Systems pharmacology of cell-signaling networks in human disease

---

A Dissertation

Presented to

The faculty of the School of Engineering and Applied Science

University of Virginia

In Partial Fulfillment

Of the requirements for the Degree

Doctor of Philosophy

By

Karin Jeanne Jensen

December 2013

APPROVAL SHEET
THE DISSERTATION IS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY IN BIOMEDICAL ENGINEERING

Karin J. Jensen

This dissertation has been read and approved by the examining committee:

Dr. Kevin A. Janes, Dissertation advisor

Dr. Jeffrey J. Saucerman, Committee chair

Dr. Michael J. Weber, Minor mentor

Dr. Jason A. Papin

Dr. Kimberly A. Kelly

Accepted for the School of Engineering and Applied Science

Dean, School of Engineering and Applied Science

December 2013

Systems pharmacology of cell-signaling networks in human disease

ABSTRACT

Cell-signaling networks are fascinatingly complex communication systems that integrate a diverse array of extracellular cues and appropriately modulate cellular responses. Deregulation of these networks can result in various human diseases, from cancer to Alzheimer's disease. A systems-level understanding of cell-signaling networks will improve our understanding of normal physiology and disease states, and most importantly, improve our ability to pharmacologically treat disease.

In this dissertation, we develop and implement experimental and computational approaches to study human cell-signaling networks. This work is comprised of four aims: 1) to develop an experimental platform to measure receptor expression, 2) to study the effect of local pathway topology on directed perturbations, 3) to apply data-driven modeling to a host-pathogen dataset to discover signaling subnetworks, and 4) to develop a high-throughput multiplex experimental assay to measure dynamic information flow in cell-signaling networks.

Many signal transduction cascades are initiated by receptors that sense extracellular stimuli and catalyze downstream signaling events. Receptor expression is therefore critical in defining cellular responsiveness. We began by developing a high-throughput qRT-PCR-based platform to measure expression of 194 signaling receptors. We then leveraged the high-throughput capabilities of the assay to probe receptor expression patterns in 40 cancer cell lines. Downstream of receptors, pathway structure is critical for processing and propagating cellular information. We next built 95 computational models to study the effect of local pathway connectivity on drug and RNAi targeting. The models revealed an important role for pathway structure in determining levels of pathway inhibition. Pathways are further interconnected with other pathways to integrate numerous cues and regulate cellular processes. We next applied data-driven modeling to study cell signaling during Coxsackievirus B3 infection of cardiomyocytes. The model implicated the ERK1/2, ERK5, and p38 pathways in regulating apoptosis and necrosis during infection. Follow-up experiments revealed dual ERK1/2 and p38 inhibition dramatically reduces cell death and virus release. In order to measure dynamic information flow throughput the network in multiple pathways simultaneously, we describe the design of a high-throughput kinase activity assay. Together these aims provide a framework to study cell-signaling networks and their pharmacological inhibition at a systems level.

ACKNOWLEDGEMENTS

This work would not have been possible without the continued support from my mentors, colleagues, friends, and family.

I thank my graduate advisor Kevin Janes for his constant and unwavering support and enthusiasm during five years of triumphs and failures in the lab.

The Janes, McManus, Weber, and Gioeli lab members provided tremendous support that was indispensable in completing this work. Undergraduate researchers Jaime Hatch and Byong Kang were instrumental in the development of the receptome profiling assay (Chapter 2) by validating all primers and profiling cell lines. Byong also performed all the bioinformatics analysis and developed software to process profiling data. Undergraduate researchers Anjun Bose and Benjamin Kuhn provided numerous contributions to several projects. Fellow graduate student Christian Moyer was a major contributor to the pathway modeling work (Chapter 3).

Finally, I thank my family and friends for their support and motivation. I especially thank my parents for their longtime support of my educational goals and for their constant encouragement. I thank my husband Paul for being my daily inspiration and supporting my scientific goals.

TABLE OF CONTENTS

LIST OF FIGURES AND TABLES

CHAPTER 1

Introduction

1.1. Complexity of cell-signaling as a challenge and opportunity

1.1.1. Introduction

Cell-signaling networks are interconnected webs of proteins that sense a cell's extracellular environment and govern cellular responses. The ability of the cell-signaling network to correctly interpret and appropriately respond to extracellular signals underlies important biological processes at the molecular, cell, tissue, and organ levels. The deregulation of this communication network is the root of diverse human diseases including cancer and Alzheimer's disease (*13, 14*). Drug discovery to combat these diseases is a challenge that spans multiple levels of complexity and scale, requiring the selection of a compound to modulate a cellular target that will translate to the amelioration of a disease across a patient population. Despite the wealth of possible targets for drug therapy, identification of new and effective drug targets remains challenging. Individual proteins of interest have been studied intensively, but understanding of signaling at the network level remains poor. At present, the PubMed database contains 67,500 publications relevant to the protein p53, yet new complexities of p53 continue to be described (*15, 16*). Poor understanding of proteins in the context of the greater network have plagued targeted drug therapies with problems of acquired drug resistance, unwanted side effects of therapies, and heterogeneous responses in patient populations (*17, 18*). An improved understanding of the roles of individual targets at the

network level will guide our selection of better therapeutic targets and combinations of targets to overcome these challenges and deliver much-needed therapies for patients.

1.1.2. Layers of biological complexity in the cell-signaling network

Cells sense and respond to an incredible array of extracellular signals: diffusible proteins, drugs, metabolites, mechanical forces, radiation, osmolarity, and many others. These diverse arrays of signals vary with time and are presented in a combinatorial manner (*19*). Cell-signaling networks have evolved to appropriately respond to these complex inputs and regulate important cellular functions in response to environmental changes, such as immunity and tissue repair (*20*). The response of a cell depends on the combination of extracellular signals received, the wiring of the signaling network, and the dynamics of signals.

It is estimated that 12.2% of human genes (of the 58.3% categorized) are dedicated to signal transduction (*21*). The largest category of signaling components is the receptors (Chapter 2). In the human genome there are 1,543 receptors that sense extracellular cues and initiate downstream signaling. The presence or absence of these receptors dictates a cell's responsiveness to particular cues. The second major group of signal transduction genes is the kinases (Chapter 5). Downstream of signaling receptors there are 518 protein kinases that catalyze signaling events (*21, 22*). Information is transmitted through the signaling network through the connections between kinases and other components in the network to form signaling pathways. Pathways are further embedded within larger subnetworks that connect multiple pathways to each other, creating a complex communication network for information processing and flow in the

cell. In order to study these complexities of the signaling network, it will be important to exploit methods that measure the components, their interactions, and the dynamics of information flow.

1.1.3. Challenges in identifying critical signaling nodes for drug development

The cell-signaling network is a promising source to look for new therapeutic targets, not only due to the vast number of potential targets and regulatory mechanisms, but also because most existing and emerging targeted molecular therapies are directed at components of the signaling network (or their mutated counterparts), such as trastuzamab for breast cancer, vemurafenib for melanoma, and imatinib for chronic myelogenous leukemia (*23-25*). These targeted therapies have shown milestone successes in treating some human diseases, but there is a desperate need to identify new targets for development. Developing drug therapies for patients is an incredibly time-consuming and costly process, averaging 12 years and $800 million per drug, with some estimates topping $2 billion (*26, 27*). Despite these enormous financial and time commitments, therapies have faced an attrition rate of about 90% from phase I to the clinic, with the majority of these failures due to lack of efficacy and safety (*28*). Perhaps the most critical step in drug development is deciding which cellular targets we should investigate for development in the first place.

Due to the enormous upfront costs for development, inhibitors continue to be developed against the same known targets, leaving many candidate enzymes unexplored (*29*). The discovery of novel clinical targets would provide new hope for patients and

energize the research community as recently observed with vemurafenib in metastatic melanoma (*30*). Historically, the focus of many drug discovery efforts was screening for drug-target interactions and studying them in isolation. By shifting this focus towards network level studies and models, we will choose better targets and preemptively thwart problems of efficacy and safety. Further, the ability to predict effective and worthwhile drug targets earlier in the process will dramatically decrease the time and financial burdens of drug discovery and development. By studying cell signaling at the network level we will identify novel and more efficacious targets to enter the drug development pipeline.

1.2. Challenges in systems-level studies of the cell-signaling network

Systems biology is an approach to studying the complex interactions in a biological system in a holistic, integrative manner to discover emergent properties of the system as a whole (*31*). Systems pharmacology applies systems biology approaches to study the effect of drugs on the network level (*32*). The systems biology approach is fueled by our ability to collect comprehensive data about biological systems and to effectively analyze and integrate these datasets. Significant advances in technology have prompted the "omics" study of biological systems, from genomics to lipidomics to metabolomics (*33-35*). The utility of these data collection efforts hinges on the development of advanced computational techniques to understand these datasets. For this reason, the combination of experimental and computational methods is necessary to study cell signaling on the network level.

1.2.1. Generation of comprehensive experimental datasets

Advances in technology have dramatically increased our ability to make measurements on the molecular and cellular level. Perhaps most notably, DNA sequencing has reduced the cost and increased the speed of generating sequencing data by four orders of magnitude over a seven year period, from a price tag of $1,000 per megabase to $0.10 per megabase (*36*). In addition to increasing assay throughput and decreasing cost, the development of assays that can measure new parameters will enrich the study of cell-signaling networks. As our appreciation for the many layers of complexity of cellular controls increases, so does our need for technology to measure controls in these layers. For example, in recent work by Bose et al, a high-throughput assay to measure phosphatase activity against specific phosphosubstrates was described. Early studies with the new assay revealed complex and dynamic regulation of phosphatases not previously appreciated (*37*). In this thesis, we will discuss the development of two experimental platforms for high-throughput data collection to probe cell-signaling networks (Chapters 2 and 5).

1.2.2. Interpretation and integration of large datasets

While many systems biology analyses of cell-signaling networks are inherently dependent on large-scale data collection efforts, datasets alone will not guide identification of new therapies. Methods to interpret and analyze datasets are as important as the datasets themselves. As has been argued recently, more data will not always lead to clarity, particularly if the type of data being collected does not address the right problem (*38*). Further, associations or connections found in large-scale data

collection efforts may not be relevant or meaningful (*39*). For these reasons, the co-development of analysis tools is necessary to make the best use of our data collection efforts and extract testable biological hypotheses. Large-scale data collection efforts coupled with analysis methods have made significant contributions to our understanding of signal transduction networks (*19, 40, 41*). It is also imperative to understand the particular caveats and limitations of different experimental techniques and define the contexts where the results may not be as accurate (see examples in Chapter 2.3-2.5). In this thesis we will discuss computational methods to analyze and interpret large datasets (Chapters 3-4).

In addition to producing comprehensive datasets and interpreting their biological meaning, understanding the layers of complexity will be dependent on our ability to integrate these disparate types of data. Computational models of cell-signaling systems have demonstrated how data integration for studying signaling at the network level can lead to new biological insights (*40, 42*). In Chapter 5 we describe data-driven modeling techniques and apply them to a dataset with multiple types of experimental measurements.

1.3. Summary

The cell-signaling network is enormously complex, demonstrated by a staggering number of components and levels of control mechanisms. The complexity of the cell-signaling network is apparent by the difficulty to develop effective therapeutics targeting signaling components. Despite this challenge, this complexity also offers a wealth of potential targets for therapy. Importantly, in order to identify the most efficacious

targets, we need to understand their roles in the larger network. Towards this goal, we will need experimental methods to measure the components and controls as well as computational tools to understand these measurements. Studying signaling complexities at the local and network levels will aid in our understanding of diseases and address unmet medical needs.

CHAPTER 2

Defining sensing and signaling capabilities of human cells by receptome profiling

2.1. Introduction

2.1.1. Signaling receptors sense the extracellular environment and initiate signal transduction in the network

Transmembrane signaling receptors are the genetically encoded sensors of the extracellular environment (*43*). A cell can display millions of receptor copies on its cell surface (*44*), yet intracellular responses can be triggered when just a few dozen receptors bind their cognate ligands (*45, 46*). Importantly, complete absence of a signaling receptor renders a cell unresponsive to its ligands (*47, 48*), meaning that the cell is "blind" to that class of environmental inputs. Transmembrane proteins are enriched in the low-abundance fractions of the transcriptome and proteome (*49*). These low-abundance transcripts and the surface proteins that they encode are also effective indicators of cell lineage (*6, 49*). The qualitative presence or absence of signaling receptors thus defines a critical facet of a cell's identity and its response capabilities.

2.1.2. Limitations of current methods to measure receptor expression

Large profiles of receptor families can be extracted from transcriptome measurements obtained by oligonucleotide microarrays (*50*), but the extracted profiles

are not definitive. Microarrays have a compressed dynamic range and poorer detection sensitivity relative to single-gene methods (*51*), and some probe sets on established platforms are still plagued with cross-hybridization artifacts (*52*). Compared to microarrays, digital transcript counting by RNA sequencing (RNA-seq) is more specific and shows substantially improved dynamic range (*53*) and sensitivity (*54*). However, RNA-seq is methodologically inefficient, because the technique must repeatedly measure high-abundance transcripts to achieve maximal sensitivity toward the low-abundance targets (*55*). There is additional evidence that the rarest transcripts identified by RNA-seq are nonfunctional (*56*), which hinders the ability of RNA-seq to determine whether signaling-competent receptors are truly present or absent in a cell population.

2.2. Design of a high-throughput assay to measure receptor expression

Gene expression measurements from microarrays or RNA-seq are often validated with quantitative reverse transcription-polymerase chain reaction (qRT-PCR) (*51, 56*). Due to its high sensitivity, wide dynamic range, and verifiable specificity, qRT-PCR is routinely viewed as a gold standard for expression studies with individual genes. Inspired by an effort aimed at characterizing the transcriptional profile of a subset of G protein-coupled receptors (*57*), here we developed and validated arrayed qRT-PCR reagents for 194 transmembrane signaling receptors in the human genome. We defined a signaling "receptome" (*58*) that includes all human receptor serine-threonine and tyrosine kinases, all cytokine and chemokine receptors, as well as all receptors of the Toll-like, Frizzled, Notch, and Patched families (Figure 2-1A). These signaling receptors bind a diverse range of macromolecular ligands and show widespread, but selective, tissue

expression. A defined panel enabled in-depth validation of gene-specific reagents that together were readily accommodated in a 96-well format for high-throughput profiling.

We designed qRT-PCR primers for each gene in the receptome and individually optimized the primers so that they produced a consistent amplicon size under the same rapid-cycling conditions. During the initial primer validation, we diagnosed correct amplicons by melt-curve analysis, gel electrophoresis, and (when necessary) sequencing. The validation experiments produced a verified list of gene-specific melting temperatures for direct assessment of receptor transcript presence-absence after each profiling experiment.

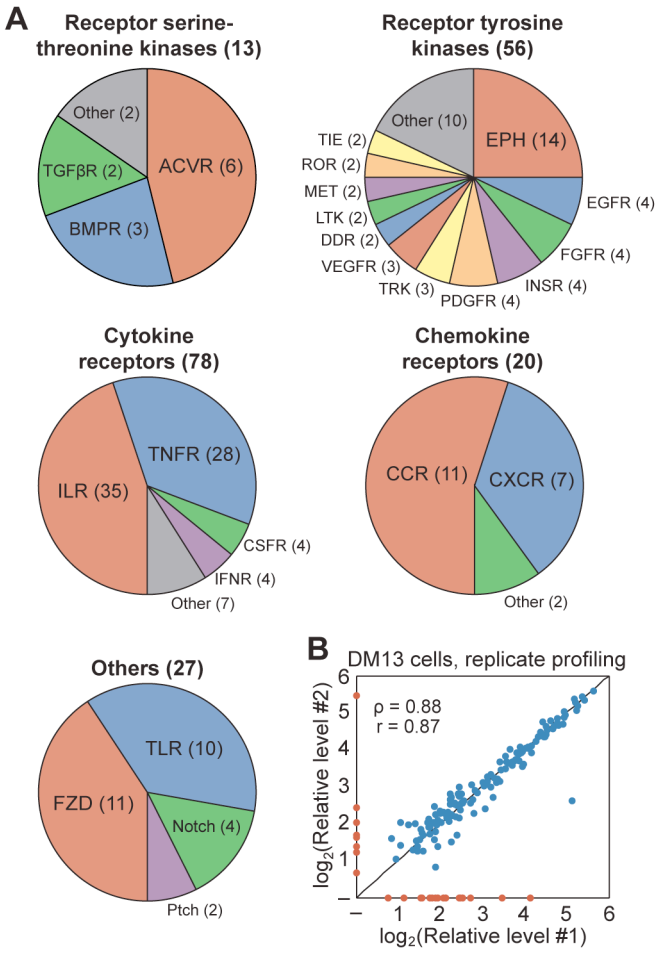


Figure 2-1. High-sensitivity profiling of a human signaling receptome by arrayed qRT-PCR. **(A)** Distribution of signaling receptor families and subfamilies comprising the receptome profiling assay: TGFβR, transforming growth factor-β receptor; BMPR, bone morphogenetic protein receptor; ACVR, activin A receptor; EPH, ephrin receptor; EGFR, epidermal growth factor receptor; FGFR, fibroblast growth factor receptor; INSR, insulin receptor; PDGFR, platelet-derived growth factor receptor; TRK, tropomyosin receptor kinase; VEGFR, vascular endothelial growth factor receptor; DDR, discoidin domain receptor; LTK, leukocyte receptor tyrosine kinase; MET, mesenchymal epithelial transition factor; ROR, retinoic acid receptor-related orphan receptor; TIE, tyrosine kinase with immunoglobulin-like and EGF-like domains; ILR, interleukin receptor; TNFR, tumor necrosis factor receptor; CSFR, colony stimulating factor receptor; IFNR, interferon receptor; CCR, chemokine (C-C motif) receptor; CXCR, chemokine (C-X-C motif) receptor; TLR, toll-like receptor; FZD, frizzled receptor; Ptch, patched. **(B)** Reproducibility of receptome profiling across assay replicates. Receptors detected in at least one assay replicate were plotted as $\log_2$ relative abundance estimated by qRT-PCR cycle threshold assuming 100% amplification efficiency. The complete set of pairwise comparisons is shown in Figure 2-2.

Because qRT-PCR of extremely low abundance targets can be sporadic (*59*), we profiled the receptome of each sample in separate duplicates. Between duplicate qRT-PCR plates, we observed strong pairwise correlations in cycle threshold (C_T) values (median Spearman $\varrho = 0.84$, Pearson $R = 0.78$) (Figure 2-1F, Figure 2-2). This indicated that plate-to-plate amplification efficiencies were comparable and average C_T values could be used as a semi-quantitative $\log_2$ measure of relative transcript abundance across independent qRT-PCR reactions. In addition, receptor transcript status could be qualitatively scored as present or absent based on whether specific amplification was detected in at least one of two replicates or not. Analysis of blank qRT-PCR reactions lacking sample indicated that the leading cause for missed detection in a replicate was competition of the desired amplicon by nonspecific primer-dimer products that arose during the late cycles of amplification. We reduced these artifacts by minimizing the primer concentration while maintaining the amplification efficiency of the desired RT-PCR product. Nonetheless, a few receptor transcripts (*EPHA8*, *ERBB2*, *NRTK1*, *IL2RB*, *IL22RA2*, *TNFRSF10A*, and *TNFRSF25*) were significantly variable ($P < 0.01$, Bonferroni-corrected binomial test) because of primer-dimer competition, reinforcing the need for duplicate measurements across the receptome.

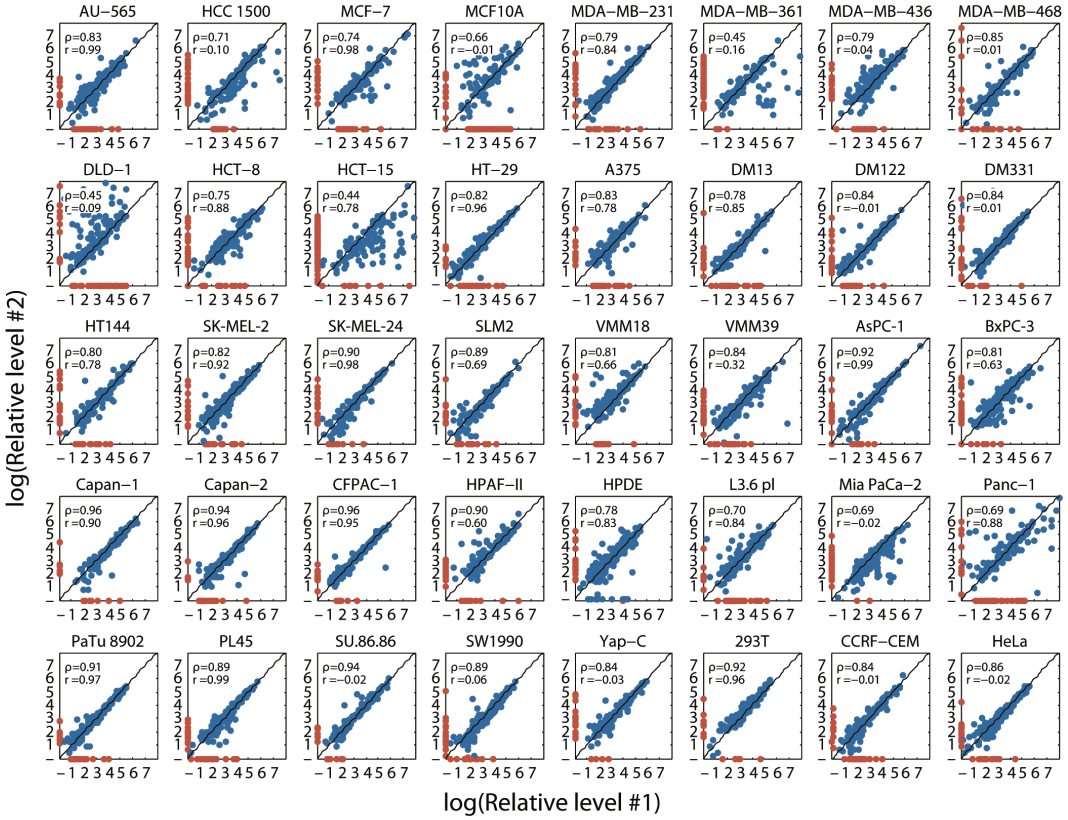


Figure 2-2. Receptome profiling is consistently reproducible across assay duplicates. Receptor transcripts detected in at least one assay replicate (red, one replicate; blue, two replicates) were plotted as $\log_2$ relative abundance estimated by qRT-PCR cycle threshold assuming 100% amplification efficiency. En dash (–) indicates not detected. The Spearman (ρ) and Pearson (r) correlation coefficients are shown inside each graph.

2.3. qRT-PCR receptome profiling is accurate, precise, and more sensitive than oligonucleotide microarrays

To investigate the sensitivity of qRT-PCR receptome profiling, we selected HT-29 colon adenocarcinoma cells treated with interferon-γ (IFN-γ), which previously served as the base condition for a large signaling dataset (*19*, *41*, *60*). We assessed transcript presence or absence in HT-29 cells exposed to IFN-γ by receptome profiling and by transcriptional profiling with conventional oligonucleotide microarrays. Compared to the

present-absent calls of the commercial microarray analysis software, we found that receptome profiling was more sensitive (Figure 2-3A). Only eight receptor transcripts were called present by microarray and absent by receptome profiling, and literature suggests that several of these receptors are false positives on the microarray. For example, both *ERBB4* and *EPOR* were called present by microarray, but *ERBB4* mRNA (*61*), ERBB4 protein (*62*), and EPOR protein and receptor signaling (*63*) are undetectable in HT-29 cells. By contrast, qRT-PCR receptome profiling detected 54 additional receptor transcripts that were called absent by microarrays. Many of these additional receptors have been detected or functionally validated in HT-29 cells previously (Table 2-1). This suggested that

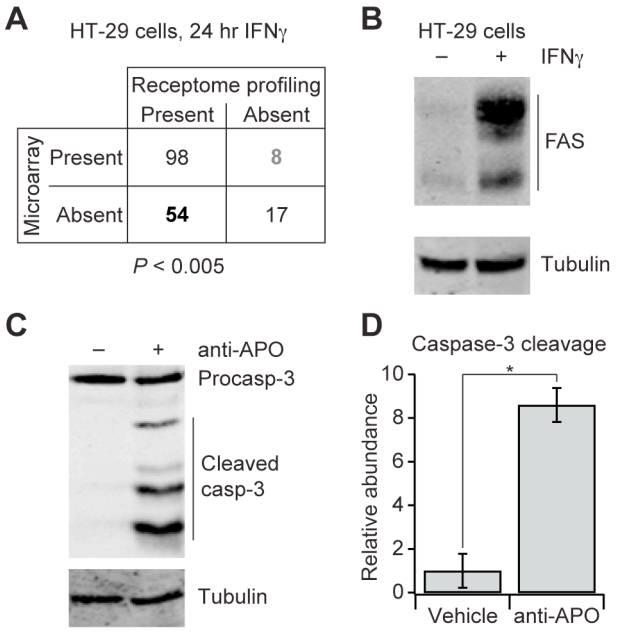


Figure 2-3. qRT-PCR receptome profiling is significantly more sensitive for detecting receptor transcripts than conventional oligonucleotide microarrays. (A) Present-absent calls for 177 receptor transcripts monitored on Affymetrix U133A microarrays were compared to receptome-profiling results for HT-29 cells sensitized with 200 U ml^{-1} IFNγ for 24 hr. Statistical significance was assessed by Fisher's exact test. (B) Detection of FAS protein in HT-29 cells with or without IFNγ sensitization. (C) Caspase-3 cleavage in IFNγ-sensitized HT-29 cells after FAS crosslinking with 1 μg ml^{-1} anti-APO for 24 hr. (D) Replicated densitometry of anti-APO-induced caspase-3 cleavage in HT-29 cells. Data are shown as the mean ± s.e.m. of three independent samples, and asterisk indicates statistical significance ($P < 0.05$) by Welch's one-sided *t* test. For (B) and (C), cells were immunoblotted for the indicated proteins with tubulin used as a loading control. All immunoblots are representative of at least three independent experiments.

conventional microarray present-absent calls largely reflect differences in detection rather than true presence or absence of a transcript.

We further evaluated the specificity of receptome profiling by analyzing receptor presence or absence through a panel of independent measurements. We selected the death receptor *FAS* as a gene with lineage-specific expression (*64*). *FAS* mRNA was predicted to be absent in IFN-γ-treated HT-29 cells by microarray but present by receptome profiling. We examined FAS abundance by immunoblotting and found that it was present and its abundance was increased by IFN-γ (Figure 2-3B), consistent with reports in other cell types (*64*). Accordingly, stimulation of IFN-γ-treated HT-29 cells with a FAS crosslinking antibody (anti-APO) resulted in a strong apoptotic response as indicated by caspase-3 cleavage (Figure 2-3C and D). Thus, qRT-PCR receptome profiling uncovered signaling capabilities missed by conventional oligonucleotide microarray methods.

Table 2-1. Literature support for present calls in HT-29 cells.

Receptor gene	Literature support
PTK7	Transmembrane and soluble form detected by immunoblotting (*65*)
TGFBR1	Detected by qRT-PCR (*66*)
CXCR1	Detected by qRT-PCR and flow cytometry, and promotes cell migration in response to CXCL8 (*67*)
CXCR2	Detected by qRT-PCR and immunohistochemistry (*68*)
CXCR4	Detected by qRT-PCR and flow cytometry, and promotes Ca^{2+} mobilization, increased ICAM1 levels, and IL-8 and GROα secretion in response to SDF-1α (*69, 70*)
CCR6	Detected by qRT-PCR (*70*)
CCR8	Detected by qRT-PCR (*70*)
IL15RA	Detected by qRT-PCR (*71*)
PRLR	Detected by Southern blotting (*72*)
OSMR	Detected by qRT-PCR (*73*)
TLR2	Detected by qRT-PCR (*74, 75*)
TLR6	Detected by qRT-PCR (*74*)
TLR7	Detected by qRT-PCR (*75*)
TLR8	Detected by qRT-PCR (*75*)
FAS	Detected by flow cytometry (*76*)
TNFRSF1A	Detected by qRT-PCR (*77*)

To assess the accuracy of absent calls, we performed reciprocal experiments with two breast epithelial lines, MDA-MB-436 and MCF10A, and tested for the expression of *CSF1R*, encoding the cytokine receptor for macrophage colony-stimulating factor (MCSF). qRT-PCR receptome profiling predicted that *CSF1R* transcripts were absent in MDA-MB-436 cells but present in MCF10A cells, whereas microarray data that did not detect *CSF1R* in either cell line (*78*). Using an antibody that recognizes CSF1R, we immunoblotted MCF10A cell lysates and detected immunoreactive bands at the predicted molecular weight of CSF1R, which were absent in MDA-MB-436 cell lysates (Figure 2-4A). We stimulated both cell lines with MCSF and monitored extracellular signal-regulated kinase 1 and 2 (ERK1/2) phosphorylation as a downstream signaling readout. Phosphorylated ERK1/2 (pERK1/2) immunoreactivity increased significantly at 15 min after MCSF stimulation in MCF10A cells (Figure 2-4A, B, and C). Conversely, no increases in pERK1/2 were observed in MDA-MB-436 cells at any time after MCSF treatment (Figure 2-4D). The lack of pERK1/2 signaling in MDA-MB-436 cells was not due to a general defect in upstream kinases, because we observed robust ERK1/2 phosphorylation upon epidermal growth factor (EGF) stimulation (Figure 2-4E). These data indicated that MDA-MB-436 cells lack *CSF1R* transcripts, validating the accuracy of the absent calls made by qRT-PCR receptome profiling.

2.4. qRT-PCR receptome profiling is more sensitive than exon arrays

Conventional oligonucleotide microarrays are heavily 3' biased and thus lack the probe density of newer arrays that target all known exons (*79*). Bioinformatic comparisons between exon-targeted and 3'-biased arrays have suggested that exon arrays are more sensitive and specific for detecting expressed transcripts than 3' arrays (*80*).

This raised the possibility that exon-array data would compare more favorably with qRT-PCR receptome profiling for predicting receptor presence or absence.

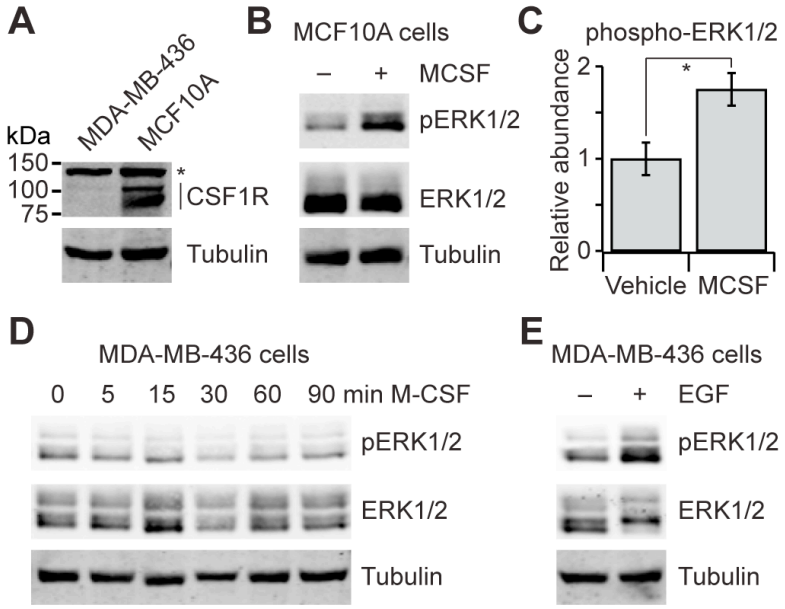


Figure 2-4. qRT-PCR receptome profiling accurately distinguishes receptor absence. (A) Detection of CSF1R protein in MCF10A cells but not in MDA-MB-436 cells. Asterisk marks a nonspecific band. (B) ERK1/2 phosphorylation in MCF10A cells following treatment with 100 ng ml^{-1} MCSF for 15 min. (C) Replicated densitometry of MCSF-induced ERK1/2 phosphorylation in MCF10A cells. Data are shown as the mean ± s.e.m. of three independent samples, and asterisk indicates statistical significance ($P < 0.05$) by Welch's one-sided t test. (D) ERK1/2 phosphorylation in MDA-MB-436 cells following treatment with 100 ng ml^{-1} MCSF for 15 min. **(E)** ERK1/2 phosphorylation in MDA-MB-436 cells following treatment with 100 ng ml^{-1} EGF for 5 min. For (A), (B), (D) and (E), cells were immunoblotted for the indicated proteins with tubulin used as a loading control. All immunoblots are representative of at least three independent experiments.

To make the direct comparison, we prepared total RNA from IFN-γ-treated HT-29 cells, MCF10A cells, and MDA-MB-436 cells and hybridized these samples to Human Exon 1.0 ST arrays. Exon arrays do not provide a discrete present-absent call, so we analyzed the receiver operating characteristics (ROC) of the background-corrected expression index (*81*) for each transcript with respect to the corresponding present-absent

call made by qRT-PCR receptome profiling (Figure 2-5A). At a false-positive rate of 10%, we found that exon arrays achieved a true-positive rate of 40–55% for signaling receptor transcripts, consistent with earlier transcriptome-wide analyses (47% true-positive rate relative to serial analysis of gene expression) (*80*). *FAS* transcripts were readily detected in IFN-γ-treated HT-29 cells below the 10% false-positive rate (Figure 2-5A), illustrating that exon arrays are more sensitive than 3' arrays for certain targets.

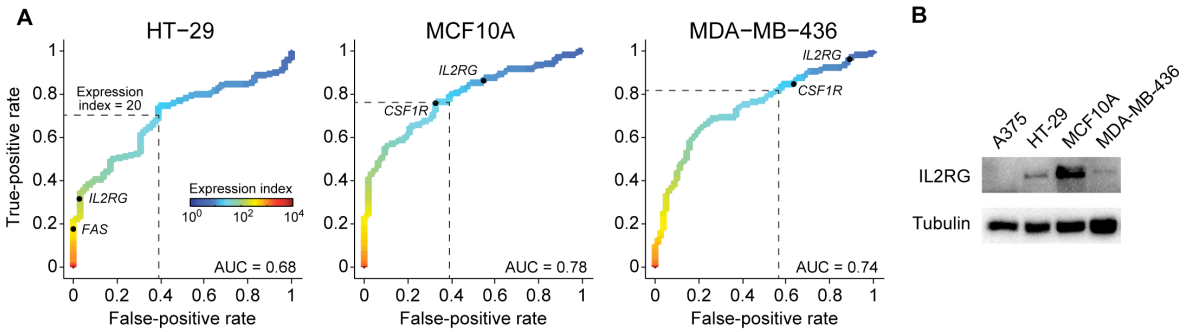


Figure 2-5. qRT-PCR receptome profiling is more sensitive for detecting receptor transcripts than exon arrays. (A to C) Receiver operating characteristic (ROC) curves relating exon array expression index to qRT-PCR present-absent calls for (A) IFNγ-sensitized HT-29 cells, (B) MCF10A cells, and (C) MDA-MB-436 cells. The dashed line indicates a representative expression index that properly distinguishes *CSF1R* expression in MCF10A and MDA-MB-436 cells. The area under the ROC curve (AUC) indicate the overall quality of the present-absent classification based on exon array data, with AUC = 1 indicating perfect classification and AUC = 0.5 indicating random guessing. (D) Detection of IL2RG protein in HT-29, MCF10A, and MDA-MB-436 cells but not in A375 cells. Cells were immunoblotted for IL2RG with tubulin used as a loading control. Immunoblots are representative of three independent experiments.

For *CSF1R*, however, we found that a false-positive rate of 40–60% must be tolerated to distinguish MCF10A and MDA-MB-436 cells properly (Figure 2-5B and C). At this relaxed expression threshold (expression index = 20), the gamma subunit of the interleukin-2 (IL-2) receptor *IL2RG* was predicted by exon arrays to be present in HT-29 cells and absent in MCF10A and MDA-MB-436 cells (Figure 2-5A, B, and C). By contrast, qRT-PCR receptome profiling predicted that *IL2RG* should be present in all

three but absent in a fourth cell line, A375 melanoma cells. Functional testing of *IL2RG* presence through IL-2 stimulation was not possible, because receptome profiling indicated that these cell lines lacked one or more of the requisite subunits for the IL-2 receptor heterotrimer (*IL2RA* and *IL2RB*). Nevertheless, we found by immunoblotting that IL2RG was detected in HT-29, MCF10A, and MDA-MB-436 cells, but not in A375 cells, and the relative abundance of IL2RG protein was consistent with its relative transcript abundance obtained by qRT-PCR receptome profiling (Figure 2-5D). These data indicated that sensitivity remains a challenge for exon-targeted microarrays when compared to receptome profiling by qRT-PCR.

2.5. qRT-PCR receptome profiling is more specific for mature transcripts than RNA-seq

A third alternative for global receptome profiling is RNA-seq (*53-55*), which is more sensitive than oligonucleotide microarrays (*54, 56*). To compare RNA-seq directly with qRT-PCR receptome profiling, we magnetically purified poly(A)$^+$ RNA from lysates of HT-29 cells treated with IFN-γ, MDA-MB-436 cells, or MCF10A cells and sequenced at two depths: 25 million (M) reads and 50M reads (IFN-γ-treated HT-29, MDA-MB-436) or 25M reads and 100M reads (MCF10A). As expected, the RNA-seq analyses were strongly correlated across duplicates (Figure 2-6A), and the 50–100M analyses detected sequences from substantially more genes than the matched 25M analyses (Figure 2-6B). The RNA-seq data provided an unbiased, comprehensive, and replicated set of measurements to compare with qRT-PCR receptome profiling.

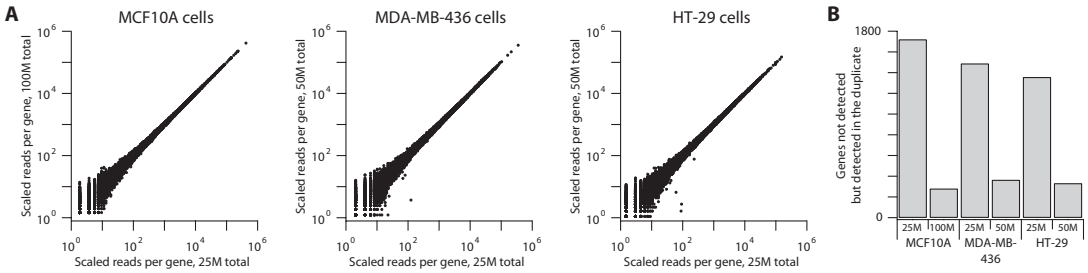


Figure 2-6. RNA-seq is consistently reproducible across assay duplicates with sensitivity that depends on sequencing depth. (A) Duplicate RNA-seq data collected from the indicated cells. HT-29 cells were IFN-g sensitized. Note that the total reads differ along the y-axis: MCF10A, 100M total reads, MDA-MB-436 and HT-29, 50M total reads. Data were scaled as the total reads per gene per million mapped reads to account for the difference in total reads between the x- and y-axes. (B) Decrease in the number of missed genes with increasing sequencing depth. Data are shown as the number of genes with zero mapped reads that were detected in the duplicate RNA-seq dataset.

We normalized the RNA-seq data to yield relative transcript abundances as reads per kilobase per million mapped reads (RPKM) (*55*) and generated ROC curves with respect to the qRT-PCR present-absent calls made by receptome profiling. For nontumorigenic MCF10A cells, there was a strong concordance between RPKM and qRT-PCR receptome profiling, which improved slightly with the depth of sequencing (Figure 2-7A). For instance, using a detection threshold of 0.3 RPKM (*6*), we found that RNA-seq could correctly distinguish the presence of *CSF1R* in MCF10A cells (~0.6 RPKM) from its absence in MDA-MB-436 cells (~0.02 RPKM) (Figure 2-7, A and B). However, for MDA-MB-436 cells, the RPKM-qRT-PCR agreement was much poorer, because false positives increased proportionally with false negatives for most RPKM thresholds (Figure 2-7B). This pattern was also observed in IFN-γ-treated HT-29 cells, with false positives increasing abruptly at thresholds as high as 10–20 RPKM (Figure 2-7C). The discrepancies in the two cancer cell lines was not resolved by deeper sequencing (Figure 2-7B and C, lower graphs), suggesting a fundamental difference between RNA-seq and qRT-PCR receptome profiling.

To determine which data type corresponded more closely to signaling competency, we selected two receptors with large RPKM values that were predicted to be absent by qRT-PCR receptome profiling. The fibroblast growth factor receptor 1-encoding gene *FGFR1* is overexpressed in some colon cancers (*82*) and was detected at ~20 RPKM in IFN-γ-treated HT-29 cells, but *FGFR1* transcripts were predicted to be absent by receptome profiling. Using a C-terminal antibody recognizing multiple splice variants of FGFR1, we detected FGFR1 in multiple colon cancer cell lines but not in IFN-γ-treated HT-29 cells (Figure 2-7D). Another discrepancy was found with the epidermal growth factor receptor family member *ERBB3*, which was sequenced at ~3 RPKM in MCF10A cells and was present by qRT-PCR, ~90 RPKM in MDA-MB-436 cells and was absent by qRT-PCR, and ~0.4 RPKM in HT-29 cells and was present by qRT-PCR. We immunoblotted for the cytoplasmic domain of ERBB3 and detected strong immunoreactivity in MCF10A cells, which was weaker in HT-29 cells and absent in MDA-MB-436 cells (Figure 2-7E), consistent with the relative abundances predicted by qRT-PCR receptome profiling. Therefore, abundant RNA-seq alignments did not necessarily correspond to functional receptors in cancer cells.

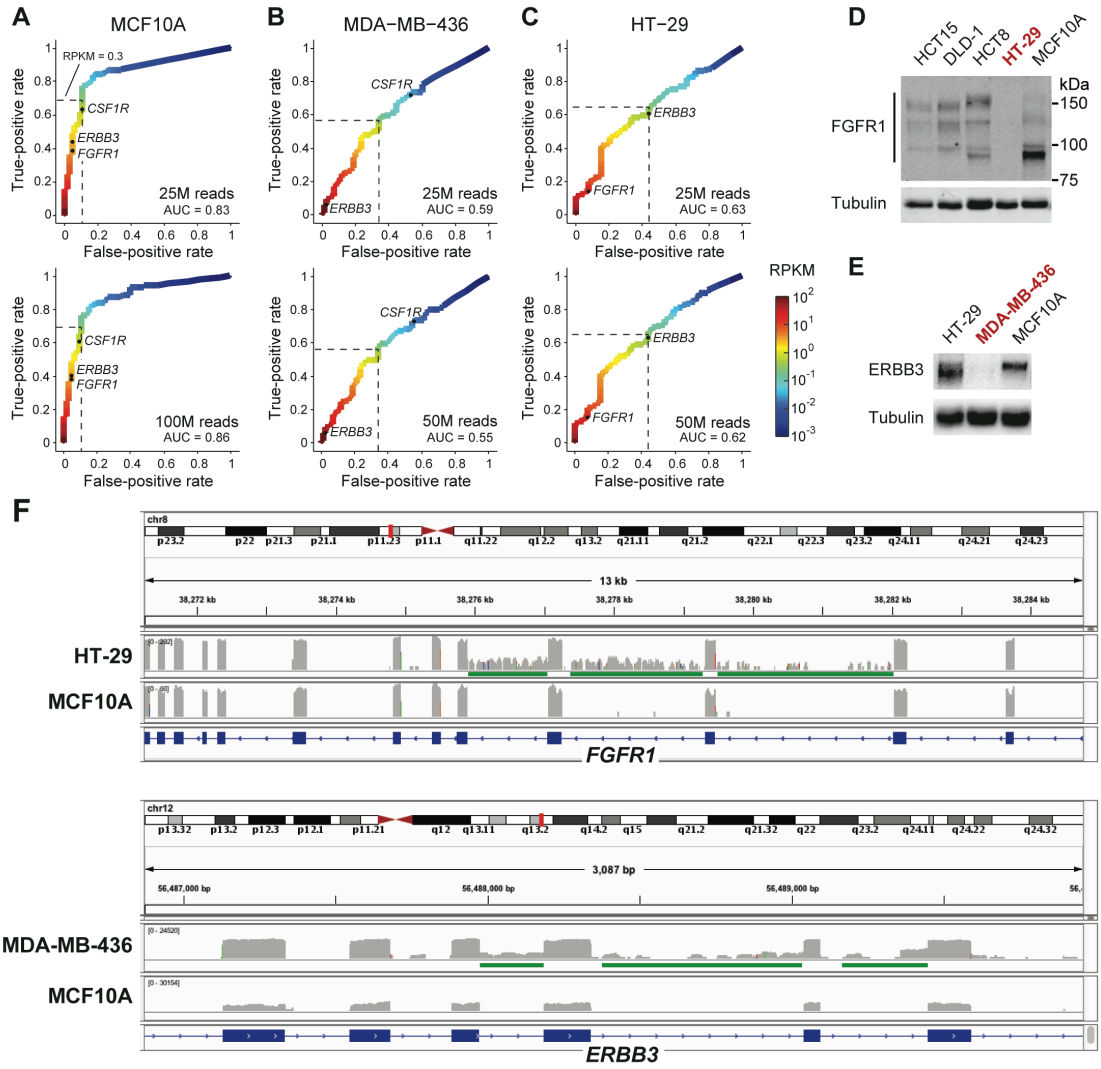
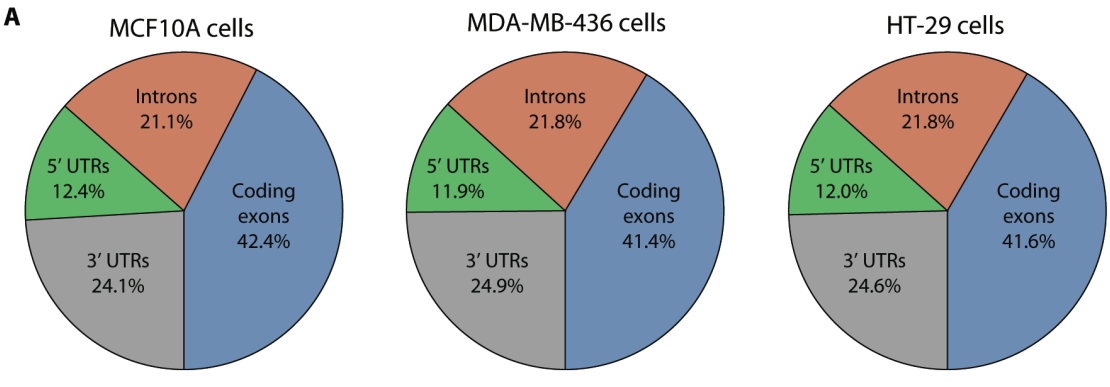


Figure 2-7. qRT-PCR receptome profiling is more specific for detecting functional receptor genes than RNA-seq. (A to **F)** Receiver operating characteristic (ROC) curves relating RNA-seq reads per kilobase per million mapped reads (RPKM) to qRT-PCR present-absent calls for MCF10A cells analyzed at 25M total reads (A) or 100M total reads (B), MDA-MB-436 cells analyzed at 25M total reads (C) or 50M total reads (D), and IFNγ-sensitized HT-29 cells analyzed at 25M total reads (E) or 50M total reads (F). The dashed line indicates a previously reported RPKM threshold for gene detection by RNA-seq (*6*). The area under the ROC curve (AUC) is shown as in Fig. 4. **(G)** Detection of FGFR1 protein in various colon cancer cell lines and MCF10A cells but not in IFNγ-sensitized HT-29 cells. **(H)** Detection of ERBB3 protein in IFNγ-sensitized HT-29 cells and MCF10A cells but not in MDA-MB-436 cells. **(I** and **J)** Coverage of RNA-seq reads across portions of the *FGFR1* **(I)** and *ERBB3* **(J)** loci for the indicated cell lines. Introns showing consistent coverage above background are underlined in green.

