NORTHWESTERN UNIVERSITY

Quantitative and Multiplexed Chemical-Genetic Phenotyping Reveals the Architecture of the Proteostasis Network in Breast Cancer

A DISSERTATION

SUBMITTED TO THE GRADUATE SCHOOL
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

for the degree

DOCTOR OF PHILOSOPHY

Field of Life Sciences

By

Sonia Brockway

EVANSTON, ILLINOIS

December 2020

ABSTRACT

Quantitative and Multiplexed Chemical-Genetic Phenotyping
Reveals the Architecture of the Proteostasis Network in Breast Cancer

Sonia Brockway

The protein homeostasis (proteostasis) network, a critical cytoprotective system that restores homeostasis in response to molecular stress, comprises distinct pathways, including the heat-shock response, unfolded protein response, oxidative stress response, and autophagy. These distinct pathways are all co-opted by tumor cells to cope with cancer-associated stress, and their activation in many cancer types is associated with decreased patient survival. The pathways, however, are typically studied in isolation, and the functional relationship within and between them remains largely unexplored. Moreover, there are few therapies that target these pathways. Thus, we sought a strategy that would simultaneously identify drug vulnerabilities dependent on the activation status of individual proteostasis factors and provide a better understanding of the functional relationship within and between the subnetworks of this complex biological system. Chemical-genetic interaction profiling in model organisms has proven powerful in providing insights into compound mechanism of action, gene function, and the relationship between networks of biological processes. However, identifying chemical-genetic interactions in mammalian systems is currently limited to low-throughput or computational methods.

Here, we develop Quantitative and Multiplexed Analysis of Phenotype by Sequencing (QMAP-Seq), which leverages next-generation sequencing for pooled high-throughput chemical-genetic profiling in mammalian cells. Using minimal automation, we treat pools of 60 cell types—comprising 12 genetic perturbations in critical proteostasis factors and their isogenic controls in five cell lines—with 1,440 compound-dose combinations, generating 86,400 chemical-genetic measurements. QMAP-Seq produces precise and accurate quantitative measures of acute drug response comparable to gold standard assays, but with increased throughput at lower cost. Moreover, QMAP-Seq reveals clinically actionable drug

vulnerabilities, including a synthetic lethal relationship between the proteasome inhibitor, Carfilzomib, and knockout of diverse proteostasis factors. Our method uncovers functional relationships involving these proteostasis factors, such as a high degree of similarity between the chemical-genetic profiles of HSF1 and HSF2, two heat-shock transcription factors with a previously unresolved functional relationship. Thus, QMAP-Seq provides a broadly accessible and scalable strategy for chemical-genetic profiling in mammalian cells, as we demonstrate for the proteostasis network.

ACKNOWLEDGMENTS

First, I would like to thank my advisor, Dr. Marc Mendillo. Marc took a chance by allowing me to join his lab after I made the difficult decision to leave my previous lab. I am grateful for the opportunity to be his first Ph.D. student and could not imagine a better scientific mentor. Marc is full of great ideas and manages to strike the perfect balance between always making himself available without micromanaging. I appreciate that he trusted me to take the lead on a systems biology project, a field that had always intrigued me, but one for which I had no prior experience. Marc gave me the freedom to pursue my research ideas and also taught me to think more critically about my research. Thank you, Marc, for helping me grow into a better scientist during my time in the lab.

I would also like to thank the members of the Mendillo Lab, both past and present, for creating a scholarly and fun lab environment. In particular, I am grateful for Dr. Kyle Metz and Seesha Takagishi for welcoming me into the lab and graciously helping me locate reagents and supplies even when they were busy. Kyle's positive attitude during our lunch and caffeine runs was indispensable, especially during the inevitable setbacks and challenges that are part of the scientific process. In addition, I would like to thank Geng Wang for helping me launch the QMAP-Seq project. His enthusiasm and creativity motivated me, and I am thrilled that he is now pursuing his dream of obtaining his Ph.D. Also, thank you to Kyle and Daniel Ansel for keeping the lab running smoothly and never complaining about my large QMAP-Seq order requests.

In addition, I would like to acknowledge my thesis committee members, Dr. Jindan Yu, Dr. Jaehyuk Choi, and Dr. Huiping Liu. My thesis committee simultaneously challenged me and supported me, and for that, I am incredibly grateful. I value their scientific expertise and suggestions, and genuinely believe that they made me and my research better. My thesis committee's encouragement was critical to keeping me moving towards the Ph.D. finish line.

Thank you to everyone who contributed to the QMAP-Seq project, including Geng Wang, Jasen Jackson, David Amici, Seesha Takagishi, Dr. Matthew Clutter, Dr. Elizabeth Bartom, and Dr. Marc Mendillo.

Thank you also to Stacy Marshall, Emily Rendleman, Emma Clark, and Didi Zha for taking the time to train me on multiple next-generation sequencing instruments.

I would also like to thank the Driskill Graduate Training Program in Life Sciences for giving me the opportunity to earn my doctorate. I value the knowledge I acquired through my coursework and the opportunities the program provided to develop my scientific skills. In addition, I am grateful for the funding and training opportunities provided by the T32 Carcinogenesis Training Grant. Thank you also to the Katten Muchin Rosenman Travel Scholarship, Biochemistry and Molecular Genetics Retreat Poster Award, Graduate School Conference Travel Grant, and Driskill Graduate Program Travel Grant for the financial support that enabled me to give a talk and poster presentation at a Gordon Research Conference, which had long been a scientific goal of mine.

Finally, I would like to thank my friends and family. I am especially thankful for my graduate school friends, Carissa Ritner, Dr. Alia Zander, and Dr. Chelsey Spriggs, for our memorable birthday lunches and laughter shared. These fellow scientists lifted me up on the long road towards a Ph.D. I am also exceptionally grateful for my husband, Blake, and son, Aidan, for their many sacrifices and unwavering love and support during my time at Northwestern University. Thank you for your patience and understanding during this journey. I am deeply appreciative of my parents, Cherian and Marina, for instilling in me the value of education and motivating me to pursue my goals. Thank you for the strength and encouragement you have given me. I would also like to thank my brother, Zubin, for his friendship and for paving the way towards the first doctorate in our family despite being younger than me. Lastly, thank you to my extended family and my husband's family for cheering me on and taking a genuine interest in my work. It is gratifying to know that others find my research as interesting as I do. None of this would be possible without all of you.

To my five-year-old, *Aidan Rohan Brockway*, who climbed on my desk during dissertation writing and asked me to explain my “pictures” (figures) to him. May your curiosity know no bounds.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1: INTRODUCTION

1.1 Proteostasis network

The protein homeostasis (proteostasis) network encompasses a critical set of cellular stress response factors that maintain proper protein function from synthesis to folding to degradation (1). Diverse internal and environmental stressors such as protein misfolding, elevated temperature, reactive oxygen species, and metabolic stress activate distinct stress response transcriptional programs that act through a variety of proteostasis factors to upregulate cytoprotective genes, which restore homeostasis and promote cell survival (2) **(Figure 1)**. When cells experience a level or duration of stress that they cannot accommodate, cell death programs are activated instead (3). This multifaced network comprises the heat-shock response, unfolded protein response, oxidative stress response, and autophagy, among other pathways.

1.2 Heat-shock response

The heat-shock response is activated in response to proteotoxic stress, a condition caused by protein misfolding and aggregation in the cytoplasm. Under basal conditions, the transcription factor HSF1 is sequestered in the cytoplasm by its association with molecular chaperones. In response to stress, HSF1 undergoes trimerization, phosphorylation, and translocation to the nucleus. It binds to heat-shock elements on the promoters of molecular chaperone genes, such as HSP70, resulting in a rapid increase in their expression (4-7). Molecular chaperones are recruited to misfolded proteins, where they assist with protein folding and clearing protein aggregates.

