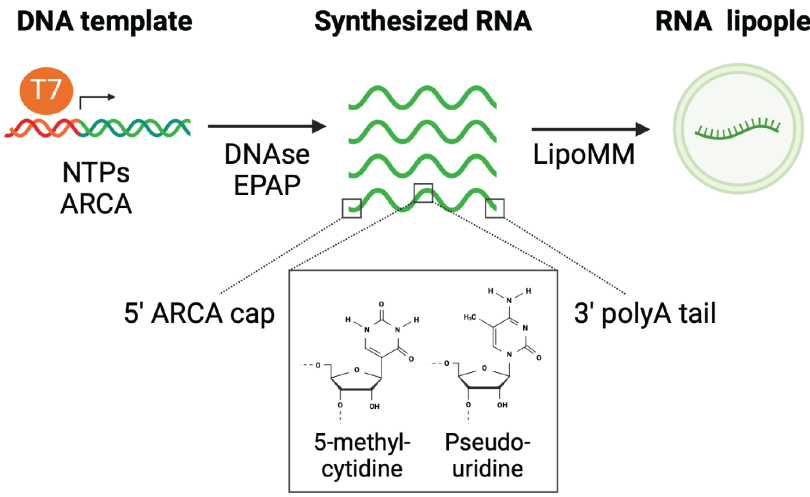


Figure 4.4: **Synthesis and transfection of modified mRNA transcripts encoding live-cell reporters in primary macrophages** Schematic diagram of modified mRNA *in vitro* synthesis and transfection. RNA is synthesized with T7 polymerase and modified nucleotides from a DNA template containing a T7 promoter. IVT mRNA is capped at its 5' end with ARCA and tailed at the 3' end with EPAP. The mRNA is transfected into cultured cells with LipoMM.

tamine requires a micelle formation step during which the negatively charged nucleic acid material is encapsulated with positively charged lipids allowing entry into a cell's plasma membrane.[181,182] We mix the two mRNA constructs before micelle formation to promote the formation of micelles with both mRNA constructs to improve the double transfection rate.

For the purposes of this work, the simultaneously transfected constructs were an mOrange fusion protein of p65 and a KTR utilizing mClover. When transfecting KTR-mClover and p65-mOrange, 62.3% cells are double transfected, and 76.3% of cells are expressing at least one of the fluorescent proteins. (Fig. 4.6b) This efficient transfection rate demonstrated the viability of the modified RNA transfection method for efficient characterization of paired signaling dynamics with live-cell reporters in primary macrophages. Furthermore, as opposed to observing many cells strongly expressing only one fluorescent construct, we observed a strong correspondence of the expression of both fluorescent constructs, indicating that micelles containing a mixture of the mRNA constructs formed. This result demonstrated the amenability of modified RNA transfections in primary macrophages, enabling the application of a set of imaging tools previously limited to immortalized cell systems.

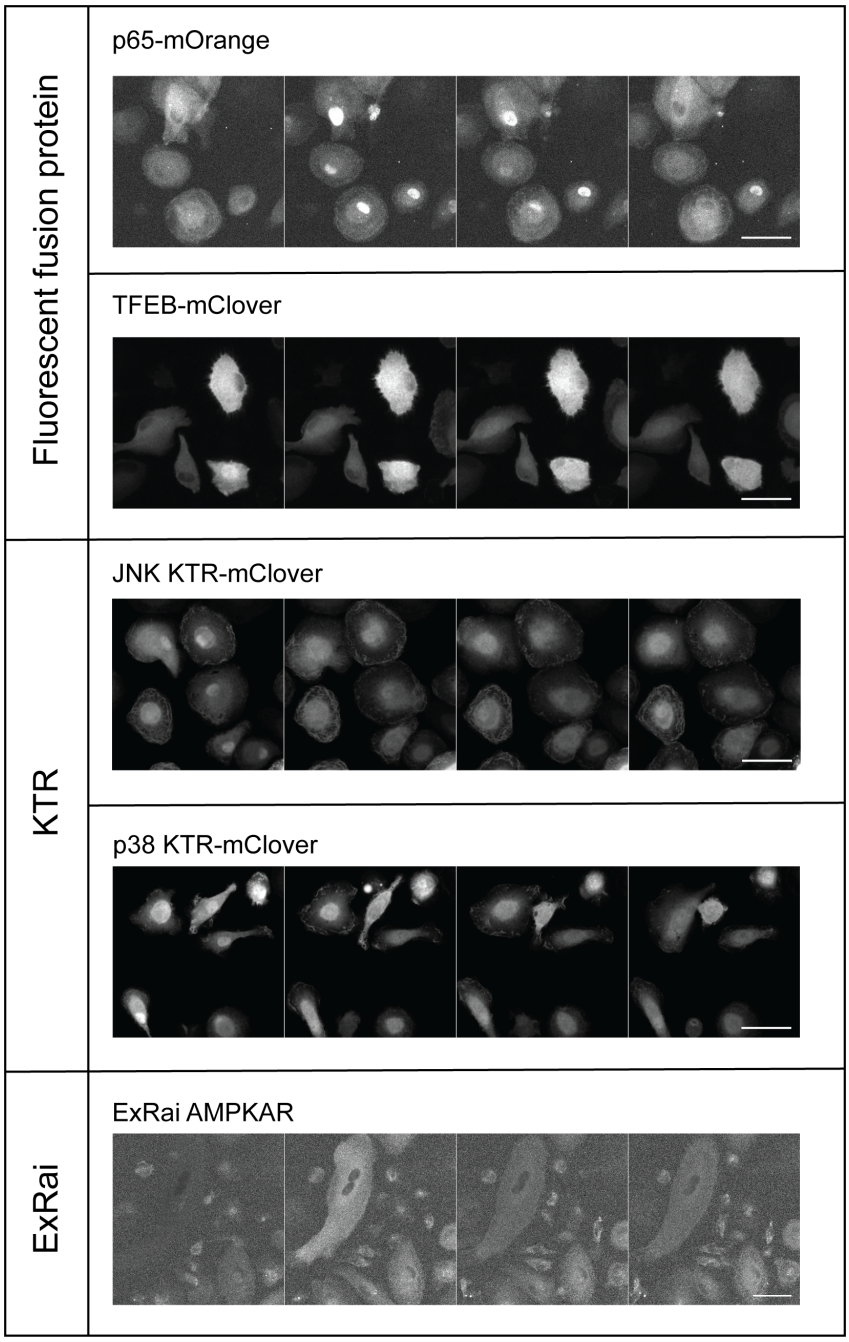


Figure 4.5: **Expression of a panel of live-cell reporters in primary macrophages with modified RNA transfection.** See section 4.5 for caption.

Immune signaling dynamics in LPS-treated primary human macrophages

As was introduced in section 1.2, dynamic patterns of signaling pathway activation are a critical mechanism by which cells encode environmental stimuli, leading to distinct cellular behaviors and phenotypes. We observed these mechanisms

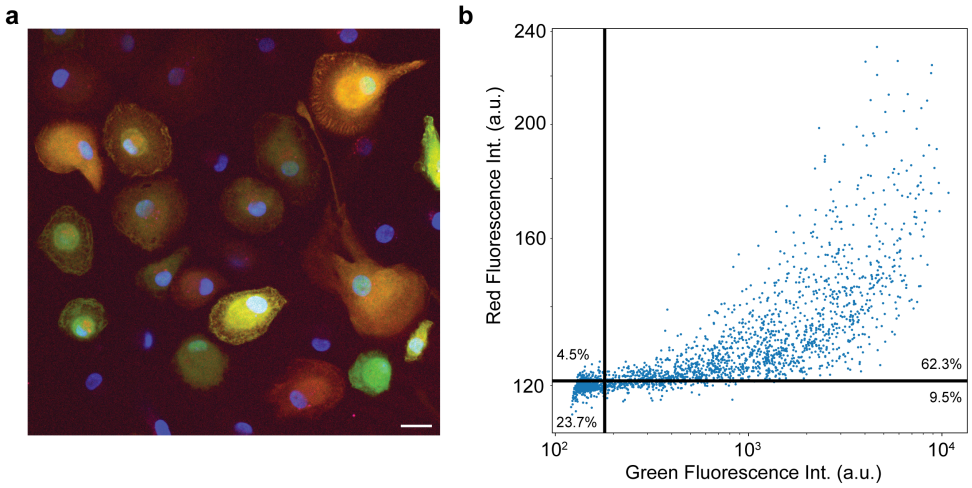


Figure 4.6: **Simultaneous transfection of two mRNA constructions in primary macrophages.** (a) Composite image of primary macrophages transfected with JNK KTR-mClover (green) and p65-mOrange (red), stained with Hoechst (blue). Scale bar corresponds to 20μm. (b) Scatter plot of fluorescence intensity values for mClover (green) and mOrange (red) in arbitrary units, plotted in log scale. (n=2490 cells, reps=2) Plotted lines are threshold values for transfected cells.

of information processing with fluorescent live-cell reporters, which offer unique insight into real-time patterns of signaling pathway activation at the single-cell level.

The Toll-like receptor 4 (TLR4) signaling pathway plays a critical role in the inflammatory response of macrophages to bacterial infection.[184,185] TLR4 is a pattern recognition receptor for LPS, a cell wall component of gram-negative bacteria.[186] Receptor binding triggers a cascade of well-characterized signaling events that diverge into two parallel signaling pathways, the nuclear factor κ-light-chain-enhancer of activated B cells (NF-κB) pathway and mitogen-activated protein kinases (MAPK) pathways, including p38 and c-Jun N-Terminal Kinase (JNK) activation.[187–189] As was covered in section 2.3, the NF-κB family of transcription factors consist of a variety of hetero-dimer constructs. The primary complex that triggers the activation of inflammatory genes in response to TLR4 activation is a hetero-dimer of p65 and p50.[9,103] To observe the translocation of p65 from the cytoplasm to the nucleus upon TLR4 activation, we expressed a fluorescent fusion protein consisting of p65 and mOrange. An analogous fusion construct with a different fluorescent protein has been previously used to measure p65 translocation dynamics.[156] The MAPK signaling pathway consists of a complex cascade of signaling

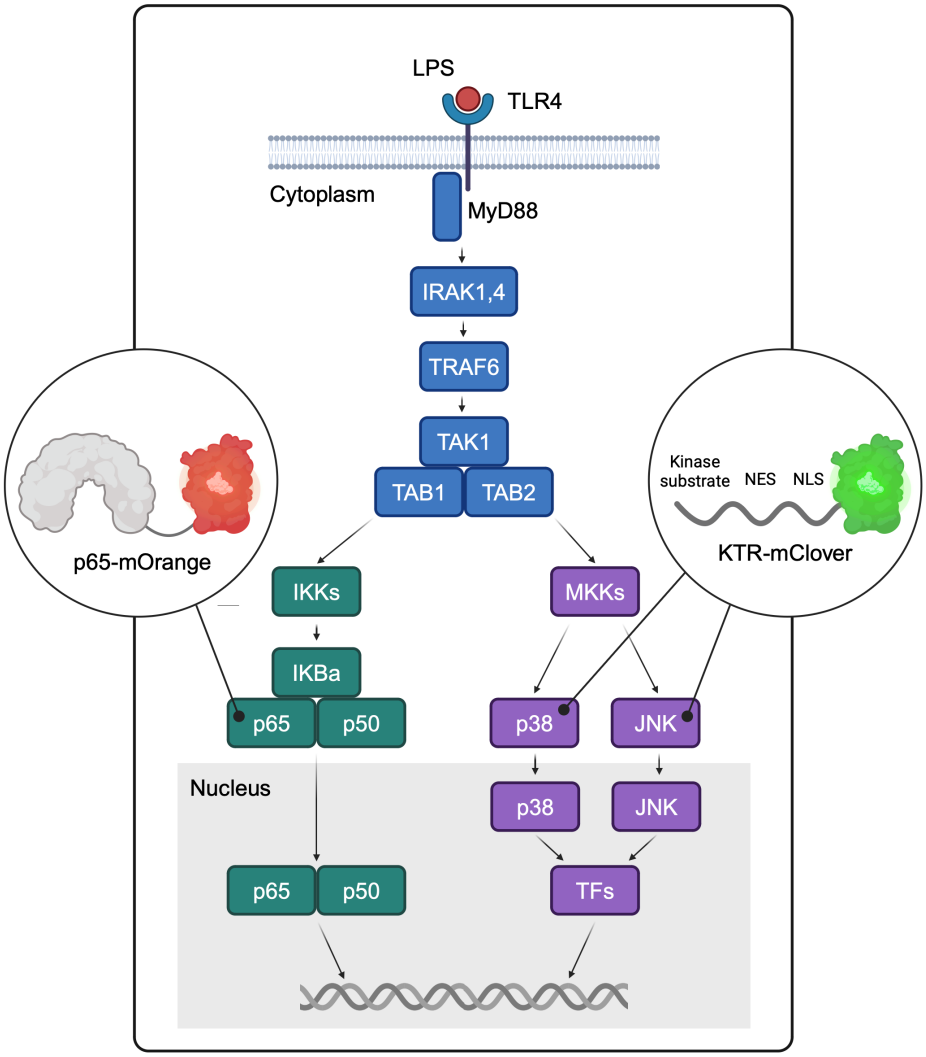


Figure 4.7: **Fluorescent live-cell reporters for dynamic observation of the TLR4 signaling pathway.** The TLR4 signaling pathway is activated by LPS binding to TLR, triggering a diverging cascade of downstream kinases. The TAK1-TABs complex is a central signaling hub that activates both NF-κB (p65) and MAPK (p38 and JNK) translocation and activation of target genes.[183] We observe p65 and MAPK activity with fluorescent live-cell reporters.

kinases that control cell growth and proliferation. MAPK proteins like p38 and JNK translocate to the nucleus upon activation, where they phosphorylate and thus activate target transcription factors.[190,191] To quantify the activation of these pathways, we expressed KTRs, which have been robustly engineered and characterized to report the translocation of p38 and JNK.[70] (Fig. 4.7)

In response to LPS treatment, we observed distinct activation of the NF-κB, p38, and

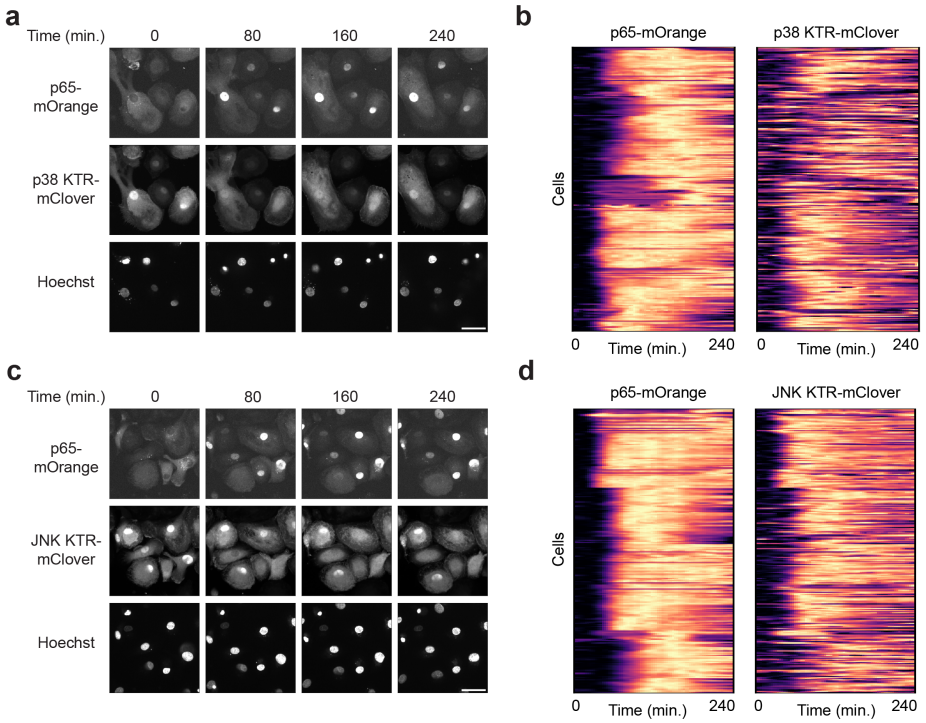


Figure 4.8: **Primary macrophages simultaneously activate NF-κB and MAPK signaling.** (a,c) Simultaneous response of p65-mOrange and p38 KTR-mClover or JNK KTR-mClover, respectively, to 100 ng/mL LPS treatment. Nuclear areas were visualized with Hoechst staining. Scale bar corresponds to 20 μm. (b,d) Single-cell nuclear p65 intensity and p38 KTR or JNK KTR, respectively, activity dynamics, presented as tandem heatmaps. KTR activity was quantified as the ratio of the cytoplasmic to nuclear fluorescence intensities.[70] Values are normalized between zero and one on a per-trace basis for each reporter.

JNK signaling pathways. As expected, LPS treatment induced sharp and sustained nuclear translocation of p65-mOrange. In the same cells, p38 and JNK activate simultaneously with p65, aligning with our molecular understanding of the NF-κB and MAPK pathways diverging from a unified upstream signal from TLR4 activation. (Fig. 4.8)

When quantifying the nuclear intensity of p65-mOrange, we observed a wide range of overlapping intensity levels in treated and untreated cells. (Fig. 4.9a) However, we observed greater resolution between LPS-treated and untreated cells in the fold-change of p65 intensity relative to levels prior to LPS-treatment. We observed largely overlapping distributions of p65 intensity, maximum p65 intensity, derivative of p65 intensity, and integral of p65 intensity from the absolute nuclear p65 intensity traces. On the other hand, these distributions were much more resolved for the traces of

fold-change of p65 intensity, in which the absolute intensity values are normalized by the initial nuclear intensity of p65. (Fig. 4.9b-e) We also compared the fold-change dynamics to values normalized by the cytoplasmic p65 intensity, as a measure of expression level, prior to LPS treatment. (Fig. 4.18) We found that the fold-change normalization of p65 intensity values better resolved distributions of metrics of p65 translocation dynamics. (Fig. 4.19)

Our observation aligns with previous work has demonstrated that the fold-change of NF-κB activity is more consequential for the expression of target genes than the absolute intensity.[45] This observation of greater differentiation between stimuli by fold-change dynamics over absolute intensities led us to perform analysis inspired by previous work that utilized information theory to demonstrate greater signal transduction in the Tgf-β pathway by fold-change levels of nuclear Smad3.[192]

As was discussed in section 1.2, the key purpose of signaling pathways is to allow a cell to assess and encode the identity and strength of environmental stimuli. Therefore, we expect different identities and concentrations of immune stimuli to elicit varying dynamic patterns of signaling activation in macrophages. The lack of resolution between dynamics metrics calculated with the absolute p65 intensity for LPS-treated and untreated cells indicates that the absolute intensity of nuclear p65 does not confer much information about the environmental concentration of LPS. However, the greater resolution in the distributions calculated with fold-change intensity implies that cells may be able to better sense environmental LPS concentrations with nuclear p65 intensity normalized by their basal activity level.

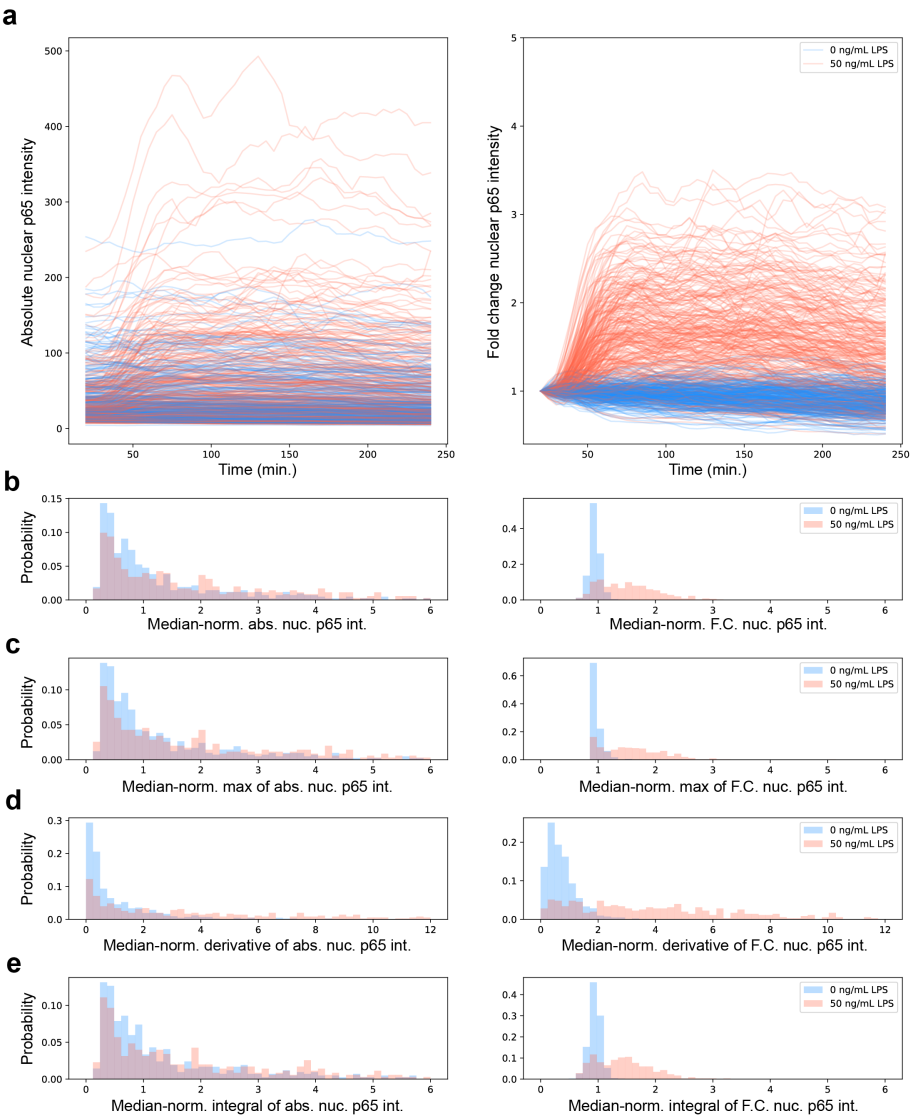


Figure 4.9: **Fold-change of nuclear p65 intensity shows greater resolution of LPS stimulation conditions than absolute intensities.** See section 4.5 for caption.