We examined *FGFR1* and *ERBB3* further by inspecting the coverage of aligned sequences across each locus. In both instances where the corresponding protein was absent despite high RPKM, we identified a subset of introns that were detected, suggesting incomplete splicing or aberrant intron retention (Figure 2-7F, green). The observed introns were not likely caused by assembly or alignment errors, because we obtained multiple paired-end reads spanning the intron-exon junctions of each retention event. MDA-MB-436 and HT-29 cells showed the same overall coverage of intronic and other noncoding RNA sequences compared to MCF10A cells (Figure 2-8A). However, when focusing on the putative false positives detected by RNA-seq in the cancer cell lines, we found that these genes were significantly enriched for intronic sequences relative to receptor transcripts that were also called present by qRT-PCR (Figure 2-8B). These data suggest that incompletely spliced RNA sequences can be discriminated more effectively by qRT-PCR-based profiling than by current implementations of RNA-seq.



B

MCF10A		Receptome profiling: Agree	Receptome profiling: Disagree
Intronic coverage	Yes	14	1
	No	173	6

$P = 0.44$

MDA-MB-436		Receptome profiling: Agree	Receptome profiling: Disagree
Intronic coverage	Yes	18	**15**
	No	147	**14**

$P < 10^{-5}$

HT-29		Receptome profiling: Agree	Receptome profiling: Disagree
Intronic coverage	Yes	19	**5**
	No	158	**12**

$P < 0.05$

Figure 2-8. Receptor discrepancies between qRT-PCR and RNA-seq in cancer cells are enriched for retained intronic sequences. **(A)** Distribution of transcript-associated RNA-seq reads in the indicated cell lines. HT-29 cells were IFN-g sensitized. **(B)** Contingency tables for the receptor genes that agree or disagree between qRT-PCR and RNA-seq with respect to the genes with partial intronic coverage as measured by RNA-seq. Statistical significance was assessed by Fisher's exact test.

2.6. Receptome profiling defines signaling signatures enriched in specific tissue lineages

To demonstrate an application of receptome profiling, we surveyed the signaling receptomes of 40 human cell lines. The collection was weighted toward pancreatic, melanocytic, breast, and colonic lineages to evaluate the link between receptome signatures and tissue origin. As expected (*83*), we found that receptome signatures clustered significantly according to lineage (Figure 2-9). Lineage enrichment was associated with the high abundance of certain signaling receptor transcripts (Figure 2-10, yellow). For example, transcripts for the receptor tyrosine kinase ERBB3 were increased among breast epithelia, which may explain why some breast cancers have amplification

of *ERBB2*, a dimerization partner of ERBB3 (*84*). Many tissue-enhanced patterns were supported by previous studies, although roughly half of the patterns uncovered by receptome profiling had not been described to our knowledge (Table 2-2).

To exploit the qualitative sensitivity of receptome profiling, we removed all quantitative information and reclustered the 40 cell lines on the basis of the presence or absence of receptor transcripts. The binary present-absent signature was sufficient to categorize much of the cell-line panel according to lineage (Figure 2-11). For the tissue types in the panel, lineage enrichment was not associated with tissue-selective presence of receptor subsets, but rather with the absence of transcripts (Figure 2-12, yellow). Using strict criteria for lineage specificity, we identified seven receptors with tissue-specific absence but only one receptor with tissue-specific presence (Table 2-3). For example, the chemokine receptor *XCR1* was absent in eight of 10 melanoma lines ($P < 0.005$, hypergeometric test), consistent with the reported loss of *XCR1* in culture compared to primary melanoma tumors (*85*). A few absent signatures could be inferred from literature reports, but most had not been reported previously (Table 2-3).

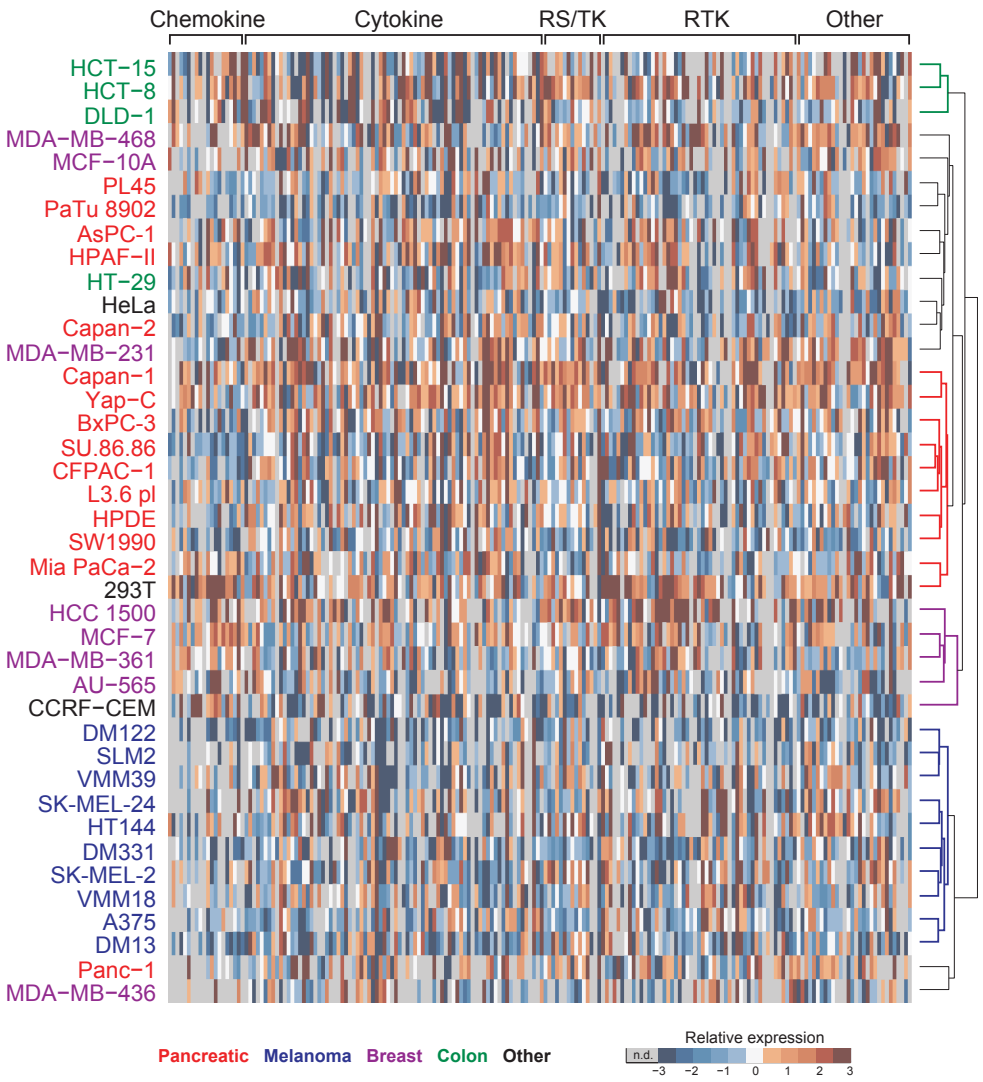


Figure 2-9. One-way hierarchical clustering of relative receptor transcript abundances obtained by receptome profiling in 40 human cell lines. One-way hierarchical clustering of receptor-specific relative abundance after normalization to *GAPDH* as a loading control. Clustering was done using a Euclidean distance metric with Ward's linkage. Dendrogram branches significantly enriched for specific lineages ($P < 0.05$) are matched to the color associated with the lineage. Enrichment analysis for cell lineages was performed by the hypergeometric test. n.d., not detected.

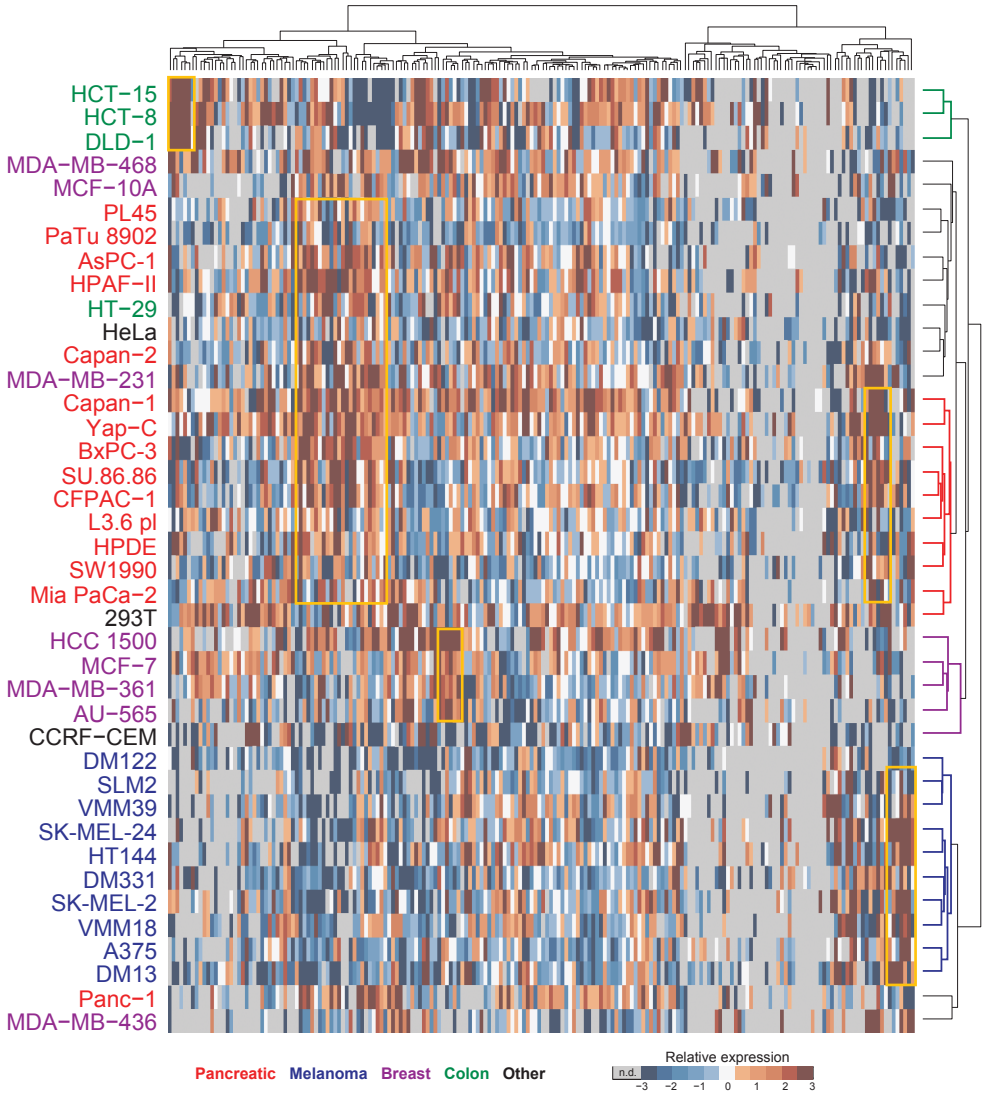


Figure 2-10. Two-way hierarchical clustering of relative receptor abundances obtained by receptome profiling in 40 human cell lines. Two-way hierarchical clustering of receptor-specific relative abundance after normalization to *GAPDH* as a loading control. Clustering was done using a Euclidean distance metric with Ward's linkage. Yellow boxes highlight local clusters of receptor patterns that are lineage specific. Dendrogram branches significantly enriched for specific lineages ($P < 0.05$) are matched to the color associated with the lineage. Enrichment analysis for cell lineages was performed by the hypergeometric test. n.d., not detected.

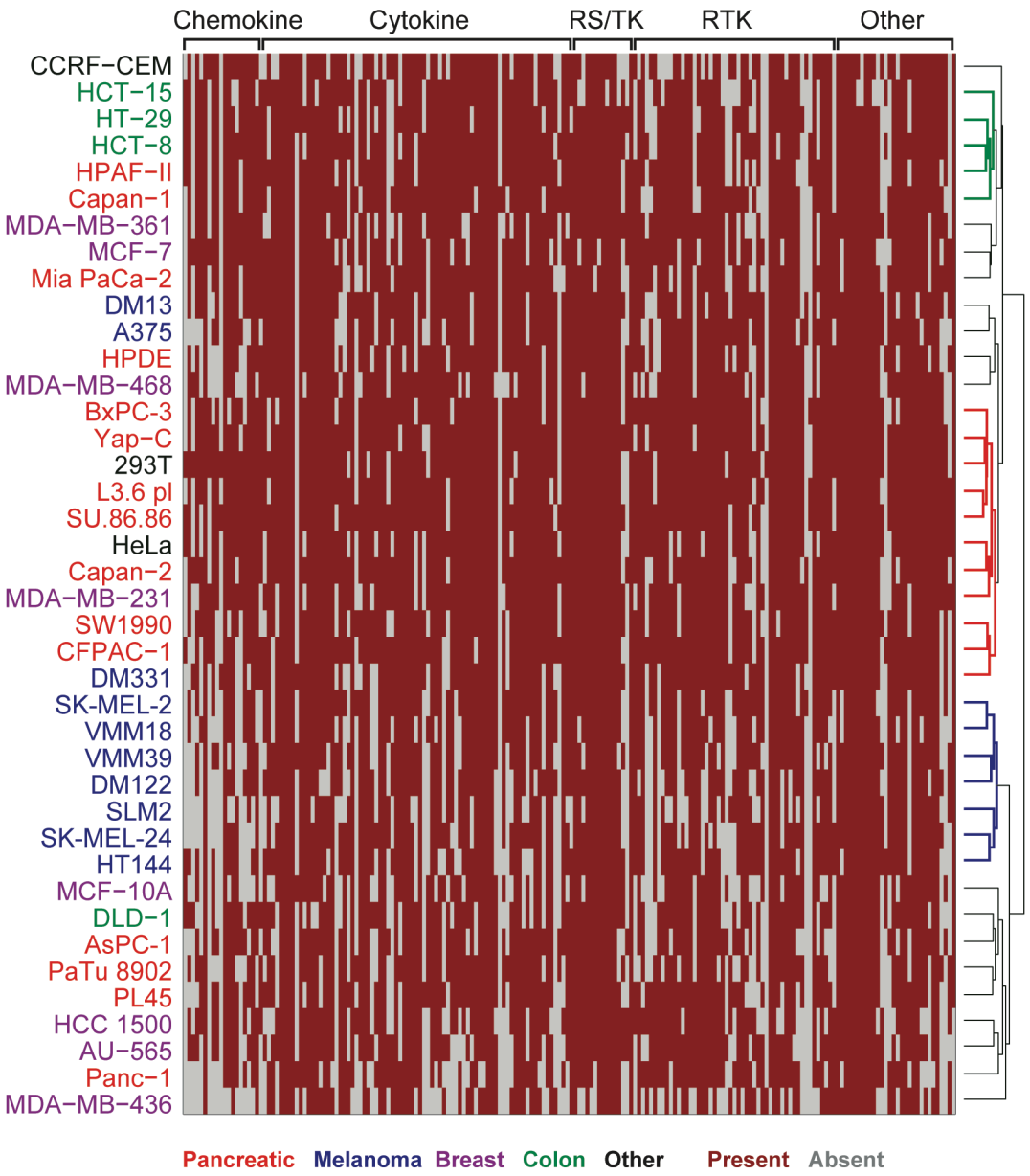


Figure 2-11. One-way hierarchical clustering of high-sensitivity present-absent calls obtained by receptome profiling in 40 human cell lines. One-way hierarchical clustering of high-sensitivity present-absent calls from receptome profiling. Clustering was done using a Euclidean distance metric with Ward's linkage. Dendrogram branches significantly enriched for specific lineages ($P < 0.05$) are matched to the color associated with the lineage. Enrichment analysis for cell lineages was performed by the hypergeometric test. n.d., not detected.

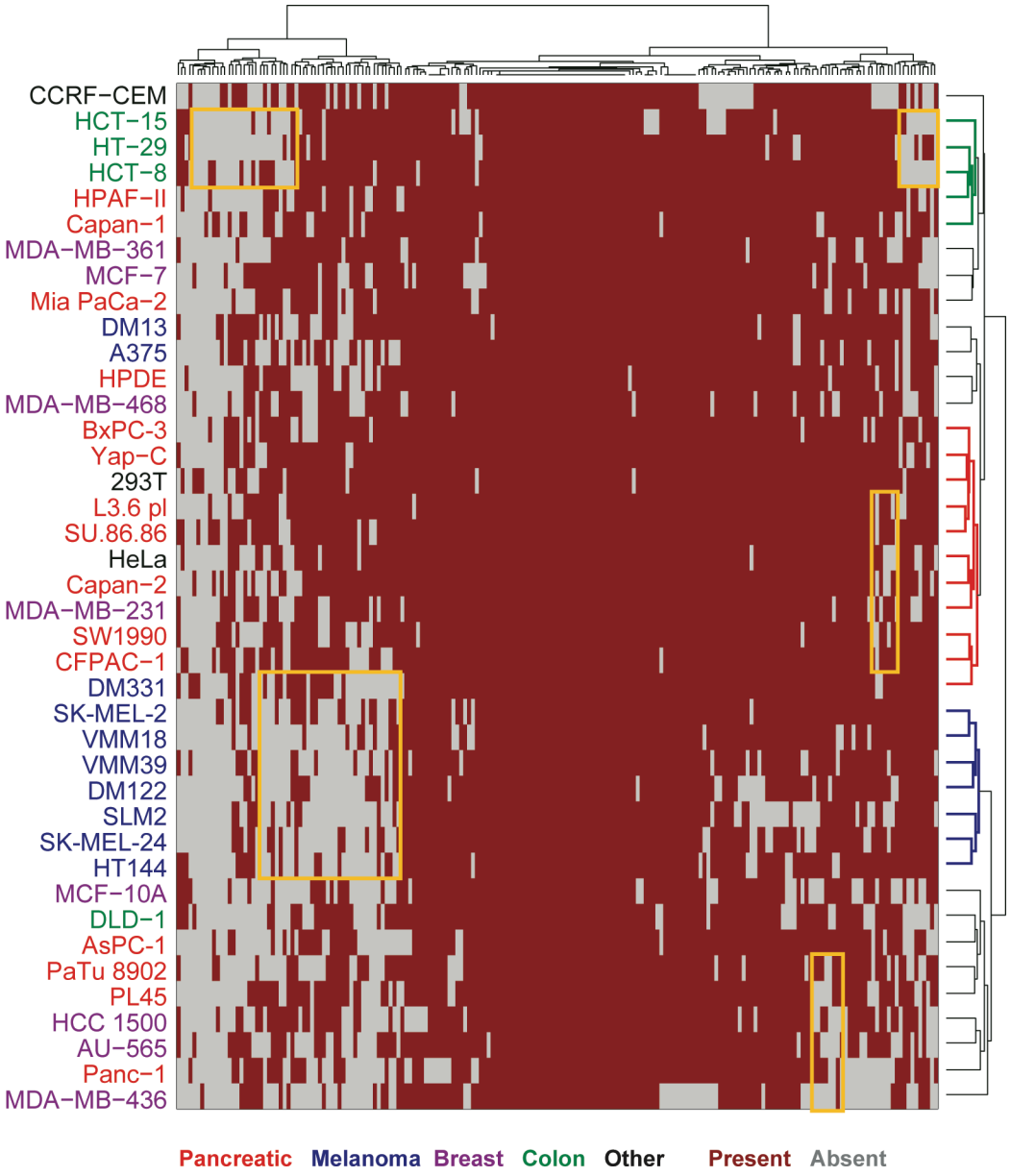


Figure 2-12. Two-way hierarchical clustering of high-sensitivity present-absent calls obtained by receptome profiling in 40 human cell lines. Two-way hierarchical clustering of high-sensitivity present-absent calls from receptome profiling. Clustering was done using a Euclidean distance metric with Ward's linkage. Yellow boxes in (B) and (D) highlight local clusters of receptor patterns that are lineage specific. Dendrogram branches significantly enriched for specific lineages ($P < 0.05$) are matched to the color associated with the lineage. Enrichment analysis for cell lineages was performed by the hypergeometric test. n.d., not detected.

2.7. Ectopic expression of IL10RA in melanoma cells engages an artificial autocrine circuit

We examined the impact of receptor absence on cell function by selecting the interleukin-10 (IL-10) receptor alpha subunit *IL10RA* for follow-up studies. *IL10RA* was called absent in nine of 10 melanoma lines by qRT-PCR receptome profiling (Table 2-3), which was notable because melanoma cells are a source of anti-inflammatory IL-10 (*86, 87*). To determine if absence of IL10RA influenced cell behavior, we used A375 melanoma cells, which lack IL10RA but constitutively secrete IL-10 (*88*). We transduced the cells with either a control luciferase-expressing lentivirus or a lentivirus encoding IL10RA. As expected, IL10RA was not detectable in control luciferase-expressing A375 cells but was present in cells transduced with IL10RA (Figure 2-13A). The IL10RA-expressing cells also showed phosphorylation of STAT3 upon stimulation with recombinant IL-10, whereas the control A375 cells were unresponsive (Figure 2-13, B and C). Therefore, the A375 melanoma cell line has all the intracellular machinery for transducing an IL-10 signal except for IL10RA, which acts as a gatekeeper for conferring cellular responsiveness.

To determine whether IL10RA had engaged an artificial autocrine circuit in A375 cells, we analyzed the concentration of IL-10 in conditioned medium by ELISA. IL-10 was readily detected in the medium conditioned by control cells but not in medium conditioned by IL10RA-expressing cells (Figure 2-13D). By contrast, *IL10* mRNA abundance was the same in the control and IL10RA-expressing cells (Figure 2-13E), suggesting that the absence of IL-10 in the medium of IL10RA-expressing cells could be

the result of autocrine trapping. We also noted a ~60% reduction in basal STAT3 phosphorylation (Figure 2-13F), which may be due to chronic IL-10 signaling causing feedback desensitization of other STAT3-activating pathways in IL10RA-expressing cells (*89*).

To test whether the IL10RA-triggered autocrine circuit was sufficient to affect cellular responses, we stimulated receptors of the tumor necrosis factor (TNF)-family that were detected in A375 cells by qRT-PCR receptome profiling. IL10RA slightly increased the resistance of A375 cells to apoptosis induced by FAS crosslinking with anti-APO (Figure 2-13G). The transcriptional signature of nuclear factor-κB (NF-κB) target genes was also altered when IL10RA-expressing A375 cells were stimulated with TNF (Figure 2-13H). Whereas some NF-κB targets, such as *NFKBIA* and *CCL2*, were unaffected by ectopic IL10RA expression, others were significantly inhibited (*IL1A* and *IL6*, Figure 2-13H). We conclude that receptor absence is important to insulate cells from secreted proteins that are meant to act as paracrine ligands. Aberrant receptor expression creates autocrine circuitry that traps paracrine factors locally and disrupts signaling, gene expression, and cellular responses.

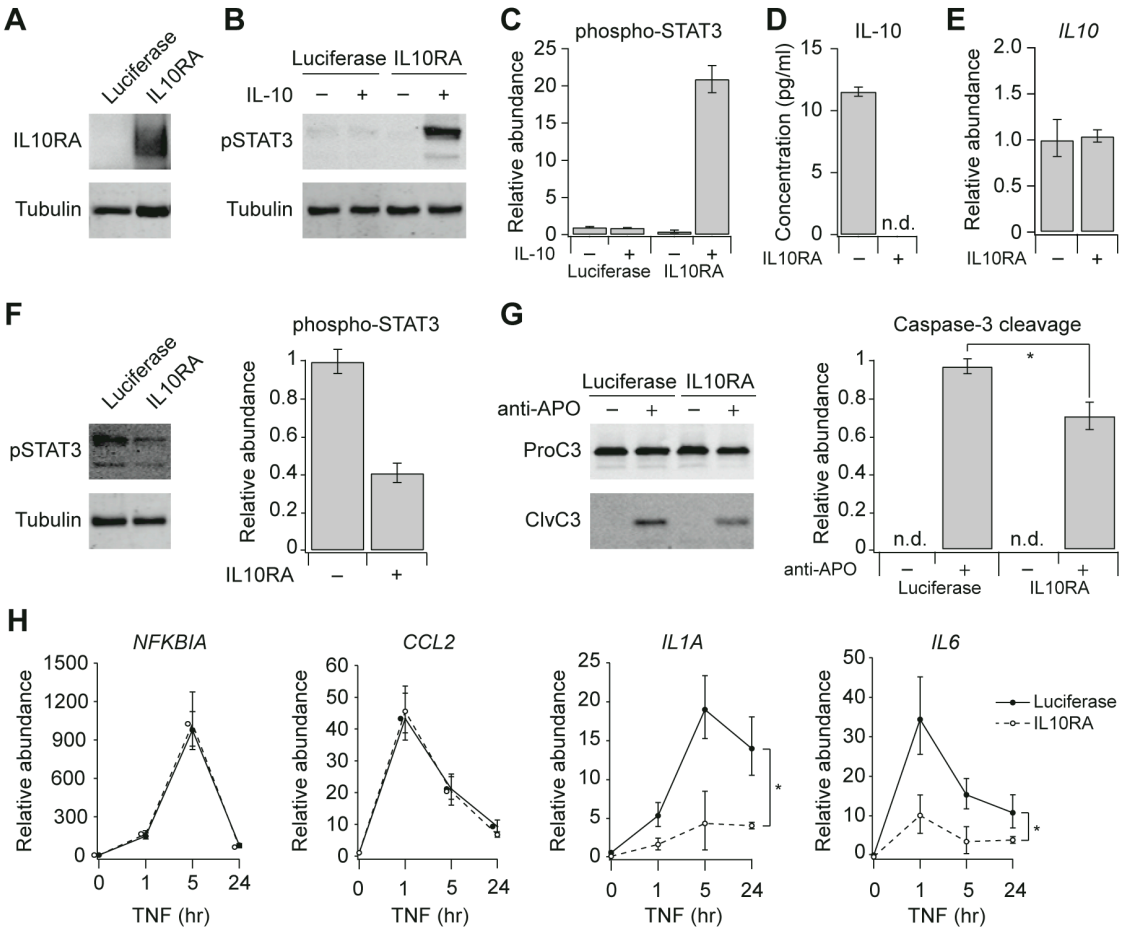


Figure 2-13. Forced expression of IL10RA in melanoma cells creates an autocrine signaling loop that alters signaling, gene expression, and cell fate. **(A)** IL10RA protein abundance after lentiviral transduction of A375 cells with V5-tagged luciferase or IL10RA. **(B)** STAT3 phosphorylation in luciferase- or IL10RA-expressing A375 cells following treatment with 20 ng ml^{-1} IL-10 for 20 min. **(C)** Replicated densitometry of IL-10-induced STAT3 phosphorylation in A375 cells. **(D)** ELISA quantification of IL-10 protein in the conditioned medium of luciferase- or IL10RA-expressing A375 cells. **(E)** qRT-PCR quantification of *IL10* mRNA abundance in luciferase- or IL10RA-expressing A375 cells. **(F** and **G)** Decrease in baseline STAT3 phosphorylation for A375 cells ectopically expressing IL10RA. **(H** and **I)** Caspase-3 cleavage in luciferase- or IL10RA-expressing A375 cells after FAS crosslinking with 1 μg ml^{-1} anti-APO for 24 hr. **(J** to **M)** qRT-PCR quantification of NF-κB target genes in luciferase- or IL10RA-expressing A375 cells following treatment with 100 ng ml^{-1} TNF for the indicated time points. For (A), (B), (F), and (H), cells were immunoblotted for the indicated proteins with tubulin or procaspase-3 used as a loading control. For (C), (D), (G), and (I), data are shown as the mean ± s.e.m. of three independent samples. For (E) and (J) to (M), data are shown as the geometric mean ± log-transformed s.e.m. of four independent samples. Asterisk indicates statistical significance ($P < 0.05$) by Welch's one-sided *t* test (I) or log-transformed two-way ANOVA with Sidák post-test correction (L and M). All immunoblots are representative of at least three independent experiments.

2.8. Perturbation of cellular receptome signatures by environmental stimuli

We explored the plasticity of cellular receptomes by profiling receptor transcript abundance in 293T embryonic kidney cells and MCF7 breast carcinoma cells after exposure to various stimuli (Figure 2-14A). We used EGF as a growth factor stimulus, IFN-γ and TNF as proinflammatory stimuli, and ionizing radiation (IR) as an environmental stress. We found that most stimulus-induced changes in the abundance of receptor transcripts were relatively minor (± twofold). This was particularly true for IR-treated samples, which gave rise to abundance changes that were highly variable across independently irradiated cultures. For the proinflammatory stimuli, however, there were several notable transcriptional responses that warranted additional analysis. In both cell lines, the TNF-superfamily receptor *TNFRSF9* was strongly induced upon TNF stimulation (Figure 2-14B), consistent with a previous report (*90*). We also observed many changes that were specific to cell type, indicating context-specific transcriptional programs. For example, abundance of the insulin receptor transcript *INSR* mildly increased in TNF-stimulated MCF7 cells (Figure 2-14B). MCF7 cells also showed selective increases in *TLR3* and *IL15RA* transcripts upon stimulation with IFN-γ, which were not observed in 293T cells (Figure 2-14B). This difference cannot be attributed to a general lack of IFN-γ responsiveness, because 293T cells abundantly express the transcripts of the required receptors and IFN-γ triggers changes in the abundance of other transcripts (Figure 2-14A) (*91*). The induction of *IL15RA* upon IFN-γ treatment of MCF7 cells agrees with a previous study (*92*), and we further found that *IL15RA* was also induced in MCF7 cells by TNF (Figure 2-14B). TNF-stimulated transcription of *IL15RA* has not been previously reported, illustrating that receptome profiling can be used as a

discovery tool to link environmental changes to transcriptional signatures of other environmental sensors.

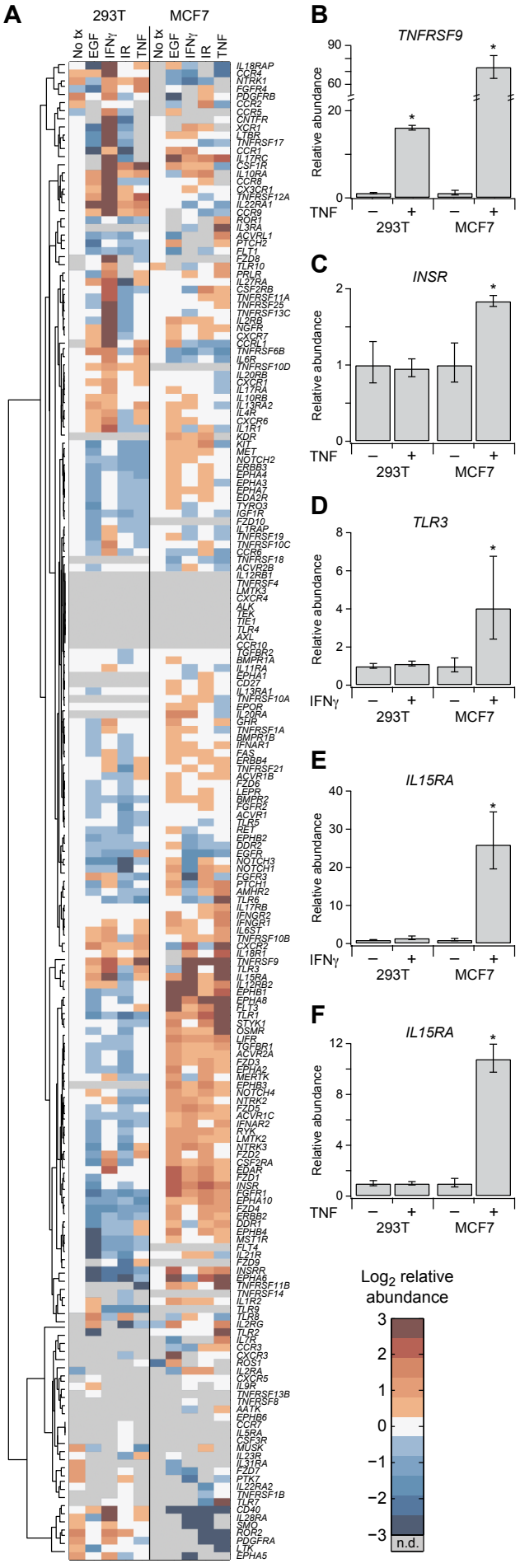


Figure 2-14. Stimulus-dependent changes in receptome profiles are dependent on cell type. (A) Receptor abundances in 293T embryonic kidney cells and MCF7 breast carcinoma cells after stimulation with 100 ng ml^{-1} EGF for 4 hr, 200 U ml^{-1} IFNγ for 4 hr, 5 Gy IR for 2 hr, or 20 ng ml^{-1} TNF for 4 hr. One-way hierarchical clustering was done using a Euclidean distance metric with Ward's linkage after normalization to *GAPDH* as a loading control. Data were centered on the cell type-matched untreated (No tx) condition or the median observed abundance if the receptor was absent under the untreated condition. **(B** to **F)** Plate-matched qRT-PCR quantification of *TNFRSF9* (B), *INSR* (C), *TLR3* (D), and *IL15RA* (E and F) in response to TNF (B, C, and F) or IFNγ (D and E). For (A), data are shown as the median cycle threshold (approximate $\log_2$ relative abundance) of three independent biological samples. For (B) to (F), data are shown as the geometric mean ± log-transformed s.e.m. of three independent biological samples. Asterisk indicates statistical significance ($P < 0.05$) by log-transformed, unpaired one-sided *t* test. n.d., not detected.

2.9. qRT-PCR receptome profiling is compatible with primary human tissues

To illustrate that receptome profiling can be applied to primary tissue samples, we profiled primary specimens of brain and skeletal muscle (Figure 2-15). Compared to the cell lines, we detected transcripts from significantly fewer types of receptors in the primary tissues ($P < 10^{-6}$, binomial test assuming 78% of receptors are present based on Figure 2-10). Although some transcripts may have been lost during sample isolation, we attributed the restricted overall expression pattern to the highly specialized tissues examined. Many specific receptors detected in one or both tissues were consistent with the known biology, including the presence of *GHR* (encoding growth hormone receptor) in muscle (*93*), *SMO* and *PTCH2* (encoding the Hedgehog receptor Smoothened and its coreceptor target Patched) and *FZD*-family (encoding the Wnt receptors of the Frizzled family) transcripts in brain (*94, 95*), and *INSR* in both brain and muscle (*96, 97*). Conversely, some receptor transcripts that were ubiquitous in cultured epithelial cells, such as *EPHA2* and *EPHB4*, were absent in the brain-muscle isolates, corroborating their reported tissue distribution (*98, 99*). We conclude that qRT-PCR receptome profiling is a versatile approach for systematic interrogation of canonical receptors involved in cell signaling.

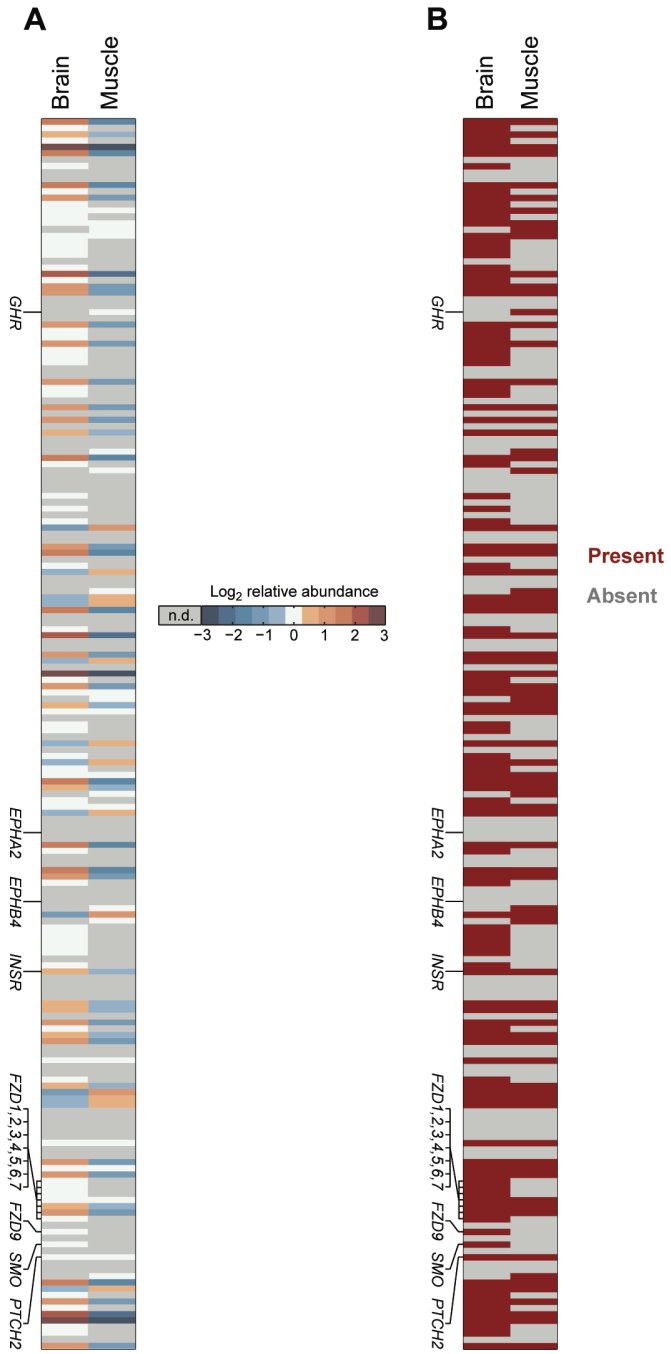


Figure 2-15. Receptome profiling is compatible with primary tissue samples. **(A)** Relative receptor abundances and **(B)** high-sensitivity present-absent calls obtained by receptome profiling in a primary human brain sample and a primary human muscle sample. n.d., not detected.

Table 2-2. Signaling receptors with abundant transcripts indicating lineage-specific gene expression.

Lineage	Receptor gene	Literature support (if available)
Pancreas	*CD40*	High abundance in pancreatic cancer (*100*).
	EPHA2	Increased abundance associated with pancreatic cancer and metastases (*101, 102*).
	EPHB2	High abundance in the developing pancreatic epithelium (*103, 104*).
	ERBB1	Increased abundance in pancreatic cancer (*105*).
	ERBB2	High abundance in the fetal pancreas during development (*106*).
	FGFR2	Required for normal pancreas development (*107*) and increased abundance in pancreatic cancer (*108*).
	FGFR3	Inhibits expansion of the immature pancreatic epithelium (*109*).
	IL1R2	Candidate biomarker for pancreatic ductal adenocarcinoma (*110*).
	MET	Increased abundance in pancreatic cancer (*111, 112*).
	RON	Increased abundance in pancreatic cancer (*113*).
	TGFBR2	Increased abundance in pancreatic cancer cell lines (*114*).
	TNFRSF10A	High abundance in many pancreatic cell lines (*115*) and increased abundance in pancreatic cancer (*116*); acts as the dominant receptor for TRAIL signaling in pancreatic cancer (*117*).
	TNFRSF10D	Increased abundance in pancreatic cancer (*116*) and pancreatic cancer cell lines (*118*).
	IL2RB	None.
	IL7R	None.
	IL15RA	None.
	IL22RA1	None.
	IL31RA	None.
	MER	None.
	ROR1	None.
	STYK1	None.
	TLR6	None.
	TNFRSF14	None.

Melanoma	*EPHA3*	Increased abundance in melanoma and implicated in cell adhesion, movement, shape, and growth (*119*).
	EPHA5	Detected in multiple melanoma cell lines (*120*).
	GHR	High abundance in skin and melanoma (*121*).
	IL1R1	High abundance in melanoma cell lines (*122*).
	IL1RAP	Autocrine IL-1 signaling important for melanoma proliferation (*122*).
	TNFRSF19	Candidate biomarker for melanoma (*123*).
	ALK7	None.
	CXCR1	None.
	DDR2	None.
	PDGFRA	None.
	TLR5	None.
Breast	*DDR1*	Increased abundance in breast cancer (*124*).
	EPHB4	Associated with the histological grade and stage of breast cancer and a survival factor in breast cancer (*125*).
	ERBB3	Important for breast tumor cell proliferation (*126, 127*).
	FGFR4	Associated with ER and PR positivity and may be involved in breast tumorigenesis (*128*); predicts resistance to tamoxifen therapy (*129*).
	EDA2R	None.
	EPHB3	None.
Colon	*TLR2*	Increased abundance and may be involved in sporadic colorectal carcinogenesis (*130*).
	CSF1R	None.
	IL10RA	None.
	IL28RA1	None.
	XCR1	None.

Table 2-3. Lineage-specific presence or absence of signaling receptors.

I. Signaling receptors absent in a lineage-selective manner.

Lineage	Receptor gene	Literature support (if available)
Melanoma	*XCR1*	Present in primary tumors but absent in cell lines (*85*).
	IL10RA	Detected in only a very small fraction of melanoma cells in animal models (*131*), and see Figure 2-13A.
	IL20RA	None.
	IL22RA2	None.
	EPHA10	None.
Breast	*TNFRSF6B*	Hormonally induced (*132*) and present in hormone-positive breast-cancer cell lines, such as MCF7 (*133*). All other breast lines in the panel are hormone negative (*78*), and all but one lack *TNFRSF6B*.
Colon	*EPHA3*	None.

II. Signaling receptors qualitatively present in a lineage-selective manner.

Lineage	Receptor gene	Literature support (if available)
Colon	*CCR1*	Widely detected in intestinal epithelial cell lines (*70*).

2.10. Sensitivity of receptome profiling may be related to method of mRNA isolation

We verified that qRT-PCR receptome profiling is substantially more sensitive for discerning receptor presence or absence than microarrays, irrespective of the microarray probe coverage along the transcript. This result was expected considering the stringency of microarray hybridization that is required to gauge specificity reliably using perfect match and mismatch probes. More surprising was the superior specificity of qRT-PCR-based profiling compared to RNA-seq when the receptomes of cancer cells were profiled. The difference here may be related to the methods used for mRNA isolation during the two measurement techniques. Our first-strand synthesis for qRT-PCR is primed with oligo$(dT)_{24}$, and high-stringency reverse transcription is performed at 50 ºC, ensuring that most, if not all, cDNAs contain at least poly$(A)_{24}$ (*134*). For RNA-seq, however, poly$(A)^+$ transcripts are isolated by magnetic separation after room-temperature annealing to oligo$(dT)_{25}$ beads, which may co-purify mRNAs with much shorter oligo(A) tails. The distinction is important, because shorter oligo(A) tails remain on transcripts undergoing nonsense-mediated decay, which is triggered when premature stop codons are encountered after aberrant splicing events, such as intron retention (*135, 136*). Nonsense-mediated decay may be specifically enhanced in cancer cells to suppress anti-tumor immune responses (*137, 138*), which could explain why we observed most RNA-seq discrepancies in transformed cells. The discrepancies can, in theory, be avoided by sequencing the transcriptome as oligo$(dT)_{24}$-primed cDNA, but this decreases the uniformity of coverage along transcripts (*55*), which is a major advantage of RNA-seq.

2.11. Importance of receptor presence/absence

Our collection of qRT-PCR receptome profiles across 40 human cell lines complements other work showing that an exceedingly small fraction of proteins is detected in a purely cell- or tissue-specific manner (*139*). Furthermore, the binary present-absent signatures indicate that receptor silencing might be just as important in defining a lineage as the receptors that are highly abundant. Receptor silencing could be an important mechanism for enabling effective paracrine communication without the complications of autocrine crosstalk. For example, forced expression of *IL10RA* in melanoma cells would not only sequester IL-10 away from neighboring immune cells but would also severely dampen the induction of anti-inflammatory signals, such as *IL6*. We confirmed these predictions by showing that the IL10RA-expressing cells had less IL-10 in the medium (Figure 2-13D) and produced fewer *IL6* transcripts in response to TNF (Figure 2-13H). Thus, IL10RA-harboring melanoma cells would be predicted to be more immunogenic overall than their naturally occurring counterparts. Besides lineage-specific silencing, it may also be worth examining receptors that are lost in individual cancer lines to get a sense of how transformed cells evolve resistance to ligands that inhibit tumor growth.

2.12. Summary

The presence or absence of signaling receptors determines a cell's ability to respond to its environment. By defining a receptome panel and validating each qRT-PCR reagent in the array individually, we provide a convenient tool for establishing the boundaries of cellular responsiveness to the ligands that activate these receptors.

Detection of the mRNA of a receptor does not always imply that this receptor will be properly translated and localized to bind ligands and transmit signals. However, we showed that lack of mature receptor transcripts was consistent with cellular unresponsiveness (Figures 2-4, A and D, and 2-13B), a finding that required the sensitivity and specificity of the profiling approach described here.

A major challenge for deciphering the microenvironment is the complex cocktail of ligands that cells encounter physiologically (*140*). We can gain a clearer understanding of the microenvironment by distinguishing the ligands that activate intracellular signaling from those that are ignored. The qRT-PCR array described here provides a straightforward and scalable way to make this discrimination. With better sensitivity than microarrays and better specificity than RNA-seq at less than $1/10^{th}$ of the cost, qRT-PCR receptome profiling could be readily incorporated into large-scale characterizations of cell lines, primary tissues, and tumors.

CHAPTER 3

Quantitative modeling of local structure in signaling cascades

3.1 Introduction

Technological advances in experimental assays have permitted extensive study of individual signaling proteins. In Chapter 2, we described an experimental method that robustly measures transcript abundance for receptors, where presence or absence of receptor expression is critical in defining signaling capabilities of the cell. Given the dynamic properties of downstream signaling networks, it is not clear whether abundance or activity of a signaling protein, or both, is most informative. Experimental measurements of protein or transcript amount (Western blots, qRT-PCRand activity (phosphoproteomics, kinase activity assays (see Chapter 5), etc) are both used extensively to collect data about cell-signaling systems. Similarly, methods to perturb signaling proteins can target either transcript abundance (RNAi) or enzymatic activity (small-molecule inhibitors). Often, these types of perturbations are used interchangeably to corroborate findings with the other method and to make conclusions about individual signaling components in the network. In this chapter we study the local pathway connectivity and its effects on perturbations of protein abundance and activity. We describe a quantitative modeling approach to characterize local network topologies and how different topologies will respond to molecular perturbations. These findings have important implications on how we interpret data about networks, notably for identifying drug targets from high-throughput screening efforts.

3.2 High-throughput screening with RNAi and small-molecule inhibitors

To identify novel targets for drug development, many large-scale cell-based screening efforts have been pursued (*141, 142*). High-throughput screens largely fall into two classes. The first class involves using RNAi knockdown of signaling molecules to identify potential targets for the future development of inhibitors. The second class of screens involves using large chemical libraries to identify drugs or combinations of drugs that could be used for an alternate purpose (for example, repurposing an anti-inflammatory drug that shows efficacy in cancer treatment when combined with a kinase inhibitor (*143*)). While both these screening techniques can identify important signaling nodes, the underlying reagents perturb signaling pathways by distinct mechanisms. RNAi identifies nodes by reducing the target transcript abundance while drug screens identify targets by reducing their catalytic activity. A key challenge for RNAi screening efforts is to decide which of tens or hundreds of targets on the "hit" list are worthy of follow-up. Given the mechanistic differences between RNAi and small molecules, it is important to determine whether these two screening methods will yield the same high-priority targets.

There exist a few anecdotal examples in the literature where RNAi knockdown and enzyme inhibition of the same target give rise to different phenotypes (*144, 145*). However, despite ongoing screening efforts, it has yet to be directly tested whether RNAi and small-molecule inhibitor screens will identify the same quality of candidate targets. This information is critical to interpret and prioritize results from both RNAi and drug screens and to translate them to therapies in the clinic. Several mechanisms have been

proposed and studied for discrepancies between RNAi and drug targeting, such as time scales of inhibition, off-target effects of the shRNA or drug, and non-enzymatic roles of a signaling protein (scaffolding, for example) (*146*). Local network connectivity, however, has not yet been considered. Examining local network connectivity will yield a simple means for excluding targets that are embedded within networks predicted to be suboptimally druggable from high-throughput screens. It may also prompt a re-examination of signaling enzymes with weak knockdown phenotypes that are predicted to be preferentially druggable.

