Inaugural dissertation for
obtaining the doctoral degree
of the
Combined Faculty of Mathematics, Engineering and Natural Sciences
of the
Ruprecht - Karls - University
Heidelberg

Presented by

M.Sc, Sophia Müller-Dott
born in Frankfurt am Main, Germany
Oral examination: 21-03-2025

From Activity Inference to Multi-Omics Network Contextualization: Deciphering Cellular Signaling and Disease Mechanisms

Referees:

Prof. Dr. Robert Russell

Prof. Dr. Julio Saez-Rodriguez

“Real knowledge is to know the extent of one's ignorance”

Confucius

Abstract

The integration of diverse omics layers with advanced computational methods can help to decipher cellular signaling and disease mechanisms. Thereby it is crucial to ensure that computational predictions truly reflect biological mechanisms and that different omics layers are cohesively integrated. This thesis focuses on evaluating approaches to infer the activity of transcription factors and kinases as well as advancing methods to uncover context-dependent signaling networks. First, to identify the most reliable strategies for activity inference, benchmarking frameworks were established to assess various inference methods. This revealed that a novel collection of signed transcription factor-gene interactions outperforms existing resources in predicting transcription factor activities. Similarly, manually curated kinase-substrate libraries combined with less complex computational models were shown to provide higher accuracy for kinase activity inference. Next, to reveal the role of these regulators in signaling pathways across diverse biological contexts, methods for network contextualization were developed, incorporating phosphoproteomics data alone and in combination with transcriptomics data. For phosphoproteomics-based network modeling, signed protein-protein interactions were incorporated to account for regulatory directionality, improving the representation of biological networks. Additionally, a multi-omics network contextualization approach was established which is able to link upstream stimuli to kinase and transcription factor activities in a cohesive manner, bridging phosphoproteomics and transcriptomics data. The network models were then applied to study the effects of metformin on colorectal cancer and the mechanisms driving hepatic stellate cell activation, uncovering condition-specific regulatory mechanisms and potential interactions between key signaling pathways. This highlights that integrating experimental data with reliable prior knowledge and advanced computational approaches can aid in understanding context-dependent signaling processes in complex biological systems.

Zusammenfassung

Das Verknüpfen verschiedener Omics-Datenebenen mit bioinformatischen Methoden kann helfen, zelluläre Signalwege und pathologische Krankheitsprozesse zu entschlüsseln. Dabei muss sichergestellt werden, dass computergestützte Vorhersagen die biologischen Mechanismen korrekt widerspiegeln und Omics-Ebenen kohärent integriert werden. Der Fokus dieser Arbeit liegt auf der Bewertung von Ansätzen zur Berechnung der Aktivität von Transkriptionsfaktoren und Kinasen sowie auf der Weiterentwicklung von Methoden zur Aufdeckung kontextabhängiger Signalnetzwerke. Zunächst wurden Bewertungskriterien festgelegt, um die zuverlässigsten Strategien zur Aktivitätsabschätzung zu identifizieren und verschiedene Inferenzmethoden zu bewerten. Es zeigte sich, dass eine neue Sammlung von Transkriptionsfaktor-Gen-Interaktionen die bestehenden Ressourcen bei der Vorhersage von Transkriptionsfaktor-Aktivitäten übertrifft. Ebenso konnten manuell erstellte Kinase-Substrat-Bibliotheken in Kombination mit weniger komplexen Modellen eine höhere Genauigkeit bei der Abschätzung der Kinaseaktivität erreichen. Um die Rolle dieser Regulatoren in Signalwegen verschiedener biologischer Kontexte zu untersuchen, wurden Methoden zur Netzwerkkontextualisierung entwickelt, die Phosphoproteomikdaten allein oder in Kombination mit Transkriptomikdaten einbeziehen. Für die phosphoproteomikbasierte Netzwerkmodellierung wurden signierte Protein-Protein-Interaktionen genutzt, um die Art der Regulation besser darzustellen und biologische Netzwerke genauer abzubilden. Zusätzlich wurde ein Multi-Omics-Ansatz zur Netzwerkkontextualisierung entwickelt, der vorgeschaltete Stimuli mit der Aktivität von Kinasen und Transkriptionsfaktoren kohärent verknüpft. Dadurch wurde die Verbindung von Phosphoproteomik- und Transkriptomik-Daten ermöglicht. Diese Netzwerkmodelle wurden verwendet, um die Wirkung von Metformin auf kolorektale Karzinome sowie die Aktivierungsmechanismen hepatischer Sternzellen zu untersuchen. Dabei konnten spezifische regulatorische Mechanismen und potenzielle Interaktionen zwischen wichtigen Signalwegen mit kontextspezifischen Variationen identifiziert werden. Dies zeigt, dass die Kombination experimenteller Daten mit zuverlässigem Vorwissen und fortschrittlichen Rechenansätzen das Verständnis kontextabhängiger Signalprozesse in komplexen biologischen Systemen verbessern kann.

Acknowledgements

Looking back on the last four years, there are so many people I want to thank who've been part of this journey. I'm really grateful to everyone who's supported me along the way.

First and foremost, I want to thank my doctoral supervisor Julio Saez-Rodriguez who has been an amazing mentor, always supportive, kind, and approachable. I truly appreciate the freedom and trust he provided throughout my PhD. Thank you Julio for all your support and guidance, I couldn't have wished for a better supervisor. I would also like to thank all past and current members of the Saez Lab who made this time very special. This group is just incredibly kind and supportive, and I truly enjoyed being part of it. A special thanks goes to Daniel Dimitrov, Ricardo Ramirez, Pau Badia i Mompel, Lorna Wessels, Jan Lanzer, Barbara Zita Peters and Leonie Küchenhoff for their friendship beyond the lab. I would also like to highlight those who patiently helped me with all my questions during my PhD, specifically Attila Gabor, Robin Fallegger, Martin Garrido-Rodriguez, Aurelien Dugourd, Denez Turei, and Pablo Rodriguez-Mier, as well as Erika Schulz and Lydia Roeder, who were always there to assist with any administrative tasks. All of you have contributed to my PhD and I am really grateful for that.

I also wish to acknowledge the scientific contributions of many people outside of the lab, without whom the completion of my thesis would not have been possible. First of all, I would like to thank my thesis advisory committee, Rob Russell and Ursula Klingmüller, for their helpful feedback and guidance throughout the years. I would also like to thank Rebecca Wade for giving her time to be part of my defense committee. Additionally, I've been lucky to work with great collaborators throughout the years and would like to particularly thank Barbora Salovska, Eirini Tsirvouli, Astrid Laegreid, Eric J. Jaehnig and Elisa Holstein for their great scientific discussions and contributions. Lastly, I would like to thank the Olsen Lab in Copenhagen, where I had the privilege to spend three months during my PhD. A special thanks to Jesper Olsen for giving me the opportunity to join his lab, and to all the Olsen Lab members for making me feel so welcome and part of the group.

Finally, I am grateful to my family, who has always supported and believed in me. I am also truly thankful for my friends outside the lab, who have always been an important part of my life and have been there for me also during my PhD. Last but not least, I would like to thank my boyfriend Philipp: Vielen Dank für deine Unterstützung bei all meinen Entscheidungen und dein Verständnis, wenn alles mal etwas stressiger war. Ich bin dir unendlich dankbar, dass du in dieser Zeit immer für mich da warst, dir all meine Präsentationen angehört hast und jeden noch so kleinen Erfolg mit mir gefeiert hast. Ich bin wirklich glücklich, dich in meinem Leben zu haben.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Systems Biology

The human body comprises nearly 37 trillion cells, forming a complex biological system that is able to operate in a highly coordinated manner (Sender et al., 2016). In this system, cells perform distinct functions, such as providing immune defense or creating protective barriers, tailored to their specific roles within the organism. These cellular functions are governed by a complex network of biomolecules, including approximately 20,000 protein-coding genes (Consortium, 2004) that are able to produce over 100,000 distinct proteins (Aebersold and Mann, 2016). These proteins can be further modulated by more than 400 types of post-translational modifications (PTMs), such as phosphorylation, acetylation, and ubiquitination, that influence their localization, stability, and activity (Khoury et al., 2011; Ramazi and Zahiri, 2021). Adding another layer of complexity, cofactors regulated by metabolic pathways play essential roles in fine-tuning protein functions (Sun et al., 2023). All together, these biomolecules form a dynamic and interconnected network that drives cellular functions and enables cells to adapt to external and internal stimuli (Aebersold and Mann, 2016). Additionally, they allow cells to communicate with one another and synchronize their processes, ensuring harmonious coordination throughout the body (Figure 1.1).

To fully understand these complex networks and be able to capture processes, such as self-regulation, adaptability, and multicellular coordination, it is crucial to study the system as a whole rather than as a collection of isolated parts (Turnbull et al., 2018). As such, systems biology emerged as a discipline dedicated to integrating biological processes to uncover how their interactions give rise to the functions of life. Systems biology empha-

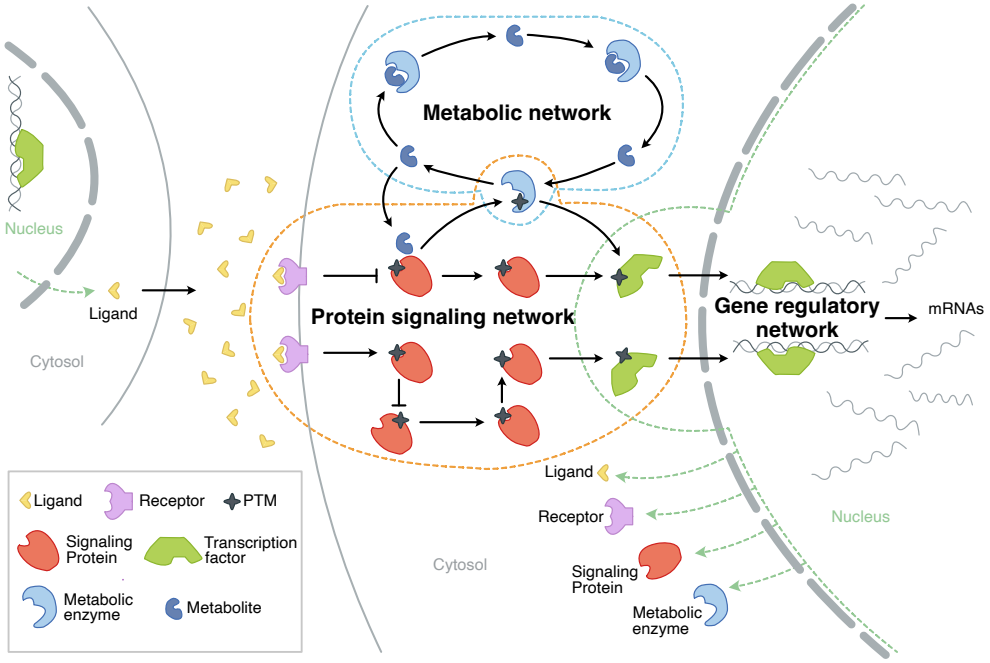

Figure 1.1. Illustration of intracellular networks driving cellular processes. Connection of signaling, metabolic, and gene regulatory networks to regulate cellular processes. Ligands are secreted from neighboring cells and bind to receptors on the cell surface leading trough intracellular signaling cascades to downstream metabolic and transcriptional regulation. *Adapted from Garrido-Rodriguez et al. (2022).*

sizes modeling dynamic, nonlinear interactions and integrating smaller, modular networks into cohesive frameworks, providing insights into processes like cell signaling, feedback loops, and metabolic regulation (Manson, 2001). For instance, mathematical models of signaling pathways can reveal how perturbations affect cellular decision-making, such as apoptosis or proliferation (Alon, 2007). Computational advances have further enabled the incorporation of context-dependent variables like environmental stimuli or cell-type specificity (Melham, 2013; Motta and Pappalardo, 2012). In systems biology, omics analyses, spanning genomics, transcriptomics, proteomics, or metabolomics for example, has emerged as a vital tool for dissecting these intricate systems, providing a comprehensive view of molecular interactions and their regulation in diverse biological contexts (Hasin et al., 2017).

1.2 Leveraging Omics Data to Profile Diseases

The emergence of high-throughput omics technologies enables a deep exploration of the molecular mechanisms underlying cellular function and diseases. These technologies generate comprehensive datasets, capturing full transcriptomes, epigenetic modifications, metabolic activities, signaling pathways, protein profiles, chromosomal arrangements,

and other molecular characteristics (Hasin et al., 2017; Joyce and Palsson, 2006; Karczewski and Snyder, 2018). By systematically profiling genes, proteins, post-translational modifications and other modalities, omics approaches offer a multidimensional approach to study cellular processes. However, while integrating these data types is essential for achieving a holistic understanding of cellular function and disease mechanisms, challenges in interpretation and cross-omics integration remain.

1.2.1 Transcriptomics

Transcriptomics examines the abundance of ribonucleic acid (RNA) molecules to uncover gene expression patterns within a given biological entity, such as a tissue, cell, or organism (Lowe et al., 2017). Early transcriptomics approaches focused on measuring gene expression at the population level, capturing the average expression across all cells in a sample, using technologies such as microarrays and RNA sequencing (RNA-seq). While microarrays rely on hybridization to predefined probes, RNA-seq provides an unbiased and high-resolution approach, capable of detecting both known and novel transcripts, as well as quantifying isoforms (Ozsolak and Milos, 2011; Wang et al., 2009). Due to its advantages and advancements in sequencing technologies, decreasing costs, and greater accessibility, RNA-seq has become the preferred method over microarrays nowadays (Mortazavi et al., 2008). To measure transcriptomics using RNA-seq, the process begins with the extraction of RNA from a tissue or cell sample, followed by the enrichment of messenger RNA (mRNA) and the synthesis of complementary DNA (cDNA) (Stark et al., 2019). The cDNA is then fragmented and sequenced to generate short reads, which are mapped to a reference genome and summarized into gene-specific counts (Conesa et al., 2016). The number of reads mapped to a particular gene reflects its expression level, providing a quantitative measure of transcriptional activity (Mortazavi et al., 2008). RNA-seq typically captures the expression of tens of thousands of genes offering a detailed snapshot of the transcriptional state of the sample under specific experimental conditions (Stark et al., 2019; Wang et al., 2009).

While transcriptomics provides valuable insights into the transcriptional activity of tissues and cells, it does not necessarily reflect their functional state given the low correlation between transcript and protein abundance (Maier et al., 2009) and the influence of additional regulatory mechanisms, including epigenetic (Fraser et al., 2021; Isbel et al., 2022), post-transcriptional (Zhao et al., 2017), and post-translational (Jensen, 2004) modifications. These regulatory layers can be explored using other omics technologies, which can complement transcriptomics data to provide a more comprehensive understanding of cellular functions and their regulation.

1.2.2 (Phospho)proteomics

Proteomics and phosphoproteomics enable the large-scale analysis of proteins and their modifications, specifically phosphorylations, within biological samples. This provides valuable insights into the molecular mechanisms underlying cellular states as proteins are the primary effectors of cellular function and morphology and are closely linked to the phenotype (Cox and Mann, 2007; Timp and Timp, 2020). High-throughput methods for proteomics are typically classified into antibody-based methods, such as protein microarrays, and mass spectrometry (MS)-based methods, which have become central to proteome analysis (Aslam et al., 2017). In recent years, advances in MS-based technologies have made it possible to quantify thousands of proteins (Beck et al., 2011; Nagaraj et al., 2011) and tens of thousands of unique phosphorylation sites across the proteome (Aebersold and Mann, 2016). To measure proteins and their modifications, proteins are first extracted from biological samples. Labeling techniques, which enable comparative quantification across samples, can be employed either at the biological sample stage, as for example stable isotope labeling by amino acids in cell culture, also known as SILAC (Ong et al., 2002), or after protein extraction like isobaric tags for relative and absolute quantitation, known as iTRAQ/TMT. After labeling, proteins are digested into peptides, typically using trypsin. For phosphoproteomics, additional enrichment steps are required to isolate phosphorylated peptides due to their low abundance. Enrichment techniques, such as immobilized metal affinity chromatography or titanium dioxide chromatography, are commonly employed to enhance the detection of phosphopeptides (Beausoleil et al., 2004; Pinkse et al., 2008). After enrichment, peptides are typically separated using liquid chromatography, enabling more accurate identification and quantification of peptides and their post-translational modifications (Bantscheff et al., 2012). After separation, peptides are introduced into the mass spectrometer, where their mass-to-charge ratios are measured, generating spectra used for peptide identification and quantification. Two principal acquisition technologies have been developed for mass spectrometry: data-dependent acquisition (DDA) and data-independent acquisition (DIA). In DDA, the mass spectrometer first performs a survey scan MS1 to detect all precursor peptide ions and then selects a predefined subset of these ions for fragmentation in a subsequent scan MS2 (Krasny and Huang, 2021). In contrast, in DIA, all precursor ions detected in the MS1 survey scan are fragmented, thanks to the recent advances in mass spectrometry technology, allowing faster scanning speeds and increased sensitivity (Gillet et al., 2012). DIA achieves this by using a series of wide isolation windows that collectively span the full mass-to-charge ratio range, capturing comprehensive fragmentation data (Krasny and Huang, 2021). The resulting spectra are then used to identify peptides, quantify their intensities, map them to the reference proteome, and ultimately summarize the data at the protein level for down-

stream proteomics analyses. With that, the abundance of proteins and phosphorylation sites in a given biological context can be measured and further analyzed.

1.2.3 Functional Analysis

To derive functional insights from omics data, the measurements are typically compared between two groups, usually in a case-control scenario. The objective is to identify genes, proteins, or phosphorylation sites whose expression or abundance levels differ between these groups. Numerous computational tools have been developed for this purpose, including limma, one of the most widely used tools (Ritchie et al., 2015). In the limma framework, the data is modeled as continuous and approximately normally distributed. Although RNA-seq data consist of discrete integer counts (Robinson and Oshlack, 2010), treating them as continuous takes advantage of the mathematical tractability of the normal distribution, which simplifies statistical analysis and computation (Law et al., 2014). Additionally, statistical methods based on approximately normally distributed data require that the standard deviations of the measurements remain more or less constant. However, omics data usually exhibits variance-mean dependence, where genes or proteins with larger abundance typically have larger standard deviations. To address this, normalization methods such as logarithmic transformation or variance-stabilizing normalization (vsn) can be applied to stabilize the variance (Holmes and Huber, 2019). After normalization, a linear model is fit to each gene, protein or phosphorylation site based on the experimental design to test the null hypothesis that no differential expression or abundance exists between experimental conditions. It calculates a moderated t-statistic for each molecule, which compares the mean expression or abundance differences between groups while accounting for variability and identifies differentially expressed or abundant molecules.

To further interpret omics data, computational methods have been developed to summarize gene- or protein-level information into higher-level functional processes. For example, enrichment analysis comprises statistical methods that organize gene- or protein-level statistics based on prior knowledge to estimate the likelihood of coordinated changes. The primary goal is to convert a disjointed list of differentially abundant molecules into a cohesive set of potential cellular processes that may be activated or repressed in a certain state of the tissue (Barry et al., 2005). These methods often rely on hypergeometric tests, rank-based statistics, or empirical likelihood estimations using permutations (Mathur et al., 2018; Väremo et al., 2013). Prior knowledge, linking genes or proteins to functional processes for enrichment analysis, is typically derived from databases like KEGG (Kanehisa et al., 2016) or Reactome (Fabregat et al., 2017), which curate col-

lections associated with signaling pathways or cellular processes. However, these sets typically group proteins as pathway members, which can limit the direct interpretation of transcriptomics data since the transcriptional state of a tissue does not necessarily reflect its functional state. Nevertheless, genes involved in a particular cellular process are usually regulated collectively and as such represent the functional footprint of cellular processes (Dugourd and Saez-Rodriguez, 2019; Szalai and Saez-Rodriguez, 2020). For instance, the functional footprint of a signaling pathway represents the genes that consistently change in expression when the pathway is perturbed, rather than merely the structural members of the pathway itself (Jiang et al., 2021; Schubert et al., 2018). Similarly, transcription factor activity can be inferred from the transcriptomic levels of its target genes, providing stronger evidence of functional relevance than the expression level of the transcription factor alone (Garcia-Alonso et al., 2018; Schubert et al., 2018). Likewise, kinase activity can be deduced from the phosphorylation state of its target proteins (Casado et al., 2013). For this, several initiatives have been undertaken to compile functional footprints for pathway and cytokine activities (Schubert et al., 2018), transcription factors (TFs) Garcia-Alonso et al., 2019, kinases (Hornbeck et al., 2012), and other perturbations (Dixit et al., 2016), providing valuable resources for these analyses.

1.3 Network Contextualization

As cells operate within complex networks of interactions, where molecules do not function independently but work coordinatively to transmit information, network biology has emerged as a powerful paradigm for understanding these interactions within biological systems (Barabási et al., 2011). To investigate the function of a cell or higher biological units, it is often advantageous to conceptualize them as systems of interacting elements. A graph-based representation is commonly used, where the system's elements are represented as nodes and their pairwise relationships are depicted as edges connecting these nodes (Pavlopoulos et al., 2011). In biological contexts, the nodes typically represent genes, proteins, or other molecules, and the directed edges represent regulatory or signaling relationships (Albert, 2007). For example, in signaling networks, these edges often represent protein-protein interactions, which are curated in databases such as STRING (Szklarczyk et al., 2019) and OmniPath (Türei et al., 2016). However, these databases usually aggregate all interactions in a cell-agnostic manner. In reality, the relevance of specific interactions is highly dependent on the biological context or cell type, as some proteins may not be expressed or active in certain conditions. To address this limitation, network contextualization has been developed to identify context-specific networks by integrating untargeted omics data with prior knowledge (Garrido-Rodriguez et al., 2022).

This approach not only enhances the biological relevance of the network but also provides a mechanistic explanation for observed patterns in the omics data. Additionally, by combining experimental data with curated prior knowledge, researchers can enhance the signal-to-noise ratio, thereby improving the robustness and interpretability of biological data analyses (Hill et al., 2016).

Multiple computational methods have been developed to model cellular signaling and infer context-specific networks, which can be categorized into different groups (Garrido-Rodriguez et al., 2022). Edge filtering and shortest path methods identify the most relevant pathways between nodes by filtering edges based on experimental data or finding the shortest paths connecting key nodes, such as specific genes or proteins of interest (Franke et al., 2006). Recursive signal propagation and heat diffusion approaches simulate the flow of information across the network, highlighting nodes and edges most influenced by the input data (Vanunu et al., 2010). Graph theory combined with statistical testing leverages mathematical properties, such as centrality or clustering coefficients, to pinpoint critical nodes or subnetworks under specific conditions (Yu et al., 2007). Bayesian networks model probabilistic dependencies between nodes, enabling inference of causal relationships and prediction of how perturbations propagate through the network (Friedman et al., 2000). Neural networks employ deep learning techniques to analyze complex patterns in signaling data, integrating large-scale omics datasets to predict signaling activity and uncover nonlinear relationships within the network Fortelny and Bock (2020). Finally, Integer linear programming (ILP) optimizes the selection of nodes and edges that best explain the observed omics data while adhering to constraints such as known interactions (Dittrich et al., 2008). More specifically, ILP is a mathematical optimization technique where variables are restricted to integer values. The goal is to maximize or minimize a linear objective function subject to a set of linear constraints. This restriction makes ILP problems inherently combinatorial, often involving the search for an optimal solution among a finite set of possibilities (Bragin et al., 2019).

These computational approaches enable the identification of context-specific networks, providing a framework for the mechanistic interpretation of omics data and advancing our understanding of cellular signaling in different biological contexts.

1.4 Evaluation of Computational Methods

Biological systems are inherently complex, characterized by a multitude of interconnected pathways and regulatory layers that are dynamically influenced by the environment. Computational tools attempt to capture this complexity through mathematical mod-

els, statistical algorithms, and machine learning techniques. However, these models often simplify reality, relying on assumptions such as linear relationships, independence of variables, or static network structures. While these assumptions make the tools computationally tractable and interpretable, they may fail to fully represent the dynamic and nonlinear nature of biological systems (Kitano, 2002).

In computational biology as well as other sciences, researchers are often faced with a choice between up to hundreds of methods for performing a certain type of data analysis (Weber et al., 2019). This can represent both an opportunity and a challenge, since each method usually proposes certain advantages and limitations, and the choice of method can heavily influence the results. Complicating matters further, methods are often presented as optimal within the context of their own evaluations, making it difficult to determine the most suitable approach for a given scenario (Norel et al., 2011).

As such, neutral comparison studies in computational sciences are essential to ensure that proposed methods perform as expected across diverse scenarios. These studies also help to establish standard practices grounded in robust, well-designed evaluations (Boulesteix et al., 2013). Typically, benchmarks for such comparisons are based on specific input data that the methods are designed to process, expected output data for validation, defined metrics for performance assessment, and performance values from a set of tools tested under the same conditions (Peters et al., 2018). For example, perturbation experiments, where specific kinases or TFs are experimentally manipulated, have been established as benchmarks for evaluating activity inference tools (Alvarez et al., 2016; Subramanian et al., 2005). Nevertheless, defining the expected output data can be challenging in biology due to the inherent noise and lack of clear ground truth in experimental datasets. Additionally, selecting appropriate input datasets is a critical design choice, as these datasets must reflect the complexity and variability of real-world biological data. To address this, diverse benchmarking evaluations are often employed, encompassing a variety of metrics and datasets that capture the range of conditions under which the methods are expected to operate. This diversity helps to ensure that performance estimates are robust and generalizable to practical applications and help identify strengths and weaknesses in tool performance (Weber et al., 2019).

For the evaluation, the choice of performance metrics usually depends on the specific task and type of data being analyzed. Tasks such as classification, clustering, or regression require tailored metrics to assess performance (Weber et al., 2019). For classification tasks, metrics such as the area under the receiver operating characteristic curve (AUROC) and area under the precision-recall curve (AUPRC) are commonly used. AUROC eval-

uates the trade-off between true positive rate (TPR) and false positive rate (FPR) across various thresholds, with a perfect classifier achieving a score of 1 and a random classifier scoring 0.5. In contrast, AUPRC measures the relationship between precision (positive predictive value) and recall (sensitivity), making it particularly valuable for imbalanced datasets where one class is significantly underrepresented (Saito and Rehmsmeier, 2015). These metrics provide a quantitative framework for comparing the performance of computational methods, enabling a rigorous assessment of their accuracy in representing biological processes. Additionally, they offer critical insights into the applicability of these methods for addressing complex biological questions.

1.5 Thesis Overview

Computational tools offer a powerful means to extract meaningful insights from omics data, aiding in the understanding of signaling mechanisms in health and disease. However, to ensure their reliability and accuracy, rigorous evaluation of computational methods is essential. Furthermore, developing advanced computational methods to incorporate multiple layers of information, including multiple omics data, provides a more comprehensive view of biological processes, enabling deeper insights into complex systems.

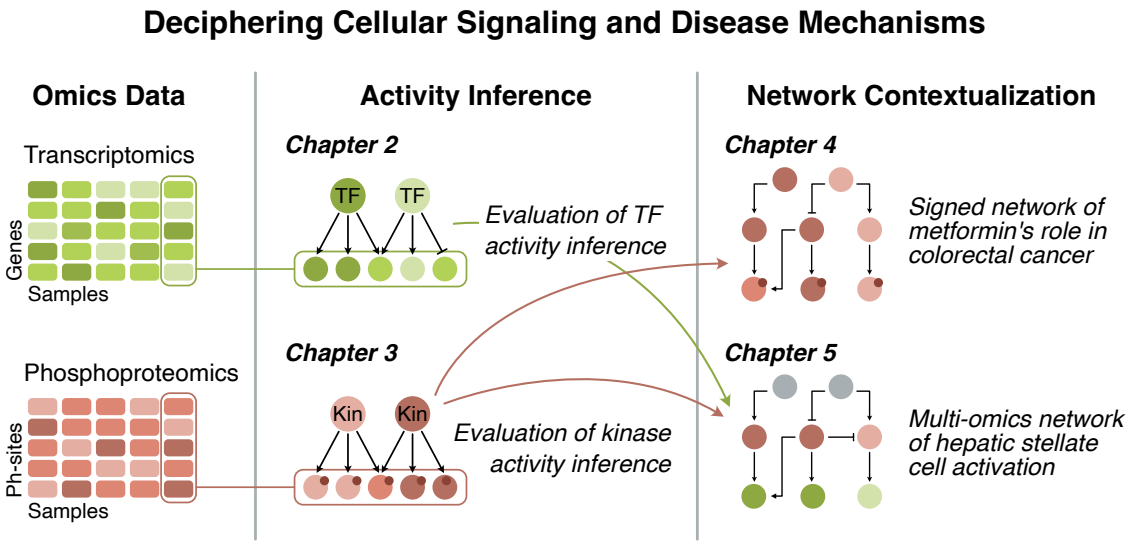

Figure 1.2. Graphical overview of this thesis. Overview of how this thesis investigates signaling through the integration of omics data and prior biological knowledge. Chapter 2 and 3 evaluate methods for the prediction of transcription factor (TF) and kinase activities from transcriptomics and phosphoproteomics data, respectively. Chapter 4 presents a network contextualization approach based on inferred kinase activities to investigate the effect of metformin on colorectal cancer cells. Lastly, Chapter 5 introduces a multi-omics network contextualization connecting TFs and kinases to upstream stimuli in a cohesive manner which was used to study hepatic stellate cell activation.

This thesis aims to enhance the understanding of signaling events by evaluating and advancing computational tools for activity estimation and network contextualization from different omics data (Figure 1.2). More specifically, in Chapter 2 ”Expanding the Coverage of Regulons for Accurate Estimation of Transcription Factor Activities”, I evaluate various TF-gene collections for TF activity estimation. Chapter 3 ”Comprehensive Evaluation of Kinase Activity Inference from Phosphoproteomics Data” focuses on the assessment of kinase activity inference tools. In Chapter 4 ”Phosphoproteomics Based Network Contextualization to Understand Metformin’s Mechanisms in Colon Cancer”, I extend a network contextualization approach for phosphoproteomics data to account for directionality and apply it to investigate the mode of action of metformin and its potential role in colon cancer. Lastly, in Chapter 5 ”Multi-Omics Network Contextualization to Investigate Hepatic Stellate Cell Activation”, I build a network contextualization model that integrates phosphoproteomics and transcriptomics data to study hepatic stellate cell activation by combining insights from signaling and gene regulation.

Chapter 2

Expanding the Coverage of Regulons for Accurate Estimation of Transcription Factor Activities

2.1 Background

Gene regulation is crucial for coordinating biological processes, governing development, cell differentiation, tissue maintenance, and a wide range of physiological functions. When this regulation is disrupted, it can lead to numerous diseases, including cancer, autoimmune disorders, neurological conditions, developmental syndromes, diabetes, and cardiovascular diseases (Lee and Young, 2013). In particular, abnormal activity of transcription factors (TFs), which are central to regulating gene expression, has been closely linked to cancer progression and can destabilize key cellular regulatory systems (Lambert et al., 2018). TFs can influence the transcription rates of their target genes, collectively referred to as the TF's regulon, by binding to specific DNA sequences. These regulatory processes are further shaped by interactions with cofactors and other proteins (Kim and Wysocka, 2023). Gene regulatory networkss (GRNs) try to capture these interactions as a simplified representation of the underlying complexity (Weidemüller et al., 2021). Coupling GRNs with activity inference algorithms can aid the interpretation of transcriptomics data by identifying highly active or inactive TFs from the experimental data. Among other things, this can help to uncover insights into aging processes (Maity et al., 2022) link TF activities to drug responses (Garcia-Alonso et al., 2018; Melms et al., 2021), and connect TF activity to the morphological characteristics of cancer (Walsh et al., 2017). However, the reliability of these analyses depends heavily on the coverage and

quality of the TF regulons used. Ensuring the use of high-quality regulons that minimize false positives while ensuring comprehensive coverage is essential for accurately identifying key TFs and their roles in regulatory networks.

In this chapter, I will present and evaluate a new set of TF regulons generated from TF-gene interaction data from public databases, text mining, and manual curation. This work was a joint project with Eirini Tsirvouli and Astrid Lægreid from the Norwegian University of Science and Technology and Miguel Vazquez from the Barcelona Supercomputing Center and has been published in a peer reviewed journal (Müller-Dott et al., 2023). I will focus specifically on the parts of the project that have been my contribution, highlighting my responsibilities of developing a workflow to determine whether interactions are activating or repressing, performing a systematic benchmark to evaluate the TF regulons and showing the value of TF activity inference in single-cell. The code to reproduce all the analysis presented can be found here `https://github.com/saezlab/CollecTRI`.

2.2 Building a High-Confidence Collection of Signed Transcription Factor-Gene Interactions

2.2.1 Collection of Transcription Factor Regulons in CollecTRI

The collection of highly reliable transcription factor (TF)-gene interactions was based on an already published compilation of available transcription regulation information from multiple databases combined with information extracted from the text-mining resource ExTRI (Vazquez et al., 2022). The initial compilation included the databases TFactS (Essaghir et al., 2010), HTRIdb (Bovolenta et al., 2012), IntAct (Kerrien et al., 2012), GOA (Huntley et al., 2015), TRRUST (Han et al., 2015), SIGNOR (Perfetto et al., 2016), CytReg (Carrasco Pro et al., 2018) and GEREDB (Huang et al., 2019)). This dataset was then expanded with updated data from SIGNOR and GOA, along with three additional resources: DoRothEA (confidence level A) (Garcia-Alonso et al., 2019), Pavlidis (Chu et al., 2021), and the NTNU Curated subset of ExTRI. DoRothEA, a meta-resource containingTF-gene interactions categorized by varying confidence levels, was filtered to include only the highest-confidence interactions (level A). The NTNU Curated subset of ExTRI provided manually curated interactions with additional information on the regulatory sign (e.g., activation or repression). Gene and protein mentions from all resources were translated into human gene symbols, including those referring to rat or mouse entities, using orthology tables. Mouse and rat TF-gene interactions were included, as regulatory mechanisms are highly conserved across these species, and text-mined data often

lacks precise species annotations in PubMed abstracts (Chen et al., 2018). Moreover, in this collection, two TF dimers, AP1 and NFKB, were treated as single transcription factors, as they are often referred to by their dimer names in the literature. Finally, this resulted in a comprehensive table where each TF-gene interaction was documented alongside the databases in which it appears, including additional details such as the mode of regulation (when available) and the PubMed identifiers (PMIDs) used to curate or extract the interaction.

From the compiled information, I then constructed signed and directed TF regulons for TF activity inference. To ensure reliability of the TF-gene interactions, I first identified unique PMIDs associated with each interaction and excluded those without any reference. Next, I focused on proteins with a direct regulatory role in gene expression, including only TFs classified as DNA-binding transcription factors (dbTFs), co-regulatory transcription factors (coTFs), or general initiation transcription factors (GTFs). These classifications were based on criteria from TFclass (Wingender et al., 2018), Lambert et al. (2018), Lovering et al. (2021), and gene ontology (GO) annotations (Gene Ontology Consortium, 2021). CoTFs and GTFs were further filtered to include only proteins annotated with the GO terms GO:0003712 and GO:0140223, or their descendant terms, using QuickGO (Binns et al., 2009). This resulted in a total of 43,175 TF-gene interactions covering 1,186 TFs (Figure 2.1).

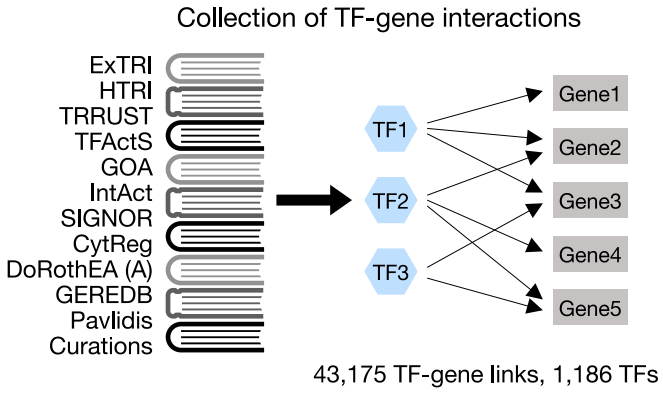

Figure 2.1. Overview of the collection of TF-gene interactions. Compilation of transcription factor (TF)-gene interactions from different resources, namly ExTRI, HTRI, TRRUST, TFactS, GOA, IntAct, SIGNOR, CytReg, DoRothEA A, GEREDB, Pavlidis, and manual curations. This collection comprises 43,175 TF-gene links covering 1,186 unique TFs. *Reprinted from Müller-Dott et al. (2023).*

2.2.2 Defining the Mode of Regulation for TF-Gene Interactions

For each TF-gene interaction, I assigned a regulatory mode indicating the sign of transcriptional regulation from the TF to its target gene. Activation was defined as an in-

crease in the target gene's expression, while repression corresponded to a decrease. I then determined the mode of regulation (MoR) using the evidence provided by the PMIDs associated with each interaction in the databases. For this, each PMID was counted as evidence only once per interaction, even if referenced in multiple databases. In the rare cases where the same PMID supported conflicting modes of regulation across databases, it was considered separately for each mode to determine the final assignment. I then explored different strategies to assign the MoR based on multiple sources of information (Figure 2.2). I compared four approaches in combination and separately:

1. Assigning the mode of regulation per TF-target interaction based on the prevalence of PMIDs associated with a specific mode.

2. Assigning a mode of regulation of a TF based on its general mode of regulation, defined by prior information about the regulatory information of the TF.

3. Assigning a positive mode of regulation of a TF based on its general mode of regulation, defined by other interactions in the regulon.

4. Assigning a default mode of regulation to TF-target interactions.

For the second approach, I integrated information from multiple sources to establish the general mode of regulation of each TF. I compiled annotations from GO terms and UniProt keywords (UniProt Consortium, 2023), structural information about the Krüppel associated boxs (KRABs) domain, and the classification of effector domains (Soto et al., 2022). Specifically, I checked if a TF was annotated with GO terms such as RNA polymerase II-specific DNA-binding transcription activator activity (GO:0001228), DNA-binding transcription repressor activity (GO:0001217), transcription coactivator activity (GO:0003713), transcription corepressor activity (GO:0003714), positive regulation of transcription by RNA polymerase II (GO:0045944), or negative regulation of transcription by RNA polymerase II (GO:0000122), including any child terms. For the UniProt keywords, I focused on annotations for Activator (KW-0010) and Repressor (KW-0678). TFs containing the KRAB domain were classified as repressors if they belonged to the IPRO3651 superfamily, except for members of the IPRO03655 ancient KRAB family, which are known to exhibit dual regulatory roles (Paysan-Lafosse et al., 2023). I determined the final regulatory mode of each TF based on the consensus among all sources and then assigned the mode of regulation to all target genes accordingly. Using this strategy, I identified 348 TFs as general activators and 232 as general repressors. This classification resulted in 10,313 TF-gene links assigned an activating mode and 3,191 links assigned a repressing mode. For comparison, the prevalence of PMIDs indicated 13,847 activat-

ing and 5,694 repressing TF-gene links. Both approaches showed a similar trend, with a majority of interactions assigned a positive regulatory mode (76% using TF classification and 71% using PMIDs).