We can quantified the degree of information transmission by absolute and fold-change nuclear p65 intensities with a metric called mutual information (MI). MI is a general measure of mutual dependence between two variables and does not assume a specified underlying distribution to the data. MI is expressed as:

$$I(X;Y) = D_{KL}(P_{(X,Y)}||P_X \times P_Y) \tag{4.1}$$

where D_{KL} is the Kullback-Leibler (KL) divergence, defined for two probabilities A and B in its discrete form as:

$$D_{KL}(A||B) = \sum_x A(x) log\left(\frac{A(x)}{B(x)}\right). \tag{4.2}$$

In this case, if X represents the concentration of LPS and Y represents a metric of signaling pathway activation, the KL-divergence measures how different the joint distribution $P(X,Y)$ is from the product of the marginal distributions $P(X)$ and $P(Y)$. If $P(X) \times P(Y) = P(X,Y)$, X and Y are independent and the KL-divergence is zero, indicating no MI between X and Y. However, when $P(X) \times P(Y)$ and $P(X,Y)$ diverge, the KL-divergence is non-zero, indicating dependence between X and Y, or said another way, MI between X and Y.

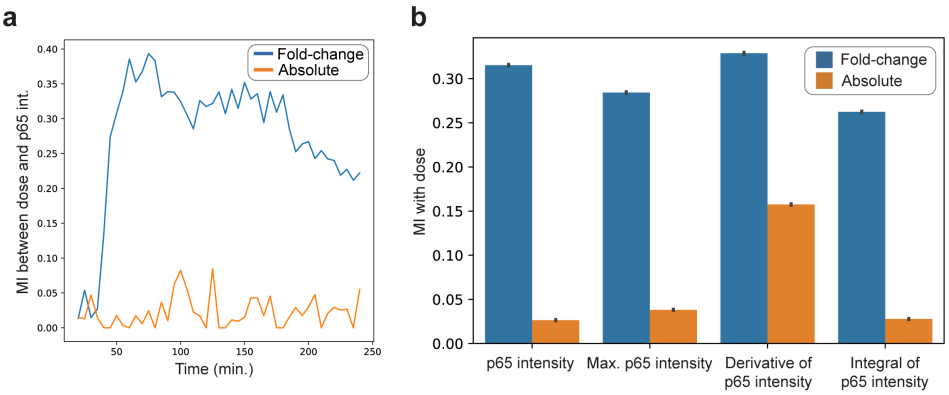


Figure 4.10: **Fold-change of p65 intensity has higher mutual information with LPS dose than absolute p65 intensity** (a) Mutual information between absolute p65 intensity or fold-change p65 intensity and LPS dose at each observed time point for four dosing conditions (0 ng/mL, 1 ng/mL, 10 ng/mL, and 50 ng/mL LPS). (b) Mutual information between metrics of p65 dynamics for absolute and fold-change intensities for the four dosing conditions. Bars are plotted with 95% confidence intervals. Confidence intervals were calculated by bootstrap sampling. For each metric, 10,000 sets of 500 cells were sampled. MI is displayed in units of bits.

In agreement with greater resolution between doses for the fold-change p65 intensity, we found the greater MI between p65 activity and the LPS dose for the fold-change time series than the absolute p65 intensity in four LPS treatment conditions, 0 ng/mL, 1 ng/mL, 10 ng/mL, and 50 ng/mL. (Fig. 4.10a) The greatest MI occurred around 40 minutes into the time series, which aligns with the typical timing of peak p65 translocation. When p65 translocation is summarized with metrics of activation, we observed a similar trend, with the fold-change time series showing greater MI with the LPS dose than the absolute intensity. (Fig. 4.10b) These results confirm that individual macrophages are able to perceive and encode the environmental

concentration of LPS. Moreover, they confirm that these metrics are able to capture the aspect of p65 translocation dynamics that cells use to transmit information.

We also quantified the MI between p65 dynamics and MAPK dynamics, which captures non-linear dependence between the two variables unlike other metrics like Pearson correlation. MI is an appropriate metric for capturing the heterogeneity shared between two signaling pathways, because it does not have inherent directionality, like conditional dependence. (Fig. 4.11a) Therefore, it does not make assumptions about the underlying architecture of the signaling network or the expected distributions for each variable. These characteristics make MI an ideal metric for quantifying information transmission in signaling networks, which operates non-linearly through molecular cascades and feedback motifs.

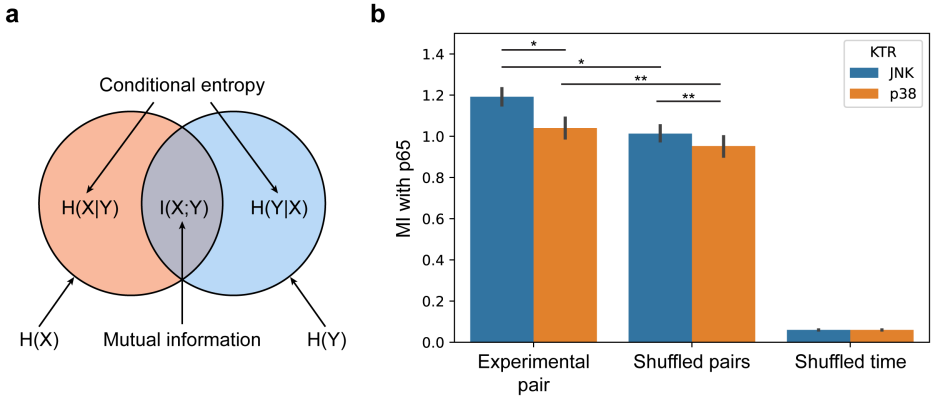


Figure 4.11: **MAPK and p65 signaling dynamics have significant mutual information in response to LPS treatment.** (a) Schematic illustrating the relationship between entropy, conditional entropy, and mutual information. (b) Mutual information (MI) between p65 dynamics and JNK or p38 dynamics. The experimentally observed pair of dynamics from single cells demonstrate the highest MI. (JNK, n = 440 cells; p38, n = 292 cells) Pairs of dynamics from two random individual cells shows moderate MI, and time series shuffled across the time dimension show no MI. Bars are plotted with 95% confidence intervals. MI is displayed in units of bits. (*p < 1×10^{-5} , **p < 0.05; unpaired t-test, two-tailed)

We treated the dynamic activation of the p65 and MAPK signaling pathways as random variables and found significant MI between p65 dynamics, and both JNK and p38 dynamics. (Fig. 4.11b) This observation corresponds with our intuition for these signaling pathways, which are initially activated by the same cascade of signaling events before diverging into two separate pathways. Therefore, we expected these activation of both the NF-κB and MAPK signaling pathways to be

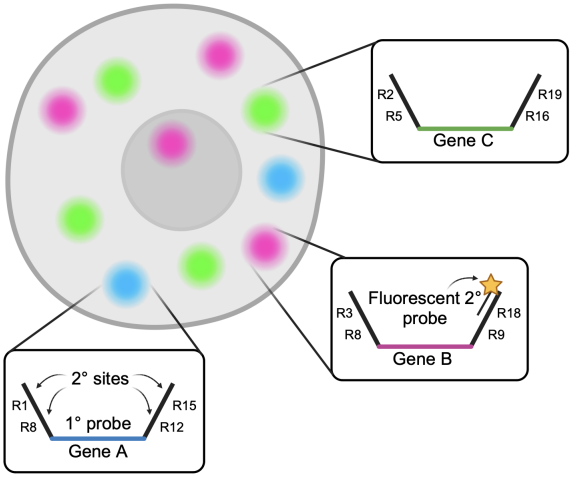
subjected to the same extrinsic noise in LPS stimulus and intrinsic noise in the cells' amenability to TLR4 signaling activation.

We demonstrated the significance of these values by comparing the experimentally observed pairs of signaling dynamics from individual cells to shuffled versions of the paired data. First, we compared the experimental pairs to pairs of dynamics from two random cells, which we refer to as the "shuffled pairs" experiment. We found a statistically significant difference between the MI values for experimental pairs and shuffled pairs. This significant difference indicates that the signaling dynamics likely encode the extrinsic noise in the environmental stimulus, and the effect of intrinsic noise in cell state on a cell's signaling response. However, there is a considerable amount of MI between shuffled pairs, which we attribute to the inherent similarity in the reactions of clonal cells to a mutually experienced stimulus. Second, we compared the experimental pairs of signaling dynamics to those shuffled across the time dimension. As expected, randomly shuffling across time abrogates any mutual information between the paired time series and serves as a negative control for our MI metric.

Spatial transcriptomics measurement in primary human macrophages with seqFISH

Spatial transcriptomics methods measure gene expression while retaining the spatial organization of the sample. Often this capability is utilized to measure gene expression in communities of cells in a tissue sample.[22,23] seqFISH measures gene expression while retaining the spatial organization of the sample by labeling mRNA transcripts with fluorescent oligo probes. It targets the mRNA molecules with single-stranded DNA oligos containing a complementary sequence flanked by binding sites for fluorescent secondary probes. It achieves signal amplification by staining each mRNA molecule with multiple primary probes. The secondary probes specifically bind with the primary probes and have been optimized to minimize off-target binding. Each primary probe contains multiple secondary binding sites that are labeled in specific imaging rounds. In this way, seqFISH utilizes a temporal barcoding scheme, where the gene identity is encoded by the imaging rounds in which an mRNA molecule is stained, to measure panels of up to thousands of genes.[32] (Fig. 4.12)

By retaining the cell-to-cell spatial organization of the sample, spatial transcriptomics method like seqFISH allows us to use the spatial organization of the sample



Codebook

Gene	A	B	C
R1	1	0	0
R2	0	0	1
R3	0	1	0
...	...	...	...
R18	0	1	0
R19	0	0	1
R20	0	0	0

Figure 4.12: **seqFISH measures gene expression by targeting DNA oligo probes to mRNA transcripts.** The seqFISH barcoding scheme targets mRNA molecules with two sets ssDNA oligo probes: (1) primary probes which bind directly to the RNA and contain flanking secondary binding sites, and (2) fluorescent secondary probes for visualization, demonstrated in the panel for gene B. Each primary probe contains four secondary binding sites, for staining in sequential imaging rounds, demonstrated in the panel for gene A. The codebook defines the stained imaging rounds that correspond with a gene identity.

to retain a cell's identity between a live-cell imaging measurement and a gene expression measurement. Leveraging this capability allows paired live-cell imaging with seqFISH to enable a high-throughput single-cell measurement that integrates signaling dynamics with the spatial expression patterns of a large panel of genes.

We assembled a panel of genes for the seqFISH measurement that are strongly differentially expressed as a function of LPS treatment. The panel was constructed based on a bulk RNA sequencing dataset that measured the transcriptional response of THP-1 cells, an immortalized macrophage line, to LPS treatment, selecting of the most strongly differentially expressed genes.[193] Consistent with an inflammatory response to LPS treatment, the panel mainly consists of a few categories of genes: cytokine genes, interferon signaling genes, metabolic genes, morphology genes, immuno-regulatory genes, and transcriptional regulatory genes.

As expected, we observed a distinct change in the transcriptomes of primary human macrophages upon LPS treatment. Principal component analysis (PCA) revealed that the transcription of cells in the LPS-treated and untreated conditions are distinct, comprising of two separate clusters. (Fig. 4.13a) The LPS-treated cluster of cells

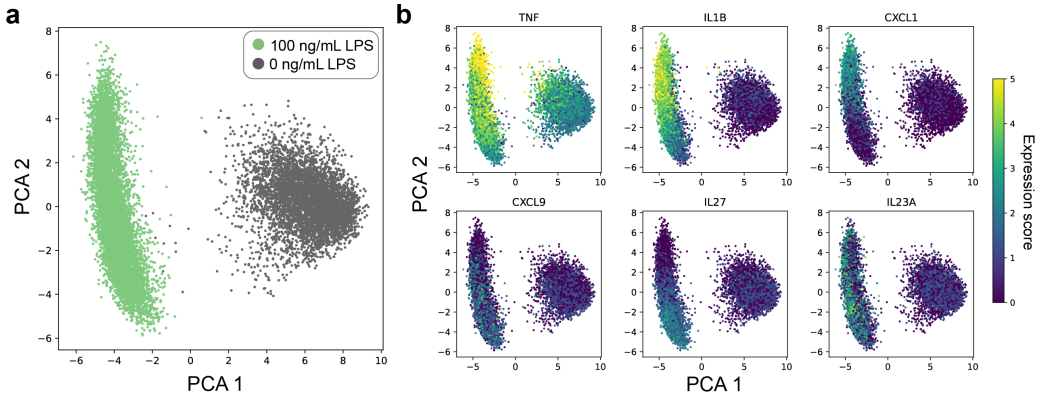


Figure 4.13: **LPS treatment strongly alters the transcriptome of primary macrophages.** (a) PCA of gene expression quantified with seqFISH. LPS-treated (green) and untreated (gray) cells cluster separately. (b) Expression of inflammatory cytokines (TNF, IL1B, CXCL1, CXCL9, IL27, and IL23A) is strongly activated by LPS treatment. Expression score is the natural logarithm of CPM-normalized counts.

strongly expressed a number of cytokines (e.g., TNF, IL1B, CXCL1, CXCL9, IL27, and IL23A), which aligns with an inflammatory cellular response. (Fig. 4.13b).

As a spatial transcriptomics methods, seqFISH has the distinct advantage of retaining the morphology of the cells in a sample, allowing a cell's transcriptome to be linked to its morphological state. We observed that the previously described round and spindle macrophages have distinct transcriptional patterns. (Fig. 4.14a) The round macrophages demonstrated higher expression of TNF, IL27, and PIM1, whereas spindle macrophages demonstrated higher expression of IL1B and CXCL1, amongst other differentially expressed genes. (Fig. 4.14b) The differential expression of these genes are a part of a larger pattern of transcriptional differences between the two morphological states. (Fig. 4.14c,d)

Our observation of subpopulations of macrophages aligned with previous work that demonstrated that TNF and IL1B are expressed by distinct subpopulations of macrophages after ischemic stroke in mice.[194,195] The expression of TNF and IL-1B in separate populations of cells likely contributes to a mechanism of the induction of inflammation that relies on synergy between TNF and IL1B.[196,197] The redundancy between the induction of inflammation by TNF and IL1B allow the cells to coordinate a rapid yet robust immune response at the population level.[198] The differential expression of TNF and IL-1B is controlled by differences in their promoter states and transcription factor requirements. IL-1B transcription in particular is influenced

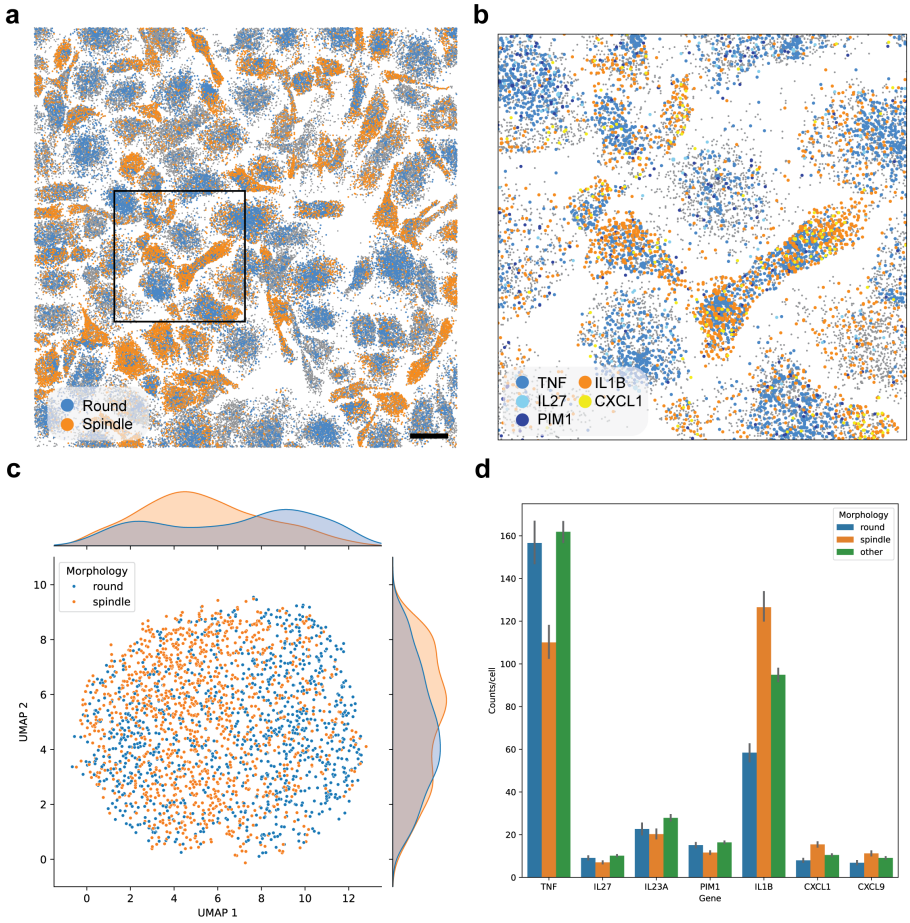


Figure 4.14: **Morphological subpopulations of macrophages have distinguishable patterns of gene expression.** (a) Scatter plot of decoded gene locations. Round cell marker genes are plotted in blue and spindle cell marker genes are plotted in orange. All other genes are plotted in gray. Scale bar corresponds to 40 μm. (b) Inset of scatter plot in (a), location denoted by the black box. The gene identity is encoded by the point color. All other genes are plotted in gray. (c) Joint plot of UMAP of gene expression data measured with seqFISH and histograms of UMAP values in each dimension. (n=2,405 cells) (d) Mean expression of round and spindle marker genes. All cells that do not meet the criteria for round or spindle cells are plotted as "other." Error bars denote the 95% confidence interval of the mean.

by the cell's metabolic state, a mechanism that may underlie the transcriptional differences between round and spindle cells.[199]

The seqFISH measurement also retained the cell-to-cell spatial organization of the sample, which allows it to explore gene expression in communities of cells unlike sequencing methods that require sample homogenization to measure gene expression. We utilized this capability to investigate the degree of spatial variation in gene

expression of the cultured macrophage samples treated with LPS. We constructed a graph consisting of the centroid locations of each cell to investigate correlations in gene expression in neighborhoods of cells in our samples. Assortativity was the metric we used to quantify the degree to which graph nodes of the same label, as opposed to different labels, are connected. (Fig. 4.15a) We performed clustering Leiden clustering on the gene expression values of LPS treated macrophages. (Fig. 4.15b) We then calculated the assortativity of the cluster labels in a graph of the cell locations to assess the similarity between the gene expression of a cell and its neighbors. (Fig. 4.15c) We found that the gene expression clusters in these graphs showed relatively high assortativity. We demonstrated the significance of this observed value by comparing it to a distribution of assortativity values for multiple iterations of the same graph with shuffled cluster labels. (Fig. 4.15d) We also found that gene expression clusters that are more similar, as measured by proximity in the UMAP of gene expression, tended to have lower assortativity than more dissimilar clusters. (Fig. 4.20a-b) For example, dissimilar cluster labels 0 and 1 have a relatively high assortativity value, reflecting spatial separation of cells with dissimilar gene expression patterns. On the other hand, similar cluster labels 1 and 2 have a relatively low assortativity value, reflecting inter-connectivity of neighborhoods of cells with similar gene expression patterns. (Fig. 4.20c-e)

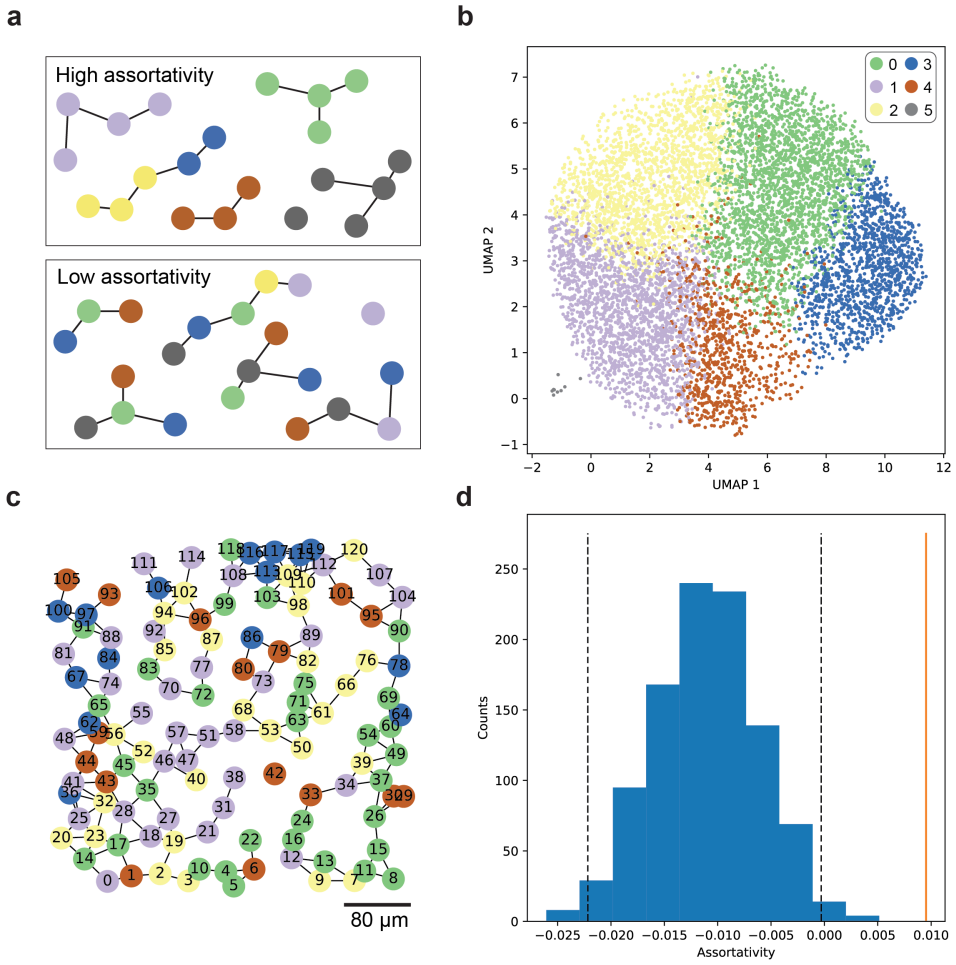


Figure 4.15: **Macrophages dosed with LPS show assortativity of gene expression.** See section 4.5 for caption.