3.3 Feedback and feed-forward loops in cell-signaling networks

Signaling pathways are not simple linear cascades of molecules that can be turned "off" and "on". Instead, the interconnectedness of pathways allows for tight regulatory control by the cell over time. Within pathways, feedback loop motifs, where an output signal is connected back to an input signal, are important architectural designs by the cell to impart complex higher-level system behaviors (*147*). Feedback loops can be described as positive or negative, where the output signal can either promote or antagonize the original input signal. Feed-forward loops have also been described in signaling systems where the input can positively or negatively affect the output of the pathway (*148*). Feedback loops can inform particular signaling behaviors in a network, such as homeostasis or amplification (*147*). While it is easy to conceptualize and predict the effect of a perturbation against a single enzyme substrate reaction, predicting the effect becomes much more complicated when the enzyme and substrate are embedded in a pathway (where each is both a substrate and an enzyme in different reactions). Local

connectivity created by feedback loops creates an extra layer of complexity, imparting additional enzyme and substrate roles for a given signaling molecule. This complexity makes it impossible to ascertain the effect of perturbations within the pathway by visual inspection of a pathway wiring. Instead, predicting outcomes from complex pathways is a challenge well suited for computational modeling.

3.4. Computational modeling of feedback and feed-forward loops in cell signaling networks

We hypothesized that local connectivity of a signaling network determines whether a signaling enzyme will be sensitive to either protein abundance or catalytic activity. To test this hypothesis, we began by computationally modeling a simple three-tiered enzymatic cascade (consisting of enzymes A, B, and C) triggered by an input stimulus and yielding an amplified output response (Figure 3-1) (*149-151*). The models were constructed using simplifying Michaelis-Menten kinetics assumptions (for review

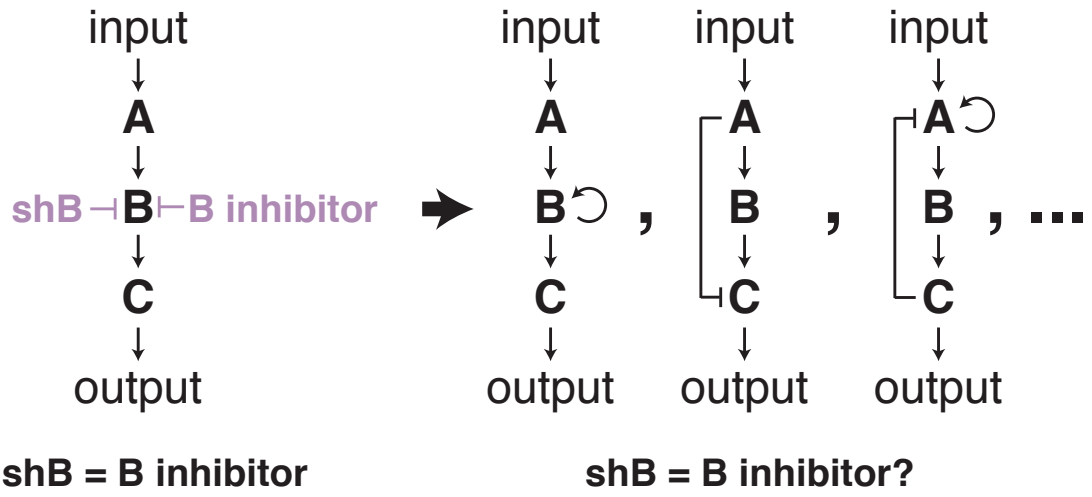


Figure 3-1. Exhaustive modeling of three-enzyme cascades. The initial model describes a three-enzyme cascade of enzymes A, B, and C, where B is targeted by RNAi or an inhibitor. Iterations of the model were built to include all possible one- and two-edge networks of positive or negative feedback.

see (*152*)). Michaelis-Menten kinetics describe an enzyme (E) and substrate (S) reaction where product (P) is formed:

$$[E] + [S] \leftrightarrow [ES] \rightarrow [E] + [P]$$

The rate of product formation (dP/dt) is described by

$$\frac{dP}{dt} = \frac{k_{cat}[E][S]}{[S] + K_M}$$

where k_{cat} is the catalytic rate and K_M is the Michaelis constant. RNAi decreases the amount of available enzyme ([E]) and subsequently decreases the numerator, whereas small-molecule inhibition increases the effective K_M and thereby increases the denominator of the equation. The parameters of the baseline model were optimized such that RNAi knockdown or competitive inhibition of the transducer enzyme B yielded identical inhibition of the pathway output at all levels of perturbation. We then adapted the model to simulate all possible iterations of one or two feedback, feed-forward, and autoregulatory loops in addition to the linear pathway (Figure 3-1).

3.5 RNAi and small-molecule inhibition are not equivalent perturbations to signaling pathways

By exhaustively simulating every one- and two-feedback configuration of a three-node signaling network we generated 95 pathway models. Perturbations were compared within each model by simulating the level of relative output (activated enzyme C compared to uninhibited pathway) over a range of 0-100% inhibition of the middle enzyme B by RNAi or a small-molecule inhibitor (Figure 3-2A). The area between these curves ("integrated difference") was used as a metric to compare the levels of discrepancy, where a negative integrated difference implies that the small-molecule is

more potent at inhibiting pathway output and a positive integrated difference implied that the RNAi is more potent (Figure 3-2A).

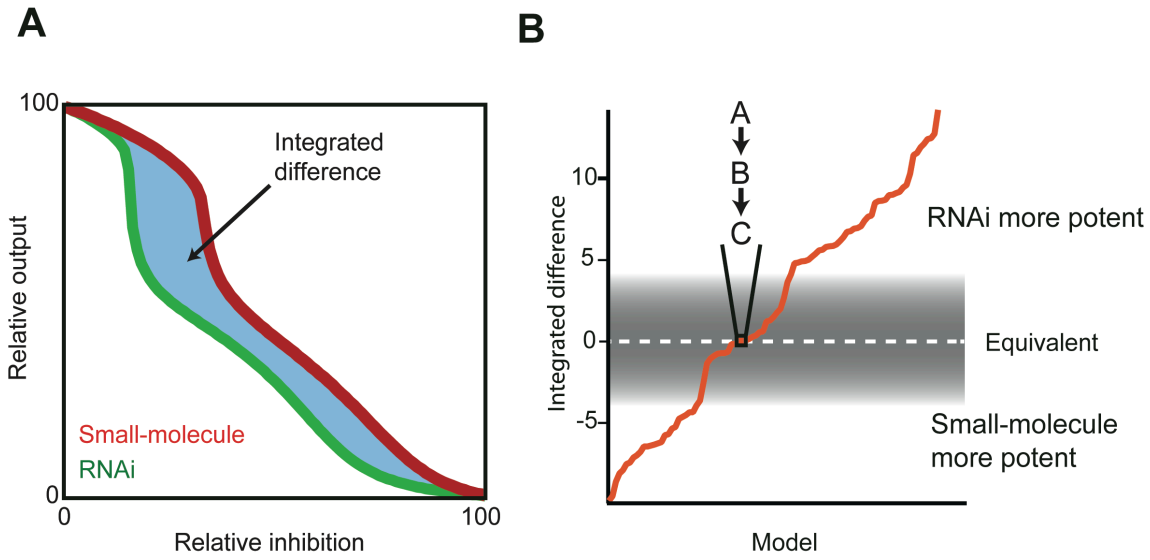


Figure 3-2. Pathway models are compared by their integrated discrepancies over 0-100% target inhibition. (A) The output of each pathway model was simulated over 0 to 100% target inhibition for small-molecule and RNAi. (B) Models were ranked by their levels of discrepancy as calculated in A.

Some network wirings predicted little or no discrepancy in pathway output (Figure 3-2B, gray), while other wirings predicted wildly different responses depending on whether enzyme B was perturbed by RNAi or small-molecule inhibition. Interestingly, some pathway configurations were sensitive to changes in protein levels but not small-molecule inhibition, whereas other pathway wirings were predicted to display the opposite behavior (Figure 3-2B). Interestingly, it is predicted that RNAi more effectively inhibits pathway output in more wirings than small-molecule inhibition (Figure 3-2B). Generally, for wirings that involve negative feedback at or above the target enzyme, the model predicts the output of the pathway will be more effectively inhibited by small-molecule inhibition. The model predicts the opposite for pathways

that include positive feedback at or above the target enzyme (Figure 3-3). Positive autoregulatory connections within a pathway also appear to be quite strong in dictating the effect of the perturbations, regardless of additional connections in the model (Figure 3-3, first three rows). These behaviors are not immediately clear upon inspection of a pathway and demonstrate the importance and utility of the models. While this simplified modeling framework has several limitations, including the assumptions made in building the models and the estimation of model parameters, it provides a starting point to begin to classify pathway wirings.

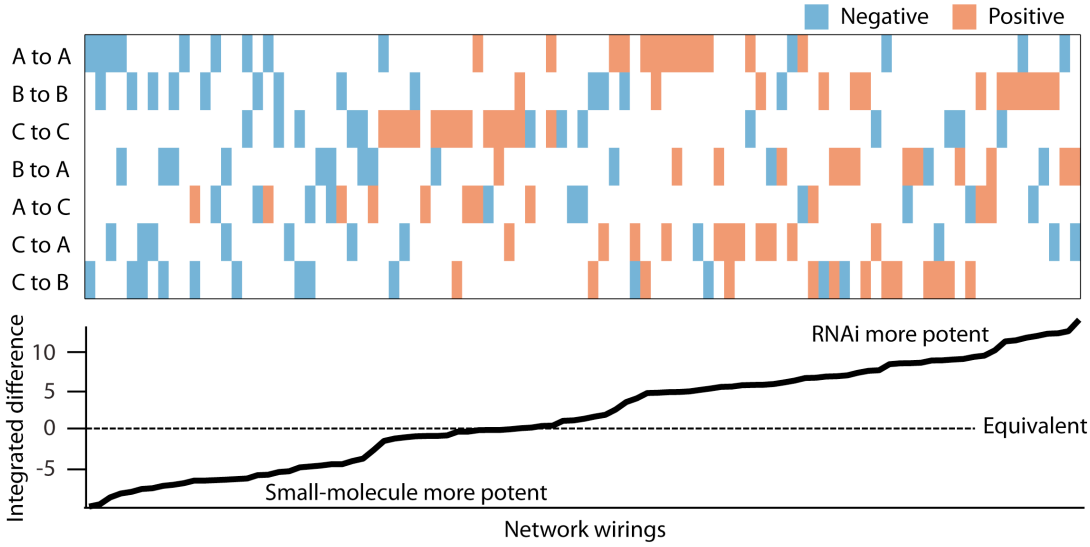


Figure 3-3. Negative and positive feedback patterns across 95 models. Enzyme connection components present in the models, where red indicates positive regulation and blue indicates negative regulation (top). Models ranked by integrated difference value (from Figure 3-2B) (bottom).

3.6. Experimental tests of model predictions

3.6.1. The Raf-MEK-ERK pathway

A network wiring that is predicted to respond differently to inhibition by small-molecule compared to RNAi is one where the third element negatively feeds back on the first (Figure 3-4A). In this model, it is predicted that small molecule inhibition of the second signaling element (B) will inhibit pathway output more potently than RNAi inhibition (Figure 3-4A). A signaling pathway that contains this subnetwork wiring is the Raf–MEK1/2–ERK1/2 pathway, where ERK1/2 hyperphosphorylates and inactivates Raf (*153, 154*), creating a negative-feedback loop (Figure 3-4B). The Raf–MEK–ERK circuit is an ideal case to test our model prediction because the pathway is relatively well insulated (*155*) and there are selective inhibitors of MEK1/2 (*156, 157*).

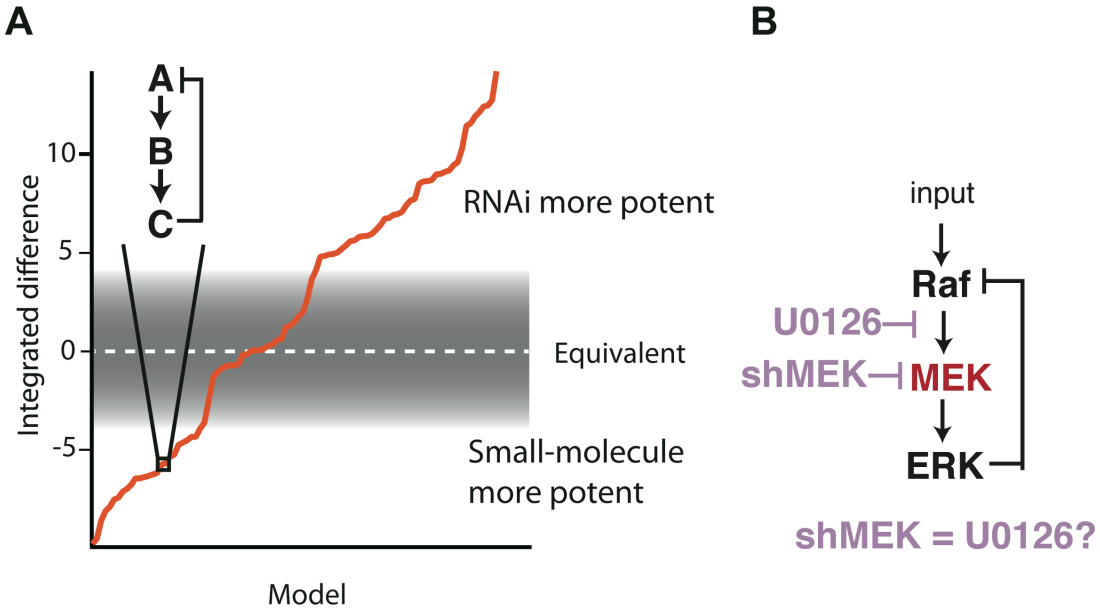


Figure 3-4. The Raf-MEK-ERK pathway is predicted to be more effectively inhibited by small-molecules than by RNAi. (A) The pathway wiring where enzyme C negatively feeds back on enzyme A is predicted to be more potently inhibited by small-molecule inhibition. (B) The Raf—MEK—ERK kinase cascade is a pathway that mirrors the structure in (A).

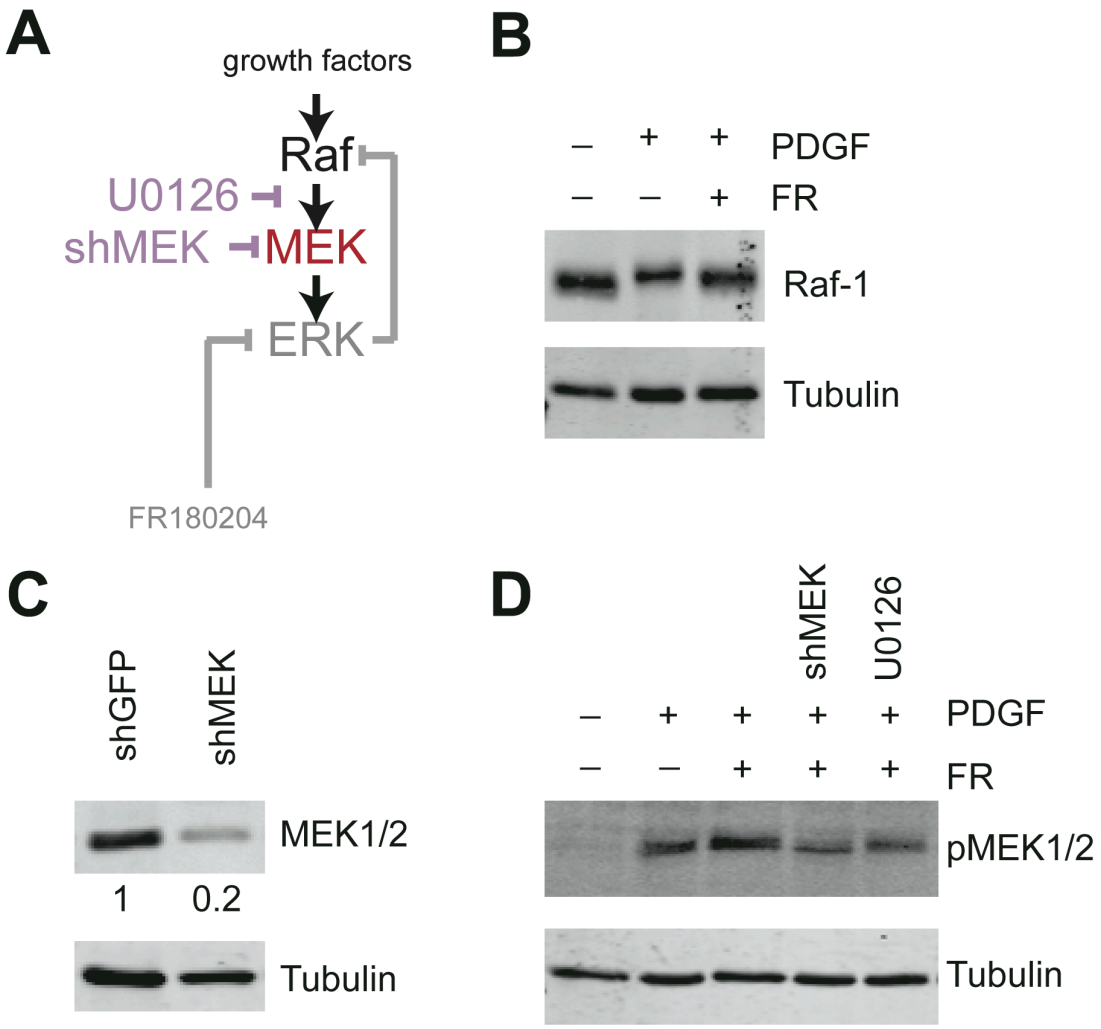


Figure 3-5. MEK1/2 inhibition by U0126 is dosed to be equivalent to RNAi knockdown in the absence of ERK1/2-Raf feedback. (A) ERK1/2-specific inhibitor FR180204 breaks negative feedback from ERK1/2 to Raf. (B) Raf hyperphosphorylation by ERK1/2 is prevented by FR180204 inhibition of ERK. (C) shMEK1/2 reduces total MEK1/2 levels by fivefold. (D) U0126 was dosed to reduce phosphorylated MEK1/2 levels to those seen with shMEK1/2 inhibition. For (B-D) 3T3 cells were serum starved for 24 hours and pretreated with 20 μM FR180204 and 2 μM U0126 for 1 hour prior to 1 ng/ml PDGF stimulation for 30 minutes (B) or 5 minutes (D).

To test this model hypothesis, we sought to compare the potency of a shMEK construct relative to the small-molecule inhibitor of MEK1/2 activation, U0126 (*156*) (Figure 3-5A). To make a fair comparison between shMEK1/2 and U0126, the "matched" U0126 dosage needed to be determined in the absence of negative feedback from ERK1/2 to Raf. This negative feedback was disrupted by adding FR180204, a selective

inhibitor of ERK1/2 (*158*) (Figure 3-5, A and B). By blocking ERK1/2 activity in response to growth factor (PDGF), FR180204 prevents the hyperphosphorylation of Raf, measured by its size upshift on a Western blot (Figure 3-5B). We engineered a stable RNA hairpin that targets both isoforms of murine MEK1/2 and reduces total MEK1/2 expression by fivefold in 3T3 fibroblasts (Figure 3-5C). A dose of U0126 (2 μM) was determined that lowered phospho-MEK1/2 to the same extent as the shMEK1/2 cells when the pathway was stimulated by a growth factor (Figure 3-5D). By matching the inhibition of U0126 to the hairpin, an appropriate comparison of mechanisms of inhibition could be made.

In vitro, MEK1/2 phosphorylation by Raf is inhibited approximately 70% by μM doses of U0126 (*157*). However, there are a few possible limitations to the dose matching evaluated in Figure 3-5D. First, the matching was performed using a phospho-MEK1/2 antibody that recognizes only phosphorylated serine 221 on MEK1/2. Raf phosphorylates MEK1/2 on both serines 217 and 221, and the bisphosphorylated MEK1/2 has been shown to be more active than the singly phosphorylated forms (*159, 160*).

To account for both of the MEK1/2 phosphorylation sites, phospho-MEK1/2 was monitored with an antibody that recognizes both the single and bis-phosphorylated forms of the protein over three concentrations of U0126 (Figure 3-6A). Unexpectedly, phospho-MEK1/2 levels increased with higher concentrations of U0126 (Figure 3-6A). This effect was not likely due to nonspecific effects of U0126 as the same effect was apparent using a second, mechanistically different inhibitor of MEK1/2, GSK1120212 (Figure 3-6B) (*159*). Despite elevated phospho-MEK1/2 levels, phospho-ERK1/2 was

still inhibited by both U0126 and GSK1120212 at higher concentrations (Figure 3-6, A and B). This effect was seen regardless of the phospho-site(s) analyzed (singly phosphorylated at serine 217 or 221, or bisphosphorylated), growth factor used for stimulation (EGF versus PDGF), or time of stimulation (data not shown).

At later times, ERK1/2 activation is limited by additional negative feedback by ERK-dependent expression of MAP kinase phosphatases (MKPs) and Sprouty proteins that down regulate upstream receptor activation (*161, 162*). Further, MEK1/2 phosphorylation can be induced by inhibition of MEK1/2 phosphorylation of ERK1/2 and disruption of the negative feedback from ERK1/2 to Raf (*163, 164*). While these previously described observations occur at longer times than the experiments here, it is possible that the 1 hour pre-treatment with MEK1/2 and ERK1/2 inhibitors could be confounding the resulting phosphorylated MEK1/2 by altering baseline signaling before growth factor stimulation.

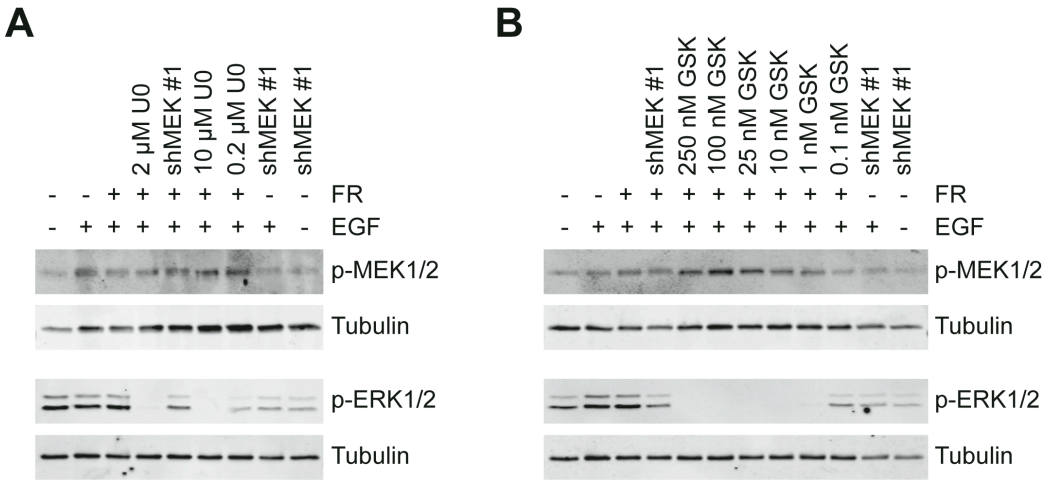


Figure 3-6. Phosphorylation of MEK1/2 increases with increased MEK1/2 inhibition by U0126 and GSK1120212. (A-B) 3T3 cells stably expressing shMEK1/2 or shGFP constructs were stimulated with 100 ng/ml EGF for 5 minutes. Cells were pretreated for 1 hour with 20 µM FR180204 and the indicated concentration of U0126 (A) or GSK1120212 (B). Samples are shGFP cells unless otherwise indicated. The phospho-MEK1/2 antibody recognizes pS217, pS221, and pS217/pS221 (Cell Signaling Technology #9121). For (A-B) 3T3 cells were serum starved for 24 hours and before drug pretreatment.

To determine if MEK1/2 and/or ERK1/2 inhibition creates artificial elevation of MEK1/2 phosphorylation, phospho-MEK1/2 was monitored in cells inhibited with 20 μM FR180204 and 0.1 nM GSK1120212 over 1 hour in the absence of pathway stimulation by growth factors (Figure 3-7). MEK1/2 phosphorylation was elevated as early as 5 minutes, and continued to rise over the 60 minute timecourse (Figure 3-7).

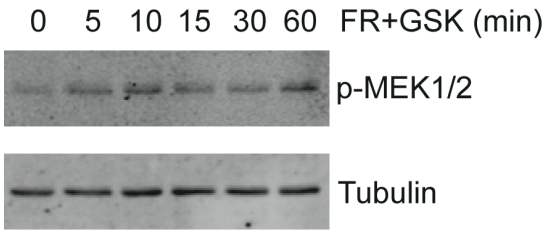


Figure 3-7. Phospho-MEK1/2 increases over time when MEK1/2 and ERK1/2 are inhibited. 3T3 cells stably expressing shGFP construct were treated with 20 μM FR180204 and 0.1 nM GSK1120212 for the indicated time following a 24 hour serum starvation.

Since previous experiments were performed by pre-treating the cells with inhibitors for 1 hour prior to growth factor stimulation, the elevated phospho-MEK1/2 levels could be an artifact from the drug pretreatment. Accounting for this artifact will be important in future experiments to more accurately evaluate the dose of MEK1/2 inhibitor that is equivalent to MEK1/2 knockdown.

Despite the potential limitations identified with our U0126 dose matching (Figures 3-6 and 3-7), we next analyzed pathway output (phospho-ERK1/2) over time in response to pathway activation by PDGF when ERK–Raf negative feedback was restored. Our preliminary data suggest that phospho-ERK1/2 responses are both qualitatively and quantitatively different for shMEK cells compared to cells treated with U0126 (Figure 3-8). In U0126-treated cells, phospho-ERK1/2 activation after PDGF stimulation is delayed and transient compared to activation in shMEK1/2 cells, which shows reduced phospho-ERK1/2 compared to controls but no differences in signaling

kinetics (Figure 3-8). The sustained phospho-ERK1/2 signaling in shMEK1/2 cells compared to U0126-treated cells thus supported our model prediction that small-molecule inhibition of MEK1/2 will more potently inhibit pathway output compared to RNAi. Levels of ERK1/2 signaling reduction may be clinically important, as in vivo models require greater than 80% inhibition of ERK1/2 signaling to demonstrate antitumor activity (*165*).

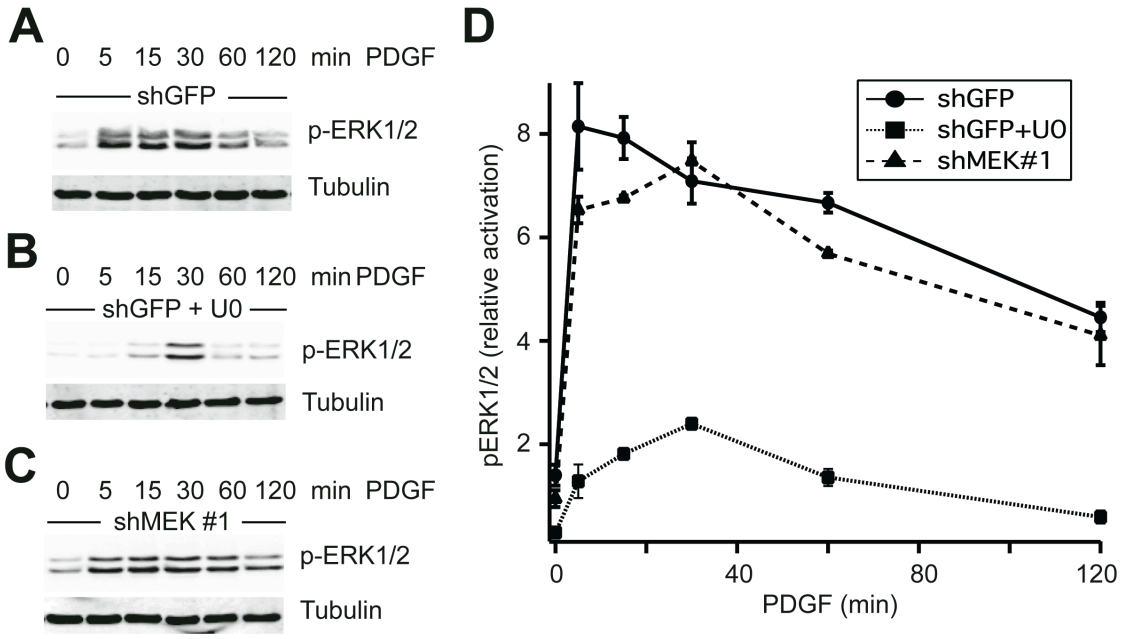


Figure 3-8. U0126 is more potent than shMEK in the presence of ERK-Raf feedback. (A-C) Stable 3T3 cell lines (shGFP or shMEK#1) were serum starved for 24 hours and pre-treated with 2 μM U0126 (B) or DMSO (A,C) for one hour and stimulated with 1 ng/ml PDGF for the indicated times. (D) Quantification of signaling timecourses in A-C for three replicates.

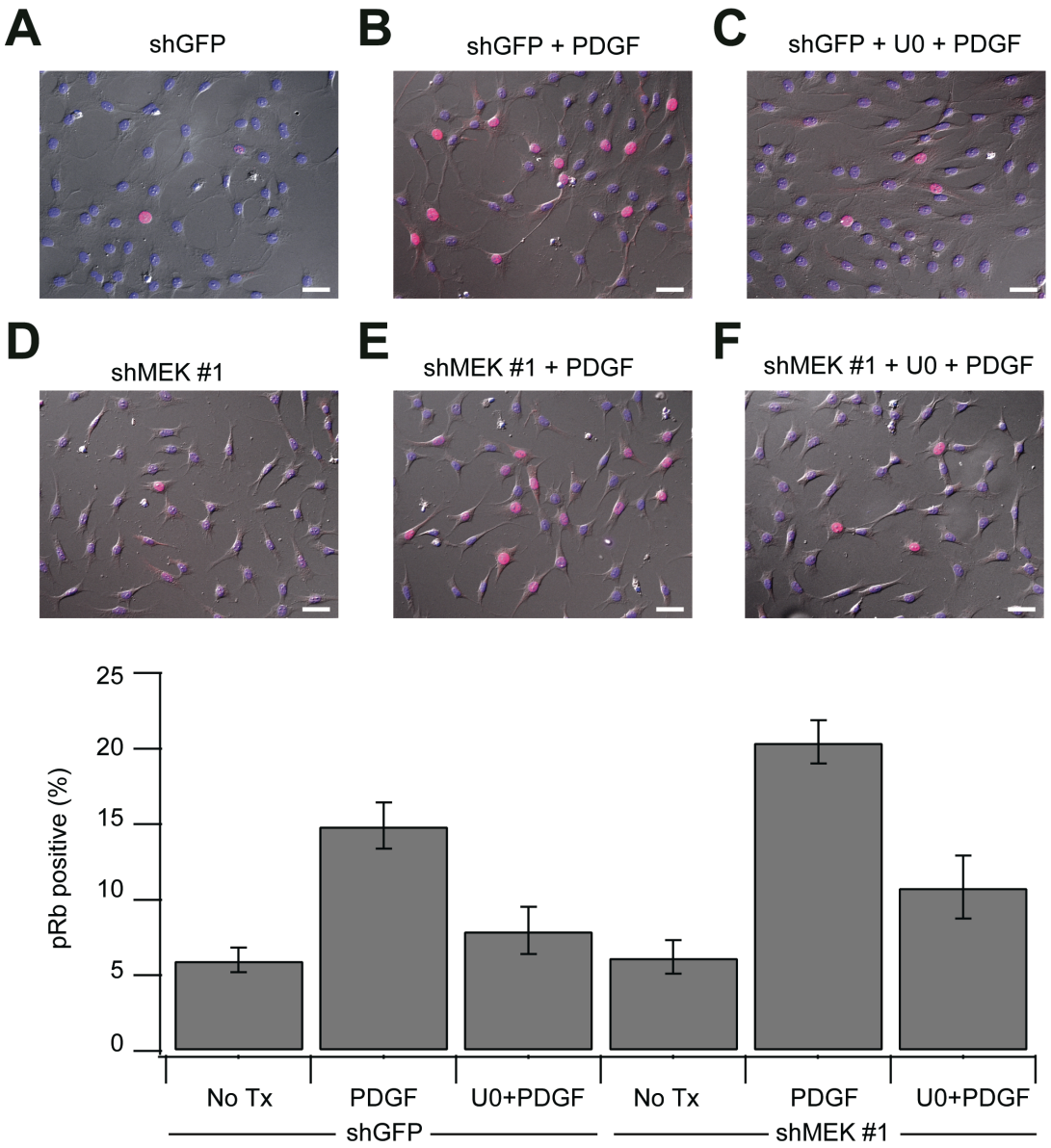


Figure 3-9. MEK1/2 inhibition by RNAi knockdown does not prevent proliferation. Stable 3T3 cell lines (shGFP or shMEK #1) were serum starved for 24 hours and pre-treated with 2 μM U0126 (C,F) or DMSO (A, B, D, E) for one hour and stimulated with 1 ng/ml PDGF for 24 hours. Cells were stained for phosphorylated Rb (pink) and nuclei were labeled by DAPI (blue). Scale bar is 40 μm. (D) Quantification of cells with nuclear phosphorylated Rb staining for three replicates.

Sustained phospho-ERK1/2 signaling has been shown to lead to cell proliferation, while transient activation does not (*166*). We examined changes in proliferation by phospho-RB immunofluorescence and found that U0126-treated cells did not proliferate

in response to PDGF, while shMEK1/2 cells show strong nuclear phospho-RB staining (Figure 3-9). This suggests that the attenuation of PDGF-induced phospho-ERK1/2 in shMEK1/2 cells is not sufficient to perturb proliferative responses to growth factor, despite substantial reductions in MEK enzyme levels (Figures 3-5C and 3-10A).

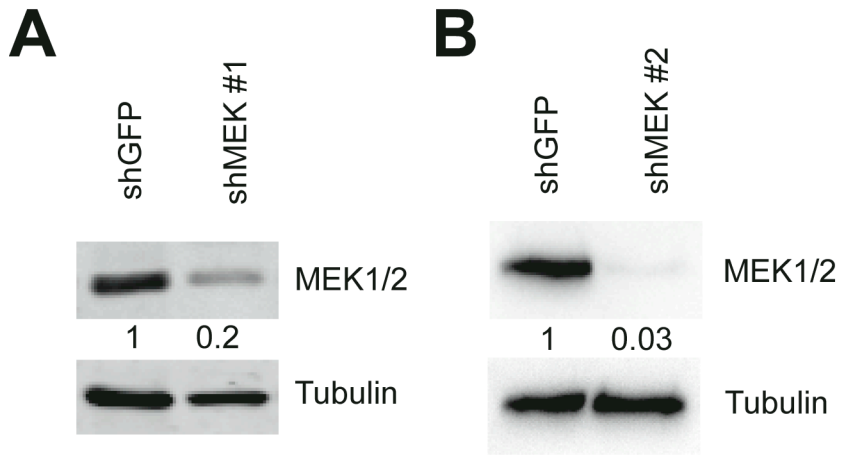


Figure 3-10. MEK1/2 hairpins #1 and #2 knockdown total MEK1/2 levels 80% and 97%, respectively. (A) shMEK #1, repeated from Fig. 3-5C as reference. (B) shMEK #2.

Whether this effect was specific to the level of inhibition or more general to behavior of MEK1/2 knockdown was tested using a second hairpin of MEK1/2 (shMEK1/2 #2). Our modeling predicts that at inhibitions near 0% or 100%, the results for the small-molecule inhibition and RNAi knockdown will converge and have the same effect on pathway output (Figure 3-2A). To test this hypothesis, we developed a second stable cell line expressing a second hairpin against MEK1/2 that was generated using two rounds lentiviral infection. This second hairpin reduced levels to 3% of control lines (Figure 3-10B). The potent MEK1/2 knockdown eliminated sustained phospho-ERK signaling relative to control (Figure 3-11, A and C), more closely resembling the result seen with U0126 treatment (Figure 3-11B). Cells expressing the more potent MEK1/2

hairpin (shMEK1/2 #2) showed reduced nuclear phospho-RB staining compared to control (Figure 3-12). This suggests that at high enough levels of MEK1/2 reduction, the RNAi will behave similar to the inhibitor.

In parallel with our experimental studies, we are also updating a specific mechanistic model of the Raf–MEK–ERK pathway. We began with a published model (*150*) and added the negative feedback from ERK to Raf that was not included in the original work. Our preliminary modeling results are consistent with our initial three-enzyme simulations, further supporting that the Raf–MEK–ERK pathway wiring is more sensitive to small-molecule inhibition than to RNAi.

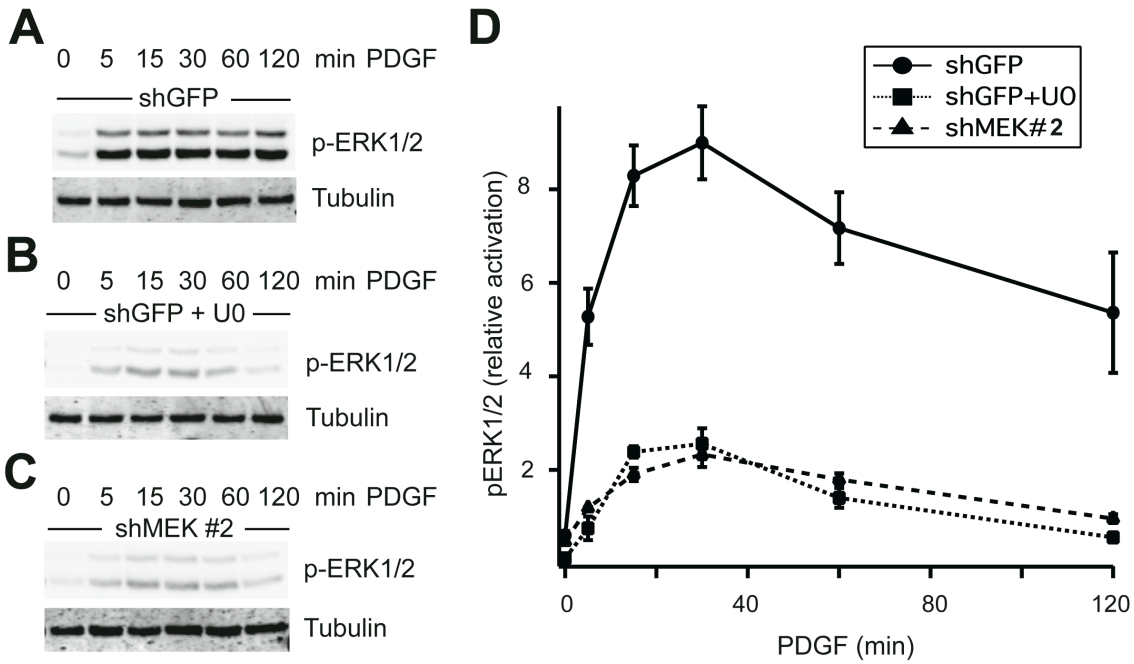


Figure 3-11. Potent MEK1/2 knockdown eliminates sustained phospho-ERK signaling in response to PDGF. (A-C) Stable 3T3 cell lines (shGFP or shMEK#1) were serum starved for 24 hours and pre-treated with 2 μM U0126 (B) or DMSO (A,C) for 1 hour and stimulated with 1 ng/ml PDGF for the indicated times. (D) Quantification of signaling timecourses in A-C for three replicates.

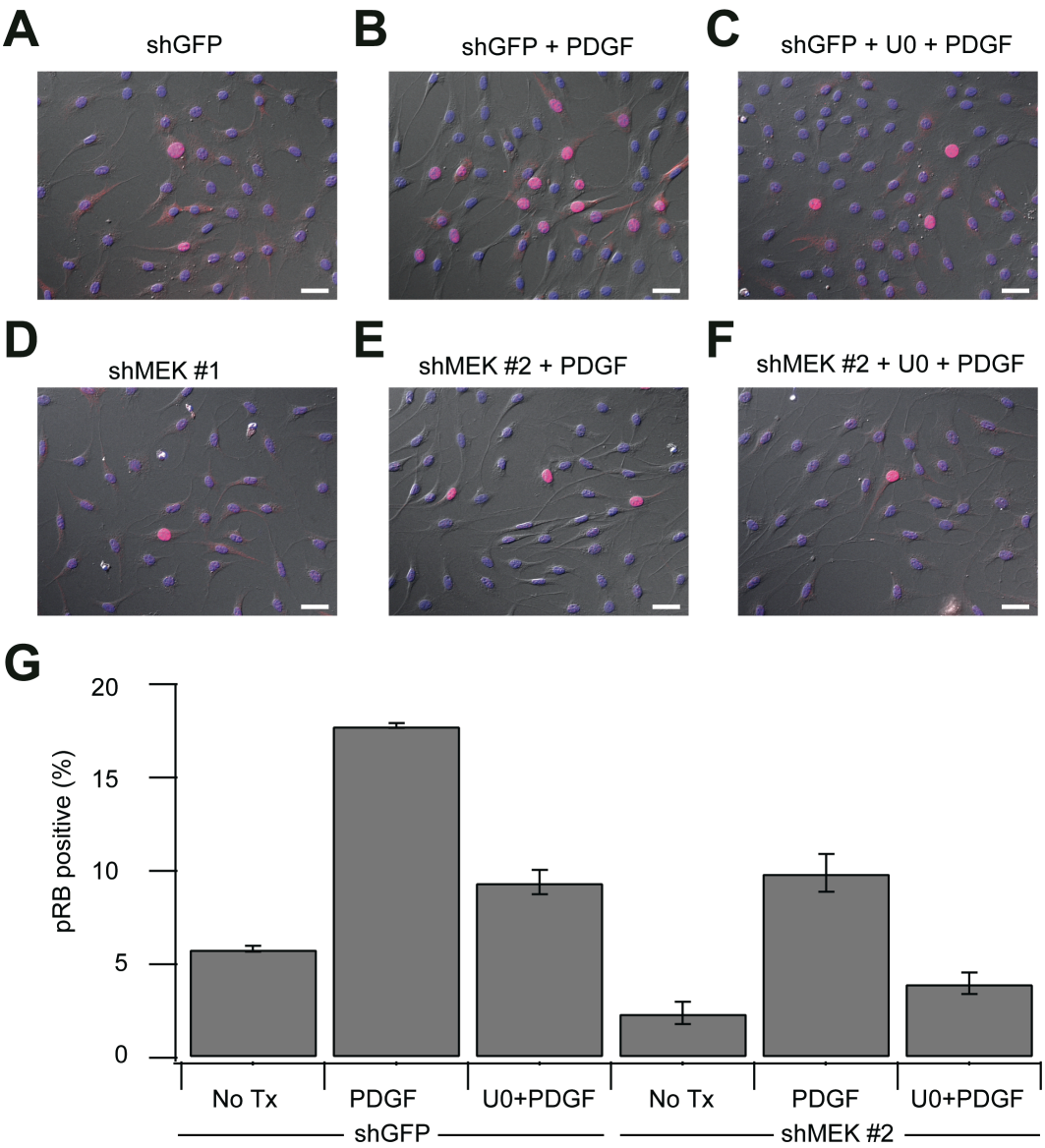


Figure 3-12. Potent MEK1/2 inhibition by RNAi knockdown prevents proliferation. Stable 3T3 cell lines (shGFP or shMEK #1) were serum starved for 24 hours and pre-treated with 2 μM U0126 (C,F) or DMSO (A, B, D, E) for one hour and stimulated with 1 ng/ml PDGF for 24 hours. Cells were stained for phosphorylated Rb (pink) and nuclei were labeled by DAPI (blue). Scale bar is 40 μm. (D) Quantification of cells with nuclear phosphorylated Rb staining for three replicates.

3.6.2. The caspase-8–caspase-3–caspase-6 pathway

Our modeling predicts that in a three-node pathway where the third element feeds back positively on the first, RNAi inhibition of the second element will inhibit pathway output more potently than small-molecule inhibition (Figure 3-13A). An example of this wiring is the caspase-8–caspase-3–caspase-6 pathway, where caspase-6 positively feeds back on caspase-8 to drive caspase cleavage and apoptosis in some cells (Figure 3-13B) (*167*).

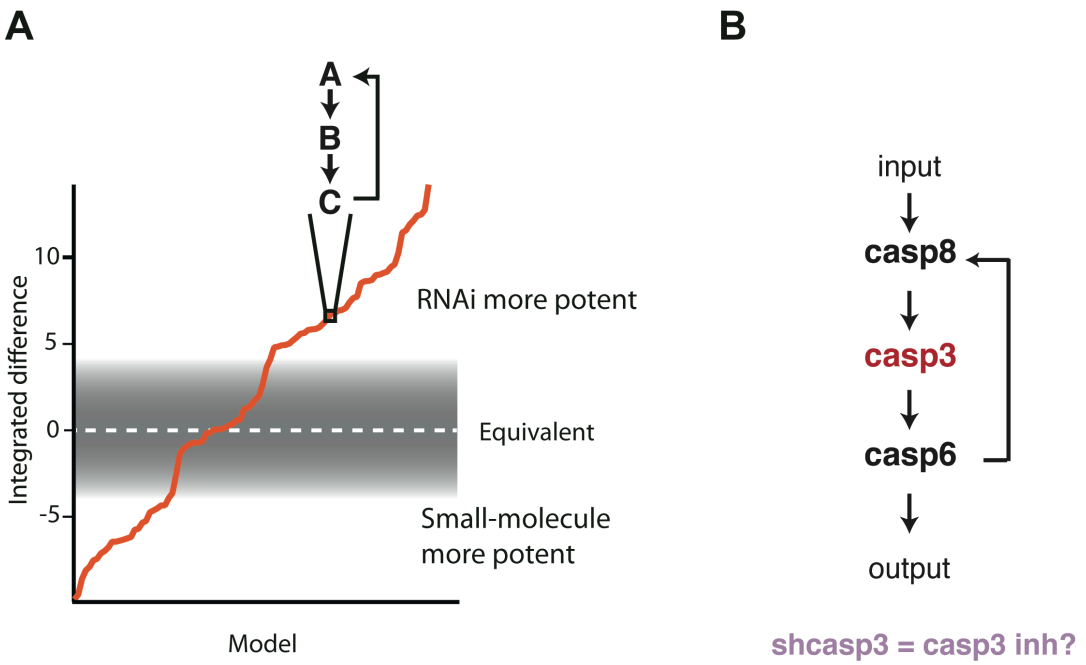


Figure 3-13. The caspase-8—caspase3—caspase-6 pathway is predicted to be more effectively inhibited by RNAi than small-molecules. (A) The pathway wiring where enzyme C positively feeds back on enzyme A is predicted to be more potently inhibited by RNAi. (B) The caspase-8—caspase3—caspase-6 cascade is a pathway that mirrors the structure in (A).

The human lymphoblastoma cell line SKW 6.4 has been shown to have strong positive feedback from caspase-6 to caspase-8 (*168*). Importantly, these cells do not engage the mitochondrial arm of apoptosis in response to anti-APO treatment (*169*), thereby insulating the caspase-8–caspase-3–caspase-6 circuit from other inputs. In

addition, there exists a complete model of the caspase-8–caspase-3–caspase-6 signaling pathway (*168*) that could test predictions as described for the Raf–MEK1/2–ERK1/2 pathway in Chapter 3.6.1. These characteristics make the SKW 6.4 line ideal for testing the model prediction.

To compare a caspase-3 inhibitor to caspase-3 knockdown directly, the positive feedback from caspase-6 to caspase-8 needs to be blocked. Since inhibitors of caspase-3 and caspase-6 are promiscuous (*170, 171*) or cell impermeable (*172*), we sought to genetically engineer a SKW 6.4 line that expresses a catalytically inactive caspase-3 (Figure 3-14A). We hypothesize that if the caspase-3 (C163A) construct is expressed at high enough levels (relative to endogenous caspase-3) it will function as a dominant negative, where the cleavage of caspase-8 by caspase-6 will be blocked. Expression of caspase-3 (C163A) would then effectively abrogate the positive feedback from caspase-6 to caspase-8 (Figure 3-14A).