HSF2, another member of the heat-shock transcription factor family, is relatively unexplored in comparison to HSF1. Whereas HSF1 is the master transcriptional regulator of the cytoprotective heat-shock response, HSF2 is thought to play a role in development and differentiation, particularly of the brain and reproductive organs (8). Unlike HSF1, which is a relatively stable protein that is constitutively and ubiquitously expressed, HSF2 is a short-lived protein and its expression varies across tissue types (9). HSF1 and HSF2 have been reported to colocalize on promoters during heat-shock (10), proteasome

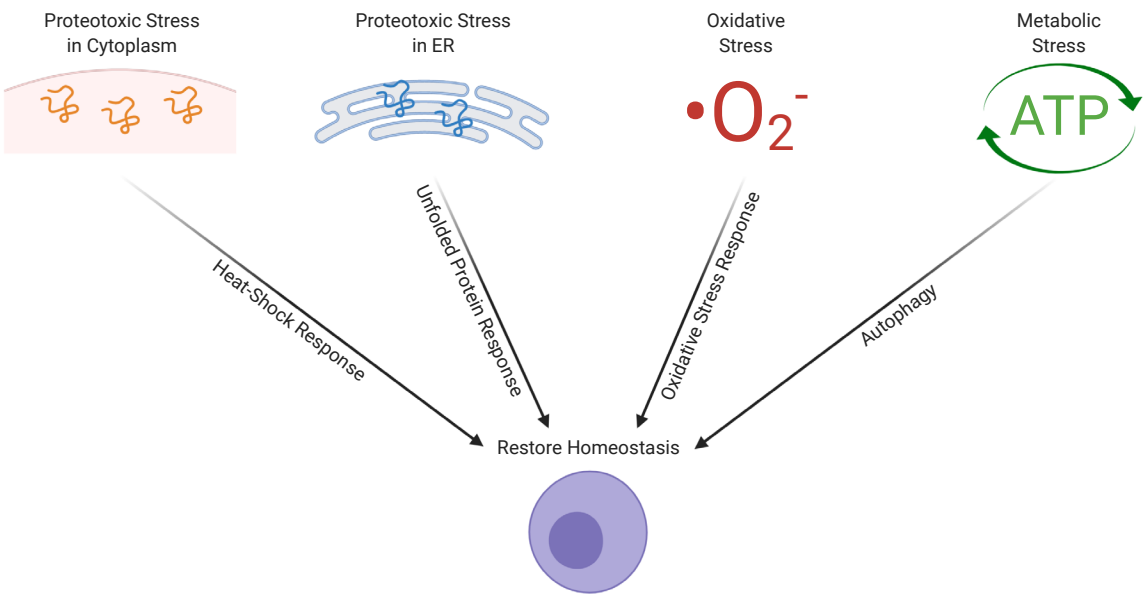


Figure 1. The subnetworks of the proteostasis network. Internal and environmental stress conditions activate stress response programs, including the heat-shock response, unfolded protein response, oxidative stress response, and autophagy, that act through proteostasis factors to restore cellular homeostasis.

inhibition (11), differentiation (10), and in nuclear stress bodies (12). Interestingly, HSF2 DNA binding was shown to be HSF1-dependent under stress conditions, while HSF1 activation was shown to be HSF2-dependent during development (12). Nevertheless, HSF1-HSF2 colocalization does not always indicate a cooperative relationship. For example, although HSF1 and HSF2 form complexes at the promoter of the heat-shock gene AIRAP during proteasome inhibition, HSF1 positively regulates AIRAP, while HSF2 negatively regulates AIRAP (13). In spermatocytes, HSF1 and HSF2 were reported to form complexes at physiological temperature, but displayed reduced complex formation during heat-shock (14), suggesting a context-dependent relationship between these two factors.

1.3 Unfolded protein response

The unfolded protein response operates under conditions of proteotoxic stress in the endoplasmic reticulum (ER) (15, 16). Because all secreted and transmembrane proteins are folded in the ER, the unfolded protein response is critical for normal cellular physiology. The ER membrane-embedded stress sensors IRE1, PERK, and ATF6 initiate the three branches of the unfolded protein response. When the endoribonuclease IRE1, which is encoded by the *ERN1* gene, senses ER stress, it splices *XBP1* mRNA to generate the active form of *XBP1* (17). Spliced *XBP1* is translated into a transcription factor that upregulates unfolded protein response target genes. When PERK senses ER stress, it phosphorylates eIF2α, which suppresses translation globally (18), but allows translation of the transcription factor ATF4 (19). ATF4 directly activates unfolded protein response target genes, including the transcription factor ATF3. The longer isoform of ATF3 represses transcription of its target genes, while the shorter isoform stimulates transcription. When the transmembrane transcription factor ATF6 senses ER stress, it is transported to the Golgi apparatus, where it is cleaved by proteases, releasing its cytosolic fragment (20). This fragment translocates to the nucleus and directly activates unfolded protein response target genes. Previous studies suggest that the three branches of the unfolded protein response have both overlapping and distinct transcriptional responses that can operate either collectively or independently (21, 22).

1.4 Oxidative stress response

The oxidative stress response is activated in response to reactive oxygen species, which oxidize cellular proteins, DNA, and lipids (23). The transcription factor NRF2, which is encoded by the *NFE2L2* gene, is the master transcriptional regulator of the oxidative stress response. Under basal conditions, NRF2 forms a complex with KEAP1, a substrate adaptor protein, and CUL3, an E3 ubiquitin ligase that targets NRF2 for ubiquitination and proteasomal degradation (23-26). Upon oxidative stress, cysteine residues on KEAP1 become oxidized, which leads to a conformational change in KEAP1. This conformational change alters the positioning of NRF2 relative to the ubiquitination machinery, such that NRF2 can no longer be ubiquitinated and degraded. This NRF2 molecule acts as a "suicide" substrate to inactivate KEAP1, which allows newly synthesized NRF2 to accumulate and translocate to the nucleus (27). NRF2 binds to antioxidant response elements together with MAF and induces the expression of antioxidant genes, which restore redox homeostasis (28-30).

1.5 Autophagy

Autophagy is one of the major protein degradation pathways in mammalian cells (31). Unlike the ubiquitin-proteasome degradation system that typically unfolds and degrades individual protein substrates, autophagy is well suited for the degradation of organelles as well as proteins that are not amenable to unfolding, including protein aggregates. Under normal conditions, low levels of basal autophagy maintain protein quality control by targeting unfolded proteins to the lysosome for degradation (32). However, autophagy is dramatically induced under conditions of acute metabolic stress. Autophagy degrades and recycles proteins, organelles, and other cellular components, which supplies the inputs for metabolic pathways and for synthesizing stress-response proteins (33, 34).

Autophagy is a complex, multistep process (35). In response to metabolic stress, the nutrient-sensing kinases AMPK and mTOR phosphorylate and activate a preinitiation complex that includes the kinase, ULK1. The preinitiation complex then phosphorylates and activates an initiation complex, which is also known as a Class III PI3K complex. The initiation complex generates PI3P, which dictates the site of nucleation of the isolation membrane. Next, the isolation membrane is elongated to form the

autophagosome, a closed double-membrane structure that surrounds cargo for degradation. Elongation is dependent on the conjugation of LC3 to the lipid phosphatidylethanolamine, which requires the E1-like activating enzyme, ATG7, the E2 ligase, ATG3, and the E3 ligase complex, ATG12/ATG5/ATG16L. The autophagosome then fuses with a lysosome to form an autolysosome. Finally, the cargo is degraded within the autolysosome by acidic hydrolases and released into the cytoplasm for recycling.

1.6 Proteostasis network in cancer

The proteostasis network is co-opted by tumor cells to cope with the myriad stresses associated with cancer, to the detriment of the organism. Notably, the heat-shock response (36-38), unfolded protein response (39), oxidative stress response (40), and autophagy (41) are activated in cancer and are associated with decreased patient survival. One of the mechanisms by which the proteostasis network is activated in cancer is through genomic alterations. Individual proteostasis factors are frequently amplified in patient tumors of varied histological origin **(Table 1)**. Most notably, HSF1, the master transcriptional regulator of the heat-shock response, is amplified in 41.1% of neuroendocrine prostate cancers (42), 34.9% of pancreatic cancers (43), and 25.7% of prostate adenocarcinomas (44). Proteostasis factors are also frequently mutated in a broad spectrum of patient tumor types **(Table 1)**. Activating gain-of-function mutations in NRF2, the master regulator of the oxidative stress response, occur in 14% of lung squamous cell carcinomas (45) and 6.1% of head and neck squamous cell carcinomas (46). Conversely, inactivating loss-of-function mutations in KEAP1, the negative regulator of NRF2, are found in 17.4% of lung adenocarcinomas (47) and 16.8% of non-small cell cancers (48).

Table 1. Genomic alterations of the proteostasis network in cancer.

Factor	Cancer Type	Mutation Frequency	Deletion Frequency	Amplification Frequency	Reference
HSF1	Neuroendocrine Prostate Cancer			41.1%	Trento/Cornell/Broad 2016
	Pancreatic Cancer		4.6%	34.9%	UTSW, Nat Commun 2015
	Prostate Adenocarcinoma	0.7%		25.7%	Fred Hutchinson CRC, Nat Med 2016
HSF2	Pancreatic Cancer		11.0%	1.8%	UTSW, Nat Commun 2015
	Neuroendocrine Prostate Cancer	0.9%		7.5%	Trento/Cornell/Broad 2016