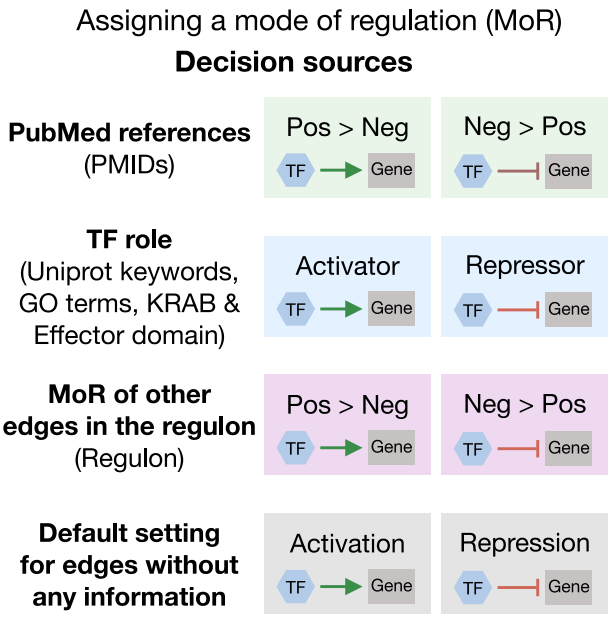

Figure 2.2. Overview of strategies for assigning a mode of regulation to TF-gene interactions. Four decision sources were used to determine whether a transcription factor (TF) activates or represses its target gene: PubMed references (PMIDs): The mode of regulation (MoR) is assigned based on the prevalence of supporting evidence for activation (Pos > Neg) or repression (Neg > Pos). TF role: The general regulatory role of the TF (activator or repressor) is determined using UniProt keywords, GO terms, and structural information such as KRAB domains and effector domain classifications. MoR of other edges in the regulon: The mode of regulation is inferred by analyzing the majority behavior of the TF's other interactions within the regulon (Pos > Neg or Neg > Pos). Default setting: For TF-gene interactions without any supporting information, a default mode of regulation is assigned as either activation or repression. *Reprinted from Müller-Dott et al. (2023).*

To evaluate how assigning a mode of regulation impacts TF-gene interactions, I tested whether the regulatory signs influenced the ability of TF regulons to capture gene expression changes resulting from TF perturbations. The rationale was that a reliable regulon, in which a TF accurately regulates its target genes, should reflect the TF's transcriptional activity through the collective expression patterns of its targets (Garcia-Alonso et al., 2019). For this analysis, I used data from KnockTF, a curated database of human RNA-seq and microarray experiments from TF knockdown and knockout studies (Feng et al., 2020). KnockTF includes datasets from various tissues and cell types. To ensure that only effective perturbations were included, I filtered for experiments where the targeted TF showed a significant decrease in expression post-knockdown or knockout. This

resulted in a dataset of 388 experiments spanning 234 TFs. Using the compiled dataset, I inferred TF activities for the different MoR strategies, ranked the TFs based on their inferred activity, and assessed the ability to recapitulate the perturbed TF using the area under the receiver operating characteristic (AUROC) and area under the precision-recall curve (AUPRC) metric (Figure 2.3). This evaluation followed the benchmarking strategy previously applied to compare inference algorithms (Badia-I-Mompel et al., 2022). Specifically, I calculated activity scores for each perturbation experiment with a univariate linear model, which measures the contribution of the regulon's targets to the overall transcriptional activity. Perturbed TFs were expected to show a strong negative activity score, reflecting their reduced function. To rank the TFs, I inverted the activity scores, assigning higher scores to the TFs with the strongest decrease in activity, consistent with expectations for knocked-out or knocked-down TFs. To address the class imbalance between perturbed and non-perturbed TFs, I employed a downsampling strategy. In each iteration, equal numbers of positive and negative classes were randomly sampled, and AUROC and AUPRC metrics were calculated. This process was repeated 1,000 times per network, generating distributions of performance metrics to ensure robust benchmarking.

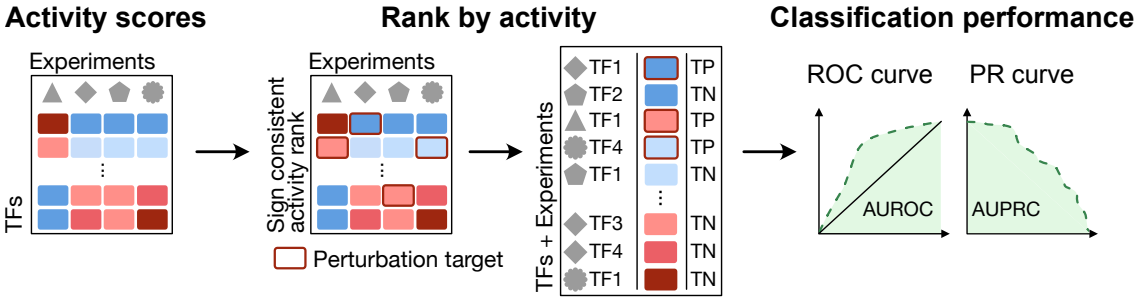

Figure 2.3. Evaluation approach for TF activity estimation. Description of benchmarking approach for the evaluation of transcription factor (TF) activity inference. TF activities are inferred from gene expression data of perturbation experiments, ranked by their activity score and aggregated across experiments. Next, the area under the Receiver operating characteristic curve (AUROC) and Precision–Recall curve metrics (AUPRC) are calculated. Here, a downsampling strategy is applied to have an equal number of perturbed and non-perturbed TFs.

Finally, applying this benchmarking strategy, I evaluated the performance of the different sign assignment strategies. Initially, I compared the first two approaches, assigning the mode of regulation per TF-target interaction based on the prevalence of PMIDs associated with a specific mode and assigning a mode of regulation of a TF based on its general mode of regulation, defined by prior information about the regulatory information of the TF and compared them to TF regulons where all interactions were assigned an activating mode of regulation. I also tested a combined approach, where the mode of

regulation was first assigned based on PMIDs, followed by assigning a general mode of regulation to TF-gene interactions without PMID evidence. Any remaining interactions were assigned an activating mode of regulation. This revealed that assigning the mode of regulation based on PMIDs significantly outperformed the regulons with all interactions rendered activating (adjusted p-value < 2.2×10^{-16}, t-value = 103 for AUROC and 69.6 for AUPRC). In contrast, using prior knowledge to assign a general mode of regulation led to a decrease in performance (adjusted p-value = 6.5×10^{-5} and 1.5×10^{-10}, t-value = 4 and 6.5 for AUROC and AUPRC, respectively) (Figure 2.4a). To address TF-gene interactions lacking information from databases, I incorporated data from other interactions within the TF's regulon. For each TF, I analyzed the interactions with an assigned mode of regulation based on PMIDs and classified the TF as activating or repressing based on the majority mode of its interactions. This approach added 1,750 activating and 154 repressing TF-gene links and significantly improved benchmark performance (adjusted p-value = 2.5×10^{-13} and 1.2×10^{-33}, t-value = 7.4 and 12.3 for AUROC and AUPRC, respectively). However, once again, incorporating additional information about the general classification of a TF based on prior knowledge did not improve performance. (Figure 2.4b). Lastly, I assessed the impact of assigning a default mode of regulation when no other information was available. Assigning a default activating mode outperformed assigning a default repressing mode (adjusted p-value < 2.2×10^{-16}, t-value = 337 and 294 for AUROC and AUPRC, respectively) (Figure 2.4c).Based on these results, the mode of regulation for each TF-gene interaction was determined using the prevalence of PMIDs, classification of the TF based on the mode of other interactions in its regulon, and a default assignment of activating mode for interactions without available information. Specifically, 19,541 TF-gene interactions were annotated based on the prevalence of PubMed references, while 1,904 interactions were assigned a mode of regulation using the TF classification inferred from other interactions within the regulon. For the remaining 21,730 interactions, an activating mode was assigned by default. This final annotation process resulted in 86% of TF-gene links being classified as activating and 14% as repressing (Figure 2.4d).

In conclusion, the mode of regulation for TF-gene interactions was assigned based on the prevalence of PMIDs, the other interactions in the regulon and a default activation for interactions without available information, resulting in 56% of TFs with a dual regulatory role, 37% of TFs exclusively linked to activating interactions, and 7% represented only by repressing links.

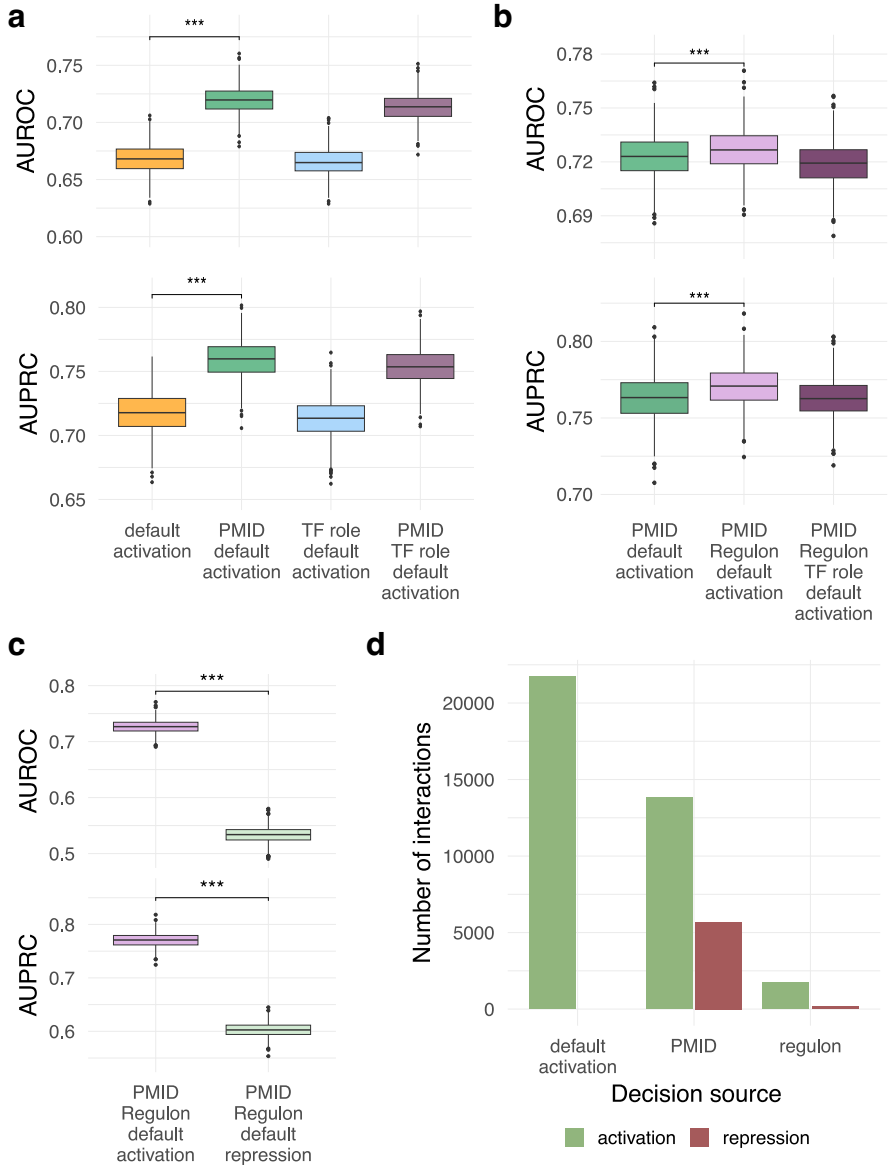

Figure 2.4. Evaluation of strategies for assigning the mode of regulation to TF-gene interactions. **a-c** Area under the receiver operating characteristic (AUROC) (top) and area under the precision-recall curve (AUPRC) (bottom) values for transcription factor (TF) activity inference using different mode of regulation (MoR) assignment strategies. The MoR was assigned based on the prevalence of PubMed IDs (PMIDs), the general regulatory role of the TF (activator or repressor) (TF role), the regulatory role of the TF based on the TF's other interactions within the regulon (regulon) and the default assignment (default activation, default repression). **d** Number of TF-gene interactions assigned as activating or repressing based on the different decision sources when combining PMIDs, regulon-based classification and default activation as the best performing combination. *Reprinted from Müller-Dott et al. (2023).*

2.2.3 Weighting Interactions Based on Binding Weights

Besides the mode of regulation, gene regulation also has a quantitative aspect that influences the interaction between TFs and their target genes (Kim and Wysocka, 2023). To

capture this aspect, I implemented a weighting scheme for TF-gene interactions, which reflects the likelihood of a TF binding to a gene's regulatory regions. This approach assumes that a TF's activity is best reflected in the coordinated expression of genes where it has a strong binding affinity, particularly within promoter regions or other regulatory elements. Consequently, genes enriched with the binding motifs of a specific TF were assigned higher weights, while genes without enriched motifs were given lower weights.

To estimate binding weights, I used two tools, FIMO (Grant et al., 2011) and MatrixRider (Grassi, 2017), focusing on known TF binding motifs and genomic regions. I first calculated binding weights for promoter regions, defined as 1,000 base pairs (bp) upstream and 100 bp downstream of the transcription start site (TSS) (Sanghi et al., 2021). Additionally, I extended the regulatory region to 10,000 bp upstream of the TSS to include proximal elements. All binding weights were shifted to positive values with a pseudo count of 1 and normalized using two strategies: (1) per TF, where weights for all target genes of a specific TF were divided by the highest binding weight for that TF, and (2) per gene, where weights for all TFs regulating a specific gene were divided by the highest weight for that gene. The binding weights inferred across tools, window sizes, and normalization strategies were highly correlated, with Pearson correlation coefficients exceeding 0.98 (Figure 2.5a). Due to this strong correlation, I focused on weights calculated with MatrixRider for the 1,000 bp upstream window in subsequent analyses. Using these weights, I inferred TF activities and evaluated their performance within the benchmarking framework described earlier. The weighted regulons performed comparably to the unweighted regulons, showing no significant improvement (t-test: adjusted p-value = 0.51 for AUROC and 0.9 for AUPRC; Figure 2.5b). Next, I explored the effect of pruning low-weight edges from the network, comparing this approach to randomly removing edges. Specifically, TF-gene interactions in the lowest 10%, 20%, and 30% quantiles of binding weights were removed, and the resulting networks were evaluated in the benchmark. However, this pruning strategy did not improve network performance for either AUROC or AUPRC in any scenario (t-test: adjusted p-value > 0.05; Figure 2.5c).

Overall, weighting TF-gene interactions did not enhance the predictive performance of the regulons and remains feasible only for a subset of TF-gene interactions.

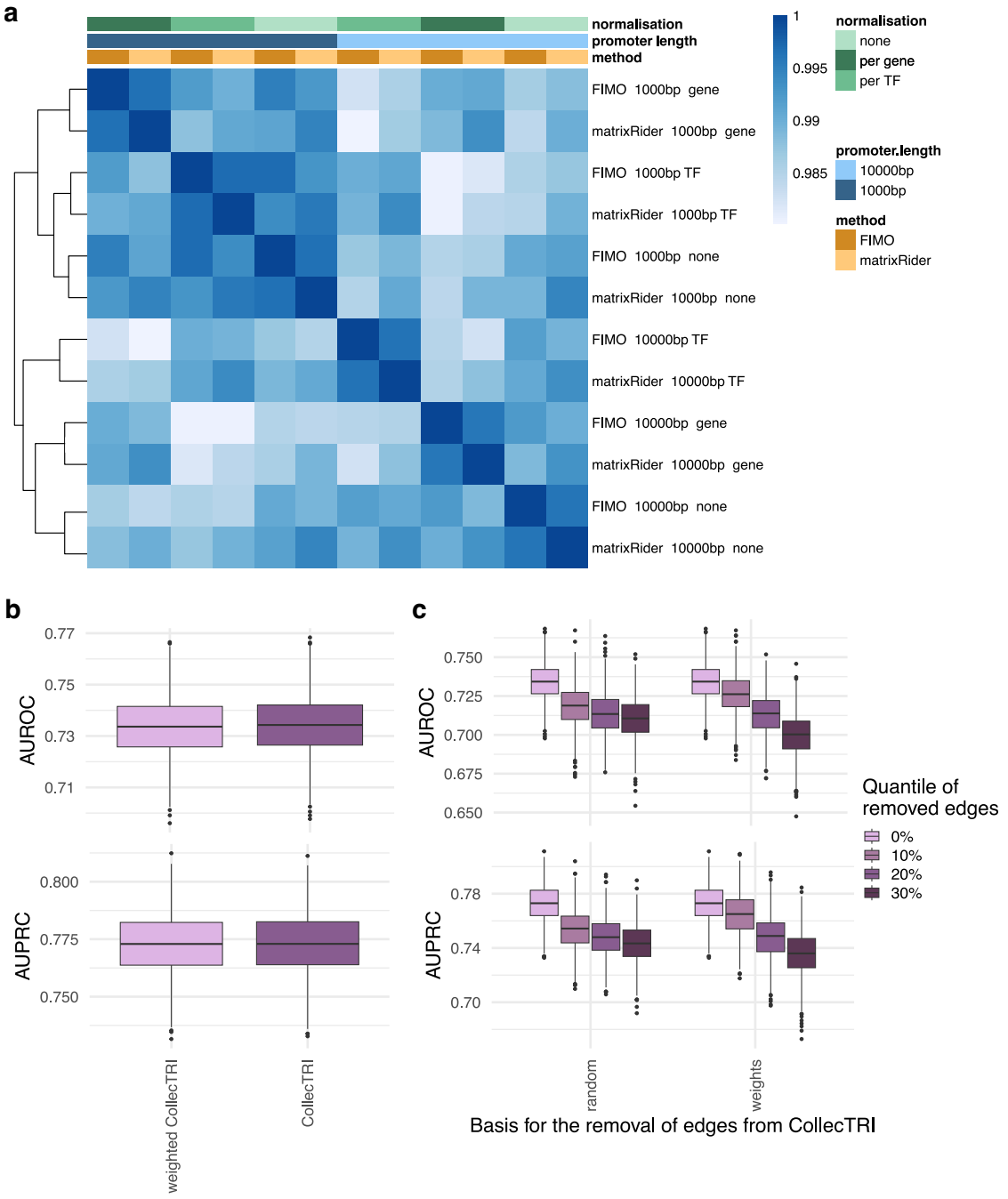

Figure 2.5. Evaluating the effect of weighting TF-gene interactions. **a** Pearson correlation of inferred binding weights across tools, window sizes, and normalization strategies. Binding weights were computed using MatrixRider and FIMO for regions 1,000 bp and 10,000 bp downstream of the transcription start site. These weights were normalized either per gene or per transcription factor (TF), and Pearson correlation was used to assess the similarity of TF-gene interaction weights across methods and conditions. **b** Predictive performance of weighted versus unweighted TF regulons derived from CollecTRI. Binding weights were calculated with MatrixRider for a 1,000 bp window and used to evaluate the impact of weighting on regulon-based predictions. **c** Predictive performance of regulons after filtering low-weight TF-gene interactions. Binding weights, prepared as in panel b, were used to examine the effect of removing the lowest 10%, 20%, and 30% of TF-gene interactions compared to randomly removing the same proportion of edges. *Reprinted from Müller-Dott et al. (2023).*

2.3 Evaluation of the Transcription Factor Regulon Collection

2.3.1 Comparison to Other Regulon Collections

Building on the constructed TF regulon collection, I evaluated its coverage and compared it to other widely used collections, including ChEA3 (Keenan et al., 2019), RegNetwork (Liu et al., 2015), Pathway Commons (Rodchenkov et al., 2020), and DoRothEA (Garcia-Alonso et al., 2019). ChEA3 provides gene set libraries based on TF-gene co-expression, TF-target associations derived from ChIP-seq experiments, and TF-gene co-occurrence data submitted through the Enrichr tool. RegNetwork is a curated database of experimentally observed and predicted transcriptional and post-transcriptional regulatory interactions. Pathway Commons compiles data on regulatory networks, signaling pathways, molecular interactions, and DNA-binding events from multiple sources. DoRothEA integrates TF-gene interactions with assigned confidence levels derived from literature-curated resources, ChIP-seq peaks, motif analyses, and gene expression data. Among these collections, only DoRothEA, also provides signed information on the direction of transcriptional regulation.

For a consistent comparison, I first filtered all TF regulons to include only interactions involving annotated dbTFs, coTFs, GTFs, as described earlier. I then analyzed the overlap of TFs and TF-gene interactions across these collections. The CollecTRI-derived regulons demonstrated broad TF coverage, including 1,186 TFs—surpassed only by the ChEA3 gene set libraries ARCHS4 (1,612), GTEx (1,578), and Enrichr (1,393). However, these ChEA3 libraries rely on co-expression and co-occurrence strategies, which are known to generate a higher rate of false-positive interactions in TF-target association studies (Huynh-Thu et al., 2010). The CollecTRI regulons included 48 new TFs not present in any of the other resources. RegNetwork and Pathway Commons also contributed unique TFs, with 80 and 42 additional TFs, respectively. Across all collections, 91.3% of TFs were shared by at least two resources. In terms of TF-gene interactions, curated resources such as RegNetwork, Pathway Commons, DoRothEA, and the CollecTRI regulons generally included fewer interactions compared to ChEA3, which relies on co-expression and co-occurrence methods. These methods often yield larger datasets but tend to include indirect or spurious regulatory relationships. Overall, there was limited overlap between the compared resources, with 63.8% of interactions unique to a single collection. For shared TFs, the overlap in target genes was low, with the mean Jaccard index of my regulons compared to other networks at just 0.01 (Figure 2.6).

In summary, the TF regulons constructed from CollecTRI provide a broader coverage of TFs compared to most other resources and additionally include signed information on transcriptional regulation, distinguishing them from most other regulon collections. Overall a low overlap between TF-gene interactions was observed across resources which is likely to influence TF activity inference.

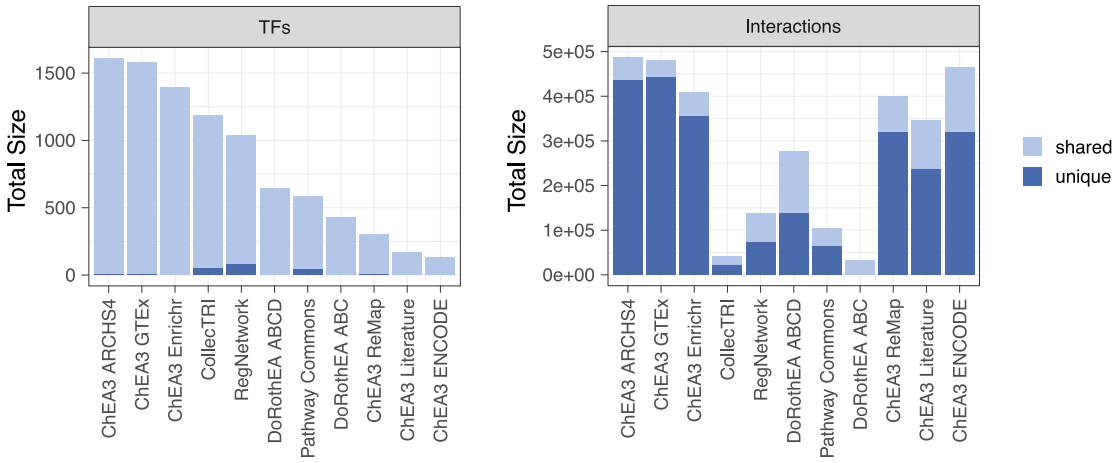

Figure 2.6. Comparison of TF coverage and TF-gene interaction across regulon collections. Total number of transcription factors (TFs) (left) and TF-gene interactions (right) represented in each resource. Any TF or interaction present in more than one resource is considered shared and highlighted in light blue. *Reprinted from Müller-Dott et al. (2023).*

Next, I used the same benchmarking pipeline as described in the previous section to systematically compare the CollecTRI-derived regulons to the other TF regulon collections. Additionally, I created a permuted version of the CollecTRI regulons as a baseline for performance, where target genes were shuffled and randomly assigned to TFs, ensuring that these regulons did not represent biological information. TF activities were inferred from the differentially expressed genes in each KnockTF experiment using the regulons provided by each resource. To ensure robustness, only TF regulons with at least five target genes measured in an experiment were included in the analysis, leading to a restricted number of TFs for each resource (Supplementary Figure A.1). Inferred TF activities were then ranked by their activity scores for each experiment, and the identification of perturbed TFs based on these scores was evaluated using AUROC and AUPRC. In this evaluation, the CollecTRI regulons achieved median AUROC and AUPRC values of 0.73 and 0.77, respectively, outperforming all other regulon collections (adjusted p-value $< 2.2\times10^{-16}$, mean t-value across tests = 271.8 for AUROC and 281.8 for AUPRC) (Figure 2.7, Supplementary Table A.1, A.2). Furthermore, all ChEA3 libraries, except for ChEA3 ARCHS4, did not perform better than the random baseline set by the permuted CollecTRI regulons (t-test: adjusted p-value > 0.05).

Overall, these results demonstrate that CollecTRI regulons outperform other TF regulon collections in identifying perturbed TFs based on inferred TF activities which indicates that the TF-gene interaction information compiled in CollecTRI provides the most reliable regulons for estimating TF activities among the resources compared.

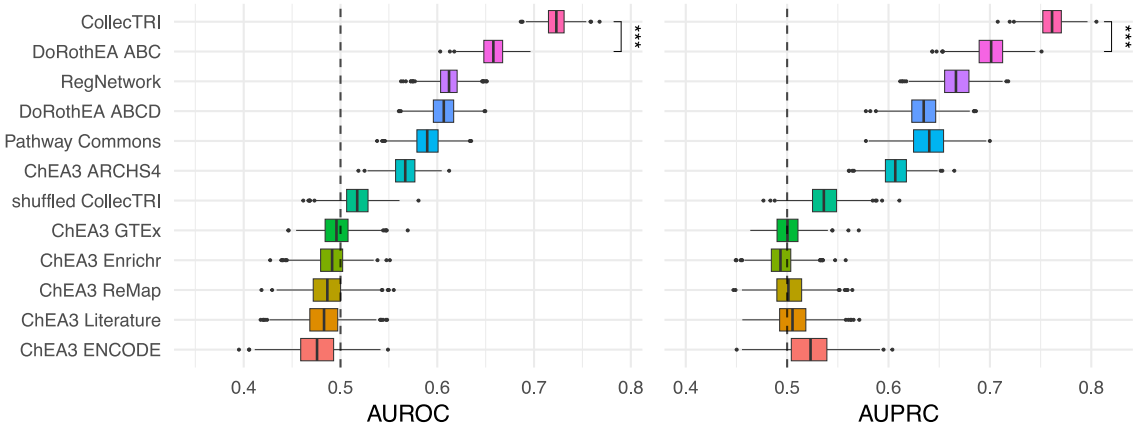

Figure 2.7. Predictive performance of TF regulon collections for TF activity inference. Performance of transcription factor (TF) regulon collections in predicting perturbed TFs from the knockTF database based on inferred TF activities. Performance is assessed using the area under the receiver operating characteristic curve (AUROC, left) and the area under the precision-recall curve (AUPRC, right). *Reprinted from Müller-Dott et al. (2023).*

2.3.2 Assessing Benchmark Fairness and Performance Robustness

To ensure that the benchmark comparison in the previous section was unbiased and reliable, I examined whether the number of target genes associated with each TF influenced the ability of the networks to predict perturbed TFs. Since the benchmark dataset predominantly includes well-studied TFs, which typically have a larger number of associated targets, I investigated whether this factor contributed to the observed performance differences. For the top three performing TF regulon collections, CollecTRI, DoRothEA ABC, and RegNetwork, I first assessed whether the number of targets differed between benchmarked TFs and those not included in the benchmark. In all three cases, TFs in the benchmark dataset had significantly more associated targets (adjusted p-values = 2.61×10^{-5}, 1.34×10^{-3}, and 2.81×10^{-4}; t-values = 4.59, 3.27, and 3.84 for CollecTRI, DoRothEA ABC, and RegNetwork, respectively) (Figure 2.8a). To further explore the relationship between the number of targets and the accuracy of TF activity inference, I calculated Pearson correlation coefficients between the activity scores and the number of targets for each experiment in the benchmark dataset. The average correlation between the number of targets and the activity score was 0.4 or lower across all resources, with CollecTRI-derived

regulons showing the lowest mean correlation of 0.19 (Figure 2.8b). These results indicate that the superior performance of the CollecTRI-derived regulons is not influenced by a bias toward TFs with a higher number of associated targets, supporting the robustness of the benchmark comparison.

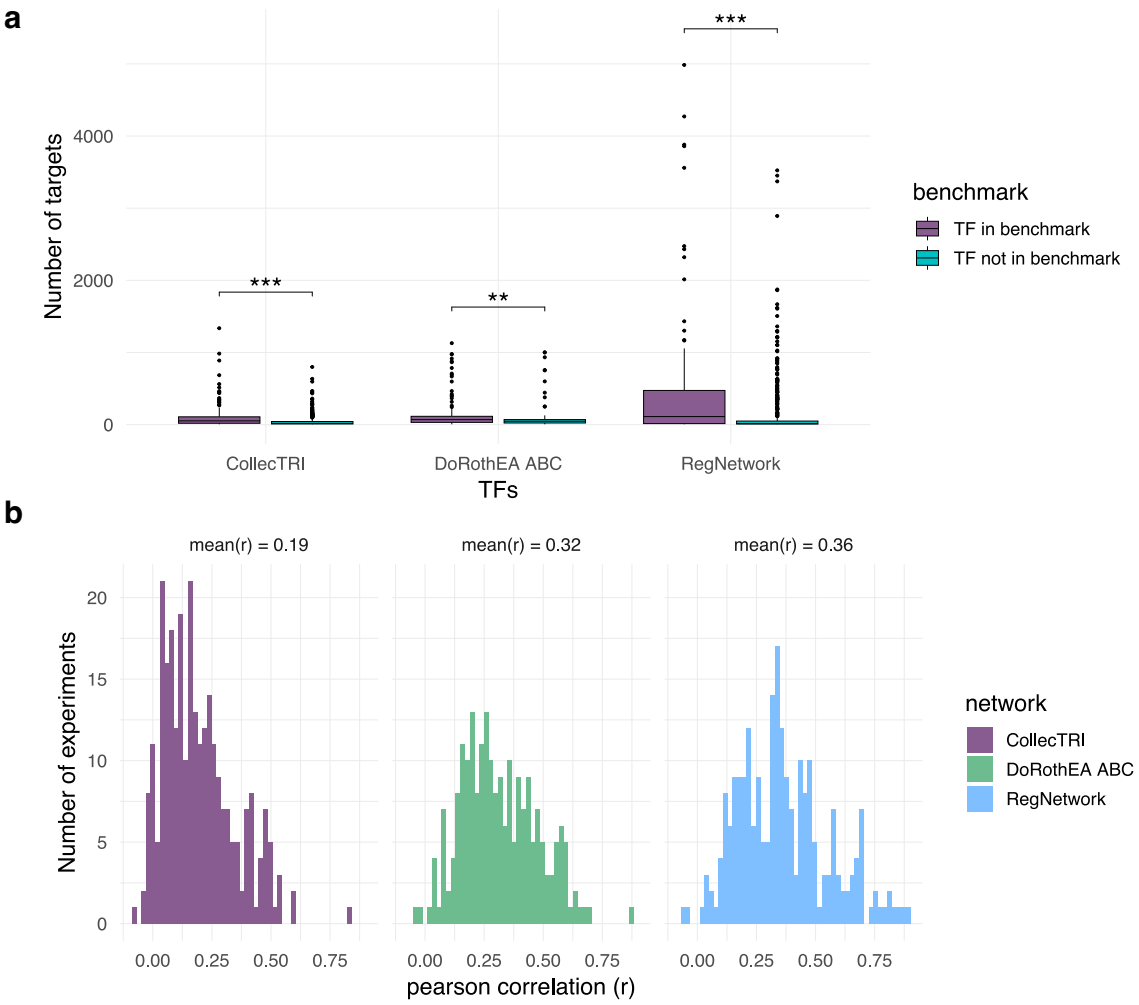

Figure 2.8. Comparison of the number of targets of a TF and its inferred activity. **a** Number of target genes for transcription factors (TFs) included in the benchmark and those not included across different regulon collections. **b** distribution of Pearson correlation coefficients (r) between the number of target genes and prediction accuracy for TF activity across benchmark experiments. Mean correlations are reported above the plots. *Reprinted from Müller-Dott et al. (2023).*

Another limitation of the benchmark is that it assumes that the perturbed TF is the most deregulated, ignoring potential off-target effects of the perturbation. To address this, I conducted another evaluation using a subset of 12 TFs for which multiple perturbation experiments were available. In this setting, I focused on one TF at a time, comparing its activity across experiments to predict whether it was perturbed or non-perturbed in each experiment. In this benchmark, the CollecTRI regulons demonstrated superior perfor-

mance for specific TFs, including REST, TP53, FLI1, NRF2F2, and SOX2, with average median AUROC and AUPRC values of 0.85 and 0.89, respectively (adjusted p-value < 1.8×10^{-10}, mean t-value across TFs = 73.4 for AUROC and 78.8 for AUPRC). Notably, REST achieved perfect classification in this setting (Figure 2.9a). Despite these successes, the overall performance was comparable to the other regulon collections tested in this modified benchmark (Figure 2.9b).

Overall, these analyses did not reveal any systematic biases or unfair advantages in the benchmark design, supporting the reliability of the comparisons across regulon collections.

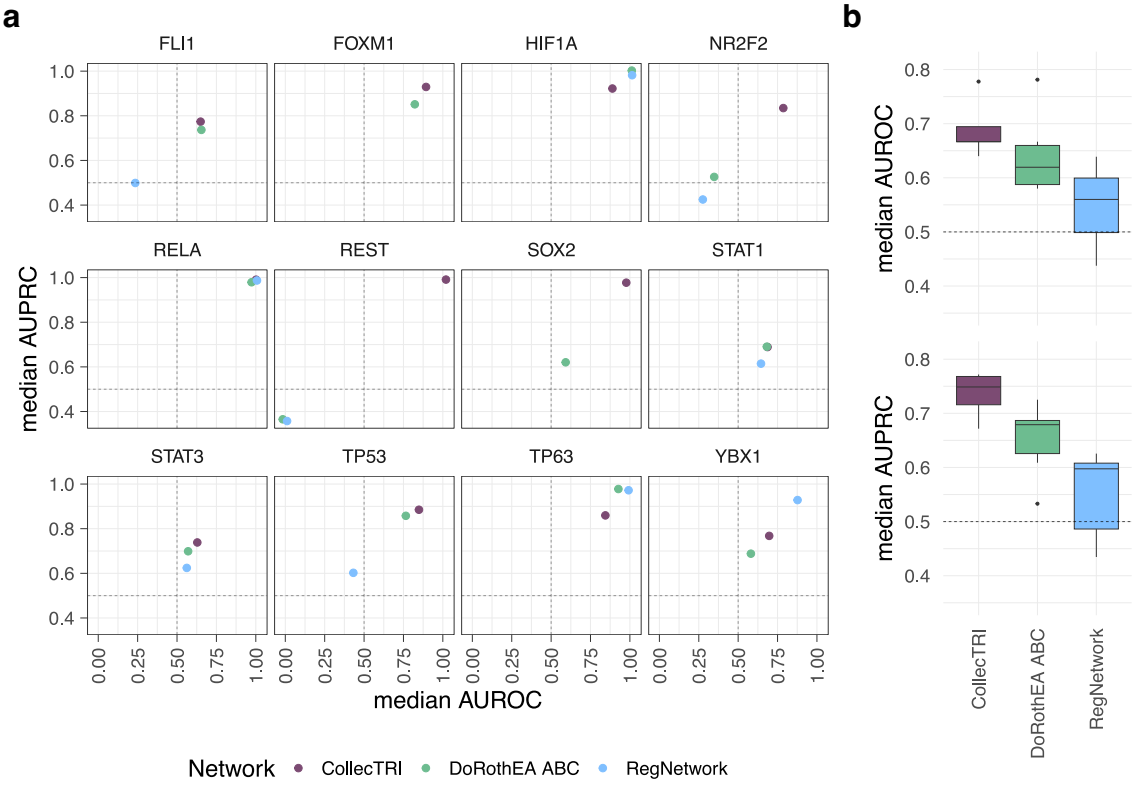

Figure 2.9. Evaluating predictive performance in identifying perturbation experiments for specific TFs. **a** Predictive performance of regulon collections in detecting perturbation experiments associated with specific TFs. Only TFs with at least five associated perturbation experiments were included in the analysis. Median AUROC and median AUPRC values are presented for each regulon collection and TF, assessing their ability to classify experiments as perturbed or non-perturbed. **b** Overall median AUROC (top) and AUPRC (bottom) values across all perturbation experiments for each regulon collection. *Reprinted from Müller-Dott et al. (2023).*

2.4 Transcription Factor Activity Estimation in Single-Cell Transcriptomics Data

TF activity estimation can provide valuable insights into cell type-specific regulatory mechanisms, particularly when compared to TF expression levels. To explore this potential, I applied the CollecTRI-derived regulons to estimate TF activities in a single-cell RNA-seq dataset of peripheral blood mononuclear cells (PBMCs). The single-cell dataset was processed using standard workflows outlined in Seurat (Hao et al., 2021). Specifically, I filtered out cells with over 5% of mitochondrial counts and unique feature counts over 2,500 or under 200d. Gene expression was then normalized by total expression, scaled to 10,000, and log-transformed. Next, I identified highly variable features (2,000) and applied a linear transformation to scale the data. I then applied principal component analysis (PCA) on the scaled data, followed by clustering using the Louvain algorithm. Cells were embedded in a K-nearest neighbor graph and visualized using a uniform manifold approximation and projection (UMAP). I then identified marker genes for each cluster using differential expression analysis which could be linked to eight distinct cell types: B cells, CD14+ monocytes, FCGR3A+ monocytes, naïve CD4+ T cells, memory CD4+ T cells, CD8+ T cells, natural killer (NK) cells, dendritic cells, and platelets, based on canonical marker gene expression (Figure 2.10a) For each cell, I then inferred TF activities using normalized gene expression counts and the CollecTRI-derived regulons. Marker TFs for each cell type were identified both through their inferred activity profiles and their expression levels. Among the 506 marker TFs identified, 93.5% were exclusively detected based on TF activity, 3.9% solely by TF expression, and 2.6% by both activity and expression (Figure 2.10b). These findings suggest that estimating TF activities can reveal cell type-specific regulatory mechanisms that are missed when relying solely on TF expression. For instance, PAX5, a key regulator of B cell development that governs their identity and function throughout B lymphopoiesis (Cobaleda et al., 2007), and EOMES, critical for NK cell maturation and functionality (Gordon et al., 2012; Kiekens et al., 2021), showed limited expression coverage in the dataset—detected in only 6.7% of B cells and 10.3% of NK cells, respectively. However, their inferred activities were consistently high across all cells of their respective types, indicating strong regulatory roles (adjusted p-value $< 2.2\times10^{-16}$, t-value = 20.7 for PAX5 and 20.2 for EOMES) (Figure 2.10c-d).