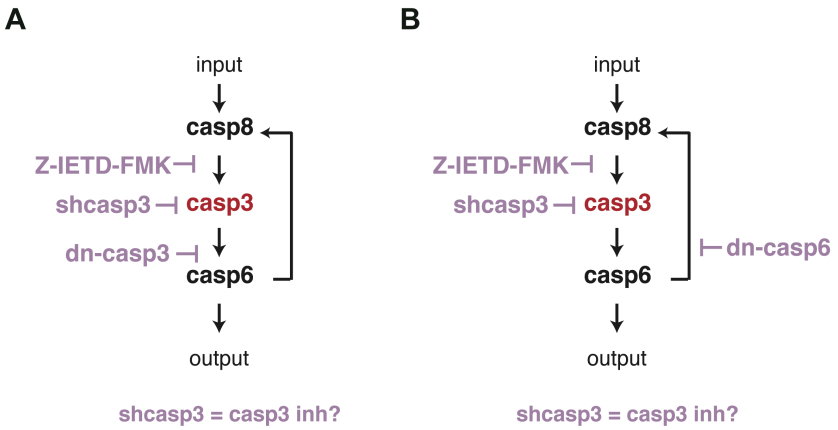


Figure 3-14. Methods to break positive feedback from caspase-6 to caspase-8. (A) Dominant-negative caspase-3 (dn-casp3) will prevent caspase-3 cleavage of caspase-6 (casp6) and break the positive feedback from caspase-6 to caspase-8 (casp8). (B) Dominant-negative caspase-6 (dn-casp6) will prevent caspase-6 cleavage of caspase-8 and break the positive feedback from caspase-6 to caspase-8.

A stable SKW 6.4 line was generated to express a dominant-negative caspase-3 construct driven by a cytomegalovirus promoter (CMV) (pLNCX2 vector), but expression was poor (Figure 3-15A). Construct expression did not improve in lines with an additional round of lentiviral infection (Figure 3-15A, 1X versus 2X). Low detection of the dominant-negative caspase-3 construct was not due to differential recognition by the total caspase-3 antibody caused by the mutation, as a second caspase-3 antibody directed against a site distant to the mutation detected similar levels (antibody #1 versus antibody #2, Figure 3-15A).

Since constitutive expression of the dominant-negative caspase-3 construct could negatively affect the selection or growth of cells, a dominant-negative caspase-3-3XFLAG was cloned into a vector driven by a doxycycline (DOX) inducible promoter. Stable SKW 6.4 cell lines were developed and induced for 24 hours. DOX induction did

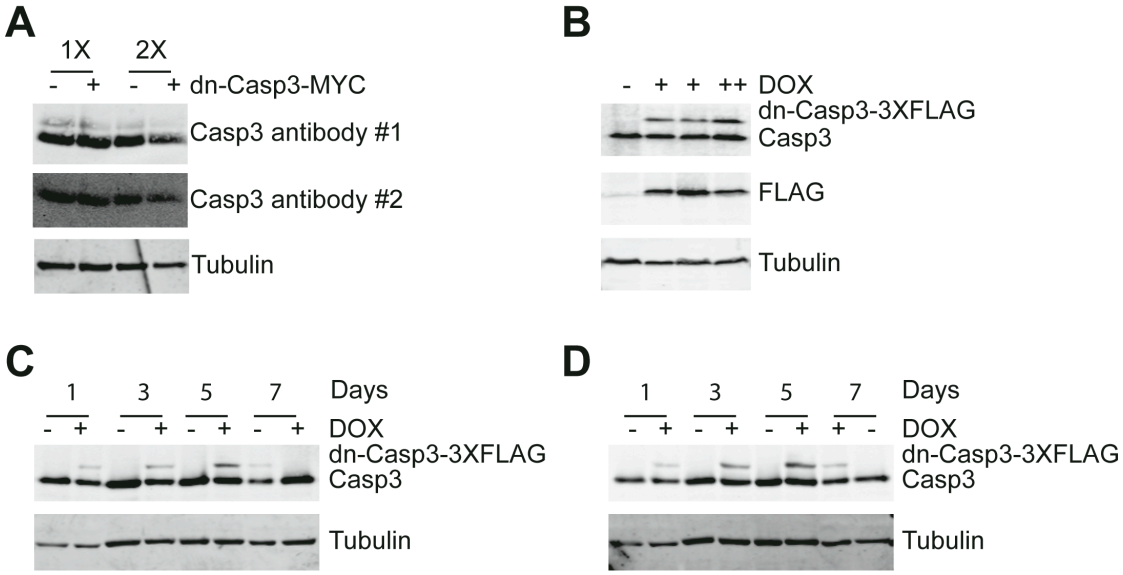


Figure 3-15. Stable SKW 6.4 cell lines expressing dominant-negative caspase-3 constructs did not achieve overexpression. (A) SKW 6.4 cells were stably infected once or twice (1X versus 2X) with pLNCX2 vector containing dominant-negative capase-3-MYC (dn-Casp3-MYC) sequence or empty pLNCX2 vector. Caspase 3 levels were monitored with two antibodies recognizing different regions of the protein. (B) SKW 6.4 cells expressing a doxycycline (DOX) inducible dominant-negative caspase-3-3XFLAG (dn-Casp3-3XFLAG) were induced with 1 μg/ml (+) or 2 μg/ml (++) DOX for 24 hours. (C) SKW 6.4 cells stably expressing DOX-inducible RNAi-resistant dn-Casp3-3XFLAG and shcaspase-3 #1 (C) or #2 (D) constructs were induced with 1 μg/ml DOX for 1, 3, 5, and 7 days.

not achieve overexpression of dominant-negative caspase-3 and doubling the standard concentration of DOX did not appreciably increase construct expression (Figure 3-15B). Increasing DOX stimulation time also did not increase expression levels (data not shown).

Since caspase-3 overexpression was not achieved by constitutive or DOX-inducible expression, we next sought to knockdown endogenous caspase-3 expression while simultaneously expressing an RNAi-resistant dominant-negative caspase-3. Stable SKW 6.4 lines were generated to inducibly express both an RNAi-resistant dominant-negative caspase-3 and caspase-3 hairpin in response to DOX. Two separate lines were generated with two unique hairpin sequences directed at caspase-3 with corresponding RNAi-resistant dominant-negative caspase-3. Dual expression did not improve ratio of dominant-negative caspase-3 relative to endogenous levels (Figure 3-15, C and D).

Since promoter strength and duration of induction did not improve expression (Figure 3-15), it is possible that the expression limitations are inherent to the caspase-3 construct. Protein expression can be limited by rare codon usage (*173*). In the caspase-3 sequence, the third position of codons is more frequently A/T (63.31%) than G/C (36.69%), which corresponds to rarer codons that can hamper protein translation (*174*). Optimization of the caspase-3 sequence for codon usage could potentially increase expression levels, but would require significant mutations to the constructs. Alternatively, negative regulators of caspase-3 transcript or protein, such as ubiquitination, could be limiting total caspase-3 protein levels. It is unclear at this time whether overexpression limitations are specific to SKW 6.4 cells, caspase-3, or the experimental systems used.

While the desired overexpression of dominant-negative caspase-3 was not achieved, it was directly tested whether the levels were sufficient to inhibit downstream signaling, indicating abrogation of the positive feedback loop from caspase-6 to caspase-8. Levels of dominant-negative caspase-3 were not sufficient to block downstream apoptotic signaling in response to three doses of anti-APO as demonstrated by PARP and lamin A cleavage (Figure 3-16). Unexpectedly, DOX-induction of the dominant-negative caspase-3 construct appeared to increase downstream apoptotic signaling in two doses of anti-APO tested (Figure 3-16). Further, selection of cells expressing higher levels of dominant-negative caspase-3 by flow cytometry (as measured by FLAG immunoreactivity) did not result in increased resistance to anti-APO stimulus.

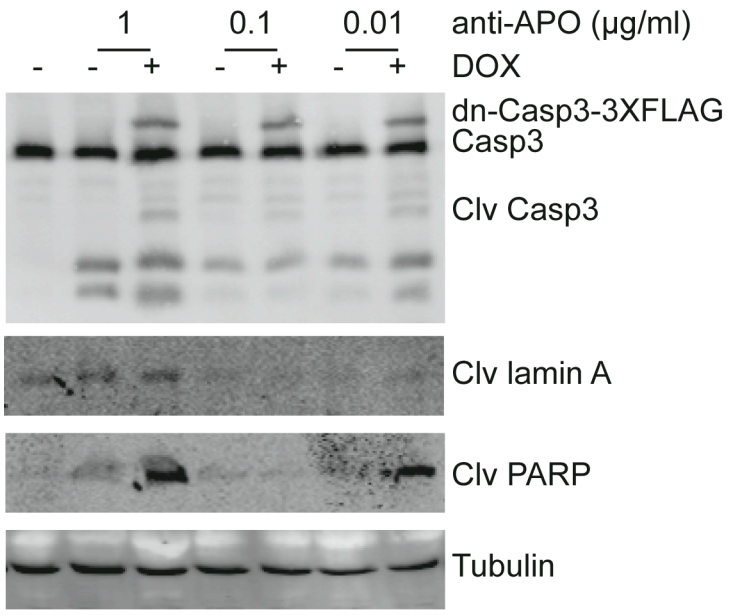


Figure 3-16. Dominant-negative caspase-3 did not block downstream apoptotic signaling. SKW 6.4 cells stably expressing the DOX-inducoble dominant-negative caspase-3-3XFLAG (dn-Casp3-3XFLAG) were induced for 24 hours with 1 µg/ml DOX. Following DOX induction, cells were treated with the indicated dose of anti-APO for 30 minutes before cell lysis.

A second option to disrupt the positive feedback loop from caspase-6 to caspase-8 is overexpression of dominant-negative caspase-6 (C163A) (Figure 3-14B). Lentiviral vectors were generated with dominant-negative caspase-6 driven by a CMV promoter (pLX302 vector). The lentivirus packaged did not produce stable cell lines for dominant-

negative caspase-6 or control (Luciferase) constructs using our standard protocol for 1X and 0.5X virus. Extended recovery time before antibiotic selection did not increase survival. Additionally, toxicity due to the polybrene was not evident over 48 hours in the dose used in the protocol (8 μg/ml). While it is presently unclear why these constructs failed to generate stable lines, it may be necessary to optimize an alternative protocol for lentivirus packaging and viral transduction for SKW 6.4 cells.

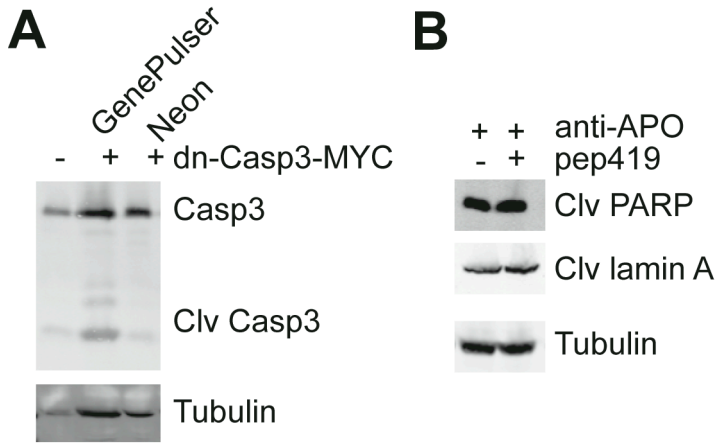


Figure 3-17. Transfection of SKW 6.4 cells by electroporation. (A) SKW 6.4 cells were electroporated with the GenePulser or Neon devices with dominant-negative caspase-3-MYC (dn-Casp3-MYC) plasmid DNA. (B) SKW 6.4 cells were electroporated with the GenePulser with or without 50 μM pep419 and stimulated with 1 μg/ml anti-APO for 30 minutes.

Efforts to transfect constructs into SKW 6.4 cells were also unsuccessful. Since SKW 6.4 cells are not readily transfected by standard lipofection methods, dominant-negative caspase-3-MYC constructs were electroporated into SKW 6.4 cells using the BioRad GenePulser XCell and Invitrogen Neon systems according to the manufacturer's

protocols. Additionally, the cell-impermeable caspase-6 specific inhibitor, pep419 (*172*), was electroporated into SKW 6.4 cells. No detectable dominant-negative caspase-3 was detected in electroporated cells (Figure 3-17A). Cells electroporated with the BioRad instrument appeared to be apoptotic as evidenced by caspase-3 cleavage (Figure 3-17A). Electroporated cells stimulated with anti-APO displayed high levels of cleaved lamin A and PARP, indicating that electroporation of dominant-negative caspase-3 nor pep419 was sufficient to inhibit pathway signaling (Figure 3-17B).

Since expression of dominant-negative caspase-3 did not reach high enough levels in SKW 6.4 cells via stable lines or transfections, we next used transient transfection of 293Ts, which are readily transfected. To block mitochondrial caspase signaling in 293Ts we simultaneously overexpressed Bcl2 with dominant-negative caspase-3 (Figure 3-18A). Unexpectedly, overexpression of Bcl2 increased caspase-3 cleavage in 293Ts and

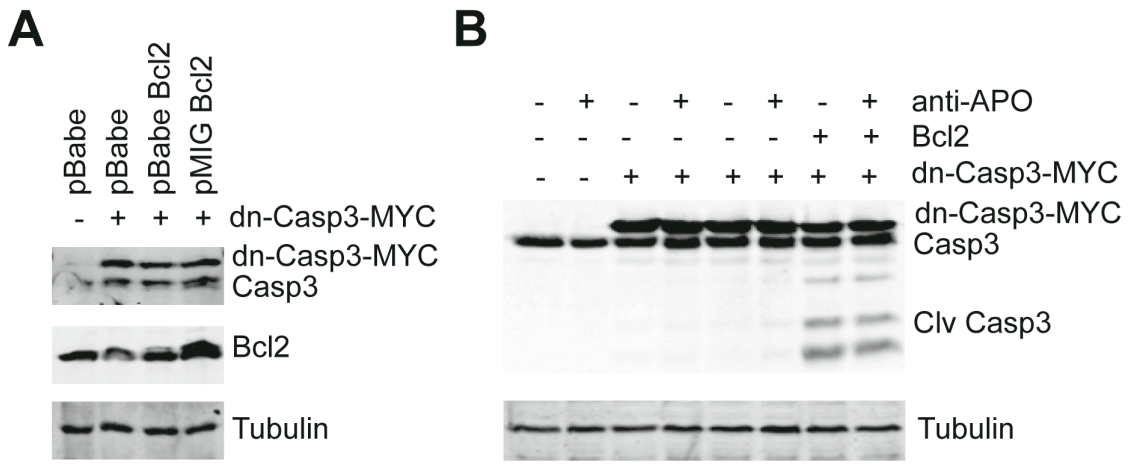


Figure 3-18. Dual transient transfection of dominant-negative caspase-3 and Bcl2 into 293T cells increased apoptotic signaling. (A) Bcl2 and dominant-negative caspase-3-MYC expression was obtained by simultaneous transfection. 293T cells were transiently transfected by lipofection with dominant-negative caspase-3-MYC or empty PLNCX2 vector and Bcl2 in pBABE or pMIG plasmids or pBABE empty vector. pMIG Bcl2 plasmid yielded higher Bcl2 expression. (B) Overexpression of Bcl2 induced caspase-3 cleavage compared to cells transfected with empty vector. Transiently transfected 293T cells were treated with 100 ng/ml anti-APO for 24 hours.

293Ts did not respond to moderate (100 ng/ml) anti-APO stimulation for 24 hours. Taken together, these results suggested that 293Ts might not be a useful system to study the caspase-8—caspase-3—caspase-6 pathway in isolation.

Increased caspase-3 cleavage was observed in SKW 6.4 and 293Ts expressing the dominant-negative caspase-3 construct. These observations raised the intriguing possibility that the construct itself could be inducing apoptotic signaling. To test this, 293Ts were transiently transfected with dominant-negative caspase-3-MYC or empty pcDNA3 plasmid and treated with varying doses of anti-APO for 30 minutes or 3 hours. Caspase-3 cleavage was evident in all samples transfected with dominant-negative caspase-3-MYC but not in those transfected with empty pcDNA3 vector (Figure 3-19A), suggesting again that 293Ts were not sensitive to anti-APO. Cleaved dominant negative caspase-3 is catalytically inactive and should not induce downstream apoptotic signaling. However, PARP cleavage seemed to increase in samples transfected with dominant-negative caspase-3-MYC (Figure 3-19A), indicating that overexpression of the dominant-negative caspase-3-MYC could be initiating apoptotic signaling. If overexpression of dominant-negative caspase-3-MYC drives apoptotic signaling, it could be preventing overexpression in stable cell lines (Figure 3-15).

To check if this observation was a general overexpression artifact, 293Ts were transiently transfected with control empty vector or Luciferase plasmids, dominant-negative caspase-3 and caspase-6 constructs, or wild-type caspase-6 (Figure 3-19B). Caspase-3 cleavage and PARP cleavage was only evident in cells transfected with the dominant-negative caspase-3 construct. The same constructs were then transiently transfected into 3T3 cells to test the generality of this observation across cell types

(Figure 3-19C). Levels of dominant-negative caspase-3-MYC relative to endogenous caspase-3 were much lower compared to 293T transfections, indicating that transfection conditions may need to be optimized for 3T3 cells. Caspase-3 cleavage was apparent in 3T3 cells transfected with dominant-negative caspase-3-MYC but not with other constructs. However, PARP cleavage was not detected in any 3T3 samples (Figure 3-19C). While these results are single observations, they raise an interesting question about cell response to caspase-3 overexpression and possibly address the difficulty obtaining caspase-3 overexpressing stable cell lines.

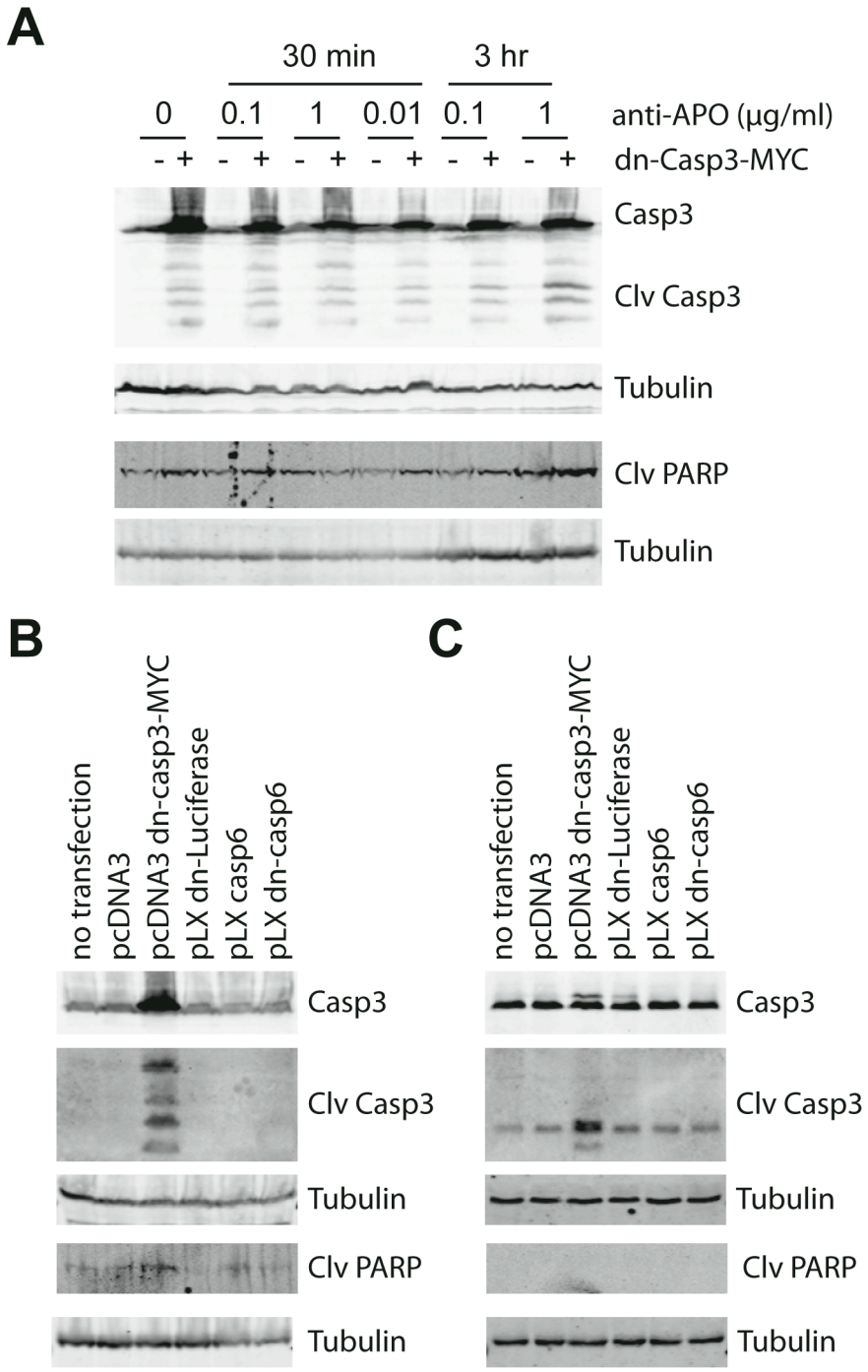


Figure 3-19. Transient transfections of dominant-negative caspase-3 into 293T and 3T3 cells increases caspase-3 cleavage. (A) 293Ts do not respond to 30 min or 3 hour exposure to anti-APO but show increased caspase-3 cleavage and PARP cleavage when transfected with dominant-negative caspase-3-MYC compared to empty pcDNA3 vector. 293Ts (B) and 3T3s (C) were transiently transfected with the indicated constructs and lysed 48 hours after transfection.

3.7. Summary

Cell-signaling networks integrate signals transmitted from receptors (Chapter 2) and process the information to make decisions about cell fate. These networks are wired so a cell can appropriately respond to a specific cue or a complex combination of cues that change over time. In contrast to simple linear cascades stemming from individual receptors, cell-signaling networks are understood to be increasingly complex and interconnected. Despite this appreciation, little work has been devoted to considering how this architecture may complicate our interpretation of experimental data. In this chapter we quantitatively modeled three-tier enzyme cascades to define the effect of local wiring on perturbations of protein amount (by RNAi) and activity (by small-molecule inhibitors). We found that local connectivity is critical in determining the behavior of a pathway in response to these mechanistically different perturbations. We believe that computational modeling of local network connectivity could identify druggable signaling nodes and prioritize drug targets, lowering the time and cost of drug development.

CHAPTER 4

A data-driven model of CVB3 infection of cardiomyocytes predicts therapeutic targets

4.1. Introduction

Connectivity within the cell-signaling network is an important consideration in determining the effect of a perturbation and thus in identifying a drug target (Chapter 3). But how can we determine unknown network topology from experimental datasets? Signaling through different pathways is not independent, because some are connected through shared components and others jointly converge upon common cellular functions. This coupling creates a hidden structure within the signaling network that regulates higher-level organizing principles of the system and how it responds to complex perturbations, such as a pathogen.

The effects of pathogens on cells are multifaceted, making it difficult to link specific stimuli or perturbations to cellular responses, and therefore difficult to determine a single protein to target with a therapy. An example of a complex perturbation is infection of a cell by a virus. Infection modifies numerous intracellular signaling pathways. These pathways work together as networks in host cells, and the resulting phenotypes are interdependent, making it difficult to link virus-induced signals and responses at a systems level. In this chapter, we apply data-driven modeling techniques to experimental data to collected from a host-cell response to a virus, Coxsackievirus B3 (CVB3), to identify how it controls the underlying signaling structure to promote its

pathogenesis (see Chapter 7.1 for a tutorial on data-driven models for cell-signaling datasets).

4.2. CVB3 and viral myocarditis

CVB3 is among the most common causes of viral myocarditis-associated heart failure in infants and young children (*175*). A major component of CVB3 pathogenesis is cell death of infected cardiomyocytes, which leads to immediate tissue damage and the subsequent release of virulent CVB3 progeny that furthers disease progression (*176*). Intervening at the early stages of CVB3 cytotoxicity could potentially reduce the severity of the disease and the need for heart transplantation in patients with viral myocarditis.

Throughout infection, CVB3 modulates various cell-signaling pathways that enable virus propagation (*175, 177*). Inhibiting these pathways may provide a therapeutic opportunity to restrict CVB3 pathogenesis. But, an important hurdle is that our understanding of how the CVB3 infection cycle intersects with the host network is fragmentary. Viruses such as CVB3 have evolved to modulate cell-signaling networks in ways that allow them simultaneously to evade host defenses, promote cell entry, and undergo replication in a changing environment (*175, 178*). Blocking individual signaling pathways in host cells often reduces CVB3 infectivity but does not prevent infection entirely (*177*). It remains unclear whether such "partly required" pathways converge upon a common set of host effectors or instead make independent contributions to pathogenesis (*134*). The challenge is that CVB3 adaptively perturbs a collection of host pathways, which must be examined concurrently with time to understand how they interact and give rise to viral functions.

4.3. CVB3-Induced Phospho-Protein Dynamics Quantitatively Predict Host-Cell Outcomes

To determine whether known CVB3-induced signaling events were sufficient to predict viral propagation and host-cell toxicity, we sought to build a predictive mathematical model based entirely on quantitative experiments. Data-driven modeling identifies higher-order statistical covariations that can be used for prediction and analysis (*179*). Unlike other modeling formalisms (*180*), data-driven approaches can accommodate variegated datasets and make predictions without detailed knowledge of the underlying biochemical mechanisms.

To build the model, we systematically assembled a host-cell signaling and response dataset in virus-infected cardiomyocytes. At five different CVB3 multiplicities of infection (M.O.I.), we profiled eight signaling phospho-proteins by ELISA at six time points over 24 hr together with six CVB3-induced host-cell readouts at three time points over 24 hr (Figures 4-1A and 4-1B). Each phospho-protein or host-cell readout was selected based on previous studies suggesting that they were critical during CVB3 pathogenesis (Table 4-1). Analyzing the information contained in this mechanism-rich signature would then allow us to examine how host-cell pathways are coordinately perturbed during CVB3 infection.

Table 4-1. Literature support for the CVB3-induced phospho-proteins and active caspases.

Phospho-proteins and caspases	Contributes to	
Caspase-2, -3, -6, -7, -8 & -9	Late-stage alterations of cellular homeostatic processes and structural integrity	(*2*)
Akt–GSK3β	Successful virus replication	(*3*)
ERK1/2	Effective virus replication and virus-mediated cytotoxicity in host cells	(*1, 4*)
CREB	Gene expression during infection	(*7*)
p38–Hsp27	Effective viral progeny release, cytotoxicity and virus-induced caspase-3 activation	(*8*)
ILK (Integrin-linked kinase)-Akt	Supporting virus infection	(*9*)
IκBα–NF-κB	Promoting virus-infected host survival	(*10*)
p38	Underpinning virus replication	(*1, 11*)
ATF2	Gene expression during infection	(*12*)

We found that CVB3-induced host-cell responses showed time and dose dependencies that were expected for end-stage readouts (Figure 4-1B). Activation of the initiator caspases, caspase-8 and caspase-9, was accelerated with increasing M.O.I., corresponding to more-complete activation of the effector caspase for apoptosis, caspase-3 (*181*). Interestingly, readouts of CVB3 propagation, such as expression of the VP1 capsid protein and the titers of released viral progeny (RVP), did not accelerate appreciably as they increased with CVB3 M.O.I.. This finding suggests intrinsic limits to the timing of the CVB3 replication cycle downstream of the M.O.I.-dependent rate processes of viral docking and internalization. The pattern of overall CVB3 cytotoxicity fell in between that of caspase and viral readouts, showing some acceleration as host-cell viability dropped with increasing CVB3 titers. Thus, CVB3 infection of cardiomyocytes elicits a collection of host-cell and viral phenotypes that are monotonic in time but differ in their kinetics and dose-dependent behaviors.

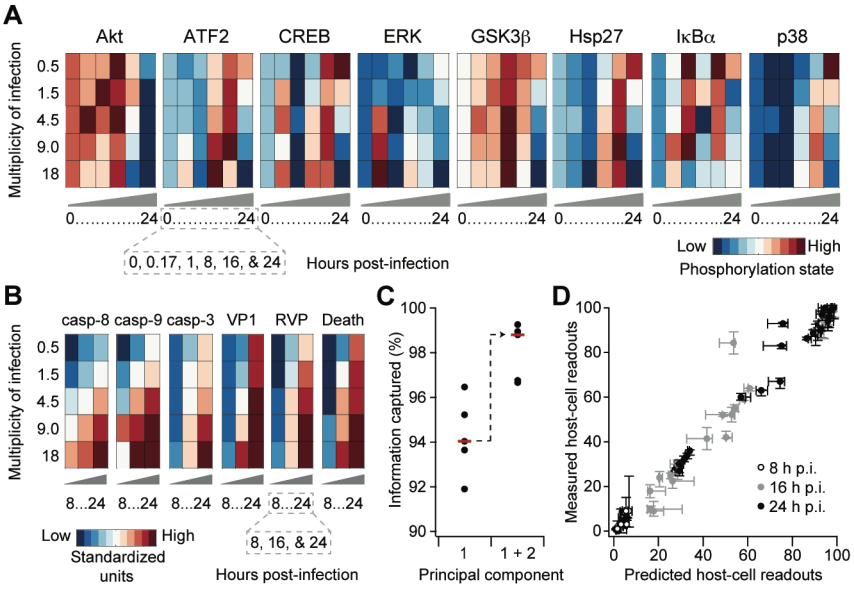


Figure 4-1. A predictive data-driven model of CVB3-induced host-cell responses. (A) Dynamic phospho-proteins signatures measured by phospho (p)-ELISA that were used as predictor variables in the data-driven model. (B) Host-cell outcomes that were to be predicted in the data-driven model. Caspase (C)-3, -8 and -9 activities were measured by activity assays with fluorogenic substrates, VP1 capsid protein expression was measured by immunoblotting, released viral progeny (RVP) titer was measured by plaque assay, and cell death was measured by MTS assay at the indicated time points. (C and D) Accurate predictions of host-cell responses with a partial least squares model using two principal components. (C) Percentage of information captured with one or two principal components. Information was measured by the percentage of variance in host-cell outcomes that was captured by the model. Note the small-but-significant increase in information capture after inclusion of the second principal component (arrow). (D) Correlation between cross-validated predictions of biological responses by partial least squares regression (x-axis) and observed biological responses (y-axis). Marker color corresponds to the post-infection (p.i.) time point at 8 (white), 16 (gray), and 24 (black) hours. HL1 cells were infected with CVB3 at one of five multiplicities of infection and then assessed for the indicated phospho-proteins and biological responses at six and three time points, respectively, over 24 hr. For (A) and (B), data are shown as the z-score standardized mean of three independent experiments as described in the Experimental Procedures. For (C), data are shown as median information captured (red) after fivefold leave-one-out cross-validation (black). For (D), data are shown as the median ± range of 3–4 biological replicates (vertical) or four model cross-validation runs (horizontal).

By comparison, we found that the dynamic patterns of protein phosphorylation stimulated by CVB3 were substantially more complex than the associated phenotypic readouts (Figure 4-1A). As before, we observed accelerated phosphorylation of some CVB3-induced pathways with increasing M.O.I., such as p38 and Hsp27, but not others, such as ERK (Figure 4-1A). In addition, biphasic activation patterns were common, and

many individual activation peaks appeared or disappeared above a critical threshold of CVB3 M.O.I. (e.g., ATF2, CREB, and IκBα). Sham infection with 0 M.O.I. did not lead to any meaningful changes when compared to 0.5 M.O.I. (Figure 4-2), confirming that the measured signaling events were due to CVB3 infection. The internal consistency of our phospho (p)-ELISA measurements was also verified by the strong correlations between p-Akt and p-GSK3β (R = 0.6), a direct substrate of Akt (*182*), and between p-p38 and p-Hsp27 (R = 0.8), a direct substrate of the MK2 kinase that is a substrate of p38 (*183*). The p-ELISA signatures thus provided a reliable starting point for connecting CVB-induced signaling to host-cell outcomes.

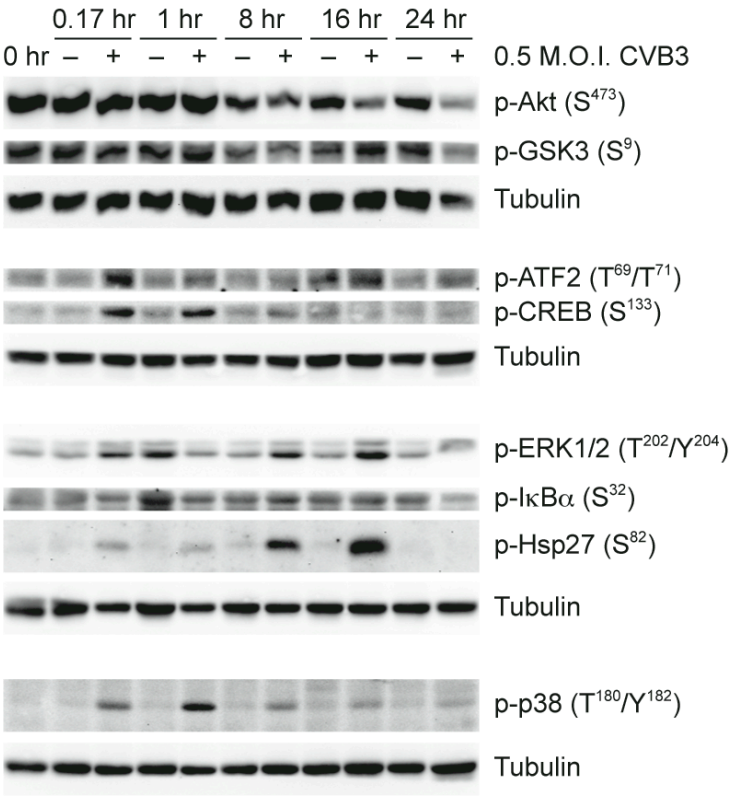


Figure 4-2. Sham infection negligibly influences phosphoprotein signatures compared to CVB3 infection at M.O.I. = 0.5. HL1 cells were infected with sham or CVB3 at M.O.I. = 0.5 for the indicated times. Samples were analyzed by immunoblotting for the indicated phospho-proteins with tubulin used as a loading control. The phospho-epitopes correspond to the sites measured by phospho-ELISA in Figure 4-1A.

One way that the observed CVB3-induced pattern of readouts could be coordinated is if each phospho-protein contributed incrementally to the pattern based on its extent of phosphorylation. Host cells would then "integrate" the intracellular state established by the level of CVB3 infection and gauge their responses accordingly. To test the feasibility of this network mechanism, we used partial-least-squares modeling to link linear combinations of measured phospho-proteins to observed CVB3-induced

readouts (*19, 179*). In a partial-least-squares model, linear combinations take the form of principal components, which are latent dimensions in the underlying dataset that are derived to be optimally efficient at predicting response outcomes (*184*).

To build the model, we first subdivided the phospho-protein time courses into early (0–8 hr) and late (8–24 hr) phases and then time-integrated each early and late phospho-protein measurement for every CVB3 M.O.I.. This subdivision allowed us to separate biphasic activation profiles into early and late peaks. Using the phospho-protein data as a set of predictor variables, we next sought a partial-least-squares model that could predict all of the CVB3-induced readouts accurately and simultaneously. We found that a model with the two leading principal components could capture all of the measured readouts to within 97% accuracy (Figure 4-1C). Importantly, this model also accurately predicted readouts for individual M.O.I. conditions that were left out of the model training during crossvalidation (Figure 4-1D). The model thus supported a network mechanism in which multiple intracellular pathways work together by independently contributing to CVB3-induced readouts.

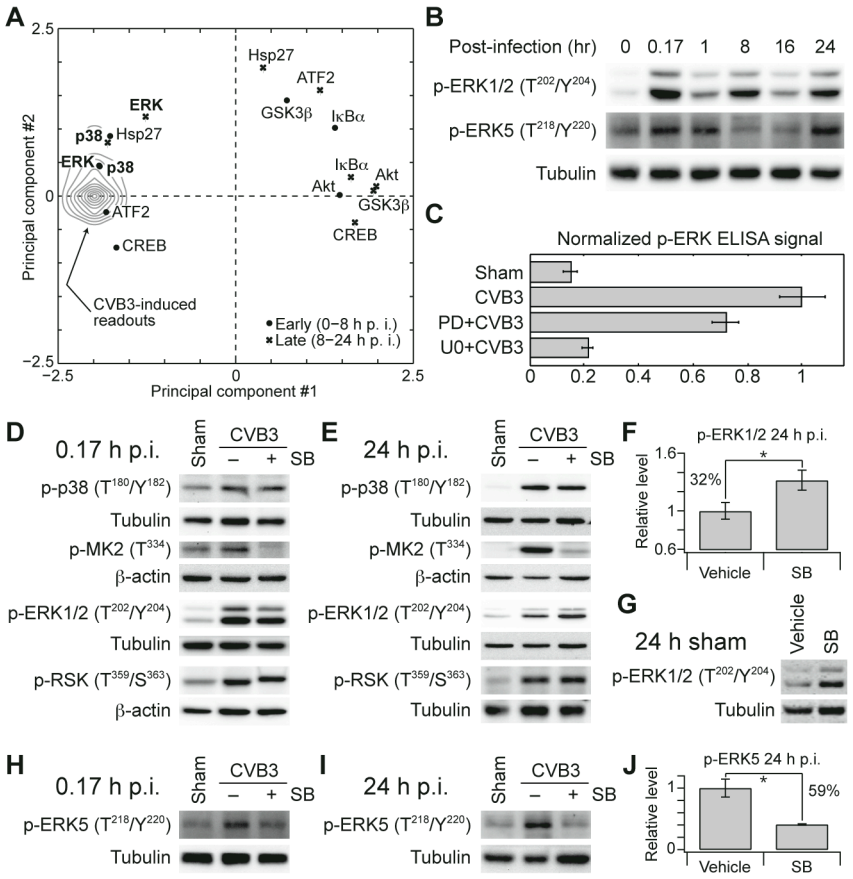


Figure 4-3. Model principal components identify crosstalk between ERK and p38 pathways. (A) Model projections of early phospho-proteins (0–8 h p.i., circles), late phospho-proteins (8–24 h p.i., clubs), and host-cell responses (contours) onto the principal components derived in Figure 1C. (B) Dynamics of phospho (p)-ERK1/2 and p-ERK5 over a 24-hr time course of CVB3 infection. HL1 cells were infected with CVB3 at M.O.I. = 9 and then assessed for p-ERKs at the indicated times p.i. by immunoblotting with tubulin used as a loading control. (C) ERK p-ELISA measurements are a convolution of p-ERK1/2 and p-ERK5. HL1 cells were pretreated with DMSO, PD to inhibit p-ERK1/2, or U0 to inhibit p-ERK1/2 and p-ERK5 and then infected with sham or CVB3 at M.O.I. = 9. The p-ERK signals were assessed by p-ELISA at 0.17 hr p.i. (D and E) Early- and late-phase activation of p-p38, p-ERK1/2, and effector kinases upon pretreatment with SB and infection with CVB3. (F) Densitometry of late-phase p-ERK1/2 in response to SB. (G) Prolonged SB inhibition leads to ERK1/2 phosphorylation independently of CVB3 infection. (H and I) Early- and late-phase phosphorylation of ERK5 upon pretreatment with SB and infection with CVB3. (J) Densitometry of late-phase p-ERK5 in response to SB. For (D–J), HL1 cells were pretreated with SB203580 (SB, 20 μM) for one hour, infected with sham or CVB3 at M.O.I. = 9, and then assessed for phospho-proteins at the indicated times post-infection by immunoblotting with either tubulin or β-actin used as a loading control. For (F) and (J), densitometry measurements were normalized to CVB3-infected cells without inhibitor and data are shown as the mean ± s.e.m. of four biological replicates. Asterisk indicates $p < 0.05$ by Welch's one-sided t test.

4.4. Intracellular crosstalk between the ERK and p38 pathways

Principal components can be further analyzed by plotting the weighted linear combinations of the original measurements that provided the basis for accurate model predictions (Figure 4-3A) (*19, 179*). In this mapping, early and late phospho-proteins are depicted together with CVB3-induced readouts. Clusters of phospho-proteins and readouts indicate measurements with close association in principal-component space and highlight correlations in the data that are most worthy of follow-up experiments (*41, 184*).

Inspection of this principal-component mapping revealed that all CVB3-induced readouts were densely clustered in one region (Figure 4-3A, contours), suggesting that they were tightly coupled. Within the cluster lay the transcription factor ATF2, which is critical for CVB3 pathogenesis in vivo (*185*), and p38, a MAPK that we recently showed is the dominant ATF2 kinase during CVB3 infection (*177*). We also found p-Hsp27 in the cluster, which was expected because of its strong concordance with p-p38 (see above). Conversely, we were surprised to find p-ERK located in the cluster together with p-p38, because the ERK1/2 and p38 pathways are generally thought to be activated by distinct stimuli and often serve antagonistic functions (*186*). Nevertheless, their tight association in the model suggested that ERKs and p38 might be functionally interlinked during CVB3 infection.

An important consideration for this prediction was the high-throughput data upon which the model was founded (*179*). The commercial p-ELISA used to assemble the p-ERK dataset is marketed as specific for ERK1/2. However, ERK1 and ERK2 share ~50% identity with ERK5, a third MAPK whose regulation is distinct (*187*). All three

ERKs have a Thr–Glu–Tyr motif that is bis-phosphorylated upon activation, and the sequence surrounding this motif is so similar that many p-ERK1/2 antibodies will cross-react with ERK5 [1]. p-ERK5 cross-reactivity is readily distinguished from p-ERK1/2 during immunoblotting (ERK5 ~ 80–100 kDa vs. ERK1 ~ 44 kDa, ERK2 ~ 42 kDa), but the ELISA format cannot resolve proteins by molecular weight. Because ERK5 signaling is important for cardiovascular tissues (*188*), we decided to investigate the individual contributions of ERK1/2 and ERK5 by independent methods.

We first monitored the kinetics of ERK1/2 and ERK5 phosphorylation by blotting with antibodies that were specific for each pathway (Figure 4-3B). Both ERK1/2 and ERK5 were strongly phosphorylated shortly after CVB3 infection at 0.17 h p.i. and also after host cytotoxicity was evident at 24 h p.i. However, ERK5 showed a more-sustained phosphorylation up to 1 h p.i. and p-ERK1/2 exhibited a second peak at 8 h p.i., illustrating differences in their regulatory kinetics. The multiphase activation of ERK1/2 was further confirmed by measuring phosphorylation of RSK, a specific ERK1/2 substrate (Figure 4-4A) (*189*). Next, we used a pair of MEK inhibitors (PD184352 [PD] and U0126 [U0]) to separate the ERK1/2 and ERK5 contributions to the ERK p-ELISA. PD at low concentrations selectively blocks MEK1/2 and ERK1/2 phosphorylation, whereas U0 inhibits MEK1/2–ERK1/2 and MEK5–ERK5 equally (*190*) (Figures 4-4B and 4-4C). Thus, the contribution of ERK5 can be inferred from the difference between PD (ERK1/2 inhibition) and U0 (ERK1/2 + ERK5 inhibition). When cells were preincubated with U0 and treated with CVB3 for 10 min, we found that the measured p-ERK ELISA signal was reduced to background levels (Figure 4-3C). By contrast,

[1] K.J.J. and K.A.J. unpublished observations.

pretreatment with PD reduced the ELISA signal by only ~30%, despite that ERK1/2 phosphorylation was completely inhibited (Figures 4-3C and 4-4C). This indicated that the p-ERK ELISA data was a convolution of ERK1/2 and ERK5 pathway activities and further implied that the predicted ERK–p38 associations (Figure 4-3A) could be between ERK1/2 and p38 or ERK5 and p38, or both.

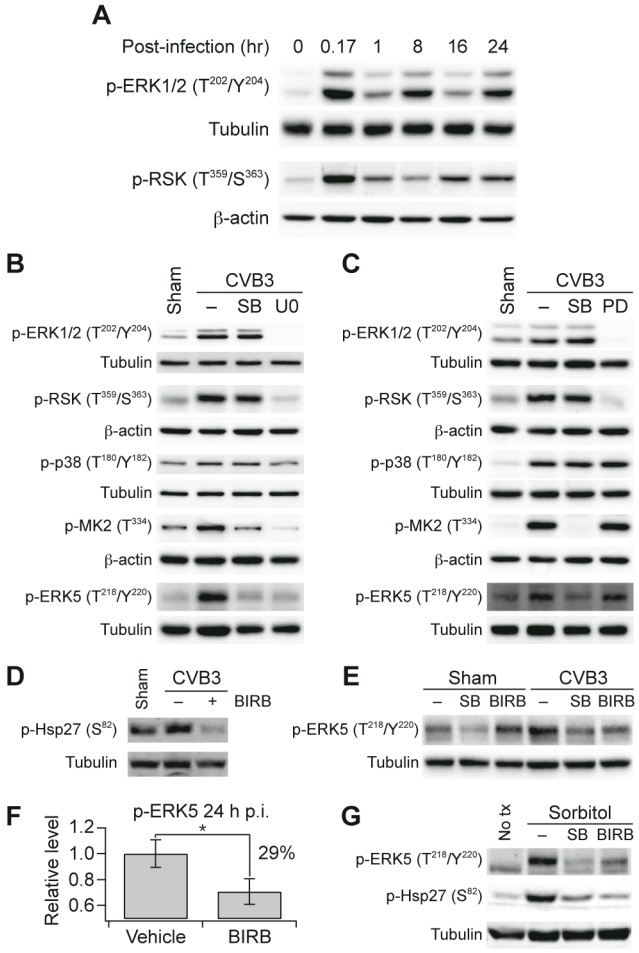


Figure 4-4. Viral and pharmacological dependencies of ERK5 phosphorylation. (A) CVB3 infection causes multiphasic activation of ERK1/2 as monitored by phospho (p)-RSK. HL1 cells were infected with CVB3 at M.O.I. = 9 and then assessed for the ERK1/2 substrate p-RSK at the indicated times p.i. by immunoblotting with tubulin used as a loading control. p-ERK1/2 is reprinted from Figure 2B for comparison. (B) U0 inhibits CVB3-induced ERK1/2 and ERK5 phosphorylation. Matched results with SB are shown for comparison. (C) PD inhibits CVB3-induced ERK1/2 phosphorylation. Matched results with SB are shown for comparison. (D) BIRB inhibits CVB3-induced p38 activity as read out by Hsp27 phosphorylation. (E) SB and BIRB both inhibit CVB3-induced ERK5 phosphorylation. (F) Densitometry of CVB3-induced p-ERK5 upon BIRB inhibition. (G) SB and BIRB both inhibit hyperosmolarity-induced ERK5 phosphorylation. 293T cells were pretreated with SB203580 (SB, 20 μM) or BIRB796 (BIRB, 20 μM) for one hour, stimulated with 400 mM sorbitol for 20 min, and then analyzed for the indicated proteins by immunoblotting with tubulin used as a loading control.For (B–E), HL1 cells were pretreated with U0126 (U0, 10 μM), PD184352 (PD, 2 μM), SB, or BIRB for one hour, infected with sham for CVB3 at (B–D) M.O.I. = 9 or (E) M.O.I. = 18 for (B,E) 0.17 h or (C,D) 24 h, and then analyzed for the indicated proteins by immunoblotting with tubulin used as a loading control. For (F), densitometry measurements were normalized to CVB3-infected cells without inhibitor and data are shown as the mean ± s.e.m of four biological replicates. Asterisk indicates $p < 0.05$ by Welch's one-sided *t* test.

We tested for crosstalk between p38 and ERKs by using SB203580 (SB), an ATP-competitive small-molecule inhibitor of p38 (*191*). We monitored p-ERK1/2, p-ERK5, and p-p38, as well as the major ERK1/2 and p38 effector kinases, RSK and MK2 (*183, 189*). We found that SB potently inhibited p38 activity in cardiomyocytes, as expected, blocking phosphorylation of MK2 at early and late times after CVB3 infection (Figure 4-3, D and E). Acute SB treatment was also specific, because we did not observe any effect on early CVB3-induced ERK1/2 phosphorylation or activity (Figure 4-3D). Upon prolonged SB treatment, however, we observed a modest-but-reproducible increase in ERK1/2 phosphorylation (Figure 4-3, E and F). We attribute this to secondary inhibition of PP1 and PP2A phosphatases, which are normally activated by p38 signaling and serve to dephosphorylate MEK1/2 upstream of ERK1/2 (*192*). Subsequent control experiments showed that SB-induced upregulation of ERK1/2 was independent of CVB3 treatment (Figure 4-3G). Thus, p38 signaling antagonizes late ERK1/2 signaling, prompting a re-evaluation of earlier p38-inhibition experiments involving CVB3 (see below) (*193*).