We also found high assortativity of gene expression cluster labels in graphs of untreated macrophages. (Fig. 4.21) This observation suggests that the macrophages create neighborhoods of similar cell states prior to treatment with an inflammatory stimulus. Dynamic measurement of the expression of a limited panel of genes could determine the extent to which these neighborhoods of similar cells exist prior to an inflammatory stimulus, and how they change over the course of treatment.

The observed spatial gene expression pattern aligns with the expected length scale of paracrine signaling, which is estimated to be 40-100 μm.[200,201] This length scale is determined by the rates of cytokine diffusion and production and the concentration of cytokines required to elicit a response in target cells. Paracrine signaling would act as a local average over cell behaviors, allowing cells to make decisions as

individuals while also collecting the information from their neighbors that allow them to act collectively.[201] Macrophages have also been shown to have quorum sensing capabilities, modulating their response to inflammatory stimuli depending on their density in a sample, further solidifying the concept of coordination between macrophage cells to inform their inflammatory response.[202,203]

Development of the integrated live-cell imaging-seqFISH experimental pipeline

To interrogate the relationship between signaling dynamics and spatial gene expression, we developed an integrated measurements workflow that pairs live-cell imaging with seqFISH. Although it is generalizable as a method, we targeted this method to investigating immune signaling in primary human macrophages. In this workflow, the macrophages are seeded onto glass slides that are compatible with live imaging and seqFISH. Time lapse images of fluorescent live-cell reporters and a nuclear stain are acquired. Then, the cells are fixed, stained with seqFISH oligo probes and imaged on the GenePS system. Images between the two microscopes were aligned using nuclear label images.

Live-cell imaging and seqFISH experiments both generate large, complex image-based data sets that require powerful image analysis methods to generate biological insights. Our integrated measurements pipeline is enabled by deep learning-based image analysis methods that increase the feasibility and reproducibility of its data analysis. Tools in the DeepCell software library, such as Caliban[204] and CellSAM[131], expedite the analysis of live-cell imaging data and Polaris[95] enables accurate quantification of gene expression from seqFISH images.

We utilized the single-cell gene expression profiles yielded by the paired seqFISH measurement to explore the degree to which transfection perturbs the immune state of the macrophages. We compared the transcriptomes of untransfected and KTR-transfected cells before and after LPS treatment to assess the effect of mRNA transfection on their basal immune state and their response to stimulus. Furthermore, we compared p65-expressing cells to these two populations to determine if exogenous expression of p65 perturbs immune signaling.

We found that the untreated and LPS-treated samples across transfection conditions group together in a PCA plot of the single-cell gene expression profiles. We did not find a clear distinction between transfected and untransfected cells or between cells over-expressing and cells not over-expressing p65. (Fig. 4.22) From this comparison, we concluded that transfection does not significantly change the gene expression profile of the macrophages. This conclusion aligns with previous work that demonstrated that modified mRNA transfection does not alter the surface marker profiles of macrophages.[178] We also concluded that the background fluorescence intensity from the fluorescent live-cell reporter does not significantly interfere with the ability of seqFISH and Polaris to measure gene expression. However, this control seqFISH measurement is limited by the panel design, which was intended to measure the transcriptional response to LPS treatment. Additional control experiments measuring gene expression with RNA sequencing would more comprehensively quantify the effect of modified mRNA transfection on the cell state of macrophages.

Paired immune signaling dynamics and spatial transcriptomics measurements in primary human macrophages

We utilized our integrated live-cell imaging-seqFISH experimental pipeline to measure p65 signaling dynamics paired with the spatial expression of our panel of inflammatory genes. This measurement builds upon our disjointed exploration of multiple aspects of immune signaling in macrophages. Our live-cell imaging experiments demonstrated capabilities of individual macrophages to perceive and encode LPS concentrations with their dynamic activation of immune signaling pathways. On the other hand, our seqFISH experiments characterized the relationship between morphological and transcriptional subpopulations form neighborhoods of cell. However, a paired measurement will enable exploration of how these subpopulations respond dynamically to inflammatory stimuli.

The seqFISH measurement revealed distinct transcriptional subpopulations of macrophages, defined by their expression of TNF, IL1B, NUPR1, and PIM1. (Fig. 4.16a-b) These transcriptional subpopulations have differing distributions of dynamic features of p65 translocation. The subpopulations defined by TNF and IL1B/GRAMD1A expression have weaker p65 translocation, particularly when measured by nuclear p65 intensity at 40 minutes and the integral of nuclear p65 intensity, in response to LPS treatment, as compared to the subpopulations defined by NUPR1, PIM1, or IL1B/ZC3H12A expression. (Fig. 4.16c) The presence of subpopulations with distinct transcriptional states and dynamic p65 signaling responses to LPS indicates a relationship between the dynamic p65 activation and transcriptional response to LPS treatment.

The transcriptional subpopulations also have varying distributions of morphological states. We observe that the subpopulation defined by TNF expression tends to have a small area and low aspect ratio. On the other hand, the subpopulation defined by IL1B and GRAMD12A expression tends to have a small area and relatively high aspect ratio. These results align with the previously discussed "round" and "spindle" subpopulations observed in unpaired seqFISH data. This measurement also revealed that the subpopulation defined by IL1B and ZC3H12A expression have a distinctly large area and the subpopulation with high NUPR1 expression is distinctly circular. (Fig. 4.17) Empirically, we did not observe individual cells undergoing dramatic morphological changes on the time scale of our dynamics measurements, suggesting that their morphology is an indicator of cell state, which is priming their response to LPS treatment.

An integrated measurement of JNK dynamics and seqFISH revealed similar transcriptionally distinct subpopulations of macrophages, similarly defined by the expression of the same marker genes (TNF, IL1B, NUPR1, and PIM1). (Fig. 4.23a-b) Observation of the same transcriptional subpopulations across both samples demonstrates the robustness of these subpopulations of macrophage responses to LPS treatment. To a lesser extent than with p65 dynamics, these subpopulations have different distributions of JNK signaling dynamics. (Fig. 4.23c) The subpopulation defined by TNF expression showed stronger JNK activation whereas the subpopulation defined by IL1B expression showed weaker activation. Our paired measurement has also revealed that the transcriptional subpopulations also have varying distributions of morphological states. (Fig. 4.23d)

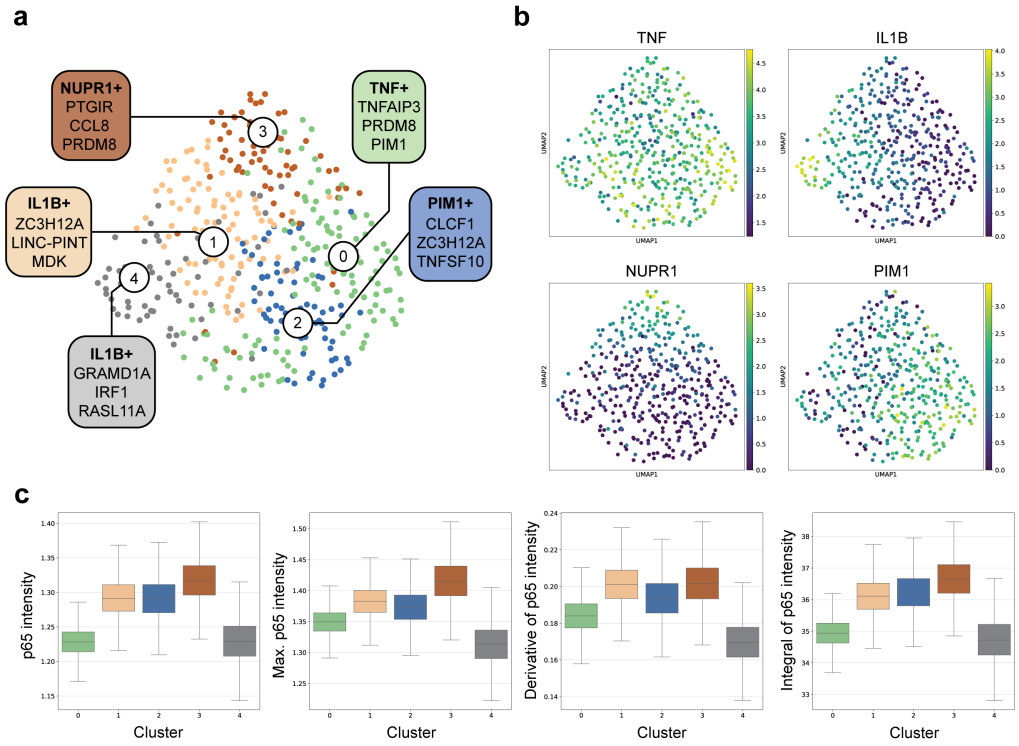


Figure 4.16: **Transcriptional subpopulations of macrophages have distinct dynamic features of p65 translocation.** (a) UMAP of gene expression data measured with seqFISH in response to 100 ng/mL LPS treatment. Leiden clustering was performed with a resolution of 0.7. Cluster stability analysis was performed to set clustering resolution parameter. Cluster marker genes are highlighted. (n=410 cells) (b) UMAP of seqFISH data, colored by cluster marker gene expression (TNF, IL1B, NUPR1, and PIM1). Expression score is the natural logarithm of CPM-normalized counts. (c) Box plots of metrics of the fold-change of p65 translocation dynamics. Boxes denote quartiles of the data and the whiskers denote the 2.5th and 97.5th percentiles of the data.

4.3 Discussion

Our integrated measurements workflow, pairing live-cell imaging with seqFISH offers insights into cellular decision making, at the single-cell level as well as population level. With this workflow, we can interrogate the capabilities of individual cells to perceive and encode environmental stimuli while also asking how they communicate to act as a collective.

The live-cell imaging measurement established the capability of individual cells in our system to distinguish between concentrations of LPS. This measurement establishes cellular perception and decision-making at the single-cell level. The signal that transmits information about environmental stimuli is distributed across multiple

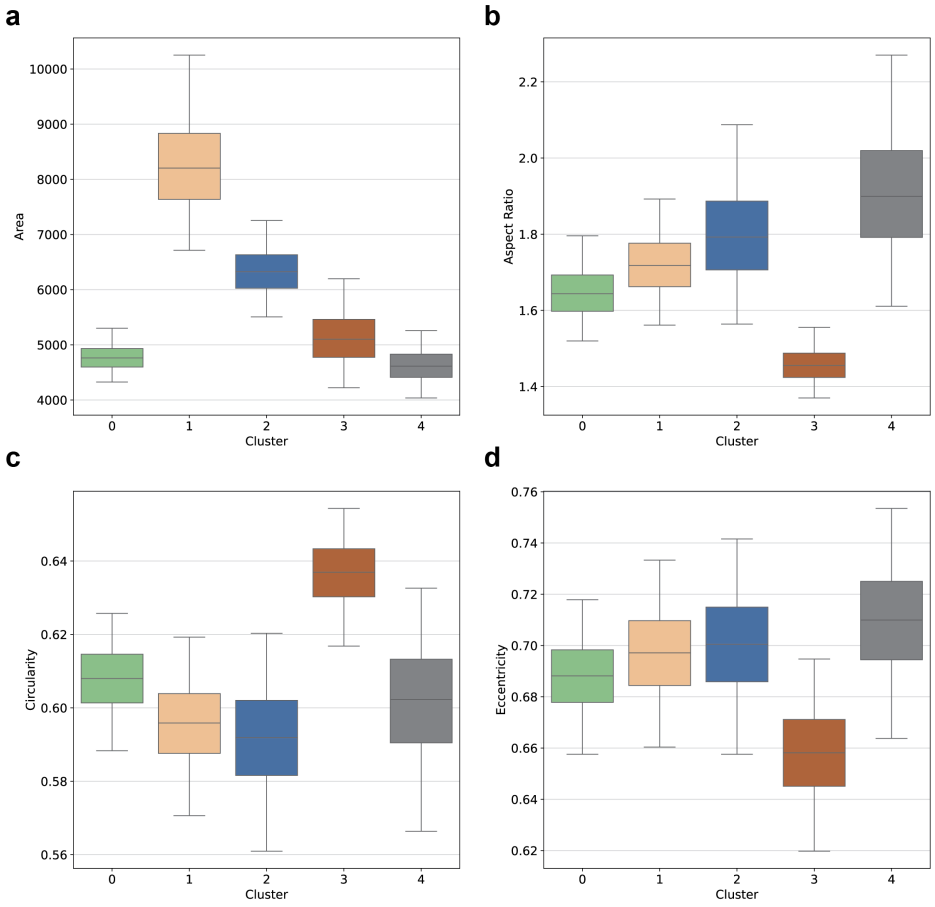


Figure 4.17: **Transcriptional subpopulations of macrophages in paired p65 dynamics-seqFISH measurement have distinct morphological features.** (a-d) Box plots of morphology metrics across gene expression clusters from Fig. 4.16. Boxes denote quartiles of the data and the whiskers denote the 2.5th and 97.5th percentiles of the data.

signaling pathways, as demonstrated by the mutual information observed between NF-κB and MAPK signaling dynamics. Such a mechanism adds redundancy to the signaling network but also allows the cells to perform diverse cellular behaviors in response to activation.

Measuring the spatial gene expression of macrophages with seqFISH revealed a few key insights for understanding single-cell heterogeneity and cell-to-cell communication in macrophages. First, the seqFISH measurement revealed morphotypes of macrophages with distinct gene expression profiles. This measurement also revealed the heterogeneity in cells states, as read out by morphology, present in a population of macrophages. Transcriptional subpopulations of cells were also

linked to differentiable patterns of morphology and signaling dynamics, indicating a relationship between a cell's state and its response to inflammatory stimuli. Such an observation could not be made if gene expression was measured with a sequencing-based method, such as single-cell RNA sequencing, which does not retain the morphological state of a cell. In particular, seqFISH is well-positioned to link morphology to gene expression as an image-based spatial transcriptomics methods, because even bead-based spatial transcriptomics methods lack the spatial resolution to capture subtle morphological details of a sample. Next, the seqFISH measurement revealed transcriptionally similar neighborhoods of cells, indicating cell-to-cell communication that could not have been observed with a method that requires sample homogenization. The characterization of paracrine signaling can be made more comprehensive in future work by adding pairs of cytokines and receptors to the seqFISH panel to rigorously characterize the spatial gene expression patterns of receptor-ligand pairs. Alternatively, immunofluorescence staining for such receptors could reveal the spatial distributions of their abundances with more precision than a spatial transcriptomics measurement.

Because of the previously characterized relationship between morphology and substrate stiffness, an interesting future direction for this work would be additional experiments exploring the interplay between the mechanical forces from substrate stiffness and macrophage cell state. We expect environmental forces to strongly influence the inflammatory state of macrophages based on previous studies.[167] Our integrated measurements workflow would add expand upon previous work by allowing investigation of the effects of substrate stiffness on macrophages' responses to inflammatory stimuli and spatial gene expression patterns before and after treatment.

A key limitation of this measurement is its utilization of p65 overexpression as a method for observing NF-κB dynamics. Overexpression of p65 is known to alter its dynamics, because the feedback mechanisms that regulate its translocation are finely tuned to its expression level.[17,205] Visualization of endogenous p65 translocation would reveal more physiologically relevant dynamics. Genetic engineering in macrophages remains challenging because of their sensitivity to double stranded breaks in their genomic DNA, so fluorescent nanobody labeling may be a more feasible alternate method for visualizing endogenous p65.[206]

In this work, we have demonstrated the viability of our integrated measurements workflow to measure immune signaling dynamics paired to spatial transcriptomics in primary macrophages. The scale of our experiment was enabled by deep learning

methods to expedite image analysis. With these methodological capabilities there are many dimensions across which this measurement can be extended. Importantly, this work is limited to measuring macrophages' response to LPS stimulation when it is well established that they demonstrate range of stimulus-specific responses.[168] Our workflow could iterate upon the current characterization of stimulus-specific responses, because it would simultaneously capture both the dynamic signaling and transcriptional responses of macrophage morphotypes to each inflammatory stimulus. Additionally, we measured gene expression at a single static time point after LPS treatment, but transcriptional dynamics are variable and gene-specific.[16,207,208] Follow-up experiments measuring gene expression at different time-points could capture known early response or late response NF-κB target genes. Finally, we demonstrated that a wide variety of live-cell reporters can be expressed in primary human macrophages with modified mRNA transfection, opening up the scope of feasible live-cell imaging experiments in this system to the range of available live-cell reporters. Therefore, this workflow could be applied to the investigation of a signal transduction in a number different signaling pathways.

4.4 Methods

Cell culture

CD14+ primary peripheral blood mononuclear cells (PBMCs) were received from the Charles River Laboratories. After thawing, the cells were cultured for 24 hours in Roswell Park Memorial Institute (RPMI) 1640 Medium (Gibco) supplemented with 2 mM L-glutamine (Gibco), 100 U/mL penicillin, 100 μg/mL streptomycin (Gibco or Caisson), and 10% fetal bovine serum (Omega Scientific or Thermo Fisher). Cells were incubated at 37°C in a humidified 5% CO_2 atmosphere in a polystyrene 6-well plate (CellTreat #229105) at a density of 1 million cells per well. After 24 hours, the cells were dosed with 50 ng/mL M-CSF (Peprotech 300-25). The cells were maintained at this M-CSF concentration for up to 9 days, replacing the media every 3 days, for differentiation from monocytes to macrophages. After 6 days of M-CSF treatment, the macrophages were dissociated from the plate with Accutase for 30 minutes at 37°C for usage in the paired imaging assay.

Modified RNA synthesis

RNA constructs were synthesized using a linear DNA oligo template that contains the coding sequence for the construct of interest preceded by a T7 promoter. This template was generated by restriction enzyme digestion or PCR. The DNA tem-

plate was purified by gel extraction (NEB T1020S). RNA was synthesized by T7 polymerase in a reaction mixture that includes modified nucleotides, pseudouridine and 5-methylcytidine at 37 °C for 2 hours (APExBIO K1064). After synthesis, the DNA template was digested with DNAse I (Thermo EN0521) and a polyA tail was enzymatically added (APExBIO K1064). The synthesized mRNA was purified with a column-based protocol (NEB T2040S). For long-term storage, the IVT RNA was aliquoted and stored at -80 °C.

Live-cell imaging in 96-well plates

The macrophages were seeded in fibronectin-functionalized (Fisher Scientific #33010018) glass-bottom 96-well plates (Cellvis P96-1.5H-N) at 80-90% confluency. Live-cell reporters were transiently expressed in the macrophages with lipofectamine designed for transfection of mRNA, Lipofectamine MessengerMax (Thermo LMRNA001). When transfecting a single mRNA construct, the cells were transfected with 100 ng of mRNA. When transfecting two mRNA constructs, the cells were transfected with 150 ng of mRNA.

Live-cell imaging on slides

The macrophages were seeded on a fibronectin-functionalized glass slide (Corning #2980-246) at 80%-90% confluence contained by a rubber gasket (Grace Bio-Labs #JTR8R-2.5). When transfecting a single mRNA construct, the cells were transfected with 150 ng of mRNA per well. When transfecting two mRNA constructs, the cells were transfected with 150 ng of mRNA per well.

Live-cell imaging conditions

The seqFISH samples were imaged with a Nikon Ti2 fluorescence microscope controlled by Nikon Elements. Images were acquired with a Nikon SOLA SE II light source, a 100X oil objective, and a Photometrics Prime 95B CMOS camera. Cells were imaged in a humidified environment at 37 °C and 5% CO_2.

Cells were treated with Hoechst 33342 (Thermo-Fisher H3570) to stain the nuclear volume at least 30 minutes prior to imaging. Cells were imaged for 15 minutes prior to LPS treatment to capture their basal signaling state. Cells were imaged at a frame rate of 5 minutes for the duration of the live-cell imaging assay.