	Melanoma		5.4%	2.7%	Roger 2016
IRE1 (*ERN1*)	Adenoid Cystic Carcinoma of the Breast			25.0%	MSKCC, J Pathol. 2015
	Neuroendocrine Prostate Cancer			22.4%	Trento/Cornell/Broad 2016
	Malignant Peripheral Nerve Sheath Tumor	13.3%		6.7%	MSKCC, Nat Genet 2014
XBP1	Malignant Peripheral Nerve Sheath Tumor		6.7%	6.7%	MSKCC, Nat Genet 2014
	Melanoma			10.8%	Roger 2016
	Neuroendocrine Prostate Cancer			7.5%	Trento/Cornell/Broad 2016
ATF3	Breast Cancer			22.6%	METABRIC, Nature 2012 & Nat Commun 2016
	Neuroendocrine Prostate Cancer			21.5%	Trento/Cornell/Broad 2016
	Melanoma		2.7%	10.8%	Roger 2016
ATF4	Melanoma	2.7%		13.5%	Roger 2016
	Pancreatic Cancer		2.8%	3.7%	UTSW, Nat Commun 2015
	Acral Melanoma	2.6%		2.6%	TGEN, Genome Res 2017
ATF6	Neuroendocrine Prostate Cancer			28.0%	Trento/Cornell/Broad 2016
	Breast Cancer			22.8%	METABRIC, Nature 2012 & Nat Commun 2016
	Bladder Urothelial Carcinoma	3.1%		15.0%	TCGA, Nature 2014
NRF2 (*NFE2L2*)	Lung Squamous Cell Carcinoma	14.0%		3.4%	TCGA, Nature 2012
	Head and Neck Squamous Cell Carcinoma	6.1%		5.7%	TCGA, Nature 2015
	Melanoma		2.7%	8.1%	Roger 2016
KEAP1	Lung Adenocarcinoma	17.4%	1.7%		TCGA, Nature 2014
	Non-Small Cell Cancer	16.8%	0.3%		MSK, Cancer Discovery 2017
	Neuroendocrine Prostate Cancer			16.8%	Trento/Cornell/Broad 2016
ATG7	Neuroendocrine Prostate Cancer	0.9%		19.6%	Trento/Cornell/Broad 2016
	Bladder Urothelial Carcinoma			11.8%	TCGA, Nature 2014
	Uterine Carcinosarcoma	9.1%			Johns Hopkins University, Nat Commun 2014

1.7 Rationale for using chemical-genetic profiling to investigate the proteostasis network

Because individual proteostasis factors are activated to varying degrees across tumors to cope with cancer-associated and drug-induced stress (36-41), but are not easily druggable (49), we reasoned that we could use chemical-genetic profiling to identify clinically relevant synthetic lethal chemical-genetic interactions dependent on the activation status of these factors. Furthermore, because the individual branches of the proteostasis network are typically studied in isolation, and the functional relationship within

and between the branches remains largely unexplored, we postulated that we could also use chemical-genetic profiling to provide functional insight into the proteostasis network. There is preliminary evidence of crosstalk between the various arms of the proteostasis network (16, 41, 50-53), but the functional relationship between the proteostasis factors has never been studied systematically.

1.8 Chemical-genetic profiling

Chemical-genetic interaction profiling in model organisms, such as yeast, has emerged as a powerful strategy to reveal functional insights into compounds, genes, and biological processes. In these studies, the mechanism of action of a compound can be deduced by comparing its chemical-genetic interaction profile (the quantitative landscape of the effects of a panel of individual genes on the efficacy of this particular compound) to the profiles of compounds with known cellular targets to identify the most similar profiles (54-58). Likewise, the function of a gene can be inferred by comparing its chemical-genetic interaction profile (the quantitative landscape of the effects of this particular gene on the efficacy of a panel of compounds) to the profiles of genes with known functions (59).

The development of highly specific and efficient genetic perturbation tools based on CRISPR-Cas9 has enabled similar types of chemical-genetic studies in mammalian systems, albeit at much smaller scales than in model organisms. Most often, chemical-genetic studies in mammalian systems involve genome-scale loss-of-function screens against one compound over the course of several weeks to identify drug targets (60) and define mechanisms of drug resistance (61, 62). Even focused studies interrogating limited numbers of chemical-genetic interactions can reveal critical insights. For example, one recent study demonstrated that the efficacy of a handful of clinical compounds was unaffected by knockout of their putative targets (63), highlighting the power of using chemical-genetic approaches to validate on-target activity of drug candidates. Despite their utility, these studies are low-throughput and thus limited to investigating small numbers of compounds.

1.9 Synthetic lethality and synthetic rescue

There is a growing interest in identifying synthetic lethal and synthetic rescue chemical-genetic interactions that can serve as the basis for cancer therapeutic strategies. Chemical-genetic synthetic lethality, a concept rooted in classical genetics (64), describes cell death resulting from the combination of a genetic variant and a chemical perturbation, where each individual perturbation is viable **(Figure 2a)**. By exploiting genetic variants (e.g., somatic mutations, copy number variations, chromosomal rearrangements, or gene expression changes) that differentiate tumor from normal tissue, synthetic lethal interactions provide a therapeutic window for selectively targeting cancer cells. The potential of synthetic lethality is best exemplified by the development and FDA approval of PARP inhibitors for patients with BRCA-mutated ovarian, breast, and prostate cancers (65, 66). There is also value in identifying synthetic rescue interactions, where a cytotoxic compound has reduced efficacy in the presence of a particular genetic variant, thus providing insights into drug resistance mechanisms (67) **(Figure 2b).**

1.10 Limitations of existing chemical-genetic approaches

The only existing strategies for identifying clinically relevant chemical-genetic interactions for more than a handful of genetic variants and compounds rely entirely on predictive approaches. Some of these predictions are based on genetic or chemical-genetic interactions identified in yeast (68, 69), while others, such as the Cancer Cell Line Encyclopedia (70), Genomics of Drug Sensitivity in Cancer (71-73), Cancer Therapeutics Response Portal (74-76), and PRISM (77, 78), are based on computational methods that correlate genetic and molecular features of human cancer cell lines with drug response. While these correlative approaches have been useful, they are limited by the fact that many features are rare and lack sufficient representation—or are not even present—in current cancer cell line collections, reducing the statistical power to detect significant correlations. Related to this, even correlations involving more common features are confounded by the multitude of additional features that also distinguish each cell line. Thus, these approaches still require direct experimental validation, ideally in a manner that tests individual features along with their corresponding isogenic controls in relevant mammalian models.

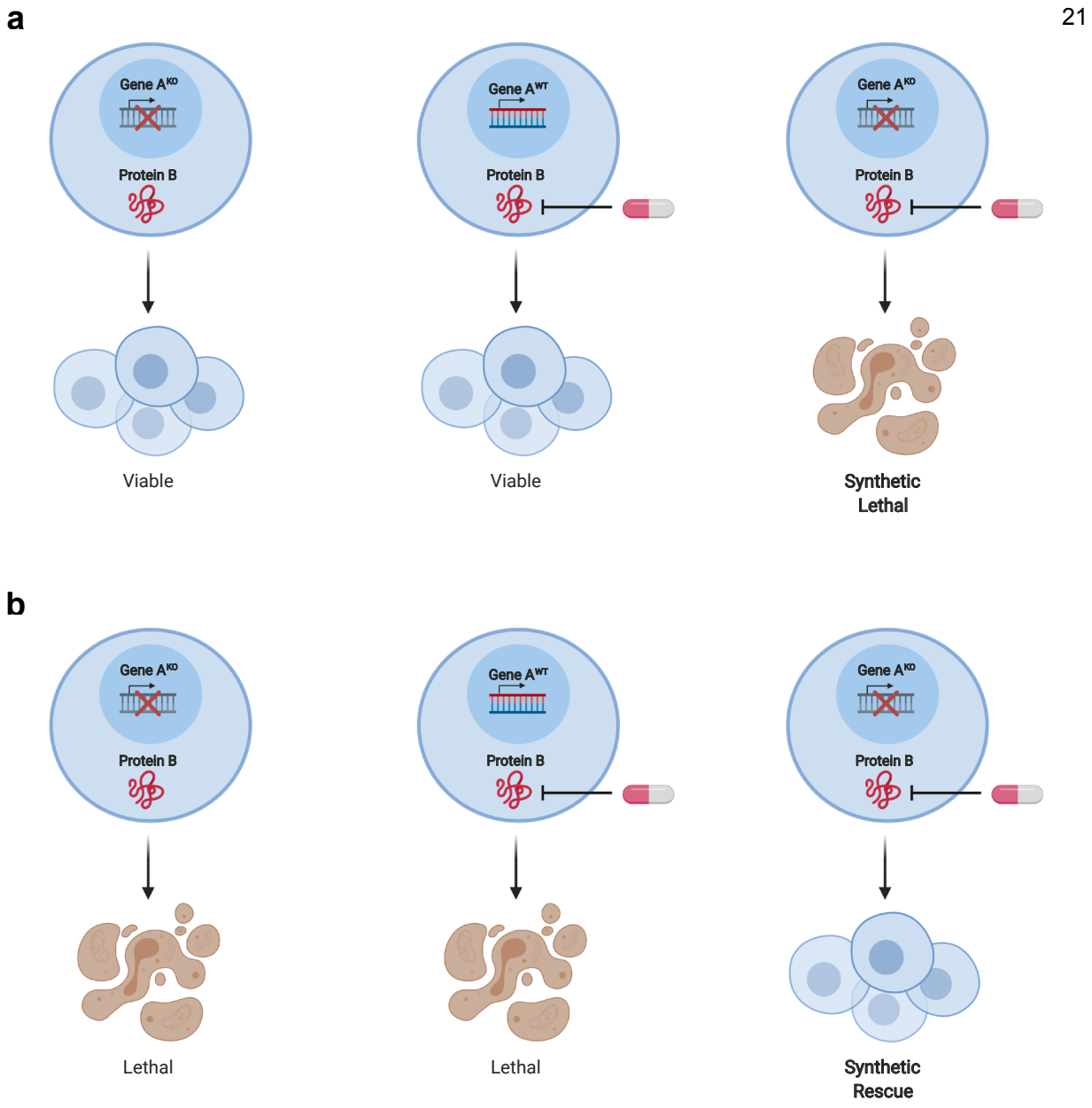


Figure 2. Synthetic lethal and synthetic rescue chemical-genetic interactions. a Schematic of a synthetic lethal chemical-genetic interaction, where the combination of a genetic perturbation and a chemical perturbation is lethal, but each perturbation alone is viable. **b** Schematic of a synthetic resuce chemical-genetic interaction, where the combination of a genetic perturbation and a chemical perturbation is viable, but each perturbation alone is lethal.