To further evaluate whether TF activities are better conserved within cell types compared to TF expression, I assessed their ability to group cells of the same type. Similarly to previously proposed (Holland et al., 2020), I calculated distance matrices for cells using

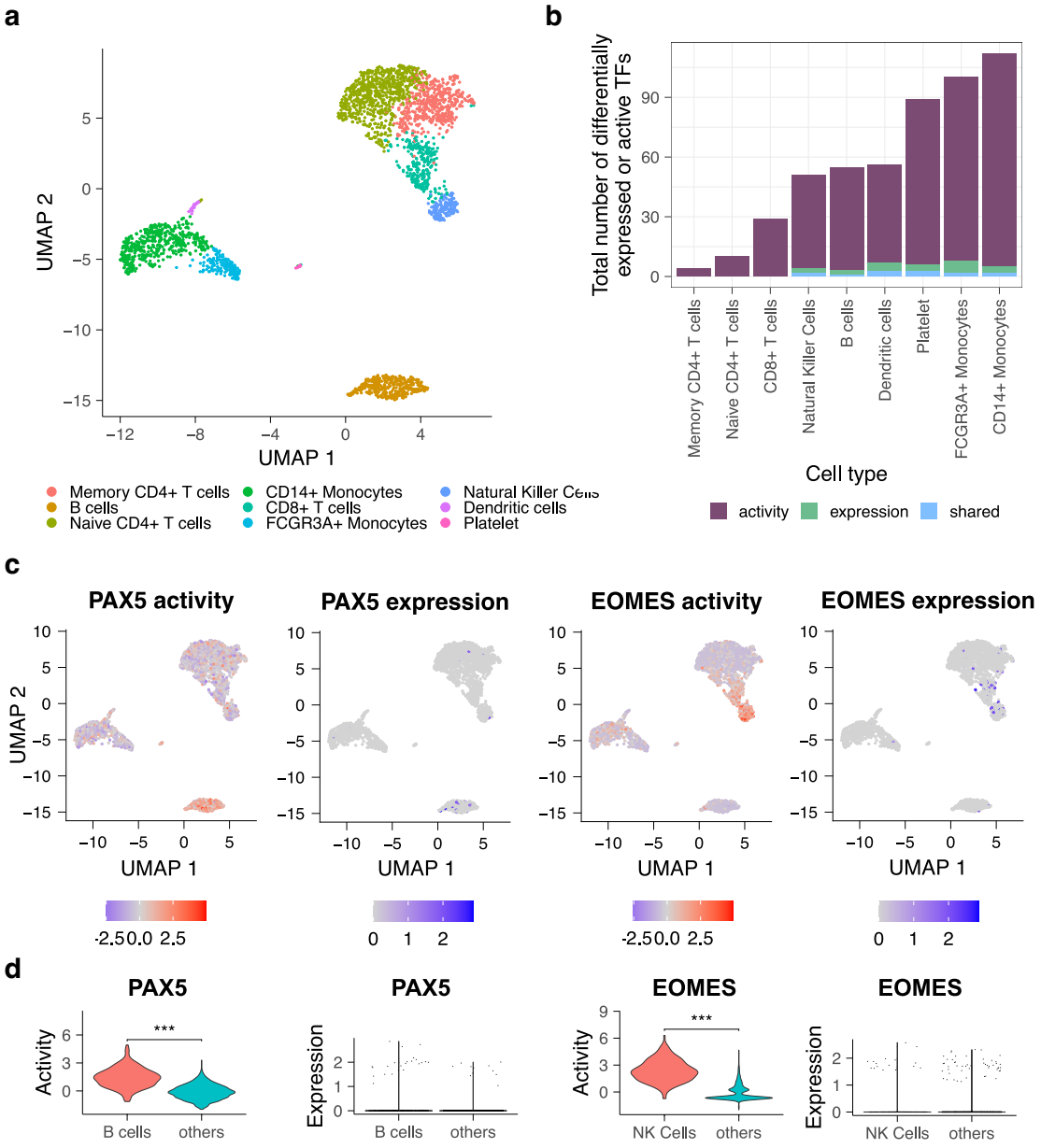

Figure 2.10. TF activity estimation in PBMCs.a uniform manifold approximation and projection (UMAP) visualization of single-cell RNA-seq data from peripheral blood mononuclear cells (PBMCs; n = 2638). Colors represent annotated cell types, including B cells, monocytes, T cells, NK cells, dendritic cells, and platelets. **b** Number of marker transcription factors (TFs) identified based on their expression or activity. **c** UMAP plots of PBMCs showing the activity and expression of PAX5 and EOMES across cells. TF activities were estimated using CollecTRI-derived regulons. **d** Comparison of the activity and expression of PAX5 and EOMES between their respective cell types (B cells for PAX5 and NK cells for EOMES) and all other cell types. Statistical testing was not performed for TF expression due to detection in fewer than 15% of cells in the corresponding cell type. *Reprinted from Müller-Dott et al. (2023).*

either TF activity or expression and computed the average silhouette width for all cells, using their cell type annotation as the reference. A higher silhouette width reflects better correspondence to annotated cell types. TF activity-based clustering produced significantly higher average silhouette widths compared to clustering by TF expression (p-value $< 2.2\times10^{-16}$, t-value = 23.8) (Supplementary Figure A.2).

In conclusion, TF activities inferred using CollecTRI-derived regulons provide more comprehensive and robust insights into the regulatory mechanisms underlying distinct cell types than TF expression alone, enabling a deeper understanding of cell type-specific regulation.

2.5 Sharing the Regulon Collection with the Community

I integrated the regulons derived from CollecTRI into the OmniPath database (Türei et al., 2021), facilitating their distribution and integration with other resources. To achieve this, I developed methods in PyPath, the database builder for OmniPath, to process the regulons. The build process standardizes gene representation using primary UniProt IDs and HGNC symbols and maps interactions to mouse and rat orthologs. Within OmniPath, I explicitly listed all variants of the ”AP1” and ”NFKB” complexes, following definitions from Bejjani et al. (2019) and Hoffmann et al. (2006). For AP1, I only included dimers formed by members of the Jun and Fos families, excluding the extended definition described in Bejjani et al. (2019). Overall, I incorporated the following information for the TF-gene interactions:

1. **TF**: Standardized to classic HGNC nomenclature, enabling users to find specific transcription factors

2. **Target gene**: Standardized to classic HGNC nomenclature, allowing users to easily search for specific genes.

3. **Mode of regulation**: Specifies whether the TF increases (activation) or decreases (repression) the expression of its target gene.

4. **Sign decision**: Indicates whether the mode of regulation (activation or repression) has been assigned to each TF-gene interaction, providing insights into regulatory dynamics.

5. **PMIDs**: Lists the supporting literature references for each TF-gene interaction, offering traceability and evidence for the assigned interactions.

6. **TF type**: Classifies transcription factors into DNA-binding TFs, co-regulators, or general initiation factors, helping users filter and interpret the regulatory roles of TFs.

With the integration into OmniPath I hoped to make the regulons easily accessible to the community through the web service at `https://omnipathdb.org/` along with other OmniPath datasets. Additionally, they can be easily accessed through the DoRothEA, decoupleR and decoupler-py packages for activity inference.

2.6 Discussion and Conclusion

In this chapter, I present a transparent and reproducible workflow for constructing transcription factor (TF) regulons by integrating TF-gene interaction data from 12 different sources, including text-mined information, manual curations, and publicly available databases. The resulting regulons offer broader coverage of TF-gene interactions compared to existing collections that primarily rely on literature-derived information. To address the limited availability of regulatory mode annotations in many public TF-gene interaction resources, I developed an evidence-driven approach to infer the mode of regulation for each TF-gene link. This approach improves TF activity inference accuracy by assigning regulatory modes to individual interactions rather than assuming uniform activation or repression based on prior TF classifications. This refinement captures the dual functionality of many TFs, which can activate or repress transcription depending on the context of regulatory elements and protein complex composition. I demonstrated that the regulons outperform other TF regulon collections in identifying perturbed TFs based on gene expression data, highlighting the high quality of the TF-gene interactions in the CollecTRI regulons. To showcase their utility, I applied these regulons to a single-cell RNA-seq dataset of peripheral blood mononuclear cells (PBMCs), identifying cell type-specific marker TFs based on their activities. This application revealed cell type-specific regulatory mechanisms that might remain undetected when relying solely on TF expression levels.

Despite the good performance of the CollecTRI-derived regulons, the benchmark dataset used for validation focuses only on a subset of TFs limiting the TF regulons actively evaluated in the benchmark. Expanding the amount of perturbation studies in the future would enable a more comprehensive evaluation of CollecTRI and other resources. Additionally, while CollecTRI regulons offer a broader coverage compared to other collections, the limited overlap between TF-gene interaction resources suggests opportunities to further expand the regulons. However, distinguishing high-quality direct interactions

from indirect or spurious relationships is challenging, especially given the reliance on literature-curated data, which may introduce biases toward well-studied TFs. Another limitation is that the regulons currently focus solely on the sign of regulation, omitting the quantitative aspects of TF-gene interactions and my attempts to incorporate TF binding weights derived from motif enrichment analysis did not improve TF activity inference. In general, the regulons currently only capture one layer of the cis-regulatory code and omit cooperative interactions between TFs and other proteins, distal regulatory interactions, and chromatin accessibility landscapes, all of which contribute to the complexity of transcriptional regulation. Additionally, the CollecTRI regulons are constructed as general TF-gene interaction networks and do not account for cell type-specific regulatory differences or TF-TF cooperativity events. Nevertheless, they provide a foundation for constructing context-specific networks using complementary data types, such as single-cell transcriptomics or chromatin accessibility data. Incorporating TF binding and proteomics data in future efforts could further enhance understanding of TF-TF regulatory interactions and context-specific mechanisms.

In conclusion, the constructed collection of TF regulons provides high coverage and confidence TF-gene interactions which are freely available to the research community. Systematic comparisons with other resources showed that CollecTRI regulons excel at recapitulating gene expression changes following TF perturbations. Furthermore, I demonstrated their applicability in biological contexts, underscoring their potential to advance the understanding of transcriptional regulation across diverse scenarios.

Chapter 3

Comprehensive Evaluation of Kinase Activity Inference from Phosphoproteomics Data

3.1 Background

Kinases play an essential role in regulating cellular processes, making their study crucial for understanding cellular function and disease mechanisms. By catalyzing the phosphorylation of threonine, serine, tyrosine, or histidine residues, kinases influence substrate proteins by modulating their activity, stability, localization, and interactions with other molecules (Manning, 2005). Dysregulated kinase activity is implicated in numerous diseases, including neurodegenerative conditions such as Alzheimer's and Parkinson's diseases (Koyano et al., 2014; Neddens et al., 2018), metabolic disorders like steatotic liver disease, obesity, and diabetes (Copps and White, 2012; Puri et al., 2008), as well as various cancer types 642011Hanahan and WeinbergHanahan and Weinberg. As a result, protein kinases are among the most targeted protein families for small-molecule therapeutics (Rodgers et al., 2018).

Advances in mass spectrometry (MS)-based phosphoproteomics have enabled large-scale profiling of global phosphorylation events, providing a detailed snapshot of cellular signaling dynamics. MS technologies can now identify and quantify up to 50,000 unique phosphopeptides, representing over 75% of cellular proteins (Sharma et al., 2014). These measurements provide a snapshot of the activity state of kinases and phosphatases. As such, phosphoproteomics can be used to infer the activity of a given kinase based on

the phosphorylation state of its targets (Dugourd and Saez-Rodriguez, 2019). To infer kinase activities, numerous computational methods have been developed with varying complexity. For instance, PTM-SEA (Krug et al., 2019) employs single-sample gene set enrichment analysis, while KSEA (Casado et al., 2013; Wiredja et al., 2017) uses a z-score to aggregate phosphorylation levels of known kinase targets relative to a background set. Despite their differences, these methods all depend on prior knowledge of kinase-substrate relationships.

Kinase target site information is typically sourced from curated databases such as PhosphoSitePlus (Hornbeck et al., 2012), SIGNOR (Lo Surdo et al., 2023), or Phospho.ELM (Dinkel et al., 2011). However, these resources cover only a small subset of measurable phosphorylation sites, often biased towards targets of well-characterized kinases (Needham et al., 2019; Savage and Zhang, 2020). This limitation poses a challenge for kinase activity inference, as reliable estimation becomes difficult when only a few substrates are available. To address this, additional targets could be included by incorporating substrates identified through large-scale *in vitro* screening assays (Mari et al., 2022; Sugiyama et al., 2019) or computational predictions, such as those generated by tools like NetworKIN (Linding et al., 2008). However, the reliability of these targets and their impact on improving predictions remain unclear. Overall, given the variety of available resources and methods for kinase activity inference, it is crucial to establish comprehensive evaluation frameworks to assess these approaches and determine the most effective strategies. In particular, evaluating diverse combinations of methods and resources is essential, extending previous comparative analyses, which have typically focused on a small subset of methods combined with a single kinase-substrate library (Hernandez-Armenta et al., 2017; Yılmaz et al., 2021).

In this chapter, I will present a systematic benchmarking strategy designed to evaluate various combinations of computational algorithms and kinase-substrate libraries for inferring kinase activities. This project was conducted in collaboration with the Zhang lab, specifically Eric J. Jaehnig from the Baylor College of Medicine, and has been published as a preprint (Müller-Dott et al., 2024) which is currently under revision in a peer-reviewed journal. I will solely focus on my contributions to the project, which include detangling prior knowledge from inference algorithms, performing kinase activity estimation across all combinations, and implementing evaluation metrics based on kinase perturbation experiments. While the manuscript also describes a complementary benchmarking approach using multi-omics data conceptualized and executed by Eric J. Jaehnig, this chapter will focus exclusively on the parts of the project for which I was responsible. The code to reproduce the results presented in this chapter can be found

here `https://github.com/saezlab/kinase_benchmark`, as well as in the package `https://benchmarkin.readthedocs.io/`.

3.2 Building an Evaluation Framework

Kinase activity inference methods are designed to identify deregulated kinases within a specific biological context, relying on phosphoproteomics data and known kinase-substrate relationships (Figure 3.1a). To evaluate the reliability of these methods, perturbation-based evaluations have been used to assess their accuracy in identifying deregulated kinases (Hernandez-Armenta et al., 2017). This approach tests whether the inferred activities of kinases align with the a priori expected deregulation of kinases. For this, a collection of perturbation experiments specifically targeting the activity of certain kinases has been previously generated (Hernandez-Armenta et al., 2017). While this collection provides a solid foundation for evaluating kinase activity inference methods, it is still limited in its coverage of kinases. To address this limitation, I extended the dataset by incorporating data from a more recent study that examined the phosphoproteomic responses of HL60 and MCF7 cells to 60 kinase inhibitors (Hijazi et al., 2020). This resulted in a collection of 212 experiments covering approximately 70 kinases where an increase or decrease of activity is expected (Figure 3.1b-c).

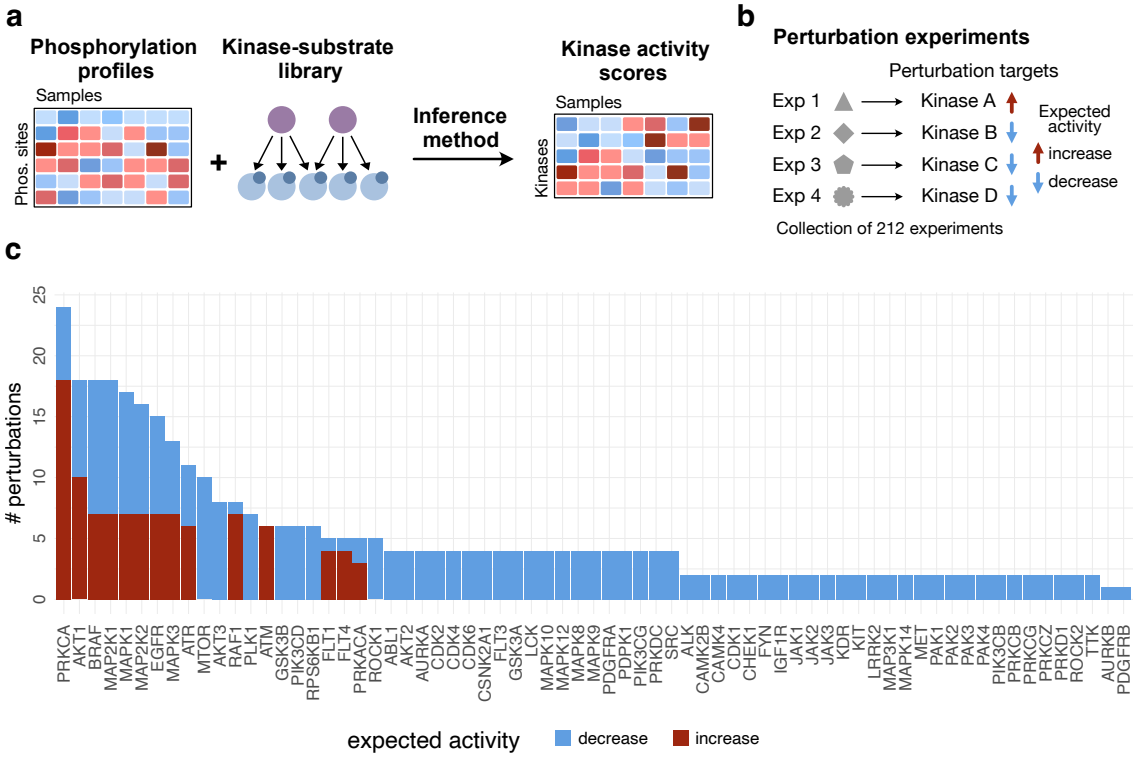

Figure 3.1. Collection of perturbation experiments for evaluating kinase activity inference. a Workflow for kinase activity inference. Phosphorylation profiles from phosphoproteomics data are combined with a kinase-substrate library to infer kinase activity scores using computational methods. The phosphorylation levels of specific sites are mapped to their corresponding kinases, enabling the estimation of kinase activities across samples. **b** Overview of perturbation experiments in the benchmark collection. For each perturbation experiment specific kinases are being targeted, resulting in an expected increase or decrease in kinase activity. **c** Number of perturbation experiments targeting each kinase, categorized by the expected change in activity.

I then implemented three metrics to evaluate the performance of kinase activity inference methods from the perturbation experiments: P_{Hit}(k), scaled rank, and area under the receiver operating characteristic curve (AUROC) (Figure 3.2). P_{Hit}(k) measures how often the perturbed kinase's activity ranks among the top k kinases in each experiment. The scaled rank assesses the rank of the perturbed kinase, adjusting for the total number of kinases by dividing the rank by the size of the inferred activity set for that experiment. Finally, AUROC scores are calculated by ranking kinases across all experiments based on their inferred activities. In this approach, true positives (TPs) correspond to the perturbed kinases, while tue negatives (TNs) include all other kinases with inferred activity in a given experiment. To address the imbalance between TPs and TNs, I subsampled the TNs 1000 times to match the number of TPs. All of these metrics assume that the highest activity change is observed in the direct target of a perturbation. However, they do not

explicitly account for potential off-target effects or downstream signaling influences.

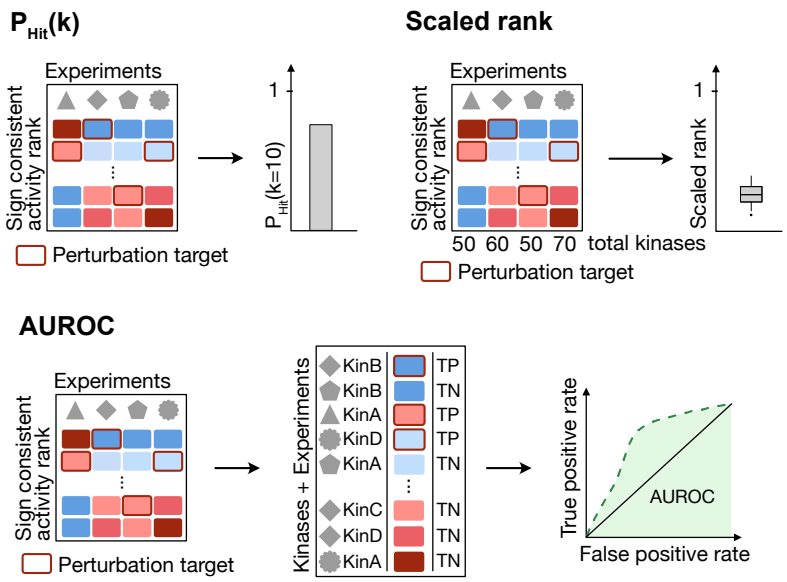

Figure 3.2. Overview evaluation metrics for kinase activity inference methods. P_{Hit}(k) measures the likelihood of the perturbed kinase being ranked among the top k kinases based on inferred activities. The scaled rank evaluates the perturbed kinase's rank relative to the total number of kinases in each experiment by dividing the rank by the total kinase count. The area under the receiver operating characteristic curve (AUROC) is determined by ranking kinases across all experiments based on their inferred activities, where perturbed kinases are considered true positives and all other kinases serve as true negatives.

3.3 Evaluation of kinase activity inference

3.3.1 Computational methods for kinase activity inference

For the evaluation, I assessed 19 different methods for predicting kinase activity scores from phosphoproteomics data, all of which rely on a set of kinase-substrate interactions. These methods include fgsea (Korotkevich et al., 2016), Fisher's exact test (Fisher), KARP (Wilkes et al., 2017), KSEA (Casado et al., 2013; Wiredja et al., 2017), the Kolmogorov-Smirnov test (KS test) (Hollander et al., 2013), the linear model implemented in RoKAI (lm RoKAI) (Yılmaz et al., 2021), the Mann-Whitney U test (MWU test), the mean, the median, a multivariate linear model (mlm) (Badia-I-Mompel et al., 2022), the normalized mean (norm mean) (Badia-I-Mompel et al., 2022), principal component analysis (PCA), PTM-SEA (Krug et al., 2019), the sum, a univariate linear model (ulm) (Badia-I-Mompel et al., 2022), the upper quantile (UQ), VIPER (Alvarez et al., 2016), the z-score as implemented by RoKAI (z-score) (Yılmaz et al., 2021), and the Chi-squared test (χ^2 test) (Table 3.1, Supplementary Table B.1). These methods differ in several aspects: whether they in-

corporate quantitative information, account for kinase promiscuity meaning the ability of sites to be phosphorylated by multiple kinases, or compute scores across multiple samples. Additionally, they can be categorized based on whether they aggregate values for a kinase's target sites or compare these values to remaining sites or an empirical null distribution.

Table 3.1. Overview of computational methods for kinase activity inference.

Method	Accounts for Magnitude	Models Kinase Promiscuity	Multi-sample Based
fgsea	Yes	No	No
Fisher	No	No	No
KARP	Yes	No	No
KSEA	Yes	No	No
Kolmogorov-Smirnov	No	No	No
Linear model - RoKAI	Yes	Yes	No
Mann-Whitney-U	No	No	No
mean	Yes	No	No
median	Yes	No	No
multivariate linear model	Yes	Yes	No
normalized mean	Yes	No	No
Principal component analysis	Yes	No	Yes
PTM-SEA	Yes	No	No
sum	Yes	No	No
univariate linear model	Yes	No	No
upper quantile	Yes	No	No
VIPER	Yes	No	No
z-score	Yes	No	No
χ^2-test	No	No	No

For each method, I inferred kinase activities based on the log fold-change of phosphorylation sites obtained from the perturbation experiments in the evaluation collection. To determine the direct targets of each kinase, I used the manually curated database PhosphoSitePlus (Hornbeck et al., 2012), a widely used resource for kinase-substrate interactions, to establish links between kinases and their downstream targets. For each experiment, I calculated kinase activities using these methods, restricting the analysis to kinases with at least five measured target phosphorylation sites. Similar to previous studies, I also investigated how the number of targets for each kinase might influence performance by using the measured number of targets in an experiment as the kinase activity (Yılmaz et al., 2021). I then applied the evaluation metrics described earlier to compare the performance of the different methods.

Across all three evaluation metrics, the z-score method demonstrated the best overall performance, achieving P_{Hit}(k) values of 0.81, 0.62, and 0.49 for k = 20, 10, and 5, respectively, an average scaled rank of 0.24, and an average AUROC of 0.79 (Figure 3.3a-c). Most methods, except for KARP, UQ, and the number of targets, achieved an average P_{Hit}(k) of at least 0.5, an average scaled rank of at least 0.3, and an average AUROC of at least 0.7. Overall, I observed strong correlations or anti-correlations between the metrics, with an absolute Pearson correlation of at least 0.9 ($p \leq 1.65 \times 10^{-7}$).

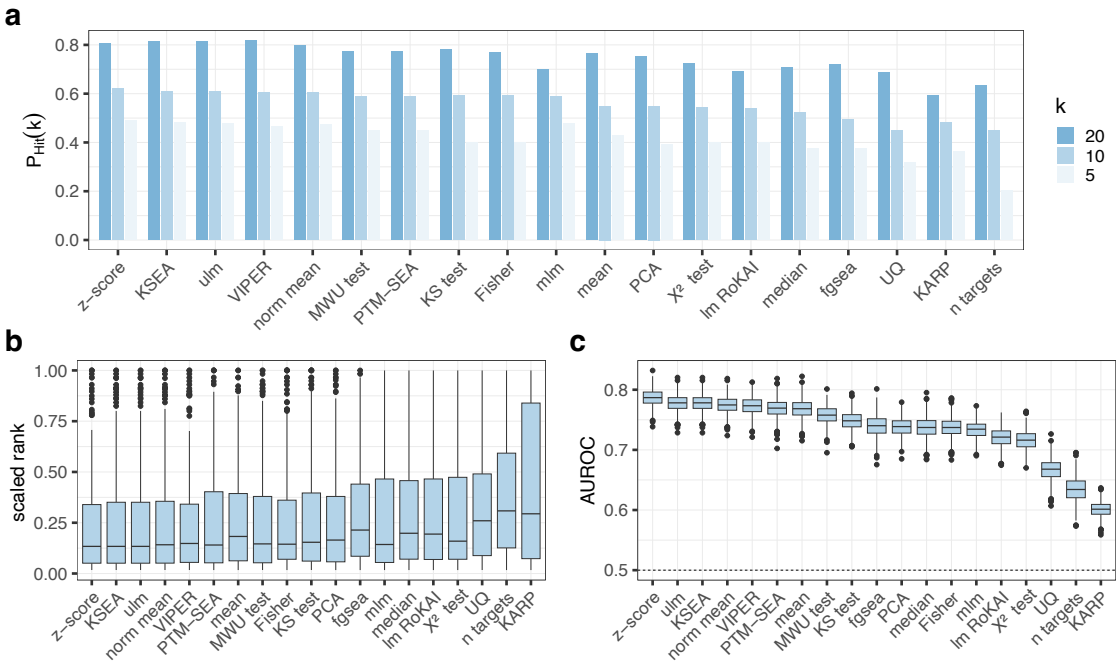

Figure 3.3. Predictive performance of kinase activity inference methods. a The probability of identifying the perturbed kinase among the top k (5, 10, 20) kinases based on inferred activities. **b** The scaled rank of the perturbed kinase activity for each experiment. **c** Area under the receiver operating characteristic (AUROC) of kinases ranked by their inferred activities across all experiments. The AUROC calculation was performed 1,000 times, each time randomly sampling a subset of the negative class equal in size to the positive class.

For the z-score, the top-performing method across metrics, I investigated whether certain kinases consistently ranked high even when they were not perturbed in an experiment, potentially introducing bias in the performance evaluation (Figure 3.4). Using the scaled rank metric, I observed that the average scaled rank of the evaluation kinases when not perturbed in an experiment was 0.52 ± 0.07. This was approximately four times higher than the scaled rank observed for the same kinases when they were perturbed, indicating that kinases in the evaluation set do not consistently rank high unless they are actually perturbed.

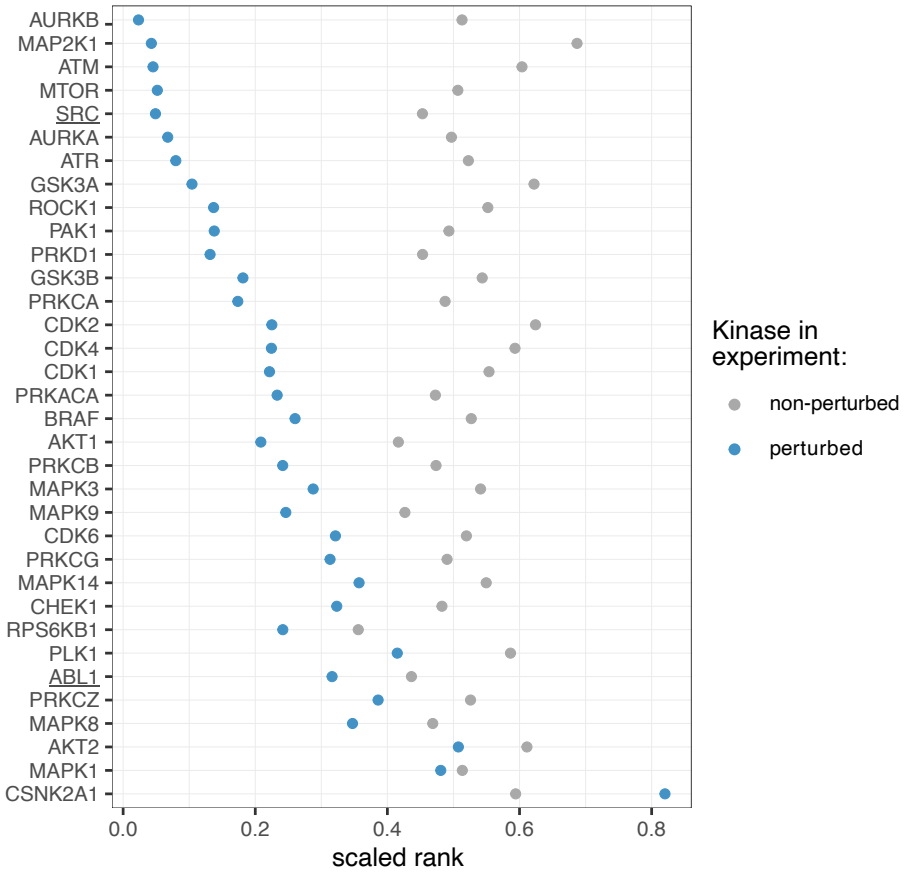

Figure 3.4. Comparison of the scaled rank of kinases in evaluation set based on perturbation status. The average scaled rank of kinases in the evaluation set is shown, categorized by whether they were perturbed or not in an experiment. Tyrosine kinases are indicated with an underline. Kinase activities were estimated using the z-score method in combination with the PhosphoSitePlus database.

Given the strong correlation observed, I chose to focus on the scaled rank as the primary evaluation metric, as it is easy to interpret and accounts for the total number of kinases, allowing for a more fair comparison between kinase-substrate libraries with differing levels of coverage as presented in the next section.

3.3.2 Comparison of Kinase-Substrate Libraries

In addition to PhosphoSitePlus, several other kinase-substrate libraries can be used for kinase activity inference. To test the impact of these libraries, I included six additional libraries and first compared them in terms of their coverage: PTMsigDB (Krug et al., 2019), GPS gold (Wang et al., 2020), OmniPath (Türei et al., 2021), iKiP-DB (Mari et al., 2022), and NetworKIN (Linding et al., 2008). This selection encompasses manually curated libraries, meta-resources, *in vitro*-based datasets, and computationally predicted interactions, providing a diverse representation of kinase-substrate relationships. PTMsigDB,

another manually curated library, provides site-specific signatures for kinases, pathways, and perturbations and incorporates information from resources like PhosphoSitePlus, NetPath (Kandasamy et al., 2010), WikiPathways (Agrawal et al., 2024), and LINCS (Keenan et al., 2018). I extracted kinase-specific signatures, excluding those from iKiP-DB, which was tested separately. GPS gold consists of a high-confidence set of kinase-substrate interactions which were used to test GPS 5.0 (Wang et al., 2020), a tool to predict kinase-substrate interactions. OmniPath represents a meta-resource which aggregates data from over 100 sources. I filtered OmniPath to retain only phosphorylation events, excluding interactions reported solely by ProtMapper (Bachman et al., 2022) or KEA3 (Kuleshov et al., 2021) libraries, due to inconsistencies. iKiP-DB reports interactions from a large-scale *in vitro* kinase study covering over 300 human kinases. Lastly, NetworKIN contains precomputed kinase-substrate interactions from the KinomeXplorer-DB (Horn et al., 2014), filtered to include only those with a NetworKIN score of at least five. I processed all of the resources into a common format, with kinases and target proteins expressed as human gene names and filtered for kinases annotated in KinHub or with the gene ontology term GO:0016301. Additional details, such as phosphorylated amino acids, their positions, and flanking sequences, were extracted where available.

I then compared the coverage of kinases across the different resources and noticed that OmniPath had the highest coverage of kinases, including 47 kinases not present in any other resource (Figure 3.5a). These unique kinases primarily originated from interactions reported by MIMP (Wagih et al., 2015) and PhosphoNetworks (Hu et al., 2014). Additionally, iKiP-DB, PhosphoSitePlus, NetworKIN, and GPS gold uniquely reported interactions for 11, 7, 2, and 1 kinases, respectively. Notably, 86.2% of all kinases were covered by at least two of the analyzed resources. Overall, all databases included serine/threonine kinases, tyrosine kinases, and dual-specificity kinases (Figure 3.5b).

Next, I compared the coverage of kinase-substrate interactions. The overlap of the interactions between resources was notably lower than the overlap in kinase coverage, with only 21.7% of interactions shared by at least two resources. iKiP-DB and NetworKIN had the lowest overlap with other resources, reporting 26,327 and 19,524 unique kinase-substrate interactions, respectively. In comparison, OmniPath, PhosphoSitePlus, PTMsigDB, and GPS gold reported 11,148, 341, 277, and 544 unique interactions, respectively (Figure 3.6a). Additionally, the median number of targets per kinase varied between resources. Manually curated resources—PhosphoSitePlus, PTMsigDB, GPS gold, and OmniPath—showed a median of 8.5 to 18 targets per kinase. In contrast, NetworKIN and iKiP-DB had a much higher median number of predicted targets per kinase, at 64 and 69, respectively (Figure 3.6b).

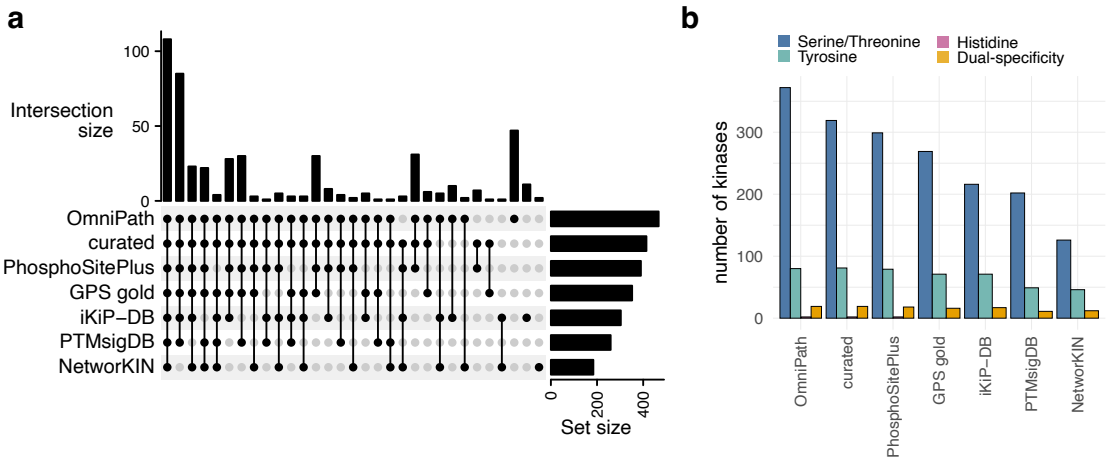

Figure 3.5. Kinase coverage across resources. **a** UpSet plot displaying the intersections of kinases across the various kinase-substrate libraries. **b** The number of kinases covered by each resource, categorized by kinase classes (serine/threonine, histidine, tyrosine, or dual-specificity). *Reprinted from Müller-Dott et al. (2024).*

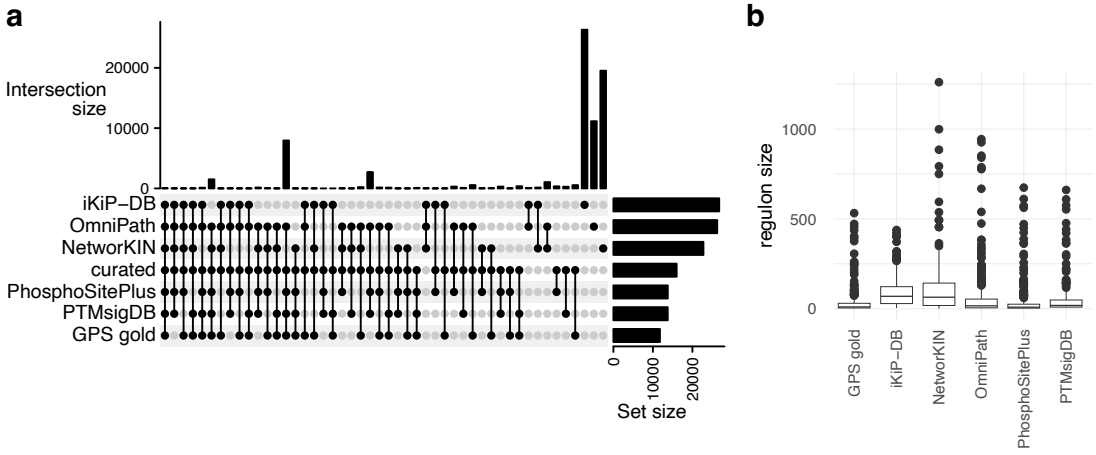

Figure 3.6. Kinase-substrate interaction coverage across resources. **a** UpSet plot showing the intersections of kinase-substrate interactions across the different kinase-substrate libraries. **b** Regulon size, representing the number of downstream phosphorylation sites assigned to each kinase in each resource. *Reprinted from Müller-Dott et al. (2024).*

Lastly, I compared the overlap of targets for each kinase between the resources by calculating the mean Jaccard index for all shared kinases between two resources. Higher Jaccard indices were observed among the curated resources, specifically PTMsigDB, GPS gold, and PhosphoSitePlus, likely due to PTMsigDB and GPS gold incorporating sites from PhosphoSitePlus. OmniPath also demonstrated moderate overlap with these curated resources, with a mean Jaccard index of 0.34. In contrast, iKiP-DB and NetworKIN

showed minimal overlap with any other resource, with their highest mean Jaccard index being just 0.03 (Figure 3.7).

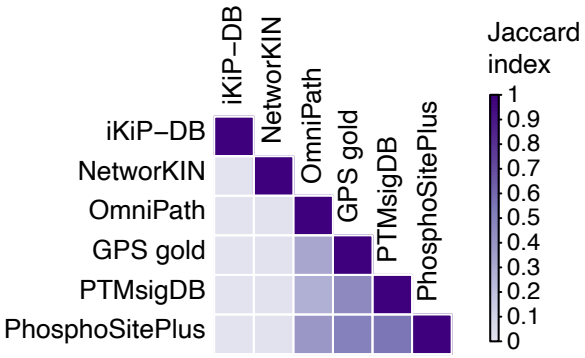

Figure 3.7. Comparison of kinase regulon overlap across resources. Mean Jaccard index of kinase regulons between kinase-substrate libraries. For all kinases shared between two libraries, the Jaccard indices of their target sets were calculated and averaged to quantify overlap. *Reprinted from Müller-Dott et al. (2024)*.

Overall, the resources showed variation in kinase coverage, with OmniPath providing the most extensive kinase representation and iKiP-DB reporting the highest number of kinase-substrate interactions. While the manually curated databases shared overlapping substrate sets, OmniPath, NetworKIN, and iKiP-DB contained significant numbers of unique substrates. These differences are likely to influence the accuracy of the predicted kinase activities derived from each resource, highlighting the importance of comparing their performance.

3.3.3 Comparison of Kinase Activity Inference across Libraries and Methods

To now evaluate the contributions of each of the different kinase-substrate libraries, I inferred kinase activity scores for the perturbation datasets by combining each library with the computational methods described previously. Due to the significant overlap among the manually curated resources, PhosphoSitePlus, PTMsigDB, and GPS gold, I created a combined resource (curated) that encompasses all interactions from these three libraries. I then first compared the inferred activity scores by assessing mean Pearson correlation coefficients, mean Spearman correlation coefficients, and the Jaccard index for the top up- and down-regulated kinases.

When comparing the results of the different computational methods across resources, most of the computational methods showed high agreement, with Pearson and Spear-

man correlation coefficients exceeding 0.77 and 0.82, respectively, in 80% of cases. For the overlap of the top 10 up- or down-regulated kinases, the average Jaccard index was 0.42, indicating that approximately six kinases were consistently identified across methods. The lowest concordance was observed for activity scores generated using the KARP score, which had Pearson correlations ranging from -0.14 to 0.03, Spearman correlations between -0.05 and 0.26, and Jaccard indices between 0.23 and 0.46 (Figure 3.8).