Preparation of seqFISH gene panel

The gene panel for seqFISH was designed using a previously published RNA sequencing dataset, demonstrating the transcriptomic changes that occurred in THP-1 cells upon LPS stimulation.[193] The panel consists of 200 of the most strongly differentially expressed genes between untreated and LPS-treated cells. The gene panel was optimized for visualization with seqFISH, taking considerations like probe binding properties, number of available probes for a genes, and expression levels into account. The codebook defining the relationship between gene identities and temporal barcodes was designed with guidance from previously published design principles to minimize the number of required staining rounds and maximize the Hamming distance between barcodes.[32]

The panel consists of the following genes: CCL4, HCAR3, ENSG00000216775, ENSG00000279320, NT5C3A, HCAR2, GBP1P1, CCL20, USP30-AS1, CXCL1, HES4, WFDC21P, CXCL3, BATF3, CCL8, CCDC194, IFITM1, HESX1, MAP1LC3A, IL27, TNFAIP6, EBI3, USP18, HMGN2P46, CCL5, CXCL10, CXCL11, TNFSF9, IL6, PLAC8, IL23A, ASNS, ENSG00000284633, TRGC1, SOCS1, MDK, GBP1, ENSG00000279118, ENSG00000277895, NOCT, PMAIP1, TNFSF10, TFPI2, AIM2, CKB, RRAS2, CD40, PKIG, PTX3, IL1B, ENSG00000270562, IL1A, TNIP3, NEURL3, IDO1, RND1, ISG20, IL15RA, SERPINB2, ETV7, MUC1, CLCF1, RUSC1-AS1, ZBP1, LAG3, TNF, PDGFRL, GBP4, IL15, HSH2D, ZBTB10, NME8, ENSG00000260244, HS3ST3B1, ACOD1, CDC42EP2, RASL11A, GCNT4, AZIN2, RNF144A, IFI44L, CASZ1, PRDM8, STK26, LAMP3, EEF2, NIPAL4, ELOVL7, OXTR, NUPR1, CFB, CATSPERB, GBP5, HELZ2, HERC5, PCDH9, GRAMD1A, ANKRD33B, TMPRSS13, SLAMF1, IRF1, MCOLN2, PRSS23, PELI1, DCLK3, SGPP2, BIRC3, SLC1A2, NECTIN3, PARP14, MDGA1, FZD4, TMCC2, SOCS2, C1S, PIM1, NIBAN1, PPM1K, TNFAIP3, IRF4, ADORA2A, CACHD1, RAPGEF2, OASL, IFIH1, ITGA1, ADM2, CD38, SOCS3, IL2RA, TTC39A, IRF1-AS1, APOL3, ADGRE1, CAMK1G, GUCY1A1, KCNA3, CCR7, C1R, ADA, ZC3H12A, IL7R, CYP7B1, LAD1, PLEKHA4, ADPGK-AS1, HAPLN3, NFIX, RNF19B, FAM186B, IFIT2, RSAD2, BRIP1, MN1, KLF5, MX1, CXCL9, PLCB4, TRAF1, TANC1, CD274, CD80, PTGIR, SYNPO2, CGN, NTN1, ADAMTS4, ABTB2, SYNPO, ABCG1, TNFAIP2, CMPK2, ITGB8, SLAMF7, GCH1, MVP-DT, IFIT3, TNFRSF9, SLC2A6, NXN, PDE4B, STAT4, PLEKHG1, PTGS2, DHX58, and ADAM19.

seqFISH sample preparation

Immediately following the live-cell imaging assay, cells were rinsed with warm 1X PBS, fixed with fresh 4% formaldehyde (Thermo #28908) in 1X PBS for 10 minutes at room temperature, and then permeabilized with 70% ethanol overnight at -20°C. The primary probe library (Spatial Genomics) was added to the sample in a flow chamber provided by Spatial Genomics and incubated overnight at 37°C. The sample was washed several times with primary wash buffer (Spatial Genomics). The nuclei of the sample were stained with staining solution (Spatial Genomics). The macrophage sample was imaged with the Spatial Genomics Gene Positioning System (GenePS). Tiled imaging and secondary probe staining were performed programmatically by this instrument.

Image analysis

Cell segmentation and tracking: The nuclei were segmented with Deepcell's deep learning model for nuclear segmentation. Nuclear segmentation was performed on images of Hoechst nuclear staining. The whole-cell areas were segmented with CellSAM.[131] Whole-cell segmentation was performed on images of live-cell reporters or on maximum intensity projections of images of multiple rounds of seqFISH staining. The nuclei were tracked through time in live-cell image sets with Caliban.[204]

Reporter activity quantification: The activity of the p65-mOrange reporter was quantified as the mean nuclear intensity. The nuclear intensity was normalized with one of two values: (1) the mean cytoplasmic intensity prior to LPS-treatment as a measure of expression level, or (2) the mean nuclear intensity prior to LPS-treatment to calculate the fold-change of nuclear p65 intensity. The activity of the KTR reporters was quantified as the mean cytoplasmic intensity of the reporter divided by the mean nuclear intensity throughout the time-lapse. For both reporters, the mean nuclear intensity was the mean value of the masked pixels for the segmented nuclei. The cytoplasmic intensity was the mean value of a masked region created by dilating the nuclear masks by 40 pixels.

Image alignment: The live-cell imaging and seqFISH measurements were collected at different magnifications of different microscopes, so linking cells between these two imaging modalities requires image alignment. Images of Hoechst nuclear labeling were acquired with the live-cell imaging configuration and the GenePS configuration. A Gaussian blur was applied to these images and the location of the

live-cell imaging fields of view was identified in the larger GenePS field of view with template matching implemented in CV2. The match was required to exceed a user-defined threshold score for downstream analysis.

seqFISH quantification: Gene expression was quantified from the seqFISH images using Polaris.[95] Gene assignment is attempted for all pixels with a spot probability greater than 0.01. The spot probability values were fit with a relaxed Bernoulli distribution during gene assignment. The gene assignments with a probability value less than 0.95 were removed from downstream analysis.

Morphology quantification: In order to measure morphology metrics of cell areas, the Python package function `skimage.measure.regionprops` was used to measure the area, perimeter, major axis length, minor axis length, eccentricity, and solidity of cells. The eccentricity is the ratio of the ratio of the distance between the focal points of the region over the major axis length. The solidity is the ratio of area of the region to area of the convex hull image. Following previously published methods for morphology quantification, the circularity was quantified as $4\pi \cdot A/P^2$, where A is the area of the region and P is the perimeter, and the aspect ratio was quantified as the ratio of the major and minor axis lengths.[174]

Morphological sub-populations: The cells were divided into defined morphological sub-populations depending on the values of two metrics. "Round" cells had a circularity value greater than 0.65. "Spindle" cells had an aspect ratio of greater than 3.5. All cells that did not meet either of these criteria were separated into an "other" category.

Timeseries analysis

Metrics for signaling dynamics: The dynamic activity of p65 and the KTR reporters was summarized with a few summary metrics. When the activity of the reporters is summarized with a single value, that intensity value was taken at 40 minutes from the beginning of the time series. This time point was selected because it has the maximum amount of mutual information with the concentration of the LPS treatment. It was empirically determined that cell death in response to LPS treatment is relatively rare in the first 150 minutes of the time series. Cell death leads to aberrant signal in the time series of quantified reporter activity. For this reason, the maximum intensity, maximum derivative, and integral of activity were calculated across the first 150 minutes of each time series.

Mutual information: The mutual information between two time series or between a time series and a particular value with `sklearn.feature_selection.mutual_info_regression`. This method estimates entropy in a non-parametric fashion with a k-nearest neighbors method, based on previous work.[209,210]

Single-cell gene expression analysis

Clustering for single-cell gene expression data: Clustering for single-cell gene expression data was performed with the Leiden algorithm implemented in `scanpy`. Cluster stability analysis was performed for resolution values from 0.1 to 1.0 with an increment of 0.1. The stability of the clusters was quantified with the adjusted Rand index score. The resolution value with the highest stability with neighboring resolution values was selected.

Marker gene analysis: Marker genes for gene expression clusters were determined by ranking the genes that characterize each cluster, as implemented in `scanpy`. The genes were ranked by their t-test score and the resulting p-value was corrected with the Benjamini-Hochberg procedure.

Assortativity analysis: The assortativity of the single-cell gene expression values was determined by creating a graph of neighboring cells. The nodes of the graph were placed at the centroids of each whole-cell area. Edges were defined between nodes closer than 46.8 μm. The threshold for defining neighbors was determined based on the typical cell area. Assortativity was calculated using a function implemented in `networkx` with the Leiden gene expression clusters as the attribute.

4.5 Figure captions

Figure 4.1: Overview of integrated measurements method in primary human macrophages (a) Schematic diagram of integrated measurements workflow that pairs live-cell imaging with seqFISH. (b) Montage of p65-mOrange translocation in response to LPS treatment in an example primary human macrophage cell. Nuclear segmentation indicated in red. Scale bar corresponds to 20 μm. (c) seqFISH staining and Polaris analysis of spatial gene expression for the same example primary macrophage cell as in (b). Gene identities are indicated with spot colors. (d) Single-cell data captured by the integrated measurements workflow for the example primary macrophage cell. The dynamic p65 translocation is plotted at the ratio of the nuclear and cytoplasmic intensity values over time. The time point of LPS treatment is indicated with the dashed red line.

Figure 4.5: Expression of a panel of live-cell reporters in primary macrophages with modified RNA transfection A panel of fluorescent fusion proteins, KTRs, and ExRai reporters were expressed in primary macrophages. Montage figures demonstrate the appearance of the reporters over time. Time increments in montage figures are 40 minutes (TFEB-mClover and ExRai AMPKAR) or 50 minutes (p65-mOrange, JNK KTR-mClover, and p38 KTR-mClover). All scale bars correspond to 40μm. Time traces quantify the activity of the signaling protein read out by each live-cell reporter.

Figure 4.9: Fold-change of nuclear p65 intensity shows greater resolution of LPS stimulation conditions than absolute intensities (a) Time series data of nuclear p65 intensity for untreated and LPS-treated cells, presented as absolute intensity (a.u.) and fold-change intensity. Time series have been smoothed with a boxcar average over a window of four frames. Fold-change is calculated using the intensity of the fourth frame. (b-e) Distributions of (b) nuclear p65 intensities after 70 minutes of imaging, (c) maximum p65 intensity, (d) maximum rate of change of p65 intensity, and (e) integral of p65 intensity over 4 hours of imaging, presented as absolute intensity and fold-change intensity for untreated and LPS-treated cells. Distribution values have been normalized to the median of the intensity values across both treatment conditions.

Figure 4.15: Macrophages dosed with LPS show assortativity of gene expression. (a) Schematic illustration of the assortativity graph metric. High assortativity means similar cells are connected, and low assortativity means dissimilar cells are connected. (b) UMAP of single-cell gene expression of primary macrophages dosed with 100 ng/mL LPS. (c) Example graph of cell centroid locations, labeled by gene expression cluster from (b). Graph defines the neighbors of each cell. Cells closer than 46.8 μm are connected. (d) Macrophages show positive assortativity, calculated with respect to gene expression clusters (0.0095), shown in orange. The distribution of assortativity values calculated for randomly shuffled cluster labels in 1000 experiments, shown in blue, has a mean of -0.011 and standard deviation of 0.0049. The dashed lines indicate the 1st and 99th percentiles of this distribution. The experimental and shuffled assortativity values are an average across graphs created for 72 fields of view, containing a total of 8,279 cells.

4.6 Supplementary figures

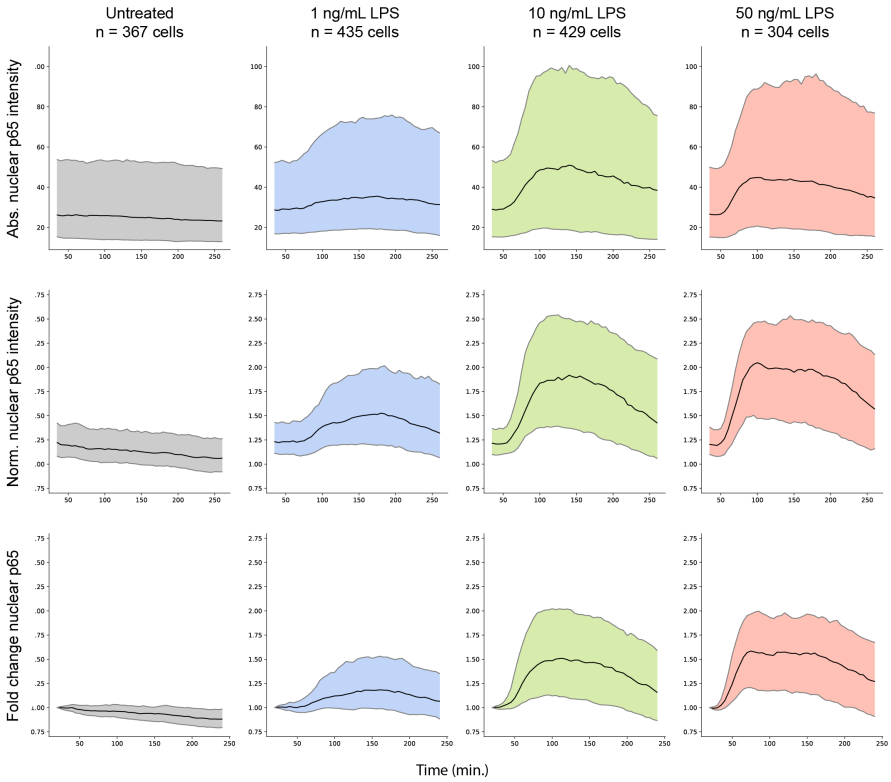


Figure 4.18: **Comparison of normalization methods for nuclear p65 dynamics in response to LPS treatment.** Primary human macrophages were treated with LPS of various concentrations at 15 minutes. (Upper) Absolute nuclear p65 intensity. (Middle) Normalized nuclear p65 intensity, relative to cytoplasmic p65 intensity, as a measure of expression level before LPS treatment. (Lower) Fold change of nuclear p65 intensity, relative to the mean nuclear p65 intensity before LPS treatment. Line plotted in black represents the mean p65 intensity for each condition. The shaded region represents the 25th-75th percentile of p65 intensities for each condition.

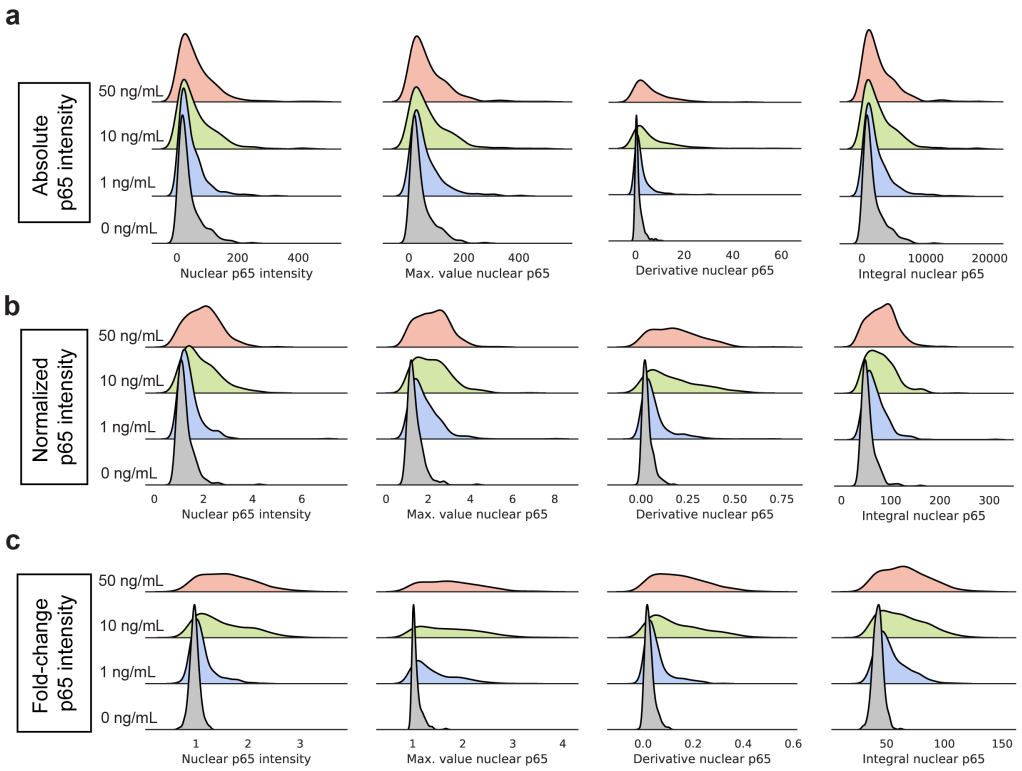


Figure 4.19: **Normalization method alters resolution of distributions of p65 dynamics metrics** (a-c) Ridge-plots for distributions of p65 dynamics metrics for absolute p65 intensity, normalized p65 intensity, and fold-change p65 intensity, respectively. Normalized p65 intensity is relative to cytoplasmic p65 intensity, as a measure of expression level. Fold-change p65 intensity is relative to the mean nuclear p65 intensity before LPS treatment.

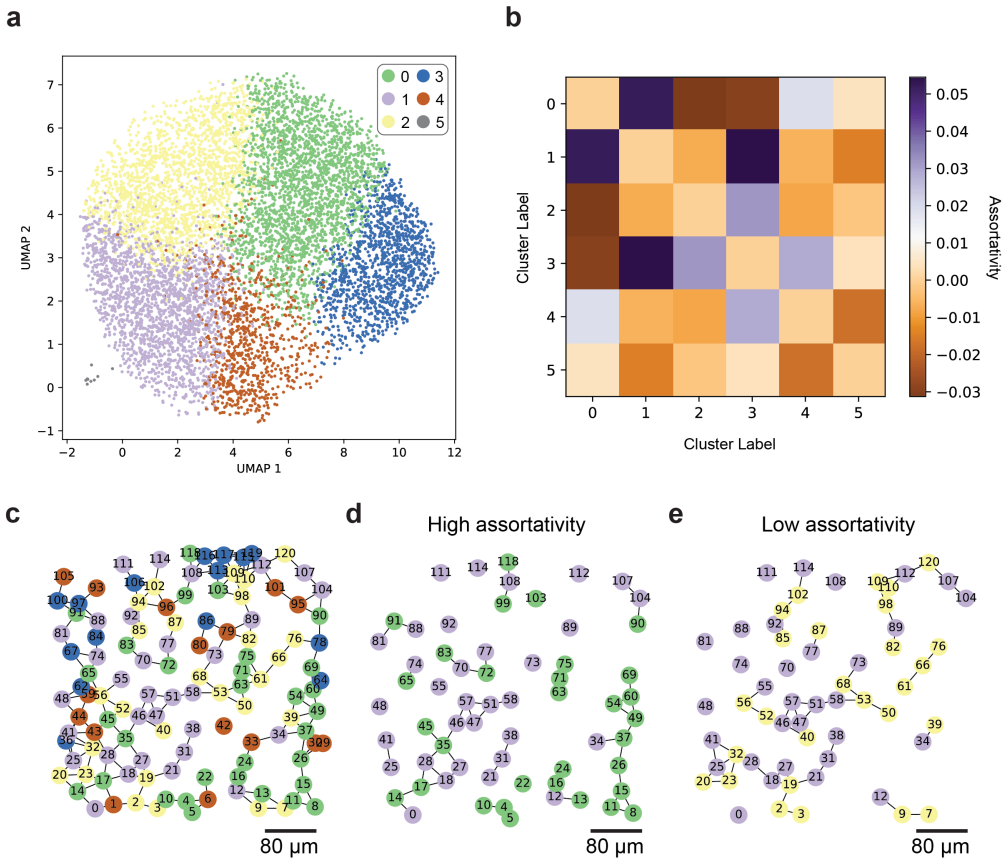


Figure 4.20: **Neighborhoods of macrophages with similar gene expression patterns.** (a) UMAP of single-cell gene expression of primary macrophages dosed with LPS. (b) Matrix of assortativity values of pairs of gene expression clusters. (c) Example graph of cell centroid locations, labeled by gene expression cluster from (a). Graph defines the neighbors of each cell. Cells closer than 46.8 μm are connected. (d) Sub-graph of graph in (c) containing two labels of interest, 0 and 1. The sub-graph shows a relatively high assortativity value of 0.023. (e) Sub-graph of graph in (c) containing two labels of interest, 1 and 2. The sub-graph shows a relatively low assortativity value of -0.003.

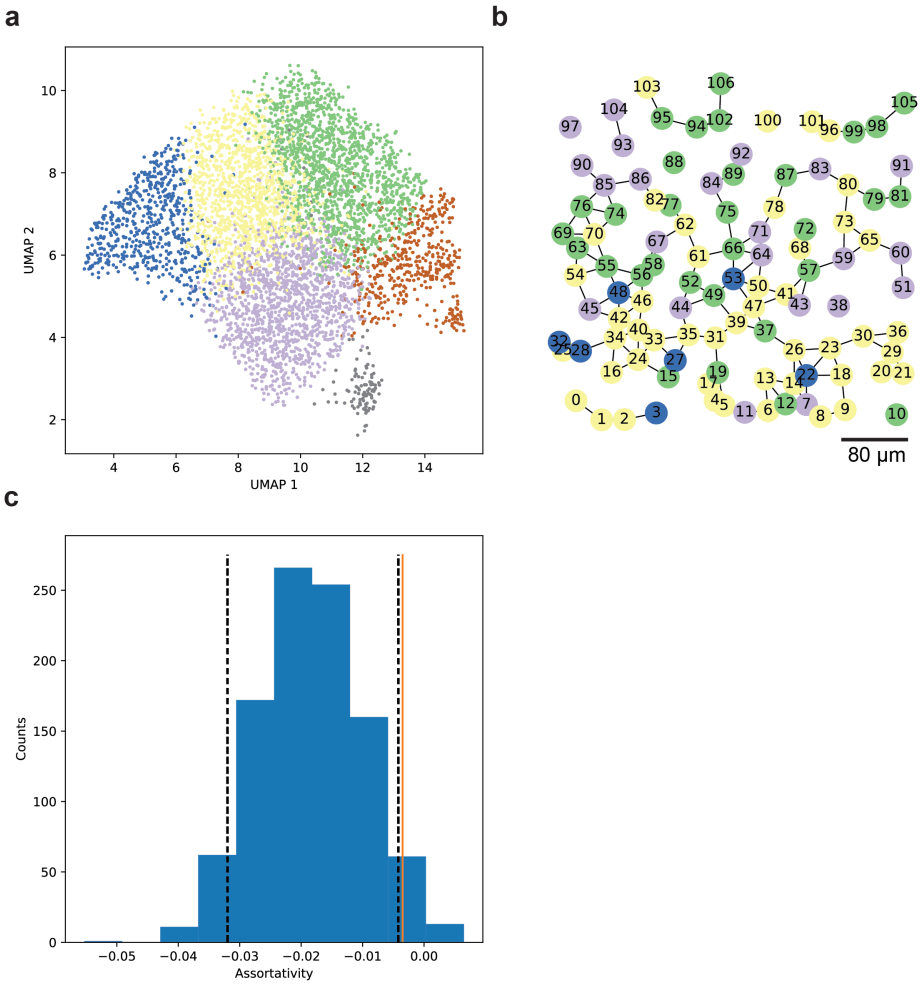


Figure 4.21: **Unreated macrophages dosed with show assortativity of gene expression.** (a) UMAP of single-cell gene expression of untreated primary macrophages. (b) Example graph of cell centroid locations, labeled by gene expression cluster from (a). Graph defines the neighbors of each cell. Cells closer than 46.8 μm are connected. (c) Macrophages show positive assortativity, calculated with respect to gene expression clusters (-0.0035), shown in orange. The distribution of assortativity values calculated for randomly shuffled cluster labels, shown in blue, has a mean of -0.018 and standard deviation of 0.0085. The dashed lines indicate the 5th and 95th percentiles of this distribution. The experimental and shuffled assortativity values are an average across graphs created for 72 fields of view, containing a total of 5,313 cells.