1.11 Development of QMAP-Seq

Here, to systematically and directly measure the contribution of individual proteostasis genes to acute drug response, we devise Quantitative and Multiplexed Analysis of Phenotype by Sequencing (QMAP-Seq). Individual cell lines and genetic perturbations are labeled with DNA barcodes, which enables pooled profiling of mixtures of cells in a single well. Unlike most chemical-genetic strategies in mammalian systems, QMAP-Seq is characterized by short-term compound treatment, which better recapitulates the timing of most high-throughput drug screening assays and enables testing of thousands of compounds in parallel. QMAP-Seq labels individual compound treatments with unique index sequences and employs massively parallel next-generation sequencing (NGS) to read out barcode abundance in response to compound treatment in a highly-multiplexed manner.

1.12 Breast cancer

Breast cancer is the leading type of cancer diagnosed in women in the United States, accounting for 30% of new female cancer cases (79). It is the second leading cause of cancer-related death within women in this country (79). Breast cancers are stratified by receptor status, which influences treatment options and prognosis. Estrogen receptor positive (ER+) breast cancers, which depend on estrogen for their growth, are treated with tamoxifen and generally have better prognosis. Human epidermal growth factor receptor 2 positive (HER2+) breast cancers, which involve amplification or overexpression of HER2, are treated with the monoclonal antibody trastuzumab or the small molecule inhibitor lapatinib. While ER+ and HER2+ breast cancers initially respond to these targeted therapies, these treatments are often hindered by the development of acquired resistance. Triple-negative (TN) breast cancers, which do not express ER, progesterone receptor (PR), or HER2, are highly heterogeneous in mutations and gene expression. No targeted therapies exist for TN breast cancers, and these cancers tend to be more aggressive and have worse prognosis.

In the present study, we utilize a diverse panel of cell lines representing ER+ (ZR-75-1), HER2+ (SKBR3), and TN (HCC-38, MDA-MB-231, BT-20) breast cancers because no single cell line would be suitable for studying all compounds due to the diversity in gene expression across breast cancer subtypes.

For example, the ER modulator tamoxifen is only appropriate to study in ER+ cells, while the HER2 inhibitor lapatinib is only suitable to study in HER2+ cells. Moreover, using a diverse panel of cell lines enables discovery of vulnerabilities unique to each breast cancer subtype.

1.13 Dissertation summary

To identify synthetic lethal and synthetic rescue chemical-genetic interactions dependent on the activation status of individual proteostasis factors and to map the proteostasis network in the context of breast cancer, we develop QMAP-Seq. We first perform QMAP-Seq with one cell line and demonstrate that it generates precise and accurate quantitative measures of compound efficacy that are concordant with established cell viability assays based on live cell imaging. We then expand QMAP-Seq to multiple cell lines, which enables the parallel measurement of 86,400 cell viability phenotypes in a single experiment. Altogether, we identify 60 sensitivity interactions and 124 resistance interactions and validate a subset of these interactions individually using an established metabolic-based cell viability assay. In particular, we observe hub drugs, such as the proteasome inhibitor, Carfilzomib, that synergize with loss of multiple nodes of the proteostasis network. Lastly, we utilize chemical-genetic profiles to systematically map the structure of the proteostasis network in breast cancer, revealing both known and previously unknown genetic relationships between the proteostasis factors. As expected, we observe a low degree of similarity between the chemical-genetic profiles for NRF2 and KEAP1 and a high degree of similarity between the profiles for IRE1 and XBP1. Shedding light on the functional relationship between HSF1 and HSF2, we observe that these factors share remarkably similar chemical-genetic profiles across the diverse compounds profiled in this study, providing evidence in support of a cooperative interaction in breast cancer. Overall, this work illustrates the power of systematic, high-throughput chemical-genetic profiling, as we demonstrate for the proteostasis network.

CHAPTER 2: RESULTS

The work from this chapter, including figures, has been reproduced with minor modifications from:

Brockway S, Wang G, Jackson JM, Amici DR, Takagishi SR, Clutter MR, Bartom ET, Mendillo ML. Quantitative and multiplexed chemical-genetic phenotyping in mammalian cells with QMAP-Seq. *Nat Commun.* (In Press).

2.1 Engineering barcoded breast cancer cell lines with inducible single gene knockouts.

Because the proteostasis network is heterogeneously activated in cancer, but is not easily druggable and is incompletely understood, we created a custom sgRNA library to disrupt a set of 10 genes that play pivotal roles in regulating the proteostasis network. While we were unable to include every possible proteostasis factor in our gene set, we selected some of the most critical factors involved in the heat-shock response (HSF1, HSF2), unfolded protein response (IRE1 (*ERN1*), XBP1, ATF3, ATF4, ATF6), oxidative stress response (NRF2 (*NFE2L2*), KEAP1), and autophagy (ATG7). As a control to validate that QMAP-Seq can detect known chemical-genetic interactions, our library also contained an sgRNA targeting SLC35F2, a solute carrier required for cellular uptake of the cytotoxic compound, YM155 (80). We engineered MDA-MB-231 triple-negative breast cancer cells with these 11 single gene knockouts and a pool of five non-targeting (NT) sgRNA controls. Because constitutive expression of Cas9 can result in off-target effects (81) and cell toxicity (82), we designed a system with doxycycline-inducible Cas9, providing temporal control over gene knockout **(Figure 3a)**. Cas9 was induced in a doxycycline dose-dependent manner **(Figure 3b)**. To enable future pooling and identification of multiple cell lines, we introduced unique 8 bp cell line barcode sequences downstream of the sgRNA within the lentiGuide-Puro plasmid **(Figure 3c)**.

To assess the efficacy of the sgRNAs, we performed Western blot analysis 96 hours after Cas9 induction and confirmed efficient whole population knockout of the proteostasis factors **(Figure 3d)**. Although we could not easily detect KEAP1 expression or knockout by Western blot, we observed the expected upregulation of NRF2 protein levels in the KEAP1 knockout cells **(Figure 3d)**. In addition, we