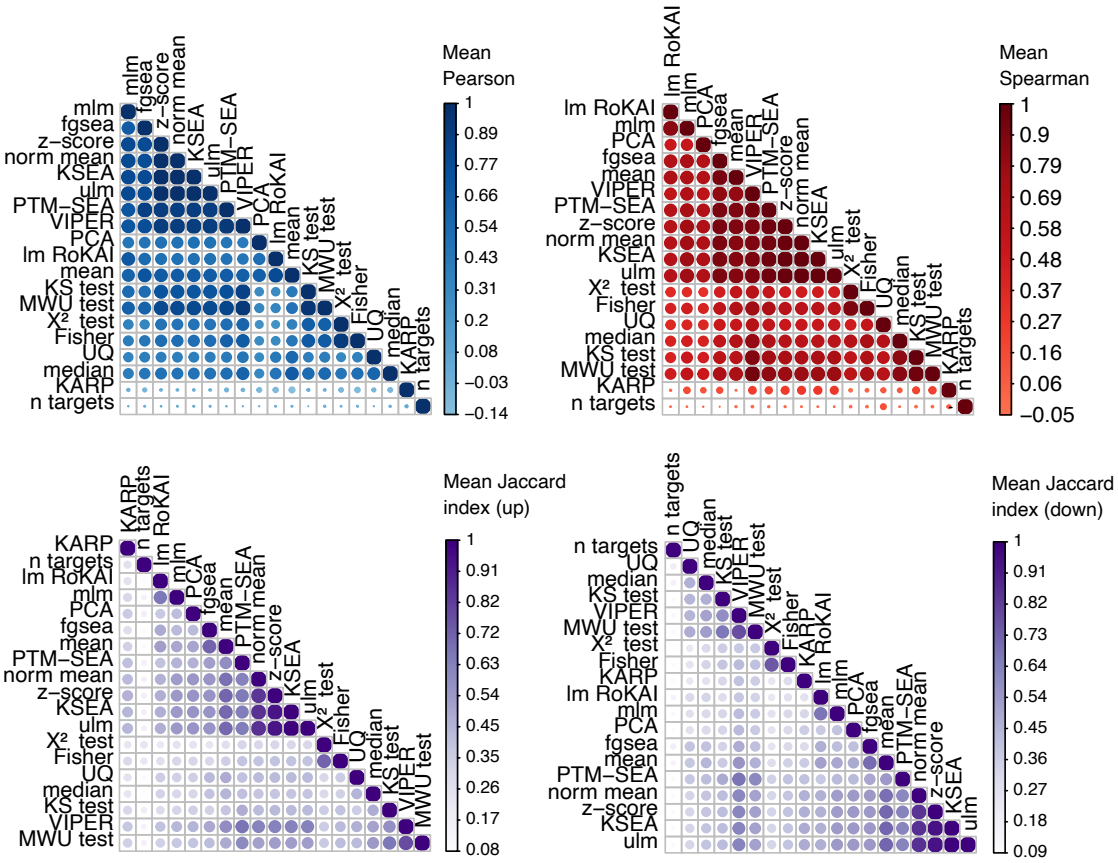

Figure 3.8. Effect of computational methods on inferred kinase activity scores. Comparison of computational methods for kinase activity inference. Based on the inferred activity scores, Pearson and Spearman correlation as well as the Jaccard index of the top 10 up- and down regulated kinases was calculated between computational methods for each kinase-substrate library and averaged across libraries. *Partially reprinted from Müller-Dott et al. (2024).*

When comparing the activity scores using different kinase-substrate libraries, I observed the highest Pearson and Spearman correlations of at least 0.88 and 0.84, respectively, among PTMsigDB, GPS gold, PhosphoSitePlus, and the curated combination. In contrast, NetworKIN and iKiP-DB showed Pearson correlations below 0.43 when compared to other libraries, which can be expected given the limited overlap in substrates between these and the other databases. Additionally, I found that the average Jaccard

index for kinase-substrate libraries across methods was 0.29, indicating that only about four of the top-scoring kinases overlapped (Figure 3.9).

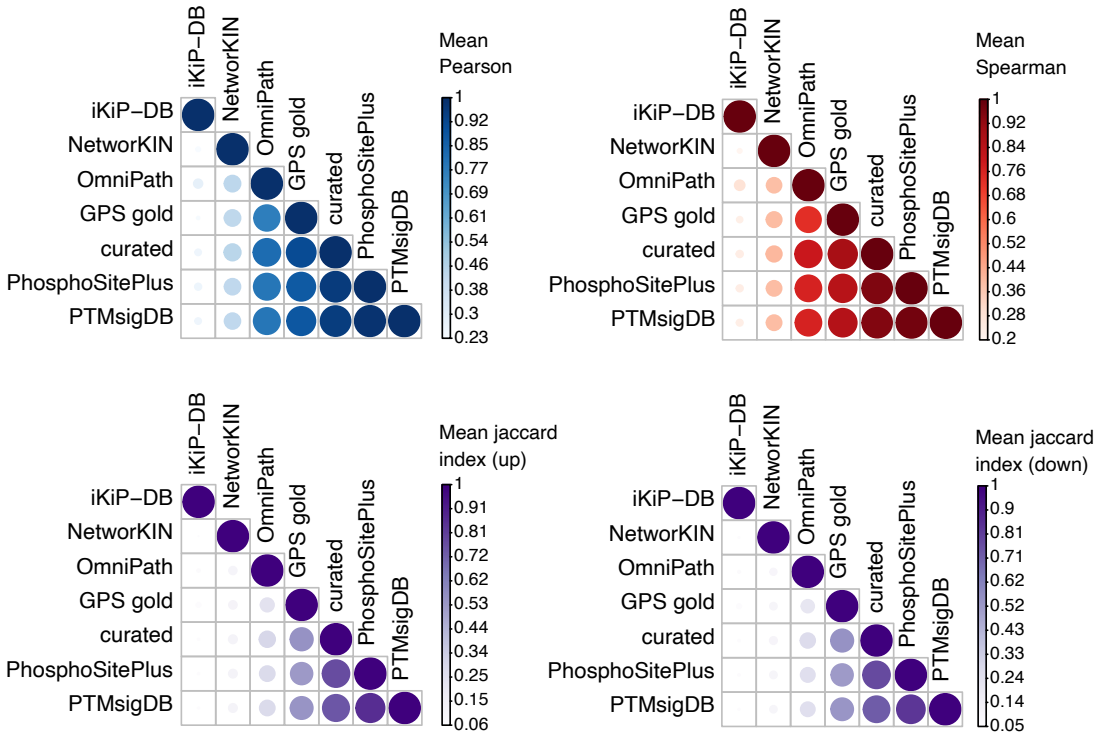

Figure 3.9. Effect of kinase-substrate libraries on inferred kinase activity scores. Comparison of kinase-substrate libraries for kinase activity inference. Based on the activity scores, Pearson and Spearman correlation as well as the Jaccard index of the top 10 up- and down regulated kinases was calculated between kinase-substrate libraries for each computational method and averaged computational methods. *Partially reprinted from Müller-Dott et al. (2024).*

To finally determine the best-performing combinations of methods and resources, I computed the scaled rank for all pairings. As previously mentioned, the scaled rank indicates the quantile in which the perturbed kinase's activity score falls, with lower values representing better performance. With that, this metric assesses how likely a perturbed kinase is to be located at the extremes of the distribution of inferred activities in each experiment, evaluating the ability of a method to assist in kinase prioritization in a real-world experiment. To provide a baseline for performance, I included a randomized kinase-substrate library, where phosphorylation sites from PhosphoSitePlus, one of the most widely used resources for kinase activity inference, were randomly reassigned to upstream kinases while preserving overlapping targets among kinases.

The results revealed distinct clusters of method-resource combinations. For methods

such as the z-score, sum, KSEA, normalized mean, PTM-SEA, and the univariate linear model, when used with PhosphoSitePlus, PTMsigDB, or OmniPath, the median scaled rank of perturbed kinases was consistently at or below the 0.24 quantile (Figure 3.10). Overall, the lowest scaled rank of 0.23 was observed with the combined curated libraries in combination with the z-score.

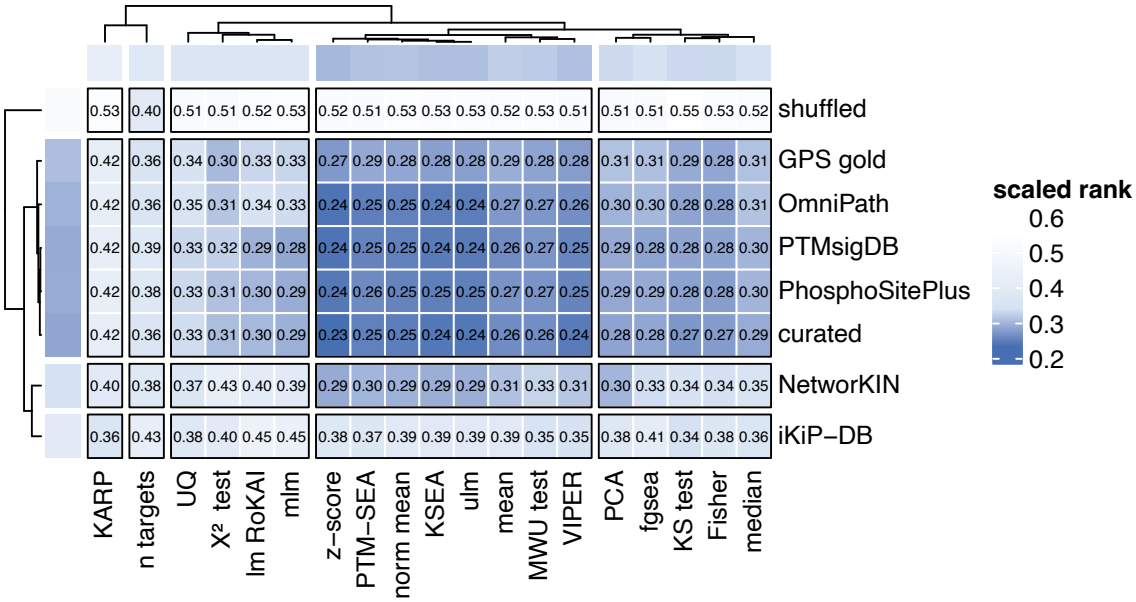

Figure 3.10. Predictive performance of computational methods and kinase-substrate libraries for activity inference. Performance of the combinations of computational methods and kinase-substrate libraries in identifying perturbed kinases from phosphoproteomics data measured by the average scaled rank. The scaled rank is calculated by ranking the kinases within an experiment based on the activities and dividing the rank of the perturbed kinase by the total number of kinases and averaging them across experiments. *Reprinted from Müller-Dott et al. (2024)*.

Lastly, I evaluated the performance of each library for kinases categorized as rich, medium, or poor, based on the number of targets identified in the curated combination (Figure 3.11). This analysis was conducted to understand whether the number of known targets influences the ability of a library to accurately infer kinase activities. Kinases classified as medium demonstrated the highest performance across libraries, with a mean scaled rank of 0.24. In contrast, a performance drop was observed for NetworKIN and iKiP-DB in the perturbation-based benchmark when evaluating poor kinases, which are likely less well-studied and have fewer known targets.

In conclusion, manually curated databases, particularly the combination of GPS gold, PTMsigDB, and PhosphoSitePlus, demonstrated the strongest performance. When stratifying kinases based on the number of targets, overall performance was higher for

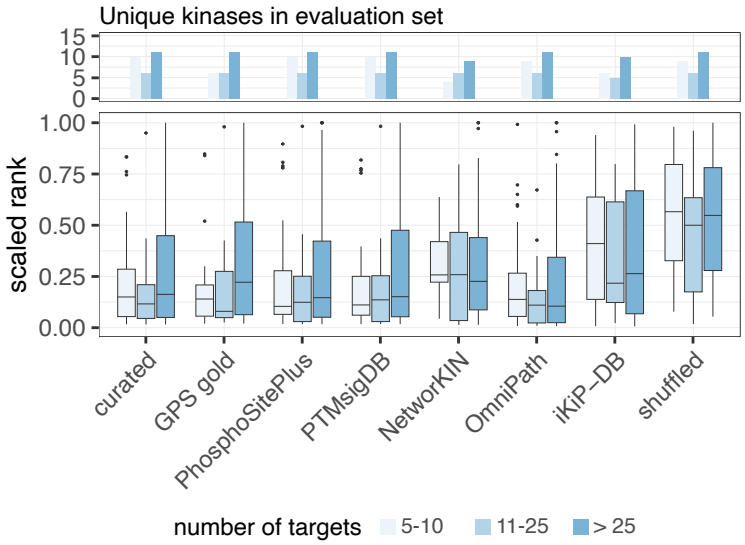

Figure 3.11. Evaluation of kinase activity inference across varying target coverage. Evaluation of kinase activity inference for kinases with 5-10, 11-25 and >25 measured targets. Kinase activity was inferred using multiple kinase-substrate libraries in combination with the z-score method and evaluated in terms of recapitulating perturbed kinases from phosphoproteomics data.

kinases with more known targets. This trend was especially evident for libraries such as NetworKIN and iKiP-DB, which may reflect a research bias toward well-studied kinases with extensive target coverage.

3.4 Adding Predicted Kinase-Substrate Interactions to Enhance Activity Inference

Next, I investigated whether the prediction accuracy could be further improved by expanding the set of kinase targets. One potential approach to enhance accuracy and increase the number of sites considered for kinase activity inference is to predict the upstream kinase of measured phosphorylation sites using tools like the large language model Phosformer (Zhou et al., 2023) or the kinase library (Johnson et al., 2023; Yaron-Barir et al., 2024). From the kinase library, kinase-substrate interactions were predicted using the percentile score for each substrate in the datasets based on the position-specific score matrices derived from the positional scanning peptide array for all Serine/Threonine and Tyrosine kinases. The highest scoring 15 kinases based on their percentile scores were then selected as upstream regulators for each phosphorylation site. Similarly, for Phosformer the highest scoring 15 kinases based on their probability score were selected as upstream regulators for each phosphorylation site.

I then compared the performance of the curated combination alone to that of the curated combination supplemented with targets predicted from the kinase library and Phosformer (Figure 3.12). Overall, incorporating predicted targets from Phosformer and the kinase library increased the number of kinases for which activity could be inferred, as well as the number of kinases included in the benchmark set. However, including these predicted targets resulted in decreased performance for both combinations (Kinase library: average scaled rank = 0.37; Phosformer: average scaled rank = 0.39). This suggests that while the predicted targets increase kinase coverage, they also introduce noise that negatively impacts the accuracy of the activity inference.

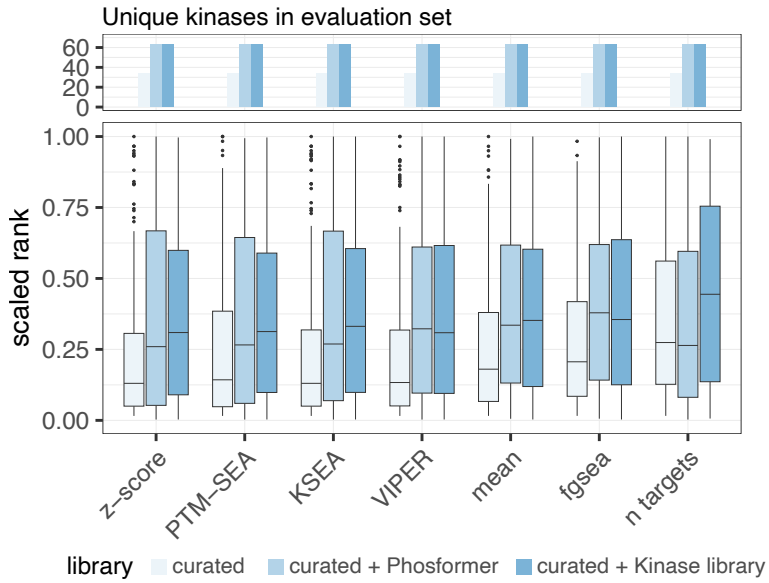

Figure 3.12. Evaluation of kinase-substrate predictions. Evaluation of kinase activity using a combination of curated kinase-substrate interactions in combination with predicted targets from the kinase library or Phosformer. Kinase activity was inferred in combination with various inference methods and evaluated in terms of recapitulating perturbed kinases from phosphoproteomics data.

Overall, increasing the number of phosphorylation sites for kinase activity inference using prediction tools did not improve the performance, making the curated combination in combination with the z-score the best performing method.

3.5 Building a Package for the Evaluation of Novel Methods

To also facilitate the evaluation of novel methods in the future, I have build the R package benchmarKIN which provides the benchmarking approach (`https://github.com/saezlab/benchmarKIN`) described in this chapter. The package includes all necessary

data and provides vignettes demonstrating how to use the benchmarking approaches for evaluating kinase activity inference. Additionally, it incorporates another benchmarking approach which is based on multi-omics tumor data that has been proposed by Eric J. Jaehnig. This approach aims to complement the perturbation-based evaluation approach, and leverages multiple omics layers to construct a gold standard set of highly active or inactive kinases using human tumor profiling data from the Clinical Proteogenomic Tumor Analysis Consortium. I hope that this package will help to simplify the process of evaluating novel methods or other kinase-substrate resources in the future.

3.6 Discussion and Conclusion

In this study, I conducted a systematic evaluation of computational methods and prior knowledge resources for kinase activity inference, focusing on identifying the most reliable approaches. By disentangling the contributions of methods and target sets, I assessed all combinations using a classical perturbation-based benchmarking framework. This revealed that simpler computational methods, such as the mean or z-score as implemented in RoKAI and KSEA, consistently performed as well as or better than more complex methods like fgsea or multivariate linear models. These findings are consistent with previous benchmarks conducted on smaller subsets of experiments (Hernandez-Armenta et al., 2017; Yılmaz et al., 2021). Additionally, manually curated target resources demonstrated the best performance in recapitulating perturbed kinases.

While the perturbation-based approach is straightforward and focuses on assessing the direct effect of the perturbation on the kinase's activity, it is limited to usually well-studied kinases that have been experimentally perturbed and profiled by phosphoproteomics. This is also demonstrated by the performance of the control metric that infers an activity solely based on the number of kinase targets. Additionally, the perturbation-based benchmark can also be confounded by downstream kinases, feedback loops, or off-target effects of perturbation drugs. While efforts have been made to better characterize the target spectrum of kinase inhibitors (Fabian et al., 2005; Klaeger et al., 2017), these are often based on binding assays, which do not always translate into changes in kinase activity. Additionally, tyrosine kinases are underrepresented in the current benchmarking dataset. This is primarily due to the limited availability of perturbation data and the scarcity of measured tyrosine phosphorylation sites, which often require specialized enrichment techniques (Engholm-Keller and Larsen, 2013). Future studies employing highly specific inhibitors, also targeting tyrosine kinases, and expanding the range of perturbation experiments would minimize these limitations. Specificity is critical because

the perturbation-based approach assumes that non-targeted kinases remain unaffected, treating them as negatives. This assumption may not hold if non-targeted kinases are indirectly influenced, reducing performance. Moreover, experimental systems lack the complexity of the tumor microenvironment, further limiting the approach.

Beyond evaluating current methods and libraries, benchmarking provides an opportunity to assess the biological relevance of predicted kinase targets. While curated libraries such as PhosphoSitePlus performed well, their coverage of kinases and phosphorylation sites remains limited, leaving much of the data generated by phosphoproteomics unused. This restricts kinase activity inference to a subset of kinases, limiting interpretability. Substrate prediction tools have been developed to address these gaps, but the tools tested here, namely Phosformer and the kinase library, did not improve performance. These tools primarily rely on sequence data and neglect critical contextual factors such as subcellular localization and regulatory interactions, limiting their accuracy. Future prediction tools that integrate modern databases, biological context, and advanced modeling are likely to improve accuracy, and the benchmarking strategy introduced here provides a robust framework to evaluate such advancements.

In conclusion, in this chapter I identified the curated target sets in combination with the z-score as implemented in RoKAI as the best-performing approach for kinase activity inference. I have made all benchmarking metrics available in the R package benchmarKIN, which includes the necessary data and vignettes to guide users through evaluating novel methods or kinase-substrate resources. This resource simplifies the benchmarking process and supports the development of improved approaches for kinase activity inference.

Chapter 4

Phosphoproteomics Based Network Contextualization to Understand Metformin's Mechanisms in Colon Cancer

4.1 Background

Colorectal cancer, being the third leading cause of cancer-related deaths worldwide, poses a major global health challenge (Colorectal Cancer Collaborators, 2022). Its development and progression are driven by complex signaling pathways that promote tumor growth, invasion, and resistance to therapy (Fearon, 2011; Koveitypour et al., 2019). Despite advances in early detection and treatment, effective therapeutic options for advanced colorectal cancer remain limited, and patient outcomes often remain poor (Arnold et al., 2017). As such, novel therapeutic strategies are needed that target the complex signaling pathways that drive tumor progression, invasion, and therapy resistance.

Metformin, a widely used antidiabetic drug, has emerged as a promising candidate in cancer prevention and treatment, including colorectal cancer (Evans et al., 2005; Meng et al., 2017). However, its mechanisms of action remain incompletely understood (Rena et al., 2017). One well-characterized mechanism involves metformin's inhibition of mitochondrial respiratory complex I, leading to 5′-AMP-activated protein kinase (AMPK) activation in an STK11 (LKB1)-dependent manner (Shaw et al., 2005). This activation suppresses mTOR signaling, leading to reduced protein synthesis, decreased cell prolif-

eration, diminished inflammation, and enhanced autophagy (Pernicova and Korbonits, 2014). Beyond the canonical AMPK pathway, metformin can activate AMPK through alternative kinases such as CAMKK2 (Hawley et al., 2005) or TAK1 (Jia et al., 2020). Moreover, it also exerts AMPK-independent effects, adding further complexity to its mechanisms of action (Stein et al., 2019).

Phosphoproteomics offers a powerful tool to study drug effects on signaling pathways by capturing its phosphorylation state. Nowadays, advances in mass spectrometry allow the measurement of around 50,000 phosphopeptides, covering over 75% of cellular proteins (Aebersold and Mann, 2016). However, the interpretation of site-specific phosphorylation data and translation into signaling networks remain challenging. To address this, network modeling approaches such as PHONEMeS have been developed, enabling the generation of contextualized networks from phosphoproteomics data by linking perturbed phosphorylation sites to upstream signaling cascades (Gjerga et al., 2021; Terfve et al., 2015). In its original implementation, PHONEMeS relied solely on unsigned kinase-substrate interactions due to the often unknown functional impact of many phosphorylation sites, limiting its ability to accurately predict downstream signaling outcomes and interpret the functional consequences of phosphorylation events.

In this chapter, I present an updated version of PHONEMeS that incorporates signed protein-protein interactions, accounting for both activating and inhibitory relationships. This updated method was applied to colon cancer cell lines treated with metformin to explore its effects on cellular signaling. This work was conducted in collaboration with the Liu lab, specifically Barbora Salovska and Erli Gao, from the Yale Cancer Biology Institute and has been published in a peer reviewed journal (Salovska et al., 2023). My contributions focused on the development and implementation of the updated network contextualization method PHONEMeS and comparing the resulting contextualized networks across different cell lines. I will explicitly refer to my work in the first person throughout this chapter. The code to reproduce the results presented in this chapter can be found here `https://github.com/saezlab/metformin_CRC`, as well as in the package `https://github.com/saezlab/PHONEMeS`.

4.2 Exploring the Phosphoproteome Response to Metformin in Colorectal Cancer Cells

To investigate the phosphoproteomic response to metformin in colorectal cancer (CRC), a set of 12 cell lines representing the heterogeneity of CRC subtypes was selected. These

cell lines were chosen based on their classification into five distinct proteomic subtypes (CPS1–CPS5), which correspond closely to clinical CRC subtypes (Roumeliotis et al., 2017). Thereby, CPS1–3 include cell lines characterized by features such as microsatellite instability, hypermutation, and ABC transporter expression, while CPS4–5 represent the colorectal stem-like subgroup (Roumeliotis et al., 2017). The selected cell lines included LoVo, RKO, and SW48 (CPS1); C2BBe1, HT115, SNU-61, SW948, and T84 (CPS2); COLO205 (CPS3); MDST8 (CPS4); and NCI-H747 and SW837 (CPS5). After 24 hours of culture, cells were treated with 10 mM metformin, and phosphoproteome measurements were taken at two time points, 30 minutes and 24 hours, to capture both early and late phospho-signaling responses. Untreated controls were also collected at the beginning of the experiment (0 hours) and after 24 hours to serve as a baseline for comparison (Figure 4.1).

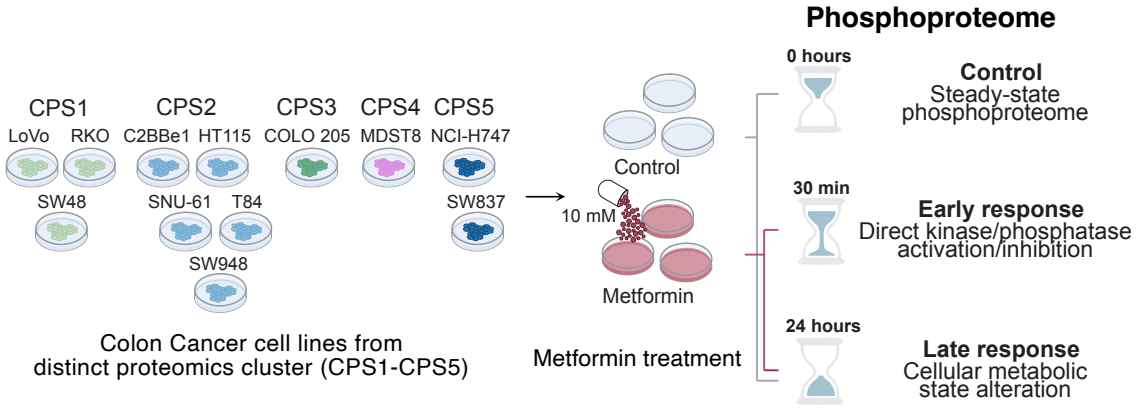

Figure 4.1. Experimental design for the phosphoproteomics analysis of metformin-treated colon cancer cell lines. Colon cancer cell lines from five distinct proteomics clusters (CPS1–CPS5) were treated with metformin (10 mM) to investigate its effects on cellular signaling. Cell lines included LoVo, RKO, and SW48 (CPS1); C2BBe1, HT115, SNU-61, T84, and SW948 (CPS2); COLO 205 (CPS3); MDST8 (CPS4); and NCI-H747 and SW837 (CPS5). Metformin-treated cells were sampled for phosphoproteomics analysis after 30 minutes (early response) and 24 hours (late response). Additionally untreated controls were collected at the start of the experiment (early control) and after 24 hours. *Adjusted from Salovska et al. (2023)*.

To investigate the effects of metformin on the phosphoproteome, I performed differential analysis using the limma R package (Ritchie et al., 2015), after data processing and normalization. Differential analysis was conducted separately for the early response (30 minutes) and the late response (24 hours), comparing metformin-treated samples to their respective controls (0 hours and 24 hours). The analysis included the standard steps of lmFit, contrasts.fit, and eBayes, to identify differentially abundant phosphorylation peptides and the resulting p-values were adjusted for multiple testing using the Benjamini-

Hochberg false discovery rate correction. I identified a distinct 'acute' response in the phosphoproteome following the 30-minute treatment, characterized by a subset of significantly regulated phosphorylation sites. However, both the number of significant phosphorylation sites and the magnitude of changes were substantially lower at 30 minutes compared to the 'late' response at 24 hours (Figure 4.2). Moreover, the acute and late responses within the cell lines were not correlated, with each time point affecting distinct sets of phosphorylation sites (Supplementary figure C.1). I also examined the regulation of AMPK and mTOR substrates, and while some sites, such as RPTOR Ser873, were also differentially abundant after 30 minutes, the majority of these phosphorylation sites was predominantly affected after 24 hours. At this later time point, thousands of significantly phosphorylated peptides ($p < 0.01$ and absolute fold change > 1.5) in 10 of the 12 cell lines were identified, representing approximately 13.8 – 41.9% of the measured phosphoproteome per cell line. Only MDST8 and NCI-H747 displayed minimal phosphorylation changes, affecting only 0.8% and 4.1% of the measured phosphoproteome, respectively. However, even among the 10 responsive cell lines, the overlap of significantly altered phosphorylation sites was limited, with only 0.08% of the measured phosphoproteome shared across cell lines.

In conclusion, most CRC cell lines exhibited significant phosphoproteomic alterations in response to metformin, especially after 24 hours. However, the degree of these changes and the specific phosphorylation sites varied greatly across the cell line panel. This variability underscores the heterogeneity in the cellular response to metformin perturbation across CRC subtypes.

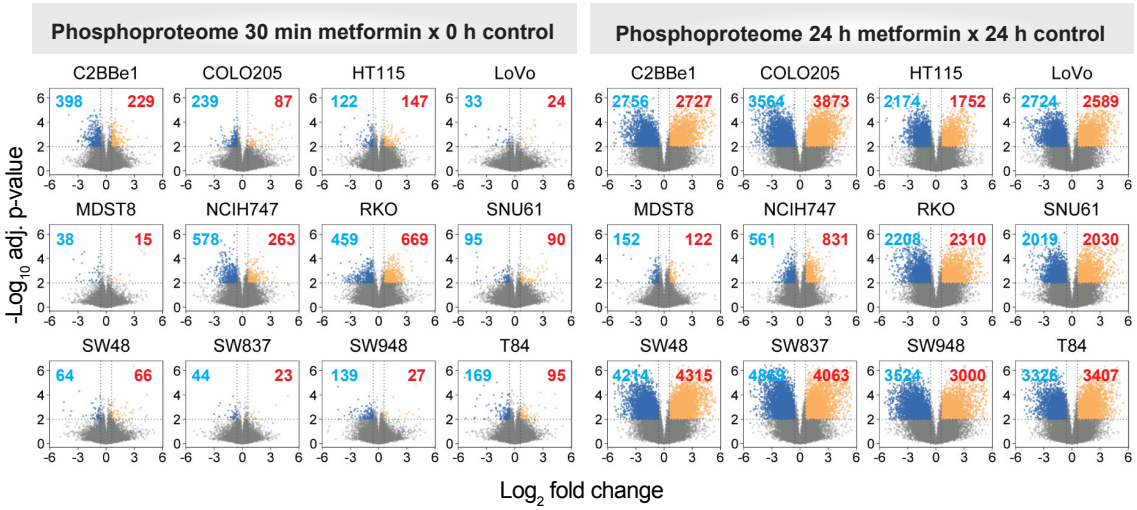

Figure 4.2. Metformin-induced phosphoproteome changes in coloncancer cell lines. The phosphoproteome of 12 colon cancer cell lines (C2BBe1, COLO205, HT115, LoVo, MDST8, NCIH747, RKO, SNU61, SW48, SW837, SW948, and T84) was analyzed 30 minutes (left panel) and 24 hours (right panel) after metformin treatment (10 mM) and compared to their respective controls (0 hour and 24 hour controls). Blue and red numbers indicate the number of significantly downregulated and upregulated phosphorylation sites, respectively (adjusted p-value < 0.01 and absolute fold change > 1.5). *Reprinted from Salovska et al. (2023)*

4.3 Uncovering Kinase Activity Profiles in Response to Metformin

To evaluate the effects of metformin on kinase activity, I used the changes in phosphorylation site abundances, as represented by t-values from the differential analysis, as input for kinase activity inference. Phosphorylation sites were assigned to kinases and phosphatases based on the OmniPath databas (Türei et al., 2021). From OmniPath, I selected only phosphorylation and dephosphorylation events and excluded interactions reported exclusively in ProtMapper (Bachman et al., 2022) due to inconsistencies. This resulted in a curated dataset of 29,445 signed kinase–phosphorylation site interactions spanning 580 kinases.

To ensure reliable kinase activity inference, I only included kinases with at least five measured targets. I then used the VIPER algorithm (Alvarez et al., 2016) implemented in the decoupleR R package (Badia-I-Mompel et al., 2022) to estimate kinase activities. I then filtered the normalized enrichment scores (NESs) generated by decoupleR to include only kinases significantly regulated in at least one cell line (p-value < 0.05). Hierarchical clustering analysis was performed on both kinases and cell lines to reveal distinct patterns

of kinase activity profiles across the cell lines (Figure 4.3a). For the cell lines, this revealed four clusters, which did not completely align with the clusters previously identified in the steady-state proteome, suggesting that the heterogeneous response to metformin is not solely determined by the basal state of the cells. Hereby, cluster 1 corresponded to RKO, SNU-61, HT115, SW948, C2bbe1, SW837, cluster 2 corresponds to LoVo, SW48, T84, cluster 3 to COLO205 and NCIH747 and cluster 4 MDST. For the kinases, one cluster (green cluster) showed a down regulation of kinase activity across cell lines and included regulators of cellular growth and proliferation, such as mTOR, CDKs, AURKA, AURKB, PLK1, and MAPKs. This cluster also encompassed tyrosine kinases like EGFR, FYN, and ABL1, as well as the dual specificity kinase DYRK1A, whose regulatory role was not apparent from phosphorylation site-level analyses alone. Additionally, a cluster of upregulated kinases (purple cluster) was observed for the cell lines, RKO, SNU-61, HT115, SW948, C2BBe1, and SW48. This group included AMPK (PRKAA1) and metabolic regulators like the insulin receptor (INSR). Furthermore, I observed an upregulation of stress-response kinases such as ATM, ATR, PRKDC (DNA-PK), and CHEK1 in these cell lines, which may explain the corresponding downregulation of CDKs. Moreover, STK11 (LKB1), a critical kinase for AMPK activation under low ATP conditions, was inferred to be activated in most cell lines (orange cluster). I additionally visualised the diversity of kinase activity responses across cell lines by generating a kinome tree using Coral (Metz et al., 2018)(Figure 4.3b). For this, kinase activity scores were mapped to branches and nodes organised based on the distinct kinase families. This revealed distinct patterns among the cell lines especially in the calcium/calmodulin-dependent protein kinase (CAMK) and the protein kinase A, G, and C (AGC) families. Interestingly, the COLO205 cell line displayed broadly reduced kinase activity, even though a global trend for phosphorylation site downregulation was not observed (Figure 4.2).

Overall, this highlights the value of kinase activity inference in phosphoproteomic analysis, as it reveals signaling changes that might not be apparent when examining phosphorylation sites alone.

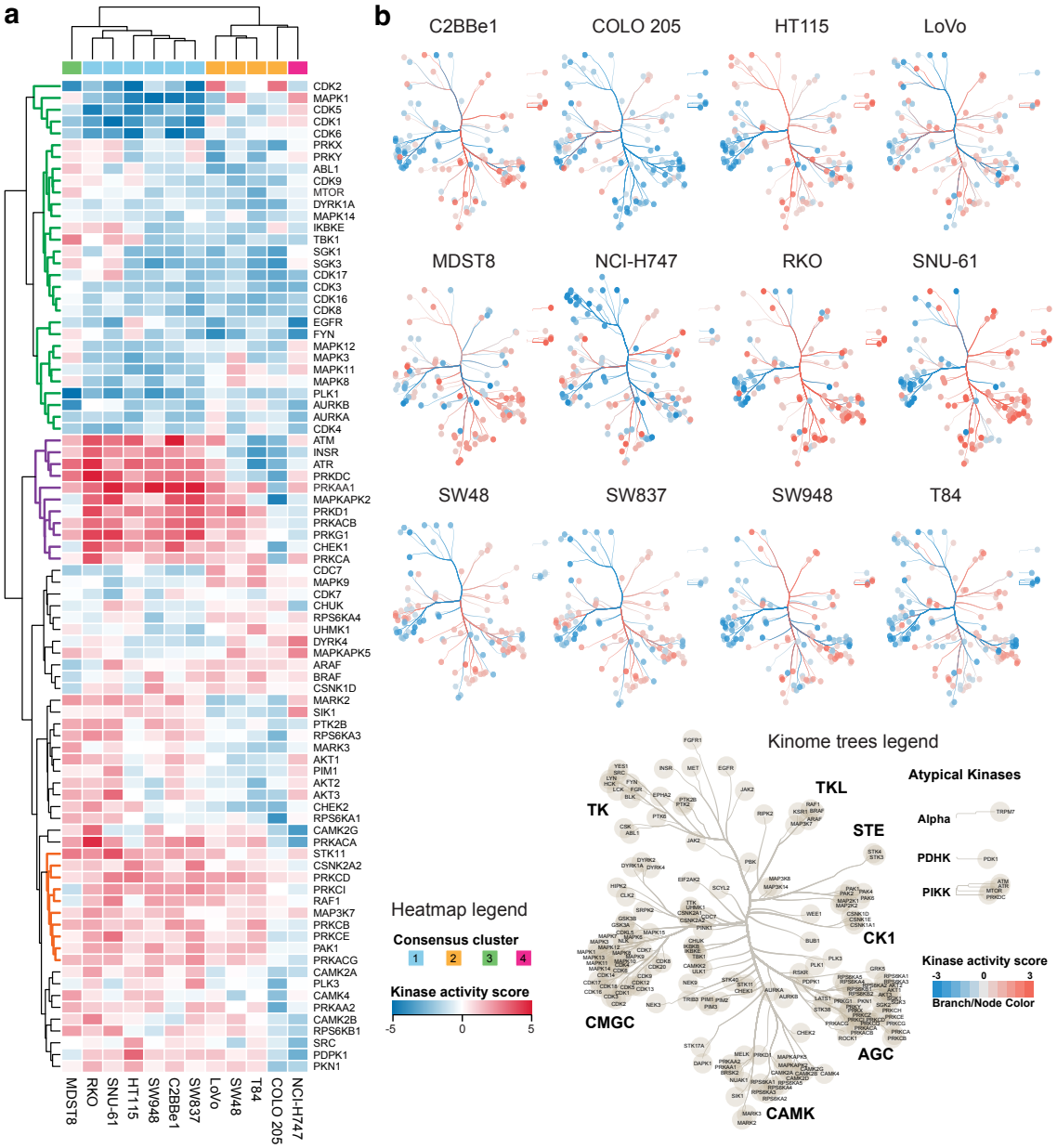

Figure 4.3. Kinase activity profiles across colorectal cancer cell lines treated with metformin. **a** Kinase activity scores for 12 colorectal cancer cell lines (C2BBe1, COLO205, HT115, LoVo, MDST8, NCIH747, RKO, SNU61, SW48, SW837, SW948, and T84), clustered based on their kinase activity profiles. Kinase activity was inferred from the t-values of the phosphorylation sites using VIPER as implemented in decoupleR. Phosphorylation sites assigned to kinases based on OmniPath. Hierarchical clustering was performed on both rows (kinases) and columns (cell lines). **b** Kinome tree for each cell line with kinase activity scores mapped to branches and nodes. These kinome trees highlight the tyrosine kinase (TK), tyrosine kinase-like (TKL), homologs of yeast Sterile 7, Sterile 11, Sterile 20 kinases (STE), casein kinase 1 (CK1), CDK, MAPK, GSK3, CLK (CMGC), protein kinase A, G, and C (AGC) and calcium/calmodulin-dependent protein kinase (CAMK) families. Kinome trees were visualised using coral. *Reprinted from Salovska et al. (2023)*

4.4 Extending the Network Model PHONEMeS to Account for Directionality of Changes

To also understand how kinases are connected within a signaling network, I developed an updated version of the network contexualization method PHONEMeS for reconstructing signaling networks from phosphoproteomics data. In contrast to the original implementation, this version incorporates protein-protein interactions and is able to account for the directionality of changes, including the upregulation or downregulation of proteins. In the original PHONEMeS implementation, which also aimed to reconstruct pathways from kinases to downstream phosphorylation sites, network contextualization relied on a prior knowledge network composed solely of kinase-substrate interactions. While this approach effectively linked kinases to deregulated phosphorylation sites, it did not consider whether these phosphorylation sites were increased or decreased in abundance. Furthermore, it lacked the ability to incorporate the functional effects of phosphorylation events, such as activation or inhibition, due to the limited availability of such information. To address these limitations, the updated PHONEMeS version incorporates signed protein-protein interactions into the prior knowledge network, enabling the inclusion of regulatory effects. In this version, kinases are linked to phosphorylation sites and connected via protein-protein interactions. Consequently, phosphorylation sites serve as endpoints in the solution networks, while their regulatory signals are propagated through the protein-protein interactions, allowing the incorporation of the direction of change. By including this information, the updated PHONEMeS offers a more precise representation of causal relationships within signaling networks.