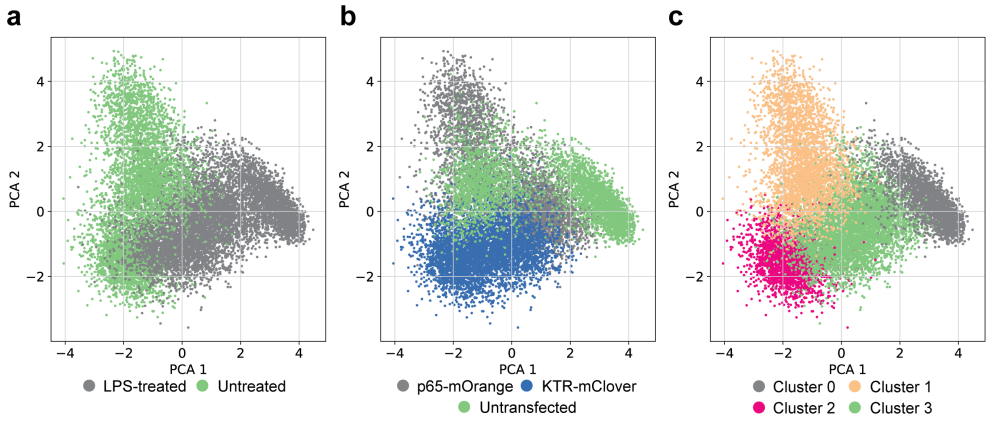


Figure 4.22: **PCA plot gene expression values of LPS-treated, transfected macrophages** (a-c) PCA of single-cell expression values of primary human macrophages measured with seqFISH for highly variable genes colored by LPS treatment condition, transfection condition, or Leiden clustering, respectively. Cluster stability analysis was performed to set clustering hyperparameter values.

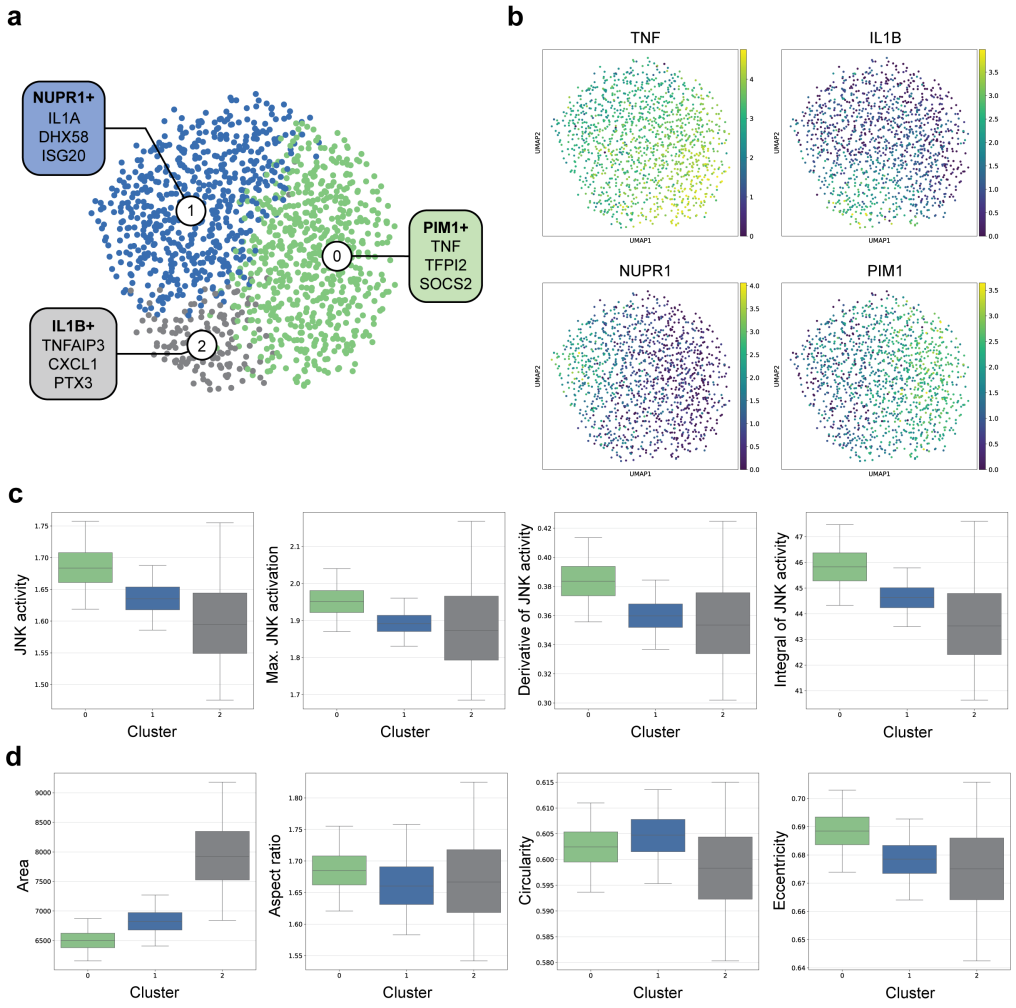


Figure 4.23: **Transcriptional sub-populations of macrophages have distinct dynamic features of JNK KTR activity.** (a) UMAP of gene expression data measured with seqFISH in response to 100 ng/mL LPS treatment. Leiden clustering was performed with a resolution of 0.5. Cluster stability analysis was performed to set clustering resolution parameter. Cluster marker genes are highlighted. (n=1,358 cells) (b) UMAP of seqFISH data, colored by cluster marker gene expression (TNF, IL1B, NUPR1, and PIM1). Expression score is the natural logarithm of CPM-normalized counts. (c) Box plots of metrics of the fold-change of JNK-KTR activity dynamics. Boxes denote quartiles of the data and the whiskers denote the 2.5th and 97.5th percentiles of the data. (d) Box plots of morphology metrics across the gene expression clusters. Boxes denote quartiles of the data and the whiskers denote the 2.5th and 97.5th percentiles of the data.

Chapter 5

CONCLUSIONS AND FUTURE DIRECTIONS

5.1 Live-cell reporters: quantifying signaling dynamics at the single-cell level

Fluorescent live-cell reporters are an important tool to capture cell-to-cell and temporal heterogeneity in signaling pathway activation with a microscope. The variety of available reporter architectures and their typically modular design allows them to be applied to explore the dynamic activation of a wide variety of signaling proteins. In this work, we demonstrated the design and initial validation of two live-cell reporters, AMPK SPARK and ExRai IKKAR/TBKAR.

AMPK SPARK is a novel live-cell reporter for AMPK activity with a few key advantages over existing reporters for AMPK. Unlike the FRET-based reporter, AMPKAR, or the cpGFP-based reporter, ExRai AMPKAR, it utilizes a single fluorescence channel, decreasing its spectral bandwidth and increasing its capacity to be multiplexed with other fluorescent live-cell reporters. However, the phase separation kinetics of its read out dynamics warrant more detailed characterization to ensure that its clustering state accurately reflects a cell's real-time of AMPK activity. The original creators of the SPARK reporter architecture have improved that responsiveness of SPARK reporters by inserting a fluorescent protein molecule in both peptide constructs of the reporter to drive stronger multivalent interactions between the reporter oligomers.[78] Further work on AMPK SPARK should involve the implementation of this incremental improvement in the SPARK design. We expect that this change will increase the dynamic range and responsiveness of AMPK SPARK.

Its multiplexing capability allows AMPK SPARK to be utilized in experiments that address immuno-metabolic signaling questions relevant to the scope of this work. In response to inflammatory stimuli, macrophages produce large amounts of cytokines and undergo highly dynamic cytoskeleton remodeling.[211–213] Therefore, the inflammatory response of macrophages is energetically expensive, so the metabolic state of macrophages is largely shaped by environmental stimuli. AMPK is a key metabolic signaling protein that regulates the balance of anabolic and catabolic cellular processes.[79,81] For this reason, we expect it to be critical to the regulation of energy expenditure during a macrophage's inflammatory response. AMPK SPARK

can be utilized to measure AMPK activity at the single-cell level simultaneously with live-cell reporters for the immune signaling. Such a measurement will illustrate the relationship between AMPK and immune signaling in individual cells and may reveal mechanisms of metabolic regulation in response to inflammatory stimuli.

Our work designing ExRai IKKAR and ExRai TBKAR with KINBERT was a valuable proof-of-concept for designing live-cell reporters with large language models. ExRai IKKAR and ExRai TBKAR demonstrated reactivity to kinase activity of IKKβ and TBK1, respectively, in response to previously characterized upstream stimuli. Although their dynamic range was limited, their responsiveness demonstrated the viability of using KINBERT to generate peptide substrate hits for kinases of interest. We propose to continue to use KINBERT to design live-cell reporters for kinases of interest. Often live-cell reporters only exist for a single kinase in a signaling pathway, because of the significant time and resource investment required to develop a live-cell reporter. We hope that KINBERT can alleviate the degree to which substrate identification can limit live-cell reporter development. With this greater throughput, we propose to characterize information transmission in cascades of signaling kinases. For example, this method could be applied to investigate the molecular mechanisms by which cells encode the identity and strength of immune stimuli in the dynamic patterns of kinase activity in the NF-κB signaling pathway.

5.2 Polaris: applying deep-learning to expedite spatial transcriptomics analysis

Spatial transcriptomics methods excel at capturing the spatial organization of patterns of gene expression. Image-based spatial transcriptomics methods generate large, complex image sets of multiple staining rounds, which typically require manual parameter tuning to extract gene expression values. With Polaris, we utilized deep learning to create an analysis pipeline for spatial transcriptomics images that mitigates the need for manual parameter tuning that requires time and expertise.

Polaris' spot detection model achieves accurate spot detection that generalizes across sample types and spatial transcriptomics assays. This capability makes Polaris an assay-agnostic, turnkey solution for extracting single-cell gene expression profiles from spatial transcriptomics images. Furthermore, Polaris does not require manual parameter tuning for accurate spot detection, greatly increasing the efficiency and reproducibility of spatial transcriptomics analysis.

Variation in fluorescence background intensities from factors like non-specific stain-

ing and autofluorescence continues to limit Polaris' performance in images from challenging tissue samples. We have observed the greatest improvements in performance from continued investment in the quantity and quality of the training data available for Polaris to learn from. Therefore, we expect that adding training data from a wider variety of tissue sample types would increase Polaris' accuracy in more challenging sample types. On the other hand, improving the quality of Polaris' consensus annotations with labels from more classical spot detection methods has previously improved its performance. Such an improvement in Polaris' training data quality demonstrates the strength of our weakly supervised training data annotation method to iteratively improve upon previous annotations.

To offer a diverging goal for this method, Polaris could be applied to spot detection for a wider variety of single-molecule experiments if its training data set included more mobile fluorescent spots. The shape of mobile fluorescent molecules departs from the circular, Gaussian spots commonly seen in spatial transcriptomics images. Because of the expected heterogeneity in the shape of mobile spots, Polaris' deep learning model could offer more robust spot detection capabilities than the more rigid mathematical operators underlying classical spot detection methods.

Polaris' gene decoding method was optimized to capitalize on the spot probability output of its spot detection model. These paired methods yielded gene expression values that strongly correlate with gene expression values measured with RNA sequencing, with manual parameter tuning. Furthermore, unlike other assay-specific analysis pipelines, Polaris' gene decoder does not make limiting assumptions about the structure or pattern of temporal barcoding to encode gene identities with imaging rounds, making it assay-agnostic and broadly applicable to the analysis of a wide variety of spatial transcriptomics images.

A known shortcoming of Polaris' gene decoder its limited ability to assign gene identities to pixels containing more than one mRNA molecule. SeqFISH+ images typically have overlapping fluorescent spots from more than one labeled mRNA molecule because they measure the expression of a large panel of up to thousands of genes in an individual cell.[32] Polaris has a mixed barcode rescue method that addresses this situation. However, the robustness of this rescue method warrants further investigation and development given the prevalence of overlapping spots in seqFISH+ images. Such an improvement to Polaris would improve the breadth of samples it could be applied to.

With Polaris, we have made an investment in model development, training data

annotation, and open-source distribution for the analysis of spatial transcriptomics images. Polaris expedites the analysis of seqFISH images in this work, but its impact reaches far beyond the scope of this work. It can also be applied to quantify single-cell gene expression from MERFISH and in situ sequencing experiments in tissue and culture-cell samples.

5.3 Integrated measurements: spatiotemporal characterization of immune signaling in primary human macrophages

In this work, we have demonstrated the value of imaging assays that capture temporal and spatial heterogeneity in living systems. These imaging methods are particularly powerful when paired, because they can reveal the morphology or spatial arrangements of sub-populations of cells with distinct signaling states. Furthermore, we have demonstrated the value of using deep learning methods to enable the scale of imaging assays.

Our expression of fluorescent live-cell reporters in primary human macrophages with modified RNA transfection demonstrated the feasibility of a few different measurements. First, it showed that a sufficient expression level of fluorescent live-cell reporters can be achieved in classically hard-to-transfect cell types with minimal impact on viability. Second, it increased the scope of measurements of signaling dynamics that are possible in primary human macrophages. Previously, these measurements were limited to immortalized human models or murine cells. Primary human macrophages are more physiologically relevant to *in vivo* macrophages, opening the possibility to collect insights about the signaling networks of these cells *in vitro* that are relevant to *in vivo* behaviors.

We leveraged seqFISH's capabilities to measure gene expression with high spatial resolution to reveal relationships between the morphology and spatial organization of primary macrophages to their gene expression patterns. This measurement can be expanded upon with panel design updated with the information gained in this work. Our seqFISH data revealed neighborhoods of cells primarily expressing TNF or IL-1B, indicating intracellular communication via cytokine expression. Promising future directions for this measurement include incorporation of pairs of cytokines and their associated receptors in the seqFISH panel to capture variability in amenability to activation by each cytokine. Such a measurement also has the potential to reveal sub-cellular localization patterns of mRNA transcripts for receptors, which may be trafficked to sub-cellular regions prior to expression. These proposed

measurements would reveal the mechanisms by which cells form patchwork neighborhoods in which cells have similar gene expression patterns as their neighbors while maintaining heterogeneity at the population level.

This work revealed a relationship between cell morphology and inflammatory signaling. This observation indicates a relationship between immune signaling and mechanotransduction and cytoskeletal regulation. YAP and TAZ are mechanosensitive transcriptional co-factors that have been linked to inflammatory signaling.[214,215] An interesting future direction for this work would be simultaneous measurement of p65 and YAP translocation. Such an experiment would more explicitly characterize the relationship between morphological regulation and immune signaling. Additionally, Piezo1 is a mechanosensitive ion channel linked to inflammatory signaling in macrophages. It plays an unknown role in regulating intracellular Ca^{2+} concentrations, so a genetically encoded calcium indicator could be paired with expression of a fluorescent p65 fusion protein in our integrate measurements workflow to observe the response to macrophages to Yoda1, Piezo1-specific agonist.[216,217]

Our integrated measurements workflow is broadly applicable to the investigation of spatiotemporal regulations of signaling pathway activation. It is particularly useful for studying pathways known to be regulated by paracrine signaling. For example, the workflow could be applied to study spatiotemporal regulation of ERK signaling. ERK is known to activate in waves across communities of cells, during processes like embryo development or wound healing.[201,218–220] Our integrated measurements workflow would link dynamic ERK activation with the activity of other signaling pathways and target gene expression at the single-cell level. Such a measurement would capture cell-to-cell heterogeneity that would indicate computation at the single-cell level within a population of cells communicating to achieve a collective goal.

BIBLIOGRAPHY

1. Billman, G. E. Homeostasis: The underappreciated and far too often ignored central organizing principle of physiology. *Frontiers in Physiology* **11.** ISSN: 1664-042X. doi:10.3389/fphys.2020.00200. http://dx.doi.org/10.3389/fphys.2020.00200 (Mar. 2020).

2. Sung, M.-H. *et al.* Switching of the relative dominance between feedback mechanisms in lipopolysaccharide-induced NF-κB signaling. *Science Signaling* **7.** ISSN: 1937-9145. doi:10.1126/scisignal.2004764. http://dx.doi.org/10.1126/scisignal.2004764 (Jan. 2014).

3. Yin, T. *et al.* Macrophage plasticity and function in cancer and pregnancy. *Frontiers in Immunology* **14.** ISSN: 1664-3224. doi:10.3389/fimmu.2023.1333549. http://dx.doi.org/10.3389/fimmu.2023.1333549 (Jan. 2024).

4. Locati, M., Curtale, G. & Mantovani, A. Diversity, mechanisms, and significance of macrophage plasticity. *Annual Review of Pathology: Mechanisms of Disease* **15,** 123–147. ISSN: 1553-4014. doi:10.1146/annurev-pathmechdis-012418-012718. http://dx.doi.org/10.1146/annurev-pathmechdis-012418-012718 (Jan. 2020).

5. Ricketts, T. D., Prieto-Dominguez, N., Gowda, P. S. & Ubil, E. Mechanisms of macrophage plasticity in the tumor environment: Manipulating activation state to improve outcomes. *Frontiers in Immunology* **12.** ISSN: 1664-3224. doi:10.3389/fimmu.2021.642285. http://dx.doi.org/10.3389/fimmu.2021.642285 (May 2021).

6. Valls, P. O. & Esposito, A. Signalling dynamics, cell decisions, and homeostatic control in health and disease. *Current Opinion in Cell Biology* **75,** 102066. ISSN: 0955-0674. doi:10.1016/j.ceb.2022.01.011. http://dx.doi.org/10.1016/j.ceb.2022.01.011 (Apr. 2022).

7. Purvis, J. E. & Lahav, G. Encoding and decoding cellular information through signaling dynamics. *Cell* **152,** 945–956. ISSN: 0092-8674. doi:10.1016/j.cell.2013.02.005. http://dx.doi.org/10.1016/j.cell.2013.02.005 (Feb. 2013).

8. Taniguchi, K. & Karin, M. NF-κB, inflammation, immunity and cancer: Coming of age. *Nature Reviews Immunology* **18,** 309–324. ISSN: 1474-1741. doi:10.1038/nri.2017.142. http://dx.doi.org/10.1038/nri.2017.142 (Jan. 2018).

9. Liu, T., Zhang, L., Joo, D. & Sun, S.-C. NF-κB signaling in inflammation. *Signal Transduction and Targeted Therapy* **2.** ISSN: 2059-3635. doi:10.1038/sigtrans.2017.23. http://dx.doi.org/10.1038/sigtrans.2017.23 (July 2017).

10. Zhang, Q. *et al.* NF-κB dynamics discriminate between TNF doses in single cells. *Cell Systems* **5,** 638–645.e5. ISSN: 2405-4712. doi:10.1016/j.cels.2017.10.011. http://dx.doi.org/10.1016/j.cels.2017.10.011 (Dec. 2017).

11. Cheng, Q. J. *et al.* NF-κB dynamics determine the stimulus specificity of epigenomic reprogramming in macrophages. *Science* **372,** 1349–1353. ISSN: 1095-9203. doi:10.1126/science.abc0269. http://dx.doi.org/10.1126/science.abc0269 (June 2021).

12. Adelaja, A. *et al.* Six distinct NFκB signaling codons convey discrete information to distinguish stimuli and enable appropriate macrophage responses. *Immunity* **54,** 916–930.e7. ISSN: 1074-7613. doi:10.1016/j.immuni.2021.04.011. http://dx.doi.org/10.1016/j.immuni.2021.04.011 (May 2021).

13. Son, M. *et al.* Spatiotemporal NF-κB dynamics encodes the position, amplitude, and duration of local immune inputs. *Science Advances* **8.** ISSN: 2375-2548. doi:10.1126/sciadv.abn6240. http://dx.doi.org/10.1126/sciadv.abn6240 (Sept. 2022).

14. Son, M., Wang, A. G., Keisham, B. & Tay, S. Processing stimulus dynamics by the NF-κB network in single cells. *Experimental and Molecular Medicine* **55,** 2531–2540. ISSN: 2092-6413. doi:10.1038/s12276-023-01133-7. http://dx.doi.org/10.1038/s12276-023-01133-7 (Dec. 2023).