A second finding from these experiments was that SB treatment potently blocked both early- and late-phase phosphorylation of ERK5 (Figure 4-3, H and J). The p38–ERK5 coupling was consistent with predictions of the model (Figure 4-3A), but such crosstalk had not previously been reported. To exclude the possibility that ERK5 inhibition was caused by off-target effects of SB, we repeated the experiments with BIRB796 (BIRB), a mechanistically distinct inhibitor of p38 (*194*) (Figure 4-4D). When cardiomyocytes were pretreated with BIRB, we observed the same blockade of CVB3-induced ERK5 phosphorylation as with SB (Figure 4-4, E and F). Last, to examine the

generality of the p38–ERK5 connection, we treated human embryonic kidney cells with sorbitol as an osmotic stress to activate both p38 and ERK5. SB and BIRB each blocked hyperosmolarity-induced ERK5 phosphorylation (Figure 4-4G), suggesting that p38 is generally required for proper activation of the MEK5–ERK5 pathway. Taken together, the molecular consequences of SB and BIRB indicate that p38 is functionally interconnected with both ERK1/2 and ERK5, as predicted by the model of CVB3 pathogenesis (Figure 4-3A).

4.5. Deconvolution of the ERK- and p38-dependent apoptotic response

An important category of host-cell responses in the starting dataset was the activity of apoptotic caspases. ERKs and p38 mapped closely to these readouts in the model and could conceivably control CVB3-induced apoptosis directly (Figure 4-3A). Both ERK1/2 and p38 have been reported to be important for proper caspase activation (*193, 195*). However, these earlier studies used a dual MEK1/2–MEK5 inhibitor (U0) and were not aware of the antagonism between p38 and ERK1/2 (Figures 4-3, E and G, and 4-4C). We thus pursued follow-up studies using gain- and loss-of-function approaches for ERK–p38 together with direct measurements of caspase processing.

We began with ERK5, as it inhibits cardiac apoptosis in other contexts (*196, 197*) but had not been previously implicated in CVB3 infection. To block ERK5 signaling, we used the specific ATP-competitive ERK5 inhibitor, XMD8-92 (XMD) (*198*). XMD treatment potently reduced phosphorylation of an ERK5 substrate (MEF2A) in cardiomyocytes and significantly increased caspase-9 and caspase-3 cleavage upon CVB3 infection ($p < 0.05$) (Figure 4-5, A and B). CVB3-induced apoptosis also

increased when endogenous ERK5 was downregulated with shRNA (Figure 4-5, C and D). We performed a reciprocal gain-of-function experiment by establishing stable lines expressing a doxycycline (DOX)-inducible mutant of MEK5 that was constitutively active (MEK5-DD) (Figure 4-5E). Upon low-level infection with CVB3 (M.O.I. = 1.5), we found that DOX treatment of MEK5-DD-expressing cells caused a significant decrease in caspase-3 cleavage ($p < 0.05$) (Figure 4-5F). Interestingly, the drop in caspase-3 cleavage was associated with changes in caspase-8 activation rather than caspase-9 activation as with XMD. We attribute this difference to the kinetics of ERK5 activation with MEK5-DD (~8 hours) versus ERK5 inhibition with XMD (< 1 hour). The MEK5-DD, XMD, and shRNA results together indicate that CVB3-induced ERK5 signaling inhibits cardiomyocyte apoptosis. This link between ERK5 and host-cell survival is unique, because virtually all other CVB3-stimulated pathways described thus far promote apoptosis rather than inhibit it (Table 4-1).

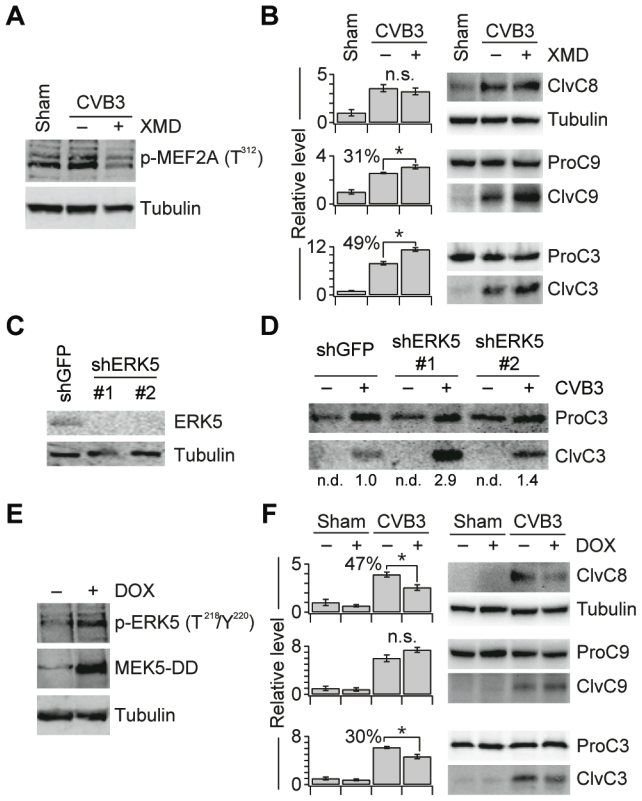


Figure 4-5. ERK5 inhibits CVB3-induced apoptosis. (A) XMD inhibits ERK5 activity in cells. HL1 cells were pretreated with XMD8-92 (XMD, 5 μM) for one hour, infected with sham or CVB3 at M.O.I. = 9, and analyzed for phospho-MEF2A at 24 h p.i. by immunoblotting with tubulin as a loading control. (B) CVB3-induced caspase-9 and caspase-3 processing is increased by XMD-mediated inhibition of ERK5. (C and D) Knockdown of endogenous ERK5 increases CVB3-induced caspase-3 cleavage. HL1 cells were transduced with the indicated shRNAs and analyzed for the indicated proteins by immunoblotting with tubulin or full-length caspase-3 used as a loading control. In (D), quantitative densitometry is shown relative to shGFP control cells infected with CVB3 at M.O.I. = 9. (E) DOX induction of MEK5-DD activates endogenous ERK5. HL1 cells stably expressing doxycycline (DOX)-inducible MEK5-DD were treated with 1 μg/ml DOX for 24 hr and analyzed for the indicated proteins by immunoblotting with tubulin used as a loading control. (F) CVB3-induced caspase-8 and caspase-3 processing is decreased upon activation of ERK5 by DOX-inducible MEK5-DD. For (A) and (B), HL1 cells were pretreated with XMD8-92 (XMD, 5 μM) for one hour, and infected with sham or CVB3 at M.O.I. = 9. For (E) and (F), HL1 cells stably expressing doxycycline (DOX)-inducible MEK5-DD were infected with CVB3 at M.O.I. = 1.5 and treated with 1 μg/ml DOX at 0 h p.i.. Samples were analyzed for the indicated active caspase-cleavage products at 24 h p.i. by immunoblotting with tubulin or full-length caspases used as a loading control. For (B) and (F), densitometry measurements were normalized to sham-infected cells without inhibitor and data are shown as the mean ± s.e.m of four biological replicates. Asterisk indicates $p < 0.05$ by Welch's one-sided t test.

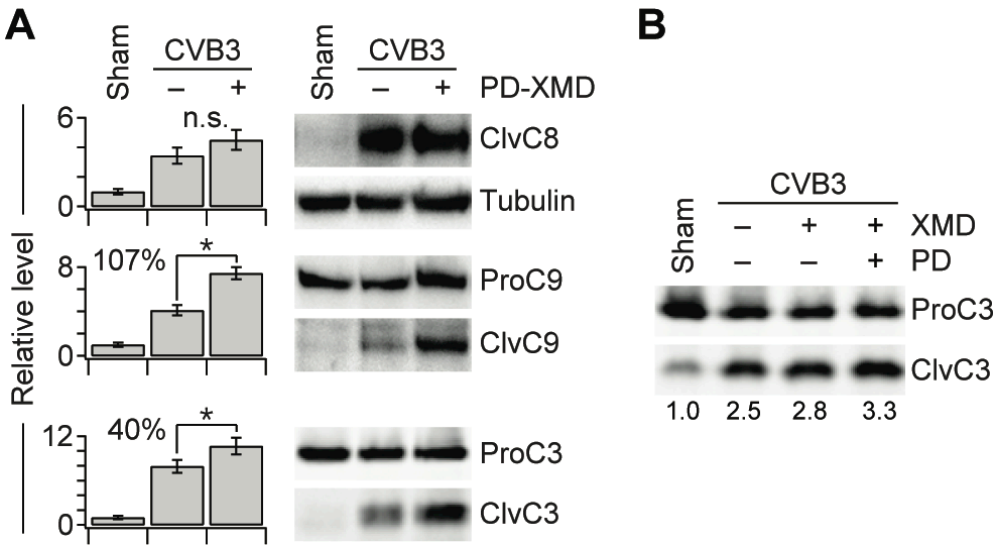


Figure 4-6. Pharmacologic perturbation of ERK5 and ERK1/2 signaling and their role in CVB3-induced caspase activation. (A and B) Pan-ERK inhibition with XMD+PD increases CVB3-induced apoptosis similarly to XMD treatment alone. In (B), cells are infected with sixfold lower CVB3 titers compared to (A), and quantitative densitometry is shown relative to sham-infected cells. HL1 cells were pretreated with XMD8-92 (XMD, 5 μM), PD184352 (PD, 2 μM), or XMD+PD for one hour, infected with sham or CVB3 at (A) M.O.I. = 9 or (B) M.O.I. = 1.5, and analyzed for the indicated active caspase-cleavage products at 24 h p.i. by immunoblotting with tubulin or full-length caspases used as a loading control. Densitometry measurements were normalized to sham-infected cells without inhibitor. For (A), data are shown as the mean ± s.e.m of four biological replicates. Asterisk indicates $p < 0.05$ by Welch's one-sided t test.

Next, we examined the p38 pathway by using a similar set of approaches. Consistent with an earlier report (*193*), we found that p38 inhibition via SB profoundly reduced caspase-3 cleavage during CVB3 infection (Figure 4-7A). We reinforced the SB result by showing that p38 inhibition with BIRB phenocopied SB in its blockade of these experiments is that p38 promotes CVB3-induced apoptosis. However, when we attempted the reciprocal gain-of-function experiment with a DOX-inducible, constitutively active mutant of MKK6 (MKK6-EE), there was no detectable change in caspase activation (Figure 4-7, B and C). The apparent contradiction prompted us to re-evaluate our experiments considering the cross-communication between p38 and ERKs (Figure 4-7E–J).

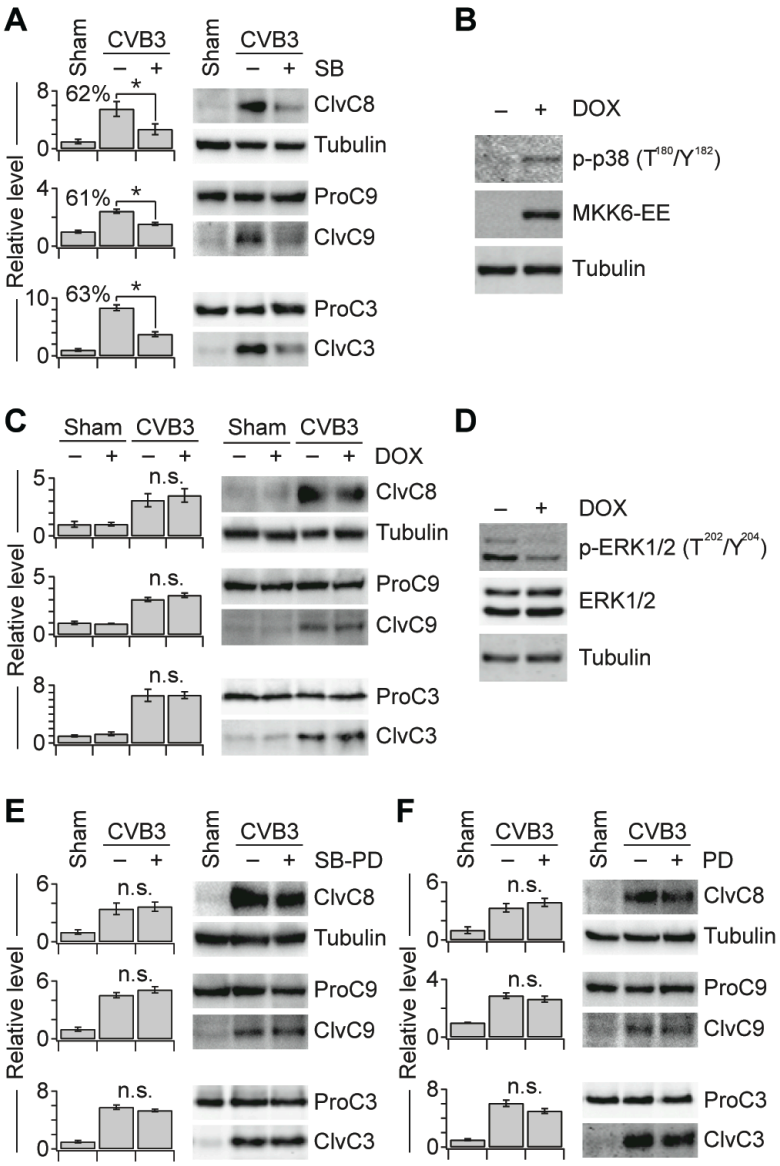


Figure 4-7. p38 inhibition blocks CVB3-induced apoptosis indirectly via ERK1/2 hyperactivation. (A) CVB3-induced caspase processing is blocked by SB-mediated inhibition of p38. (B–D) p38 activation via DOX-inducible MKK6-EE does not affect CVB3-induced apoptosis but suppresses ERK1/2 phosphorylation. For (B) and (D), HL1 cells stably expressing doxycycline (DOX)-inducible MKK6-EE were treated with 1 μg/ml DOX for 8 hr and analyzed for the indicated proteins by immunoblotting with tubulin used as a loading control. (E) SB-mediated inhibition of caspase processing is blocked by co-inhibition of ERK1/2 signaling with PD. (F) ERK1/2 inhibition with PD does not affect CVB3-induced apoptosis. For (A), (E), and (F), HL1 cells were pretreated with SB203580 (SB, 20 μM), PD184352 (PD, 2 μM), or SB+PD for one hour, and infected with sham or CVB3 at M.O.I. = 9. For (C), HL1 cells stably expressing DOX-inducible MKK6-EE were infected with CVB3 at M.O.I. = 1.5 and treated with 1 µg/ml DOX at 8 h p.i.. Samples were analyzed for the indicated active caspase-cleavage products at 24 h p.i. by immunoblotting with tubulin or full-length caspases used as a loading control. Densitometry measurements were normalized to sham-infected cells without inhibitor and data are shown as the mean ± s.e.m of four biological replicates. Asterisk indicates $p < 0.05$ by Welch's one-sided t test.

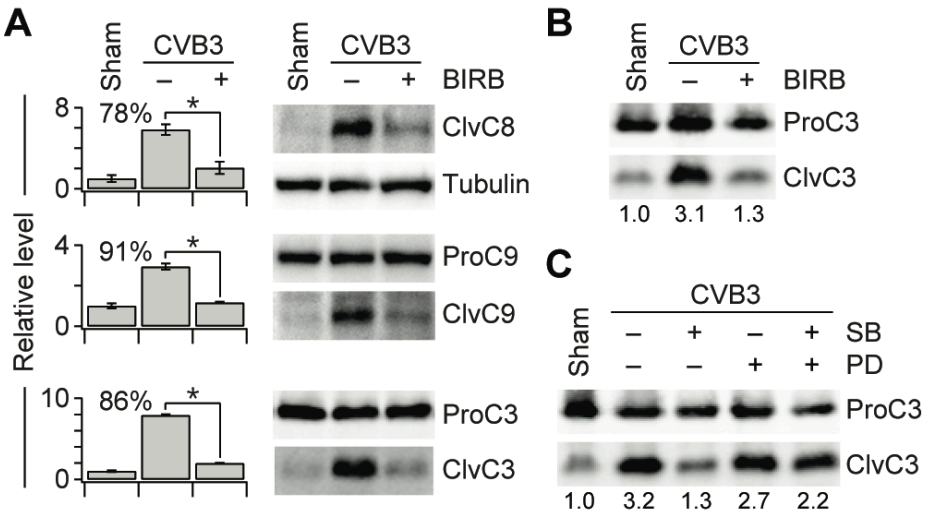


Figure 4-8. Pharmacologic perturbation of p38 and ERK1/2 signaling and their role in CVB3-induced caspase activation. (A and B) CVB3-induced caspase processing is blocked by BIRB-mediated inhibition of p38. In (B), cells are infected with sixfold lower CVB3 titers compared to (A). (C) SB-mediated inhibition of caspase processing is blocked by co-inhibition of ERK1/2 signaling with PD. Cells are infected with sixfold lower CVB3 titers compared to Figure 4E. HL1 cells were pretreated with BIRB796 (BIRB, 20 μM), SB203580 (SB, 20 μM), PD184352 (PD, 2 μM), or SB+PD for one hour, infected with sham or CVB3 at (A) M.O.I. = 9 or (B,C) M.O.I. = 1.5, and analyzed for the indicated active caspase-cleavage products at 24 h p.i. by immunoblotting with tubulin or full-length caspases used as a loading control. Densitometry measurements were normalized to sham-infected cells without inhibitor. For (A), data are shown as the mean ± s.e.m of four biological replicates. Asterisk indicates $p < 0.05$ by Welch's one-sided *t* test.

We reasoned that secondary inhibition of ERK5 would partially offset the observed SB–BIRB phenotype rather than cause it (Figures 4-3H–J and 4-5). Therefore, our attention turned to ERK1/2, which becomes hyperactivated upon prolonged p38 inhibition (Figure 4-3, F and G). This negative regulation of ERK1/2 by p38 was further strengthened by the reduced p-ERK1/2 observed in DOX-treated MKK6-EE cells (Figure 4-7D). To determine whether the consequences of p38 inhibition were mediated through ERK1/2, we combined SB with PD to block ERK1/2 hyperactivation and found that CVB3-induced apoptosis occurred normally (Figures 4-7E and 4-8C). Remarkably, ERK1/2 inhibition by itself did not substantially affect apoptosis of CVB3-infected cells (Figures 4-7F and 4-8C), suggesting that ERK1/2 acted as a pro-survival signal only

when p38 function was blocked. The p38-specific role of ERK1/2 was re-emphasized in ERK5-inhibited cells, where PD+XMD increased apoptosis as with XMD alone (Figures 4-5B, and 4-6, A and B). Upon this re-evaluation of earlier studies using SB (*193*), we conclude that p38 signaling does not directly control CVB3-induced apoptosis.

4.6. p38 signaling contributes to CVB3-induced pathogenesis by stimulating necrosis

Apoptosis is but one facet of the host-cell response to CVB3 infection, raising the question of whether other aspects of pathogenesis could require p38 signaling (*199, 200*). In the original dataset, overall CVB3 cytotoxicity was measured via tetrazolium reduction. However, this method was inadequate to read out cytotoxicity in the presence of signaling perturbations, which could also affect proliferation and metabolism. We therefore switched to a fluorescent amine-reactive dye that intensely labels cells with compromised plasma-membrane integrity irrespective of the mechanism of cell death (*201*).

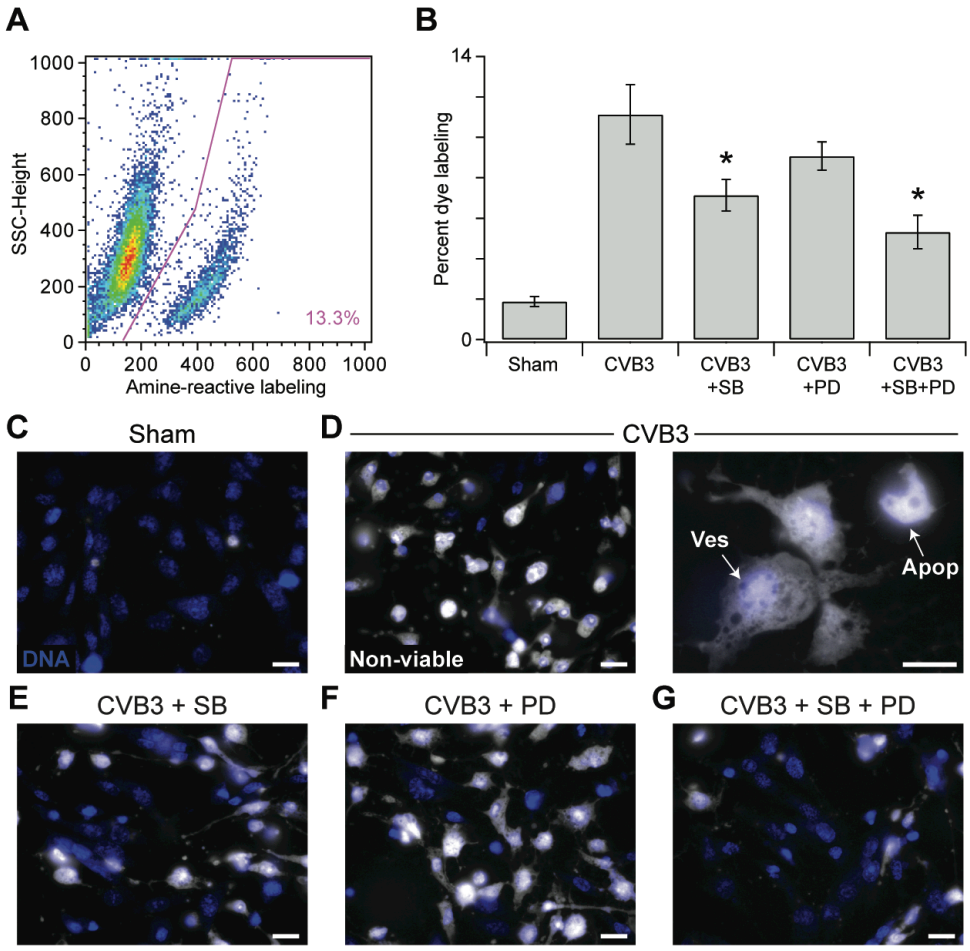


Figure 4-9. p38 inhibition improves viability of CVB3-infected cells independently of ERK1/2 signaling. (A) Representative flow-cytometry profile of CVB3-infected cardiomyocytes labeled with amine-reactive dye. Dye-positive cells (pink) are considered non-viable. (B) SB and SB+PD inhibit CVB3-induced cytotoxicity. (C–G) Representative images of amine-labeled adherent CVB3-infected cardiomyocytes pretreated with SB, PD, or SB+PD. Cells were pretreated with SB203580 (SB, 20 μM), PD184352 (PD, 2 μM), or SB+PD for one hour and infected with sham or CVB3 at M.O.I. = 9. Cells labeled with amine-reactive dye, and monitored by flow cytometry or fluorescence microscopy. For (B), data are shown as the mean ± s.e.m. of four biological replicates. Asterisk indicates $p < 0.05$ by Welch's one-sided t test. For (C–G), scale bar is 20 μm.

We found that CVB3 infection caused a dramatic increase in the percentage of dye-labeled, non-viable cells as compared to sham infection (Figures 4-9A–D). The actual extent of cytotoxicity was much greater than the flow-cytometry estimate (compare Figures 4-9, B and D), because many infected cells were so damaged that they were unavoidably lost during the suspension preparation. As with the earlier apoptosis

experiments (Figure 4-7), we found that SB pretreatment strongly decreased the extent of CVB3-induced cytotoxicity, whereas PD did not have a significant impact ($p > 0.05$) (Figures 4-9B–F). Surprisingly, when SB and PD were combined, we observed a clear improvement in overall cell viability even though caspase activation was unaffected under these conditions (Figures 4-7E and 4-9G). The pronounced result of dual p38–ERK1/2 inhibition was also reflected in significantly reduced titers of released viral progeny ($p < 0.001$) (Figure 4-10, A and B). This raised the possibility that p38 could control alternative death pathways that were distinct from apoptosis but critically important for CVB3 pathogenesis.

We closely examined the morphology of CVB3-infected cells by microscopy and noted a mixture of phenotypes indicative of discrete single-cell outcomes (Figure 4-9D). Some cells had a rounded appearance with condensed nuclei, suggesting an apoptotic fate. Others, however, remained fully spread and had aberrant lamellipodia-like projections (Figure 4-9D, right). These cells also had an intact nucleus along with intracellular vesicles that remained dye impermeant. Our observations suggested that a fraction of CVB3-infected cells undergo a vesiculated form of cell death with certain hallmarks of necrosis (*200*).

To determine whether CVB3 infection was associated with biochemical readouts of necrosis, we examined the chromatin protein HMGB1, which is released extracellularly by necrotic cells (*202*). We validated the marker by stimulating cardiomyocytes with hydrogen peroxide (a recognized inducer of necrosis) and observing pronounced HMGB1 release (Figure 4-11A). Importantly, we found that HMGB1 was clearly detected in supernatants from CVB3-infected cells (Figure 4-10C). HMGB1

release was unaffected by the apoptosis inhibitor DEVD-CHO but was slightly reduced by the necrosis inhibitor Necrostatin-1, likely as a result of CVB3-induced autocrine TNF signaling (*177, 203*). Thus, HMGB1 is a reliable marker of necrosis stimulated by CVB3.

Upon p38 inhibition with SB or BIRB, we observed near-complete blockade of HMGB1 release, suggesting potent inhibition of necrosis (Figures 4-10D and 4-11B). Conversely, necrosis was negligibly affected in CVB3-infected cells treated with PD to inhibit ERK1/2, consistent with the earlier labeling results (Figures 4-9, B and F, and 4-10E). In stark contrast to the apoptotic readouts (Figure 4-7E), we did not observe any reversion of necrosis when CVB3-infected cells were pretreated with SB + PD (Figure 4-10F). Last, to test whether p38 signaling was sufficient to drive virus-induced necrosis, we returned to the inducible MKK6-EE cells and found that DOX treatment substantially augmented HMGB1 release during CVB3 infection (Figure 4-10G). We conclude that p38 signaling is a critical component of a necrosis pathway, which promotes CVB3 propagation independently of ERK-dependent apoptosis.

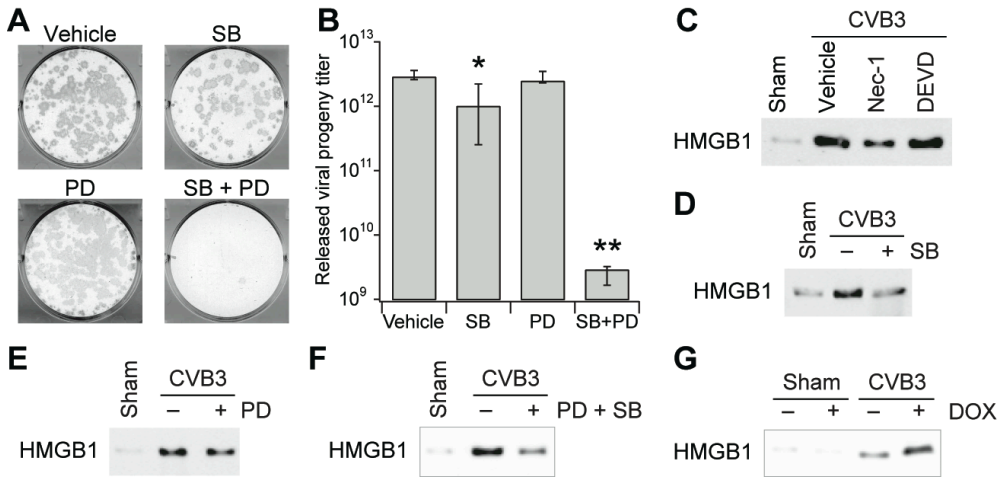


Figure 4-10. p38 controls CVB3-induced necrosis. (A and B) SB+PD markedly inhibits released viral progeny (RVP) in CVB3-infected cells. (C) HMGB1 is a reliable marker of CVB3-induced necrosis. (D–F) SB blocks CVB3-induced HMGB1 release independently of ERK1/2 pathway inhibition with PD. Note that PD does not affect the inhibition of HMGB1 release caused by SB. (G) p38 activation via DOX-inducible MKK6-EE increases CVB3-induced HMGB1 release. For (A–F), cells were pretreated with SB203580 (SB, 20 μM), PD184352 (PD, 2 μM), SB+PD, Necrostatin-1 (Nec-1, 50 μM), or DEVD-CHO (DEVD, 0.1 μM) for one hour and infected with sham or CVB3 at M.O.I. = 9. For (A and B), RVP titers were determined by plaque assay. For (D–G), culture supernatants were concentrated and analyzed for HMGB1 release by immunoblotting. For (G), HL1 cells stably expressing doxycycline (DOX)-inducible MKK6-EE were infected with CVB3 at M.O.I. = 1.5 and treated with 1 μg/ml DOX at 8 h p.i.. For (B), data are shown as the median ± range of four biological replicates. Single asterisk indicates $p < 0.05$ and double asterisk indicates $p < 0.001$ by Welch's two-sided *t* test.

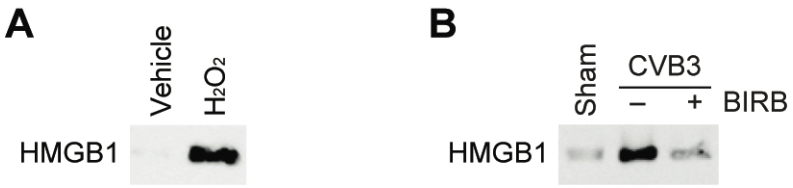


Figure 4-11. Monitoring CVB3-induced necrosis with HMGB1. (A) Validation of HMGB1 as a marker of necrosis caused by H_2O_2. HL1 cells were treated with 2 mM H_2O_2 for 24 h and analyzed for HMGB1 release by immunoblotting. (B) BIRB inhibits CVB3-induced necrosis. HL1 cells were pretreated with BIRB796 (BIRB, 20 μM) for one hour, infected with sham for CVB3 at M.O.I. = 9 for 24 h, and then analyzed for HMGB1 release by immunoblotting.

4.7. Summary

Viruses such as CVB3 activate many host-cell signaling pathways and evoke many host-cell responses. Our study began with a holistic, systems approach to monitor these events dynamically and as a function of CVB3 titer. By analyzing the data to make quantitative predictions of host-cell outcome, we quickly converged on ERKs and p38 as key pathways for CVB3 pathogenesis. Early-phase ERK1/2 activation stems directly from CVB3 docking to host membranes, whereas late-phase activation occurs due to cleavage of upstream signaling molecules by viral proteases (*195, 204*). Late-phase p38 and ERK5 signaling probably lies downstream of autocrine proinflammatory cytokines, which are induced during the final stages of the viral life cycle (Figure 4-13) (*177*). Despite differences in

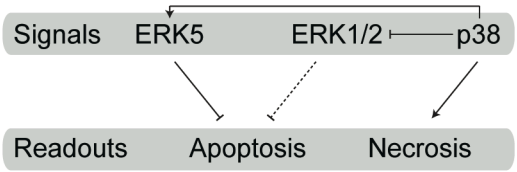


Figure 4-12. Model for ERK–p38 signaling, apoptosis, and necrosis induced by CVB3. Dashed line indicates context-dependent inhibition of apoptosis by ERK1/2.

activation, our work here shows that ERKs and p38 are strongly interconnected (Figure 4-12). These dependencies are important for interpreting the results of "single-pathway" perturbations that propagate through the network (*193, 195*).

Notably, we were able to uncover a role for ERK5 in CVB3 pathogenesis by modeling a dataset that did not measure ERK5 explicitly. We have shown elsewhere that quantitatively accurate signaling measurements are critical for data-driven models to reflect underlying biological mechanisms (*145, 179*). Our results here using a pan-ERK p-ELISA indicate that measurements of specific proteins may not be as important. This is encouraging, because many modern signaling assays increase overall throughput by relaxing the specificity constraints of traditional approaches (*205*).

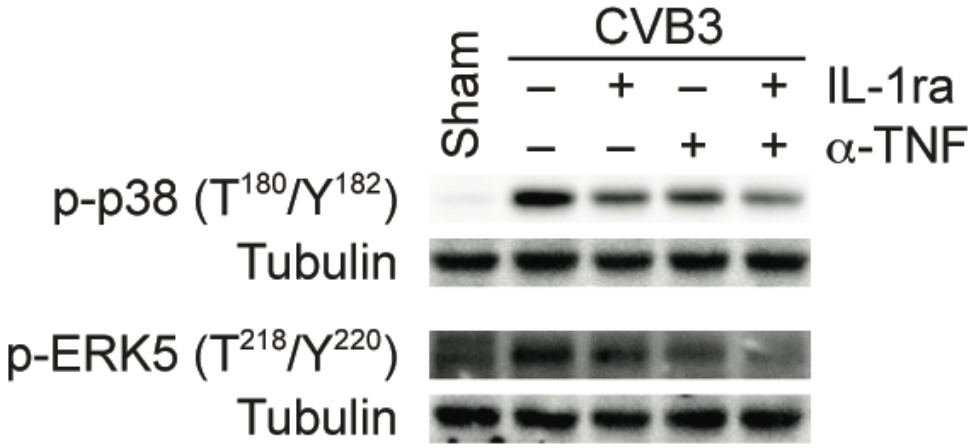


Figure 4-13. Late-phase ERK5 and p38 signaling is inhibited upon blockade of autocrine proinflammatory cytokines. HL1 cells were pretreated with TNF-neutralizing antibody (α-TNF, 1 mM), the IL-1 receptor antagonist IL-1ra (0.5 mM), or α-TNF + IL-1ra for one hour, infected with sham or CVB3 at M.O.I. = 9, and then analyzed for p-p38 and p-ERK5 at 24 h p.i. by immunoblotting with tubulin used as a loading control.

Similarly, our work shows that agglomerated cell-outcome data may be sufficient for viral-host modeling and discovering overlooked phenotypes. The importance of CVB3-induced necrosis as a host-cell fate was revealed here without direct necrotic

readouts in the model (Figure 4-12). This information was presumably embedded in the overall cytotoxicity measure, which depends strongly on the level of necrosis (Figures 5A–D). Interestingly, the associated RVP titers appear to be influenced by apoptosis and necrosis reciprocally. When both apoptosis and necrosis are blocked upon p38 inhibition with SB, there is a slight reduction in RVP. However, when apoptosis is restored in p38-inhibited cells by blocking ERK1/2 hyperactivation, RVP is dramatically reduced (Figure 4-10B). Thus, necrosis may be the preferred outcome for CVB3, which is counteracted by the host-cell drive to die by apoptosis. To isolate necrosis specifically requires targeting an upstream mediator (p38) and resetting the other secondary consequences of pathway inhibition (e.g., ERK1/2) (Figure 4-12). Such combinatorial anti-viral strategies would be difficult to predict without the aid of a systems model for the host-cell response to CVB3 infection.

CHAPTER 5

Design of a high-throughput multiplex kinase activity assay

5.1 Introduction

The human cell-signaling network is comprised of numerous signaling pathways and subnetworks (Chapters 3-4) that integrate complex input signals from receptors (Chapter 2) and other stimuli to execute outputs such as gene expression, cell growth or cell death (*206*). A complete systems-level understanding of the network as a whole will be important in understanding complex disease states and developing therapies to combat them. Methods that can probe the network more generally to understand how the subnetworks and pathways function together will be important tools for understanding network-level emergent phenomena such as acquired drug resistance and drug synergy.

5.2 Kinases are critical signaling nodes

Kinases are essential regulators of the cell-signaling network. The human "kinome" consists of 518 kinases that catalyze the phosphorylation of serine, threonine, or tyrosine amino acid residues of their protein substrates (*22*). Phosphorylation of a substrate can alter its stability, activity, or localization within the cell (*206*). Accordingly, the deregulation of kinase signaling has been shown to contribute to a multitude of diseases such as cancer and Alzheimer's disease (*13, 14*). Kinases often sit at critical nodes in the signaling network where they function as important activity hubs, integrating many inputs into numerous outputs. Importantly, while there are numerous

techniques available that can measure the extent of protein phosphorylation (mass spectrometry, Western blot, ELISA, etc), these assays are not good readouts of a kinase's enzymatic activity (*207*). Phosphorylation is a net measurement of kinase activity and the antagonists of kinases, the phosphatases, which dephosphorylate protein substrates (*208*). Methods that measure kinase activity directly will provide important insights to the regulation of these important components of cell-signaling networks.

5.3 Design of a high-throughput multiplex kinase activity assay

To address these needs, our group developed a high-throughput kinase activity assay (*209*). This assay design was successful in answering important biological questions (*19, 41, 134, 210*), but it faced several drawbacks. Due to the radioactive readout of the assay, sub-physiological levels of ATP were required (due to radioactivity limits), which hurt the sensitivity of the assay. Additionally, despite increased throughput compared to earlier assays, our assay was only capable of measuring kinases singly. To address these drawbacks and improve both the sensitivity and throughput of the assay, we are developing a nonradioactive, multiplex kinase assay (Figure 5-1). Acquiring multi-kinase measurements from a single sample will increase efficiency, making it easier for activities to be measured in high-throughput experiments. In this chapter we present preliminary feasibility experiments towards the design of a multiplex, high-throughput kinase activity assay.

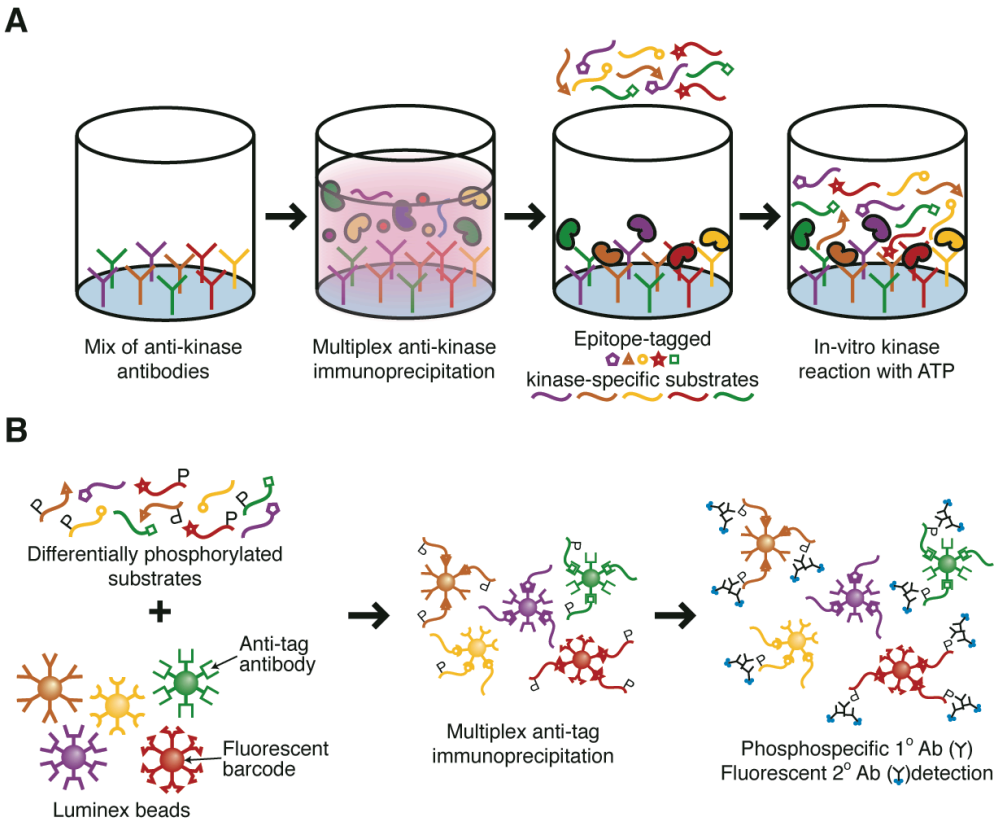


Figure 5-1. Design of high-throughput multiplex kinase activity assay. (A) Anti-kinase antibodies are precipitated onto a 96-well plate and incubated with cell lysate to immunoprecipitate endogenous kinases. An in vitro kinase reaction is performed in the well with epitope-tagged kinase-specific substrates and ATP. (B) Differentially phosphorylated substrates from the kinase reaction are incubated with Luminex beads with anti-tag antibodies. The substrates are immunoprecipitated onto the beads and labeled with phospho-specific antibodies.

Multiplexing kinases requires that the kinases are specific. Because kinase specificity is determined by the amino acid sequence surrounding the target phosphorylation site (*211*) and docking sites (*212*), multiple kinase activities can be measured in the same reaction, which not only increases the throughput of the assay, but also substantially decreases the amount of sample required. For our assay, we have chosen five kinases with distinct sequence recognition motifs and/or docking sites that are embedded in five central cell-signaling pathways: extracellular-regulated kinase

(ERK1/2), IκB kinase (IKK), Akt, mitogen-activated protein kinase-associated protein kinase 2 (MK2), and c-jun N-terminal kinase (JNK) (Table 5-1).

Table 5-1. Published substrate consensus sequences and docking sites for five assay kinases. Phosphorylation site is shown in red and Φ indicates any hydrophobic amino acid.

Kinase	Substrate consensus	Substrate docking site	Reference
ERK	PX(**S/T**)P	FXFP	(*213, 214*)
Akt	RXRXX(**S/T**)		(*215*)
JNK	(**S/T**)P	(K/R)XXXLXL	(*216*)
IKK	(pT/pS)(pT/pS)X(Y/F)X**S**(L/I)X(D/E)		(*217*)
MK2	(L/I/F)XR(Q/S/T)L(**S/T**)Φ		(*218*)

In our proposed design, mixtures of anti-kinase antibodies are coated on 96-well plates and used to immunoprecipitate endogenous kinases from a cell lysate. Then, mixtures of independently tagged substrates for the plate-bound kinases are added with ATP, and kinase-mediated substrate phosphorylation is allowed to proceed in vitro (Figure 5-1A). Epitope tags fused to the recombinant substrate provide a means to isolate individual substrates on fluorescent beads coated with anti-tag antibodies (Figure 5-1B). Kinase activity is read out by the extent of substrate phosphorylation, which will be marked by antibodies against the phosphorylation site on the substrate (Figure 5-1B). The identity of the substrate on the bead is indicated by the bead's fluorescence and anti-tag antibody, which can be tracked along with kinase activity on a Luminex instrument. By using fluorescence to quantify substrate phosphorylation (instead a traditional radioactive ^{32}P label), higher and more-physiological ATP concentrations can be used to significantly enhance the sensitivity of the assay. Using physiological ATP levels has been shown to considerably increase measured activity levels of certain kinases (*219,*

220) and should significantly improve the sensitivity of the assay. By increasing the throughput of the assay, using physiological levels of ATP (1 mM or higher), as well as obtaining an amplification of signal from the sandwich ELISA (Figure 5-1B), we anticipate that we will be able to increase the sensitivity by greater than tenfold compared to our previous kinase activity assay (*209*).

5.4. Reagent development for proposed assay design

5.4.1. Antibodies

We have validated commercially available anti-kinase antibodies for microplate immunoprecipitation of our five kinases (*209*) (Table 5-2). Anti-phospho antibodies for the substrates were also validated for use by Western blot (Table 5-2 and Figure 5-5).

Table 5-2. Validated antibodies for kinase assay.

Kinase	IP antibody	Substrate	Phospho site	Phospho antibody	Tag	Tag Antibody
MK2	StressGen KAP-MA015	Hsp27	S82	CST 2401	FLAG	Sigma F1804
IKK	BD 559675	IKB	S32	CST 2859	HA	Roche 11867423001
ERK	Upstate 06-182	RSK	T573	CST 9346	AU1	Covance MMS-130P
Akt	Upstate 05-591	GSK3	S21	CST 9316	VSVG	Roche 11667351001
JNK	Santa Cruz c-17	cjun	S73	CST 9164	Glu-Glu	Covance MMS-115P

5.4.2. DNA constructs

Due to the requirement for substrates to be separated by the Luminex instrument for a multiplex assay (Figure 5-1B), substrates used in the assay need to be differentially tagged. For this purpose, we designed DNA constructs to contain N-terminal epitope tags (HA, FLAG, Glu-Glu, VSVG, AU1) in triplicate. Epitope tagging was chosen because there are high affinity antibodies available for these short sequences (*221, 222*).

By using the epitope tag sequence in triplicate, we increased the avidity of the antibody-tag immunoprecipitation of the recombinant substrate to the Luminex bead (Figure 5-1B). Sequences for triple-epitope tags were cloned into the multiple cloning site (MCS) pGEX-4T-1 glutathione S-transferase (GST) fusion vectors. By cloning the epitope tag first and separately from the substrate sequence, these vectors can easily accept alternative substrate sequences without the need to clone additional epitope-tagged vectors. Subsequently, we cloned kinase-specific substrate sequences into the epitope-tag vectors to generate the following tagged substrates: 3XAU1-RSK(386-752) for ERK1/2, 3XHA-IκBα(1-62) for IKK, 3XVSVG-GSKα(1-97) for Akt, 3XFLAG-Hsp27 for MK2, and 3XGluGlu-cjun(1-79) for JNK. These substrate sequences have been shown to be good readouts of kinase activity for the five kinases described (*215, 216, 223-225*). Truncations of substrates excluded kinase domains of the substrates and allowed for easier protein purification.

5.4.3. Recombinant protein substrates

Recombinant substrates were expressed and purified for use in the assay. Plasmid constructs described in Chapter 5.4.2 were transformed into chemically competent low copy C41 bacteria for expression. Sequences cloned into the MCS of pGEX-4T1 vectors are under control of the lac operon to control expression. The pGEX-4T1 vector constitutively expresses the LacI repressor protein that prevents activity at the Ptac promoter for the MCS. Addition of Isopropyl β-D-1-thiogalactopyranoside (IPTG) relieves the inhibition by LacI and allows expression at the Ptac promoter. IPTG is a molecular imitator of allolactose, a lactose metabolite that drives transcription at the Ptac

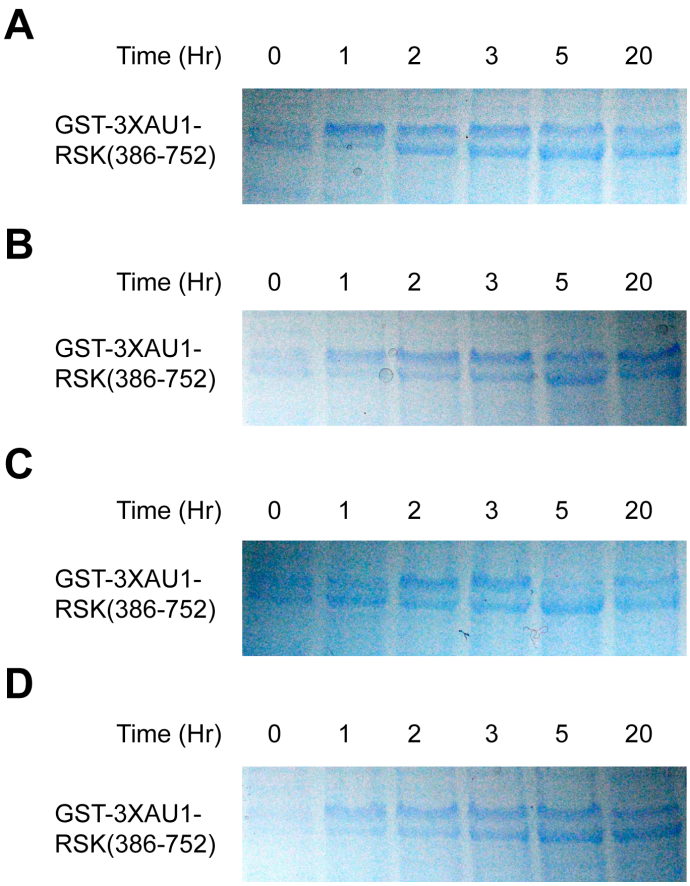


Figure 5-2. Example of induction optimization for 3XAU1-RSK(386-752). Bacterial lysates were run on SDS-PAGE gel after being induced by IPTG (A) 0.4 mM (B) 1 mM (C) 2 mM and (D) 5 mM.

promoter. IPTG is preferable for use in induction experiments because IPTG cannot be hydrolyzed by β-galactosidase like allolactase. Since the IPTG is not hydrolyzed, its concentration is maintained during the induction for constant expression. The amount and duration of IPTG induction was optimized to yield maximal protein production (Table 5-3) (example shown in Figure 5-2).