The updated PHONEMeS is built on the causal reasoning framework of CARNIVAL, an integer linear programming (ILP)-based network contextualization method originally designed to connect upstream perturbations to the deregulation of downstream transcription factors within a prior knowledge network (Liu et al., 2019). In the updated PHONEMeS implementation, I adopted the objective function used in CARNIVAL (Equation 4.1) which optimizes the inclusion of input nodes, such as perturbed kinases and phosphorylation sites identified from experimental data, while penalizing the overall size of the network.

$$\sum_{v \in V_d} |d_v| \left[1 - \sigma_v (V_{\text{act}}(v) - V_{\text{inh}}(v))\right] + \beta \cdot \sum_{e \in E} \left(E_{\text{act}}(e) + E_{\text{inh}}(e)\right) \tag{4.1}$$

where:

V_d : Input node selection
d_v : Measurement or perturbation value for node v
σ_v : Sign (d_v)
$V_{\text{act}}(v)$: Node activation indicator for node v, $V_{\text{act}}(v) \in \{0, 1\}$
$V_{\text{inh}}(v)$: node inhibition indicator for node v, $V_{\text{inh}}(v) \in \{0, 1\}$
β : Regularisation weight
$E_{\text{act}}(e)$: Edge activation indicator for edge e, $E_{\text{act}}(v) \in \{0, 1\}$
$E_{\text{inh}}(e)$: Edge inhibition indicator for edge e, $E_{\text{inh}}(v) \in \{0, 1\}$

I also used the seven constraints implemented in CARNIVAL which among other things ensure that the resulting networks are acyclic and maintain sign-consistent paths (Supplementary Equations C.1-C.9). Additionally, I introduced a preprocessing step to reduce the solution space by pruning nodes located n steps upstream and downstream of the input phosphorylation sites and kinases. This step enhances computational efficiency while preserving biologically relevant interactions.

In addition to the methodological implementation, I incorporated additional functionalities to enhance the analysis of phosphoproteomics data in the updated PHONEMeS method. First, to address cases where the direct perturbation targets might be unknown, I included a module for selecting the perturbed nodes for the network contextualization based on kinase activity inference. This module infers kinase activities from differentially abundant phosphorylation sites, as described above, to select the most deregulated kinases. Second, to improve the interpretability of the contextualized networks generated by PHONEMeS, I added a feature to link deregulated phosphorylation sites to their corresponding proteins within the network, whenever present. For simplified visualization, I also implemented the option to exclude phosphorylation sites and focus solely on protein-protein interactions. Furthermore, targeted analysis is supported through the extraction of n-step subnetworks upstream and/or downstream of specific proteins, enabling a more detailed exploration. These functionalities allow a better investigation of signaling events derived from phosphoproteomics data.

Taken together, the updated PHONEMeS approach integrates both sign information and protein-protein interactions, offering a more comprehensive framework for network contextualization. I then applied this in the next section to investigate the mechanisms of metformin signaling in colorectal cancer cell lines.

4.5 Network Contextualization of Metformin-Induced Signaling Changes

Building on the functionalities of the updated PHONEMeS, signaling networks were reconstructed for each cell line by integrating phosphoproteomics data with prior knowledge of signaling. Given that the 24-hour time point exhibited more pronounced effects of metformin treatment, the network contextualization focused exclusively on this time point. For this, the top 15% of phosphorylation sites, ranked by absolute t-values, were selected as deregulated, and their corresponding t-values were provided as input weights for the objective function. Since the direct targets of metformin remain unknown, deregulated kinases were defined as those within the top 15% of absolute activity. Kinases with a positive activity score were classified as upregulated, while those with a negative score were classified as downregulated. In cases where AMPK (PRKAA1) ranked among the top 15%, it was manually excluded to allow for the exploration of its upstream regulation. Next, kinases with an absolute score below 0.5 were removed from the prior knowledge network, based on the assumption that their activity is not influenced by metformin. Additionally, the prior knowledge network, containing protein-protein and kinase-substrate interactions from OmniPath, was pruned by removing nodes more than 50 steps upstream or downstream of the selected kinases and phosphorylation sites, ensuring a focus on the most relevant interactions.

After initial network contextualization an iterative refinement process was implemented to ensure the coherence between PHONEMeS-inferred activity and experimental data. For nodes where the inferred activity from PHONEMeS did not align with the kinase activity estimated by decoupleR, the following adjustments were made: nodes with a score smaller than 2 were excluded from the network, while nodes with a score higher than 2 were added to the input list of deregulated kinases. PHONEMeS was re-run with the adapted inputs and the iterative process was repeated until all inferred activities from PHONEMeS and decoupleR were consistent, yielding a final network solution.

For the final networks, protein-protein interaction subnetworks were generated for each cell line by excluding kinase-phosphorylation site interactions. To investigate AMPK regulation across the cell lines, subnetworks were extracted containing all nodes and edges two steps upstream and downstream of AMPK. Next, a backbone network was created, incorporating nodes from all cell line-specific subnetworks while highlighting cell line-specific components (Figure 4.4a). The reconstructed networks revealed substantial differences in metformin response across cell lines (Figure 4.4b). Firstly, AMPK activation was not identified in MDST8, downregulated in COLO205, and activated to varying de-

grees in the other cell lines. Secondly, the upstream signals leading to AMPK activation also differed. In three cell lines, STK11 was identified as the primary upstream regulator, which typically plays a role in the context of low energy conditions (Shaw et al., 2005). In nine cell lines, CAMKK1 and CAMKK2 were detected as the upstream regulator of AMPK which can occur under conditions of calcium flux (Hawley et al., 2005). Lastly, in five cell lines MAP3K7 (TAK1) emerged as the upstream regulator, typically responsible for AMPK activation during lysosomal injury by various agents including metformin treatment (Jia et al., 2020). Downstream signaling from AMPK was equally heterogeneous with half of the cell lines exhibiting mTOR inactivation, often accompanied by inhibition of SGK1, an autophagy inhibitor, as the most plausible pathway explaining signaling outcomes in four cell lines. Other downstream effectors were less commonly shared, appearing in no more than two cell lines.

Altogether, PHONEMeS revealed diverse metformin-induced responses in colorectal cancer cell lines where context-dependent kinase activities orchestrate AMPK activation and downstream signaling.

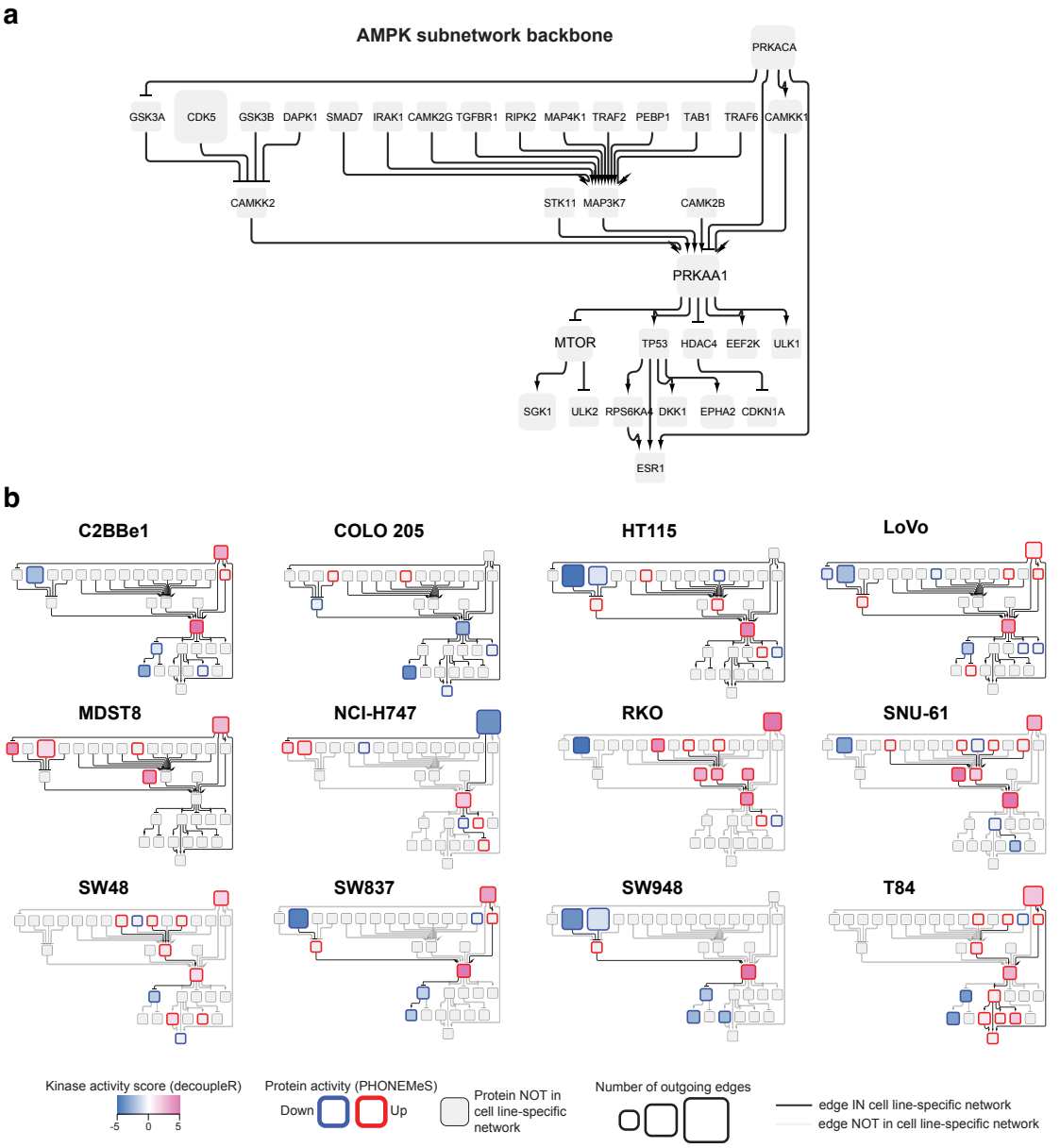

Figure 4.4. AMPK subnetworks of colorectal cancer cell lines treated with metformin. a AMPK subnetwork backbone of all interactions identified in any of the contextualised networks capturing metformin signaling of the 12 colorectal cancer cell lines (C2BBe1, COLO205, HT115, LoVo, MDST8, NCIH747, RKO, SNU61, SW48, SW837, SW948, and T84). Nodes represent proteins, and edges represent protein-protein interactions. Contextualised networks were generated by PHONEMeS through combining phosphoproteomics data with prior biological knowledge. Subnetworks around AMPK were then generated by including two nodes up- and downstream of AMPK. **b** Cell line-specific AMPK subnetworks reconstructed using PHONEMeS. Each panel represents a different colorectal cancer cell line. Nodes are colored based on their activity score inferred from decoupleR and PHONEMeS. Grey nodes represent proteins present in the backbone but not included in the specific cell line's network. *Reprinted from Salovska et al. (2023)*

4.6 Providing the Updated Network Model to the Community

Lastly, to also enable other researchers to apply the updated PHONEMeS version for network contextualization of their own phosphoproteomics data, I have developed an R package that provides all functionalities described above `https://github.com/saezlab/PHONEMeS`. The package includes the updated PHONEMeS implementation, preprocessing tools, a curated prior knowledge network, and comprehensive vignettes demonstrating how to use the method.

I hope this package will allow other researchers to effectively apply PHONEMeS to their datasets, uncover complex signaling networks, and support the interpretation of phosphoproteomics data.

4.7 Discussion and Conclusion

In this chapter, I developed and applied an updated version of PHONEMeS to contextualize signaling networks in colorectal cancer cell lines treated with metformin. This approach integrated phosphoproteomics data with prior knowledge networks, incorporating protein-protein interactions and the directional changes in the abundance and activity of phosphorylation sites and proteins, particularly kinases. Using this framework, I reconstructed context-specific signaling networks, linking deregulated kinases, selected based on their inferred activities, to their downstream targets, focusing on the 24-hour time point when metformin-induced effects were most pronounced. I then focused on the investigation of signaling events surrounding AMPK, uncovering both similarities and differences across the cell lines.

While the updated PHONEMeS is a valuable tool for constructing and investigating signed signaling networks, the method still relies on the quality and coverage of the prior knowledge network. While OmniPath is a comprehensive meta-database that integrates information from multiple sources, incomplete or inaccurate interactions within its dataset may limit the accuracy of the inferred networks. Additionally, the selection of deregulated kinases and phosphorylation sites is based on thresholds, which could introduce biases or exclude biologically relevant nodes with less pronounced changes. Lastly, the approach does not model feedback loops and only provides a static snapshot of the signaling state. Incorporating time-series data could help address these limitations by capturing dynamic signaling changes and providing insights into feedback mechanisms over time. Further-

more, integrating additional omics layers, such as transcriptomics or proteomics, could refine signal modeling by providing complementary data to better prioritize interactions and enhance the biological relevance of the networks.

In conclusion, this chapter highlights the value of PHONEMeS to uncover context-dependent signaling mechanisms by integrating phosphoproteomics data with prior biological knowledge. The updated PHONEMeS version provides a more accurate and biologically meaningful framework for investigating complex signaling pathways due to its inclusion of the directionality of activity and abundance changes. Its application to metformin-treated colorectal cancer cell lines revealed diverse and context-specific signaling dynamics, offering insights into metformin's mechanisms and potential therapeutic effects.

Chapter 5

Multi-Omics Network Contextualization to Investigate Hepatic Stellate Cell Activation

5.1 Background

Hepatic stellate cell (HSC) activation is a key driver of liver fibrosis by altering the composition of the extracellular matrix (ECM) (Friedman, 2008; Kisseleva and Brenner, 2020). These changes affect the liver architecture and contribute to the progression of liver cirrhosis, a major risk factor for the development of hepatocellular carcinoma (Breitkopf-Heinlein and Martinez-Chantar, 2024; Forner et al., 2012). While under physiological conditions, HSCs play an essential role in liver repair and vitamin A storage, dysregulation of HSC activation can result in extensive ECM production, leading to scar tissue formation that impairs liver function (Tsuchida and Friedman, 2017). Transforming growth factor beta (TGFβ) is well-established as a central mediator of HSC activation (Dooley and ten Dijke, 2012), driving their transdifferentiationinto a myofibroblast-like phenotype. However, TGFβ signal transduction in HSCs is highly context-dependent and underlies multiple levels of regulation (Dewidar et al., 2015). Recent evidence has also suggested that the growth arrest-specific 6 (GAS6) via the AXL receptor may be involved, though the precise mechanisms underlying its role in this process and its effect on TGFβ signaling remain unclear (Bárcena et al., 2015; Tutusaus et al., 2020).

Network biology has become a prominent field for deciphering the complex interactions within biological systems. In particular, network contextualization facilitates the

integration of context-specific experimental data into generalistic biological interaction networks. In recent years, numerous methods have been developed to incorporate omics data into large-scale signaling network models (Garrido-Rodriguez et al., 2022), including CARNIVAL (Liu et al., 2019), NicheNet (Browaeys et al., 2019), CausalR (Bradley and Barrett, 2017), and TieDIE (Paull et al., 2013). Nevertheless, despite the growing availability of tools for network contextualization, many approaches are not able to integrate multiple omics layers into a unified and cohesive framework. Additionally, most tools typically only perform traditional single-sample analysis and are not capable of performing joint inference across multiple conditions. To overcome these limitations, CORNETO, a framework for knowledge-driven network inference, was recently developed (Rodriguez-Mier et al., 2024). CORNETO redefines the joint inference task as a constrained optimization problem with a penalty that enforces structured sparsity, enabling simultaneous network inference across multiple samples or conditions. Moreover, it allows for the customization of cost functions and constraints to tailor the optimization problem to specific biological questions or data modalities.

In this chapter, I utilised the CORNETO framework, leveraging its ability to jointly model multiple conditions and its flexibility in designing constraints, to extend the objective function of CARNIVAL to allow the integration of multiple omics layers. Specifically, I integrated kinases as an intermediate layer to bridge upstream signaling events with downstream gene regulatory mechanisms. This extended approach was applied to investigate the roles of TGFβ and GAS6 in HSC activation, allowing the identification of potential crosstalk between these pathways. This project was carried out in collaboration with the Klingmüller lab, specifically with Elisa Holstein, from the DKFZ. My main contributions involved extending the CORNETO framework for multi-omics data integration and performing the computational analyses. I will explicitly refer to my work in the first person throughout this chapter. All code used to reproduce the results presented here is available at `https://github.com/saezlab/HSC_multiomics`. The repository is currently private but can be accessed upon request and will be made publicly available upon publication of this project.

5.2 Temporal Profiling of Hepatic Stellate Cell Activation Using Multi-Omics Data

For the investigation of the molecular mechanisms underlying hepatic stellate cell (HSC) activation, I conducted the computational analysis of time-series transcriptomics and phosphoproteomics data. In this dataset, HSCs were stimulated with growth arrest-specific 6

(GAS6), transforming growth factor beta (TGFβ), and their combination (GAS6 + TGFβ), in which the cells were pre stimulated with GAS6 for four hours followed by TGFβ stimulation, with measurements taken at multiple time points to capture phosphorylation dynamics on a short time scale (5 min, 30 min and 60 min for GAS6 and 30 min and 60 min for TGFβ) as well as transcriptomic and proteomic changes on a long time scale (up to 52 hours for proteomics and up to 28 hours for transcriptomics) and assess both individual and combined effects of these stimulations (Figure 5.1). phosphate-buffered salines (PBSs) treatment was used as a control to provide a baseline for comparison against the stimulated conditions.

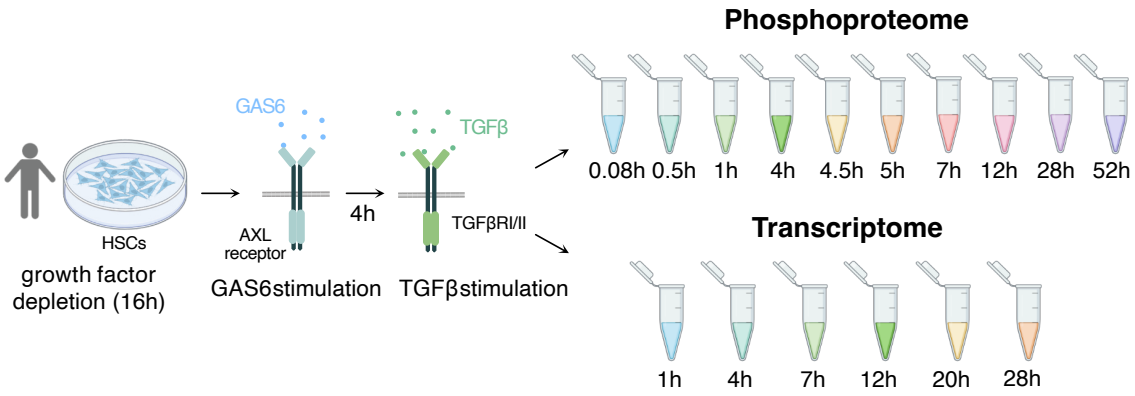

Figure 5.1. Experimental design for time-series analysis of hepatic stellate cell activation. Hepatic stellate cells (HSCs) were subjected to growth factor depletion for 16 hours prior to stimulation with either growth arrest-specific 6 (GAS6) via the AXL receptor, transforming growth factor beta (TGFβ) via TGFβ receptors I/II, or a combination of GAS6 and TGFβ. GAS6 stimulation was performed for 4 hours, after which TGFβ stimulation was initiated. For the phosphoproteome 10 time points were collected (0.08, 0.5, 1, 4, 4.5, 5, 7, 12, 28, and 52 hours), while for the transcriptome 6 time points were collected (1, 4, 7, 12, 20, and 28 hours).

Following data processing and normalization, I conducted differential analysis using the limma R package (Ritchie et al., 2015). For transcriptomics data, I applied trimmed Mean of M-values (TMM) normalization to account for differences in library sizes and RNA composition across samples (Robinson and Oshlack, 2010), while for phosphoproteomics data, I employed variance stabilizing normalization (VSN) to ensure homogeneity in variance across phosphorylation intensities (Huber et al., 2002). For the differential analysis, I performed time-matched comparisons between PBS and the treatment conditions for both transcriptomics and phosphoproteomics data. This revealed a total of 71, 951, and 1,142 deregulated phosphorylation sites (adjusted p-value ≤ 0.05 and absolute logFC ≥ 1) and 7, 3,509, and 3,577 deregulated genes for GAS6, TGFβ, and GAS6 + TGFβ, respectively (Figure 5.2, Supplementary Table D.1, D.2). GAS6 induced signal

transduction led to early change of the phosphoproteome, with the highest number of deregulated phosphorylation sites observed at 0.08 hours, and almost no effect on gene regulation, with just 7 deregulated genes identified across all time points (Supplementary Table D.1). In contrast, both TGFβ and the GAS6 + TGFβ combination exhibited a progressive increase in deregulated phosphorylation sites over time and substantially influenced gene expression, with an average of 877 and 894 deregulated genes per time point, respectively.

Overall, GAS6 primarily induced an early and transient signaling response in HSCs with minimal effect on gene expression, whereas TGFβ and the combination of GAS6 + TGFβ resulted in sustained signaling activity and substantial transcriptional changes.

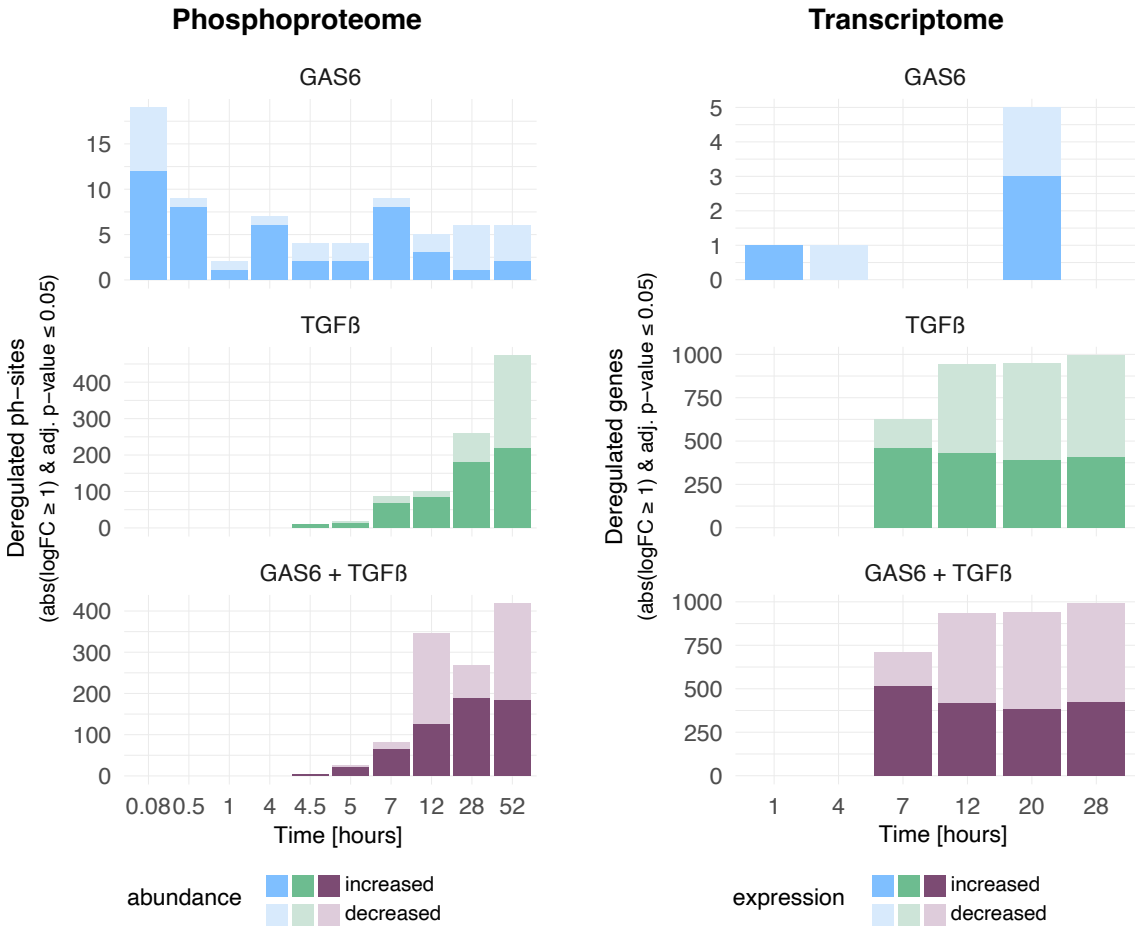

Figure 5.2. Time-course analysis of deregulated phosphorylation sites and genes across treatments. Number of significantly deregulated phosphorylation sites (ph-sites) and genes over time for each treatment condition: GAS6, TGFβ, and the combination of GAS6 + TGFβ. GAS6 stimulation was performed for 4 hours, after which TGFβ stimulation was initiated. Deregulation was determined using adjusted p-value ≤ 0.05 and absolute logFC ≥ 1.

5.3 Identification of Key Regulators Driving Hepatic Stellate Cell Activation

To identify kinases and transcription factors (TFs) affected by TGFb, GAS6 and their stimulation I inferred their activities per time point using the T-values from the differential analysis of the phosphorylation sites and genes, respectively. In line with the best-performing method for kinase activity inference identified in the previous chapter, I used a curated set of kinase-substrate interaction resources, including PhosphoSitePlus (Hornbeck et al., 2012), PTMsigDB (Liberzon et al., 2011), and the gold-standard set of GPS 5.0 (Wang et al., 2020) and inferred activities using the z-score as implemented in RoKAI (Yılmaz et al., 2021). Similarly, I estimated TF activities using the CollecTRI meta-resource introduced in the previous chapter, linking genes to their TFs, combined with the univariate linear model (ULM) (Badia-I-Mompel et al., 2022). In total I inferred an activity for 81 kinases of which 26 were significantly up- or downregulated compared to PBS in any condition at least one time point. For TFs I inferred activities for 685 TFs of which 328 were significantly up- or downregulated (adjusted p-value ≤ 0.05) in at least one point.

Besides other deregulated kinases, AKT1, AKT2, and PIM1 exhibited strong activation in response to GAS6 stimulation within the first 30 minutes (inferred activity score $\geq$ 5). In contrast, reduced activity was observed for HIPK2 and CSNK2A1 following GAS6 treatment within 30 minutes and 4 hours, respectively (inferred activity score $\leq$ -5). For TGFβ and the combination of GAS6 and TGFβ, the activity of MAPKAPK2 increased the most to an inferred activity score of 11.6 and 11.2, respectively, after 7 hours (Figure 5.3).

For TFs, a strong activation of SMAD1, SMAD2, SMAD3, and SMAD4 exhibited strong activation following TGFβ and GAS6 + TGFβ stimulation, consistently reaching activity scores above 10.7 after 7 hours of treatment. Similarly, FOXC1 showed notable activation under the same conditions. In contrast, while initial changes in transcription factor activities, such as BMAL2, SATB2, RELA, or SMAD1, were observed following GAS6 stimulation, none of these factors exhibited a consistent increase or decrease in activation over time (Figure 5.4)

Next, to evaluate the overall contribution of kinases and TFs on HSC activation, I calculated the area under the curve (AUC) for their activity scores over time, serving as a measure of the average activation. This approach avoids focusing on a single time point, which could overlook important dynamic changes. For GAS6, where most changes

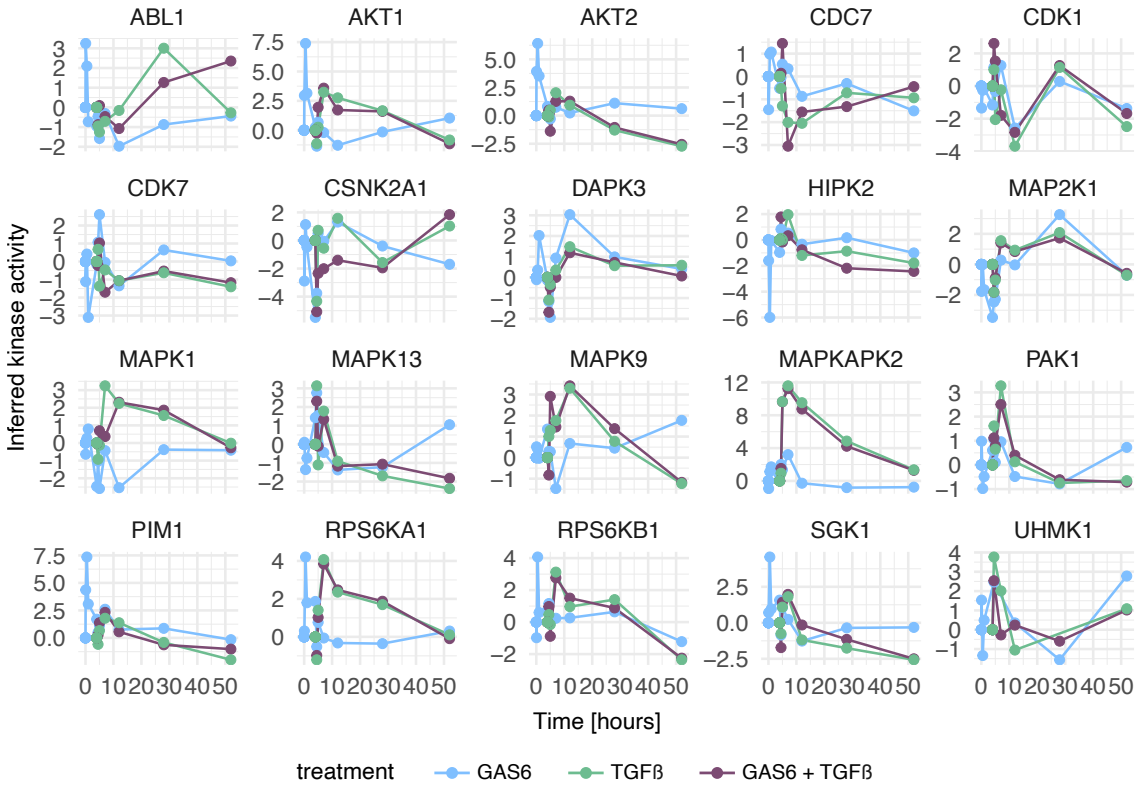

Figure 5.3. Kinase activity inference over time across treatments. Inferred kinase activity over time in response to GAS6 (blue), TGFβ (green), and the combination of GAS6 + TGFβ (purple). Kinase activities were inferred using a curated set of kinase-substrate interaction resources, including kinase-substrate interaction databases, and activities were estimated using the z-score as implemented in RoKAI. The 20 displayed kinases were selected based on the highest absolute activity score at any time point and treatment.

occurred within the first 4 hours, I restricted the AUC calculation to this early response window. In contrast, for TGFβ and the combination of GAS6 and TGFβ, I calculated AUCs over the first 28 hours, corresponding to the time frame where both transcriptomics and phosphoproteomics measurements were available. These AUC values were then used to identify key regulators for the multi-omics network model described in the next section.

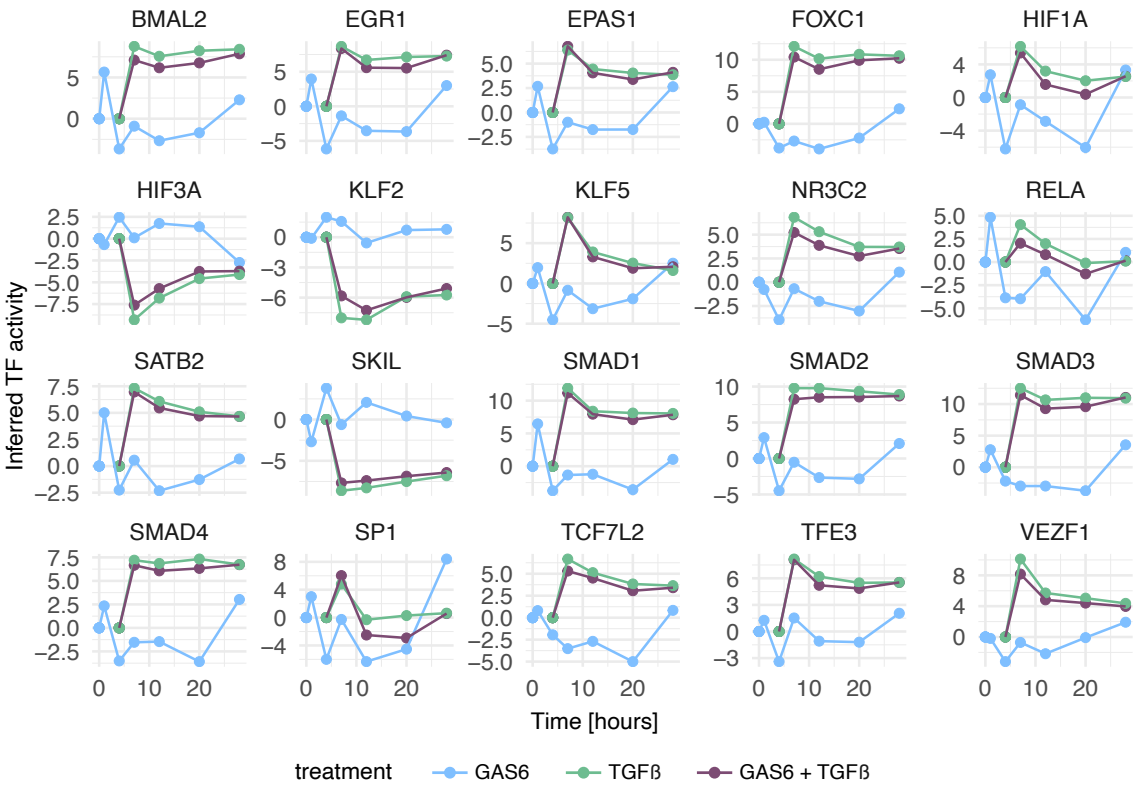

Figure 5.4. Transcription factor activity inference over time across treatments. Inferred transcription factor (TF) activity over time in response to GAS6 (blue), TGFβ (green), and the combination of GAS6 + TGFβ (purple). TF activities were inferred using CollecTRI combined with the univariate linear model implemented in decoupler. The 20 displayed TFs were selected based on the highest absolute activity score at any time point and treatment.

5.4 Enhancing Network Contextualization with Multi-Omics Integration

To explore the interplay between signaling and gene regulation, I extended the CORNETO implementation of CARNIVAL, a network contextualization method which connects upstream perturbations to deregulation of downstream transcription factors in a prior knowledge network (Liu et al., 2019). This network consists of signed and directed protein–protein interactions, where nodes represent proteins and edges capture the regulatory relationships between them. Edge signs indicated activation or inhibition, adding contextual information about the regulatory nature of each interaction. The original contextualization is formulated as an integer linear programming (ILP) problem, guided by an objective function that includes a penalty for network size and a term for incorporating deregulated TFs, which can be selected based on the experimental data (Equation 5.1).

$$\sum_{v \in V_d} |d_v| \big[1 - \sigma_v (V_{\text{act}}(v) - V_{\text{inh}}(v))\big] + \beta \cdot \sum_{e \in E} \big(E_{\text{act}}(e) + E_{\text{inh}}(e)\big) \tag{5.1}$$

where:

V_d : Input node selection
d_v : Measurement or perturbation value for node v
σ_v : Sign (d_v)
$V_{\text{act}}(v)$: Node activation indicator for node v, $V_{\text{act}}(v) \in \{0, 1\}$
$V_{\text{inh}}(v)$: node inhibition indicator for node v, $V_{\text{inh}}(v) \in \{0, 1\}$
β : Regularisation weight
$E_{\text{act}}(e)$: Edge activation indicator for edge e, $E_{\text{act}}(v) \in \{0, 1\}$
$E_{\text{inh}}(e)$: Edge inhibition indicator for edge e, $E_{\text{inh}}(v) \in \{0, 1\}$

To incorporate additional information from the multi-omics data, I included selected kinases from the experimental data as input for the contextualization process and for that extended the objective function with an additional penalty term for excluding these kinases from the solution network (Equation 5.2). Hereby, kinases were only considered in the network if a downstream effect on gene regulation could be observed.

$$\sum_{v \in V_d} |d_v| \big[1 - \sigma_v (V_{\text{act}}(v) - V_{\text{inh}}(v))\big] + \beta_k \cdot \sum_{v \in V_k} |d_v| \big[1 - \sigma_v (V_{\text{act}}(v) - V_{\text{inh}}(v))\big] + \beta \cdot R \tag{5.2}$$

where:

V_d : Input node selection
d_v : Measurement or perturbation value for node v
σ_v : Sign (d_v)
$V_{\text{act}}(v)$: Node activation indicator v, $V_{\text{act}}(v) \in \{0, 1\}$
$V_{\text{inh}}(v)$: node inhibition indicator v, $V_{\text{inh}}(v) \in \{0, 1\}$
β_k : Kinase regularisation weight
V_k : Kinase node selection
β : Regularisation weight
R : Size regularisation term $\sum_{e \in E} \big(E_{\text{act}}(e) + E_{\text{inh}}(e)\big)$

Furthermore, in addition to the original seven constraints implemented in CARNI-

VAL (Supplementary Equations C.1-C.9), I included two additional constraints to ensure kinases and TFs could only be included with their correct regulatory signs (Equation 5.3).

$$\big(V_{\text{act}}(v) - V_{\text{inh}}(v)\big) \cdot \sigma_v \geq 0 \tag{5.3}$$

where:

σ_v : Sign measurement or perturbation value for node
$V_{\text{act}}(v)$: Node activation indicator v, $V_{\text{act}}(v) \in \{0, 1\}$
$V_{\text{inh}}(v)$: node inhibition indicator v, $V_{\text{inh}}(v) \in \{0, 1\}$

This approach now enabled the integration of information from both transcriptomics and phosphoproteomics data, which I applied to investigate the effects of GAS6, TGFβ, and their combination on HSC activation in the following section. For the contextualization, I also made use of CORNETO's multi-condition contextualization feature, which prioritizes interactions shared across conditions while still accounting for condition-specific variations.

5.5 Insights into Regulatory Networks Governing Hepatic Stellate Cell Activation

Finally, to explore the regulatory mechanisms underlying HSC activation, I generated three contextualized networks to link changes in signaling to gene regulation in response to GAS6, TGFβ, and their combined stimulation. For this, I selected kinases and TFs with absolute AUCs above the 80th and 90th quantiles, respectively, to focus on the most influential regulators based on their activity over time as input for the network contextualization. Additionally, I selected AXL for GAS6, TGFBRI/II for TGFβ, and both AXL and TGFBRI/II for the combined stimulation as the initial stimuli for the networks.

I then tested 10 different β weights for the objective function, defining the penalty for the overall size of the networks: 0.01, 0.02, 0.05, 0.1, 0.2, 0.5, 1, 2, 5 and 10, corresponding to $\log_{10}(\beta)$ values of -2, -1.7, -1.3, -1, -0.7, -0.3, 0, 0.3, 0.7, and 1. A higher β weight corresponds to a higher size penalty, typically resulting in smaller networks. These weights were evaluated based on the mean percentage of kinases and TFs included and the overall network size (Figure 5.5). As expected, a higher β coefficient resulted in fewer edges and lower inclusion of kinases and TFs in the contextualised networks. The

networks stabilized in size and inclusion of kinases and TFs for a β coefficient below 0.5, which was then selected as the final parameter for the network contextualisation.