15. Lee, K.-S. *et al.* Functional role of NF-κB in expression of human endothelial nitric oxide synthase. *Biochemical and Biophysical Research Communications* **448,** 101–107. ISSN: 0006-291X. doi:10.1016/j.bbrc.2014.04.079. http://dx.doi.org/10.1016/j.bbrc.2014.04.079 (May 2014).

16. Lane, K. *et al.* Measuring signaling and RNA-Seq in the same cell links gene expression to dynamic patterns of NF-kB activation. *Cell Systems* **4,** 458–469.e5. ISSN: 2405-4712. doi:10.1016/j.cels.2017.03.010. http://dx.doi.org/10.1016/j.cels.2017.03.010 (Apr. 2017).

17. Tay, S. *et al.* Single-cell NF-κB dynamics reveal digital activation and analogue information processing. *Nature* **466,** 267–271. ISSN: 1476-4687. doi:10.1038/nature09145. http://dx.doi.org/10.1038/nature09145 (June 2010).

18. Kellogg, R. A. & Tay, S. Noise facilitates transcriptional control under dynamic inputs. *Cell* **160,** 381–392. ISSN: 0092-8674. doi:10.1016/j.cell.2015.01.013. http://dx.doi.org/10.1016/j.cell.2015.01.013 (Jan. 2015).

19. Palla, G. *et al.* Squidpy: A scalable framework for spatial omics analysis. *Nature Methods* **19,** 171–178. doi:10.1038/s41592-021-01358-2. https://doi.org/10.1038/s41592-021-01358-2 (Jan. 2022).

20. Bressan, D., Battistoni, G. & Hannon, G. J. The dawn of spatial omics. *Science* **381.** ISSN: 1095-9203. doi:10.1126/science.abq4964. http://dx.doi.org/10.1126/science.abq4964 (Aug. 2023).

21. Lowe, R., Shirley, N., Bleackley, M., Dolan, S. & Shafee, T. Transcriptomics technologies. *PLOS Computational Biology* **13,** e1005457. ISSN: 1553-7358. doi:10.1371/journal.pcbi.1005457. http://dx.doi.org/10.1371/journal.pcbi.1005457 (May 2017).

22. Williams, C. G., Lee, H. J., Asatsuma, T., Vento-Tormo, R. & Haque, A. An introduction to spatial transcriptomics for biomedical research. *Genome Medicine* **14.** ISSN: 1756-994X. doi:10.1186/s13073-022-01075-1. http://dx.doi.org/10.1186/s13073-022-01075-1 (June 2022).

23. Moses, L. & Pachter, L. Museum of spatial transcriptomics. *Nature Methods* **19,** 534–546. ISSN: 1548-7105. doi:10.1038/s41592-022-01409-2. http://dx.doi.org/10.1038/s41592-022-01409-2 (Mar. 2022).

24. Rao, A., Barkley, D., França, G. S. & Yanai, I. Exploring tissue architecture using spatial transcriptomics. *Nature* **596,** 211–220. ISSN: 1476-4687. doi:10.1038/s41586-021-03634-9. http://dx.doi.org/10.1038/s41586-021-03634-9 (Aug. 2021).

25. Ståhl, P. L. *et al.* Visualization and analysis of gene expression in tissue sections by spatial transcriptomics. *Science* **353,** 78–82. doi:10.1126/science.aaf2403. https://doi.org/10.1126/science.aaf2403 (July 2016).

26. Vickovic, S. *et al.* High-definition spatial transcriptomics for in situ tissue profiling. *Nature Methods* **16,** 987–990. ISSN: 15487105. doi:10.1038/s41592-019-0548-y. http://dx.doi.org/10.1038/s41592-019-0548-y (2019).

27. Rodriques, S. G. *et al.* Slide-seq: A scalable technology for measuring genome-wide expression at high spatial resolution. *Science* **363,** 1463–1467. doi:10.1126/science.aaw1219. https://doi.org/10.1126/science.aaw1219 (Mar. 2019).

28. Stickels, R. R. *et al.* Highly sensitive spatial transcriptomics at near-cellular resolution with Slide-seqV2. *Nature Biotechnology* **39,** 313–319. doi:10.1038/s41587-020-0739-1. https://doi.org/10.1038/s41587-020-0739-1 (Dec. 2020).

29. Ke, R. *et al.* In situ sequencing for RNA analysis in preserved tissue and cells. *Nature Methods* **10,** 857–860. doi:10.1038/nmeth.2563. https://doi.org/10.1038/nmeth.2563 (July 2013).

30. Chen, K. H., Boettiger, A. N., Moffitt, J. R., Wang, S. & Zhuang, X. Spatially resolved, highly multiplexed RNA profiling in single cells. *Science* **348.** doi:10.1126/science.aaa6090. https://doi.org/10.1126/science.aaa6090 (Apr. 2015).

31. Codeluppi, S. *et al.* Spatial organization of the somatosensory cortex revealed by osmFISH. *Nature Methods* **15,** 932–935. ISSN: 15487105. doi:10.1038/s41592-018-0175-z. http://dx.doi.org/10.1038/s41592-018-0175-z (2018).

32. Eng, C.-H. L. *et al.* Transcriptome-scale super-resolved imaging in tissues by RNA seqFISH+. *Nature* **568,** 235–239. doi:10.1038/s41586-019-1049-y. https://doi.org/10.1038/s41586-019-1049-y (Mar. 2019).

33. Goh, J. J. L. *et al.* Highly specific multiplexed RNA imaging in tissues with split-FISH. *Nature Methods* **17,** 689–693. ISSN: 15487105. doi:10.1038/s41592-020-0858-0. http://dx.doi.org/10.1038/s41592-020-0858-0 (2020).

34. Hsieh, W.-C. *et al.* Spatial multi-omics analyses of the tumor immune microenvironment. *Journal of Biomedical Science* **29.** ISSN: 1423-0127. doi:10.1186/s12929-022-00879-y. http://dx.doi.org/10.1186/s12929-022-00879-y (Nov. 2022).

35. Janesick, A. *et al.* High resolution mapping of the tumor microenvironment using integrated single-cell, spatial and in situ analysis. *Nature Communications* **14.** ISSN: 2041-1723. doi:10.1038/s41467-023-43458-x. http://dx.doi.org/10.1038/s41467-023-43458-x (Dec. 2023).

36. Hunter, M. V., Moncada, R., Weiss, J. M., Yanai, I. & White, R. M. Spatially resolved transcriptomics reveals the architecture of the tumor-microenvironment interface. *Nature Communications* **12.** ISSN: 2041-1723. doi:10.1038/s41467-021-26614-z. http://dx.doi.org/10.1038/s41467-021-26614-z (Nov. 2021).

37. Arora, R. *et al.* Spatial transcriptomics reveals distinct and conserved tumor core and edge architectures that predict survival and targeted therapy response. *Nature Communications* **14.** ISSN: 2041-1723. doi:10.1038/s41467-023-40271-4. http://dx.doi.org/10.1038/s41467-023-40271-4 (Aug. 2023).

38. Zhang, M. *et al.* Spatially resolved cell atlas of the mouse primary motor cortex by MERFISH. *Nature* **598,** 137–143. ISSN: 1476-4687. doi:10.1038/s41586-021-03705-x. http://dx.doi.org/10.1038/s41586-021-03705-x (Oct. 2021).

39. Zhang, M. *et al.* Molecularly defined and spatially resolved cell atlas of the whole mouse brain. *Nature* **624,** 343–354. ISSN: 1476-4687. doi:10.1038/s41586-023-06808-9. http://dx.doi.org/10.1038/s41586-023-06808-9 (Dec. 2023).

40. Cunningham, R. P. & Porat-Shliom, N. Liver zonation – revisiting old questions with new technologies. *Frontiers in Physiology* **12.** ISSN: 1664-042X. doi:10.3389/fphys.2021.732929. http://dx.doi.org/10.3389/fphys.2021.732929 (Sept. 2021).

41. Paris, J. & Henderson, N. C. Liver zonation, revisited. *Hepatology* **76,** 1219–1230. ISSN: 1527-3350. doi:10.1002/hep.32408. http://dx.doi.org/10.1002/hep.32408 (Mar. 2022).

42. Yu, S. *et al.* Spatial transcriptome profiling of normal human liver. *Scientific Data* **9.** ISSN: 2052-4463. doi:10.1038/s41597-022-01676-w. http://dx.doi.org/10.1038/s41597-022-01676-w (Oct. 2022).

43. Hildebrandt, F. *et al.* Spatial Transcriptomics to define transcriptional patterns of zonation and structural components in the mouse liver. *Nature Communications* **12.** ISSN: 2041-1723. doi:10.1038/s41467-021-27354-w. http://dx.doi.org/10.1038/s41467-021-27354-w (Dec. 2021).

44. Baruzzo, G., Cesaro, G. & Di Camillo, B. Identify, quantify and characterize cellular communication from single-cell RNA sequencing data with scSeqComm. *Bioinformatics* **38** (ed Mathelier, A.) 1920–1929. ISSN: 1367-4811. doi:10.1093/bioinformatics/btac036. http://dx.doi.org/10.1093/bioinformatics/btac036 (Jan. 2022).

45. Lee, R. E., Walker, S. R., Savery, K., Frank, D. A. & Gaudet, S. Fold change of nuclear NF-κB determines TNF-induced transcription in single cells. *Molecular Cell* **53,** 867–879. ISSN: 1097-2765. doi:10.1016/j.molcel.2014.01.026. http://dx.doi.org/10.1016/j.molcel.2014.01.026 (Mar. 2014).

46. Maity, A. & Wollman, R. Information transmission from NFkB signaling dynamics to gene expression. *PLOS Computational Biology* **16** (ed Igoshin, O. A.) e1008011. ISSN: 1553-7358. doi:10.1371/journal.pcbi.1008011. http://dx.doi.org/10.1371/journal.pcbi.1008011 (Aug. 2020).

47. Specht, E. A., Braselmann, E. & Palmer, A. E. A critical and comparative review of fluorescent tools for live-cell imaging. *Annual Review of Physiology* **79,** 93–117. ISSN: 1545-1585. doi:10.1146/annurev-physiol-022516-034055. http://dx.doi.org/10.1146/annurev-physiol-022516-034055 (Feb. 2017).

48. Maryu, G. *et al.* Live-cell imaging with genetically encoded protein kinase activity reporters. *Cell Structure and Function* **43,** 61–74. ISSN: 1347-3700. doi:10.1247/csf.18003. http://dx.doi.org/10.1247/csf.18003 (2018).

49. Pargett, M. & Albeck, J. G. Live-cell imaging and analysis with multiple genetically encoded reporters. *Current Protocols in Cell Biology* **78.** ISSN: 1934-2616. doi:10.1002/cpcb.38. http://dx.doi.org/10.1002/cpcb.38 (Mar. 2018).

50. Heim, R. & Tsien, R. Y. Engineering green fluorescent protein for improved brightness, longer wavelengths and fluorescence resonance energy transfer. *Current Biology* **6,** 178–182. ISSN: 0960-9822. doi:10.1016/s0960-

9822(02)00450-5. http://dx.doi.org/10.1016/S0960-9822(02)00450-5 (Feb. 1996).

51. Lippincott-Schwartz, J., Snapp, E. & Kenworthy, A. Studying protein dynamics in living cells. *Nature Reviews Molecular Cell Biology* **2,** 444–456. ISSN: 1471-0080. doi:10.1038/35073068. http://dx.doi.org/10.1038/35073068 (June 2001).

52. Saito, Y. *et al.* Machine-learning-guided mutagenesis for directed evolution of fluorescent proteins. *ACS Synthetic Biology* **7,** 2014–2022. ISSN: 2161-5063. doi:10.1021/acssynbio.8b00155. http://dx.doi.org/10.1021/acssynbio.8b00155 (Aug. 2018).

53. Biswas, S., Khimulya, G., Alley, E. C., Esvelt, K. M. & Church, G. M. Low-N protein engineering with data-efficient deep learning. *Nature Methods* **18,** 389–396. ISSN: 1548-7105. doi:10.1038/s41592-021-01100-y. http://dx.doi.org/10.1038/s41592-021-01100-y (Apr. 2021).

54. Shaner, N. C., Steinbach, P. A. & Tsien, R. Y. A guide to choosing fluorescent proteins. *Nature Methods* **2,** 905–909. ISSN: 1548-7105. doi:10.1038/nmeth819. http://dx.doi.org/10.1038/nmeth819 (Nov. 2005).

55. Sampattavanich, S. *et al.* Encoding growth factor identity in the temporal dynamics of FOXO3 under the combinatorial control of ERK and AKT kinases. *Cell Systems* **6,** 664–678.e9. ISSN: 2405-4712. doi:10.1016/j.cels.2018.05.004. http://dx.doi.org/10.1016/j.cels.2018.05.004 (June 2018).

56. Mehta, S. & Zhang, J. Reporting from the field: Genetically encoded fluorescent reporters uncover signaling dynamics in living biological systems. *Annual Review of Biochemistry* **80,** 375–401. ISSN: 1545-4509. doi:10.1146/annurev-biochem-060409-093259. http://dx.doi.org/10.1146/annurev-biochem-060409-093259 (July 2011).

57. Ni, Q., Mehta, S. & Zhang, J. Live-cell imaging of cell signaling using genetically encoded fluorescent reporters. *The FEBS Journal* **285,** 203–219. ISSN: 1742-4658. doi:10.1111/febs.14134. http://dx.doi.org/10.1111/febs.14134 (July 2017).

58. Forster, T. Energiewanderung und fluoreszenz. *Die Naturwissenschaften* **33,** 166–175. ISSN: 1432-1904. doi:10.1007/bf00585226. http://dx.doi.org/10.1007/BF00585226 (1946).

59. Miyawaki, A. Visualization of the spatial and temporal dynamics of intracellular signaling. *Developmental Cell* **4,** 295–305. ISSN: 1534-5807. doi:10.1016/s1534-5807(03)00060-1. http://dx.doi.org/10.1016/S1534-5807(03)00060-1 (Mar. 2003).

60. Ni, Q., Titov, D. V. & Zhang, J. Analyzing protein kinase dynamics in living cells with FRET reporters. *Methods* **40,** 279–286. ISSN: 1046-2023. doi:10.1016/j.ymeth.2006.06.013. http://dx.doi.org/10.1016/j.ymeth.2006.06.013 (Nov. 2006).

61. Aoki, K., Komatsu, N., Hirata, E., Kamioka, Y. & Matsuda, M. Stable expression of FRET biosensors: A new light in cancer research. *Cancer Science* **103,** 614–619. ISSN: 1349-7006. doi:10.1111/j.1349-7006.2011.02196.x. http://dx.doi.org/10.1111/j.1349-7006.2011.02196.x (Feb. 2012).

62. Oldach, L. & Zhang, J. Genetically encoded fluorescent biosensors for live-cell visualization of protein phosphorylation. *Chemistry and Biology* **21,** 186–197. ISSN: 1074-5521. doi:10.1016/j.chembiol.2013.12.012. http://dx.doi.org/10.1016/j.chembiol.2013.12.012 (Feb. 2014).

63. Chung, C.-I. *et al.* Intrabody-based FRET probe to visualize endogenous histone acetylation. *Scientific Reports* **9.** ISSN: 2045-2322. doi:10.1038/s41598-019-46573-2. http://dx.doi.org/10.1038/s41598-019-46573-2 (July 2019).

64. Nakaoka, S., Sasaki, K., Ito, A., Nakao, Y. & Yoshida, M. A genetically encoded FRET probe to detect intranucleosomal histone H3K9 or H3K14 acetylation using BRD4, a BET family member. *ACS Chemical Biology* **11,** 729–733. ISSN: 1554-8937. doi:10.1021/cb501046t. http://dx.doi.org/10.1021/cb501046t (May 2015).

65. Shih, W. M., Gryczynski, Z., Lakowicz, J. R. & Spudich, J. A. A FRET-based sensor reveals large ATP hydrolysis–induced conformational changes and three distinct states of the molecular motor myosin. *Cell* **102,** 683–694. ISSN: 0092-8674. doi:10.1016/s0092-8674(00)00090-8. http://dx.doi.org/10.1016/S0092-8674(00)00090-8 (Sept. 2000).

66. Heyduk, T. Measuring protein conformational changes by FRET/LRET. *Current Opinion in Biotechnology* **13,** 292–296. ISSN: 0958-1669. doi:10.1016/s0958-1669(02)00332-4. http://dx.doi.org/10.1016/S0958-1669(02)00332-4 (Aug. 2002).

67. Allen, M. D. & Zhang, J. Subcellular dynamics of protein kinase A activity visualized by FRET-based reporters. *Biochemical and Biophysical Research Communications* **348,** 716–721. ISSN: 0006-291X. doi:10.1016/j.bbrc.2006.07.136. http://dx.doi.org/10.1016/j.bbrc.2006.07.136 (Sept. 2006).

68. Depry, C. & Zhang, J. in *Signal Transduction Protocols* 285–294 (Humana Press, 2011). ISBN: 9781617791604. doi:10.1007/978-1-61779-160-4_16. http://dx.doi.org/10.1007/978-1-61779-160-4_16.

69. Asher, W. B. *et al.* Single-molecule FRET imaging of GPCR dimers in living cells. *Nature Methods* **18,** 397–405. ISSN: 1548-7105. doi:10.1038/s41592-021-01081-y. http://dx.doi.org/10.1038/s41592-021-01081-y (Mar. 2021).

70. Regot, S., Hughey, J. J., Bajar, B. T., Carrasco, S. & Covert, M. W. High-sensitivity measurements of multiple kinase activities in live single cells. *Cell* **157,** 1724–1734. ISSN: 0092-8674. doi:10.1016/j.cell.2014.04.039. http://dx.doi.org/10.1016/j.cell.2014.04.039 (June 2014).

71. Kudo, T. *et al.* Live-cell measurements of kinase activity in single cells using translocation reporters. *Nature Protocols* **13,** 155–169. ISSN: 1750-2799. doi:10.1038/nprot.2017.128. http://dx.doi.org/10.1038/nprot.2017.128 (Dec. 2017).

72. Nagai, T., Sawano, A., Park, E. S. & Miyawaki, A. Circularly permuted green fluorescent proteins engineered to sense Ca 2+. *Proceedings of the National Academy of Sciences* **98,** 3197–3202. ISSN: 1091-6490. doi:10.1073/pnas.051636098. http://dx.doi.org/10.1073/pnas.051636098 (Mar. 2001).

73. Nakai, J., Ohkura, M. & Imoto, K. A high signal-to-noise Ca2+ probe composed of a single green fluorescent protein. *Nature Biotechnology* **19,** 137–141. ISSN: 1546-1696. doi:10.1038/84397. http://dx.doi.org/10.1038/84397 (Feb. 2001).

74. Mehta, S. *et al.* Single-fluorophore biosensors for sensitive and multiplexed detection of signalling activities. *Nature Cell Biology* **20,** 1215–1225. ISSN: 1476-4679. doi:10.1038/s41556-018-0200-6. http://dx.doi.org/10.1038/s41556-018-0200-6 (Sept. 2018).

75. Zhang, L. *et al.* Clinical and translational values of spatial transcriptomics. *Signal Transduction and Targeted Therapy* **7.** doi:10.1038/s41392-022-00960-w. https://doi.org/10.1038/s41392-022-00960-w (Apr. 2022).

76. Schmitt, D. L. *et al.* Spatial regulation of AMPK signaling revealed by a sensitive kinase activity reporter. *Nature Communications* **13.** ISSN: 2041-1723. doi:10.1038/s41467-022-31190-x. http://dx.doi.org/10.1038/s41467-022-31190-x (July 2022).

77. Zhang, Q. *et al.* Visualizing dynamics of cell signaling in vivo with a phase separation-based kinase reporter. *Molecular Cell* **69,** 334–346.e4. ISSN: 1097-2765. doi:10.1016/j.molcel.2017.12.008. http://dx.doi.org/10.1016/j.molcel.2017.12.008 (Jan. 2018).

78. Li, X. *et al.* ATM-SPARK: A GFP phase separation–based activity reporter of ATM. *Science Advances* **9.** ISSN: 2375-2548. doi:10.1126/sciadv.ade3760. http://dx.doi.org/10.1126/sciadv.ade3760 (Mar. 2023).

79. Herzig, S. & Shaw, R. J. AMPK: Guardian of metabolism and mitochondrial homeostasis. *Nature Reviews Molecular Cell Biology* **19,** 121–135. ISSN: 1471-0080. doi:10.1038/nrm.2017.95. http://dx.doi.org/10.1038/nrm.2017.95 (Oct. 2017).

80. Richter, E. A. & Ruderman, N. B. AMPK and the biochemistry of exercise: implications for human health and disease. *Biochemical Journal* **418,** 261–275. ISSN: 1470-8728. doi:10.1042/bj20082055. http://dx.doi.org/10.1042/BJ20082055 (Feb. 2009).

81. Trefts, E. & Shaw, R. J. AMPK: Restoring metabolic homeostasis over space and time. *Molecular Cell* **81,** 3677–3690. ISSN: 1097-2765. doi:10.1016/j.molcel.2021.08.015. http://dx.doi.org/10.1016/j.molcel.2021.08.015 (Sept. 2021).

82. Najar, M. A. *et al.* A complete map of the Calcium/calmodulin-dependent protein kinase kinase 2 (CAMKK2) signaling pathway. *Journal of Cell Communication and Signaling* **15,** 283–290. ISSN: 1873-961X. doi:10.1007/s12079-020-00592-1. http://dx.doi.org/10.1007/s12079-020-00592-1 (Nov. 2020).