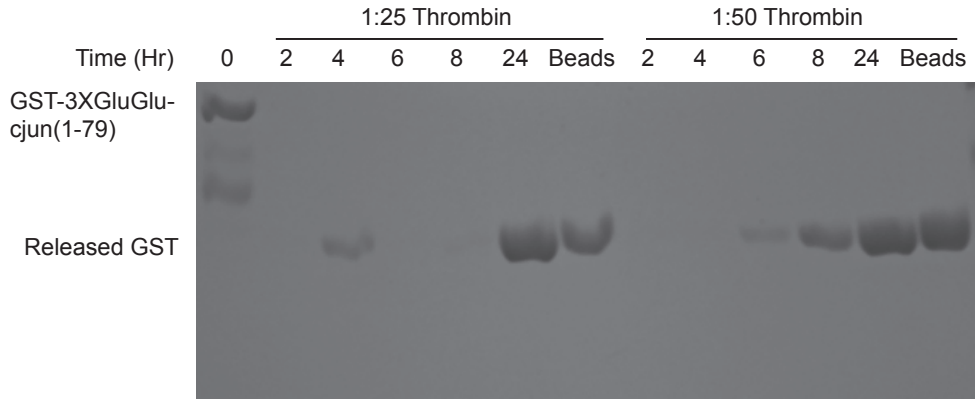


Figure 5-3. Example of thrombin optimization for 3XGlu-Glu-cjun(1-79). Release of GST and 3XGluGlu-cjun (not shown) from glutathione agarose beads were monitored over 24 hours for with varying concentrations of thrombin enzyme (not all dilutions shown).

Table 5-3. Optimized protein purification conditions.

Substrate	[IPTG](mM)	Induction time	Thrombin (µl)	Thrombin time
Hsp27	0.4mM	4 hour	2ul	6 hour
IκBα	0.4mM	1 hour	2ul	6 hour
RSK	0.4mM	5 hour	12ul	4 hour
GSK3α	2mM	6 hour	8ul	2 hour
c-jun	1mM	1 hour	2ul	6 hour

Addition of the GST moiety to the desired substrate provides an affinity handle to purify the substrate from the bacteria culture after induction. GST is a powerful affinity handle because its affinity for its substrate, glutathione, is in the submillimolar range (*226*). Using the GST handle, the recombinant substrates are purified from the bacterial lysate using glutathione agarose beads. After the purification, the GST can be removed from the substrate using thrombin enzyme that specifically recognizes and cleaves the sequence between GST and the substrate. The amount of thrombin enzyme and cleavage time were optimized to maximize GST cleavage and simultaneously minimize cleavage products of the substrate (Table 5-3). Following the thrombin digest, thrombin enzyme is

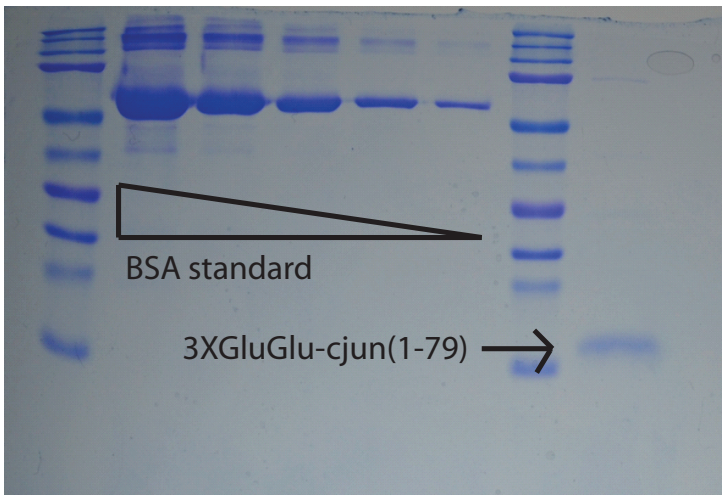


Figure 5-4. Example of recombinant substrate quantification. BSA standard (10 μg, 5 μg, 2.5 μg, 1.25 μg, 0.625 μg) was run on a SDS-PAGE protein gel alongside purified recombinant substrate and stained with Coomasie blue stain. Image analysis and densitometry was preformed in ImageJ software (*5*).

removed from the purified substrate with p-aminobenzamidine beads to yield purified substrate. Purified substrate was quantified by SDS-PAGE (example shown in Figure 5-4). We have shown that the purified recombinant substrates are all phosphorylated in vitro by immunoprecipitated cellular kinases in our microplate format (Figure 5-5).

5.4.4. Luminex beads conjugated with anti-epitope tag antibodies

Customized Luminex beads are necessary to immunoprecipitate the appropriate substrates to differentially labeled beads. Anti-epitope tag antibodies were chemically coupled by EDC and Sulfo-NHS to carboxylated Luminex beads to form a stable amide linkage (*227*). Each epitope tag antibody was coupled to a different bead identification region for discrimination by the Luminex instrument. Each bead region is defined by varying concentrations of two dyes to give each region a unique identity. The amount of anti-epitope tag antibody that could be chemically coupled the Luminex beads was

optimized and customized beads were prepared with each of the five epitope tag antibodies.

5.5. Preliminary results with high-throughput multiplex kinase activity assay

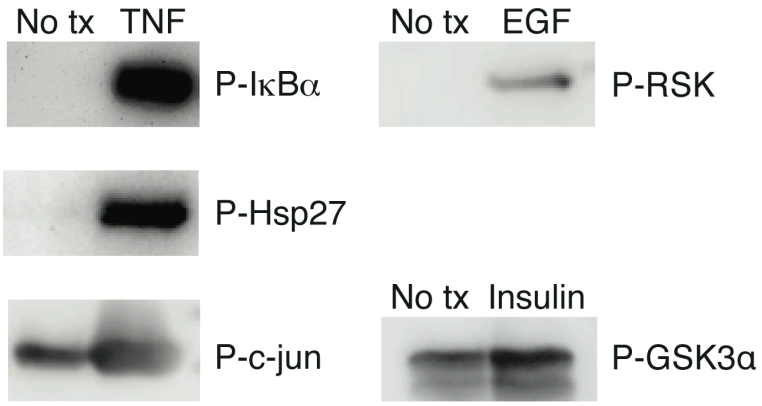


Figure 5-5. Recombinant substrates are phosphorylated in vitro by immunoprecipitated endogenous kinases. IKK, MK2, and JNK kinases were active in phosphorylating their substrates IκBα, Hsp27, and c-jun in TNF-stimulated lysates. ERK kinase from EGF-stimulated lysate was active in phosphorylating its substrate RSK. Akt kinase from insulin-stimulated lysate was active in phosphorylating GSK3α.

Multiplexing is dependent on kinase specificity within the assay. Since many kinases have numerous substrates in vivo, substrates used in the assay must only be recognized by the intended kinase. While the five kinases chosen for the assay have distinct sequence recognition motifs and/or docking sites (Table 5-1), it is important to validate that there is no crosstalk in the assay format. First, it is necessary to check whether a kinase is phosphorylating substrates other than its intended target. To test this directly, assays will be performed with mismatched immunoprecipitation antibodies and substrates (Figure 5-6). In our preliminary experiments with MK2 and IKK, immunoprecipitation of IKK did not result in Hsp27 phosphorylation (Figure 5-6A). However, immunoprecipitation of MK2 resulted in IκBα phosphorylation (Figure 5-6B), suggesting

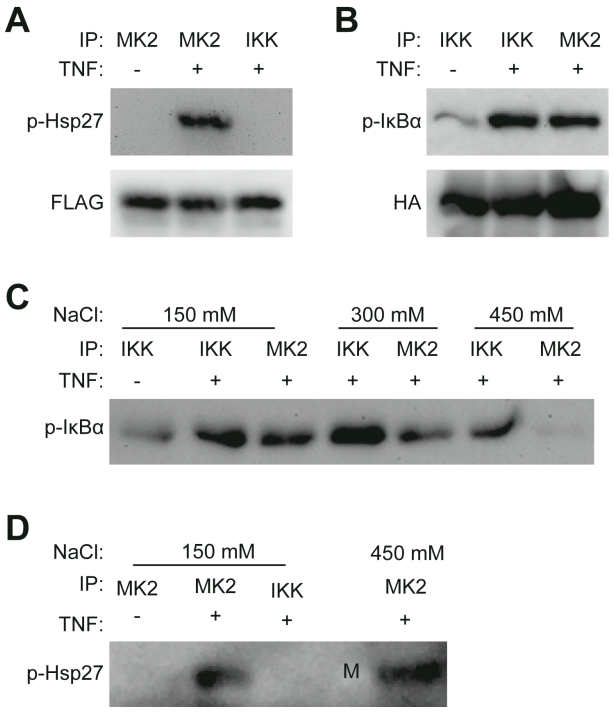


Figure 5-6. Apparent crosstalk from MK2 to IκBα is resolved by higher stringency washes. (A) Hsp27 is phosphorylated by MK2 in TNF-stimulated lysates, but not by IKK. (B) IκBα is phosphorylated in both IKK and MK2 assays with TNF-stimulated lysates. (C) Increasing NaCl in the wash buffer abolishes apparent MK2 activity against IκBα. (D) High stringency wash condition does not affect MK2 assay against Hsp27. M denotes marker lane.

that MK2 may have activity towards IκBα or that IKK was being immunoprecipitated with MK2. Increasing the stringency of the washes following the immunoprecipitation abrogated IκBα phosphorylation in the MK2 assay (Figure 5-6C). The high stringency wash did not reduce phosphorylated Hsp27 in the MK2 assay (Figure 5-6D). This suggested that IKK was immunoprecipitating in complex with MK2 and that this complex was disrupted by the high salt washes. It will also be important to validate that no kinase is activating another kinase. In vitro, ERK has been shown to phosphorylate MK2 (*228*). While this does not happen in cells, this interaction or similar interactions among the kinases being assayed will be important to exclude when multiplexing.

In addition to our kinase assay protocol, we have developed a protocol for immunoprecipitating differentially-phosphorylated substrates onto custom Luminex beads for measurement on the Luminex machine. In our preliminary experiments, we

have shown detection of increased substrate phosphorylation by Luminex for all substrates. We have also observed no significant reduction in measured activity when two kinases are multiplexed (Figure 5-7, A and B). Multiplexing slightly decreased the measured levels of Akt activity but the relative increase in signal induced by TNF was unchanged (Figure 5-7A). While several assay parameters remain to be optimized, our preliminary results suggest that our proposed format will accommodate the five kinases outlined here in a multiplex format.

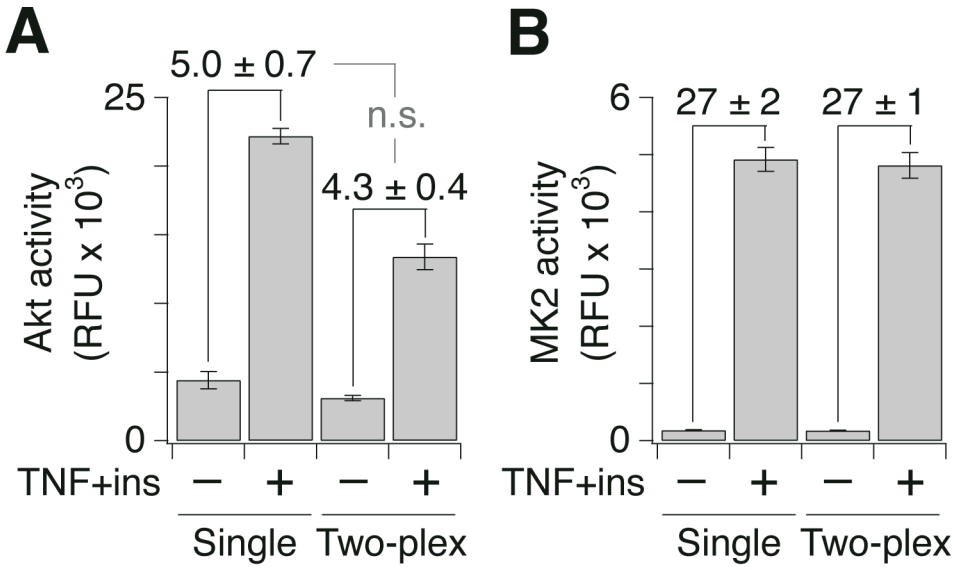


Figure 5-7. Preliminary assay shows no loss of signal when two kinases are multiplexed. (A) Akt activity in HT29 cells stimulated with TNF and insulin in single and two-plex assay formats. (B) MK2 activity in HT29 cells stimulated with TNF and insulin in single and two-plex assay formats.

5.6. Summary

Kinases are an important class of signaling proteins in human cell-signaling networks that are key integrators of signals. Existing kinase activity assays measure kinase function individually (*209, 229*) or indirectly (*230*). Measuring multiple key signaling pathways simultaneously will help us understand the complex effect of a stimulus or drug. In order to probe the network more generally, we sought to develop a high-throughput experimental method to measure kinase activity. Our design is also more sensitive than previous kinase activity assays (*209*) due to the physiological ATP concentrations allowed by the non-radioactive format and signal amplification by sandwich ELISA. Together, these innovations will substantially increase throughput of measuring kinase activity and will require far less sample than previously required.

Another strength of our assay design is that it is generalizable. We specifically designed our triple epitope-tag plasmids to be amenable to easy interchange of substrate protein constructs. In the future, we believe that our platform will be able to accommodate additional five-kinase panels to build larger kinase activity datasets or to answer different biological questions by studying a new set of kinases. This assay will allow us to probe multiple signaling subnetworks simultaneously and study signaling at the network level.

CHAPTER 6

Conclusions and Future Directions

6.1. Utilizing systems approaches to understand cell-signaling networks and treat disease

Systems approaches continue to make important contributions to our understanding of the cell-signaling network (*231, 232*). Host-pathogen interactions, cancer, and other diseases are increasingly appreciated as complex cellular perturbations that will benefit from systems-level analysis (*233, 234*). Similarly, the treatment of these diseases is moving beyond the study of a single drug's interaction against its intended target to the global effect of a drug or combination of drugs. The field of systems pharmacology is gaining recognition as a promising approach to study the complexity of drugs within the network and design better therapies (*32*). In this chapter we will discuss future directions for the projects presented in this thesis and how they will help answer some of these questions about cell-signaling at the network level.

6.2. High-throughput experimental assays for the generation of large datasets

6.2.1. Receptome profiling

Our receptome profiling assay (Chapter 2) demonstrated the importance of receptor presence/absence and the inadequacy of existing platforms to accurately measure their expression and definitively conclude true absence. Advantages of our platform include its low cost and that it does not require specialized equipment, making it

desirable for high-throughput profiling and accessible for other labs and researchers. Further, we provide a framework for the design and validation of additional primer sets to expand the platform to include additional receptor families. Since the assay is compatible with any human cells, both established cell lines and primary cells, it will find use in the study of diverse human diseases.

In our initial application of the receptome assay, we profiled 40 cancer cell lines, with a focus on breast, melanocytic, and pancreatic lines (Figure 2-9). This analysis uncovered receptors that were overexpressed or selectively silenced in cancer types that were not previously appreciated (Tables 2-2 and 2-3). The presence or absence of a single receptor can change the baseline signaling, gene expression, and response to stimuli (as demonstrated with IL10RA in melanoma, Chapter 2-7). Continued study of these receptors in particular cancers could identify receptor expression that may be important in disease progression or receptors that could be therapeutically targeted. There is precedence for targeting receptor overexpression in several cancers demonstrated by the success of anti-EGFR, anti-VEGFR, and anti-HER2 therapies, among others (*235*). Receptor upregulation has also been shown recently to be an important mechanism of acquired or innate drug resistance in certain cancers (*236-238*). Receptome profiling could identify sets of receptors that change expression levels in response to drug treatment or when cells become resistant to a drug. Additionally, our receptome profiling has been demonstrated to be compatible with primary tissue samples (Figure 2-15). By profiling primary tumor samples, distinct patient signatures could be associated with drug response, which could provide a rapid and inexpensive clinical test.

Receptome profiling could also be useful in the characterization of cell lines as models for disease. The ability to grow human cells outside the body in culture and the development of stable cell lines has contributed greatly to many facets of medical science. Despite their utility and near universal use, limitations of cell line models are becoming more apparent, from misclassification to contamination to lack of representation of human disease (*239*). The limitations of cell line models are also taking blame for the failure to translate findings from models to patients. In one example of a misclassification, a particularly aggressive breast cancer line was actually melanoma (MDA-MB-435) (*240*). Especially concerning is that research articles continued to employ MDA-MB-435s a breast cancer model after it was reclassified (*241, 242*). Cell origins can be analyzed and compared with genetic analysis. However, genetic identity may not capture heterogeneity in disease presentation or the plasticity of cellular phenotypes across varying growth conditions or other variables (Chapter 2.8). Gene expression data can also be used to compare cell lines, but differences in receptor expression might not be captured in microarray or RNAseq comparisons due to the discrepancies between the methods as described in Chapter 2.4-2.5. Receptome profiling could provide a fast and inexpensive way to compare receptor profiles of cell line model systems to patient tumors to identify more representative model systems for experiments.

6.2.2 High-throughput multiplex kinase activity assay

Once complete, our high-throughput multiplex kinase activity assay will provide a useful tool for measuring information flow throughout the cell-signaling network by simultaneously measuring the activity of five kinases in canonical pathways. Our assay

and associated reagents are designed to be adaptable to alternate panels of five kinases. The ability to probe kinase activity at this scale, with improved sensitivity and lower sample requirement, will allow for the collection of large datasets that could serve as the foundation for data-driven models (Chapter 4). Interesting cell-signaling questions could be asked at the network level, such as why certain drug pairs demonstrate synergy when used in combination. While many screening efforts have been developed to identify synergistic drug pairs, the mechanisms of the synergy can be difficult to elucidate (*243*).

A second application of interest would be to use the assay to study complex perturbations to the cell-signaling network and how they alter the network over time. Particularly of interest in our group is the study of CVB3 in chronic infections. While CVB3 can result in acute infection of cardiomyocytes (Chapter 4), chronic myocarditis can develop from the mere presence of CVB3 RNA without an active infection (*176, 244*). The molecular mechanism by which CVB3 elicits this damage to the heart remains unknown and a challenging signaling problem at the network level. By monitoring the dynamics of multiple pathways simultaneously, our kinase assay will provide important insight to these biological questions.

6.3. Computational modeling of signaling networks to interpret and integrate datasets

6.3.1. Pathway connectivity analysis and target druggability

In Chapter 3, we introduced a set of computational models to predict the discrepancy between RNAi and small-molecule inhibition due to local network connectivity. Our models predicted a wide range of discrepancies dependent on pathway

topologies. From these results we hypothesize that sensitivity of a signaling enzyme to protein levels or catalytic activity would be exploited by evolution and cancer-cell evolution. Specifically, we predict that subnetworks sensitive to changes in activity but not protein levels will predominantly be activated by mutation. There are interesting biological examples in the literature to support the conclusions from the models. For example, in the topology of the Raf-MEK1/2-ERK1/2 pathway (Figure 3-4B), MEK1/2 is predicted to be more effectively inhibited by perturbations to enzyme activity compared to perturbations of enzyme level. In the literature, there are no reports of MEK1/2 being transcriptionally regulated in any context, but an activating mutation in MEK1/2 has recently been reported in vemurafenib-resistant melanoma (*245*). This evidence is consistent with our model predictions that for MEK1/2, catalytic activity is more important than enzyme levels for pathway inhibition. Conversely, we predict that subnetworks sensitive to protein levels but not activity will only be upregulated or downregulated. For example, in the caspase-8-caspase-3-caspase-6 pathway (Figure 3-13B), our modeling predicts that targeting protein level is more effective than catalytic activity. Caspase-3 levels have been shown to be transcriptionally regulated (*246*), and most inactivating caspase-3 mutations in human cancers are missense mutations that effectively reduce protein levels (*247*). The appreciation for the mechanistic differences between RNAi and small-molecule inhibition in different pathway wirings could provide important insight in prioritizing RNAi screens and may shed light on susceptible targets in cancer cell evolution.

6.3.2. Discovery of signaling subnetworks by data-driven modeling

Despite the importance of connectivity in cell-signaling pathways and networks (Chapter 3), connectivity is not always readily apparent from experimental data. Crosstalk amongst signaling pathways will be important to understand in disease contexts because activation of compensatory pathways is thought to be a major contributor to acquired drug resistance (*248*). Additionally, by identifying pathways that are perturbed together in a disease context, we can better identify combinations of drug targets to treat a disease (Chapter 4). Data-driven modeling approaches will continue to be important in developing biological hypotheses by integrating disparate types of experimental data, especially as technological advances allow us to measure new biological mechanisms.

In Chapter 4, we utilized data-driven modeling to understand the complexity of cell death during CVB3 infection of cardiomyocytes. We described a signaling subnetwork where the ERK1/2, ERK5, and p38 pathways function together as regulators of apoptosis and necrosis in the CVB3-infected cardiomyocytes. In our experimental follow-up experiments we found that by simultaneously inhibiting p38 and ERK1/2 with SB203580 and PD184352, cardiomyocyte death and virus progeny release were greatly reduced (Figures 4-9 and 4-10, A and B). These results suggest a possible combination treatment for patients with acute infection. Animal models of CVB3 infection could be used to address some of these early considerations in justifying p38 and ERK1/2 inhibition as a plausible therapy. Clinically, inhibitors of the p38 pathway have been explored for inflammatory conditions (*249, 250*) and inhibitors targeting the ERK1/2 pathway have shown promise in several cancers (*251*).

Testing whether or not this signaling subnetwork is exclusive to CVB3 infection also warrants further study. CVB3 is one of 29 serotypes of coxsackievirus that are responsible for a diverse range of medical conditions, from pancreatitis to hand, foot, and mouth disease (*252, 253*). Interestingly, coxsackieviruses have also been implicated with type I diabetes (*254, 255*). Elucidating signaling mechanisms of CVB3 may provide insight into these other coxsackievirus-induced conditions and contribute to developing therapies.

6.4 Summary

The human cell-signaling network is a fascinatingly complex system, comprised of a staggering number of components that are controlled by numerous mechanisms. This complexity introduces both a challenge and an opportunity to identify critical signaling nodes in human diseases. Systems approaches to study this complexity at the network level will be important in identifying more effective therapeutics. By understanding the role of individual nodes in the greater network, we can improve our decisions about which targets should enter the costly and time-consuming drug development pipeline. Increasing the predictability of drug target identification and development will decrease the time and cost of finding new therapies that will be more efficacious and less toxic for patients suffering from disease.

CHAPTER 7

Appendices

7.1. Tutorial: Modeling the latent dimensions of multivariate cell-signaling datasets

7.1.1. Overview of data-driven models for cell-signaling datasets

Cellular signal transduction is coordinated by modifications of many proteins within cells. Protein modifications are not independent, because some are connected through shared signaling cascades and others jointly converge upon common cellular functions. This coupling creates a hidden structure within a signaling network that can point to higher-level organizing principles of interest to systems biology. One can identify important covariations within large-scale datasets by using mathematical models that extract latent dimensions—the key structural elements of a measurement set. In this tutorial, we introduce two principal components-based methods for identifying and interpreting latent dimensions. Principal components analysis provides a starting point for unbiased inspection of the major sources of variation within a dataset. Partial least squares regression reorients these dimensions toward a specific hypothesis of interest. Both approaches have been used widely in studies of cell signaling, and they should be standard analytical tools once highly multivariate datasets become straightforward to accumulate.

7.1.2. Introduction

Biology is now awash with large-scale measurements of cell signaling (*205*). High-throughput technologies can readily measure signaling-protein levels and

modification states such as phosphorylation. Moreover, we can observe dozens to thousands of post-translational modifications and how they change with time under different environmental conditions and perturbations (*256-260*). These modifications to signaling proteins propagate information flow through the cell. The question is how best to use these data to uncover patterns of regulation that may suggest how the underlying network operates.

If a large-scale dataset revolves around a single perturbation or stimulus, then "hit lists" ranking the largest-magnitude changes may suffice for gene or pathway discovery. However, when numerous perturbations or stimuli are involved concurrently, it is much harder to link stimulus- or perturbation-induced changes within the cell to phenotypic outcomes. For example, a specific small-molecule inhibitor should strongly block the activity of the target enzyme, but what happens if multiple inhibitors are combined and the cells are also challenged with a microbial pathogen? To make these types of inferences, we need data reflecting complex biological scenarios and a simplified representation of the measurements—we need a data-driven model (*179*).

Complex datasets benefit from models that address the fundamental challenge of dimensionality (*261*). When large spreadsheets of measurements are recast as a vector algebra (*179*), each experimental condition appears as a projection along a set of dimensions defined by the measured variables (see below). If we could inspect the condition-specific projections along all measured variables (e.g., post-translational modifications), then we could possibly discern patterns within the measurements. The problem is that when interpreting highly multivariate datasets with hundreds or thousands of dimensions, we struggle to have intuition beyond the three dimensions that we can see

(*262, 263*). Data-driven models simplify dimensions according to specific quantitative criteria, identifying a small number of "latent variables" that comprise a reduced dimensional space for prediction and analysis.

Dimensionality reduction of signaling data remains an active area of research (*264, 265*), but here we will review two established methods: principal components analysis (PCA) and partial least squares regression (PLSR). PCA and PLSR have been applied to signal transduction over the past several years, but they have a much longer history in data-rich fields such as spectroscopy, econometrics, and food science (*266, 267*). The main distinction between the two methods lies in the overarching goal of the resulting model. PCA is an unsupervised method, meaning that dimensions are reduced based on intrinsic features of the data. Thus, PCA allows the data to "speak for itself", but the corollary is that method is deaf to user input regarding the types of relationships that latent dimensions should uncover. It is here that PLSR excels. As a supervised method, PLSR starts with a hypothetical relationship between variables (dimensions) that are independent and those that are dependent. The algorithm then reduces dimensions to retain the hypothesized relationship as much as can be supported by the data by creating a linear regression model. In contrast to PCA, there are countless user-defined hypotheses that can be tested with PLSR, and models can even predict dependent variables given new input data. Thus, PCA is most useful as an explanatory tool for unbiased discovery of patterns within datasets (*268, 269*), whereas PLSR acts as a predictive tool for linking multivariate inputs to outputs (*19, 134, 270-273*). The limitation of both of these methods is that the predictions are correlative—they provide a guide for causation that must be tested subsequently with mechanistic experiments.

The goal of this tutorial is to provide readers with a working knowledge of PCA–PLSR and their application to cell-signaling datasets. We begin by reviewing the basics of vector and matrix algebra that are essential for understanding how latent dimensions are identified. Then, we will provide a detailed introduction to PCA and PLSR, focusing on the underlying mathematics and the technical considerations important for real biological applications. In parallel, we will walk through an example of dimensionality reduction using a primary signaling dataset from a recent publication (*177*). We conclude with a brief discussion of more advanced approaches and refer interested readers to further literature on these topics.

7.1.3. Vector and matrix algebra

Latent variables align with the important changes in a high-dimensional dataset. A model using latent variables must know where these changes "point" and how far each data point "moves" with respect to them. Therefore, it is natural to treat latent variables as *vectors*, which are defined as mathematical entities with a direction and a magnitude. Using vectors lends to a geometric interpretation where we can compare the angle and distance between the two vectors. Importantly, vector algebra easily generalizes to beyond three dimensions, which is essential for modeling large multivariate datasets. As we expand into higher dimensions, the utility of the vector description and its geometric interpretations becomes even more valuable.

The magnitude of a vector is often referred to as the vector *norm*. For a vector x in n dimensions,

$$x = \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix}$$

the Euclidean norm ($\|x\|$) is calculated as follows:

$$\|x\| = \sqrt{x_1^2 + x_2^2 + \ldots + x_n^2}$$

The vector norm is useful for normalization, since dividing a vector by its norm yields a scaled vector of unit length (or a *unit vector*).

Vectors with the same number of dimensions can be grouped together as a *matrix*. We refer to a matrix X with m rows and n columns as an $m \times n$ matrix. To help distinguish between vectors and matrices, we often name matrices with capital letters and vectors with lowercase letters. (Individual matrix elements use lowercase letters with two subscripts that refer to the row and column of the matrix.) X can thus be viewed as a collection of m row vectors (r), each with n dimensions, or as a collection of n column vectors (c), each with m dimensions:

$$X = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{m1} & x_{m2} & \cdots & x_{mn} \end{bmatrix} = \begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_m \end{bmatrix} = \begin{bmatrix} c_1 & c_2 & \cdots & c_n \end{bmatrix}$$

Matrices directly correspond to the data tables of large-scale experiments. The columns indicate the signaling *variables* measured during the experiments: phospho-proteins, catalytic activities, etc. The rows indicate the experimental *observations* where the variables were measured: a particular growth factor stimulation, time point, genetic

background, etc. Each condition can thus be treated as a row vector, whose individual elements indicate how strongly that row vector projects upon the dimensions specified by the variables.

Treating data tables as matrices allows linear algebra to be performed on the matrix en route to building a data-driven model. However, while a matrix may appear similar to an array, it is important to recall that algebraic operations on matrices and vectors differ from those on arrays. When a matrix A is multiplied by a vector x (Ax), the result is a projection of the matrix onto that vector space. The vector samples the matrix across the shared dimension, combining the products of the individual elements. By extension, when two matrices are multiplied, we combine the products of the row elements of the first matrix with the corresponding column elements in the second matrix across the shared dimension. Thus, matrix and vector multiplication is defined only when the two elements share the inner dimension of the product. For an $m \times p$ matrix A and a $p \times n$ matrix B, the product AB exists and yields an $m \times n$ matrix (with the product summed across the shared inner dimension, p). By contrast, BA does not exist unless $m = n$. A real number can multiply a vector or matrix of any size by scalar multiplication, which simply scales each element of the vector or matrix by the real number.

One convenient transformation that we can perform on a matrix or vector without loss of information is to take its *transpose*, denoted by the superscript ^{T}. The transpose of a matrix is the same matrix where the rows in the matrix become the columns in the transposed matrix and the columns in the original matrix become the rows (for a square matrix, this is equivalent to reflecting the matrix across its main diagonal). So while a 2 × 3 matrix C cannot be multiplied by itself, C can be multiplied by C^{T}:

$$C = \begin{bmatrix} 2 & 1 & 3 \\ 1 & 6 & 5 \end{bmatrix} \quad C^T = \begin{bmatrix} 2 & 1 \\ 1 & 6 \\ 3 & 5 \end{bmatrix} \quad CC^T = \begin{bmatrix} 14 & 23 \\ 23 & 62 \end{bmatrix}$$

If a row or column of a matrix is simply a linear combination of the others, then there is no additional information provided by this row or column. The *rank* of a matrix tells us how many dimensions are truly distinct from one another. If each row of a matrix is an observation, the rank will tell us how many observations provide unique data. The rank also indicates how many variables (columns) are unique. The rank of the above matrix *C* is two, because one column can always be expressed as a linear combination of the other two; for example:

$$\frac{13}{11}\begin{bmatrix} 2 \\ 1 \end{bmatrix} + \frac{7}{11}\begin{bmatrix} 1 \\ 6 \end{bmatrix} = \begin{bmatrix} 3 \\ 5 \end{bmatrix}$$

Vectors in a matrix can be linearly independent, but that does not mean that their directions are non-overlapping. For instance, the second and third column vectors of *C* are independent, but their projection onto one another is nonzero (1×3 + 6×5 = 33), meaning that they partly point in the same direction. When vectors have a zero projection onto one another, they are said to be *orthogonal*. The simplest example of orthogonal vectors is the Cartesian pair of unit vectors: [1 0] and [0 1]. Other two-dimensional vectors can be orthogonal—such as [2 1] and [1 –2]—and the principle of orthogonality extends naturally to higher dimensions. If we can extract orthogonal vectors from a data matrix, these can serve as new axes onto which the data can be projected. The new axes act as latent variables in our experiment and are defined as a

weighted combination of the variables that we measured. We are interested in choosing vectors that are orthogonal because the latent variables they define cannot be projected onto one another, making them more interpretable. This analysis can simplify the data by projecting it onto a smaller set of orthogonal vectors, which identify groups of variables that fluctuate together and describe the data more efficiently.

One way to identify orthogonal vectors is by calculating the *eigenvectors* of a data matrix. The eigenvectors (x) of a square $n \times n$ matrix A are special, in that they return a scaled version of themselves when multiplied by A:

$$Ax = \lambda x$$

where λ is a scalar called an eigenvalue. The eigenvector can be used as a new axis onto which we can project the data, with the eigenvalue indicating how strongly the data is projected onto that "axis". Eigenvectors by definition are always orthogonal to one another; therefore, given a matrix of rank n, there can be up to n eigenvectors with their associated eigenvalues.

We defined eigenvectors for a square matrix where the number of observations equals the number of variables, but what if the data matrix is not square? Here, one can factorize a matrix into the product of three matrices by the process of *singular value decomposition* (SVD):

$$X = U\Sigma V^T$$

U and V^T are the left and right matrices of *singular vectors*, which are mutually orthogonal and are conceptually similar to eigenvalues. $\sum$ is a diagonal matrix comprised

of *singular values*, which are analogous to eigenvalues and correspond to the weightings of the singular vectors. By convention, the singular values in matrix $\sum$ are organized in decreasing order from top left to bottom right. Focusing on the singular vectors with the largest associated singular values, it is possible to extract systemic patterns from large biological datasets (*268*).

7.1.4. Principal components analysis

SVD is not the only way to break down a data matrix into digestible pieces. *Principal components analysis* (PCA) decomposes an $m \times n$ matrix X into the product of two matrices:

$$X = \begin{bmatrix} t_{11} & t_{12} & \cdots & t_{1rank(X)} \\ t_{21} & t_{22} & \cdots & t_{2rank(X)} \\ \vdots & \vdots & \ddots & \vdots \\ t_{m1} & t_{m2} & \cdots & t_{mrank(X)} \end{bmatrix} \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ p_{rank(X)1} & p_{rank(X)2} & & p_{rank(X)n} \end{bmatrix} = TP^T$$

where T is comprised of row vectors called *scores* (t) and P is comprised of column vectors called *loadings* (p). The scores and loadings vectors are ranked by their contribution to the overall variance in the dataset, so that models of the data can be built from the leading score-loading pairs:

$$X \approx \begin{bmatrix} t_{11} \\ t_{21} \\ \vdots \\ t_{m1} \end{bmatrix} \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1n} \end{bmatrix} + \begin{bmatrix} t_{12} \\ t_{22} \\ \vdots \\ t_{m2} \end{bmatrix} \begin{bmatrix} p_{21} & p_{22} & \cdots & p_{2n} \end{bmatrix} + \ldots = t_1 p_1^T + t_2 p_2^T + \ldots$$

The $t_i p_i^T$ vector product indicates the variance captured by the i^{th} *principal component* of the data. Principal components act as the latent variables of a PCA model.

Just as with eigenvectors and singular vectors, a defining characteristic of principal components is their orthogonality: each t is orthogonal from all other score vectors and each p^T is orthogonal from all other loadings vectors. While orthogonality constrains the direction of the vectors with respect to one another, it does not uniquely define the direction of any one vector. This yields an infinite number of possible solutions to the principal-component problem. PCA arrives at a solution by further imposing two criteria: principal components must 1) have loadings vectors of unit length (for convenience) and 2) must explain as much of the variance of the original or residual data as possible. Principal components are calculated iteratively, with first principal component maximizing the variance captured, and the second principal component calculated with the residual information not captured by the first principal component ($X - t_1 p_1^T$). By solving for the principal components that capture the maximum variance, new axes are identified that more efficiently represent the information contained in the original dataset. The early principal components are comprised of loadings vectors paired with the large-magnitude score vectors, indicating that the data is most strongly projected onto these principal component "axes". Later principal components carry small-magnitude score vectors, such as noise among observations, which can be omitted from the PCA model.

There are several conceptual and mathematical similarities between PCA and SVD. The major distinction is that SVD is performed directly on X, whereas PCA is calculated using the covariance matrix of X. Loadings vectors (the columns of P) are the

eigenvectors of the column covariance matrix (calculated as X^TX if X has been preprocessed appropriately; see below), and row vectors (of the scores matrix T) are the eigenvectors of the row covariance matrix (calculated as XX^T). Indeed, if X is preprocessed by column centering (by subtracting the mean of each column from each entry in the column), then P^T obtained by PCA is equivalent to V^T obtained by SVD, with $T = U\sum$. The correspondence breaks down when X is not preprocessed before PCA, because X^TX then carries with it the mean values of the different variables in X (see below).

To illustrate the application of PCA in practice, we selected a primary signaling dataset of phosphoprotein measurements collected by phospho-ELISA (*177*) (Figure 7-1A). HL1 murine cardiomyocytes were pretreated with one of seven signaling inhibitors or 15 inhibitor pairs and then infected with the cardiotoxic virus, coxsackievirus B3 (CVB3). At six time points over 24 hr after CVB3 infection, cells were analyzed for nine phosphoproteins: Akt, ATF2, CREB, ERK, GSK3β, HSP27, IκBα, JNK, and P38. In parallel, the extent of viral progeny release (VPR) was measured by plaque assay as a measure of productive virus infection and propagation within host cells. The VPR measurements will be used later when building a supervised model with PLSR.

We first simplify the dynamic nature of the phosphoprotein measurements by taking the time integral. Although not required for PCA, the time integral provides a simple metric that captures the magnitude and duration of the signaling event (*274*). Integrated signals have been informative in many data-driven models and are thus used here for simplicity (*19, 60, 270, 271*). However, we note that many other metrics have

been explored and could be useful in specific circumstances, such as for multiphasic signaling events (*19, 271*).

Next, we must preprocess the data so that each phosphoprotein has an equal opportunity to contribute to the principal components of the model. As mentioned above, data-driven models are built by retaining the largest variations or covariations within a dataset (*179, 266, 267*). Unprocessed data carries with it the heterogeneous characteristics of the original signaling measurements (*205*), which can bias a model. In the original phospho-ELISA measurements of Figure 7-1A, for example, the mean of the AKT time integral (taken across all inhibitor conditions) is 174.2 and its variance is 1360. By contrast, the mean of P38 is 1043.5 and its variance is 247,880. Using these raw data, P38 will contribute much more to the model relative to AKT, simply because the mean values are larger and they vary more. The contribution is disproportionate, however, because the larger and more-variable P38 measurements may simply reflect a more-sensitive assay. Thus, some form of preprocessing is required to place all the data on equal footing.

The most-common way to preprocess the measurement is to center each variable by subtracting its mean and then scale the centered variable by dividing by its standard deviation. This standardization turns each variable column into a nondimensional z-score, which indicates the number of standard deviations that an observation lies away from the mean. Centering is not appropriate when baseline data are included in the model, because the act of centering will move the baseline from zero to a negative value (*19, 266*). However, in this example, all samples were infected with CVB3 and thus z-scoring is appropriate without uninfected baseline data.

After preprocessing the phospho-ELISA dataset (X), we must now define scores and loadings vectors according to the eigenvalue profile as introduced above. Numerically, this can be achieved by the nonlinear iterative partial least squares (NIPALS) algorithm (*267*). NIPALS starts by selecting a row vector from X and defining it as a provisional score vector (t_1 for the first principal component). X is then projected onto t_i to define a provisional loadings vector (p_1):

$$p_1^T = \frac{t_1^T X}{t_1^T t_1}$$

Next, X is projected onto this loadings vector after normalization ($p_{1,norm}$) to define a new score vector ($t_{i,new}$):

$$p_{1,norm} = \frac{p_1}{\|p_1\|}$$
$$t_{1,new} = X p_{1,norm}$$

If $t_{1,new} = t_1$ within a numerical tolerance, then the NIPALS algorithm has converged, and t_1 and $p_{1,norm}$ represent the scores and loadings vectors for the first principal component. The contribution of the first principal component is then subtracted from X, and the second principal component is derived from the residual in the same way that the first was derived from X. This iterative procedure can continue up to the rank of X, but often the first several principal components (eigenvalues of the covariance matrix) contain the bulk of the standardized variation within the data. For example, the two leading principal components of our phospho-ELISA dataset capture 45% of the overall variation, and the first five capture 87%.

We can graphically portray the projections of the signaling variables onto the PCA principal components via a loadings plot (Figure 7-1B). Variables that are similarly loaded along a principal component share common features. For example, the phospho-kinases ERK, GSK3β, and JNK are all loaded strongly along the second principal component, suggesting that they are regulated together in the context of the viral infection. We would thus expect to see these proteins show a similar behavior pattern under the different treatments. This is most clearly visible in the first two inhibitor combinations (Figure 7-1, A and B, single asterisks) where ERK, GSK3β, and JNK share a strong relative activation. However, the weaker covariation of the three kinases in other inhibitor combinations also informs the analysis and helps to determine their final projection. The simplest explanation for the ERK–GSK3β–JNK grouping is that these pathways are activated by a common upstream signaling event, possibly via one of the initial stages of viral infection. Identifying this signaling event could be a motivation for future experimental studies, which focus on signaling proteins upstream of ERK, GSK3β, and JNK.

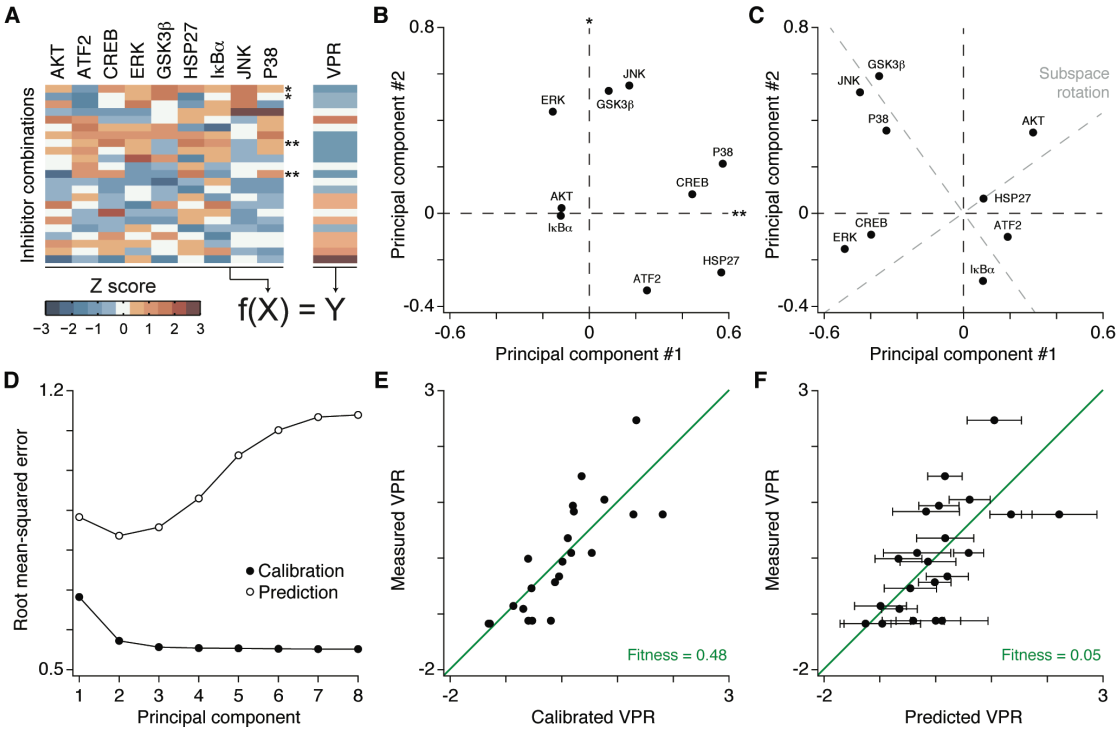


Figure 7-1. PCA–PLSR modeling of phosphoprotein signaling data and cell outcomes. (A) Time-integrated phospho-ELISA data and VPR for a pairwise inhibitor screen in HL1 cardiomyocytes infected with coxsackievirus B3(*1*). Color maps show the variable-specific z-score across all inhibitor conditions, and asterisks indicate specific observations that illustrate the principal components of the PCA model. VPR is regressed as a function of phosphoprotein data in the PLSR model. (B) Loading plot for the first two principal components of a PCA model of the phosphoprotein signaling data. Asterisks link principal component loadings to specific observations in (A). (C) Loading plot for the first two principal components of PLSR model linking phosphoprotein signaling data to VPR. Note how the loadings have shifted relative to the PCA loadings in (B). Gray dashed lines indicate a subspace rotation that aligns principal components more clearly with the starting measured variables. (D) Defining the optimal number of PLSR principal components by leave-one-out cross-validation. Root-mean-squared error of fitted observations (calibration, solid) is shown in comparison to the error of observations omitted during model training (prediction, hollow). Note that calibration errors decrease monotonically with increasing principal components, whereas prediction errors show a minimum with two principal components. (E), (F) PLSR models with acceptable calibration fitness may have poor prediction fitness. Measured VPR data are plotted versus the fitted calibration values or the leave-one-out prediction values of the model. Standard errors on the predictions were estimated by jackknifing.

Another example of grouped phospho-proteins is P38, CREB, and HSP27, which are loaded along the first principal component in part because they are all strongly activated by the eighth and twelfth inhibitor combinations (Figure 7-1, A and B, double asterisks). Note also that ATF2 is positively loaded on the first principal component but negatively loaded on the second principal component, because it is strongly activated by conditions 8 and 12 but inhibited by conditions 1 and 2. AKT and IκBα are not strongly loaded on either principal component, and inspection of the primary data indicates why—neither is consistently activated or inhibited by conditions 1+2 or 8+12. Accordingly, these two phospho-proteins are strongly loaded on the third principal component as part of the residual information not captured by the first two principal components (not shown).

Data-driven modeling results should always be followed up in this way with respect to the primary dataset. It provides a reality check for the model and a more objective appreciation for patterns in the measurements. Although conditions 1+2 or 8+12 could be rationalized as discriminating conditions after PCA, they probably would not be the first patterns identified by visual inspection. Therein lies a major strength of data-driven modeling approaches (*179*).

7.1.5. Partial least squares regression

PLSR extends naturally from PCA with a few important additions. First, rather than decomposing a single matrix into scores and loadings, the data is split into two blocks

that specify a hypothesized relationship between them. The independent block (X) contains the input variables, and the dependent block (Y) contains the output variables. The hypothesis is that the input variables determine the output variables. In our example, we retain the phospho-ELISA measurements as the independent block and further assign the VPR measurements of viral propagation as the dependent block (Figure 7-1A). Thus, our working hypothesis is that the phospho-proteins in X quantitatively control the VPR outcomes in Y.

Both the independent and dependent blocks can now be decomposed into their respective scores-loadings vectors:

$$X = TP^T$$
$$Y = UQ^T$$

The relationship between the two blocks is specified as a linear regression between the scores of the independent block and the dependent block:

$$U \approx TB \Rightarrow Y \approx TBQ^T$$

(Note that, in our example, Y has a rank of one and therefore $Y \approx TB$.) The regression coefficient matrix (B) allows predictions of Y to be made from condition-specific score projections of X.

We could decompose X and Y separately by PCA and regress the principal components afterwards (this approach is termed principal components regression (*267*)). However, PLSR models improve the relationship between the two blocks by implementing a key algorithmic change in the NIPALS algorithm. Rather than have two

PCA-type decompositions of X and Y occur independently, the scores vectors t and u are exchanged during each iteration to define a principal component. X is first projected onto a provisional score vector selected from Y (u_1 for the first principal component) to define a provisional loadings vector (p_1):

$$p_1^T = \frac{u_1^T X}{u_1^T u_1}$$

X is projected onto this loadings vector after normalization to define a provisional score vector for X (t_1), and then Y is projected onto t_1 to define a provisional loadings vector for Y (q_1):

$$p_{1,norm} = \frac{p_1}{\|p_1\|}$$
$$t_1 = X p_{1,norm}$$
$$q_1^T = \frac{t_1^T Y}{t_1^T t_1}$$

Last, Y is projected onto the normalized q_1 to define a new score vector ($u_{1,new}$) that will be projected on X in the next iteration:

$$q_{1,norm} = \frac{q_1}{\|q_1\|}$$
$$u_{1,new} = Y q_{1,norm}$$

The algorithm stops when score vector (t_1) for X has stabilized as described for PCA above. The projection of X onto u_1 and Y onto t_1 is what biases the decomposition of each

block toward the variance of the other. The result is a principal components-based model that is optimized to capture the overall variation in *Y* by using the variation in *X*.