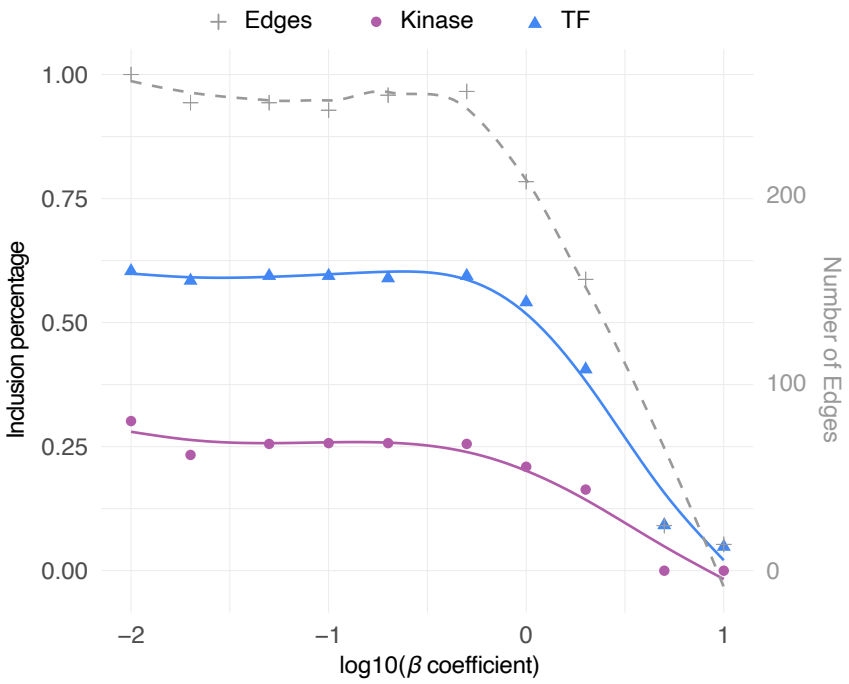

Figure 5.5. Optimization of the size penalty in the network contextualization. The effect of varying β weights ($\log_{10}$-transformed) on the percentage of kinases and transcription factors (TFs) included (left y-axis) and the total number of edges (right y-axis) in the contextualized networks.

After generating the contextualized networks, I constructed a backbone network, consolidating all interactions found across the individual solutions, to facilitate the network comparison. This backbone provided a unified view of the signaling and regulatory pathways active during HSC activation. The resulting backbone network consisted of 137 proteins, of which 40, 53, and 85 were identified in the GAS6, TGFβ, and GAS6 + TGFβ networks, respectively (Figures D.1–D.3).

Among other findings, the contextualized networks revealed a potential crosstalk between GAS6 and TGFβ signaling through PRKCA and ATF1 (Figure 5.6). PRKCA has been previously implicated in fibrotic processes and has been linked to the indirect regulation of SMAD3 and MAPK14, both of which are critical mediators of TGFβ-driven fibrosis (Giarratana et al., 2024; Xue et al., 2018). Similarly, ATF1 is associated with cellular stress responses and survival, with its downstream effects on MAPK8, also known as JNK, signaling being particularly interesting, as it plays a pivotal role in promoting fibrogenesis (Hao et al., 2024). This suggests that the combined stimulation of AXL and TGFβ may fine-tune these pathways by modulating the activity of PRKCA and ATF1, potentially shaping the fibrotic response in HSCs.

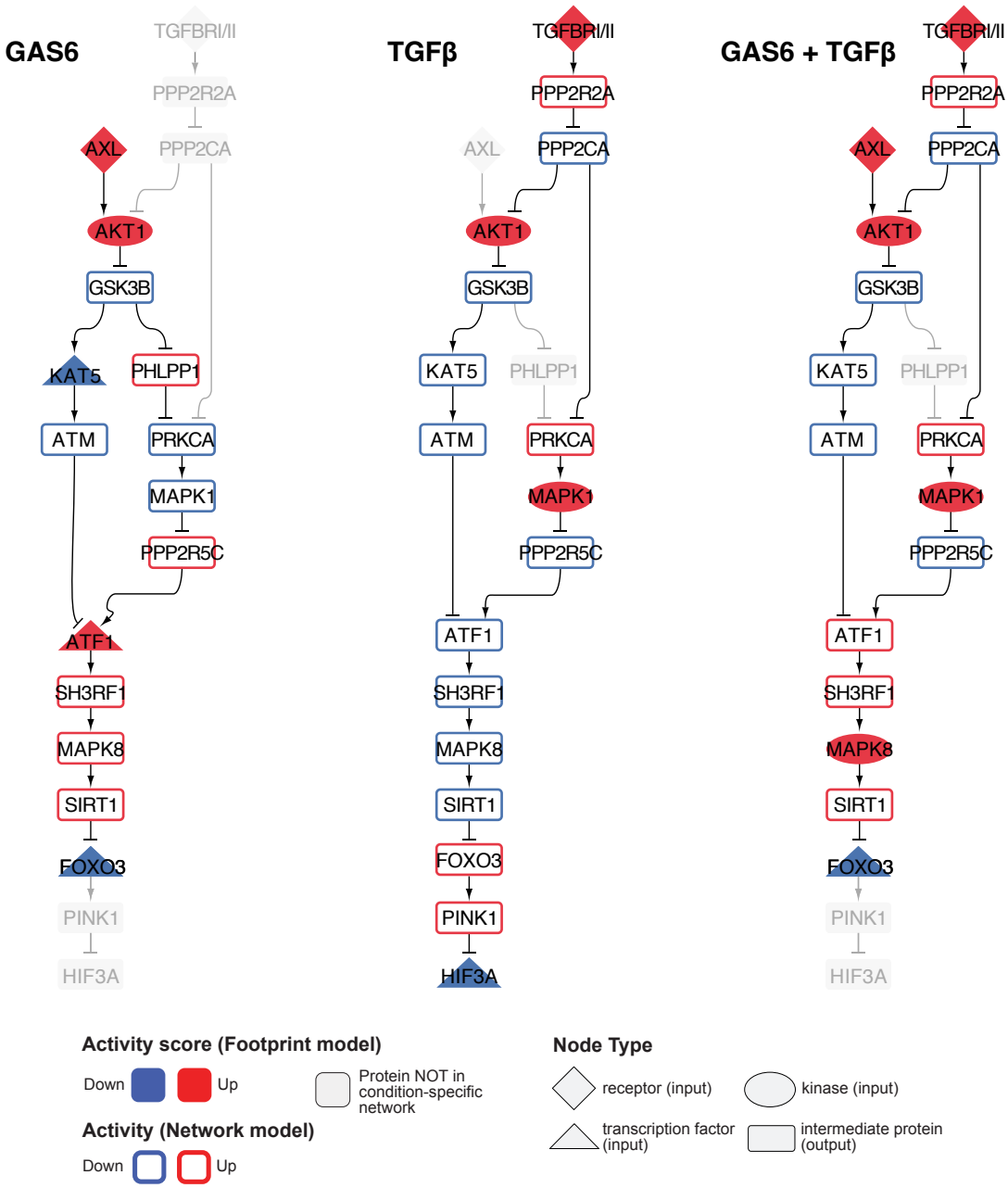

Figure 5.6. Contextualized subnetworks for GAS6, TGFβ, and GAS6 + TGFβ signaling. Contextualized networks connecting signaling events to transcriptional regulation induced by GAS6, TGFβ, and their combined stimulation. In these networks, nodes represent proteins, and arrows denote regulatory relationships. All conditions were jointly modeled, linking the targeted receptors (diamonds) stimulated by GAS6, TGFβ, or their combination to deregulated kinases (circles) and transcription factors (triangles) identified through footprint-based activity inference. The activity of intermediate proteins (rectangles) was inferred within the network model. Nodes from all conditions were integrated to construct a backbone network.

Overall, this analysis highlights the potential of network contextualization approaches to generate new hypotheses by integrating prior knowledge with context-dependent omics

data. In the context of HSC activation, the multi-omics network contextualization uncovered potential interactions and regulatory mechanisms involved in the crosstalk between GAS6 and TGFβ signaling which can also aid the prioritization of relevant regulators for further investigation.

5.6 Discussion and Conclusion

In this chapter, I presented the development and application of a multi-omics network contextualization approach to investigate the regulatory mechanisms driving hepatic stellate cell (HSC) activation. To integrate multiple omics layers into the network model, I incorporated the inclusion of kinases identified from phosphoproteomics data into the objective function. Additionally, I evaluated various β coefficients, which define the penalty for overall network size, to optimize the inclusion of key regulators and the overall network size. By applying this approach to integrate prior knowledge with context-dependent transcriptomics and phosphoproteomics data, I generated contextualized networks for GAS6, TGFβ, and their combined stimulation. These networks link upstream signaling events to downstream gene regulation, offering new hypotheses about the role of GAS6 in HSC activation and its potential crosstalk with TGFβ.

While network contextualization approaches are valuable tools for generating new biological hypotheses, they typically only provide a snapshot of the regulatory interactions. Although I attempted to incorporate the time-course of the omics data by identifying regulators across the entire time span rather than focusing on a single time point, this approach remains limited in capturing time-dependent regulation or dynamic changes in the signaling pathways. Furthermore, feedback loops, which are essential in many biological processes (Alon, 2007), cannot be modeled with this approach. In addition to time, many other elements, such as spatial dynamics or complex formation, are often simplified or disregarded when scaling up models of signaling. Nevertheless, network contextualization allows to reduce the complexity of signaling networks, producing a smaller, more focused subset of interactions. This network can then serve as a foundation for more fine-grained analyses, such as dynamic modeling, to investigate time-dependent mechanisms, feedback loops, or complex formation in greater detail.

Additionally, while incorporating prior knowledge into network inference methods can help to reduce false positive rates compared to solely data-driven approaches (Praveen and Fröhlich, 2013), the accuracy of the inferred networks depends significantly on the quality and completeness of the prior knowledge networks. As such, the choice of database used can significantly impact the network results (Mubeen et al., 2019). Furthermore,

prior knowledge is often biased toward well-characterized pathways, which may lead to the omission of interactions in less-studied areas of biology (Garrido-Rodriguez et al., 2022). As the repository of biological knowledge grows, these biases and limitations are expected to lessen, enabling more reliable and comprehensive network inferences in the future.

Lastly, while this approach is helpful for generating new hypotheses, the predicted interactions and regulatory mechanisms still need to be validated experimentally. The contextualized networks identified PRKCA and ATF1, as well as their downstream regulators MAPK8 and FOXO3, as potential mediators of crosstalk between GAS6 and TGFβ signaling. To confirm their biological relevance, targeted experiments such as knockdown or overexpression studies of these regulators should still be performed in the future.

In conclusion, this chapter highlights the value of integrating multi-omics data with network-based contextualization to investigate complex cellular processes, such as HSC activation. The incorporation of kinases into the network model and the use of multi-condition contextualization provide a robust framework for systematically exploring signaling and transcriptional regulation. The contextualized networks improve our understanding of HSC activation and the roles of GAS6 and TGFβ, offering a foundation for more detailed analyses, such as dynamic modeling.

Chapter 6

Concluding Remarks and Outlook

Cellular signaling and disease mechanisms are shaped by complex regulatory networks involving kinases and transcription factors, as well as other proteins or signaling molecules. Deciphering these processes requires high-throughput multi-omics approaches to capture diverse molecular layers, combined with computational frameworks to integrate and contextualize this data. As the scale and complexity of biological data grows, the need for interdisciplinary research connecting molecular biology and computational modeling becomes increasingly important. This thesis aims to bridge these fields by evaluating and enhancing methods for activity inference and network contextualization, leveraging benchmarking frameworks and network-based approaches to uncover regulatory mechanisms underlying cellular signaling and disease progression.

In the chapter ”Expanding the Coverage of Regulons for Accurate Estimation of Transcription Factor Activities”, I evaluated transcription factor (TF) regulon collections to improve the accurate prediction of TF activities. By leveraging an expanded set of regulons that combine TF-gene interactions from multiple databases, including text-mining-derived information, I developed a knowledge-based approach to assign regulatory modes to these interactions. I demonstrated that these signed regulons outperform existing databases in both predictive accuracy and coverage, highlighting their utility in TF activity estimation. Furthermore, I demonstrated their application in single-cell data, showing that they outperform TF expression alone in identifying marker TFs for specific cell types. Beyond single-cell applications, these regulons offer a powerful tool for uncovering the role of gene regulation in various biological contexts. Future efforts could focus on refining these regulons by incorporating cell-type or context-specific information to further improve their accuracy and better capture the complexity and dynamics of regulatory networks.

In the chapter ”Comprehensive Evaluation of Kinase Activity Inference from Phosphoproteomics Data”, I presented a comprehensive evaluation framework for kinase activity inference methods using phosphoproteomics data. While kinase activity inference is pivotal for understanding signaling pathways, its accuracy and robustness often depend on the choice of kinase-substrate libraries and the algorithms employed. To address this, I systematically benchmarked various inference methods across multiple datasets, highlighting the influence of library selection and algorithmic approaches on inferred activities. This revealed that simpler computational approaches, such as the z-score, performed as well as or better than more complex models in predicting kinase activities. Additionally, manually curated kinase-substrate libraries outperformed those based on *in vitro* experiments or computational predictions. To enable broader use, I implemented this benchmarking framework into a package, facilitating the evaluation of novel methods in the future. This benchmark could be further improved in the future by incorporating additional datasets from kinase perturbation studies, providing an even more robust foundation for assessing kinase activity inference methods.

In the chapter ”Phosphoproteomics Based Network Contextualization to Understand Metformin’s Mechanisms in Colon Cancer”, I refined the network contextualization method PHONEMeS, which predicts signaling networks by integrating phosphoproteomics data with prior knowledge resources. I extended this method to incorporate protein-protein interactions and include the regulatory effects between these interactions, allowing for the incorporation of directionality in signaling changes. To demonstrate its utility, I applied this approach to analyze the response of diverse colorectal cancer (CRC) cell lines to Metformin, identifying distinct signaling responses across cell lines. These findings underscore the potential of PHONEMeS to reveal context-specific signaling mechanisms. Future improvements could be the incorporation of functional weights or additional omics layers, further refining the accuracy and biological relevance of contextualized networks.

In the chapter ”Multi-Omics Network Contextualization to Investigate Hepatic Stellate Cell Activation”, I developed a multi-omics network contextualization approach that links signaling events to gene regulation in a cohesive manner. I first predicted the activities of key regulators, such as kinases and transcription factors (TFs), and then connected them within a prior knowledge network. This contextualization builds on the CARNIVAL method, originally designed for predicting signaling networks from transcriptomics data, which I extended by incorporating kinases as an intermediate layer in its objective function. I demonstrated this approach in the context of hepatic stellate cell (HSC) activation, using TGFβ and GAS6 stimulation to link upstream perturbations through kinases to deregulated TFs. This analysis uncovered potential crosstalk between TGFβ and GAS6

signaling pathways, which should be validated in the future. Additionally, the resulting networks provide a foundation for future advancements, such as dynamic modeling, to explore time-dependent regulatory mechanisms in greater detail.

In summary, this thesis focused on deciphering cellular signaling and disease mechanisms by integrating diverse omics data and developing new computational approaches. Kinase and transcription factor activities are key regulators of signaling pathways, yet their accurate inference depends on the quality of prior knowledge and computational frameworks. Through benchmarking, I demonstrated the value of curated resources and simpler computational methods for robust predictions. Building on this, I refined network contextualization techniques to decipher signaling events from phosphoproteomics and multi-omics data. For phosphoproteomics data, I incorporated the mode of regulation by adding directionality to protein-protein interactions, enabling more accurate modeling of signaling dynamics. For multi-omics data, I integrated kinases as an intermediate layer to link upstream signaling events with downstream gene regulation, bridging phosphoproteomic and transcriptomic information. These refinements provided deeper insights into regulatory mechanisms and facilitated the generation of context-specific networks across diverse biological conditions.

All of the approaches presented in this thesis rely heavily on prior biological knowledge. To improve these methods in the future, it is essential to obtain comprehensive and reliable information about TF-gene interactions, kinase-substrate interactions, and protein-protein interactions. Expanding these datasets will not only enhance the accuracy and applicability of computational methods but also address current biases toward well-studied proteins and pathways. Special attention should be given to understudied proteins and areas like the ”dark phosphoproteome” (Needham et al., 2019), which remain largely unexplored and could hold critical regulatory insight. Moreover, the integration of emerging technologies and data types, such as single-cell or spatial omics data, could further refine these methods and expand their scope. By contextualizing regulatory networks at higher resolution and in specific cellular environments, these approaches could provide even more actionable insights into disease mechanisms and cellular processes. Lastly, advancing these approaches requires a close collaboration between biologists and computational scientists. Close interdisciplinary interactions will be crucial for developing methods that are biologically meaningful and for interpreting the results within their proper experimental and biological context.

Appendix A

Expanding the Coverage of Regulons for Accurate Estimation of Transcription Factor Activities

A.1 Supplementary Figures

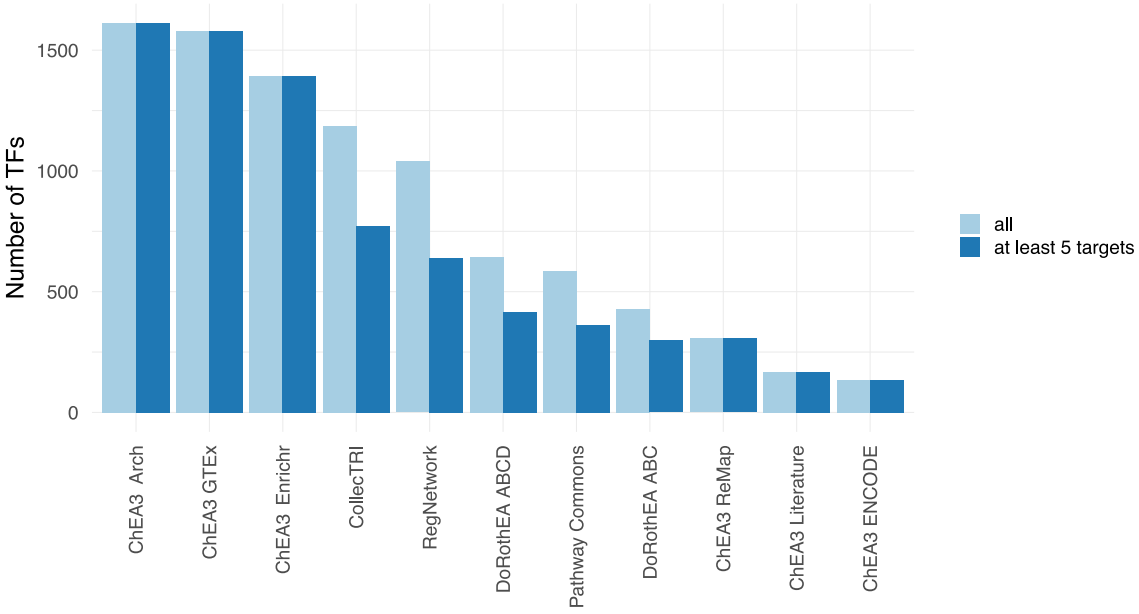

Figure A.1. TF coverage across resources. The total number of transcription factors (TFs) and the subset of TFs with at least five target genes identified within each resource.*Reprinted from Müller-Dott et al. (2023)*

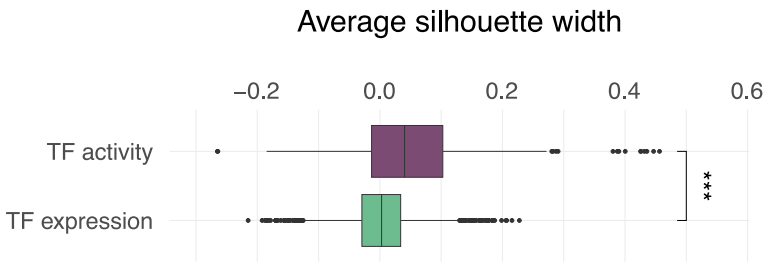

Figure A.2. Silhouette comparison between TF activity and expression. Average silhouette width of cell clusters based on transcription factor (TF) activity and expression. Higher silhouette widths indicate better alignment of clusters with the annotated cell types. *Reprinted from Müller-Dott et al. (2023)*

A.2 Supplementary Tables

Table A.1. AUROC comparison across TF regulon collections. Area under the receiver operating characteristic curve (AUROC) values were compared between different transcription factor (TF) regulon collections using a two-sample t-test. AUROC values represent the performance in recapitulating perturbed TFs from transcriptomics data. P-values from the t-test were adjusted using Benjamini-Hochberg.

GRN 1	GRN 2	AUROC Adj. P Value	AUROC T Value
ChEA3 ARCHS4	ChEA3 ENCODE	$<2.2x10^{-16}$	111.2
ChEA3 ARCHS4	ChEA3 ENCODE	$<2.2x10^{-16}$	111.2
ChEA3 ARCHS4	ChEA3 Enrichr	$<2.2x10^{-16}$	114.5
ChEA3 ARCHS4	ChEA3 GTEx	$<2.2x10^{-16}$	102.5
ChEA3 ARCHS4	ChEA3 Literature	$<2.2x10^{-16}$	121.2
ChEA3 ARCHS4	ChEA3 ReMap	$<2.2x10^{-16}$	118.2
ChEA3 ARCHS4	Shuffled CollecTRI	$<2.2x10^{-16}$	78.7
ChEA3 Enrichr	ChEA3 ENCODE	$<2.2x10^{-16}$	19.8
ChEA3 Enrichr	ChEA3 Literature	$<2.2x10^{-16}$	21.2
ChEA3 Enrichr	ChEA3 ReMap	$<2.2x10^{-16}$	12.8
ChEA3 GTEx	ChEA3 ENCODE	$<2.2x10^{-16}$	26.5
ChEA3 GTEx	ChEA3 Enrichr	$<2.2x10^{-16}$	8.6
ChEA3 GTEx	ChEA3 Literature	$<2.2x10^{-16}$	28.5
ChEA3 GTEx	ChEA3 ReMap	$<2.2x10^{-16}$	20.5
ChEA3 Literature	ChEA3 ENCODE	0.97	0.1
ChEA3 ReMap	ChEA3 ENCODE	$<2.2x10^{-16}$	8.1
ChEA3 ReMap	ChEA3 Literature	$<2.2x10^{-16}$	8.5
CollecTRI	ChEA3 ARCHS4	$<2.2x10^{-16}$	271.9
CollecTRI	ChEA3 ENCODE	$<2.2x10^{-16}$	300.7
CollecTRI	ChEA3 Enrichr	$<2.2x10^{-16}$	364.6
CollecTRI	ChEA3 GTEx	$<2.2x10^{-16}$	344.6
CollecTRI	ChEA3 Literature	$<2.2x10^{-16}$	331.2
CollecTRI	ChEA3 ReMap	$<2.2x10^{-16}$	342.5
CollecTRI	DoRothEA ABC	$<2.2x10^{-16}$	109.9
CollecTRI	DoRothEA ABCD	$<2.2x10^{-16}$	185.6
CollecTRI	Pathway Commmons	$<2.2x10^{-16}$	218.4

Continued on the next page...

Continued from previous page.

GRN 1	GRN 2	AUROC Adj. P Value	AUROC T Value
CollecTRI	RegNetwork	$<2.2x10^{-16}$	193.8
CollecTRI	Shuffled CollecTRI	$<2.2x10^{-16}$	327.1
DoRothEA ABC	ChEA3 ARCHS4	$<2.2x10^{-16}$	149.1
DoRothEA ABC	ChEA3 ENCODE	$<2.2x10^{-16}$	217.4
DoRothEA ABC	ChEA3 Enrichr	$<2.2x10^{-16}$	249.7
DoRothEA ABC	ChEA3 GTEx	$<2.2x10^{-16}$	234.4
DoRothEA ABC	ChEA3 Literature	$<2.2x10^{-16}$	237.3
DoRothEA ABC	ChEA3 ReMap	$<2.2x10^{-16}$	241.2
DoRothEA ABC	DoRothEA ABCD	$<2.2x10^{-16}$	71.2
DoRothEA ABC	Pathway Commmons	$<2.2x10^{-16}$	109.7
DoRothEA ABC	RegNetwork	$<2.2x10^{-16}$	73.3
DoRothEA ABC	Shuffled CollecTRI	$<2.2x10^{-16}$	214.2
DoRothEA ABCD	ChEA3 ARCHS4	$<2.2x10^{-16}$	76.2
DoRothEA ABCD	ChEA3 ENCODE	$<2.2x10^{-16}$	165.2
DoRothEA ABCD	ChEA3 Enrichr	$<2.2x10^{-16}$	182.6
DoRothEA ABCD	ChEA3 GTEx	$<2.2x10^{-16}$	169.1
DoRothEA ABCD	ChEA3 Literature	$<2.2x10^{-16}$	180
DoRothEA ABCD	ChEA3 ReMap	$<2.2x10^{-16}$	180.4
DoRothEA ABCD	Pathway Commmons	$<2.2x10^{-16}$	42.4
DoRothEA ABCD	Shuffled CollecTRI	$<2.2x10^{-16}$	147.3
Pathway Commmons	ChEA3 ARCHS4	$<2.2x10^{-16}$	28.8
Pathway Commmons	ChEA3 ENCODE	$<2.2x10^{-16}$	128.3
Pathway Commmons	ChEA3 Enrichr	$<2.2x10^{-16}$	134.3
Pathway Commmons	ChEA3 GTEx	$<2.2x10^{-16}$	122.6
Pathway Commmons	ChEA3 Literature	$<2.2x10^{-16}$	138.9
Pathway Commmons	ChEA3 ReMap	$<2.2x10^{-16}$	136.7
Pathway Commmons	Shuffled CollecTRI	$<2.2x10^{-16}$	100.6
RegNetwork	ChEA3 ARCHS4	$<2.2x10^{-16}$	80
RegNetwork	ChEA3 ENCODE	$<2.2x10^{-16}$	169
RegNetwork	ChEA3 Enrichr	$<2.2x10^{-16}$	189.3
RegNetwork	ChEA3 GTEx	$<2.2x10^{-16}$	175

Continued on the next page...

Continued from previous page.

GRN 1	GRN 2	AUROC Adj. P Value	AUROC T Value
RegNetwork	ChEA3 Literature	$<2.2x10^{-16}$	184.9
RegNetwork	ChEA3 ReMap	$<2.2x10^{-16}$	185.9
RegNetwork	DoRothEA ABCD	0.48	0.8
RegNetwork	Pathway Commmons	$<2.2x10^{-16}$	44.6
RegNetwork	Shuffled CollecTRI	$<2.2x10^{-16}$	152.9
Shuffled CollecTRI	ChEA3 ENCODE	$<2.2x10^{-16}$	46.9
Shuffled CollecTRI	ChEA3 Enrichr	$<2.2x10^{-16}$	33.3
Shuffled CollecTRI	ChEA3 GTEx	$<2.2x10^{-16}$	24
Shuffled CollecTRI	ChEA3 Literature	$<2.2x10^{-16}$	50.5
Shuffled CollecTRI	ChEA3 ReMap	$<2.2x10^{-16}$	43.6

Table A.2. AUPRC comparison across TF regulon collections. Area under the precision recall curve (AUPRC) values were compared between different transcription factor (TF) regulon collections using a two-sample t-test. AUPRC values represent the performance in recapitulating perturbed TFs from transcriptomics data. P-values from the t-test were adjusted using Benjamini-Hochberg

GRN 1	GRN 2	AUPRC Adj. P Value	AUPRC T Value
ChEA3 ARCHS4	ChEA3 ENCODE	$<2.2x10^{-16}$	132
ChEA3 ARCHS4	ChEA3 Enrichr	$<2.2x10^{-16}$	185
ChEA3 ARCHS4	ChEA3 GTEx	$<2.2x10^{-16}$	192.3
ChEA3 ARCHS4	ChEA3 Literature	$<2.2x10^{-16}$	183.8
ChEA3 ARCHS4	ChEA3 ReMap	$<2.2x10^{-16}$	160.2
ChEA3 ARCHS4	shuffled CollecTRI	$<2.2x10^{-16}$	138
ChEA3 Enrichr	ChEA3 ENCODE	$<2.2x10^{-16}$	4.5
ChEA3 Enrichr	ChEA3 Literature	$<2.2x10^{-16}$	20.6
ChEA3 Enrichr	ChEA3 ReMap	$<2.2x10^{-16}$	6.5
ChEA3 GTEx	ChEA3 ENCODE	$<2.2x10^{-16}$	10.5
ChEA3 GTEx	ChEA3 Enrichr	$<2.2x10^{-16}$	8.1
ChEA3 GTEx	ChEA3 Literature	$<2.2x10^{-16}$	13.4

Continued on the next page...

Continued from previous page.

GRN 1	GRN 2	AUPRC Adj. P Value	AUPRC T Value
ChEA3 GTEx	ChEA3 ReMap	$<2.2x10^{-16}$	13.7
ChEA3 Literature	ChEA3 ENCODE	$<2.2x10^{-16}$	20.3
ChEA3 ReMap	ChEA3 ENCODE	0.52	0.7
ChEA3 ReMap	ChEA3 Literature	$<2.2x10^{-16}$	24.8
CollecTRI	ChEA3 ARCHS4	$<2.2x10^{-16}$	237.8
CollecTRI	ChEA3 ENCODE	$<2.2x10^{-16}$	304.9
CollecTRI	ChEA3 Enrichr	$<2.2x10^{-16}$	431.1
CollecTRI	ChEA3 GTEx	$<2.2x10^{-16}$	437.6
CollecTRI	ChEA3 Literature	$<2.2x10^{-16}$	396.6
CollecTRI	ChEA3 ReMap	$<2.2x10^{-16}$	377.3
CollecTRI	DoRothEA ABC	$<2.2x10^{-16}$	83.2
CollecTRI	DoRothEA ABCD	$<2.2x10^{-16}$	160.5
CollecTRI	Pathway Commmons	$<2.2x10^{-16}$	170.6
CollecTRI	RegNetwork	$<2.2x10^{-16}$	139.5
CollecTRI	shuffled CollecTRI	$<2.2x10^{-16}$	360.7
DoRothEA ABC	ChEA3 ARCHS4	$<2.2x10^{-16}$	131.1
DoRothEA ABC	ChEA3 ENCODE	$<2.2x10^{-16}$	225.9
DoRothEA ABC	ChEA3 Enrichr	$<2.2x10^{-16}$	303.5
DoRothEA ABC	ChEA3 GTEx	$<2.2x10^{-16}$	310
DoRothEA ABC	ChEA3 Literature	$<2.2x10^{-16}$	290.9
DoRothEA ABC	ChEA3 ReMap	$<2.2x10^{-16}$	270.5
DoRothEA ABC	DoRothEA ABCD	$<2.2x10^{-16}$	69.6
DoRothEA ABC	Pathway Commmons	$<2.2x10^{-16}$	87.9
DoRothEA ABC	RegNetwork	$<2.2x10^{-16}$	55.6
DoRothEA ABC	shuffled CollecTRI	$<2.2x10^{-16}$	251.9
DoRothEA ABCD	ChEA3 ARCHS4	$<2.2x10^{-16}$	56.9
DoRothEA ABCD	ChEA3 ENCODE	$<2.2x10^{-16}$	170
DoRothEA ABCD	ChEA3 Enrichr	$<2.2x10^{-16}$	229.5
DoRothEA ABCD	ChEA3 GTEx	$<2.2x10^{-16}$	236.2
DoRothEA ABCD	ChEA3 Literature	$<2.2x10^{-16}$	224.9
DoRothEA ABCD	ChEA3 ReMap	$<2.2x10^{-16}$	203.1

Continued on the next page...

Continued from previous page.

GRN 1	GRN 2	AUPRC Adj. P Value	AUPRC T Value
DoRothEA ABCD	Pathway Commmons	$<2.2x10^{-16}$	24.1
DoRothEA ABCD	shuffled CollecTRI	$<2.2x10^{-16}$	183
Pathway Commmons	ChEA3 ARCHS4	$<2.2x10^{-16}$	26.3
Pathway Commmons	ChEA3 ENCODE	$<2.2x10^{-16}$	140.7
Pathway Commmons	ChEA3 Enrichr	$<2.2x10^{-16}$	183.4
Pathway Commmons	ChEA3 GTEx	$<2.2x10^{-16}$	189.7
Pathway Commmons	ChEA3 Literature	$<2.2x10^{-16}$	185.1
Pathway Commmons	ChEA3 ReMap	$<2.2x10^{-16}$	164.1
Pathway Commmons	shuffled CollecTRI	$<2.2x10^{-16}$	144.4
RegNetwork	ChEA3 ARCHS4	$<2.2x10^{-16}$	65.3
RegNetwork	ChEA3 ENCODE	$<2.2x10^{-16}$	173.6
RegNetwork	ChEA3 Enrichr	$<2.2x10^{-16}$	229.1
RegNetwork	ChEA3 GTEx	$<2.2x10^{-16}$	235.6
RegNetwork	ChEA3 Literature	$<2.2x10^{-16}$	225.8
RegNetwork	ChEA3 ReMap	$<2.2x10^{-16}$	204.8
RegNetwork	DoRothEA ABCD	$<2.2x10^{-16}$	10.7
RegNetwork	Pathway Commmons	$<2.2x10^{-16}$	33.1
RegNetwork	shuffled CollecTRI	$<2.2x10^{-16}$	185.4
shuffled CollecTRI	ChEA3 ENCODE	$<2.2x10^{-16}$	21
shuffled CollecTRI	ChEA3 Enrichr	$<2.2x10^{-16}$	33
shuffled CollecTRI	ChEA3 GTEx	$<2.2x10^{-16}$	40.4
shuffled CollecTRI	ChEA3 Literature	$<2.2x10^{-16}$	49.1
shuffled CollecTRI	ChEA3 ReMap	$<2.2x10^{-16}$	24.1

Appendix B

Comprehensive Evaluation of Kinase Activity Inference from Phosphoproteomics Data

B.1 Supplementary Tables

Table B.1. Description of computational methods for kinase activity inference.

Method	Description
fgsea	Fast gene set enrichment infers kinase activities using a weighted running sum method. It begins by ranking molecular features for each sample and calculates an enrichment score by traversing the ranked list. The running sum statistic increases when a feature belongs to the target set and decreases when it does not. The size of the increment is proportional to the correlation between the feature and the phenotype.
	Continued on the next page...

Continued from previous page.

Method	Description
Fisher's exact test	Fisher's exact test assesses the over-representation of molecular features associated with a kinase's target sites compared to non-target sites. This method relies on a contingency table that categorizes phosphorylation sites into four groups: kinase targets versus non-targets and deregulated versus non-deregulated sites.
KARP	KARP calculates a K-score by taking the ratio of the sum of molecular features for a kinase's target sites to the sum of molecular features across all phosphorylation sites. This score is then adjusted to account for the imbalance in known targets by multiplying it by the square root of the ratio of measured targets for the kinase to the total number of known targets in the given resource.
KSEA	KSEA calculates a z-score to quantify the difference between the mean molecular features of a kinase's known targets and the mean molecular features of all phosphorylation sites. This difference is normalized by the square root of the number of identified targets and the standard deviation of the molecular features across all phosphorylation sites.
Kologomorov-Smirnov	The Kolmogorov-Smirnov test compares the running sums of molecular features for a kinase's targets and non-targets. Features are ranked, and the running sum statistic increases uniformly when a feature belongs to the target list. Unlike fgsea, the increment size remains constant regardless of the feature's correlation.

Continued on the next page...

Continued from previous page.

Method	Description
Linear model - RoKAI	The linear model implemented in RoKAI simultaneously analyzes the molecular readouts of all phosphorylation sites across all kinases. Each phosphorylation site is modeled as the sum of activities of its associated kinases, with weights for non-targets set to zero. Kinase activities are inferred using least squares optimization with ridge regularization to ensure stability and prevent overfitting.
Mann-Whitney-U	The Mann-Whitney U test, also called the Wilcoxon rank-sum test, compares the ranks of molecular features between a kinase's targets and non-targets. Phosphorylation sites are ranked together based on their molecular features, and the U-statistic is computed from the sum of ranks for the target and non-target groups.
mean	The mean represents the average value of the molecular features across all target sites of a kinase.
median	The median is the middle value of the molecular features for all target sites of a kinase, determined after ranking these features in ascending or descending order.
multivariate linear model	The multivariate linear model, as implemented in decoupler, simultaneously models the molecular readouts of all features across all kinases. Each phosphorylation site is represented as the sum of activities of its associated kinases, with weights for non-targets set to zero to focus solely on the relevant kinase associations.

Continued on the next page...

Continued from previous page.

Method	Description
normalized mean	The normalized mean is calculated by first generating a random null distribution of means through random permutations of target features. The average value of the molecular features for all target sites of a kinase is then normalized by subtracting the mean of the null distribution and dividing by its standard deviation.
principal component analysis	Principal Component Analysis (PCA) is conducted across samples using only the molecular features of a kinase's target sites. The kinase score is determined by the variance explained by the first principal component.
PTM-SEA	PTM-SEA calculates an enrichment score using the same approach as fgsea, based on a weighted running sum. Additionally, it computes a normalized enrichment score by comparing the enrichment score to a null distribution generated through random permutations of target features.
sum	The sum represents the total of the molecular feature values across all target sites of a kinase.
univariate linear model	The univariate linear model, as implemented in decoupler, models the molecular readouts of all features for each kinase individually. Non-target weights are set to zero, and the t-value obtained from the fitted model is used to represent the activity of the kinase.
upper quantile	The upper quantile represents the value below which 75% of the molecular feature values for all target sites of a kinase are distributed.

Continued on the next page...

Continued from previous page.

Method	Description
VIPER	VIPER estimates kinase activities by calculating a three-tailed enrichment score, which evaluates the ranking of all phosphorylation sites relative to a kinase's targets based on their molecular features. A normalized enrichment score is then derived by comparing the observed enrichment score to a null distribution generated through random permutations.
z-score - RoKAI	The z-score, as implemented in RoKAI, calculates the mean of the molecular features for a kinase's known targets. This value is then normalized by dividing it by the square root of the number of identified targets for the kinase and the standard deviation of the molecular features across all phosphorylation sites.
X^2-test	The Chi-square test, similar in purpose to Fisher's exact test, evaluates whether kinase target sites are more associated with deregulation compared to non-target sites. It does so by comparing the observed and expected counts in the contingency table used for Fisher's exact test. Unlike Fisher's exact test, the Chi-square test is better suited for larger sample sizes and uses a large-sample approximation to determine statistical significance.