83. Shackelford, D. B. & Shaw, R. J. The LKB1–AMPK pathway: Metabolism and growth control in tumour suppression. *Nature Reviews Cancer* **9,** 563–575. ISSN: 1474-1768. doi:10.1038/nrc2676. http://dx.doi.org/10.1038/nrc2676 (Aug. 2009).

84. Tsou, P., Zheng, B., Hsu, C. H., Sasaki, A. T. & Cantley, L. C. A fluorescent reporter of AMPK activity and cellular energy stress. *Cell Metabolism* **13,** 476–486. ISSN: 15504131. doi:10.1016/j.cmet.2011.03.006 (2011).

85. Huang, P. S. *et al.* High thermodynamic stability of parametrically designed helical bundles. *Science* **346,** 481–485. ISSN: 10959203. doi:10.1126/science.1257481 (2014).

86. Thomson, A. R. *et al.* Computational design of water-soluble. *Science* **280,** 485–488 (2014).

87. Durocher, D., Henckel, J., Fersht, A. R. & Jackson, S. P. The FHA Domain Is a Modular Phosphopeptide Recognition Motif. *Molecular Cell* **4,** 387–394. ISSN: 1097-2765. doi:10.1016/s1097-2765(00)80340-8. http://dx.doi.org/10.1016/s1097-2765(00)80340-8 (Sept. 1999).

88. Durocher, D. *et al.* The molecular basis of FHA domain: phosphopeptide binding specificity and implications for phospho-dependent signaling mechanisms. *Molecular Cell* **6,** 1169–1182. ISSN: 1097-2765. doi:10.1016/s1097-2765(00)00114-3. http://dx.doi.org/10.1016/S1097-2765(00)00114-3 (Nov. 2000).

89. Kim, J. H. *et al.* High cleavage efficiency of a 2A peptide derived from porcine teschovirus-1 in human cell lines, zebrafish and mice. *PLoS ONE* **6** (ed Thiel, V.) e18556. ISSN: 1932-6203. doi:10.1371/journal.pone.0018556. http://dx.doi.org/10.1371/journal.pone.0018556 (Apr. 2011).

90. Bagur, R. & Hajnóczky, G. Intracellular Ca2+ sensing: Its role in calcium homeostasis and signaling. *Molecular Cell* **66,** 780–788. ISSN: 1097-2765. doi:10.1016/j.molcel.2017.05.028. http://dx.doi.org/10.1016/j.molcel.2017.05.028 (June 2017).

91. Liu, C.-m. & Hermann, T. E. Characterization of ionomycin as a calcium ionophore. *Journal of Biological Chemistry* **253,** 5892–5894. ISSN: 0021-9258. doi:10.1016/s0021-9258(17)34550-7. http://dx.doi.org/10.1016/S0021-9258(17)34550-7 (Sept. 1978).

92. Racioppi, L. & Means, A. R. Calcium/calmodulin-dependent protein kinase kinase 2: Roles in signaling and pathophysiology. *Journal of Biological Chemistry* **287,** 31658–31665. ISSN: 0021-9258. doi:10.1074/jbc.r112.356485. http://dx.doi.org/10.1074/jbc.R112.356485 (Sept. 2012).

93. Anderson, K. A. *et al.* Hypothalamic CaMKK2 contributes to the regulation of energy balance. *Cell Metabolism* **7,** 377–388. ISSN: 1550-4131. doi:10.1016/j.cmet.2008.02.011. http://dx.doi.org/10.1016/j.cmet.2008.02.011 (May 2008).

94. Pajak, B. *et al.* 2-Deoxy-d-Glucose and its analogs: From diagnostic to therapeutic agents. *International Journal of Molecular Sciences* **21,** 234. ISSN: 1422-0067. doi:10.3390/ijms21010234. http://dx.doi.org/10.3390/ijms21010234 (Dec. 2019).

95. Laubscher, E. *et al.* Accurate single-molecule spot detection for image-based spatial transcriptomics with weakly supervised deep learning. *Cell Systems* **15,** 475–482.e6. ISSN: 2405-4712. doi:10.1016/j.cels.2024.04.006. http://dx.doi.org/10.1016/j.cels.2024.04.006 (May 2024).

96. Kong, F., You, H., Zheng, K., Tang, R. & Zheng, C. The crosstalk between pattern-recognition receptor signaling and calcium signaling. *International Journal of Biological Macromolecules* **192,** 745–756. ISSN: 0141-8130. doi:10.1016/j.ijbiomac.2021.10.014. http://dx.doi.org/10.1016/j.ijbiomac.2021.10.014 (Dec. 2021).

97. Yoo, H., Triandafillou, C. & Drummond, D. A. Cellular sensing by phase separation: Using the process, not just the products. *Journal of Biological Chemistry* **294,** 7151–7159. ISSN: 0021-9258. doi:10.1074/jbc.tm118.001191. http://dx.doi.org/10.1074/jbc.TM118.001191 (May 2019).

98. Alberti, S. Phase separation in biology. *Current Biology* **27,** R1097–R1102. ISSN: 0960-9822. doi:10.1016/j.cub.2017.08.069. http://dx.doi.org/10.1016/j.cub.2017.08.069 (Oct. 2017).

99. Alberti, S., Gladfelter, A. & Mittag, T. Considerations and challenges in studying liquid-liquid phase separation and biomolecular condensates. *Cell* **176,** 419–434. ISSN: 0092-8674. doi:10.1016/j.cell.2018.12.035. http://dx.doi.org/10.1016/j.cell.2018.12.035 (Jan. 2019).

100. Honda, K., Takaoka, A. & Taniguchi, T. Type I inteferon gene induction by the interferon regulatory factor family of transcription factors. *Immunity* **25,** 349–360. ISSN: 1074-7613. doi:10.1016/j.immuni.2006.08.009. http://dx.doi.org/10.1016/j.immuni.2006.08.009 (Sept. 2006).

101. Yanai, H. *et al.* Revisiting the role of IRF3 in inflammation and immunity by conditional and specifically targeted gene ablation in mice. *Proceedings of the National Academy of Sciences* **115,** 5253–5258. ISSN: 1091-6490. doi:10.1073/pnas.1803936115. http://dx.doi.org/10.1073/pnas.1803936115 (Apr. 2018).

102. Jefferies, C. A. Regulating IRFs in IFN driven disease. *Frontiers in Immunology* **10.** ISSN: 1664-3224. doi:10.3389/fimmu.2019.00325. http://dx.doi.org/10.3389/fimmu.2019.00325 (Mar. 2019).

103. Dorrington, M. G. & Fraser, I. D. C. NF-kB signaling in macrophages: Dynamics, crosstalk, and signal integration. *Frontiers in Immunology* **10.** ISSN: 1664-3224. doi:10.3389/fimmu.2019.00705. http://dx.doi.org/10.3389/fimmu.2019.00705 (Apr. 2019).

104. Balka, K. R. *et al.* TBK1 and IKKϵ act redundantly to mediate STING-induced NF-kB responses in myeloid cells. *Cell Reports* **31,** 107492. ISSN: 2211-1247. doi:10.1016/j.celrep.2020.03.056. http://dx.doi.org/10.1016/j.celrep.2020.03.056 (Apr. 2020).

105. Kravchenko, V. V., Mathison, J. C., Schwamborn, K., Mercurio, F. & Ulevitch, R. J. IKKi/IKKϵ plays a key role in integrating signals induced by pro-inflammatory stimuli. *Journal of Biological Chemistry* **278,** 26612–26619. ISSN: 0021-9258. doi:10.1074/jbc.m303001200. http://dx.doi.org/10.1074/jbc.M303001200 (July 2003).

106. Yum, S., Li, M., Fang, Y. & Chen, Z. J. TBK1 recruitment to STING activates both IRF3 and NF-κB that mediate immune defense against tumors and viral infections. *Proceedings of the National Academy of Sciences* **118.** ISSN: 1091-6490. doi:10.1073/pnas.2100225118. http://dx.doi.org/10.1073/pnas.2100225118 (Mar. 2021).

107. Park, M. & Hong, J. Roles of NF-κB in cancer and inflammatory diseases and their therapeutic approaches. *Cells* **5,** 15. ISSN: 2073-4409. doi:10.3390/cells5020015. http://dx.doi.org/10.3390/cells5020015 (Mar. 2016).

108. Petro, T. M. IFN Regulatory Factor 3 in health and disease. *The Journal of Immunology* **205,** 1981–1989. ISSN: 1550-6606. doi:10.4049/jimmunol.

2000462. http://dx.doi.org/10.4049/jimmunol.2000462 (Oct. 2020).

109. Thompson, C. D., Matta, B. & Barnes, B. J. Therapeutic targeting of IRFs: Pathway-dependence or structure-based? *Frontiers in Immunology* **9.** ISSN: 1664-3224. doi:10.3389/fimmu.2018.02622. http://dx.doi.org/10.3389/fimmu.2018.02622 (Nov. 2018).

110. Yu, H., Lin, L., Zhang, Z., Zhang, H. & Hu, H. Targeting NF-κB pathway for the therapy of diseases: mechanism and clinical study. *Signal Transduction and Targeted Therapy* **5.** ISSN: 2059-3635. doi:10.1038/s41392-020-00312-6. http://dx.doi.org/10.1038/s41392-020-00312-6 (Sept. 2020).

111. Di Francesco, B. *et al.* NF-κB: A druggable target in acute myeloid leukemia. *Cancers* **14,** 3557. ISSN: 2072-6694. doi:10.3390/cancers14143557. http://dx.doi.org/10.3390/cancers14143557 (July 2022).

112. Li, Y. *et al.* Inhibition of NF-κB signaling unveils novel strategies to overcome drug resistance in cancers. *Drug Resistance Updates* **73,** 101042. ISSN: 1368-7646. doi:10.1016/j.drup.2023.101042. http://dx.doi.org/10.1016/j.drup.2023.101042 (Mar. 2024).

113. Lipniacki, T., Paszek, P., Brasier, A. R., Luxon, B. & Kimmel, M. Mathematical model of NF-κB regulatory module. *Journal of Theoretical Biology* **228,** 195–215. ISSN: 0022-5193. doi:10.1016/j.jtbi.2004.01.001. http://dx.doi.org/10.1016/j.jtbi.2004.01.001 (May 2004).

114. Ihekwaba, A., Benson, N., Kell, D., Broomhead, D. & Grimley, R. Sensitivity analysis of parameters controlling oscillatory signalling in the NF-κB pathway: The roles of IKK and IκBα. *Systems Biology* **1,** 93–103. ISSN: 1741-248X. doi:10.1049/sb:20045009. http://dx.doi.org/10.1049/sb:20045009 (June 2004).

115. Cheong, R., Hoffmann, A. & Levchenko, A. Understanding NF-κB signaling via mathematical modeling. *Molecular Systems Biology* **4.** ISSN: 1744-4292. doi:10.1038/msb.2008.30. http://dx.doi.org/10.1038/msb.2008.30 (Jan. 2008).

116. Basak, S., Behar, M. & Hoffmann, A. Lessons from mathematically modeling the NF-κB pathway. *Immunological Reviews* **246,** 221–238. ISSN: 1600-065X. doi:10.1111/j.1600-065x.2011.01092.x. http://dx.doi.org/10.1111/j.1600-065X.2011.01092.x (Mar. 2012).

117. Korwek, Z. *et al.* Nonself RNA rewires IFN-β signaling: A mathematical model of the innate immune response. *Science Signaling* **16.** ISSN: 1937-9145. doi:10.1126/scisignal.abq1173. http://dx.doi.org/10.1126/scisignal.abq1173 (Dec. 2023).

118. Burkart, S. S. *et al.* High-resolution kinetic characterization of the RIG-I-signaling pathway and the antiviral response. *Life Science Alliance* **6,** e202302059. ISSN: 2575-1077. doi:10.26508/lsa.202302059. http://dx.doi.org/10.26508/lsa.202302059 (Aug. 2023).

119. Hoffmann, A., Levchenko, A., Scott, M. L. & Baltimore, D. The IκB-NF-κB signaling module: Temporal control and selective gene activation. *Science* **298,** 1241–1245. ISSN: 1095-9203. doi:10.1126/science.1071914. http://dx.doi.org/10.1126/science.1071914 (Nov. 2002).

120. Suzuki, T. *et al.* Cell type-specific subcellular localization of phospho-TBK1 in response to cytoplasmic viral DNA. *PLoS ONE* **8** (ed Guo, H.) e83639. ISSN: 1932-6203. doi:10.1371/journal.pone.0083639. http://dx.doi.org/10.1371/journal.pone.0083639 (Dec. 2013).

121. Diella, F., Gould, C. M., Chica, C., Via, A. & Gibson, T. J. Phospho.ELM: A database of phosphorylation sites. *Nucleic Acids Research* **36,** D240–D244. ISSN: 1362-4962. doi:10.1093/nar/gkm772. http://dx.doi.org/10.1093/nar/gkm772 (Dec. 2007).

122. Hornbeck, P. V. *et al.* PhosphoSitePlus, 2014: Mutations, PTMs and recalibrations. *Nucleic Acids Research* **43,** D512–D520. ISSN: 0305-1048. doi:10.1093/nar/gku1267. http://dx.doi.org/10.1093/nar/gku1267 (Dec. 2014).

123. Ma, R. *et al.* KinasePhos 3.0: redesign and expansion of the prediction on kinase-specific phosphorylation sites. *Genomics, Proteomics and Bioinformatics* **21,** 228–241. ISSN: 1672-0229. doi:10.1016/j.gpb.2022.06.004. http://dx.doi.org/10.1016/j.gpb.2022.06.004 (Feb. 2023).

124. Komal, A., Noreen, M. & El-Kott, A. F. TLR3 agonists: RGC100, ARNAX, and poly-IC: A comparative review. *Immunologic Research* **69,** 312–322. ISSN: 1559-0755. doi:10.1007/s12026-021-09203-6. http://dx.doi.org/10.1007/s12026-021-09203-6 (June 2021).

125. Tak, P. P. & Firestein, G. S. NF-κB: A key role in inflammatory diseases. *Journal of Clinical Investigation* **107,** 7–11. ISSN: 0021-9738. doi:10.1172/jci11830. http://dx.doi.org/10.1172/JCI11830 (Jan. 2001).

126. Karin, M., Yamamoto, Y. & Wang, Q. M. The IKK NF-κB system: A treasure trove for drug development. *Nature Reviews Drug Discovery* **3,** 17–26. ISSN: 1474-1784. doi:10.1038/nrd1279. http://dx.doi.org/10.1038/nrd1279 (Jan. 2004).

127. Prescott, J. & Cook, S. Targeting IKKβ in cancer: Challenges and opportunities for the therapeutic utilisation of IKKβ inhibitors. *Cells* **7,** 115. ISSN: 2073-4409. doi:10.3390/cells7090115. http://dx.doi.org/10.3390/cells7090115 (Aug. 2018).

128. Hasan, M. & Yan, N. Therapeutic potential of targeting TBK1 in autoimmune diseases and interferonopathies. *Pharmacological Research* **111,** 336–342. ISSN: 1043-6618. doi:10.1016/j.phrs.2016.04.008. http://dx.doi.org/10.1016/j.phrs.2016.04.008 (Sept. 2016).

129. Hasan, M. *et al.* Cutting edge: Inhibiting TBK1 by compound II ameliorates autoimmune disease in mice. *The Journal of Immunology* **195,** 4573–4577. ISSN: 1550-6606. doi:10.4049/jimmunol.1500162. http://dx.doi.org/10.4049/jimmunol.1500162 (Nov. 2015).

130. Bodewes, I. L. *et al.* TBK1: A key regulator and potential treatment target for interferon positive Sjögren's syndrome, systemic lupus erythematosus and systemic sclerosis. *Journal of Autoimmunity* **91,** 97–102. ISSN: 0896-8411. doi:10.1016/j.jaut.2018.02.001. http://dx.doi.org/10.1016/j.jaut.2018.02.001 (July 2018).

131. Israel, U. *et al.* A foundation model for cell segmentation. doi:10.1101/2023.11.17.567630. http://dx.doi.org/10.1101/2023.11.17.567630 (Nov. 2023).

132. Bochinski, E., Eiselein, V. & Sikora, T. *High-speed tracking-by-detection without using image information* in *2017 14th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)* (IEEE, Aug. 2017). doi:10.1109/avss.2017.8078516. http://dx.doi.org/10.1109/AVSS.2017.8078516.

133. Asp, M., Bergenstråhle, J. & Lundeberg, J. Spatially resolved transcriptomes—next generation tools for tissue exploration. *BioEssays* **42,** 1–16. ISSN: 15211878. doi:10.1002/bies.201900221 (2020).

134. Axelrod, S. *et al.* Starfish: Scalable pipelines for image-based transcriptomics. *Journal of Open Source Software* **6,** 2440. ISSN: 2475-9066. doi:10.21105/joss.02440 (2021).

135. Cisar, C., Keener, N., Ruffalo, M. & Paten, B. A unified pipeline for FISH spatial transcriptomics. *Cell Genomics,* 100384. doi:10.1016/j.xgen.2023.100384. https://doi.org/10.1016/j.xgen.2023.100384 (Aug. 2023).

136. Valen, D. A. V. *et al.* Deep learning automates the quantitative analysis of individual cells in live-cell imaging experiments. *PLOS Computational Biology* **12** (ed Meier-Schellersheim, M.) e1005177. doi:10.1371/journal.pcbi.1005177. https://doi.org/10.1371/journal.pcbi.1005177 (Nov. 2016).

137. Stringer, C., Wang, T., Michaelos, M. & Pachitariu, M. Cellpose: A generalist algorithm for cellular segmentation. *Nature Methods* **18,** 100–106. doi:10.1038/s41592-020-01018-x (2020).

138. Greenwald, N. F. *et al.* Whole-cell segmentation of tissue images with human-level performance using large-scale data annotation and deep learning. *Nature Biotechnology* **40,** 555–565. doi:10.1038/s41587-021-01094-0. https://doi.org/10.1038/s41587-021-01094-0 (Nov. 2021).

139. Pachitariu, M. & Stringer, C. Cellpose 2.0: How to train your own model. *Nature Methods* **19,** 1634–1641. doi:10.1038/s41592-022-01663-4. https://doi.org/10.1038/s41592-022-01663-4 (Nov. 2022).

140. Mabaso, M., Withey, D. & Twala, B. Spot detection methods in fluorescence microscopy imaging: A review. *Image Analysis and Stereology* **37,** 173–190. ISSN: 18545165. doi:10.5566/ias.1690 (2018).

141. Van der Walt, S. *et al.* scikit-image: Image processing in Python. *PeerJ* **2,** e453. doi:10.7717/peerj.453. https://doi.org/10.7717/peerj.453 (June 2014).

142. Allan, D. B., Caswell, T., Keim, N. C., van der Wel, C. M. & Verweij, R. W. *soft-matter/trackpy: Trackpy v0.5.0* 2021. doi:10.5281/ZENODO.4682814. https://zenodo.org/record/4682814.

143. Gudla, P., Nakayama, K., Pegoraro, G. & Mistelli, T. SpotLearn: Convolutional neural network for detection of fluorescence in situ hybridization (FISH) signals in high- throughput imaging approaches. *Cold Spring Harbor Symposia on Quantitative Biology* **82,** 57–70. doi:10.1101/sqb.2017.82.033761.SpotLearn (2017).

144. Eichenberger, B. T., Zhan, Y., Rempfler, M., Giorgetti, L. & Chao, J. A. deepBlink : threshold-independent detection and localization of diffraction-limited spots. *Nucleic Acids Research* **49,** 7292–7297 (2021).

145. Wollmann, T. & Rohr, K. Deep Consensus Network: Aggregating predictions to improve object detection in microscopy images. *Medical Image Analysis* **70,** 102019. ISSN: 13618423. doi:10.1016/j.media.2021.102019. https://doi.org/10.1016/j.media.2021.102019 (2021).

146. Ratner, A. *et al.* Snorkel: Rapid training data creation with weak supervision. *The VLDB Journal* **29,** 709–730. doi:10.1007/s00778-019-00552-1. https://doi.org/10.1007/s00778-019-00552-1 (July 2019).

147. Moon, T. K. The Expectation-Maximization algorithm. *IEEE Signal Processing Magazine* **13,** 47–60. doi:10.1109/79.543975. https://doi.org/10.1109/79.543975 (1996).

148. Moffitt, J. R. *et al.* High-throughput single-cell gene-expression profiling with multiplexed error-robust fluorescence in situ hybridization. *Proceedings of the National Academy of Sciences* **113,** 11046–11051. doi:10.1073/pnas.1612826113. https://doi.org/10.1073/pnas.1612826113 (Sept. 2016).

149. Petukhov, V. *et al.* Cell segmentation in imaging-based spatial transcriptomics. *Nature Biotechnology* **40,** 345–354. ISSN: 1546-1696. doi:10.1038/s41587-021-01044-w. https://doi.org/10.1038/s41587-021-01044-w (2022).

150. Boersma, S. *et al.* Translation and replication dynamics of single RNA viruses. *Cell* **183,** 1930–1945.e23. ISSN: 0092-8674. doi:https://doi.org/10.1016/j.cell.2020.10.019. https://www.sciencedirect.com/science/article/pii/S0092867420313829 (2020).

151. Smal, I., Loog, M., Niessen, W. & Meijering, E. Quantitative comparison of spot detection methods in fluorescence microscopy. *IEEE Transactions on Medical Imaging* **29,** 282–301. ISSN: 02780062. doi:10.1109/TMI.2009.2025127 (2010).

152. Ruusuvuori, P. *et al.* Evaluation of methods for detection of fluorescence labeled subcellular objects in microscope images. *BMC Bioinformatics* **11.** doi:10.1186/1471-2105-11-248. https://doi.org/10.1186/1471-2105-11-248 (May 2010).

153. Hoffman, M. D., Blei, D. M., Wang, C. & Paisley, J. Stochastic variational inference. *Jounral of Machine Learning Research* **14,** 1303–1347. ISSN: 1 (2013).