PLSR typically leads to a dramatic reorganization of the loadings vectors relative to those obtained by PCA (*179*). For our phospho-ELISA–VPR model, we see many new groupings of signaling variables (Figure 7-1C). P38 now maps with JNK and GSK3β in the second quadrant of the loadings plot, because all three phospho-kinases have a notable anticorrelation with VPR (Figures 7-1, A and C). ERK and CREB also are slightly anticorrelated with VPR. However, some of the later inhibitor combinations have high ERK–CREB and high VPR, which causes these two proteins to be loaded less heavily on the second principal component compared to P38, JNK, and GSK3β. As with PCA, these types of interpretations are critical for getting the most out of a PLSR model. The difference is that the PLSR loadings highlight specific groups of correlations between the independent and dependent blocks, whereas PCA loadings highlight coordinated variation more generally.

Interpretability of a data-driven model is often improved by "subspace rotation" (*268*). Even though PCA and PLSR principal components are uniquely determined, the factorization is degenerate (just as the number eight can be factored into 8×1, 4×2, etc.). We can exploit this property by rotating the scores and loadings vectors so that they coincide with specific variables and conditions more directly. For example, note in the loadings plot of the phospho-ELISA–VPR model that most of the phospho-proteins are diagonally situated along the first two principal components. By rotating the axes counterclockwise by ~35° (Figure 7-1C, gray), we define a new coordinate system that is equally predictive (provided that the scores vectors and regression coefficients are

similarly rotated) but cleaner in its loadings. In the new coordinate system, the first axis consists of positive contributions from AKT and negative contributions from ERK–CREB, whereas the second axis consists of positive contributions from P38–JNK–GSK3β and negative contributions from IκBα–ATF2. These types of rotations are valuable when seeking to assign biological functions to latent dimensions (*19*, *134*).

One question related to interpretability is how many latent variables should be retained in the model. More principal components will obviously capture more of the training data, but there is a danger of overfitting. Fortunately, PLSR has an objective strategy that uses crossvalidation to identify the optimal number of latent dimensions. Before explaining the crossvalidation procedure, however, we must clarify the distinction between model calibration and model prediction. Model calibration involves the $Y_{cal} \approx TBQ^T$ values associated with observations that were used during the model training. Thus, Y_{cal} indicates fitted values, whose root mean-squared error from the measured values (Y_{meas}) will always decrease when more principal components are used (Figure 7-1D, solid). Conversely, prediction involves the $Y_{pred} \approx TBQ^T$ values associated with observations that were omitted from the model training. Y_{pred} indicates true predictions as long as the input observations are independent from the training set. Importantly, Y_{pred} error *increases* beyond a critical number of principal components (Figure 7-1D, hollow), which is a red flag for overfitting.

During "leave-one-out" crossvalidation, each observation is withheld individually and a PLSR model is built on the remaining $m - 1$ observations. Then, the trained model is asked to predict the one observation withheld from the training, and the withholding-training-prediction cycle is repeated for all m observations. This procedure generates a

crossvalidated prediction for each observation. It also allows an estimate of prediction uncertainty based on how much the calibration fluctuates when the other observations are withheld. The "jackknifed" standard error of the prediction (SE_{Ypred}) (*275*) is defined as:

$$SE_{Ypred} = \sqrt{\frac{m-1}{m}\sum_{i=1}^{m-1}\left(Y - Y_{cal,i}\right)^2}$$

where $Y_{cal,i}$ represents the model calibration from the i^{th} crossvalidation run that includes Y. Other error models exist for estimating Y_{pred} uncertainty when Y has not been measured explicitly (*266*).

To assess overall model stability and predictive ability, we must check how much the TBQ^T values degrade when our observations are used for crossvalidated prediction rather than for calibration (Figure 7-1, E and F). Root mean-squared error is less useful here, because it applies to standardized variables and its magnitude is difficult to interpret. The squared Pearson correlation is often used to compare measured data with calibrations or predictions (*276*). However, we have found that PLSR models often retain a good Pearson correlation even when the predictions are no longer quantitatively accurate (*145*). We thus developed a more-stringent fitness metric (*60*) that quantifies the extent of one-to-one mapping between measurements and calibrations-predictions:

$$\text{Fitness} = 1 - \left[\frac{\sum_{i=1}^{m}(Y_{pred,i} - Y_i)^2}{\sum_{i=1}^{m}\left(Y_{pred,i}\right)^2 - \frac{\left(\sum_{i=1}^{m} Y_{pred,i}\right)^2}{m}}\right]$$

When $Y_{pred} = Y$ for all m observations, the fitness is one, with the value falling to below zero with increasing deviations from this line (Figure 7-1, E and F, green). In our example, we see a reasonable fitness of the calibration (we consider ≥ ~0.5 to be acceptable), but the performance degrades substantially when crossvalidated predictions are considered. Note in Figure 7-1F that the crossvalidated predictions of VPR are still strongly correlated with the measured values ($R = 0.64$), even though the quantitative accuracy is poor. Thus, we conclude from our analysis that this particular PLSR model can be used as a descriptive tool for concisely visualizing patterns in the data, but it should not be used as a predictive tool for new experiments. Publication-quality PLSR models should be strongly and stably predictive (*19, 134, 256, 270, 271, 273*).

7.1.6. Advanced topics and further reading

In this tutorial, we focused on standard approaches for identifying latent dimensions, but there are other more-sophisticated ways of organizing the initial dataset or defining the optimization criterion. One challenge with the flat data-table format is that it struggles to capture the organization of highly systematic datasets (*205*). For example, if multiple signaling measurements are collected at various time points per condition, how is the temporal information kept separate from the signaling information for each condition? One workaround is to unfold the time component as additional variables—either directly or as time-derived metrics (*19*)—and then perform the decomposition. However, the stability of such decompositions can be problematic, and

the resulting latent dimensions do not completely reflect the organization of the original dataset.

To address this problem, Bro developed a multiway variant of PLSR that decomposes data cubes or hypercubes in a way akin to principal components (*277*). Score vectors are retained, but loading vectors are replaced with weight vectors that capture the latent dimensions associated with each "way" of the dataset:

$$X \approx \sum_{i=1}^{n} t_i \prod_{j=1}^{m} w_{i,j}^{T}$$

where n is the number of latent dimensions in the model, m is the number of ways in the dataset (e.g., $m = 2$ for datasets organized by signaling variables and time variables), and $w_{i,j}$ is the weight vector for the i^{th} latent dimension along the j^{th} way of the dataset. The end result is a more-heavily constrained model that allows a direct mapping to the organization of the starting dataset.

There are also several common variants of PCA, which differ from the standard method in the constraints on the latent dimensions. For example, independent components analysis (ICA) requires latent variables to be minimally statistically dependent rather than orthogonal (*278*). The ICA constraint leads to independent components with minimal overlap, in contrast to orthogonal principal components, which are often a weighted blend of the measurement variables. ICA may be valuable in circumstances where collections of measured signaling variables are coordinating distinct functions within the cell.

One simple alteration to PCA is to constrain the elements of the loadings vectors to be positive. This non-negative matrix factorization (*279*) forces latent dimensions to

build upon one another when combined, rather than offset. The non-negativity constraint is useful in fields such as face recognition and linguistics. However, its utility for signal transduction has not been explored, likely because many signaling pathways are known to block the activation of others, undermining the system as a purely additive one.

Other more-elaborate variants of PCA have been developed that incorporate prior biological knowledge. Network component analysis (NCA) uses connectivity information about measured variables to bias the decomposition toward groups of variables with direct interconnections (*280*). In doing so, NCA identifies latent dimensions that are supported by network topology and thus hopefully more mechanistic. The drawback is that NCA requires more observations than PCA, and there are particular connectivity restrictions that may not hold for all networks. However, these types of approaches may become feasible as large-scale systematic measurements of signal transduction become more commonplace (*205*).

Last, we remind that PCA–PLSR models, like all models, are as valuable when they fail as when they succeed. A recent variant of PLSR, called model breakpoint analysis, embraces this view and allows additional predictions to be extracted from data-driven models (*145*). As discussed above, PLSR is fundamentally a linear model, which is valid when there exists a linear mapping between the independent and dependent blocks. By starting with a valid PLSR model, perturbing the data in the independent block with a nonlinear mask applied to each column, and then retraining, one can identify critical failures (breakpoints) where the model abruptly stops making accurate predictions. Model breakpoint analysis delves into these failed models to identify signaling variables whose loadings change coincidently with the breakpoint. The

analysis highlights a small set of variables that might not otherwise be evident in a standard loadings analysis. Model breakpoints and other engineering-inspired "failure analyses" (*281*) will likely become more important as still-higher dimensional data spaces are reduced to their latent variables.

7.1.7. Conclusions

In classical signal processing, electrical engineers work seamlessly between the spatial-time domain and the frequency domain (*282*). Although the information is equivalent, sometimes there are clear analytical advantages to working in one domain compared to the other. The same could be argued for the latent-variable space provided by principal components. By performing an eigenvalue-type decomposition of the original data space, latent dimensions capture the variation and covariation that we most often care about and display this information in an intuitive way.

Signal transduction is fundamentally a multivariate process. Now that experimental platforms have developed to embrace this complexity, the statistical approaches introduced here should become even more important. Already, most signaling experiments are multivariate, be it a mass spectrometry experiment or simply a large panel of immunoblots. Fortunately, signaling biochemists can leverage the tools from other data-rich fields (*266, 267*) that faced the same challenges decades earlier.

7.2. Methods

7.2.1. Chapter 2 Methods

Cell culture

293T, CCRF-CEM, DLD-1, HCT-8, HCT-15, HT-29, AU-565, HCC1500, MCF-7, MDA-MB-231, MDA-MB-361, MDA-MB-468, AsPC-1, BxPC-3, Capan-1, Capan-2, CFPAC-1, HPAF-II, L3.6 pl, Mia PaCa-2, Panc-1, PaTu 8902, PL45, SU.86.86, SW1990, and Yap-C cells were cultured according to ATCC recommendations. MDA-MB-231, MDA-MB-361, and MDA-MB-468 cells were cultured without CO_2. The 5E clone of MCF10A cells was cultured as described (*283, 284*). MDA-MB-436 cells were cultured in L-15 medium with 10% FBS without CO_2. HPDE cells were cultured as described (*285*). A375, HT144, SK-MEL-2, and SLM2 cells were cultured in RPMI medium with 10% FBS and 5% CO_2. HeLa cells were cultured in DMEM medium with 10% FBS and 5% CO_2. DM13, DM122, DM331, SK-MEL-24, VMM18, and VMM39 cells were cultured in RPMI medium with 5% FBS and 5% CO_2.

Primary tissues

Brain and muscle samples were obtained as anonymized, snap-frozen cadaver tissue through the Biorepository and Tissue Research Facility at the University of Virginia.

Primer design

The receptors were selected from the Human Plasma Membrane Receptome (*58*), and receptor RefSeq mRNA sequences were obtained from the National Center for

Biotechnology Information (NCBI). Primers were designed using Primer3 (*286*) with each search constrained to a product size of 150–200 bp, primer sequence lengths of 18–22 bp, and GC content 40–60%. The specificity of primer targets was verified by using NCBI's Basic Local Alignment Search Tool (BLAST). The generality of primers sets was confirmed with the NCBI single-nucleotide polymorphism database to ensure that no reported polymorphism was located at the 3' end of any primer.

Quantitative RT-PCR (qRT-PCR)

RNA from cultured cells was isolated with the RNeasy Plus Mini kit (Qiagen) according to the manufacturer's protocol. RNA from primary tissues was isolated with RNA STAT-60 (Tel-Test) after homogenization on a TissueLyser LT (Qiagen). First-strand cDNA synthesis and qRT-PCR were performed as described (*134*). For qRT-PCR experiments other than those used to generate receptome profiles, data were normalized to the geometric mean of three housekeeping genes, and stability of the normalization was qualitatively assessed with a fourth housekeeping gene among the following candidates: *GAPDH*, *HINT1*, *PPIA*, *PRDX6*, *B2M*, and *GUSB*.

Primer validation

Each primer set was tested on cDNA together with a no reverse-transcription sample to control for genomic contamination and a blank sample to control for primer dimers. The size of any amplicon above a melting temperature of 77 °C was verified by gel electrophoresis to confirm the expected amplicon size. Ambiguous amplicons were gel

purified and analyzed by conventional DNA sequencing. Primer concentration was initially set at 10 pmol per 15 μl reaction and was empirically adjusted to optimize the specificity of amplification.

Receptor profiling assay

Primers were lyophilized in 96-well low-profile PCR plates (Bio-Rad) for 24 hours at 0.110 mbar (Labconco). 10 μl of reverse-transcribed cDNA was diluted in 740 μL of H_2O and mixed on ice with 750 μL of 2× master mix: 2× PCR Buffer II (Applied Biosystems), 8 mM $MgCl_2$, 400 μM dNTP's, 300 μg/ml BSA, 10% glycerol, 0.5× SYBR green (Invitrogen), and 0.05 U/ml Taq polymerase (Roche) (*134*). 15 μL of the master mix-cDNA mixture was loaded into each well of the lyophilized plate, and qRT-PCR was performed on a CFX96 real-time PCR instrument (Bio-Rad) with the following amplification protocol: denaturation at 95 ºC for 90 s; amplification cycles of 95 ºC for 10 s, 60 ºC for 10 s, and 72 ºC for 12 s repeated 40 times; a fusion step of 65 ºC to 95 ºC increased at a rate of 0.1 ºC s^{-1} and measured at 0.5 ºC increments. For each plate, *GAPDH* was used as the loading control and a blank well with no primer served as a negative control.

Raw receptor-profiling data was extracted using CFX Manager 2.0 (Bio-Rad). The baseline for cycle threshold values (C_T) was set at 25 RFU and the baseline for the melting temperature (T_m) estimate was set at 15 –d(RFU) dT^{-1}. These data were exported and compared against a database of T_m ranges for each receptor amplicon to make

present-absent calls based on specificity of the amplification. Absent calls were made only if both duplicate runs were called absent.

For relative quantification, C_T values were normalized to *GAPDH* from each plate and the C_T values from duplicate runs of each receptor were averaged. If only one run was called present, we used the C_T value from that run. For clustering, absent genes were nominally assigned a C_T value that was three cycles higher than the highest C_T value observed for that gene.

Oligonucleotide microarrays

HT-29 cells were plated at 50,000 cells cm^{-2} for 24 hr and stimulated with 200 U ml^{-1} IFN-γ (Roche) for 24 hr. RNA was isolated with the RNeasy Mini Kit (Qiagen), and overall RNA integrity was confirmed with a Bioanalyzer (Agilent). Biotin-labeled cRNA was prepared using the T7-based BioArray HighYield RNA Transcript Labeling Kit (Enzo), and samples were hybridized to GeneChip Human Genome U133A Arrays (Affymetrix) and scanned according to the manufacturer's recommendations.

Processing and analysis of microarray data

The scanned images of HT-29 microarrays were analyzed using Expression Console 1.1 (Affymetrix). The Microarray Suite (MAS) 5.0 algorithm was used to determine present-absent calls of the HT-29 data and the raw breast cancer cell line microarray data (ArrayExpress #E-TABM-157) (*78*). The HT-29 data were then compared to the receptome profiling data as follows. For genes that have multiple probes, a present call

was made if at least one probe with an "_at" or "_a_at" suffix designation was called present for at least one biological sample. An absent call was made if all "_at" and "_a_at" probes for a gene for all three biological samples were called absent. If there were no "_at" or "_a_at" probes for a gene, the present-absent calls were made by using probes with an "_s_at" suffix designation, and if there were no "_s_at" probes, the calls were made by using probes with an "_x_at" suffix.

Exon array

RNA was isolated with the RNeasy Plus Mini Kit (Qiagen), and overall RNA integrity was confirmed on a Bioanalyzer (Agilent). Sense-strand cDNA synthesis was performed with the Ambion WT Expression Kit (Applied Biosystems). Briefly, double-stranded cDNA was synthesized with engineered primers containing a T7 promoter sequence. The cDNA was used as a template for antisense cRNA synthesis by in vitro transcription using T7 RNA polymerase. The cRNA was reverse transcribed with random primers to synthesize single-stranded, sense-strand cDNA. The cDNA was then fragmented, labeled, and hybridized to a GeneChip Human Exon 1.0 ST array (Affymetrix) according the manufacturer's recommendations. The chips were scanned on a GeneChip Scanner 3000 7G (Affymetrix).

Processing and analysis of exon array data

Exon array data were analyzed with the R package JETTA (*81*). The gene expression index for each transcript cluster was calculated for the core probe set after median-GC background correction and normalization by median scaling. Background correction was performed on the exon array data relative to an earlier dataset of 178 human cell lines (GSE29682).

RNA-seq

All RNA sequencing data was generated by the Genomics Services Lab at the HudsonAlpha Institute for Biotechnology (Huntsville, AL). RNA was isolated with the RNeasy Plus Mini Kit (Qiagen), and 1 μg of total RNA was enriched for poly(A)$^+$ transcripts with oligo(dT)$_{25}$ Dynabeads (Invitrogen). Each cDNA library was prepared with the NEBNext first-strand synthesis, second-strand synthesis, end repair, dA tailing, and quick ligation modules (New England Biolabs). Libraries were indexed with standard Illumina-type adapters and sequenced on an Illumina HiSeq 2000 using version 3 reagents that generate 180–200M reads per lane. Samples were 50-bp paired-end sequenced in duplicate at 25M and 50M or 100M reads per sample.

Processing and analysis of RNA-seq data

RNA-seq reads were filtered for signal to noise, assessed for overall quality with FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/), and then mapped using STAR 2.2.0 (*287*) against the human genome build hg19. Reads that map to each gene

were counted with HTSeq (http://www-huber.embl.de/users/anders/HTSeq/) under the union set, whereby reads that do not completely overlap a gene are still counted. The percentage of reads mapping to exons, introns, and untranslated regions was calculated by intersecting the data with features from the hg19 assembly by using BEDTools (*288*). RPKM calculations were performed by normalizing to the median transcript length for each gene and the total library size of each sample. Receptor genes with partial intronic coverage were counted manually using the Integrative Genomics Viewer (*289*).

ROC analysis

ROC curves were generated in R using the ROCR package (*290*).

Plasmids and viral transduction

V5-tagged Luciferase and IL10RA vectors were prepared by recombination of donor plasmids into the lentiviral destination vector pLX302 (*291*) by using LR Clonase (Invitrogen). Donor plasmids were verified by sequencing and recombined plasmids were verified by restriction digest. Lentiviruses were packaged as previously described (*292*). Stably transduced A375 cells were selected with 1 μg ml^{-1} puromycin until control plates had cleared.

IL-10 ELISA

A375 cells were plated at 50,000 cells cm^{-2} for 24 hr and conditioned medium was collected. After centrifuging to remove dead cells, supernatants were analyzed for IL-10 by ELISA (R&D Systems) according to the manufacturer's instructions.

Cell stimulation

HT-29 cells were plated at 50,000 cells cm^{-2} for 24 hr, pretreated with 200 U ml^{-1} IFN-γ (Roche) for 24 hr, and stimulated with 1 μg ml^{-1} FAS crosslinking antibody (APO-1-3, Axxora) for 24 hr. MDA-MB-436 cells were plated at 50,000 cells cm^{-2} for 24 hr and treated with 100 ng ml^{-1} MCSF (Peprotech) or 100 ng ml^{-1} EGF (Peprotech) for the indicated times before lysis. MCF10A-5E cells were plated at 25,000 cells cm^{-2} for 24 hr and treated with 100 ng ml^{-1} MCSF (Peprotech) for 15 min. A375 cells were plated at 50,000 cells cm^{-2} for 24 hr and treated with 50 ng ml^{-1} IL-10 (Peprotech) for 20 min, 1 μg ml^{-1} FAS crosslinking antibody (APO-1-3, Axxora) for 24 hr, or 100 ng ml^{-1} TNF (Peprotech) for the indicated times. 293T cells were plated at 50,000 cells cm^{-2} and MCF7 cells at 25,000 cells cm^{-2} for 24 hr before stimulation with 100 ng ml^{-1} EGF (Peprotech) for 4 hr, 200 U ml^{-1} IFN-γ (Roche) for 4 hr, 5 Gy IR (^{60}Co) for 2 hr, or 20 ng ml^{-1} TNF (Peprotech) for 4 hr.

Western blot analysis

Cells were lysed in RIPA buffer (50 mM Tris [pH 7.5], 150 mM NaCl, 5 mM EDTA, 1% Triton X-100, 0.1% SDS, 0.5% sodium deoxycholate). Equal amounts of clarified

lysates (20 μg) were subjected to SDS-polyacrylamide gel electrophoresis and transferred onto PVDF membranes (Millipore). Membranes were blocked for 1 h in 0.5× blocking solution (Li-Cor) diluted with PBS. Membranes were incubated overnight with primary antibodies recognizing the following proteins or epitopes: FAS (Cell Signaling; 1:1000), CSF1R (Santa Cruz; 1:1000), phosphorylated ERK1/2 (T^{202}/Y^{204}, Cell Signaling; 1:1000), ERK1/2 (Millipore; 1:1000), caspase-3 (Cell Signaling; 1:1000), phosphorylated STAT3 (Y^{705}, Cell Signaling; 1:1000), ERBB3 (Cell Signaling; 1:1000), FGFR1 (Cell Signaling; 1:1000), IL10RA (Millipore; 1:1000), IL-2Rγ (Santa Cruz; 1:1000), or α-tubulin (Abcam; 1:20000 or Cell Signaling; 1:1000). Subsequently, membranes were incubated with secondary IRDye conjugated antibodies (Li-Cor; 1:20,000) or with horseradish peroxidase-conjugated secondary antibodies (Jackson Immunoresearch; 1:10,000). Protein bands were detected by an Odyssey infrared scanner (Li-Cor) or by enhanced chemiluminescence (Pierce) on a ChemiDoc MP camera-based detection system (BioRad). Densitometry of bands was performed in ImageJ.

Hierarchical clustering

Hierarchical clustering was performed with the clustergram function in MATLAB by the unweighted pair group method with a Euclidean distance metric and Ward's linkage.

Statistical analysis

Comparison of the receptor present-absent calls for receptome profiling and microarrays was performed by the Fisher exact test. qRT-PCR time courses were compared by two-

way ANOVA after log transformation to allow for parametric analysis (*293*), using the Sidák correction to account for multiple-hypothesis testing. Individual cell stimulations were compared by unpaired one-sided *t* test after log transformation. Lineage enrichment within the clustered receptome profiles was determined by the hypergeometric test. Receptors with lineage-specific absence must be highly enriched for absence ($P < 0.01$, hypergeometric test), absent in ≥75% of the cell lines comprising that lineage, and present in ≥50% of all cell lines tested. Receptors with lineage-specific presence must be highly enriched for presence ($P < 0.01$, hypergeometric test), present in ≥75% of the cell lines representing that lineage, and absent in ≥50% of the cell lines tested.

7.2.2. Chapter 3 Methods

Computational models of signaling pathways

Differential equation models of three-tiered enzyme cascades were built in MATLAB. The models were built under the assumption of Michaelis-Menten kinetics and the existence of constitutive enzymes that deactivate pathway components. Parameters in the models were optimized to yield the smallest difference between linear cascades when inhibited by RNAi, competitive inhibitor, and a noncompetitive inhibitor.

Cell culture

3T3 and SKW6.4 cells were cultured according to ATCC recommendations.

Plasmids and viral transduction

Hairpins that target both isoforms of MEK were cloned into pSLIK vectors. Hairpin #1 targets the sequence ggccttctacagcgacggcga and hairpin #2 targets the sequence cggcgagatcagcatctgcat. Lentiviruses were packaged as previously described (*292*). DOX-inducible caspase-3 was cloned by PCR into the Tet-tight entry vector pEN_TTmiRc2 (*296*). To generate caspase-3 (C163A) and caspase-6 (C163A), C^{163} was mutated to alanine by site-directed mutagenesis (Quikchange II XL, Stratagene). Stably transduced 3T3 and SKW6.4 cells were selected with 2 μg ml^{-1} puromycin or 300 μg ml^{-1} G418 until control plates had cleared.

Cell Stimulation

3T3 cells were plated at 18,000 cells cm^{-2} overnight followed by a 24 hr serum starvation. Cells were stimulated with the specified amounts of FAS crosslinking antibody (APO-1-3, Axxora), PDGF (Peprotech), EGF (Peprotech), U0126 (Calbiochem), FR180204 (Tocris), or GSK1120212 (Selleck) for the times indicated. SKW6.4 cells were suspended at $4*10^5$ cells/ml and treated with 50 μM pep419 (Millipore) or the indicated amounts of FAS crosslinking antibody (APO-1-3, Axxora).

Western blot analysis

Cells were lysed in RIPA buffer (50 mM Tris [pH 7.5], 150 mM NaCl, 5 mM EDTA, 1% Triton X-100, 0.1% SDS, 0.5% sodium deoxycholate). Equal amounts of clarified lysates (20 μg) were subjected to SDS-polyacrylamide gel electrophoresis and transferred onto PVDF membranes (Millipore). Membranes were blocked for 1 h in 0.5× blocking solution (Li-Cor) diluted with PBS. Membranes were incubated overnight with primary

antibodies recognizing the following proteins or epitopes: phosphorylated ERK1/2 (T^{202}/Y^{204}, Cell Signaling; 1:1000), ERK1/2 (Millipore; 1:1000), caspase-3 (Cell Signaling; 1:1000), cleaved Lamin A (Cell Signaling Technology, 1:1000), cleaved PARP (BD Biosciences, 1:1000), MEK1/2 (Cell Signaling Technology, 1:000), phosphorylated MEK1/2 (Cell Signaling Technology, 1:1000), FLAG (Sigma, 1:1000), Raf-1 (Santa Cruz, 1:1000) or α-tubulin (Abcam; 1:20000 or Cell Signaling; 1:1000). Subsequently, membranes were incubated with secondary IRDye conjugated antibodies (Li-Cor; 1:20,000). Protein bands were detected by an Odyssey infrared scanner (Li-Cor). Densitometry of bands was performed in ImageJ.

Immunofluorescence

3T3 cells were allowed to adhere overnight to glass coverslips treated with poly-D-lysine (Sigma) followed by a 24 hr serum starvation and stimulation with PDGF (Peprotech) as indicated. Immunofluorescence was performed as previously described with pRB antibody (Cell Signaling Technology, 1:1000).

7.2.3. Chapter 4 Methods

Plasmids

DOX-inducible MEK5 (*294*) and MKK6-EE (*295*) were cloned by PCR into the Tet-tight entry vector pEN_TTmiRc2 (*296*). To generate MEK5-DD, S^{311} and T^{315} of MEK5 were both mutated to aspartate by site-directed mutagenesis (Quikchange II XL, Stratagene). All entry vectors were verified by sequencing, and lentiviral vectors were cloned by LR

recombination into pSLIK neo (*296*). pLKO.1 puro shERK5 lentiviral vectors (TRCN0000023234 and TRCN0000023236) were obtained from Open Biosystems.

Cells and Viruses

HL1 cells were provided by Dr. William Claycomb (Louisiana State University Health Sciences Center, New Orleans, USA) (*297*). 293T cells were obtained from ATCC. CVB3 (Kandolf strain) was propagated in HeLa cells, and virus titers were determined by plaque assay. Retroviruses and lentiviruses were packaged as previously described (*292*). Stably transduced HL1 cells were selected with 4 μg/ml puromycin or 150 μg/ml G418 until control plates had cleared.

Viral Infection and Perturbations

HL1 cells were sham-infected with PBS or infected with CVB3 at M.O.I. = 0.5, 1.5, 4.5, 9, or 18 and cell extracts were prepared at 0, 0.17, 1, 8, 16, and 24 hr. For perturbation experiments, the following chemical inhibitors were added one hour before infection: SB203580 (20 μM, Tocris Biosciences), BIRB796 (5 μM, Selleck Chemicals), XMD8-92 (5 μM, Axon Medchem), U0126 (20 μM, Tocris Biosciences), PD184352 (2 μM, Santa Cruz Biotechnology), DEVD-CHO (0.1 μM, EMD), and Necrostatin-1 (50 μM, Calbiochem).

Plaque Assays

CVB3 titers from triplicate cell supernatants were determined on monolayers of HeLa cells by an agar overlay plaque assay as described elsewhere (*177*).

p-ELISA

Cell lysates were normalized to protein concentration and analyzed by p-ELISA (Biosource) for the phosphorylation levels of Akt (S^{473}), ATF2 (T^{69}/T^{71}), CREB (S^{133}), ERK1/2 (T^{185}/Y^{187}), GSK3β (S^{9}), Hsp27 (S^{82}), IκBα (S^{32}), and p38 MAPK (T^{180}/Y^{182}) according to the manufacturer's instruction.

Caspase-3, -8 and -9 Activity Assays

Caspase activities were measured according to the manufacturer's instruction (R&D Systems) as described elsewhere (*193*). Fluorescence was measured at excitation and emission wavelengths of 485 nm and 535 nm, respectively, using a ***Tecan GENios*** fluorescent reader.

Western blot analysis

Western blot analysis was performed as described previously (*177*) with one of the following primary antibodies: anti-p-ERK1/2 (T^{202}/Y^{204}, Cell Signaling, 1:1000), anti-p-ERK5 (T^{218}/Y^{220}, Cell Signaling, 1:1000), anti-p-p38 (T^{180}/T^{182}, Cell Signaling, 1:1000), anti-p-MAPKAPK2 (T^{334}, Cell Signaling, 1:1000), anti-VP1 (Dako, 1:1000), anti-cleaved

caspase-8 (Cell Signaling, 1:1000), anti-caspase-9 (Cell Signaling, 1:1000), anti-caspase-3 (Cell Signaling, 1:1000), anti-β-actin (Sigma, 1:5000), anti-p-Akt (S^{473}, Cell Signaling, 1:1000), anti-p-GSK3β (S^{9}, Cell Signaling, 1:1000), anti-p-ATF2 (T^{69}/T^{71}, Cell Signaling, 1:1000), anti-p-CREB (S^{133}, Cell Signaling, 1:1000), anti-p-IκBα (S^{32}, Cell Signaling, 1:1000), anti-p-Hsp27 (S^{82}, Cell Signaling, 1:1000), anti-p-RSK (T^{359}/S^{363}, 1:1000), anti-MEK5 (StressGen, 1:1000), anti-HA (Roche, 1:1000), anti-HMGB-1 (Epitomics, 1:1000), anti-p-MEF2A (T^{312}, Abcam, 1:1000), anti-cleaved caspase-8 (Cell Signaling, 1:1000), or anti-tubulin (Cell Signaling, 1:5000 or Abcam, 1:20000) for 1 hr or overnight, followed by incubation for 1 hr with horseradish peroxidase-conjugated secondary antibodies (Santa Cruz) or infrared dye-conjugated secondary antibodies (Licor). Immunoreactive bands were visualized by enhanced chemiluminescence (Pierce, Rockford, IL) on ChemiGenius2 or ChemiDoc MP camera-based detection systems or by infrared fluorescence on an Odyssey infrared imaging system. Where indicated, band intensities were quantified by densitometry with ImageJ, and all blotting results were replicated with at least one additional set of independent biological samples.

Cell Viability Assays

HL1 cells were grown in 12-well plates and infected with CVB3 (M.O.I. = 9) for 16 and 24 hr after pretreatment with inhibitors. The MTS solutions (1:5) were added to wells for 2.5 hr and then transferred to 96-well plates. Cell viabilities of infected cells and non-infected were assessed by MTS assay (CellTiter 96; Promega, Inc., Madison, WI). Amine-reactive labeling was performed with the LIVE/DEAD fixable violet dead stain (Invitrogen) according to the manufacturer's recommendations. For flow cytometry,

cells were labeled in suspension, washed with PBS + 0.1% Tween-20, and analyzed on a BD FACSCalibur equipped with 407 nm violet laser excitation. For microscopy, adherent cells were labeled, washed with PBS, permeabilized with 0.3% Triton X-100 in PBS, and counterstained with DRAQ-5 before imaging by widefield microscopy as described previously (*292*).

Partial Least Squares Regression

Phospho-proteins (predictor variables) and readouts (response variables) were standardized as z-scores, and the phospho-protein time course was time-integrated over early (0–8 hr) and late (8–24 hr) phases. Partial least squares regression was performed with the "plsregress" function in MATLAB by standard approaches (*19, 179*). The stability of the model was assessed by fivefold leave-one-out cross-validation.

Statistical analysis

All hypothesis testing was performed with Welch's one- or two-sided *t* test at a significance level of $\alpha = 0.05$.

7.2.4. Chapter 5 Methods

Cell culture

HT29 cells were cultured according to ATCC recommendations.

Plasmids

Sequences for triple epitope tags (FLAG, HA, Glu-Glu, AU1, and VSVG) were cloned into pGEX-4T1 vector (GE Healthcare). Sequences for substrates (IκBα(1-62), Hsp27, RSK(386-752), c-jun(1-79), GSK3α (1-97)) were subsequently cloned in to the pGEX-4T1 vectors containing the triple epitope tags by PCR. All vectors were verified by sequencing.

Recombinant protein purification

pGEX-4T1 vectors were transformed into chemically competent C41 bacteria and induced with isopropyl ß-D-1-thiogalactopyranoside (IPTG) (induction conditions are detailed in Table 5-3). Pelleted cells were lysed in 7.5ml TNE lysis buffer (50 mM Tris pH 7.4, 150 mM NaCl, 1 mM EDTA, 10 μg/ml leupeptin, 10 μg/ml aprotinin, 1 μg/ml pepstatin) per 250 ml culture, and lysed with lysozyme and deoxycholate. Lysates were clarified by centrifugation and incubated at 4*C overnight with glutathione-coated agarose beads (Sigma). The beads were then washed three times with ice-cold PBS with 0.05% Triton X-100 and twice with PBS. Proteins were cleaved from the beads by thrombin digest (EMD) for the indicated times (Table 5-3) or eluted with 10 mM glutathione in 50 mM Tris pH 8.0.

Western blot analysis

Cells were lysed in RIPA buffer (50 mM Tris [pH 7.5], 150 mM NaCl, 5 mM EDTA, 1% Triton X-100, 0.1% SDS, 0.5% sodium deoxycholate). Equal amounts of clarified lysates (20 μg) were subjected to SDS-polyacrylamide gel electrophoresis and transferred

onto PVDF membranes (Millipore). Membranes were incubated overnight with primary antibodies recognizing the following proteins or epitopes: anti-p-Hsp27 (Cell Signaling, 1:1000), anti-p- IκBα (Cell Signaling, 1:1000), anti-p-c-jun (Cell Signaling, 1:1000), anti-p-GSK3α (Cell Signaling, 1:1000), anti-pRSK (Cell Signaling, 1:1000), anti-FLAG (Sigma, 1:1000), or anti-HA (Roche, 1:1000), followed by incubation for 1 hr with horseradish peroxidase-conjugated secondary antibodies (Santa Cruz). Immunoreactive bands were visualized by enhanced chemiluminescence (Pierce, Rockford, IL) on ChemiDoc MP camera-based detection system.

Cell stimulation and lysis

HT29 cells were plated at 50,000 cells cm^2 and stimulated with 100 ng ml^{-1} TNF (Peprotech) for 15 minutes, or 100 ng ml^{-1}EGF (Peprotech) or 500 ng ml^{-1} insulin (Sigma) for 5 minutes. Cells were lysed as previously described (Janes MCP 2003).

In vitro kinase activity assay

Protein A/G microtiter plates were coated overnight with 10 μg/ml anti-kinase antibodies (see Table 5-2) and washed three times with blocking buffer (1% bovine serum albumin (Sigma) in 50 mM Tris-HCL (pH 7.5), 150 mM NaCl, 0.05% Triton X-100). HT29 cell lysates (200 μg) were then added overnight then washed two times with wash buffer (50 mM Tris-HCl (pH 7.5), 15 mM $MgCl_2$, 5 mM ß-glycerophosphate, 1 mM EGTA, 0.2 mM Na_3VO_4, 0.2 mM DTT). The wells were resuspended in 20 μl kinase assay buffer (kinase wash buffer plus 1 mM ATP). Recombinant substrate (1 μg in 20 μl) was then

added to the wells to initiate the reaction. The kinase reactions were allowed to proceed for 120 min at 37° C, then terminated by 60 μl 20 mM EDTA.

Readout of kinase activity assay on Luminex

Filter microplates were prewet (Millipore) with 150 ul blocking buffer (1% bovine serum albumin in PBS) and vacuum-filtered. Conjugated microsphere (2,500 per well per kinase) were added in blocking buffer. Wells were filter-washed 2X with wash buffer (0.05% Tween-20 in PBS) and 25 ng kinase assay sample was added per well. Plates were shaken at 1,100 rpm for 30 sec then at 300 rpm for 1 hr in the dark. Wells were vacuum-filtered and washed 2X with wash buffer. Biotinylated secondary antibody (1:10,000) was added to the wells. Plates were shaken at 1,100 rpm for 30 sec then at 300 rpm for 1 hr in the dark. Wells were vacuum-filtered and washed 2X with wash buffer. Streptavidin-PE (1:10,000) was added in blocking buffer. Plates were shaken at 1,100 rpm for 30 sec then at 300 rpm for 10 min in the dark. Wells were vacuum-filtered and washed 2X with wash buffer. Wells were resuspended in blocking buffer and incubated overnight. Microplates were read on a Luminex 100 system on the high PMT setting.

Preparation of antibody conjugated Luminex beads

Anti-tag antibodies (see Table 5-2) were coupled by EDC/Sulfo-NHS to carboxylated microspheres (Luminex) according to the manufacturer's protocol.

7.3. References

1. F. S. Garmaroudi *et al*., Pairwise network mechanisms in the host signaling response to coxsackievirus B3 infection. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 17053 (Sep 28, 2010).
2. C. M. Carthy *et al*., Caspase activation and specific cleavage of substrates after coxsackievirus B3-induced cytopathic effect in HeLa cells. *Journal of virology* **72**, 7669 (Sep, 1998).
3. M. Esfandiarei *et al*., Protein kinase B/Akt regulates coxsackievirus B3 replication through a mechanism which is not caspase dependent. *Journal of virology* **78**, 4289 (Apr, 2004).
4. H. Luo *et al*., Coxsackievirus B3 replication is reduced by inhibition of the extracellular signal-regulated kinase (ERK) signaling pathway. *Journal of virology* **76**, 3365 (Apr, 2002).
5. C. A. Schneider, W. S. Rasband, K. W. Eliceiri, NIH Image to ImageJ: 25 years of image analysis. *Nature methods* **9**, 671 (Jul, 2012).
6. D. Ramskold, E. T. Wang, C. B. Burge, R. Sandberg, An abundance of ubiquitously expressed genes revealed by tissue transcriptome sequence data. *PLoS Comput Biol* **5**, e1000598 (Dec, 2009).
7. D. Yang *et al*., Viral myocarditis: identification of five differentially expressed genes in coxsackievirus B3-infected mouse heart. *Circulation research* **84**, 704 (Apr 2, 1999).
8. X. Si *et al*., Stress-activated protein kinases are involved in coxsackievirus B3 viral progeny release. *Journal of virology* **79**, 13875 (Nov, 2005).
9. M. Esfandiarei *et al*., Novel role for integrin-linked kinase in modulation of coxsackievirus B3 replication and virus-induced cardiomyocyte injury. *Circ Res* **99**, 354 (Aug 18, 2006).
10. M. Esfandiarei *et al*., Coxsackievirus B3 activates nuclear factor kappa B transcription factor via a phosphatidylinositol-3 kinase/protein kinase B-dependent pathway to improve host cell viability. *Cell Microbiol* **9**, 2358 (Oct, 2007).
11. D. Marchant *et al*., Bosentan enhances viral load via endothelin-1 receptor type-A-mediated p38 mitogen-activated protein kinase activation while improving cardiac function during coxsackievirus-induced myocarditis. *Circulation research* **104**, 813 (Mar 27, 2009).
12. A. M. Reimold, J. Kim, R. Finberg, L. H. Glimcher, Decreased immediate inflammatory gene induction in activating transcription factor-2 mutant mice. *International immunology* **13**, 241 (Feb, 2001).
13. P. Blume-Jensen, T. Hunter, Oncogenic kinase signalling. *Nature* **411**, 355 (May 17, 2001).
14. A. Takashima, GSK-3 is essential in the pathogenesis of Alzheimer's disease. *Journal of Alzheimer's disease : JAD* **9**, 309 (2006).
15. J. Y. Lee *et al*., Tumor suppressor p53 plays a key role in induction of both tristetraprolin and let-7 in human cancer cells. *Nucleic acids research* **41**, 5614 (Jun 1, 2013).

16. F. J. Lopez-Diaz *et al*., Coordinate Transcriptional and Translational Repression of p53 by TGF-beta1 Impairs the Stress Response. *Molecular cell* **50**, 552 (May 23, 2013).
17. M. C. Ravnan, M. S. Matalka, Vemurafenib in patients with BRAF V600E mutation-positive advanced melanoma. *Clinical therapeutics* **34**, 1474 (Jul, 2012).
18. W. E. Evans, H. L. McLeod, Pharmacogenomics--drug disposition, drug targets, and side effects. *The New England journal of medicine* **348**, 538 (Feb 6, 2003).
19. K. A. Janes *et al*., A systems model of signaling identifies a molecular basis set for cytokine-induced apoptosis. *Science* **310**, 1646 (Dec 9, 2005).
20. I. L. Weissman, Stem cells: units of development, units of regeneration, and units in evolution. *Cell* **100**, 157 (Jan 7, 2000).
21. J. C. Venter *et al*., The sequence of the human genome. *Science* **291**, 1304 (Feb 16, 2001).
22. G. Manning, D. B. Whyte, R. Martinez, T. Hunter, S. Sudarsanam, The protein kinase complement of the human genome. *Science* **298**, 1912 (Dec 6, 2002).
23. C. L. Vogel *et al*., Efficacy and safety of trastuzumab as a single agent in first-line treatment of HER2-overexpressing metastatic breast cancer. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* **20**, 719 (Feb 1, 2002).
24. P. B. Chapman *et al*., Improved survival with vemurafenib in melanoma with BRAF V600E mutation. *The New England journal of medicine* **364**, 2507 (Jun 30, 2011).
25. B. J. Druker *et al*., Activity of a specific inhibitor of the BCR-ABL tyrosine kinase in the blast crisis of chronic myeloid leukemia and acute lymphoblastic leukemia with the Philadelphia chromosome. *The New England journal of medicine* **344**, 1038 (Apr 5, 2001).
26. C. P. Adams, V. V. Brantner, Estimating the cost of new drug development: is it really 802 million dollars? *Health affairs* **25**, 420 (Mar-Apr, 2006).
27. J. A. DiMasi, R. W. Hansen, H. G. Grabowski, The price of innovation: new estimates of drug development costs. *Journal of health economics* **22**, 151 (Mar, 2003).
28. I. Kola, J. Landis, Can the pharmaceutical industry reduce attrition rates? *Nature reviews. Drug discovery* **3**, 711 (Aug, 2004).
29. O. Fedorov, S. Muller, S. Knapp, The (un)targeted cancer kinome. *Nature chemical biology* **6**, 166 (Mar, 2010).
30. J. A. Sosman *et al*., Survival in BRAF V600-mutant advanced melanoma treated with vemurafenib. *The New England journal of medicine* **366**, 707 (Feb 23, 2012).
31. H. Kitano, Systems biology: a brief overview. *Science* **295**, 1662 (Mar 1, 2002).
32. J. van der Greef, R. N. McBurney, Innovation: Rescuing drug discovery: in vivo systems pathology and systems pharmacology. *Nature reviews. Drug discovery* **4**, 961 (Dec, 2005).
33. K. Dettmer, B. D. Hammock, Metabolomics--a new exciting field within the "omics" sciences. *Environmental health perspectives* **112**, A396 (May, 2004).
34. M. R. Wenk, Lipidomics: new tools and applications. *Cell* **143**, 888 (Dec 10, 2010).

35. O. Morozova, M. A. Marra, Applications of next-generation sequencing technologies in functional genomics. *Genomics* **92**, 255 (Nov, 2008).
36. J. Shendure, E. Lieberman Aiden, The expanding scope of DNA sequencing. *Nature biotechnology* **30**, 1084 (Nov, 2012).
37. A. K. Bose, K. A. Janes, A high-throughput assay for phosphoprotein-specific phosphatase activity in cellular extracts. *Molecular & cellular proteomics : MCP* **12**, 797 (Mar, 2013).
38. M. B. Yaffe, The scientific drunk and the lamppost: massive sequencing efforts in cancer discovery and treatment. *Science signaling* **6**, pe13 (Apr 2, 2013).
39. D. Venet, J. E. Dumont, V. Detours, Most random gene expression signatures are significantly associated with breast cancer outcome. *PLoS computational biology* **7**, e1002240 (Oct, 2011).
40. B. Schoeberl, C. Eichler-Jonsson, E. D. Gilles, G. Muller, Computational modeling of the dynamics of the MAP kinase cascade activated by surface and internalized EGF receptors. *Nature biotechnology* **20**, 370 (Apr, 2002).
41. K. A. Janes *et al.*, The response of human epithelial cells to TNF involves an inducible autocrine cascade. *Cell* **124**, 1225 (Mar 24, 2006).
42. K. A. Ryall *et al.*, Network reconstruction and systems analysis of cardiac myocyte hypertrophy signaling. *The Journal of biological chemistry* **287**, 42259 (Dec 7, 2012).
43. J. Downward, The ins and outs of signalling. *Nature* **411**, 759 (Jun 14, 2001).
44. R. N. Fabricant, J. E. De Larco, G. J. Todaro, Nerve growth factor receptors on human melanoma cells in culture. *Proc Natl Acad Sci U S A* **74**, 565 (Feb, 1977).
45. T. Uyemura, H. Takagi, T. Yanagida, Y. Sako, Single-molecule analysis of epidermal growth factor signaling that leads to ultrasensitive calcium response. *Biophys J* **88**, 3720 (May, 2005).
46. T. Tani *et al.*, Trafficking of a ligand-receptor complex on the growth cones as an essential step for the uptake of nerve growth factor at the distal end of the axon: a single-molecule analysis. *J Neurosci* **25**, 2181 (Mar 2, 2005).
47. M. A. Bothwell, A. L. Schechter, K. M. Vaughn, Clonal variants of PC12 pheochromocytoma cells with altered response to nerve growth factor. *Cell* **21**, 857 (Oct, 1980).
48. R. J. Davis, Independent mechanisms account for the regulation by protein kinase C of the epidermal growth factor receptor affinity and tyrosine-protein kinase activity. *J Biol Chem* **263**, 9462 (Jul 5, 1988).
49. E. Lundberg *et al.*, Defining the transcriptome and proteome in three functionally different human cell lines. *Mol Syst Biol* **6**, 450 (Dec 21, 2010).
50. I. Ben-Shlomo, R. Rauch, O. Avsian-Kretchmer, A. J. Hsueh, Matching receptome genes with their ligands for surveying paracrine/autocrine signaling systems. *Mol Endocrinol* **21**, 2009 (Aug, 2007).
51. L. Shi *et al.*, The MicroArray Quality Control (MAQC) project shows inter- and intraplatform reproducibility of gene expression measurements. *Nat Biotechnol* **24**, 1151 (Sep, 2006).
52. K. A. Janes, RUNX1 and its understudied role in breast cancer. *Cell Cycle* **10**, 3461 (Oct 15, 2011).