Appendix C

Phosphoproteomics Based Network Contextualization to Understand Metformin's Mechanisms in Colon Cancer

C.1 Supplementary Figures

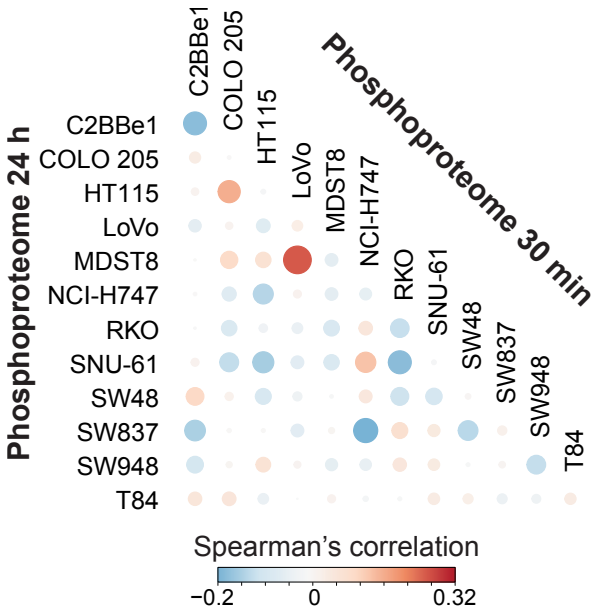

Figure C.1. Comparison between the acute and late response of metformin in each cell line. Spearman correlation between the 30-min and 24-h log2 fold changes at the phosphoproteome level for each colorectal cancer cell line. *Reprinted from Salovska et al. (2023).*

C.2 Supplementary Equations

1. Vertex activation/inhibition exclusivity:

$$V_{\text{act}}(v) + V_{\text{inh}}(v) \leq 1, \quad \forall v \in V. \tag{C.1}$$

where:

$V_{\text{act}}(v)$: Node activation indicator for node v, $V_{\text{act}}(v) \in \{0, 1\}$
$V_{\text{inh}}(v)$: Node inhibition indicator for node v, $V_{\text{inh}}(v) \in \{0, 1\}$
V : Set of all nodes in the graph

2. Edge activation/inhibition exclusivity:

$$E_{\text{act}}(e) + E_{\text{inh}}(e) \leq 1, \quad \forall e \in E. \tag{C.2}$$

where:

$E_{\text{act}}(e)$: Edge activation indicator for edge e, $E_{\text{act}}(e) \in \{0, 1\}$
$E_{\text{inh}}(e)$: Edge inhibition indicator for edge e, $E_{\text{inh}}(e) \in \{0, 1\}$
E : Set of all edges in the graph

3. Sign consistency:

For each edge $e \in E$ from s_e to t_e:

- If interaction$(e) = 1$ (activation):

$$E_{\text{act}}(e) \leq V_{\text{act}}(s_e), \quad E_{\text{inh}}(e) \leq V_{\text{inh}}(s_e). \tag{C.3}$$

where:

s_e : Source node of edge e
t_e : Target node of edge e
$V_{\text{act}}(s_e)$: Node activation indicator for source node s_e, $V_{\text{act}}(s_e) \in \{0, 1\}$
$V_{\text{inh}}(s_e)$: Node inhibition indicator for source node s_e, $V_{\text{inh}}(s_e) \in \{0, 1\}$
$E_{\text{act}}(e)$: Edge activation indicator for edge e, $E_{\text{act}}(e) \in \{0, 1\}$
$E_{\text{inh}}(e)$: Edge inhibition indicator for edge e, $E_{\text{inh}}(e) \in \{0, 1\}$

- If interaction$(e) = -1$ (inhibition):

$$E_{\text{act}}(e) \leq V_{\text{inh}}(s_e), \quad E_{\text{inh}}(e) \leq V_{\text{act}}(s_e). \tag{C.4}$$

where:

s_e : Source node of edge e
t_e : Target node of edge e
$V_{\text{act}}(s_e)$: Node activation indicator for source node s_e, $V_{\text{act}}(s_e) \in \{0, 1\}$
$V_{\text{inh}}(s_e)$: Node inhibition indicator for source node s_e, $V_{\text{inh}}(s_e) \in \{0, 1\}$
$E_{\text{act}}(e)$: Edge activation indicator for edge e, $E_{\text{act}}(e) \in \{0, 1\}$
$E_{\text{inh}}(e)$: Edge inhibition indicator for edge e, $E_{\text{inh}}(e) \in \{0, 1\}$

4. Acyclicity constraints:

$$V_{\text{pos}}(t_e) - V_{\text{pos}}(s_e) \geq 1 - M\left[1 - (E_{\text{act}}(e) + E_{\text{inh}}(e))\right], \quad \forall e \in E. \tag{C.5}$$

where:

$V_{\text{pos}}(t_e)$: Position of target vertex t_e in the acyclic graph
$V_{\text{pos}}(s_e)$: Position of source vertex s_e in the acyclic graph
M : A large constant used to enforce acyclicity
$E_{\text{act}}(e)$: Edge activation indicator for edge e, $E_{\text{act}}(e) \in \{0, 1\}$
$E_{\text{inh}}(e)$: Edge inhibition indicator for edge e, $E_{\text{inh}}(e) \in \{0, 1\}$

5. Signal propagation as a tree:

$$\sum_{e \in \text{In}(v)} (E_{\text{act}}(e) + E_{\text{inh}}(e)) \leq 1, \quad \forall v \in V. \tag{C.6}$$

where:

$\text{In}(v)$: Set of edges ending at vertex v
$E_{\text{act}}(e)$: Edge activation indicator for edge e, $E_{\text{act}}(e) \in \{0, 1\}$
$E_{\text{inh}}(e)$: Edge inhibition indicator for edge e, $E_{\text{inh}}(e) \in \{0, 1\}$
V : Set of all vertices in the graph

6. Vertex activation/inhibition conditions:

For each non-perturbed vertex $v \in V \setminus V_p$:

$$V_{\text{act}}(v) \leq \sum_{e \in \text{In}(v)} E_{\text{act}}(e), \quad V_{\text{inh}}(v) \leq \sum_{e \in \text{In}(v)} E_{\text{inh}}(e). \tag{C.7}$$

where:

$\text{In}(v)$: Set of edges ending at vertex v
$E_{\text{act}}(e)$: Edge activation indicator for edge e, $E_{\text{act}}(e) \in \{0, 1\}$
$E_{\text{inh}}(e)$: Edge inhibition indicator for edge e, $E_{\text{inh}}(e) \in \{0, 1\}$
$V_{\text{act}}(v)$: Node activation indicator for node v, $V_{\text{act}}(v) \in \{0, 1\}$
$V_{\text{inh}}(v)$: Node inhibition indicator for node v, $V_{\text{inh}}(v) \in \{0, 1\}$
V_p : Set of perturbed vertices

7. Perturbed inputs:

For each perturbed vertex v provided as input:

- If activation (perturbation value $= +1$):

$$V_{\text{act}}(v) = 1, \quad V_{\text{inh}}(v) = 0. \tag{C.8}$$

- If inhibition (perturbation value $= -1$):

$$V_{\text{act}}(v) = 0, \quad V_{\text{inh}}(v) = 1. \tag{C.9}$$

where:

$V_{\text{act}}(v)$: Node activation indicator for node v, $V_{\text{act}}(v) \in \{0, 1\}$
$V_{\text{inh}}(v)$: Node inhibition indicator for node v, $V_{\text{inh}}(v) \in \{0, 1\}$
v : Perturbed vertex

Appendix D

Multi-Omics Network Contextualization to Investigate Hepatic Stellate Cell Activation

D.1 Supplementary Figures

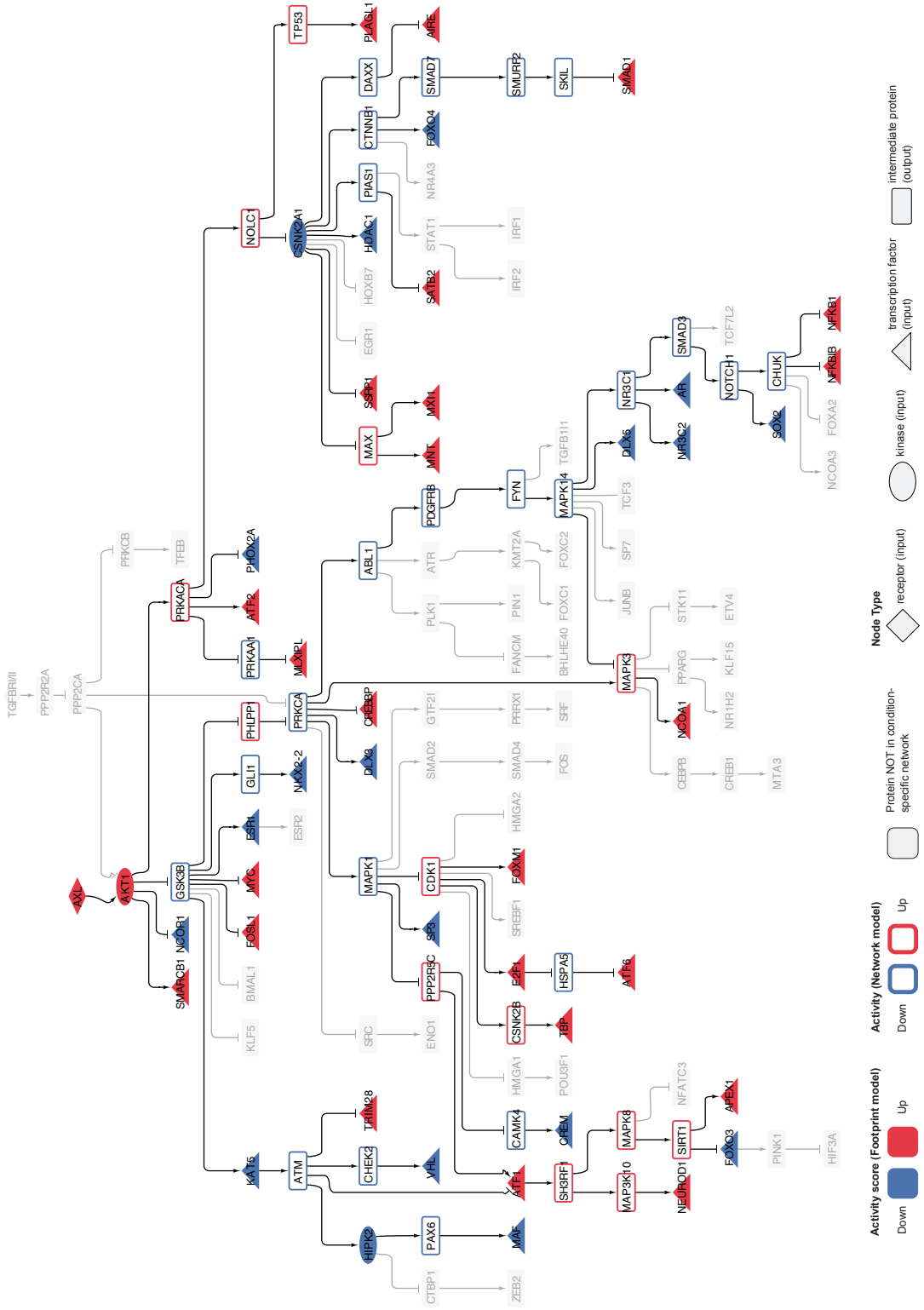

Figure D.1. Contextualized signaling network for GAS6 stimulation in hepatic stellate cells. Contextualized network connecting signaling events to transcriptional regulation induced by GAS6 stimulation. In the network, nodes represent proteins, and arrows denote regulatory relationships. The network was jointly modeled with two other conditions, namely TGFβ and a combined stimulation of TGFβ and GAS6, linking the targeted receptors AXL (diamond) to deregulated kinases (circles) and transcription factors (triangles) identified through footprint-based activity inference. The activity of intermediate proteins (rectangles) was inferred within the network model. Nodes from all conditions were integrated to construct a backbone network.

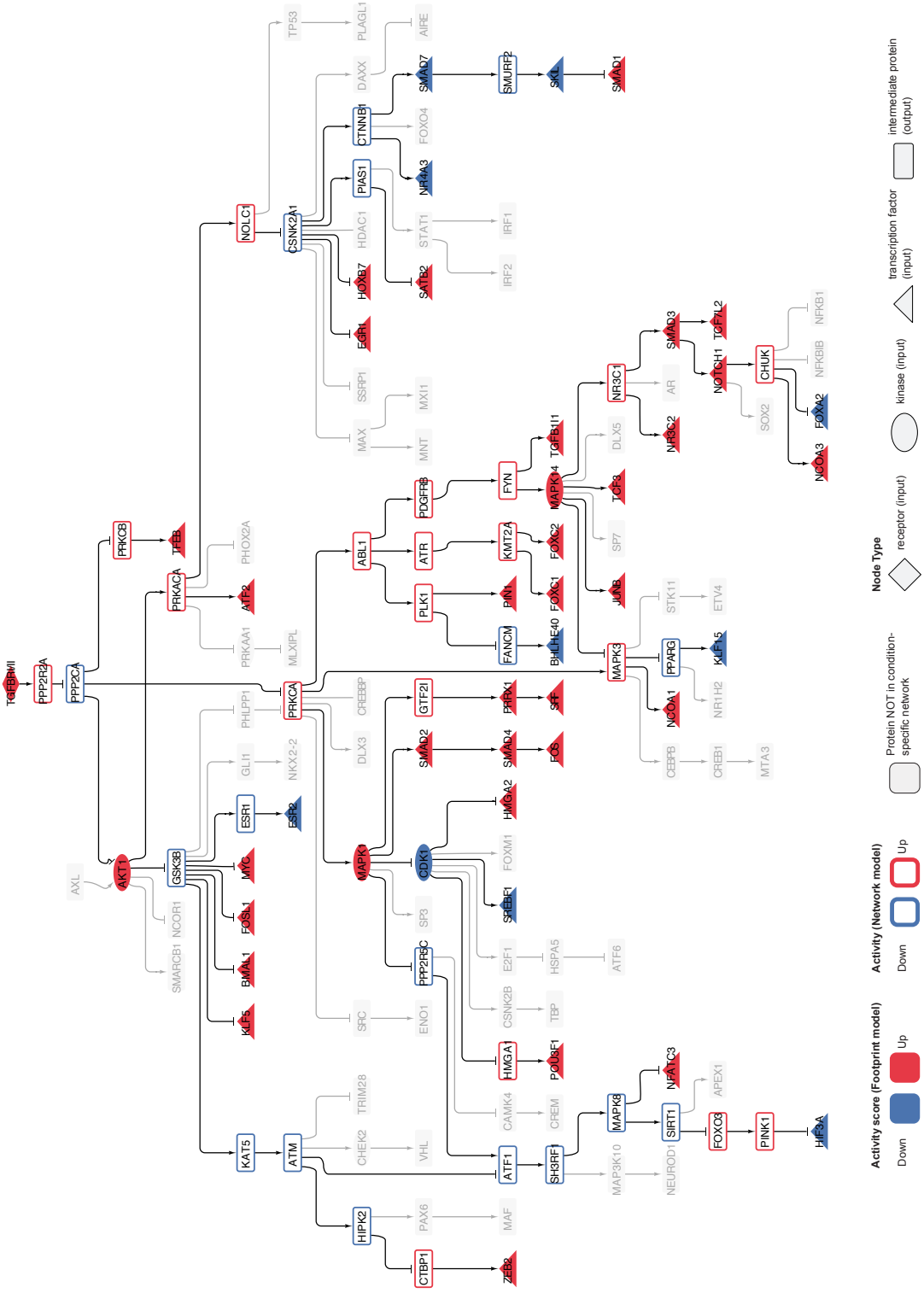

Figure D.2. Contextualized signaling network for TGFβ stimulation in hepatic stellate cells. Contextualized network connecting signaling events to transcriptional regulation induced by TGFβ stimulation. In the network, nodes represent proteins, and arrows denote regulatory relationships. The network was jointly modeled with two other conditions, namely GAS6 and a combined stimulation of TGFβ and GAS6, linking the targeted receptors TGFßRI/II (diamond) to deregulated kinases (circles) and transcription factors (triangles) identified through footprint-based activity inference. The activity of intermediate proteins (rectangles) was inferred within the network model. Nodes from all conditions were integrated to construct a backbone network.

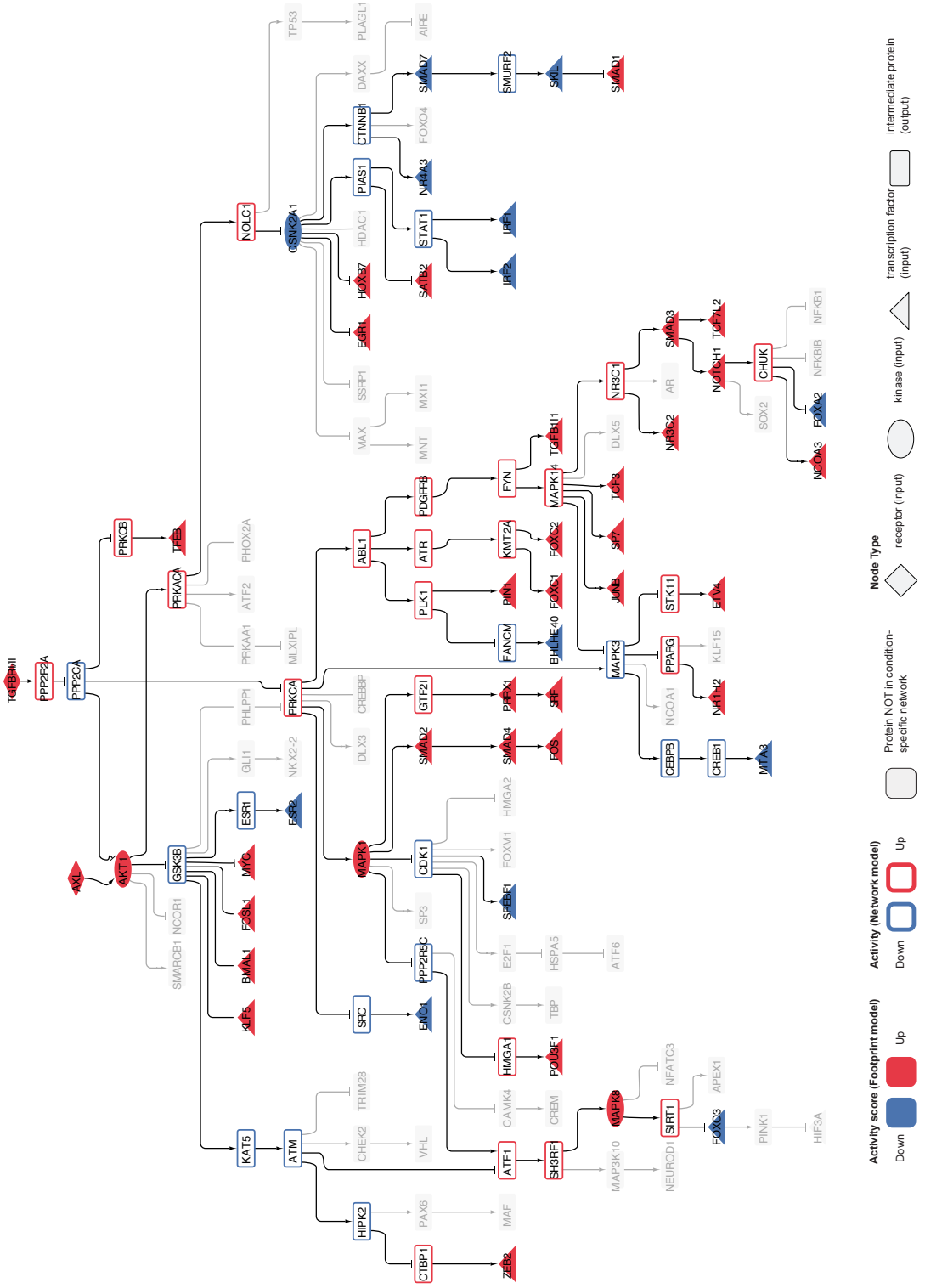

Figure D.3. Contextualized signaling network for GAS6 and TGFβ stimulation in hepatic stellate cells. Contextualized network connecting signaling events to transcriptional regulation induced by GAS6 and TGFβ stimulation. In the network, nodes represent proteins, and arrows denote regulatory relationships. The network was jointly modeled with two other conditions, namely GAS6 and TGFβ, linking the targeted receptors AXL and TGFßRI/II (diamonds) to deregulated kinases (circles) and transcription factors (triangles) identified through footprint-based activity inference. The activity of intermediate proteins (rectangles) was inferred within the network model. Nodes from all conditions were integrated to construct a backbone network.

D.2 Supplementary Tables

Table D.1. Top differentially expressed genes per stimulation. Top 10 significantly differentially expressed genes based on their absolute logFC across time points per stimulation (GAS6, TGFß, GAS6+TGFß). Differential expression analysis was performed using limma and p-values were adjusted using Benjamini-Hochberg.

Gene	logFC	adj. P-value	Time [h]	Stimulation
SLC39A10	1.07	0.03	1	GAS6
TMEM132E	-1.17	0.045	4	GAS6
MT1E	1.49	0.009	20	GAS6
MT1F	1.41	0.011	20	GAS6
MT1X	1.28	0.03	20	GAS6
SLC39A10	-1.06	0.034	20	GAS6
KRT15	-1	0.049	20	GAS6
PMEPA1	5.61	$<2.2x10^{-16}$	12	TGFß
ODAPH	5.59	$<2.2x10^{-16}$	12	TGFß
FOXS1	5.52	$<2.2x10^{-16}$	12	TGFß
FOXS1	6.22	$<2.2x10^{-16}$	20	TGFß
PRG4	6.1	$<2.2x10^{-16}$	20	TGFß
PMEPA1	5.71	$<2.2x10^{-16}$	20	TGFß
PRG4	6.95	$<2.2x10^{-16}$	28	TGFß
ISLR2	6.19	$<2.2x10^{-16}$	28	TGFß
FOXS1	6.04	$<2.2x10^{-16}$	28	TGFß
PMEPA1	5.84	$<2.2x10^{-16}$	28	TGFß
PMEPA1	5.51	$<2.2x10^{-16}$	7	TGFß + GAS6
FOXS1	5.66	$<2.2x10^{-16}$	12	TGFß + GAS6
ODAPH	5.58	$<2.2x10^{-16}$	12	TGFß + GAS6
FOXS1	6.23	$<2.2x10^{-16}$	20	TGFß + GAS6
PRG4	5.84	$<2.2x10^{-16}$	20	TGFß + GAS6
PMEPA1	5.56	$<2.2x10^{-16}$	20	TGFß + GAS6
PRG4	6.95	$<2.2x10^{-16}$	28	TGFß + GAS6
FOXS1	6.34	$<2.2x10^{-16}$	28	TGFß + GAS6
ISLR2	6.03	$<2.2x10^{-16}$	28	TGFß + GAS6
PMEPA1	5.74	$<2.2x10^{-16}$	28	TGFß + GAS6

Table D.2. Top differentially abundant phosphorylation sites per stimulation. Top 10 significantly differentially abundant phosphorylation sites based on their absolute logFC across time points per stimulation (GAS6, TGFß, GAS6+TGFß). Differential analysis was performed using limma and p-values were adjusted using Benjamini-Hochberg.

Phosphorylation site	logFC	adj. P-value	Time [h]	Stimulation
EHBP1_S428	-5.46	0.016	0.08	GAS6
UBAP2L_S467	-4.6	$<2.2x10^{-16}$	0.08	GAS6
NCOR2_S943	-4.55	0.009	0.08	GAS6
NUP155_S992	-4.35	0.013	0.08	GAS6
SH3RF3_S404	4.68	0.011	4	GAS6
MAP1A_S1146	-7.88	0.014	7	GAS6
DEK_S19	5.2	0.014	7	GAS6
EML3_T881	4.59	$<2.2x10^{-16}$	7	GAS6
EHBP1_S428	-4.92	0.028	12	GAS6
ARHGEF28_S513	-4.25	$<2.2x10^{-16}$	12	GAS6
MAP1B_S1819	5.91	0.036	4.5	TGFß
HIVEP1_S1884	-5.83	0.04	5	TGFß
H2AZ1_T104	5.47	0.015	7	TGFß
NUAK1_S388	5.04	$<2.2x10^{-16}$	12	TGFß
YEATS2_S471	8.34	0.01	28	TGFß
VPS50_S28	-5.24	$<2.2x10^{-16}$	28	TGFß
ETV3_S159	6.41	0.009	52	TGFß
GAPVD1_S946	-5.4	0.013	52	TGFß
HLA-A_S359	5.05	0.014	52	TGFß
CHERP_S830	-4.96	0.003	52	TGFß
EML3_T881	5	$<2.2x10^{-16}$	7	TGFß + GAS6
YEATS2_S471	-7.87	0.007	12	TGFß + GAS6
CORO7_S879	-6.84	0.006	12	TGFß + GAS6
CORO7_S880	-6.84	0.006	12	TGFß + GAS6
CORO7_Y883	-6.84	0.006	12	TGFß + GAS6
ERCC5_S382	5.12	$<2.2x10^{-16}$	12	TGFß + GAS6
YEATS2_S471	7.24	0.014	28	TGFß + GAS6
VPS50_S28	-6.45	$<2.2x10^{-16}$	28	TGFß + GAS6
RAPH1_T1153	5.7	0.046	52	TGFß + GAS6
ABI1_Y213	5.26	0.039	52	TGFß + GAS6

Glossary

AUC area under the curve

AUPRC area under the precision-recall curve

AUROC area under the receiver operating characteristic curve

cDNA complementary DNA

coTF co-regulatory transcription factor

CRC colorectal cancer

dbTF DNA-binding transcription factor

DDA data-dependent acquisition

DIA data-independent acquisition

ECM extracellular matrix

GAS6 growth arrest-specific 6

GO gene ontology

GRN Gene regulatory networks

GTF general initiation transcription factor

HSC hepatic stellate cell

ILP integer linear programming

KRAB Krüppel associated box

MoR mode of regulation

mRNA messenger RNA

MS mass spectrometry

NES normalized enrichment score

PBMC peripheral blood mononuclear cell

PBS phosphate-buffered saline

PCA principal component analysis

PMID PubMed identifier

PTM post-translational modification

RNA ribonucleic acid

RNA-seq RNA sequencing

TF transcription factor

TGFβ transforming growth factor beta

TN tue negative

TP true positive

TSS transcription start site

UMAP uniform manifold approximation and projection

vsn variance-stabilizing normalization

Bibliography

Aebersold, R. and M. Mann (2016, September). Mass-spectrometric exploration of proteome structure and function. *Nature 537*(7620), 347–355.

Agrawal, A., H. Balcı, K. Hanspers, S. L. Coort, M. Martens, D. N. Slenter, F. Ehrhart, D. Digles, A. Waagmeester, I. Wassink, T. Abbassi-Daloii, E. N. Lopes, A. Iyer, J. M. Acosta, L. G. Willighagen, K. Nishida, A. Riutta, H. Basaric, C. T. Evelo, E. L. Willighagen, M. Kutmon, and A. R. Pico (2024, January). WikiPathways 2024: next generation pathway database. *Nucleic Acids Res 52*(D1), D679–D689.

Albert, R. (2007, November). Network inference, analysis, and modeling in systems biology. *Plant Cell 19*(11), 3327–3338.

Alon, U. (2007, June). Network motifs: theory and experimental approaches. *Nature Reviews Genetics 8*(6), 450–461.

Alvarez, M. J., Y. Shen, F. M. Giorgi, A. Lachmann, B. B. Ding, B. H. Ye, and A. Califano (2016, August). Functional characterization of somatic mutations in cancer using network-based inference of protein activity. *Nat Genet 48*(8), 838–847.

Arnold, M., M. S. Sierra, M. Laversanne, I. Soerjomataram, A. Jemal, and F. Bray (2017, April). Global patterns and trends in colorectal cancer incidence and mortality. *Gut 66*(4), 683–691.

Aslam, B., M. Basit, M. A. Nisar, M. Khurshid, and M. H. Rasool (2017, February). Proteomics: Technologies and their applications. *J Chromatogr Sci 55*(2), 182–196.

Bachman, J. A., P. K. Sorger, and B. M. Gyori (2022, October). Assembling a corpus of phosphoproteomic annotations using ProtMapper to normalize site information from databases and text mining.

Badia-I-Mompel, P., J. Vélez Santiago, J. Braunger, C. Geiss, D. Dimitrov, S. Müller-Dott, P. Taus, A. Dugourd, C. H. Holland, R. O. Ramirez Flores, and J. Saez-Rodriguez

(2022, March). decoupler: ensemble of computational methods to infer biological activities from omics data. *Bioinform Adv 2*(1), vbac016.

Bantscheff, M., S. Lemeer, M. M. Savitski, and B. Kuster (2012, September). Quantitative mass spectrometry in proteomics: critical review update from 2007 to the present. *Anal Bioanal Chem 404*(4), 939–965.

Barabási, A.-L., N. Gulbahce, and J. Loscalzo (2011, January). Network medicine: a network-based approach to human disease. *Nat Rev Genet 12*(1), 56–68.

Bárcena, C., M. Stefanovic, A. Tutusaus, L. Joannas, A. Menéndez, C. García-Ruiz, P. Sancho-Bru, M. Marí, J. Caballeria, C. V. Rothlin, J. C. Fernández-Checa, P. G. de Frutos, and A. Morales (2015, September). Gas6/Axl pathway is activated in chronic liver disease and its targeting reduces fibrosis via hepatic stellate cell inactivation. *J Hepatol 63*(3), 670–678.

Barry, W. T., A. B. Nobel, and F. A. Wright (2005, May). Significance analysis of functional categories in gene expression studies: a structured permutation approach. *Bioinformatics 21*(9), 1943–1949.

Beausoleil, S. A., M. Jedrychowski, D. Schwartz, J. E. Elias, J. Villén, J. Li, M. A. Cohn, L. C. Cantley, and S. P. Gygi (2004, August). Large-scale characterization of HeLa cell nuclear phosphoproteins. *Proc Natl Acad Sci U S A 101*(33), 12130–12135.

Beck, M., A. Schmidt, J. Malmstroem, M. Claassen, A. Ori, A. Szymborska, F. Herzog, O. Rinner, J. Ellenberg, and R. Aebersold (2011, November). The quantitative proteome of a human cell line. *Mol Syst Biol 7*, 549.

Bejjani, F., E. Evanno, K. Zibara, M. Piechaczyk, and I. Jariel-Encontre (2019, August). The AP-1 transcriptional complex: Local switch or remote command? *Biochim Biophys Acta Rev Cancer 1872*(1), 11–23.

Binns, D., E. Dimmer, R. Huntley, D. Barrell, C. O'Donovan, and R. Apweiler (2009, November). QuickGO: a web-based tool for gene ontology searching. *Bioinformatics 25*(22), 3045–3046.

Boulesteix, A.-L., S. Lauer, and M. J. A. Eugster (2013, April). A plea for neutral comparison studies in computational sciences. *PLOS ONE 8*(4), e61562.

Bovolenta, L. A., M. L. Acencio, and N. Lemke (2012, August). HTRIdb: an open-access database for experimentally verified human transcriptional regulation interactions. *BMC Genomics 13*, 405.

Bradley, G. and S. J. Barrett (2017, November). CausalR: extracting mechanistic sense from genome scale data. *Bioinformatics 33*(22), 3670–3672.

Bragin, M. A., P. B. Luh, B. Yan, and X. Sun (2019). A scalable solution methodology for mixed-integer linear programming problems arising in automation. *IEEE Transactions on Automation Science and Engineering 16*(2), 531–541.

Breitkopf-Heinlein, K. and M. L. Martinez-Chantar (2024, September). Targeting hepatic stellate cells to combat liver fibrosis: where do we stand? *Gut 73*(9), 1411–1413.

Browaeys, R., W. Saelens, and Y. Saeys (2019, December). NicheNet: modeling intercellular communication by linking ligands to target genes. *Nature Methods 17*(2), 159–162.

Carrasco Pro, S., A. Dafonte Imedio, C. S. Santoso, K. A. Gan, J. A. Sewell, M. Martinez, R. Sereda, S. Mehta, and J. I. Fuxman Bass (2018, October). Global landscape of mouse and human cytokine transcriptional regulation. *Nucleic Acids Res 46*(18), 9321–9337.

Casado, P., J.-C. Rodriguez-Prados, S. C. Cosulich, S. Guichard, B. Vanhaesebroeck, S. Joel, and P. R. Cutillas (2013, March). Kinase-substrate enrichment analysis provides insights into the heterogeneity of signaling pathway activation in leukemia cells. *Sci Signal 6*(268), rs6.

Chen, L., A. E. Fish, and J. A. Capra (2018, October). Prediction of gene regulatory enhancers across species reveals evolutionarily conserved sequence properties. *PLoS Comput Biol 14*(10), e1006484.

Chu, E. C.-P., A. Morin, T. H. C. Chang, T. Nguyen, Y.-C. Tsai, A. Sharma, C. C. Liu, and P. Pavlidis (2021, October). Experiment level curation of transcriptional regulatory interactions in neurodevelopment. *PLoS Comput Biol 17*(10), e1009484.

Cobaleda, C., A. Schebesta, A. Delogu, and M. Busslinger (2007, May). Pax5: the guardian of B cell identity and function. *Nat Immunol 8*(5), 463–470.

Colorectal Cancer Collaborators (2022, July). Global, regional, and national burden of colorectal cancer and its risk factors, 1990-2019: a systematic analysis for the global burden of disease study 2019. *Lancet Gastroenterol Hepatol 7*(7), 627–647.

Conesa, A., P. Madrigal, S. Tarazona, D. Gomez-Cabrero, A. Cervera, A. McPherson, M. W. Szcześniak, D. J. Gaffney, L. L. Elo, X. Zhang, and A. Mortazavi (2016, January). A survey of best practices for RNA-seq data analysis. *Genome Biol 17*, 13.

Consortium, I. H. G. S. (2004, October). Finishing the euchromatic sequence of the human genome. *Nature 431*(7011), 931–945.

Copps, K. D. and M. F. White (2012, October). Regulation of insulin sensitivity by serine/threonine phosphorylation of insulin receptor substrate proteins IRS1 and IRS2. *Diabetologia 55*(10), 2565–2582.

Cox, J. and M. Mann (2007, August). Is proteomics the new genomics? *Cell 130*(3), 395–398.

Dewidar, B., J. Soukupova, I. Fabregat, and S. Dooley (2015, October). TGF-β in hepatic stellate cell activation and liver fibrogenesis: Updated. *Current Pathobiology Reports 3*(4), 291–305.

Dinkel, H., C. Chica, A. Via, C. M. Gould, L. J. Jensen, T. J. Gibson, and F. Diella (2011, January). Phospho.ELM: a database of phosphorylation sites–update 2011. *Nucleic Acids Res 39*(Database issue), D261–7.

Dittrich, M. T., G. W. Klau, A. Rosenwald, T. Dandekar, and T. Müller (2008, July). Identifying functional modules in protein-protein interaction networks: an integrated exact approach. *Bioinformatics 24*(13), i223–31.

Dixit, A., O. Parnas, B. Li, J. Chen, C. P. Fulco, L. Jerby-Arnon, N. D. Marjanovic, D. Dionne, T. Burks, R. Raychowdhury, B. Adamson, T. M. Norman, E. S. Lander, J. S. Weissman, N. Friedman, and A. Regev (2016, December). Perturb-Seq: Dissecting molecular circuits with scalable Single-Cell RNA profiling of pooled genetic screens. *Cell 167*(7), 1853–1866.e17.

Dooley, S. and P. ten Dijke (2012, January). TGF-β in progression of liver disease. *Cell Tissue Res 347*(1), 245–256.

Dugourd, A. and J. Saez-Rodriguez (2019, June). Footprint-based functional analysis of multiomic data. *Curr Opin Syst Biol 15*, 82–90.

Engholm-Keller, K. and M. R. Larsen (2013, March). Technologies and challenges in large-scale phosphoproteomics. *Proteomics 13*(6), 910–931.

Essaghir, A., F. Toffalini, L. Knoops, A. Kallin, J. van Helden, and J.-B. Demoulin (2010, June). Transcription factor regulation can be accurately predicted from the presence of target gene signatures in microarray gene expression data. *Nucleic Acids Res 38*(11), e120.

Evans, J. M. M., L. A. Donnelly, A. M. Emslie-Smith, D. R. Alessi, and A. D. Morris (2005, June). Metformin and reduced risk of cancer in diabetic patients. *BMJ 330*(7503), 1304–1305.

Fabian, M. A., W. H. Biggs, 3rd, D. K. Treiber, C. E. Atteridge, M. D. Azimioara, M. G. Benedetti, T. A. Carter, P. Ciceri, P. T. Edeen, M. Floyd, J. M. Ford, M. Galvin, J. L. Gerlach, R. M. Grotzfeld, S. Herrgard, D. E. Insko, M. A. Insko, A. G. Lai, J.-M. Lélias, S. A. Mehta, Z. V. Milanov, A. M. Velasco, L. M. Wodicka, H. K. Patel, P. P. Zarrinkar, and D. J. Lockhart (2005, March). A small molecule-kinase interaction map for clinical kinase inhibitors. *Nat Biotechnol 23*(3), 329–336.

Fabregat, A., K. Sidiropoulos, G. Viteri, O. Forner, P. Marin-Garcia, V. Arnau, P. D'Eustachio, L. Stein, and H. Hermjakob (2017, March). Reactome pathway analysis: a high-performance in-memory approach. *BMC Bioinformatics 18*(1), 142.

Fearon, E. R. (2011). Molecular genetics of colorectal cancer. *Annu Rev Pathol 6*, 479–507.

Feng, C., C. Song, Y. Liu, F. Qian, Y. Gao, Z. Ning, Q. Wang, Y. Jiang, Y. Li, M. Li, J. Chen, J. Zhang, and C. Li (2020, January). KnockTF: a comprehensive human gene expression profile database with knockdown/knockout of transcription factors. *Nucleic Acids Res 48*(D1), D93–D100.

Forner, A., J. M. Llovet, and J. Bruix (2012, March). Hepatocellular carcinoma. *The Lancet 379*(9822), 1245–1255.

Fortelny, N. and C. Bock (2020, August). Knowledge-primed neural networks enable biologically interpretable deep learning on single-cell sequencing data. *Genome Biol 21*(1), 190.

Franke, L., H. van Bakel, L. Fokkens, E. D. de Jong, M. Egmont-Petersen, and C. Wijmenga (2006, June). Reconstruction of a functional human gene network, with an application for prioritizing positional candidate genes. *Am J Hum Genet 78*(6), 1011–1025.

Fraser, L. C. R., R. J. Dikdan, S. Dey, A. Singh, and S. Tyagi (2021, October). Reduction in gene expression noise by targeted increase in accessibility at gene loci. *Proc Natl Acad Sci U S A 118*(42).

Friedman, N., M. Linial, I. Nachman, and D. Pe'er (2000). Using bayesian networks to analyze expression data. *J Comput Biol 7*(3-4), 601–620.

Friedman, S. L. (2008, January). Hepatic stellate cells: protean, multifunctional, and enigmatic cells of the liver. *Physiol Rev 88*(1), 125–172.

Garcia-Alonso, L., C. H. Holland, M. M. Ibrahim, D. Turei, and J. Saez-Rodriguez (2019, August). Benchmark and integration of resources for the estimation of human transcription factor activities. *Genome Res 29*(8), 1363–1375.

Garcia-Alonso, L., F. Iorio, A. Matchan, N. Fonseca, P. Jaaks, G. Peat, M. Pignatelli, F. Falcone, C. H. Benes, I. Dunham, G. Bignell, S. S. McDade, M. J. Garnett, and J. Saez-Rodriguez (2018, February). Transcription factor activities enhance markers of drug sensitivity in cancer. *Cancer Res 78*(3), 769–780.

Garrido-Rodriguez, M., K. Zirngibl, O. Ivanova, S. Lobentanzer, and J. Saez-Rodriguez (2022, July). Integrating knowledge and omics to decipher mechanisms via large-scale models of signaling networks. *Molecular Systems Biology*.

Gene Ontology Consortium (2021, January). The gene ontology resource: enriching a GOld mine. *Nucleic Acids Res 49*(D1), D325–D334.

Giarratana, A. O., C. M. Prendergast, M. M. Salvatore, and K. M. Capaccione (2024, June). TGF-β signaling: critical nexus of fibrogenesis and cancer. *Journal of Translational Medicine 22*(1), 1–16.

Gillet, L. C., P. Navarro, S. Tate, H. Röst, N. Selevsek, L. Reiter, R. Bonner, and R. Aebersold (2012, June). Targeted data extraction of the MS/MS spectra generated by data-independent acquisition: a new concept for consistent and accurate proteome analysis. *Mol Cell Proteomics 11*(6), O111.016717.

Gjerga, E., A. Dugourd, L. Tobalina, A. Sousa, and J. Saez-Rodriguez (2021, April). PHONEMeS: Efficient modeling of signaling networks derived from Large-Scale mass spectrometry data. *J Proteome Res 20*(4), 2138–2144.

Gordon, S. M., J. Chaix, L. J. Rupp, J. Wu, S. Madera, J. C. Sun, T. Lindsten, and S. L. Reiner (2012, January). The transcription factors t-bet and eomes control key checkpoints of natural killer cell maturation. *Immunity 36*(1), 55–67.

Grant, C. E., T. L. Bailey, and W. S. Noble (2011, April). FIMO: scanning for occurrences of a given motif. *Bioinformatics 27*(7), 1017–1018.

Grassi, E. (2017). MatrixRider.

Han, H., H. Shim, D. Shin, J. E. Shim, Y. Ko, J. Shin, H. Kim, A. Cho, E. Kim, T. Lee, H. Kim, K. Kim, S. Yang, D. Bae, A. Yun, S. Kim, C. Y. Kim, H. J. Cho, B. Kang,

S. Shin, and I. Lee (2015, June). TRRUST: a reference database of human transcriptional regulatory interactions. *Sci Rep 5*, 11432.

Hanahan, D. and R. A. Weinberg (2011, March). Hallmarks of cancer: the next generation. *Cell 144*(5), 646–674.