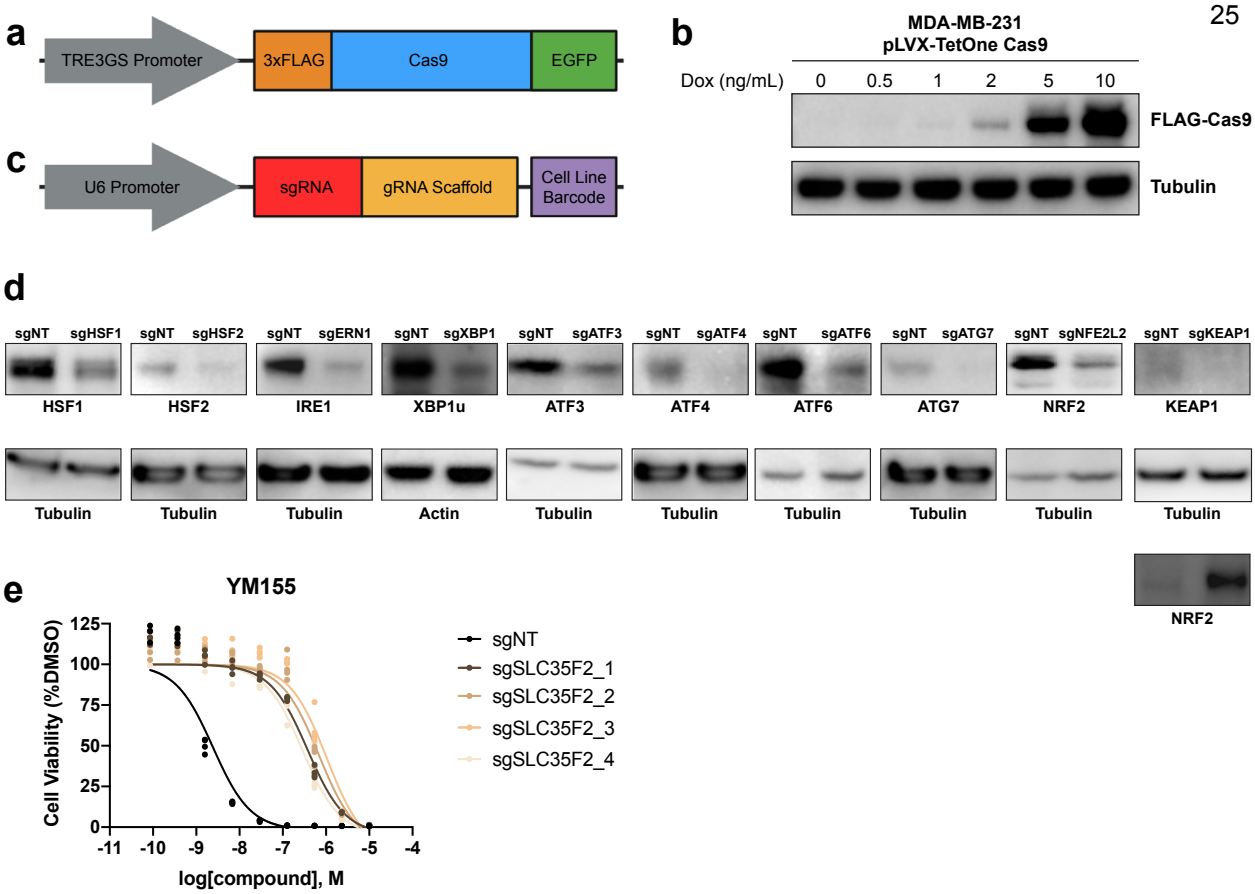


Figure 3. Engineering barcoded breast cancer cell lines with inducible single gene knockouts. **a** Schematic of doxycycline-inducible Cas9 plasmid (pLVX-TetOne Cas9). **b** MDA-MB-231 cells transduced with pLVX-TetOne Cas9 were treated with doxycycline for 48 hours. Cells were then harvested for Western blot analysis with FLAG antibody and Tubulin antibody (loading control). Results are representative of two independent experiments. **c** Schematic of lentiGuide-Puro plasmid engineered with an 8 bp cell line barcode downstream of the gRNA scaffold. **d** MDA-MB-231 cells that stably express pLVX-TetOne Cas9 were transduced with the indicated sgRNAs expressed from lentiGuide-Puro. Cells were treated with 10 ng/mL doxycycline for 96 hours to induce Cas9 expression and then harvested for Western blot analysis with the indicated antibodies. Experiment was performed once. **e** Dose-response curves as measured using CellTiter-Glo for MDA-MB-231 sgNT and sgSLC35F2 cells 72 hours after treatment with YM155. Each data point represents one of four biologically independent replicates.

used an ATP-based cell viability assay to confirm the ability of the sgRNAs that target SLC35F2 to confer resistance to YM155 **(Figure 3e)**.

2.2 QMAP-Seq generates precise and accurate quantitative measures of drug response.

We sought to develop and apply QMAP-Seq to quantify the response of a mixed pool of MDA-MB-231 cells possessing our panel of proteostasis factor knockouts to treatment with 89 compounds targeting diverse biological processes at four doses in duplicate **(Figure 4a)**. Our experimental workflow involved inducing Cas9 to initiate knockout, treating with either DMSO control or compound for 72 hours, and then preparing crude cell lysates. Previous studies have demonstrated the utility of spike-in standards for quantification when performing RNA-Seq (83) and ChIP-Seq (84, 85). To enable a quantitative assay, we introduced 293T cell spike-in standards composed of predetermined numbers of cells for each of five unique sgNT barcodes into each sample. Spike-in cell numbers were customized for each experiment to cover the expected range of cell numbers for any individual perturbation at the time of cell lysis (see “Materials and Methods” for details). We next amplified the 768 samples, corresponding to distinct compound-dose-replicate combinations, using unique sets of i5 and i7 indexed-primers **(Figure 4b, Table 2, Table 3)**. To facilitate Illumina sequencing, our PCR primers incorporated P5 and P7 adaptors complementary to flow cell oligos. To improve sequence diversity, we utilized a mix of P5 primers with varying stagger lengths. After PCR amplification, we pooled and purified the PCR products followed by Illumina sequencing with a single 164 bp read to sequence the sgRNA and cell line barcodes **(Figure 4b)**, a 6 bp index read to sequence the i7 index, and a 6 bp index read to sequence the i5 index.

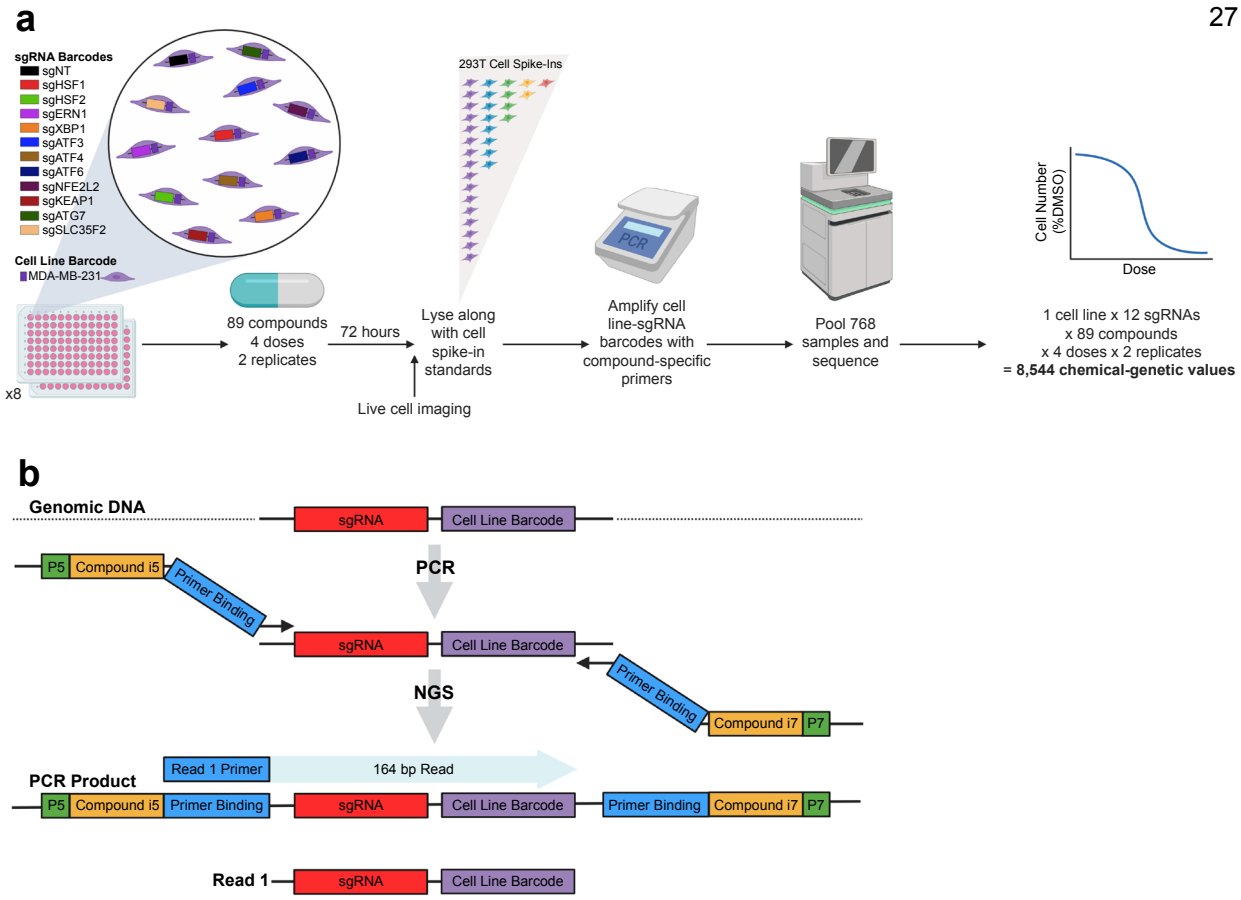


Figure 4. Experimental workflow for QMAP-Seq with one cell line. a Experimental workflow for QMAP-Seq with one cell line. **b** Schematic of QMAP-Seq library preparation using unique sets of i5/i7 indexed primers followed by next-generation sequencing of amplicons.

Table 2. P5 5001-5016 staggered PCR primer sequences. i5 Index sequences indicated in bold.