53. Z. Wang, M. Gerstein, M. Snyder, RNA-Seq: a revolutionary tool for transcriptomics. *Nat Rev Genet* **10**, 57 (Jan, 2009).
54. M. Sultan *et al*., A global view of gene activity and alternative splicing by deep sequencing of the human transcriptome. *Science* **321**, 956 (Aug 15, 2008).
55. A. Mortazavi, B. A. Williams, K. McCue, L. Schaeffer, B. Wold, Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat Methods* **5**, 621 (Jul, 2008).
56. D. Hebenstreit *et al*., RNA sequencing reveals two major classes of gene expression levels in metazoan cells. *Mol Syst Biol* **7**, 497 (2011).
57. A. Hansen *et al*., Sensitive and specific method for detecting G protein-coupled receptor mRNAs. *Nat Methods* **4**, 35 (Jan, 2007).
58. I. Ben-Shlomo, S. Yu Hsu, R. Rauch, H. W. Kowalski, A. J. Hsueh, Signaling receptome: a genomic and evolutionary perspective of plasma membrane receptors involved in signal transduction. *Sci STKE* **2003**, RE9 (Jun 17, 2003).
59. E. E. Karrer *et al*., In situ isolation of mRNA from individual plant cells: creation of cell-specific cDNA libraries. *Proc Natl Acad Sci U S A* **92**, 3814 (Apr 25, 1995).
60. S. Gaudet *et al*., A Compendium of Signals and Responses Triggered by Prodeath and Prosurvival Cytokines. *Mol Cell Proteomics* **4**, 1569 (Oct, 2005).
61. W. K. Wu *et al*., Expression of ErbB receptors and their cognate ligands in gastric and colon cancer cell lines. *Anticancer Res* **29**, 229 (Jan, 2009).
62. H. J. Cho *et al*., Conjugated linoleic acid inhibits cell proliferation and ErbB3 signaling in HT-29 human colon cell line. *American journal of physiology. Gastrointestinal and liver physiology* **284**, G996 (Jun, 2003).
63. S. Elliott *et al*., Lack of expression and function of erythropoietin receptors in the kidney. *Nephrology, dialysis, transplantation : official publication of the European Dialysis and Transplant Association - European Renal Association*, (Dec 13, 2011).
64. S. Nagata, P. Golstein, The Fas death factor. *Science* **267**, 1449 (Mar 10, 1995).
65. H. W. Na, W. S. Shin, A. Ludwig, S. T. Lee, The Cytosolic Domain of Protein-tyrosine Kinase 7 (PTK7), Generated from Sequential Cleavage by a Disintegrin and Metalloprotease 17 (ADAM17) and gamma-Secretase, Enhances Cell Proliferation and Migration in Colon Cancer Cells. *The Journal of biological chemistry* **287**, 25001 (Jul 20, 2012).
66. F. E. Ahmed *et al*., Differences in mRNA and microRNA microarray expression profiles in human colon adenocarcinoma HT-29 cells treated with either Intensity-modulated Radiation Therapy (IMRT), or Conventional Radiation Therapy (RT). *Cancer genomics & proteomics* **6**, 109 (Mar-Apr, 2009).
67. A. Sturm *et al*., CXCL8 modulates human intestinal epithelial cells through a CXCR1 dependent pathway. *Cytokine* **29**, 42 (Jan 7, 2005).
68. A. Li, M. L. Varney, R. K. Singh, Constitutive expression of growth regulated oncogene (gro) in human colon carcinoma cells with different metastatic potential and its role in regulating their metastatic phenotype. *Clinical & experimental metastasis* **21**, 571 (2004).
69. N. J. Jordan *et al*., Expression of functional CXCR4 chemokine receptors on human colonic epithelial cells. *J Clin Invest* **104**, 1061 (Oct, 1999).

70. M. B. Dwinell, L. Eckmann, J. D. Leopard, N. M. Varki, M. F. Kagnoff, Chemokine receptor expression by human intestinal epithelial cells. *Gastroenterology* **117**, 359 (Aug, 1999).
71. A. C. Stevens *et al.*, Interleukin-15 signals T84 colonic epithelial cells in the absence of the interleukin-2 receptor beta-chain. *The American journal of physiology* **272**, G1201 (May, 1997).
72. M. Nagano *et al.*, Expression of prolactin and growth hormone receptor genes and their isoforms in the gastrointestinal tract. *The American journal of physiology* **268**, G431 (Mar, 1995).
73. J. Dambacher *et al.*, Interleukin 31 mediates MAP kinase and STAT1/3 activation in intestinal epithelial cells and its expression is upregulated in inflammatory bowel disease. *Gut* **56**, 1257 (Sep, 2007).
74. G. Melmed *et al.*, Human intestinal epithelial cells are broadly unresponsive to Toll-like receptor 2-dependent bacterial ligands: implications for host-microbial interactions in the gut. *J Immunol* **170**, 1406 (Feb 1, 2003).
75. J. Xu, Y. Yang, C. Wang, B. Jiang, Rotavirus and coxsackievirus infection activated different profiles of toll-like receptors and chemokines in intestinal epithelial cells. *Inflammation research : official journal of the European Histamine Research Society ... [et al.]* **58**, 585 (Sep, 2009).
76. M. T. Abreu-Martin *et al.*, Fas activates the JNK pathway in human colonic epithelial cells: lack of a direct role in apoptosis. *Am J Physiol* **276**, G599 (Mar, 1999).
77. H. Yin *et al.*, Cyclooxygenase-independent effects of aspirin on HT-29 human colon cancer cells, revealed by oligonucleotide microarrays. *Biotechnology letters* **28**, 1263 (Aug, 2006).
78. R. M. Neve *et al.*, A collection of breast cancer cell lines for the study of functionally distinct cancer subtypes. *Cancer Cell* **10**, 515 (Dec, 2006).
79. T. A. Clark *et al.*, Discovery of tissue-specific exons using comprehensive human exon microarrays. *Genome Biol* **8**, R64 (2007).
80. K. Kapur, Y. Xing, Z. Ouyang, W. H. Wong, Exon arrays provide accurate assessments of gene expression. *Genome Biol* **8**, R82 (2007).
81. J. Seok, W. Xu, H. Gao, R. W. Davis, W. Xiao, JETTA: junction and exon toolkits for transcriptome analysis. *Bioinformatics* **28**, 1274 (May 1, 2012).
82. T. Sato *et al.*, Overexpression of the fibroblast growth factor receptor-1 gene correlates with liver metastasis in colorectal cancer. *Oncol Rep* **21**, 211 (Jan, 2009).
83. D. T. Ross *et al.*, Systematic variation in gene expression patterns in human cancer cell lines. *Nat Genet* **24**, 227 (Mar, 2000).
84. M. A. Olayioye, R. M. Neve, H. A. Lane, N. E. Hynes, The ErbB signaling network: receptor heterodimerization in development and cancer. *EMBO J* **19**, 3159 (Jul 3, 2000).
85. H. Seidl *et al.*, Profiles of chemokine receptors in melanocytic lesions: de novo expression of CXCR6 in melanoma. *Human pathology* **38**, 768 (May, 2007).
86. Q. Chen, V. Daniel, D. W. Maher, P. Hersey, Production of IL-10 by melanoma cells: examination of its role in immunosuppression mediated by melanoma. *Int J Cancer* **56**, 755 (Mar 1, 1994).

87. S. Kruger-Krasagakes *et al*., Expression of interleukin 10 in human melanoma. *Br J Cancer* **70**, 1182 (Dec, 1994).
88. H. Sumimoto, F. Imabayashi, T. Iwata, Y. Kawakami, The BRAF-MAPK signaling pathway is essential for cancer-immune evasion in human melanoma cells. *J Exp Med* **203**, 1651 (Jul 10, 2006).
89. B. Sen, B. Saigal, N. Parikh, G. Gallick, F. M. Johnson, Sustained Src inhibition results in signal transducer and activator of transcription 3 (STAT3) activation and cancer cell survival via altered Janus-activated kinase-STAT3 binding. *Cancer Res* **69**, 1958 (Mar 1, 2009).
90. J. Wang, M. Lopez-Fraga, A. Rynko, D. D. Lo, TNFR and LTbetaR agonists induce follicle-associated epithelium and M cell specific genes in rat and human intestinal epithelial cells. *Cytokine* **47**, 69 (Jul, 2009).
91. P. T. Sarkis, S. Ying, R. Xu, X. F. Yu, STAT1-independent cell type-specific regulation of antiviral APOBEC3G by IFN-alpha. *J Immunol* **177**, 4530 (Oct 1, 2006).
92. K. Kitaya, T. Yasuo, T. Yamaguchi, S. Fushiki, H. Honjo, Genes regulated by interferon-gamma in human uterine microvascular endothelial cells. *Int J Mol Med* **20**, 689 (Nov, 2007).
93. J. F. Martini *et al*., Quantitative analysis by polymerase chain reaction of growth hormone receptor gene expression in human liver and muscle. *Endocrinology* **136**, 1355 (Apr, 1995).
94. R. A. Ihrie *et al*., Persistent sonic hedgehog signaling in adult brain determines neural stem cell positional identity. *Neuron* **71**, 250 (Jul 28, 2011).
95. N. C. Inestrosa, E. Arenas, Emerging roles of Wnts in the adult nervous system. *Nat Rev Neurosci* **11**, 77 (Feb, 2010).
96. J. C. Bruning *et al*., Role of brain insulin receptor in control of body weight and reproduction. *Science* **289**, 2122 (Sep 22, 2000).
97. J. C. Bruning *et al*., A muscle-specific insulin receptor knockout exhibits features of the metabolic syndrome of NIDDM without altering glucose tolerance. *Mol Cell* **2**, 559 (Nov, 1998).
98. R. A. Lindberg, T. Hunter, cDNA cloning and characterization of eck, an epithelial cell receptor protein-tyrosine kinase in the eph/elk family of protein kinases. *Mol Cell Biol* **10**, 6316 (Dec, 1990).
99. B. D. Bennett *et al*., Cloning and characterization of HTK, a novel transmembrane tyrosine kinase of the EPH subfamily. *J Biol Chem* **269**, 14211 (May 13, 1994).
100. S. He *et al*., Expression of the co-signaling molecules CD40-CD40L and their growth inhibitory effect on pancreatic cancer in vitro. *Oncology reports* **28**, 262 (Jul, 2012).
101. S. V. Mudali *et al*., Patterns of EphA2 protein expression in primary and metastatic pancreatic carcinoma and correlation with genetic status. *Clinical & experimental metastasis* **23**, 357 (2006).
102. M. S. Duxbury, H. Ito, M. J. Zinner, S. W. Ashley, E. E. Whang, EphA2: a determinant of malignant cellular behavior and a potential therapeutic target in pancreatic adenocarcinoma. *Oncogene* **23**, 1448 (Feb 19, 2004).

103. A. Villasenor, D. C. Chong, M. Henkemeyer, O. Cleaver, Epithelial dynamics of pancreatic branching morphogenesis. *Development (Cambridge, England)* **137**, 4295 (Dec, 2010).
104. J. M. van Eyll *et al.*, Eph receptors and their ephrin ligands are expressed in developing mouse pancreas. *Gene Expr Patterns* **6**, 353 (Apr, 2006).
105. T. Troiani *et al.*, Targeting EGFR in Pancreatic Cancer Treatment. *Current drug targets*, (Mar 29, 2012).
106. M. R. Kritzik *et al.*, Expression of ErbB receptors during pancreatic islet development and regrowth. *The Journal of endocrinology* **165**, 67 (Apr, 2000).
107. M. A. Pulkkinen, B. Spencer-Dene, C. Dickson, T. Otonkoski, The IIIb isoform of fibroblast growth factor receptor 2 is required for proper growth and branching of pancreatic ductal epithelium but not for differentiation of exocrine or endocrine cells. *Mech Dev* **120**, 167 (Feb, 2003).
108. S. Nomura *et al.*, FGF10/FGFR2 signal induces cell migration and invasion in pancreatic cancer. *British journal of cancer* **99**, 305 (Jul 22, 2008).
109. S. Arnaud-Dabernat *et al.*, FGFR3 is a negative regulator of the expansion of pancreatic epithelial cells. *Diabetes* **56**, 96 (Jan, 2007).
110. F. Ruckert *et al.*, Examination of apoptosis signaling in pancreatic cancer by computational signal transduction analysis. *PloS one* **5**, e12243 (2010).
111. M. F. Di Renzo, R. Poulsom, M. Olivero, P. M. Comoglio, N. R. Lemoine, Expression of the Met/hepatocyte growth factor receptor in human pancreatic cancer. *Cancer research* **55**, 1129 (Mar 1, 1995).
112. M. Ebert, M. Yokoyama, H. Friess, M. W. Buchler, M. Korc, Coexpression of the c-met proto-oncogene and hepatocyte growth factor in human pancreatic cancer. *Cancer research* **54**, 5775 (Nov 15, 1994).
113. R. M. Thomas *et al.*, The RON receptor tyrosine kinase mediates oncogenic phenotypes in pancreatic cancer cells and is increasingly expressed during pancreatic cancer progression. *Cancer Res* **67**, 6075 (Jul 1, 2007).
114. T. Jonson *et al.*, Altered expression of TGFB receptors and mitogenic effects of TGFB in pancreatic carcinomas. *International journal of oncology* **19**, 71 (Jul, 2001).
115. K. W. Wagner *et al.*, Death-receptor O-glycosylation controls tumor-cell sensitivity to the proapoptotic ligand Apo2L/TRAIL. *Nat Med* **13**, 1070 (Sep, 2007).
116. A. D. Sanlioglu *et al.*, High TRAIL death receptor 4 and decoy receptor 2 expression correlates with significant cell death in pancreatic ductal adenocarcinoma patients. *Pancreas* **38**, 154 (Mar, 2009).
117. J. Lemke *et al.*, TRAIL signaling is mediated by DR4 in pancreatic tumor cells despite the expression of functional DR5. *Journal of molecular medicine (Berlin, Germany)* **88**, 729 (Jul, 2010).
118. S. M. Ibrahim *et al.*, Pancreatic adenocarcinoma cell lines show variable susceptibility to TRAIL-mediated cell death. *Pancreas* **23**, 72 (Jul, 2001).
119. R. Chiari *et al.*, Identification of a tumor-specific shared antigen derived from an Eph receptor and presented to CD4 T cells on HLA class II molecules. *Cancer research* **60**, 4855 (Sep 1, 2000).

120. B. Mosch, D. Pietzsch, J. Pietzsch, Irradiation affects cellular properties and Eph receptor expression in human melanoma cells. *Cell Adh Migr* **6**, 113 (Mar 1, 2012).
121. M. Ginarte, T. Garcia-Caballero, V. Fernandez-Redondo, A. Beiras, J. Toribio, Expression of growth hormone receptor in benign and malignant cutaneous proliferative entities. *Journal of cutaneous pathology* **27**, 276 (Jul, 2000).
122. Y. Qin *et al.*, Constitutive aberrant endogenous interleukin-1 facilitates inflammation and growth in human melanoma. *Mol Cancer Res* **9**, 1537 (Nov, 2011).
123. R. A. Spanjaard, K. M. Whren, C. Graves, J. Bhawan, Tumor necrosis factor receptor superfamily member TROY is a novel melanoma biomarker and potential therapeutic target. *International journal of cancer. Journal international du cancer* **120**, 1304 (Mar 15, 2007).
124. K. T. Barker *et al.*, Expression patterns of the novel receptor-like tyrosine kinase, DDR, in human breast tumours. *Oncogene* **10**, 569 (Feb 2, 1995).
125. Q. Wu *et al.*, Expression of Ephb2 and Ephb4 in breast carcinoma. *Pathology oncology research : POR* **10**, 26 (2004).
126. T. Holbro *et al.*, The ErbB2/ErbB3 heterodimer functions as an oncogenic unit: ErbB2 requires ErbB3 to drive breast tumor cell proliferation. *Proc Natl Acad Sci U S A* **100**, 8933 (Jul 22, 2003).
127. N. R. Lemoine *et al.*, Expression of the ERBB3 gene product in breast cancer. *British journal of cancer* **66**, 1116 (Dec, 1992).
128. S. Jaakkola *et al.*, Amplification of fgfr4 gene in human breast and gynecological cancers. *International journal of cancer. Journal international du cancer* **54**, 378 (May 28, 1993).
129. D. Meijer *et al.*, Fibroblast growth factor receptor 4 predicts failure on tamoxifen therapy in patients with recurrent breast cancer. *Endocrine-related cancer* **15**, 101 (Mar, 2008).
130. Y. Nihon-Yanagi, K. Terai, T. Murano, T. Matsumoto, S. Okazumi, Tissue expression of Toll-like receptors 2 and 4 in sporadic human colorectal cancer. *Cancer immunology, immunotherapy : CII* **61**, 71 (Jan, 2012).
131. M. L. Garcia-Hernandez, R. Hernandez-Pando, P. Gariglio, J. Berumen, Interleukin-10 promotes B16-melanoma growth by inhibition of macrophage functions and induction of tumour and vascular cell proliferation. *Immunology* **105**, 231 (Feb, 2002).
132. H. F. Chen, J. S. Chen, C. T. Shun, Y. F. Tsai, H. N. Ho, Decoy receptor 3 expression during the menstrual cycle and pregnancy, and regulation by sex steroids in endometrial cells in vitro. *Hum Reprod* **24**, 1350 (Jun, 2009).
133. Z. Ge, A. J. Sanders, L. Ye, Y. Wang, W. G. Jiang, Expression of death decoy receptor-3 (DcR3) in human breast cancer and its functional effects on breast cancer cells in vitro. *J Exp Ther Oncol* **9**, 109 (2011).
134. K. Miller-Jensen, K. A. Janes, J. S. Brugge, D. A. Lauffenburger, Common effector processing mediates cell-specific responses to stimuli. *Nature* **448**, 604 (Aug 2, 2007).

135. C. J. Decker, R. Parker, A turnover pathway for both stable and unstable mRNAs in yeast: evidence for a requirement for deadenylation. *Genes Dev* **7**, 1632 (Aug, 1993).
136. P. A. Frischmeyer, H. C. Dietz, Nonsense-mediated mRNA decay in health and disease. *Hum Mol Genet* **8**, 1893 (1999).
137. J. El-Bchiri *et al.*, Nonsense-mediated mRNA decay impacts MSI-driven carcinogenesis and anti-tumor immunity in colorectal cancers. *PLoS One* **3**, e2583 (2008).
138. F. Pastor, D. Kolonias, P. H. Giangrande, E. Gilboa, Induction of tumour immunity by targeted inhibition of nonsense-mediated mRNA decay. *Nature* **465**, 227 (May 13, 2010).
139. F. Ponten *et al.*, A global view of protein expression in human cells, tissues, and organs. *Mol Syst Biol* **5**, 337 (2009).
140. K. A. Janes, Paring down signaling complexity. *Nat Biotechnol* **28**, 681 (Jul, 2010).
141. J. P. MacKeigan, L. O. Murphy, J. Blenis, Sensitized RNAi screen of human kinases and phosphatases identifies new regulators of apoptosis and chemoresistance. *Nature cell biology* **7**, 591 (Jun, 2005).
142. A. Friedman, N. Perrimon, A functional RNAi screen for regulators of receptor tyrosine kinase and ERK signalling. *Nature* **444**, 230 (Nov 9, 2006).
143. D. G. Roller *et al.*, Synthetic lethal screening with small-molecule inhibitors provides a pathway to rational combination therapies for melanoma. *Molecular cancer therapeutics* **11**, 2505 (Nov, 2012).
144. Z. A. Knight, H. Lin, K. M. Shokat, Targeting the cancer kinome through polypharmacology. *Nature reviews. Cancer* **10**, 130 (Feb, 2010).
145. K. A. Janes, H. C. Reinhardt, M. B. Yaffe, Cytokine-induced signaling networks prioritize dynamic range over signal strength. *Cell* **135**, 343 (Oct 17, 2008).
146. W. A. Weiss, S. S. Taylor, K. M. Shokat, Recognizing and exploiting differences between RNAi and small-molecule inhibitors. *Nature chemical biology* **3**, 739 (Dec, 2007).
147. O. Brandman, T. Meyer, Feedback loops shape cellular signals in space and time. *Science* **322**, 390 (Oct 17, 2008).
148. S. Chatterjee *et al.*, Tumor VEGF:VEGFR2 autocrine feed-forward loop triggers angiogenesis in lung cancer. *The Journal of clinical investigation* **123**, 1732 (Apr 1, 2013).
149. R. Heinrich, B. G. Neel, T. A. Rapoport, Mathematical models of protein kinase signal transduction. *Mol Cell* **9**, 957 (May, 2002).
150. B. Schoeberl, C. Eichler-Jonsson, E. D. Gilles, G. Muller, Computational modeling of the dynamics of the MAP kinase cascade activated by surface and internalized EGF receptors. *Nat Biotechnol* **20**, 370 (Apr, 2002).
151. W. W. Chen *et al.*, Input-output behavior of ErbB signaling pathways as revealed by a mass action model trained against dynamic data. *Mol Syst Biol* **5**, 239 (2009).
152. A. Cornish-Bowden, *Fundamentals of enzyme kinetics*. (Butterworths, London ; Boston, 1979), pp. xiii, 230 p.
153. M. K. Dougherty *et al.*, Regulation of Raf-1 by direct feedback phosphorylation. *Molecular cell* **17**, 215 (Jan 21, 2005).

154. R. Fritsche-Guenther *et al*., Strong negative feedback from Erk to Raf confers robustness to MAPK signalling. *Molecular systems biology* **7**, 489 (May 24, 2011).
155. A. Schulze, B. Nicke, P. H. Warne, S. Tomlinson, J. Downward, The transcriptional response to Raf activation is almost completely dependent on Mitogen-activated Protein Kinase Kinase activity and shows a major autocrine component. *Molecular biology of the cell* **15**, 3450 (Jul, 2004).
156. M. F. Favata *et al*., Identification of a novel inhibitor of mitogen-activated protein kinase kinase. *The Journal of biological chemistry* **273**, 18623 (Jul 17, 1998).
157. S. P. Davies, H. Reddy, M. Caivano, P. Cohen, Specificity and mechanism of action of some commonly used protein kinase inhibitors. *The Biochemical journal* **351**, 95 (Oct 1, 2000).
158. M. Ohori *et al*., Identification of a selective ERK inhibitor and structural determination of the inhibitor-ERK2 complex. *Biochemical and biophysical research communications* **336**, 357 (Oct 14, 2005).
159. A. G. Gilmartin *et al*., GSK1120212 (JTP-74057) is an inhibitor of MEK activity and activation with favorable pharmacokinetic properties for sustained in vivo pathway inhibition. *Clinical cancer research : an official journal of the American Association for Cancer Research* **17**, 989 (Mar 1, 2011).
160. D. R. Alessi *et al*., Identification of the sites in MAP kinase kinase-1 phosphorylated by p74raf-1. *The EMBO journal* **13**, 1610 (Apr 1, 1994).
161. H. Hanafusa, S. Torii, T. Yasunaga, E. Nishida, Sprouty1 and Sprouty2 provide a control mechanism for the Ras/MAPK signalling pathway. *Nature cell biology* **4**, 850 (Nov, 2002).
162. O. Bermudez, G. Pages, C. Gimond, The dual-specificity MAP kinase phosphatases: critical roles in development and cancer. *American journal of physiology. Cell physiology* **299**, C189 (Aug, 2010).
163. N. Ishii *et al*., Enhanced Inhibition of ERK Signaling by a Novel Allosteric MEK Inhibitor, CH5126766, That Suppresses Feedback Reactivation of RAF Activity. *Cancer research* **73**, 4050 (Jul 1, 2013).
164. H. Huynh, K. C. Soo, P. K. Chow, E. Tran, Targeted inhibition of the extracellular signal-regulated kinase kinase pathway with AZD6244 (ARRY-142886) in the treatment of hepatocellular carcinoma. *Molecular cancer therapeutics* **6**, 138 (Jan, 2007).
165. G. Bollag *et al*., Clinical efficacy of a RAF inhibitor needs broad target blockade in BRAF-mutant melanoma. *Nature* **467**, 596 (Sep 30, 2010).
166. L. O. Murphy, S. Smith, R. H. Chen, D. C. Fingar, J. Blenis, Molecular interpretation of ERK signal duration by immediate early gene products. *Nature cell biology* **4**, 556 (Aug, 2002).
167. S. Inoue, G. Browne, G. Melino, G. M. Cohen, Ordering of caspases in cells undergoing apoptosis by the intrinsic pathway. *Cell death and differentiation* **16**, 1053 (Jul, 2009).
168. M. Bentele *et al*., Mathematical modeling reveals threshold mechanism in CD95-induced apoptosis. *J Cell Biol* **166**, 839 (Sep 13, 2004).
169. C. Scaffidi *et al*., Two CD95 (APO-1/Fas) signaling pathways. *Embo J* **17**, 1675 (Mar 16, 1998).

170. A. B. Berger, K. B. Sexton, M. Bogyo, Commonly used caspase inhibitors designed based on substrate specificity profiles lack selectivity. *Cell research* **16**, 961 (Dec, 2006).
171. C. Pop, G. S. Salvesen, Human caspases: activation, specificity, and regulation. *The Journal of biological chemistry* **284**, 21777 (Aug 14, 2009).
172. K. Stanger *et al.*, Allosteric peptides bind a caspase zymogen and mediate caspase tetramerization. *Nature chemical biology* **8**, 655 (Jul, 2012).
173. B. L. Lampson *et al.*, Rare codons regulate KRas oncogenesis. *Current biology : CB* **23**, 70 (Jan 7, 2013).
174. Y. Nakamura, T. Gojobori, T. Ikemura, Codon usage tabulated from international DNA sequence databases: status for the year 2000. *Nucleic acids research* **28**, 292 (Jan 1, 2000).
175. M. Esfandiarei, B. M. McManus, Molecular biology and pathogenesis of viral myocarditis. *Annu Rev Pathol* **3**, 127 (2008).
176. C. Kawai, From myocarditis to cardiomyopathy: mechanisms of inflammation and cell death: learning from the past for the future. *Circulation* **99**, 1091 (Mar 2, 1999).
177. F. S. Garmaroudi *et al.*, Pairwise network mechanisms in the host signaling response to coxsackievirus B3 infection. *Proc Natl Acad Sci U S A* **107**, 17053 (Sep 28, 2010).
178. D. Ribet, P. Cossart, Pathogen-mediated posttranslational modifications: A re-emerging field. *Cell* **143**, 694 (Nov 24, 2010).
179. K. A. Janes, M. B. Yaffe, Data-driven modelling of signal-transduction networks. *Nat Rev Mol Cell Biol* **7**, 820 (Nov, 2006).
180. B. B. Aldridge, J. M. Burke, D. A. Lauffenburger, P. K. Sorger, Physicochemical modelling of cell signalling pathways. *Nat Cell Biol* **8**, 1195 (Nov, 2006).
181. S. J. Riedl, Y. Shi, Molecular mechanisms of caspase regulation during apoptosis. *Nat Rev Mol Cell Biol* **5**, 897 (Nov, 2004).
182. D. A. Cross, D. R. Alessi, P. Cohen, M. Andjelkovich, B. A. Hemmings, Inhibition of glycogen synthase kinase-3 by insulin mediated by protein kinase B. *Nature* **378**, 785 (Dec 21-28, 1995).
183. J. Rouse *et al.*, A novel kinase cascade triggered by stress and heat shock that stimulates MAPKAP kinase-2 and phosphorylation of the small heat shock proteins. *Cell* **78**, 1027 (Sep 23, 1994).
184. K. J. Jensen, K. A. Janes, Modeling the latent dimensions of multivariate signaling datasets. *Physical biology* **9**, 045004 (Aug, 2012).
185. A. M. Reimold, J. Kim, R. Finberg, L. H. Glimcher, Decreased immediate inflammatory gene induction in activating transcription factor-2 mutant mice. *Int Immunol* **13**, 241 (Feb, 2001).
186. Z. Xia, M. Dickens, J. Raingeaud, R. J. Davis, M. E. Greenberg, Opposing effects of ERK and JNK-p38 MAP kinases on apoptosis. *Science* **270**, 1326 (Nov 24, 1995).
187. S. Nishimoto, E. Nishida, MAPK signalling: ERK5 versus ERK1/2. *EMBO Rep* **7**, 782 (Aug, 2006).

188. C. P. Regan *et al*., Erk5 null mice display multiple extraembryonic vascular and embryonic cardiovascular defects. *Proc Natl Acad Sci U S A* **99**, 9248 (Jul 9, 2002).
189. T. W. Sturgill, L. B. Ray, E. Erikson, J. L. Maller, Insulin-stimulated MAP-2 kinase phosphorylates and activates ribosomal protein S6 kinase II. *Nature* **334**, 715 (Aug 25, 1988).
190. S. P. Davies, H. Reddy, M. Caivano, P. Cohen, Specificity and mechanism of action of some commonly used protein kinase inhibitors. *Biochem J* **351**, 95 (Oct 1, 2000).
191. J. C. Lee *et al*., A protein kinase involved in the regulation of inflammatory cytokine biosynthesis. *Nature* **372**, 739 (Dec 22-29, 1994).
192. J. Westermarck, S. P. Li, T. Kallunki, J. Han, V. M. Kahari, p38 mitogen-activated protein kinase-dependent activation of protein phosphatases 1 and 2A inhibits MEK1 and MEK2 activity and collagenase 1 (MMP-1) gene expression. *Molecular and cellular biology* **21**, 2373 (Apr, 2001).
193. X. Si *et al*., Stress-activated protein kinases are involved in coxsackievirus B3 viral progeny release. *Journal of virology* **79**, 13875 (Nov, 2005).
194. C. Pargellis *et al*., Inhibition of p38 MAP kinase by utilizing a novel allosteric binding site. *Nat Struct Biol* **9**, 268 (Apr, 2002).
195. H. Luo *et al*., Coxsackievirus B3 replication is reduced by inhibition of the extracellular signal-regulated kinase (ERK) signaling pathway. *Journal of virology* **76**, 3365 (Apr, 2002).
196. C. Yan *et al*., Activation of extracellular signal-regulated kinase 5 reduces cardiac apoptosis and dysfunction via inhibition of a phosphodiesterase 3A/inducible cAMP early repressor feedback loop. *Circ Res* **100**, 510 (Mar 2, 2007).
197. T. E. Kimura *et al*., Targeted deletion of the extracellular signal-regulated protein kinase 5 attenuates hypertrophic response and promotes pressure overload-induced apoptosis in the heart. *Circ Res* **106**, 961 (Mar 19, 2010).
198. Q. Yang *et al*., Pharmacological inhibition of BMK1 suppresses tumor growth through promyelocytic leukemia protein. *Cancer Cell* **18**, 258 (Sep 14, 2010).
199. C. M. Carthy *et al*., Caspase activation and specific cleavage of substrates after coxsackievirus B3-induced cytopathic effect in HeLa cells. *Journal of virology* **72**, 7669 (Sep, 1998).
200. J. P. Yuan *et al*., Coxsackievirus B3-induced apoptosis and caspase-3. *Cell Res* **13**, 203 (Jun, 2003).
201. S. P. Perfetto *et al*., Amine reactive dyes: an effective tool to discriminate live and dead cells in polychromatic flow cytometry. *J Immunol Methods* **313**, 199 (Jun 30, 2006).
202. P. Scaffidi, T. Misteli, M. E. Bianchi, Release of chromatin protein HMGB1 by necrotic cells triggers inflammation. *Nature* **418**, 191 (Jul 11, 2002).
203. A. Degterev *et al*., Chemical inhibitor of nonapoptotic cell death with therapeutic potential for ischemic brain injury. *Nat Chem Biol* **1**, 112 (Jul, 2005).
204. M. Huber *et al*., Cleavage of RasGAP and phosphorylation of mitogen-activated protein kinase in the course of coxsackievirus B3 replication. *Journal of virology* **73**, 3587 (May, 1999).

205. J. G. Albeck *et al*., Collecting and organizing systematic sets of protein data. *Nat Rev Mol Cell Biol* **7**, 803 (Nov, 2006).
206. T. Hunter, Signaling--2000 and beyond. *Cell* **100**, 113 (Jan 7, 2000).
207. T. Pawson, R. Linding, Network medicine. *FEBS letters* **582**, 1266 (Apr 9, 2008).
208. T. Hunter, Protein kinases and phosphatases: the yin and yang of protein phosphorylation and signaling. *Cell* **80**, 225 (Jan 27, 1995).
209. K. A. Janes *et al*., A high-throughput quantitative multiplex kinase assay for monitoring information flow in signaling networks: application to sepsis-apoptosis. *Molecular & cellular proteomics : MCP* **2**, 463 (Jul, 2003).
210. S. Gaudet *et al*., A compendium of signals and responses triggered by prodeath and prosurvival cytokines. *Molecular & cellular proteomics : MCP* **4**, 1569 (Oct, 2005).
211. Z. Songyang *et al*., Use of an oriented peptide library to determine the optimal substrates of protein kinases. *Current biology : CB* **4**, 973 (Nov 1, 1994).
212. P. M. Holland, J. A. Cooper, Protein modification: docking sites for kinases. *Current biology : CB* **9**, R329 (May 6, 1999).
213. F. A. Gonzalez, D. L. Raden, R. J. Davis, Identification of substrate recognition determinants for human ERK1 and ERK2 protein kinases. *The Journal of biological chemistry* **266**, 22159 (Nov 25, 1991).
214. D. Jacobs, D. Glossip, H. Xing, A. J. Muslin, K. Kornfeld, Multiple docking sites on substrate proteins form a modular system that mediates recognition by ERK MAP kinase. *Genes & development* **13**, 163 (Jan 15, 1999).
215. T. Obata *et al*., Peptide and protein library screening defines optimal substrate motifs for AKT/PKB. *The Journal of biological chemistry* **275**, 36108 (Nov 17, 2000).
216. B. Derijard *et al*., JNK1: a protein kinase stimulated by UV light and Ha-Ras that binds and phosphorylates the c-Jun activation domain. *Cell* **76**, 1025 (Mar 25, 1994).
217. J. E. Hutti *et al*., IkappaB kinase beta phosphorylates the K63 deubiquitinase A20 to cause feedback inhibition of the NF-kappaB pathway. *Molecular and cellular biology* **27**, 7451 (Nov, 2007).
218. I. A. Manke *et al*., MAPKAP kinase-2 is a cell cycle checkpoint kinase that regulates the G2/M transition and S phase progression in response to UV irradiation. *Molecular cell* **17**, 37 (Jan 7, 2005).
219. M. D. Shults, K. A. Janes, D. A. Lauffenburger, B. Imperiali, A multiplexed homogeneous fluorescence-based assay for protein kinase activity in cell lysates. *Nature methods* **2**, 277 (Apr, 2005).
220. Z. A. Knight, K. M. Shokat, Features of selective kinase inhibitors. *Chemistry & biology* **12**, 621 (Jun, 2005).
221. J. W. Jarvik, C. A. Telmer, Epitope tagging. *Annual review of genetics* **32**, 601 (1998).
222. C. E. Fritze, T. R. Anderson, Epitope tagging: general method for tracking recombinant proteins. *Methods in enzymology* **327**, 3 (2000).
223. P. P. Roux, S. A. Richards, J. Blenis, Phosphorylation of p90 ribosomal S6 kinase (RSK) regulates extracellular signal-regulated kinase docking and RSK activity. *Molecular and cellular biology* **23**, 4796 (Jul, 2003).

224. R. Geleziunas *et al*., Human T-cell leukemia virus type 1 Tax induction of NF-kappaB involves activation of the IkappaB kinase alpha (IKKalpha) and IKKbeta cellular kinases. *Molecular and cellular biology* **18**, 5157 (Sep, 1998).
225. C. McCormick, D. Ganem, The kaposin B protein of KSHV activates the p38/MK2 pathway and stabilizes cytokine mRNAs. *Science* **307**, 739 (Feb 4, 2005).
226. W. B. Jakoby, D. M. Ziegler, The enzymes of detoxication. *The Journal of biological chemistry* **265**, 20715 (Dec 5, 1990).
227. G. T. Hermanson, *Bioconjugate techniques*. (Academic Press, San Diego, 1996), pp. xxv, 785 p.
228. D. Stokoe *et al*., MAPKAP kinase-2; a novel protein kinase activated by mitogen-activated protein kinase. *The EMBO journal* **11**, 3985 (Nov, 1992).
229. K. A. Janes *et al*., A high-throughput quantitative multiplex kinase assay for monitoring information flow in signaling networks: application to sepsis-apoptosis. *Mol Cell Proteomics* **2**, 463 (Jul, 2003).
230. K. Kubota *et al*., Sensitive multiplexed analysis of kinase activities and activity-based kinase identification. *Nat Biotechnol* **27**, 933 (Oct, 2009).
231. E. C. Butcher, E. L. Berg, E. J. Kunkel, Systems biology in drug discovery. *Nature biotechnology* **22**, 1253 (Oct, 2004).
232. P. Rajasethupathy, S. J. Vayttaden, U. S. Bhalla, Systems modeling: a pathway to drug discovery. *Current opinion in chemical biology* **9**, 400 (Aug, 2005).
233. Q. Xue, K. Miller-Jensen, Systems biology of virus-host signaling network interactions. *BMB reports* **45**, 213 (Apr, 2012).
234. J. J. Hornberg, F. J. Bruggeman, H. V. Westerhoff, J. Lankelma, Cancer: a Systems Biology disease. *Bio Systems* **83**, 81 (Feb-Mar, 2006).
235. A. M. Scott, J. P. Allison, J. D. Wolchok, Monoclonal antibodies in cancer therapy. *Cancer immunity* **12**, 14 (2012).
236. E. J. Morris *et al*., Discovery of a Novel ERK Inhibitor with Activity in Models of Acquired Resistance to BRAF and MEK Inhibitors. *Cancer discovery* **3**, 742 (Jul, 2013).
237. H. Shi, X. Kong, A. Ribas, R. S. Lo, Combinatorial treatments that overcome PDGFRbeta-driven resistance of melanoma cells to V600EB-RAF inhibition. *Cancer research* **71**, 5067 (Aug 1, 2011).
238. R. S. Lo, Receptor tyrosine kinases in cancer escape from BRAF inhibitors. *Cell research* **22**, 945 (Jun, 2012).
239. M. Lacroix, Persistent use of "false" cell lines. *International journal of cancer. Journal international du cancer* **122**, 1 (Jan 1, 2008).
240. J. M. Rae, C. J. Creighton, J. M. Meck, B. R. Haddad, M. D. Johnson, MDA-MB-435 cells are derived from M14 melanoma cells--a loss for breast cancer, but a boon for melanoma research. *Breast cancer research and treatment* **104**, 13 (Jul, 2007).
241. M. Christgen, U. Lehmann, MDA-MB-435: the questionable use of a melanoma cell line as a model for human breast cancer is ongoing. *Cancer biology & therapy* **6**, 1355 (Sep, 2007).
242. M. Lacroix, MDA-MB-435 cells are from melanoma, not from breast cancer. *Cancer chemotherapy and pharmacology* **63**, 567 (Feb, 2009).

243. J. Jia *et al*., Mechanisms of drug combinations: interaction and network perspectives. *Nature reviews. Drug discovery* **8**, 111 (Feb, 2009).
244. K. Klingel *et al*., Ongoing enterovirus-induced myocarditis is associated with persistent heart muscle infection: quantitative analysis of virus replication, tissue damage, and inflammation. *Proceedings of the National Academy of Sciences of the United States of America* **89**, 314 (Jan 1, 1992).
245. C. M. Emery *et al*., MEK1 mutations confer resistance to MEK and B-RAF inhibition. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 20411 (Dec 1, 2009).
246. E. Devarajan *et al*., Down-regulation of caspase 3 in breast cancer: a possible mechanism for chemoresistance. *Oncogene* **21**, 8843 (Dec 12, 2002).
247. S. Ghavami *et al*., Apoptosis and cancer: mutations within caspase genes. *Journal of medical genetics* **46**, 497 (Aug, 2009).
248. J. S. Logue, D. K. Morrison, Complexity in the signaling network: insights from the use of targeted inhibitors in cancer therapy. *Genes & development* **26**, 641 (Apr 1, 2012).
249. C. Dominguez, D. A. Powers, N. Tamayo, p38 MAP kinase inhibitors: many are made, but few are chosen. *Current opinion in drug discovery & development* **8**, 421 (Jul, 2005).
250. M. R. Lee, C. Dominguez, MAP kinase p38 inhibitors: clinical results and an intimate look at their interactions with p38alpha protein. *Current medicinal chemistry* **12**, 2979 (2005).
251. A. Akinleye, M. Furqan, N. Mukhi, P. Ravella, D. Liu, MEK and the inhibitors: from bench to bedside. *Journal of hematology & oncology* **6**, 27 (2013).
252. S. Huber, A. I. Ramsingh, Coxsackievirus-induced pancreatitis. *Viral immunology* **17**, 358 (2004).
253. R. Osterback *et al*., Coxsackievirus A6 and hand, foot, and mouth disease, Finland. *Emerging infectious diseases* **15**, 1485 (Sep, 2009).
254. M. S. Horwitz *et al*., Diabetes induced by Coxsackie virus: initiation by bystander damage and not molecular mimicry. *Nature medicine* **4**, 781 (Jul, 1998).
255. H. Jaidane, D. Hober, Role of coxsackievirus B4 in the pathogenesis of type 1 diabetes. *Diabetes & metabolism* **34**, 537 (Dec, 2008).
256. B. D. Cosgrove *et al*., Cytokine-associated drug toxicity in human hepatocytes is associated with signaling network dysregulation. *Mol Biosyst* **6**, 1195 (Jul, 2010).
257. J. Saez-Rodriguez *et al*., Discrete logic modelling as a means to link protein signalling networks with functional analysis of mammalian signal transduction. *Mol Syst Biol* **5**, 331 (2009).
258. A. Wolf-Yadlin, S. Hautaniemi, D. A. Lauffenburger, F. M. White, Multiple reaction monitoring for robust quantitative proteomic analysis of cellular signaling networks. *Proc Natl Acad Sci U S A* **104**, 5860 (Apr 3, 2007).
259. S. C. Bendall *et al*., Single-cell mass cytometry of differential immune and drug responses across a human hematopoietic continuum. *Science* **332**, 687 (May 6, 2011).
260. J. M. Irish *et al*., Single cell profiling of potentiated phospho-protein networks in cancer cells. *Cell* **118**, 217 (Jul 23, 2004).

261. D. R. Catchpoole, P. Kennedy, D. B. Skillicorn, S. Simoff, The curse of dimensionality: a blessing to personalized medicine. *J Clin Oncol* **28**, e723 (Dec 1, 2010).
262. P. J. Huber, Experiences With 3-Dimensional Scatterplots. *J Amer Statistical Assoc* **82**, 448 (Jun, 1987).
263. E. R. Tufte, *Envisioning Information*. (Graphics Press, Cheshire, CT, 1990), pp. 126.
264. P. Qiu *et al*., Extracting a cellular hierarchy from high-dimensional cytometry data with SPADE. *Nat Biotechnol* **29**, 886 (2011).
265. P. E. Barbano *et al*., A mathematical tool for exploring the dynamics of biological networks. *Proc Natl Acad Sci U S A* **104**, 19169 (Dec 4, 2007).
266. H. Martens, M. Martens, *Multivariate Analysis of Quality: An Introduction*. (John Wiley & Sons, Chichester, U.K., 2001), pp. 445.
267. P. Geladi, B. R. Kowalski, Partial Least-Squares Regression - a Tutorial. *Anal Chim Acta* **185**, 1 (JUL 31, 1986).
268. O. Alter, P. O. Brown, D. Botstein, Singular value decomposition for genome-wide expression data processing and modeling. *Proc Natl Acad Sci U S A* **97**, 10101 (Aug 29, 2000).
269. F. G. Kuruvilla, P. J. Park, S. L. Schreiber, Vector algebra in the analysis of genome-wide expression data. *Genome Biol* **3**, RESEARCH0011 (2002).
270. N. Kumar, A. Wolf-Yadlin, F. M. White, D. A. Lauffenburger, Modeling HER2 effects on cell behavior from mass spectrometry phosphotyrosine data. *PLoS Comput Biol* **3**, e4 (Jan 5, 2007).
271. D. Kumar, R. Srikanth, H. Ahlfors, R. Lahesmaa, K. V. Rao, Capturing cell-fate decisions from the molecular signatures of a receptor-dependent signaling response. *Mol Syst Biol* **3**, 150 (2007).
272. M. L. Kemp, L. Wille, C. L. Lewis, L. B. Nicholson, D. A. Lauffenburger, Quantitative network signal combinations downstream of TCR activation can predict IL-2 production response. *J Immunol* **178**, 4984 (Apr 15, 2007).
273. A. Gordus *et al*., Linear combinations of docking affinities explain quantitative differences in RTK signaling. *Mol Syst Biol* **5**, 235 (2009).
274. A. R. Asthagiri, C. A. Reinhart, A. F. Horwitz, D. A. Lauffenburger, The role of transient ERK2 signals in fibronectin- and insulin-mediated DNA synthesis. *J Cell Sci* **113**, 4499 (Dec, 2000).
275. B. Efron, R. J. Tibshirani, *An Introduction to the Bootstrap*. (Chapman and Hall, London, 1993), pp. 436.
276. S. Wold, M. Sjostrom, L. Eriksson, PLS-regression: a basic tool of chemometrics. *Chemometrics Intell. Lab. Syst.* **58**, 109 (Oct 28, 2001).
277. R. Bro, Multiway calibration. Multilinear PLS. *J Chemometr* **10**, 47 (Jan-Feb, 1996).
278. W. Liebermeister, Linear modes of gene expression determined by independent component analysis. *Bioinformatics* **18**, 51 (Jan, 2002).
279. D. D. Lee, H. S. Seung, Learning the parts of objects by non-negative matrix factorization. *Nature* **401**, 788 (Oct 21, 1999).

280. J. C. Liao *et al*., Network component analysis: reconstruction of regulatory signals in biological systems. *Proc Natl Acad Sci U S A* **100**, 15522 (Dec 23, 2003).
281. A. Abdi, M. B. Tahoori, E. S. Emamian, Fault diagnosis engineering of digital circuits can identify vulnerable molecules in complex cellular pathways. *Sci Signal* **1**, ra10 (2008).
282. A. V. Oppenheim, A. S. Willsky, S. H. Hawab, *Signals and Systems*. Prentice-Hall Signal Processing Series (Prentice Hall, Upper Saddle River, NJ, ed. 2nd, 1996), pp. 957.
283. J. Debnath, S. K. Muthuswamy, J. S. Brugge, Morphogenesis and oncogenesis of MCF-10A mammary epithelial acini grown in three-dimensional basement membrane cultures. *Methods* **30**, 256 (Jul, 2003).
284. K. A. Janes, C. C. Wang, K. J. Holmberg, K. Cabral, J. S. Brugge, Identifying single-cell molecular programs by stochastic profiling. *Nat Methods* **7**, 311 (Apr, 2010).
285. T. Furukawa *et al*., Long-term culture and immortalization of epithelial cells from normal adult human pancreatic ducts transfected by the E6E7 gene of human papilloma virus 16. *The American journal of pathology* **148**, 1763 (Jun, 1996).
286. S. Rozen, H. Skaletsky, Primer3 on the WWW for general users and for biologist programmers. *Methods Mol Biol* **132**, 365 (2000).
287. A. Dobin *et al*., STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15 (Jan 1, 2013).
288. A. R. Quinlan, I. M. Hall, BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841 (Mar 15, 2010).
289. J. T. Robinson *et al*., Integrative genomics viewer. *Nat Biotechnol* **29**, 24 (Jan, 2011).
290. T. Sing, O. Sander, N. Beerenwinkel, T. Lengauer, ROCR: visualizing classifier performance in R. *Bioinformatics* **21**, 3940 (Oct 15, 2005).
291. X. Yang *et al*., A public genome-scale lentiviral expression library of human ORFs. *Nat Methods* **8**, 659 (Aug, 2011).
292. L. Wang, J. S. Brugge, K. A. Janes, Intersection of FOXO- and RUNX1-mediated gene expression programs in single breast epithelial cells during morphogenesis and tumor progression. *Proc Natl Acad Sci U S A* **108**, E803 (Oct 4, 2011).
293. M. Bengtsson, A. Stahlberg, P. Rorsman, M. Kubista, Gene expression profiling in single cells from the pancreatic islets of Langerhans reveals lognormal distribution of mRNA levels. *Genome Res* **15**, 1388 (Oct, 2005).
294. J. S. Boehm *et al*., Integrative genomic approaches identify IKBKE as a breast cancer oncogene. *Cell* **129**, 1065 (Jun 15, 2007).
295. J. Raingeaud, A. J. Whitmarsh, T. Barrett, B. Derijard, R. J. Davis, MKK3- and MKK6-regulated gene expression is mediated by the p38 mitogen-activated protein kinase signal transduction pathway. *Molecular and cellular biology* **16**, 1247 (Mar, 1996).
296. K. J. Shin *et al*., A single lentiviral vector platform for microRNA-based conditional RNA interference and coordinated transgene expression. *Proc Natl Acad Sci U S A* **103**, 13759 (Sep 12, 2006).

297. W. C. Claycomb *et al*., HL-1 cells: a cardiac muscle cell line that contracts and retains phenotypic characteristics of the adult cardiomyocyte. *Proc Natl Acad Sci U S A* **95**, 2979 (Mar 17, 1998).