Hao, M., Y. Lu, Z. Yao, P. Wang, Z. Chen, L. Zhou, Q. Yang, X. Sang, K. Wang, and G. Cao (2024, October). The role of JNK signaling pathway in organ fibrosis. *Journal of Advanced Research*.

Hao, Y., S. Hao, E. Andersen-Nissen, W. M. Mauck, 3rd, S. Zheng, A. Butler, M. J. Lee, A. J. Wilk, C. Darby, M. Zager, P. Hoffman, M. Stoeckius, E. Papalexi, E. P. Mimitou, J. Jain, A. Srivastava, T. Stuart, L. M. Fleming, B. Yeung, A. J. Rogers, J. M. McElrath, C. A. Blish, R. Gottardo, P. Smibert, and R. Satija (2021, June). Integrated analysis of multimodal single-cell data. *Cell 184*(13), 3573–3587.e29.

Hasin, Y., M. Seldin, and A. Lusis (2017, May). Multi-omics approaches to disease. *Genome Biol 18*(1), 83.

Hawley, S. A., D. A. Pan, K. J. Mustard, L. Ross, J. Bain, A. M. Edelman, B. G. Frenguelli, and D. G. Hardie (2005, July). Calmodulin-dependent protein kinase kinase-beta is an alternative upstream kinase for AMP-activated protein kinase. *Cell Metab 2*(1), 9–19.

Hernandez-Armenta, C., D. Ochoa, E. Gonçalves, J. Saez-Rodriguez, and P. Beltrao (2017, June). Benchmarking substrate-based kinase activity inference using phosphoproteomic data. *Bioinformatics 33*(12), 1845–1851.

Hijazi, M., R. Smith, V. Rajeeve, C. Bessant, and P. R. Cutillas (2020, April). Reconstructing kinase network topologies from phosphoproteomics data reveals cancer-associated rewiring. *Nat Biotechnol 38*(4), 493–502.

Hill, S. M., L. M. Heiser, T. Cokelaer, M. Unger, N. K. Nesser, D. E. Carlin, Y. Zhang, A. Sokolov, E. O. Paull, C. K. Wong, K. Graim, A. Bivol, H. Wang, F. Zhu, B. Afsari, L. V. Danilova, A. V. Favorov, W. S. Lee, D. Taylor, C. W. Hu, B. L. Long, D. P. Noren, A. J. Bisberg, HPN-DREAM Consortium, G. B. Mills, J. W. Gray, M. Kellen, T. Norman, S. Friend, A. A. Qutub, E. J. Fertig, Y. Guan, M. Song, J. M. Stuart, P. T. Spellman, H. Koeppl, G. Stolovitzky, J. Saez-Rodriguez, and S. Mukherjee (2016, April). Inferring causal molecular networks: empirical assessment through a community-based effort. *Nat Methods 13*(4), 310–318.

Hoffmann, A., G. Natoli, and G. Ghosh (2006, October). Transcriptional regulation via the NF-kappaB signaling module. *Oncogene 25*(51), 6706–6716.

Holland, C. H., J. Tanevski, J. Perales-Patón, J. Gleixner, M. P. Kumar, E. Mereu, B. A. Joughin, O. Stegle, D. A. Lauffenburger, H. Heyn, B. Szalai, and J. Saez-Rodriguez (2020, February). Robustness and applicability of transcription factor and pathway analysis tools on single-cell RNA-seq data. *Genome Biol 21*(1), 36.

Hollander, M., D. A. Wolfe, and E. Chicken (2013, November). *Nonparametric Statistical Methods: Hollander/nonparametric statistical methods* (3 ed.). Wiley Series in Probability and Statistics. Hoboken, NJ: Wiley-Blackwell.

Holmes, S. and W. Huber (2019, February). *Modern Statistics for Modern Biology*. Cambridge University Press.

Horn, H., E. M. Schoof, J. Kim, X. Robin, M. L. Miller, F. Diella, A. Palma, G. Cesareni, L. J. Jensen, and R. Linding (2014, June). KinomeXplorer: an integrated platform for kinome biology studies. *Nat Methods 11*(6), 603–604.

Hornbeck, P. V., J. M. Kornhauser, S. Tkachev, B. Zhang, E. Skrzypek, B. Murray, V. Latham, and M. Sullivan (2012, January). PhosphoSitePlus: a comprehensive resource for investigating the structure and function of experimentally determined post-translational modifications in man and mouse. *Nucleic Acids Res 40*(Database issue), D261–70.

Hu, J., H.-S. Rho, R. H. Newman, J. Zhang, H. Zhu, and J. Qian (2014, January). PhosphoNetworks: a database for human phosphorylation networks. *Bioinformatics 30*(1), 141–142.

Huang, T., X. Huang, B. Shi, and M. Yao (2019, August). GEREDB: Gene expression regulation database curated by mining abstracts from literature. *J Bioinform Comput Biol 17*(4), 1950024.

Huber, W., A. von Heydebreck, H. Sültmann, A. Poustka, and M. Vingron (2002). Variance stabilization applied to microarray data calibration and to the quantification of differential expression. *Bioinformatics 18 Suppl 1*, S96–104.

Huntley, R. P., T. Sawford, P. Mutowo-Meullenet, A. Shypitsyna, C. Bonilla, M. J. Martin, and C. O'Donovan (2015, January). The GOA database: gene ontology annotation updates for 2015. *Nucleic Acids Res 43*(Database issue), D1057–63.

Huynh-Thu, V. A., A. Irrthum, L. Wehenkel, and P. Geurts (2010, September). Inferring regulatory networks from expression data using tree-based methods. *PLoS One 5*(9).

Isbel, L., R. S. Grand, and D. Schübeler (2022, December). Generating specificity in genome regulation through transcription factor sensitivity to chromatin. *Nat Rev Genet 23*(12), 728–740.

Jensen, O. N. (2004, February). Modification-specific proteomics: characterization of post-translational modifications by mass spectrometry. *Curr Opin Chem Biol 8*(1), 33–41.

Jia, J., B. Bissa, L. Brecht, L. Allers, S. W. Choi, Y. Gu, M. Zbinden, M. R. Burge, G. Timmins, K. Hallows, C. Behrends, and V. Deretic (2020, August). AMPK is activated during lysosomal damage via a galectin-ubiquitin signal transduction system. *Autophagy 16*(8), 1550–1552.

Jiang, P., Y. Zhang, B. Ru, Y. Yang, T. Vu, R. Paul, A. Mirza, G. Altan-Bonnet, L. Liu, E. Ruppin, L. Wakefield, and K. W. Wucherpfennig (2021, October). Systematic investigation of cytokine signaling activity at the tissue and single-cell levels. *Nat Methods 18*(10), 1181–1191.

Johnson, J. L., T. M. Yaron, E. M. Huntsman, A. Kerelsky, J. Song, A. Regev, T.-Y. Lin, K. Liberatore, D. M. Cizin, B. M. Cohen, N. Vasan, Y. Ma, K. Krismer, J. T. Robles, B. van de Kooij, A. E. van Vlimmeren, N. Andrée-Busch, N. F. Käufer, M. V. Dorovkov, A. G. Ryazanov, Y. Takagi, E. R. Kastenhuber, M. D. Goncalves, B. D. Hopkins, O. Elemento, D. J. Taatjes, A. Maucuer, A. Yamashita, A. Degterev, M. Uduman, J. Lu, S. D. Landry, B. Zhang, I. Cossentino, R. Linding, J. Blenis, P. V. Hornbeck, B. E. Turk, M. B. Yaffe, and L. C. Cantley (2023, January). An atlas of substrate specificities for the human serine/threonine kinome. *Nature 613*(7945), 759–766.

Joyce, A. R. and B. Ø. Palsson (2006, March). The model organism as a system: integrating 'omics' data sets. *Nat Rev Mol Cell Biol 7*(3), 198–210.

Kandasamy, K., S. S. Mohan, R. Raju, S. Keerthikumar, G. S. S. Kumar, A. K. Venugopal, D. Telikicherla, J. D. Navarro, S. Mathivanan, C. Pecquet, S. K. Gollapudi, S. G. Tattikota, S. Mohan, H. Padhukasahasram, Y. Subbannayya, R. Goel, H. K. C. Jacob, J. Zhong, R. Sekhar, V. Nanjappa, L. Balakrishnan, R. Subbaiah, Y. L. Ramachandra, B. A. Rahiman, T. S. K. Prasad, J.-X. Lin, J. C. D. Houtman, S. Desiderio, J.-C. Renauld, S. N. Constantinescu, O. Ohara, T. Hirano, M. Kubo, S. Singh, P. Khatri, S. Draghici, G. D. Bader, C. Sander, W. J. Leonard, and A. Pandey (2010, January). NetPath: a public resource of curated signal transduction pathways. *Genome Biol 11*(1), R3.

Kanehisa, M., Y. Sato, M. Kawashima, M. Furumichi, and M. Tanabe (2016, January). KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Res 44*(D1), D457–62.

Karczewski, K. J. and M. P. Snyder (2018, May). Integrative omics for health and disease. *Nat Rev Genet 19*(5), 299–310.

Keenan, A. B., S. L. Jenkins, K. M. Jagodnik, S. Koplev, E. He, D. Torre, Z. Wang, A. B. Dohlman, M. C. Silverstein, A. Lachmann, M. V. Kuleshov, A. Ma'ayan, V. Stathias, R. Terryn, D. Cooper, M. Forlin, A. Koleti, D. Vidovic, C. Chung, S. C. Schürer, J. Vasiliauskas, M. Pilarczyk, B. Shamsaei, M. Fazel, Y. Ren, W. Niu, N. A. Clark, S. White, N. Mahi, L. Zhang, M. Kouril, J. F. Reichard, S. Sivaganesan, M. Medvedovic, J. Meller, R. J. Koch, M. R. Birtwistle, R. Iyengar, E. A. Sobie, E. U. Azeloglu, J. Kaye, J. Osterloh, K. Haston, J. Kalra, S. Finkbiener, J. Li, P. Milani, M. Adam, R. Escalante-Chong, K. Sachs, A. Lenail, D. Ramamoorthy, E. Fraenkel, G. Daigle, U. Hussain, A. Coye, J. Rothstein, D. Sareen, L. Ornelas, M. Banuelos, B. Mandefro, R. Ho, C. N. Svendsen, R. G. Lim, J. Stocksdale, M. S. Casale, T. G. Thompson, J. Wu, L. M. Thompson, V. Dardov, V. Venkatraman, A. Matlock, J. E. Van Eyk, J. D. Jaffe, M. Papanastasiou, A. Subramanian, T. R. Golub, S. D. Erickson, M. Fallahi-Sichani, M. Hafner, N. S. Gray, J.-R. Lin, C. E. Mills, J. L. Muhlich, M. Niepel, C. E. Shamu, E. H. Williams, D. Wrobel, P. K. Sorger, L. M. Heiser, J. W. Gray, J. E. Korkola, G. B. Mills, M. LaBarge, H. S. Feiler, M. A. Dane, E. Bucher, M. Nederlof, D. Sudar, S. Gross, D. F. Kilburn, R. Smith, K. Devlin, R. Margolis, L. Derr, A. Lee, and A. Pillai (2018, January). The library of integrated Network-Based cellular signatures NIH program: System-Level cataloging of human cells response to perturbations. *Cell Syst 6*(1), 13–24.

Keenan, A. B., D. Torre, A. Lachmann, A. K. Leong, M. L. Wojciechowicz, V. Utti, K. M. Jagodnik, E. Kropiwnicki, Z. Wang, and A. Ma'ayan (2019, July). ChEA3: transcription factor enrichment analysis by orthogonal omics integration. *Nucleic Acids Res 47*(W1), W212–W224.

Kerrien, S., B. Aranda, L. Breuza, A. Bridge, F. Broackes-Carter, C. Chen, M. Duesbury, M. Dumousseau, M. Feuermann, U. Hinz, C. Jandrasits, R. C. Jimenez, J. Khadake, U. Mahadevan, P. Masson, I. Pedruzzi, E. Pfeiffenberger, P. Porras, A. Raghunath, B. Roechert, S. Orchard, and H. Hermjakob (2012, January). The IntAct molecular interaction database in 2012. *Nucleic Acids Res 40*(Database issue), D841–6.

Khoury, G. A., R. C. Baliban, and C. A. Floudas (2011, September). Proteome-wide post-

translational modification statistics: frequency analysis and curation of the swiss-prot database. *Scientific Reports 1*(1), 1–5.

Kiekens, L., W. Van Loocke, S. Taveirne, S. Wahlen, E. Persyn, E. Van Ammel, Z. De Vos, P. Matthys, F. Van Nieuwerburgh, T. Taghon, P. Van Vlierberghe, B. Vandekerckhove, and G. Leclercq (2021, September). T-BET and EOMES accelerate and enhance functional differentiation of human natural killer cells. *Front Immunol 12*, 732511.

Kim, S. and J. Wysocka (2023, February). Deciphering the multi-scale, quantitative cis-regulatory code. *Mol Cell 83*(3), 373–392.

Kisseleva, T. and D. Brenner (2020, October). Molecular and cellular mechanisms of liver fibrosis and its regression. *Nature Reviews Gastroenterology & Hepatology 18*(3), 151–166.

Kitano, H. (2002, March). Systems biology: A brief overview. *Science 295*(5560), 1662–1664.

Klaeger, S., S. Heinzlmeir, M. Wilhelm, H. Polzer, B. Vick, P.-A. Koenig, M. Reinecke, B. Ruprecht, S. Petzoldt, C. Meng, J. Zecha, K. Reiter, H. Qiao, D. Helm, H. Koch, M. Schoof, G. Canevari, E. Casale, S. R. Depaolini, A. Feuchtinger, Z. Wu, T. Schmidt, L. Rueckert, W. Becker, J. Huenges, A.-K. Garz, B.-O. Gohlke, D. P. Zolg, G. Kayser, T. Vooder, R. Preissner, H. Hahne, N. Tõnisson, K. Kramer, K. Götze, F. Bassermann, J. Schlegl, H.-C. Ehrlich, S. Aiche, A. Walch, P. A. Greif, S. Schneider, E. R. Felder, J. Ruland, G. Médard, I. Jeremias, K. Spiekermann, and B. Kuster (2017, December). The target landscape of clinical kinase drugs. *Science 358*(6367).

Korotkevich, G., V. Sukhov, N. Budin, B. Shpak, M. N. Artyomov, and A. Sergushichev (2016, June). Fast gene set enrichment analysis.

Koveitypour, Z., F. Panahi, M. Vakilian, M. Peymani, F. Seyed Forootan, M. H. Nasr Esfahani, and K. Ghaedi (2019, December). Signaling pathways involved in colorectal cancer progression. *Cell Biosci 9*, 97.

Koyano, F., K. Okatsu, H. Kosako, Y. Tamura, E. Go, M. Kimura, Y. Kimura, H. Tsuchiya, H. Yoshihara, T. Hirokawa, T. Endo, E. A. Fon, J.-F. Trempe, Y. Saeki, K. Tanaka, and N. Matsuda (2014, June). Ubiquitin is phosphorylated by PINK1 to activate parkin. *Nature 510*(7503), 162–166.

Krasny, L. and P. H. Huang (2021, February). Data-independent acquisition mass spectrometry (DIA-MS) for proteomic applications in oncology. *Mol Omics 17*(1), 29–42.

Krug, K., P. Mertins, B. Zhang, P. Hornbeck, R. Raju, R. Ahmad, M. Szucs, F. Mundt, D. Forestier, J. Jane-Valbuena, H. Keshishian, M. A. Gillette, P. Tamayo, J. P. Mesirov, J. D. Jaffe, S. A. Carr, and D. R. Mani (2019, March). A curated resource for phosphosite-specific signature analysis. *Mol Cell Proteomics 18*(3), 576–593.

Kuleshov, M. V., Z. Xie, A. B. K. London, J. Yang, J. E. Evangelista, A. Lachmann, I. Shu, D. Torre, and A. Ma'ayan (2021, July). KEA3: improved kinase enrichment analysis via data integration. *Nucleic Acids Res 49*(W1), W304–W316.

Lambert, S. A., A. Jolma, L. F. Campitelli, P. K. Das, Y. Yin, M. Albu, X. Chen, J. Taipale, T. R. Hughes, and M. T. Weirauch (2018, February). The human transcription factors. *Cell 172*(4), 650–665.

Law, C. W., Y. Chen, W. Shi, and G. K. Smyth (2014, February). voom: Precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biol 15*(2), R29.

Lee, T. I. and R. A. Young (2013, March). Transcriptional regulation and its misregulation in disease. *Cell 152*(6), 1237–1251.

Linding, R., L. J. Jensen, A. Pasculescu, M. Olhovsky, K. Colwill, P. Bork, M. B. Yaffe, and T. Pawson (2008, January). NetworKIN: a resource for exploring cellular phosphorylation networks. *Nucleic Acids Res 36*(Database issue), D695–9.

Liu, A., P. Trairatphisan, E. Gjerga, A. Didangelos, J. Barratt, and J. Saez-Rodriguez (2019, November). From expression footprints to causal pathways: contextualizing large signaling networks with CARNIVAL. *npj Systems Biology and Applications 5*(1), 1–10.

Liu, Z.-P., C. Wu, H. Miao, and H. Wu (2015, September). RegNetwork: an integrated database of transcriptional and post-transcriptional regulatory networks in human and mouse. *Database (Oxford) 2015*.

Lo Surdo, P., M. Iannuccelli, S. Contino, L. Castagnoli, L. Licata, G. Cesareni, and L. Perfetto (2023, January). SIGNOR 3.0, the SIGnaling network open resource 3.0: 2022 update. *Nucleic Acids Res 51*(D1), D631–D637.

Lovering, R. C., P. Gaudet, M. L. Acencio, A. Ignatchenko, A. Jolma, O. Fornes, M. Kuiper, I. V. Kulakovskiy, A. Lægreid, M. J. Martin, and C. Logie (2021, October). A GO catalogue of human DNA-binding transcription factors. *Biochim Biophys Acta Gene Regul Mech 1864*(11-12), 194765.

Lowe, R., N. Shirley, M. Bleackley, S. Dolan, and T. Shafee (2017, May). Transcriptomics technologies. *PLoS Comput Biol 13*(5), e1005457.

Maier, T., M. Güell, and L. Serrano (2009, December). Correlation of mRNA and protein in complex biological samples. *FEBS Lett 583*(24), 3966–3973.

Maity, A. K., X. Hu, T. Zhu, and A. E. Teschendorff (2022, June). Inference of age-associated transcription factor regulatory activity changes in single cells. *Nat Aging 2*(6), 548–561.

Manning, G. (2005, December). Genomic overview of protein kinases. *WormBook*, 1–19.

Manson, S. M. (2001, August). Simplifying complexity: a review of complexity theory. *Geoforum 32*(3), 405–414.

Mari, T., K. Mösbauer, E. Wyler, M. Landthaler, C. Drosten, and M. Selbach (2022, June). In vitro Kinase-to-Phosphosite database (iKiP-DB) predicts kinase activity in phosphoproteomic datasets. *J Proteome Res 21*(6), 1575–1587.

Mathur, R., D. Rotroff, J. Ma, A. Shojaie, and A. Motsinger-Reif (2018, May). Gene set analysis methods: a systematic comparison. *BioData Min 11*, 8.

Melham, T. (2013, April). Modelling, abstraction, and computation in systems biology: A view from computer science. *Progress in Biophysics and Molecular Biology 111*(2–3), 129–136.

Melms, J. C., J. Biermann, H. Huang, Y. Wang, A. Nair, S. Tagore, I. Katsyv, A. F. Rendeiro, A. D. Amin, D. Schapiro, C. J. Frangieh, A. M. Luoma, A. Filliol, Y. Fang, H. Ravichandran, M. G. Clausi, G. A. Alba, M. Rogava, S. W. Chen, P. Ho, D. T. Montoro, A. E. Kornberg, A. S. Han, M. F. Bakhoum, N. Anandasabapathy, M. Suárez-Fariñas, S. F. Bakhoum, Y. Bram, A. Borczuk, X. V. Guo, J. H. Lefkowitch, C. Marboe, S. M. Lagana, A. Del Portillo, E. J. Tsai, E. Zorn, G. S. Markowitz, R. F. Schwabe, R. E. Schwartz, O. Elemento, A. Saqi, H. Hibshoosh, J. Que, and B. Izar (2021, July). A molecular single-cell lung atlas of lethal COVID-19. *Nature 595*(7865), 114–119.

Meng, F., L. Song, and W. Wang (2017, February). Metformin improves overall survival of colorectal cancer patients with diabetes: A Meta-Analysis. *J Diabetes Res 2017*, 5063239.

Metz, K. S., E. M. Deoudes, M. E. Berginski, I. Jimenez-Ruiz, B. A. Aksoy, J. Hammerbacher, S. M. Gomez, and D. H. Phanstiel (2018, September). Coral: Clear and customizable visualization of human kinome data. *Cell Syst 7*(3), 347–350.e1.

Mortazavi, A., B. A. Williams, K. McCue, L. Schaeffer, and B. Wold (2008, July). Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat Methods 5*(7), 621–628.

Motta, S. and F. Pappalardo (2012, October). Mathematical modeling of biological systems. *Briefings in Bioinformatics 14*(4), 411–422.

Mubeen, S., C. T. Hoyt, A. Gemünd, M. Hofmann-Apitius, H. Fröhlich, and D. Domingo-Fernández (2019, November). The impact of pathway database choice on statistical enrichment analysis and predictive modeling. *Front Genet 10*, 1203.

Müller-Dott, S., E. J. Jaehnig, K. P. Munchic, W. Jiang, T. M. Yaron-Barir, S. R. Savage, M. Garrido-Rodriguez, J. L. Johnson, A. Lussana, E. Petsalaki, J. T. Lei, A. Dugourd, K. Krug, L. C. Cantley, D. R. Mani, B. Zhang, and J. Saez-Rodriguez (2024, July). Comprehensive evaluation of phosphoproteomic-based kinase activity inference.

Müller-Dott, S., E. Tsirvouli, M. Vazquez, R. O. Ramirez Flores, P. Badia-I-Mompel, R. Fallegger, D. Türei, A. Lægreid, and J. Saez-Rodriguez (2023, November). Expanding the coverage of regulons from high-confidence prior knowledge for accurate estimation of transcription factor activities. *Nucleic Acids Res 51*(20), 10934–10949.

Nagaraj, N., J. R. Wisniewski, T. Geiger, J. Cox, M. Kircher, J. Kelso, S. Pääbo, and M. Mann (2011, November). Deep proteome and transcriptome mapping of a human cancer cell line. *Mol Syst Biol 7*, 548.

Neddens, J., M. Temmel, S. Flunkert, B. Kerschbaumer, C. Hoeller, T. Loeffler, V. Niederkofler, G. Daum, J. Attems, and B. Hutter-Paier (2018, June). Phosphorylation of different tau sites during progression of alzheimer's disease. *Acta Neuropathol Commun 6*(1), 52.

Needham, E. J., B. L. Parker, T. Burykin, D. E. James, and S. J. Humphrey (2019, January). Illuminating the dark phosphoproteome. *Sci Signal 12*(565).

Norel, R., J. J. Rice, and G. Stolovitzky (2011, October). The self-assessment trap: can we all be better than average? *Mol Syst Biol 7*, 537.

Ong, S.-E., B. Blagoev, I. Kratchmarova, D. B. Kristensen, H. Steen, A. Pandey, and M. Mann (2002, May). Stable isotope labeling by amino acids in cell culture, silac, as a simple and accurate approach to expression proteomics. *Molecular amp; Cellular Proteomics 1*(5), 376–386.

Ozsolak, F. and P. M. Milos (2011, February). RNA sequencing: advances, challenges and opportunities. *Nat Rev Genet 12*(2), 87–98.

Paull, E. O., D. E. Carlin, M. Niepel, P. K. Sorger, D. Haussler, and J. M. Stuart (2013, November). Discovering causal pathways linking genomic events to transcriptional states using tied diffusion through interacting events (TieDIE). *Bioinformatics 29*(21), 2757–2764.

Pavlopoulos, G. A., M. Secrier, C. N. Moschopoulos, T. G. Soldatos, S. Kossida, J. Aerts, R. Schneider, and P. G. Bagos (2011, April). Using graph theory to analyze biological networks. *BioData Min 4*, 10.

Paysan-Lafosse, T., M. Blum, S. Chuguransky, T. Grego, B. L. Pinto, G. A. Salazar, M. L. Bileschi, P. Bork, A. Bridge, L. Colwell, J. Gough, D. H. Haft, I. Letunić, A. Marchler-Bauer, H. Mi, D. A. Natale, C. A. Orengo, A. P. Pandurangan, C. Rivoire, C. J. A. Sigrist, I. Sillitoe, N. Thanki, P. D. Thomas, S. C. E. Tosatto, C. H. Wu, and A. Bateman (2023, January). InterPro in 2022. *Nucleic Acids Res 51*(D1), D418–D427.

Perfetto, L., L. Briganti, A. Calderone, A. Cerquone Perpetuini, M. Iannuccelli, F. Langone, L. Licata, M. Marinkovic, A. Mattioni, T. Pavlidou, D. Peluso, L. L. Petrilli, S. Pirrò, D. Posca, E. Santonico, A. Silvestri, F. Spada, L. Castagnoli, and G. Cesareni (2016, January). SIGNOR: a database of causal relationships between biological entities. *Nucleic Acids Res 44*(D1), D548–54.

Pernicova, I. and M. Korbonits (2014, March). Metformin–mode of action and clinical implications for diabetes and cancer. *Nat Rev Endocrinol 10*(3), 143–156.

Peters, B., S. E. Brenner, E. Wang, D. Slonim, and M. G. Kann (2018, November). Putting benchmarks in their rightful place: The heart of computational biology. *PLOS Computational Biology 14*(11), e1006494.

Pinkse, M. W. H., S. Mohammed, J. W. Gouw, B. van Breukelen, H. R. Vos, and A. J. R. Heck (2008, February). Highly robust, automated, and sensitive online TiO2-based phosphoproteomics applied to study endogenous phosphorylation in drosophila melanogaster. *J Proteome Res 7*(2), 687–697.

Praveen, P. and H. Fröhlich (2013, June). Boosting probabilistic graphical model inference by incorporating prior knowledge from multiple sources. *PLOS ONE 8*(6), e67410.

Puri, P., F. Mirshahi, O. Cheung, R. Natarajan, J. W. Maher, J. M. Kellum, and A. J. Sanyal (2008, February). Activation and dysregulation of the unfolded protein response in nonalcoholic fatty liver disease. *Gastroenterology 134*(2), 568–576.

Ramazi, S. and J. Zahiri (2021, January). Post-translational modifications in proteins: resources, tools and prediction methods. *Database 2021*.

Rena, G., D. G. Hardie, and E. R. Pearson (2017, August). The mechanisms of action of metformin. *Diabetologia 60*(9), 1577–1585.

Ritchie, M. E., B. Phipson, D. Wu, Y. Hu, C. W. Law, W. Shi, and G. K. Smyth (2015, April). limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res 43*(7), e47.

Robinson, M. D. and A. Oshlack (2010, March). A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biology 11*(3), 1–9.

Rodchenkov, I., O. Babur, A. Luna, B. A. Aksoy, J. V. Wong, D. Fong, M. Franz, M. C. Siper, M. Cheung, M. Wrana, H. Mistry, L. Mosier, J. Dlin, Q. Wen, C. O'Callaghan, W. Li, G. Elder, P. T. Smith, C. Dallago, E. Cerami, B. Gross, U. Dogrusoz, E. Demir, G. D. Bader, and C. Sander (2020, January). Pathway commons 2019 update: integration, analysis and exploration of pathway data. *Nucleic Acids Res 48*(D1), D489–D497.

Rodgers, G., C. Austin, J. Anderson, A. Pawlyk, C. Colvis, R. Margolis, and J. Baker (2018, May). Glimmers in illuminating the druggable genome. *Nat Rev Drug Discov 17*(5), 301–302.

Rodriguez-Mier, P., M. Garrido-Rodriguez, A. Gabor, and J. Saez-Rodriguez (2024, October). Unified knowledge-driven network inference from omics data.

Roumeliotis, T. I., S. P. Williams, E. Gonçalves, C. Alsinet, M. Del Castillo Velasco-Herrera, N. Aben, F. Z. Ghavidel, M. Michaut, M. Schubert, S. Price, J. C. Wright, L. Yu, M. Yang, R. Dienstmann, J. Guinney, P. Beltrao, A. Brazma, M. Pardo, O. Stegle, D. J. Adams, L. Wessels, J. Saez-Rodriguez, U. McDermott, and J. S. Choudhary (2017, August). Genomic determinants of protein abundance variation in colorectal cancer cells. *Cell Rep 20*(9), 2201–2214.

Saito, T. and M. Rehmsmeier (2015, March). The Precision-Recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. *PLOS ONE 10*(3), e0118432.

Salovska, B., E. Gao, S. Müller-Dott, W. Li, C. C. Cordon, S. Wang, A. Dugourd, G. Rosenberger, J. Saez-Rodriguez, and Y. Liu (2023, February). Phosphoproteomic analysis of metformin signaling in colorectal cancer cells elucidates mechanism of action and potential therapeutic opportunities. *Clin Transl Med 13*(2), e1179.

Sanghi, A., J. J. Gruber, A. Metwally, L. Jiang, W. Reynolds, J. Sunwoo, L. Orloff, H. Y. Chang, M. Kasowski, and M. P. Snyder (2021, September). Chromatin accessibility associates with protein-RNA correlation in human cancer. *Nat Commun 12*(1), 5732.

Savage, S. R. and B. Zhang (2020, July). Using phosphoproteomics data to understand cellular signaling: a comprehensive guide to bioinformatics resources. *Clin Proteomics 17*, 27.

Schubert, M., B. Klinger, M. Klünemann, A. Sieber, F. Uhlitz, S. Sauer, M. J. Garnett, N. Blüthgen, and J. Saez-Rodriguez (2018, January). Perturbation-response genes reveal signaling footprints in cancer gene expression. *Nat Commun 9*(1), 20.

Sender, R., S. Fuchs, and R. Milo (2016, August). Revised estimates for the number of human and bacteria cells in the body. *PLOS Biology 14*(8), e1002533.

Sharma, K., R. C. J. D'Souza, S. Tyanova, C. Schaab, J. R. Wiśniewski, J. Cox, and M. Mann (2014, September). Ultradeep human phosphoproteome reveals a distinct regulatory nature of tyr and Ser/Thr-based signaling. *Cell Rep 8*(5), 1583–1594.

Shaw, R. J., K. A. Lamia, D. Vasquez, S.-H. Koo, N. Bardeesy, R. A. Depinho, M. Montminy, and L. C. Cantley (2005, December). The kinase LKB1 mediates glucose homeostasis in liver and therapeutic effects of metformin. *Science 310*(5754), 1642–1646.

Soto, L. F., Z. Li, C. S. Santoso, A. Berenson, I. Ho, V. X. Shen, S. Yuan, and J. I. Fuxman Bass (2022, February). Compendium of human transcription factor effector domains. *Mol Cell 82*(3), 514–526.

Stark, R., M. Grzelak, and J. Hadfield (2019, November). RNA sequencing: the teenage years. *Nat Rev Genet 20*(11), 631–656.

Stein, B. D., D. Calzolari, K. Hellberg, Y. S. Hu, L. He, C.-M. Hung, E. Q. Toyama, D. S. Ross, B. F. Lillemeier, L. C. Cantley, J. R. Yates, 3rd, and R. J. Shaw (2019, December). Quantitative in vivo proteomics of metformin response in liver reveals AMPK-Dependent and -independent signaling networks. *Cell Rep 29*(10), 3331–3348.e7.

Subramanian, A., P. Tamayo, V. K. Mootha, S. Mukherjee, B. L. Ebert, M. A. Gillette, A. Paulovich, S. L. Pomeroy, T. R. Golub, E. S. Lander, and J. P. Mesirov (2005, Oc-

tober). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A 102*(43), 15545–15550.

Sugiyama, N., H. Imamura, and Y. Ishihama (2019, July). Large-scale discovery of substrates of the human kinome. *Sci Rep 9*(1), 10503.

Sun, Y., T. Zhang, B. Lu, X. Li, and L. Jiang (2023, April). Application of cofactors in the regulation of microbial metabolism: A state of the art review. *Frontiers in Microbiology 14*.

Szalai, B. and J. Saez-Rodriguez (2020, December). Why do pathway methods work better than they should? *FEBS Lett 594*(24), 4189–4200.

Szklarczyk, D., A. L. Gable, D. Lyon, A. Junge, S. Wyder, J. Huerta-Cepas, M. Simonovic, N. T. Doncheva, J. H. Morris, P. Bork, L. J. Jensen, and C. v. Mering (2019, January). STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res 47*(D1), D607–D613.

Terfve, C. D. A., E. H. Wilkes, P. Casado, P. R. Cutillas, and J. Saez-Rodriguez (2015, September). Large-scale models of signal propagation in human cells derived from discovery phosphoproteomic data. *Nature Communications 6*(1), 1–11.

Timp, W. and G. Timp (2020, January). Beyond mass spectrometry, the next step in proteomics. *Sci Adv 6*(2), eaax8978.

Tsuchida, T. and S. L. Friedman (2017, May). Mechanisms of hepatic stellate cell activation. *Nature Reviews Gastroenterology & Hepatology 14*(7), 397–411.

Türei, D., T. Korcsmáros, and J. Saez-Rodriguez (2016, November). OmniPath: guidelines and gateway for literature-curated signaling pathway resources. *Nat Methods 13*(12), 966–967.

Türei, D., A. Valdeolivas, L. Gul, N. Palacio-Escat, M. Klein, O. Ivanova, M. Ölbei, A. Gábor, F. Theis, D. Módos, T. Korcsmáros, and J. Saez-Rodriguez (2021, March). Integrated intra- and intercellular signaling knowledge for multicellular omics analysis. *Mol Syst Biol 17*(3), e9923.

Turnbull, L., M.-T. Hütt, A. A. Ioannides, S. Kininmonth, R. Poeppl, K. Tockner, L. J. Bracken, S. Keesstra, L. Liu, R. Masselink, and A. J. Parsons (2018, June). Connectivity and complex systems: learning from a multi-disciplinary perspective. *Applied Network Science 3*(1).

Tutusaus, A., E. de Gregorio, B. Cucarull, H. Cristóbal, C. Aresté, I. Graupera, M. Coll, A. Colell, G. Gausdal, J. B. Lorens, P. García de Frutos, A. Morales, and M. Marí (2020). A functional role of GAS6/TAM in nonalcoholic steatohepatitis progression implicates AXL as therapeutic target. *Cell Mol Gastroenterol Hepatol 9*(3), 349–368.

UniProt Consortium (2023, January). UniProt: the universal protein knowledgebase in 2023. *Nucleic Acids Res 51*(D1), D523–D531.

Vanunu, O., O. Magger, E. Ruppin, T. Shlomi, and R. Sharan (2010, January). Associating genes and protein complexes with disease via network propagation. *PLoS Comput Biol 6*(1), e1000641.

Väremo, L., J. Nielsen, and I. Nookaew (2013, April). Enriching the gene set analysis of genome-wide data by incorporating directionality of gene expression and combining statistical hypotheses and methods. *Nucleic Acids Res 41*(8), 4378–4391.

Vazquez, M., M. Krallinger, F. Leitner, M. Kuiper, A. Valencia, and A. Laegreid (2022, January). ExTRI: Extraction of transcription regulation interactions from literature. *Biochim Biophys Acta Gene Regul Mech 1865*(1), 194778.

Wagih, O., J. Reimand, and G. D. Bader (2015, June). MIMP: predicting the impact of mutations on kinase-substrate phosphorylation. *Nat Methods 12*(6), 531–533.

Walsh, L. A., M. J. Alvarez, E. Y. Sabio, M. Reyngold, V. Makarov, S. Mukherjee, K.-W. Lee, A. Desrichard, Ş. Turcan, M. G. Dalin, V. K. Rajasekhar, S. Chen, L. T. Vahdat, A. Califano, and T. A. Chan (2017, August). An integrated systems biology approach identifies TRIM25 as a key determinant of breast cancer metastasis. *Cell Rep 20*(7), 1623–1640.

Wang, C., H. Xu, S. Lin, W. Deng, J. Zhou, Y. Zhang, Y. Shi, D. Peng, and Y. Xue (2020, February). GPS 5.0: An update on the prediction of kinase-specific phosphorylation sites in proteins. *Genomics Proteomics Bioinformatics 18*(1), 72–80.

Wang, Z., M. Gerstein, and M. Snyder (2009, January). RNA-Seq: a revolutionary tool for transcriptomics. *Nat Rev Genet 10*(1), 57–63.

Weber, L. M., W. Saelens, R. Cannoodt, C. Soneson, A. Hapfelmeier, P. P. Gardner, A.-L. Boulesteix, Y. Saeys, and M. D. Robinson (2019, June). Essential guidelines for computational method benchmarking. *Genome Biology 20*(1), 1–12.

Weidemüller, P., M. Kholmatov, E. Petsalaki, and J. B. Zaugg (2021, December). Transcription factors: Bridge between cell signaling and gene regulation. *Proteomics 21*(23-24), e2000034.

Wilkes, E. H., P. Casado, V. Rajeeve, and P. R. Cutillas (2017, September). Kinase activity ranking using phosphoproteomics data (KARP) quantifies the contribution of protein kinases to the regulation of cell viability. *Mol Cell Proteomics 16*(9), 1694–1704.

Wingender, E., T. Schoeps, M. Haubrock, M. Krull, and J. Dönitz (2018, January). TFClass: expanding the classification of human transcription factors to their mammalian orthologs. *Nucleic Acids Res 46*(D1), D343–D347.

Wiredja, D. D., M. Koyutürk, and M. R. Chance (2017, November). The KSEA app: a web-based tool for kinase activity inference from quantitative phosphoproteomics. *Bioinformatics 33*(21), 3489–3491.

Xue, X., J. Ren, X. Sun, Y. Gui, Y. Feng, B. Shu, W. Wei, Q. Lu, Y. Liang, W. He, J. Yang, and C. Dai (2018, July). Protein kinase Cα drives fibroblast activation and kidney fibrosis by stimulating autophagic flux. *J Biol Chem 293*(28), 11119–11130.

Yaron-Barir, T. M., B. A. Joughin, E. M. Huntsman, A. Kerelsky, D. M. Cizin, B. M. Cohen, A. Regev, J. Song, N. Vasan, T.-Y. Lin, J. M. Orozco, C. Schoenherr, C. Sagum, M. T. Bedford, R. M. Wynn, S.-C. Tso, D. T. Chuang, L. Li, S. S.-C. Li, P. Creixell, K. Krismer, M. Takegami, H. Lee, B. Zhang, J. Lu, I. Cossentino, S. D. Landry, M. Uduman, J. Blenis, O. Elemento, M. C. Frame, P. V. Hornbeck, L. C. Cantley, B. E. Turk, M. B. Yaffe, and J. L. Johnson (2024, May). The intrinsic substrate specificity of the human tyrosine kinome. *Nature 629*(8014), 1174–1181.

Yılmaz, S., M. Ayati, D. Schlatzer, A. E. Çiçek, M. R. Chance, and M. Koyutürk (2021, February). Robust inference of kinase activity using functional networks. *Nat Commun 12*(1), 1177.

Yu, H., P. M. Kim, E. Sprecher, V. Trifonov, and M. Gerstein (2007, April). The importance of bottlenecks in protein networks: correlation with gene essentiality and expression dynamics. *PLoS Comput Biol 3*(4), e59.

Zhao, B. S., I. A. Roundtree, and C. He (2017, January). Post-transcriptional gene regulation by mRNA modifications. *Nat Rev Mol Cell Biol 18*(1), 31–42.

Zhou, Z., W. Yeung, N. Gravel, M. Salcedo, S. Soleymani, S. Li, and N. Kannan (2023, February). Phosformer: an explainable transformer model for protein kinase-specific phosphorylation predictions. *Bioinformatics 39*(2).