Name	Sequence
P5 5001 0nt stagger	AATGATACGGCGACCACCGAGATCTACAC**CGCCAT**ACACTCTTTCCCTACACGACGCTCTTCCGATCTTTGTGGAAAGGACGAAACACCG
P5 5001 1nt stagger	AATGATACGGCGACCACCGAGATCTACAC**CGCCAT**ACACTCTTTCCCTACACGACGCTCTTCCGATCTCTTGTGGAAAGGACGAAACACCG
P5 5001 2nt stagger	AATGATACGGCGACCACCGAGATCTACAC**CGCCAT**ACACTCTTTCCCTACACGACGCTCTTCCGATCTGCTTGTGGAAAGGACGAAACACCG
P5 5001 3nt stagger	AATGATACGGCGACCACCGAGATCTACAC**CGCCAT**ACACTCTTTCCCTACACGACGCTCTTCCGATCTAGCTTGTGGAAAGGACGAAACACCG
P5 5001 4nt stagger	AATGATACGGCGACCACCGAGATCTACAC**CGCCAT**ACACTCTTTCCCTACACGACGCTCTTCCGATCTCAACTTGTGGAAAGGACGAAACACCG
P5 5001 6nt stagger	AATGATACGGCGACCACCGAGATCTACAC**CGCCAT**ACACTCTTTCCCTACACGACGCTCTTCCGATCTTGCACCTTGTGGAAAGGACGAAACACCG
P5 5001 7nt stagger	AATGATACGGCGACCACCGAGATCTACAC**CGCCAT**ACACTCTTTCCCTACACGACGCTCTTCCGATCTACGCAACTTGTGGAAAGGACGAAACACCG
P5 5001 8nt stagger	AATGATACGGCGACCACCGAGATCTACAC**CGCCAT**ACACTCTTTCCCTACACGACGCTCTTCCGATCTGAAGACCCTTGTGGAAAGGACGAAACACCG
P5 5002 0nt stagger	AATGATACGGCGACCACCGAGATCTACAC**ATTTTA**ACACTCTTTCCCTACACGACGCTCTTCCGATCTTTGTGGAAAGGACGAAACACCG
P5 5002 1nt stagger	AATGATACGGCGACCACCGAGATCTACAC**ATTTTA**ACACTCTTTCCCTACACGACGCTCTTCCGATCTCTTGTGGAAAGGACGAAACACCG
P5 5002 2nt stagger	AATGATACGGCGACCACCGAGATCTACAC**ATTTTA**ACACTCTTTCCCTACACGACGCTCTTCCGATCTGCTTGTGGAAAGGACGAAACACCG
P5 5002 3nt stagger	AATGATACGGCGACCACCGAGATCTACAC**ATTTTA**ACACTCTTTCCCTACACGACGCTCTTCCGATCTAGCTTGTGGAAAGGACGAAACACCG
P5 5002 4nt stagger	AATGATACGGCGACCACCGAGATCTACAC**ATTTTA**ACACTCTTTCCCTACACGACGCTCTTCCGATCTCAACTTGTGGAAAGGACGAAACACCG
P5 5002 6nt stagger	AATGATACGGCGACCACCGAGATCTACAC**ATTTTA**ACACTCTTTCCCTACACGACGCTCTTCCGATCTTGCACCTTGTGGAAAGGACGAAACACCG
P5 5002 7nt stagger	AATGATACGGCGACCACCGAGATCTACAC**ATTTTA**ACACTCTTTCCCTACACGACGCTCTTCCGATCTACGCAACTTGTGGAAAGGACGAAACACCG
P5 5002 8nt stagger	AATGATACGGCGACCACCGAGATCTACAC**ATTTTA**ACACTCTTTCCCTACACGACGCTCTTCCGATCTGAAGACCCTTGTGGAAAGGACGAAACACCG
P5 5003 0nt stagger	AATGATACGGCGACCACCGAGATCTACAC**GCAACG**ACACTCTTTCCCTACACGACGCTCTTCCGATCTTTGTGGAAAGGACGAAACACCG
P5 5003 1nt stagger	AATGATACGGCGACCACCGAGATCTACAC**GCAACG**ACACTCTTTCCCTACACGACGCTCTTCCGATCTCTTGTGGAAAGGACGAAACACCG
P5 5003 2nt stagger	AATGATACGGCGACCACCGAGATCTACAC**GCAACG**ACACTCTTTCCCTACACGACGCTCTTCCGATCTGCTTGTGGAAAGGACGAAACACCG
P5 5003 3nt stagger	AATGATACGGCGACCACCGAGATCTACAC**GCAACG**ACACTCTTTCCCTACACGACGCTCTTCCGATCTAGCTTGTGGAAAGGACGAAACACCG
P5 5003 4nt stagger	AATGATACGGCGACCACCGAGATCTACAC**GCAACG**ACACTCTTTCCCTACACGACGCTCTTCCGATCTCAACTTGTGGAAAGGACGAAACACCG
P5 5003 6nt stagger	AATGATACGGCGACCACCGAGATCTACAC**GCAACG**ACACTCTTTCCCTACACGACGCTCTTCCGATCTTGCACCTTGTGGAAAGGACGAAACACCG
P5 5003 7nt stagger	AATGATACGGCGACCACCGAGATCTACAC**GCAACG**ACACTCTTTCCCTACACGACGCTCTTCCGATCTACGCAACTTGTGGAAAGGACGAAACACCG
P5 5003 8nt stagger	AATGATACGGCGACCACCGAGATCTACAC**GCAACG**ACACTCTTTCCCTACACGACGCTCTTCCGATCTGAAGACCCTTGTGGAAAGGACGAAACACCG
P5 5004 0nt stagger	AATGATACGGCGACCACCGAGATCTACAC**TAGGGC**ACACTCTTTCCCTACACGACGCTCTTCCGATCTTTGTGGAAAGGACGAAACACCG
P5 5004 1nt stagger	AATGATACGGCGACCACCGAGATCTACAC**TAGGGC**ACACTCTTTCCCTACACGACGCTCTTCCGATCTCTTGTGGAAAGGACGAAACACCG
P5 5004 2nt stagger	AATGATACGGCGACCACCGAGATCTACAC**TAGGGC**ACACTCTTTCCCTACACGACGCTCTTCCGATCTGCTTGTGGAAAGGACGAAACACCG
P5 5004 3nt stagger	AATGATACGGCGACCACCGAGATCTACAC**TAGGGC**ACACTCTTTCCCTACACGACGCTCTTCCGATCTAGCTTGTGGAAAGGACGAAACACCG
P5 5004 4nt stagger	AATGATACGGCGACCACCGAGATCTACAC**TAGGGC**ACACTCTTTCCCTACACGACGCTCTTCCGATCTCAACTTGTGGAAAGGACGAAACACCG
P5 5004 6nt stagger	AATGATACGGCGACCACCGAGATCTACAC**TAGGGC**ACACTCTTTCCCTACACGACGCTCTTCCGATCTTGCACCTTGTGGAAAGGACGAAACACCG
P5 5004 7nt stagger	AATGATACGGCGACCACCGAGATCTACAC**TAGGGC**ACACTCTTTCCCTACACGACGCTCTTCCGATCTACGCAACTTGTGGAAAGGACGAAACACCG
P5 5004 8nt stagger	AATGATACGGCGACCACCGAGATCTACAC**TAGGGC**ACACTCTTTCCCTACACGACGCTCTTCCGATCTGAAGACCCTTGTGGAAAGGACGAAACACCG
P5 5005 0nt stagger	AATGATACGGCGACCACCGAGATCTACAC**GCCCTA**ACACTCTTTCCCTACACGACGCTCTTCCGATCTTTGTGGAAAGGACGAAACACCG
P5 5005 1nt stagger	AATGATACGGCGACCACCGAGATCTACAC**GCCCTA**ACACTCTTTCCCTACACGACGCTCTTCCGATCTCTTGTGGAAAGGACGAAACACCG
P5 5005 2nt stagger	AATGATACGGCGACCACCGAGATCTACAC**GCCCTA**ACACTCTTTCCCTACACGACGCTCTTCCGATCTGCTTGTGGAAAGGACGAAACACCG
P5 5005 3nt stagger	AATGATACGGCGACCACCGAGATCTACAC**GCCCTA**ACACTCTTTCCCTACACGACGCTCTTCCGATCTAGCTTGTGGAAAGGACGAAACACCG
P5 5005 4nt stagger	AATGATACGGCGACCACCGAGATCTACAC**GCCCTA**ACACTCTTTCCCTACACGACGCTCTTCCGATCTCAACTTGTGGAAAGGACGAAACACCG