154. Gataric, M. *et al.* PoSTcode: Probabilistic image-based spatial transcriptomics decoder. *bioRxiv*. doi:10.1101/2021.10.12.464086. https://doi.org/10.1101/2021.10.12.464086 (Oct. 2021).

155. Liu, J. *et al.* Concordance of MERFISH spatial transcriptomics with bulk and single-cell RNA sequencing. *Life Science Alliance* **6,** e202201701. doi:10.26508/lsa.202201701. https://doi.org/10.26508/lsa.202201701 (Dec. 2022).

156. Feldman, D. *et al.* Optical pooled screens in human cells. *Cell* **179,** 787–799.e17. doi:10.1016/j.cell.2019.09.016. https://doi.org/10.1016/j.cell.2019.09.016 (Oct. 2019).

157. Zhao, P., Zhu, J., Ma, Y. & Zhou, X. Modeling zero inflation is not necessary for spatial transcriptomics. *Genome Biology* **23.** doi:10.1186/s13059-022-02684-0. https://doi.org/10.1186/s13059-022-02684-0 (May 2022).

158. Satija, R., Farrell, J. A., Gennert, D., Schier, A. F. & Regev, A. Spatial reconstruction of single-cell gene expression data. *Nature Biotechnology* **33,** 495–502. doi:10.1038/nbt.3192. https://doi.org/10.1038/nbt.3192 (Apr. 2015).

159. Beliveau, B. J. *et al.* OligoMiner provides a rapid, flexible environment for the design of genome-scale oligonucleotide in situ hybridization probes. *Proceedings of the National Academy of Sciences* **115,** E2183–E2192. doi:10.

1073/pnas.1714530115. https://doi.org/10.1073/pnas.1714530115 (Feb. 2018).

160. Lionnet, T. *et al.* A transgenic mouse for in vivo detection of endogenous labeled mRNA. *Nature Methods* **8,** 165–170. ISSN: 1548-7105. doi:10.1038/nmeth.1551. http://dx.doi.org/10.1038/nmeth.1551 (Jan. 2011).

161. Thompson, R. E., Larson, D. R. & Webb, W. W. Precise nanometer localization analysis for individual fluorescent probes. *Biophysical Journal* **82,** 2775–2783. doi:10.1016/s0006-3495(02)75618-x. https://doi.org/10.1016/s0006-3495(02)75618-x (May 2002).

162. Abadi, M. *et al.* TensorFlow: large-scale machine learning on heterogeneous distributed systems. *CoRR* **abs/1603.04467.** arXiv: 1603.04467. http://arxiv.org/abs/1603.04467 (2016).

163. Mass, E., Nimmerjahn, F., Kierdorf, K. & Schlitzer, A. Tissue-specific macrophages: How they develop and choreograph tissue biology. *Nature Reviews Immunology* **23,** 563–579. ISSN: 1474-1741. doi:10.1038/s41577-023-00848-y. http://dx.doi.org/10.1038/s41577-023-00848-y (Mar. 2023).

164. Wynn, T. A., Chawla, A. & Pollard, J. W. Macrophage biology in development, homeostasis and disease. *Nature* **496,** 445–455. ISSN: 1476-4687. doi:10.1038/nature12034. http://dx.doi.org/10.1038/nature12034 (Apr. 2013).

165. Park, M. D., Silvin, A., Ginhoux, F. & Merad, M. Macrophages in health and disease. *Cell* **185,** 4259–4279. ISSN: 0092-8674. doi:10.1016/j.cell.2022.10.007. http://dx.doi.org/10.1016/j.cell.2022.10.007 (Nov. 2022).

166. Ross, E. A., Devitt, A. & Johnson, J. R. Macrophages: The good, the bad, and the gluttony. *Frontiers in Immunology* **12.** ISSN: 1664-3224. doi:10.3389/fimmu.2021.708186. http://dx.doi.org/10.3389/fimmu.2021.708186 (Aug. 2021).

167. McWhorter, F. Y., Wang, T., Nguyen, P., Chung, T. & Liu, W. F. Modulation of macrophage phenotype by cell shape. *Proceedings of the National Academy of Sciences* **110,** 17253–17258. ISSN: 1091-6490. doi:10.1073/pnas.1308887110. http://dx.doi.org/10.1073/pnas.1308887110 (Oct. 2013).

168. Singh, A., Sen, S., Adelaja, A. & Hoffmann, A. Stimulus-response signaling dynamics characterize macrophage polarization states. doi:10.1101/2022.03.27.485991. http://dx.doi.org/10.1101/2022.03.27.485991 (Mar. 2022).

169. Hourani, T. *et al.* Label-free macrophage phenotype classification using machine learning methods. *Scientific Reports* **13.** ISSN: 2045-2322. doi:10.1038/s41598-023-32158-7. http://dx.doi.org/10.1038/s41598-023-32158-7 (Mar. 2023).

170. Rostam, H. M., Reynolds, P. M., Alexander, M. R., Gadegaard, N. & Ghaemmaghami, A. M. Image based Machine Learning for identification of macrophage subsets. *Scientific Reports* **7.** ISSN: 2045-2322. doi:10.1038/s41598-017-03780-z. http://dx.doi.org/10.1038/s41598-017-03780-z (June 2017).

171. Selig, M. *et al.* Prediction of six macrophage phenotypes and their IL-10 content based on single-cell morphology using artificial intelligence. *Frontiers in Immunology* **14.** ISSN: 1664-3224. doi:10.3389/fimmu.2023.1336393. http://dx.doi.org/10.3389/fimmu.2023.1336393 (Jan. 2024).

172. Kesapragada, M. *et al.* A data-driven approach to establishing cell motility patterns as predictors of macrophage subtypes and their relation to cell morphology. doi:10.1101/2022.11.29.518400. http://dx.doi.org/10.1101/2022.11.29.518400 (Dec. 2022).

173. Eligini, S. *et al.* Human monocyte-derived macrophages are heterogenous: Proteomic profile of different phenotypes. *Journal of Proteomics* **124,** 112–123. ISSN: 1874-3919. doi:10.1016/j.jprot.2015.03.026. http://dx.doi.org/10.1016/j.jprot.2015.03.026 (June 2015).

174. Phillip, J. M., Han, K.-S., Chen, W.-C., Wirtz, D. & Wu, P.-H. A robust unsupervised machine-learning method to quantify the morphological heterogeneity of cells and nuclei. *Nature Protocols* **16,** 754–774. ISSN: 1750-2799. doi:10.1038/s41596-020-00432-x. http://dx.doi.org/10.1038/s41596-020-00432-x (Jan. 2021).

175. Keller, A.-A., Maeß, M. B., Schnoor, M., Scheiding, B. & Lorkowski, S. in *Methods in Molecular Biology* 187–195 (Springer New York, 2018). ISBN: 9781493978373. doi:10.1007/978-1-4939-7837-3_18. http://dx.doi.org/10.1007/978-1-4939-7837-3_18.

176. Maess, M. B., Wittig, B. & Lorkowski, S. Highly efficient transfection of human THP-1 macrophages by nucleofection. *Journal of Visualized Experiments.* ISSN: 1940-087X. doi:10.3791/51960. http://dx.doi.org/10.3791/51960 (Sept. 2014).

177. Maess, M. B., Keller, A.-A., Rennert, K., Mosig, A. & Lorkowski, S. Optimization of the transfection of human THP-1 macrophages by application of Nunc UpCell technology. *Analytical Biochemistry* **479,** 40–42. ISSN: 0003-2697. doi:10.1016/j.ab.2014.12.023. http://dx.doi.org/10.1016/j.ab.2014.12.023 (June 2015).

178. Moradian, H., Roch, T., Lendlein, A. & Gossen, M. mRNA transfection-induced activation of primary human monocytes and macrophages: Dependence on carrier system and nucleotide modification. *Scientific Reports* **10.** ISSN: 2045-2322. doi:10.1038/s41598-020-60506-4. http://dx.doi.org/10.1038/s41598-020-60506-4 (Mar. 2020).

179. Zhang, X. & Mosser, D. M. Macrophage activation by endogenous danger signals. *The Journal of Pathology* **214,** 161–178. ISSN: 1096-9896. doi:10.1002/path.2284. http://dx.doi.org/10.1002/path.2284 (Dec. 2007).

180. Karikó, K., Buckstein, M., Ni, H. & Weissman, D. Suppression of RNA recognition by Toll-like receptors: The impact of nucleoside modification and the evolutionary origin of RNA. *Immunity* **23,** 165–175. ISSN: 1074-7613. doi:10.1016/j.immuni.2005.06.008. http://dx.doi.org/10.1016/j.immuni.2005.06.008 (Aug. 2005).

181. Chong, Z. X., Yeap, S. K. & Ho, W. Y. Transfection types, methods and strategies: A technical review. *PeerJ* **9,** e11165. ISSN: 2167-8359. doi:10.7717/peerj.11165. http://dx.doi.org/10.7717/peerj.11165 (Apr. 2021).

182. Fus-Kujawa, A. *et al.* An overview of methods and tools for transfection of eukaryotic cells in vitro. *Frontiers in Bioengineering and Biotechnology* **9.** ISSN: 2296-4185. doi:10.3389/fbioe.2021.701031. http://dx.doi.org/10.3389/fbioe.2021.701031 (July 2021).

183. Xu, Y.-R. & Lei, C.-Q. TAK1-TABs complex: A central signalosome in inflammatory responses. *Frontiers in Immunology* **11.** ISSN: 1664-3224. doi:10.3389/fimmu.2020.608976. http://dx.doi.org/10.3389/fimmu.2020.608976 (Jan. 2021).

184. Kawasaki, T. & Kawai, T. Toll-like receptor signaling pathways. *Frontiers in Immunology* **5.** ISSN: 1664-3224. doi:10.3389/fimmu.2014.00461. http://dx.doi.org/10.3389/fimmu.2014.00461 (Sept. 2014).

185. Kuzmich, N. *et al.* TLR4 signaling pathway modulators as potential therapeutics in inflammation and sepsis. *Vaccines* **5,** 34. ISSN: 2076-393X. doi:10.3390/vaccines5040034. http://dx.doi.org/10.3390/vaccines5040034 (Oct. 2017).

186. Bertani, B. & Ruiz, N. Function and biogenesis of lipopolysaccharides. *EcoSal Plus* **8** (ed Slauch, J. M.) ISSN: 2324-6200. doi:10.1128/ecosalplus.esp-0001-2018. http://dx.doi.org/10.1128/ecosalplus.ESP-0001-2018 (Dec. 2018).

187. Oeckinghaus, A., Hayden, M. S. & Ghosh, S. Crosstalk in NF-κB signaling pathways. *Nature Immunology* **12,** 695–708. ISSN: 1529-2916. doi:10.1038/ni.2065. http://dx.doi.org/10.1038/ni.2065 (July 2011).

188. Wu, X. & Schauss, A. G. Mitigation of inflammation with foods. *Journal of Agricultural and Food Chemistry* **60,** 6703–6717. ISSN: 1520-5118. doi:10.1021/jf3007008. http://dx.doi.org/10.1021/jf3007008 (Apr. 2012).

189. Alto, N. M. & Orth, K. Subversion of cell signaling by pathogens. *Cold Spring Harbor Perspectives in Biology* **4,** a006114–a006114. ISSN: 1943-0264. doi:10.1101/cshperspect.a006114. http://dx.doi.org/10.1101/cshperspect.a006114 (Sept. 2012).

190. Plotnikov, A., Zehorai, E., Procaccia, S. & Seger, R. The MAPK cascades: Signaling components, nuclear roles and mechanisms of nuclear translocation. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research* **1813,** 1619–1633. ISSN: 0167-4889. doi:10.1016/j.bbamcr.2010.12.012. http://dx.doi.org/10.1016/j.bbamcr.2010.12.012 (Sept. 2011).

191. Braicu *et al.* A comprehensive review on MAPK: A promising therapeutic target in cancer. *Cancers* **11,** 1618. ISSN: 2072-6694. doi:10.3390/cancers11101618. http://dx.doi.org/10.3390/cancers11101618 (Oct. 2019).

192. Frick, C. L., Yarka, C., Nunns, H. & Goentoro, L. Sensing relative signal in the Tgf-β/Smad pathway. *Proceedings of the National Academy of Sciences* **114.** ISSN: 1091-6490. doi:10.1073/pnas.1611428114. http://dx.doi.org/10.1073/pnas.1611428114 (Mar. 2017).

193. Muralidharan, S. *et al.* Immunolipidomics reveals a globoside network during the resolution of pro-inflammatory response in human macrophages. *Frontiers in Immunology* **13.** ISSN: 1664-3224. doi:10.3389/fimmu.2022.926220. http://dx.doi.org/10.3389/fimmu.2022.926220 (June 2022).

194. Clausen, B. H. *et al.* Interleukin-1beta and tumor necrosis factor-alpha are expressed by different subsets of microglia and macrophages after ischemic stroke in mice. *Journal of Neuroinflammation* **5.** ISSN: 1742-2094. doi:10.1186/1742-2094-5-46. http://dx.doi.org/10.1186/1742-2094-5-46 (Oct. 2008).

195. Clausen, B. H. *et al.* Characterization of the TNF and IL-1 systems in human brain and blood after ischemic stroke. *Acta Neuropathologica Communications* **8.** ISSN: 2051-5960. doi:10.1186/s40478-020-00957-y. http://dx.doi.org/10.1186/s40478-020-00957-y (June 2020).

196. Panzer, S., Madden, M. & Matsuki, K. Interaction of IL-1β, IL-6 and tumour necrosis factor-alpha (TNF-α) in human T cells activated by murine antigens. *Clinical and Experimental Immunology* **93,** 471–478. ISSN: 1365-2249. doi:10.1111/j.1365-2249.1993.tb08203.x. http://dx.doi.org/10.1111/j.1365-2249.1993.tb08203.x (Sept. 1993).

197. Saperstein, S., Chen, L., Oakes, D., Pryhuber, G. & Finkelstein, J. IL-1β augments TNF-α–mediated inflammatory responses from lung epithelial cells. *Journal of Interferon and Cytokine Research* **29,** 273–284. ISSN: 1557-7465. doi:10.1089/jir.2008.0076. http://dx.doi.org/10.1089/jir.2008.0076 (May 2009).

198. Xue, Q. *et al.* Analysis of single-cell cytokine secretion reveals a role for paracrine signaling in coordinating macrophage responses to TLR4 stimulation. *Science Signaling* **8.** ISSN: 1937-9145. doi:10.1126/scisignal.aaa2155. http://dx.doi.org/10.1126/scisignal.aaa2155 (June 2015).

199. Adamik, J. *et al.* Distinct mechanisms for induction and tolerance regulate the immediate early genes encoding interleukin 1β and tumor necrosis factor α. *PLoS ONE* **8** (ed Yan, C.) e70622. ISSN: 1932-6203. doi:10.1371/journal.pone.0070622. http://dx.doi.org/10.1371/journal.pone.0070622 (Aug. 2013).

200. Francis, K. & Palsson, B. O. Effective intercellular communication distances are determined by the relative time constants for cyto/chemokine secretion and diffusion. *Proceedings of the National Academy of Sciences* **94,** 12258–12262. ISSN: 1091-6490. doi:10.1073/pnas.94.23.12258. http://dx.doi.org/10.1073/pnas.94.23.12258 (Nov. 1997).

201. Handly, L. N., Pilko, A. & Wollman, R. Paracrine communication maximizes cellular response fidelity in wound signaling. *eLife* **4.** ISSN: 2050-084X. doi:10.7554/elife.09652. http://dx.doi.org/10.7554/eLife.09652 (Oct. 2015).

202. Muldoon, J. J., Chuang, Y., Bagheri, N. & Leonard, J. N. Macrophages employ quorum licensing to regulate collective activation. *Nature Communications* **11.** ISSN: 2041-1723. doi:10.1038/s41467-020-14547-y. http://dx.doi.org/10.1038/s41467-020-14547-y (Feb. 2020).

203. Postat, J. & Bousso, P. Quorum sensing by monocyte-derived populations. *Frontiers in Immunology* **10.** ISSN: 1664-3224. doi:10.3389/fimmu.2019.02140. http://dx.doi.org/10.3389/fimmu.2019.02140 (Sept. 2019).

204. Schwartz, M. S. *et al.* Caliban: Accurate cell tracking and lineage construction in live-cell imaging experiments with deep learning. doi:10.1101/803205. http://dx.doi.org/10.1101/803205 (Oct. 2019).

205. De Lorenzi, R., Gareus, R., Fengler, S. & Pasparakis, M. GFP-p65 knock-in mice as a tool to study NF-κB dynamics in vivo. *genesis* **47,** 323–329. ISSN: 1526-968X. doi:10.1002/dvg.20468. http://dx.doi.org/10.1002/dvg.20468 (Mar. 2009).

206. Morales, A. J. *et al.* A type I IFN-dependent DNA damage response regulates the genetic program and inflammasome activation in macrophages. *eLife* **6.**

ISSN: 2050-084X. doi:10.7554/elife.24655. http://dx.doi.org/10.7554/eLife.24655 (Mar. 2017).

207. Iwanaszko, M., Brasier, A. R. & Kimmel, M. The dependence of expression of NF-κB-dependent genes: Statistics and evolutionary conservation of control sequences in the promoter and in the 3' UTR. *BMC Genomics* **13,** 182. ISSN: 1471-2164. doi:10.1186/1471-2164-13-182. http://dx.doi.org/10.1186/1471-2164-13-182 (2012).

208. Zhao, M. *et al.* Transcriptional outcomes and kinetic patterning of gene expression in response to NF-κB activation. *PLOS Biology* **16** (ed Kadonaga, J.) e2006347. ISSN: 1545-7885. doi:10.1371/journal.pbio.2006347. http://dx.doi.org/10.1371/journal.pbio.2006347 (Sept. 2018).

209. Kraskov, A., Stögbauer, H. & Grassberger, P. Estimating mutual information. *Physical Review E* **69.** ISSN: 1550-2376. doi:10.1103/physreve.69.066138. http://dx.doi.org/10.1103/PhysRevE.69.066138 (June 2004).

210. Ross, B. C. Mutual information between discrete and continuous data sets. *PLoS ONE* **9** (ed Marinazzo, D.) e87357. ISSN: 1932-6203. doi:10.1371/journal.pone.0087357. http://dx.doi.org/10.1371/journal.pone.0087357 (Feb. 2014).

211. Peters, A. The energy request of inflammation. *Endocrinology* **147,** 4550–4552. ISSN: 1945-7170. doi:10.1210/en.2006-0815. http://dx.doi.org/10.1210/en.2006-0815 (Oct. 2006).

212. Lacourt, T. E., Vichaya, E. G., Chiu, G. S., Dantzer, R. & Heijnen, C. J. The high costs of low-grade inflammation: Persistent fatigue as a consequence of reduced cellular-energy availability and non-adaptive energy expenditure. *Frontiers in Behavioral Neuroscience* **12.** ISSN: 1662-5153. doi:10.3389/fnbeh.2018.00078. http://dx.doi.org/10.3389/fnbeh.2018.00078 (Apr. 2018).

213. Wang, H. & Ye, J. Regulation of energy balance by inflammation: Common theme in physiology and pathology. *Reviews in Endocrine and Metabolic Disorders* **16,** 47–54. ISSN: 1573-2606. doi:10.1007/s11154-014-9306-8. http://dx.doi.org/10.1007/s11154-014-9306-8 (Dec. 2014).

214. Meli, V. S. *et al.* YAP-mediated mechanotransduction tunes the macrophage inflammatory response. *Science Advances* **6.** ISSN: 2375-2548. doi:10.1126/sciadv.abb8471. http://dx.doi.org/10.1126/sciadv.abb8471 (Dec. 2020).

215. Meli, V. S., Veerasubramanian, P. K., Downing, T. L., Wang, W. & Liu, W. F. Mechanosensation to inflammation: Roles for YAP/TAZ in innate immune cells. *Science Signaling* **16.** ISSN: 1937-9145. doi:10.1126/scisignal.adc9656. http://dx.doi.org/10.1126/scisignal.adc9656 (May 2023).

216. Solis, A. G. *et al.* Mechanosensation of cyclical force by PIEZO1 is essential for innate immunity. *Nature* **573,** 69–74. ISSN: 1476-4687. doi:10.1038/s41586-019-1485-8. http://dx.doi.org/10.1038/s41586-019-1485-8 (Aug. 2019).

217. Atcha, H. *et al.* Mechanically activated ion channel Piezo1 modulates macrophage polarization and stiffness sensing. *Nature Communications* **12.** ISSN: 2041-1723. doi:10.1038/s41467-021-23482-5. http://dx.doi.org/10.1038/s41467-021-23482-5 (May 2021).

218. Corson, L. B., Yamanaka, Y., Lai, K.-M. V. & Rossant, J. Spatial and temporal patterns of ERK signaling during mouse embryogenesis. *Development* **130,** 4527–4537. ISSN: 0950-1991. doi:10.1242/dev.00669. http://dx.doi.org/10.1242/dev.00669 (Oct. 2003).

219. Aoki, K. *et al.* Propagating wave of ERK activation orients collective cell migration. *Developmental Cell* **43,** 305–317.e5. ISSN: 1534-5807. doi:10.1016/j.devcel.2017.10.016. http://dx.doi.org/10.1016/j.devcel.2017.10.016 (Nov. 2017).

220. Watson, U., Jain, R., Asthana, S. & Saini, D. K. in *G Protein-Coupled Receptors: Emerging Paradigms in Activation, Signaling and Regulation Part A* 111–140 (Elsevier, 2018). doi:10.1016/bs.ircmb.2018.02.004. http://dx.doi.org/10.1016/bs.ircmb.2018.02.004.