P5 5005 6nt stagger	AATGATACGGCGACCACCGAGATCTACAC**GCCCTA**ACACTCTTTCCCTACACGACGCTCTTCCGATCTTGCACCTTGTGGAAAGGACGAAACACCG
P5 5005 7nt stagger	AATGATACGGCGACCACCGAGATCTACAC**GCCCTA**ACACTCTTTCCCTACACGACGCTCTTCCGATCTACGCAACTTGTGGAAAGGACGAAACACCG
P5 5005 8nt stagger	AATGATACGGCGACCACCGAGATCTACAC**GCCCTA**ACACTCTTTCCCTACACGACGCTCTTCCGATCTGAAGACCCTTGTGGAAAGGACGAAACACCG
P5 5006 0nt stagger	AATGATACGGCGACCACCGAGATCTACAC**CGTTGC**ACACTCTTTCCCTACACGACGCTCTTCCGATCTTTGTGGAAAGGACGAAACACCG
P5 5006 1nt stagger	AATGATACGGCGACCACCGAGATCTACAC**CGTTGC**ACACTCTTTCCCTACACGACGCTCTTCCGATCTCTTGTGGAAAGGACGAAACACCG
P5 5006 2nt stagger	AATGATACGGCGACCACCGAGATCTACAC**CGTTGC**ACACTCTTTCCCTACACGACGCTCTTCCGATCTGCTTGTGGAAAGGACGAAACACCG
P5 5006 3nt stagger	AATGATACGGCGACCACCGAGATCTACAC**CGTTGC**ACACTCTTTCCCTACACGACGCTCTTCCGATCTAGCTTGTGGAAAGGACGAAACACCG
P5 5006 4nt stagger	AATGATACGGCGACCACCGAGATCTACAC**CGTTGC**ACACTCTTTCCCTACACGACGCTCTTCCGATCTCAACTTGTGGAAAGGACGAAACACCG
P5 5006 6nt stagger	AATGATACGGCGACCACCGAGATCTACAC**CGTTGC**ACACTCTTTCCCTACACGACGCTCTTCCGATCTTGCACCTTGTGGAAAGGACGAAACACCG
P5 5006 7nt stagger	AATGATACGGCGACCACCGAGATCTACAC**CGTTGC**ACACTCTTTCCCTACACGACGCTCTTCCGATCTACGCAACTTGTGGAAAGGACGAAACACCG
P5 5006 8nt stagger	AATGATACGGCGACCACCGAGATCTACAC**CGTTGC**ACACTCTTTCCCTACACGACGCTCTTCCGATCTGAAGACCCTTGTGGAAAGGACGAAACACCG
P5 5007 0nt stagger	AATGATACGGCGACCACCGAGATCTACAC**TAAAAT**ACACTCTTTCCCTACACGACGCTCTTCCGATCTTTGTGGAAAGGACGAAACACCG
P5 5007 1nt stagger	AATGATACGGCGACCACCGAGATCTACAC**TAAAAT**ACACTCTTTCCCTACACGACGCTCTTCCGATCTCTTGTGGAAAGGACGAAACACCG
P5 5007 2nt stagger	AATGATACGGCGACCACCGAGATCTACAC**TAAAAT**ACACTCTTTCCCTACACGACGCTCTTCCGATCTGCTTGTGGAAAGGACGAAACACCG
P5 5007 3nt stagger	AATGATACGGCGACCACCGAGATCTACAC**TAAAAT**ACACTCTTTCCCTACACGACGCTCTTCCGATCTAGCTTGTGGAAAGGACGAAACACCG
P5 5007 4nt stagger	AATGATACGGCGACCACCGAGATCTACAC**TAAAAT**ACACTCTTTCCCTACACGACGCTCTTCCGATCTCAACTTGTGGAAAGGACGAAACACCG
P5 5007 6nt stagger	AATGATACGGCGACCACCGAGATCTACAC**TAAAAT**ACACTCTTTCCCTACACGACGCTCTTCCGATCTTGCACCTTGTGGAAAGGACGAAACACCG
P5 5007 7nt stagger	AATGATACGGCGACCACCGAGATCTACAC**TAAAAT**ACACTCTTTCCCTACACGACGCTCTTCCGATCTACGCAACTTGTGGAAAGGACGAAACACCG
P5 5007 8nt stagger	AATGATACGGCGACCACCGAGATCTACAC**TAAAAT**ACACTCTTTCCCTACACGACGCTCTTCCGATCTGAAGACCCTTGTGGAAAGGACGAAACACCG
P5 5008 0nt stagger	AATGATACGGCGACCACCGAGATCTACAC**ATGGCG**ACACTCTTTCCCTACACGACGCTCTTCCGATCTTTGTGGAAAGGACGAAACACCG
P5 5008 1nt stagger	AATGATACGGCGACCACCGAGATCTACAC**ATGGCG**ACACTCTTTCCCTACACGACGCTCTTCCGATCTCTTGTGGAAAGGACGAAACACCG
P5 5008 2nt stagger	AATGATACGGCGACCACCGAGATCTACAC**ATGGCG**ACACTCTTTCCCTACACGACGCTCTTCCGATCTGCTTGTGGAAAGGACGAAACACCG
P5 5008 3nt stagger	AATGATACGGCGACCACCGAGATCTACAC**ATGGCG**ACACTCTTTCCCTACACGACGCTCTTCCGATCTAGCTTGTGGAAAGGACGAAACACCG
P5 5008 4nt stagger	AATGATACGGCGACCACCGAGATCTACAC**ATGGCG**ACACTCTTTCCCTACACGACGCTCTTCCGATCTCAACTTGTGGAAAGGACGAAACACCG
P5 5008 6nt stagger	AATGATACGGCGACCACCGAGATCTACAC**ATGGCG**ACACTCTTTCCCTACACGACGCTCTTCCGATCTTGCACCTTGTGGAAAGGACGAAACACCG
P5 5008 7nt stagger	AATGATACGGCGACCACCGAGATCTACAC**ATGGCG**ACACTCTTTCCCTACACGACGCTCTTCCGATCTACGCAACTTGTGGAAAGGACGAAACACCG
P5 5008 8nt stagger	AATGATACGGCGACCACCGAGATCTACAC**ATGGCG**ACACTCTTTCCCTACACGACGCTCTTCCGATCTGAAGACCCTTGTGGAAAGGACGAAACACCG
P5 5009 0nt stagger	AATGATACGGCGACCACCGAGATCTACAC**AGCGCC**ACACTCTTTCCCTACACGACGCTCTTCCGATCTTTGTGGAAAGGACGAAACACCG
P5 5009 1nt stagger	AATGATACGGCGACCACCGAGATCTACAC**AGCGCC**ACACTCTTTCCCTACACGACGCTCTTCCGATCTCTTGTGGAAAGGACGAAACACCG
P5 5009 2nt stagger	AATGATACGGCGACCACCGAGATCTACAC**AGCGCC**ACACTCTTTCCCTACACGACGCTCTTCCGATCTGCTTGTGGAAAGGACGAAACACCG
P5 5009 3nt stagger	AATGATACGGCGACCACCGAGATCTACAC**AGCGCC**ACACTCTTTCCCTACACGACGCTCTTCCGATCTAGCTTGTGGAAAGGACGAAACACCG
P5 5009 4nt stagger	AATGATACGGCGACCACCGAGATCTACAC**AGCGCC**ACACTCTTTCCCTACACGACGCTCTTCCGATCTCAACTTGTGGAAAGGACGAAACACCG
P5 5009 6nt stagger	AATGATACGGCGACCACCGAGATCTACAC**AGCGCC**ACACTCTTTCCCTACACGACGCTCTTCCGATCTTGCACCTTGTGGAAAGGACGAAACACCG
P5 5009 7nt stagger	AATGATACGGCGACCACCGAGATCTACAC**AGCGCC**ACACTCTTTCCCTACACGACGCTCTTCCGATCTACGCAACTTGTGGAAAGGACGAAACACCG
P5 5009 8nt stagger	AATGATACGGCGACCACCGAGATCTACAC**AGCGCC**ACACTCTTTCCCTACACGACGCTCTTCCGATCTGAAGACCCTTGTGGAAAGGACGAAACACCG
P5 5010 0nt stagger	AATGATACGGCGACCACCGAGATCTACAC**GCACAT**ACACTCTTTCCCTACACGACGCTCTTCCGATCTTTGTGGAAAGGACGAAACACCG
P5 5010 1nt stagger	AATGATACGGCGACCACCGAGATCTACAC**GCACAT**ACACTCTTTCCCTACACGACGCTCTTCCGATCTCTTGTGGAAAGGACGAAACACCG

P5 5010 2nt stagger	AATGATACGGCGACCACCGAGATCTACAC**GCACAT**ACACTCTTTCCCTACACGACGCTCTTCCGATCTGCTTGTGGAAAGGACGAAACACCG
P5 5010 3nt stagger	AATGATACGGCGACCACCGAGATCTACAC**GCACAT**ACACTCTTTCCCTACACGACGCTCTTCCGATCTAGCTTGTGGAAAGGACGAAACACCG
P5 5010 4nt stagger	AATGATACGGCGACCACCGAGATCTACAC**GCACAT**ACACTCTTTCCCTACACGACGCTCTTCCGATCTCAACTTGTGGAAAGGACGAAACACCG
P5 5010 6nt stagger	AATGATACGGCGACCACCGAGATCTACAC**GCACAT**ACACTCTTTCCCTACACGACGCTCTTCCGATCTTGCACCTTGTGGAAAGGACGAAACACCG
P5 5010 7nt stagger	AATGATACGGCGACCACCGAGATCTACAC**GCACAT**ACACTCTTTCCCTACACGACGCTCTTCCGATCTACGCAACTTGTGGAAAGGACGAAACACCG
P5 5010 8nt stagger	AATGATACGGCGACCACCGAGATCTACAC**GCACAT**ACACTCTTTCCCTACACGACGCTCTTCCGATCTGAAGACCCTTGTGGAAAGGACGAAACACCG
P5 5011 0nt stagger	AATGATACGGCGACCACCGAGATCTACAC**CGCGTA**ACACTCTTTCCCTACACGACGCTCTTCCGATCTTTGTGGAAAGGACGAAACACCG
P5 5011 1nt stagger	AATGATACGGCGACCACCGAGATCTACAC**CGCGTA**ACACTCTTTCCCTACACGACGCTCTTCCGATCTCTTGTGGAAAGGACGAAACACCG
P5 5011 2nt stagger	AATGATACGGCGACCACCGAGATCTACAC**CGCGTA**ACACTCTTTCCCTACACGACGCTCTTCCGATCTGCTTGTGGAAAGGACGAAACACCG
P5 5011 3nt stagger	AATGATACGGCGACCACCGAGATCTACAC**CGCGTA**ACACTCTTTCCCTACACGACGCTCTTCCGATCTAGCTTGTGGAAAGGACGAAACACCG
P5 5011 4nt stagger	AATGATACGGCGACCACCGAGATCTACAC**CGCGTA**ACACTCTTTCCCTACACGACGCTCTTCCGATCTCAACTTGTGGAAAGGACGAAACACCG
P5 5011 6nt stagger	AATGATACGGCGACCACCGAGATCTACAC**CGCGTA**ACACTCTTTCCCTACACGACGCTCTTCCGATCTTGCACCTTGTGGAAAGGACGAAACACCG
P5 5011 7nt stagger	AATGATACGGCGACCACCGAGATCTACAC**CGCGTA**ACACTCTTTCCCTACACGACGCTCTTCCGATCTACGCAACTTGTGGAAAGGACGAAACACCG
P5 5011 8nt stagger	AATGATACGGCGACCACCGAGATCTACAC**CGCGTA**ACACTCTTTCCCTACACGACGCTCTTCCGATCTGAAGACCCTTGTGGAAAGGACGAAACACCG
P5 5012 0nt stagger	AATGATACGGCGACCACCGAGATCTACAC**GATTAT**ACACTCTTTCCCTACACGACGCTCTTCCGATCTTTGTGGAAAGGACGAAACACCG
P5 5012 1nt stagger	AATGATACGGCGACCACCGAGATCTACAC**GATTAT**ACACTCTTTCCCTACACGACGCTCTTCCGATCTCTTGTGGAAAGGACGAAACACCG
P5 5012 2nt stagger	AATGATACGGCGACCACCGAGATCTACAC**GATTAT**ACACTCTTTCCCTACACGACGCTCTTCCGATCTGCTTGTGGAAAGGACGAAACACCG
P5 5012 3nt stagger	AATGATACGGCGACCACCGAGATCTACAC**GATTAT**ACACTCTTTCCCTACACGACGCTCTTCCGATCTAGCTTGTGGAAAGGACGAAACACCG
P5 5012 4nt stagger	AATGATACGGCGACCACCGAGATCTACAC**GATTAT**ACACTCTTTCCCTACACGACGCTCTTCCGATCTCAACTTGTGGAAAGGACGAAACACCG
P5 5012 6nt stagger	AATGATACGGCGACCACCGAGATCTACAC**GATTAT**ACACTCTTTCCCTACACGACGCTCTTCCGATCTTGCACCTTGTGGAAAGGACGAAACACCG
P5 5012 7nt stagger	AATGATACGGCGACCACCGAGATCTACAC**GATTAT**ACACTCTTTCCCTACACGACGCTCTTCCGATCTACGCAACTTGTGGAAAGGACGAAACACCG
P5 5012 8nt stagger	AATGATACGGCGACCACCGAGATCTACAC**GATTAT**ACACTCTTTCCCTACACGACGCTCTTCCGATCTGAAGACCCTTGTGGAAAGGACGAAACACCG
P5 5013 0nt stagger	AATGATACGGCGACCACCGAGATCTACAC**ATGACC**ACACTCTTTCCCTACACGACGCTCTTCCGATCTTTGTGGAAAGGACGAAACACCG
P5 5013 1nt stagger	AATGATACGGCGACCACCGAGATCTACAC**ATGACC**ACACTCTTTCCCTACACGACGCTCTTCCGATCTCTTGTGGAAAGGACGAAACACCG
P5 5013 2nt stagger	AATGATACGGCGACCACCGAGATCTACAC**ATGACC**ACACTCTTTCCCTACACGACGCTCTTCCGATCTGCTTGTGGAAAGGACGAAACACCG
P5 5013 3nt stagger	AATGATACGGCGACCACCGAGATCTACAC**ATGACC**ACACTCTTTCCCTACACGACGCTCTTCCGATCTAGCTTGTGGAAAGGACGAAACACCG
P5 5013 4nt stagger	AATGATACGGCGACCACCGAGATCTACAC**ATGACC**ACACTCTTTCCCTACACGACGCTCTTCCGATCTCAACTTGTGGAAAGGACGAAACACCG
P5 5013 6nt stagger	AATGATACGGCGACCACCGAGATCTACAC**ATGACC**ACACTCTTTCCCTACACGACGCTCTTCCGATCTTGCACCTTGTGGAAAGGACGAAACACCG
P5 5013 7nt stagger	AATGATACGGCGACCACCGAGATCTACAC**ATGACC**ACACTCTTTCCCTACACGACGCTCTTCCGATCTACGCAACTTGTGGAAAGGACGAAACACCG
P5 5013 8nt stagger	AATGATACGGCGACCACCGAGATCTACAC**ATGACC**ACACTCTTTCCCTACACGACGCTCTTCCGATCTGAAGACCCTTGTGGAAAGGACGAAACACCG
P5 5014 0nt stagger	AATGATACGGCGACCACCGAGATCTACAC**CTGATA**ACACTCTTTCCCTACACGACGCTCTTCCGATCTTTGTGGAAAGGACGAAACACCG
P5 5014 1nt stagger	AATGATACGGCGACCACCGAGATCTACAC**CTGATA**ACACTCTTTCCCTACACGACGCTCTTCCGATCTCTTGTGGAAAGGACGAAACACCG
P5 5014 2nt stagger	AATGATACGGCGACCACCGAGATCTACAC**CTGATA**ACACTCTTTCCCTACACGACGCTCTTCCGATCTGCTTGTGGAAAGGACGAAACACCG
P5 5014 3nt stagger	AATGATACGGCGACCACCGAGATCTACAC**CTGATA**ACACTCTTTCCCTACACGACGCTCTTCCGATCTAGCTTGTGGAAAGGACGAAACACCG
P5 5014 4nt stagger	AATGATACGGCGACCACCGAGATCTACAC**CTGATA**ACACTCTTTCCCTACACGACGCTCTTCCGATCTCAACTTGTGGAAAGGACGAAACACCG
P5 5014 6nt stagger	AATGATACGGCGACCACCGAGATCTACAC**CTGATA**ACACTCTTTCCCTACACGACGCTCTTCCGATCTTGCACCTTGTGGAAAGGACGAAACACCG
P5 5014 7nt stagger	AATGATACGGCGACCACCGAGATCTACAC**CTGATA**ACACTCTTTCCCTACACGACGCTCTTCCGATCTACGCAACTTGTGGAAAGGACGAAACACCG
P5 5014 8nt stagger	AATGATACGGCGACCACCGAGATCTACAC**CTGATA**ACACTCTTTCCCTACACGACGCTCTTCCGATCTGAAGACCCTTGTGGAAAGGACGAAACACCG
P5 5015 0nt stagger	AATGATACGGCGACCACCGAGATCTACAC**TCACGG**ACACTCTTTCCCTACACGACGCTCTTCCGATCTTTGTGGAAAGGACGAAACACCG
P5 5015 1nt stagger	AATGATACGGCGACCACCGAGATCTACAC**TCACGG**ACACTCTTTCCCTACACGACGCTCTTCCGATCTCTTGTGGAAAGGACGAAACACCG
P5 5015 2nt stagger	AATGATACGGCGACCACCGAGATCTACAC**TCACGG**ACACTCTTTCCCTACACGACGCTCTTCCGATCTGCTTGTGGAAAGGACGAAACACCG
P5 5015 3nt stagger	AATGATACGGCGACCACCGAGATCTACAC**TCACGG**ACACTCTTTCCCTACACGACGCTCTTCCGATCTAGCTTGTGGAAAGGACGAAACACCG
P5 5015 4nt stagger	AATGATACGGCGACCACCGAGATCTACAC**TCACGG**ACACTCTTTCCCTACACGACGCTCTTCCGATCTCAACTTGTGGAAAGGACGAAACACCG
P5 5015 6nt stagger	AATGATACGGCGACCACCGAGATCTACAC**TCACGG**ACACTCTTTCCCTACACGACGCTCTTCCGATCTTGCACCTTGTGGAAAGGACGAAACACCG
P5 5015 7nt stagger	AATGATACGGCGACCACCGAGATCTACAC**TCACGG**ACACTCTTTCCCTACACGACGCTCTTCCGATCTACGCAACTTGTGGAAAGGACGAAACACCG
P5 5015 8nt stagger	AATGATACGGCGACCACCGAGATCTACAC**TCACGG**ACACTCTTTCCCTACACGACGCTCTTCCGATCTGAAGACCCTTGTGGAAAGGACGAAACACCG
P5 5016 0nt stagger	AATGATACGGCGACCACCGAGATCTACAC**TATTGG**ACACTCTTTCCCTACACGACGCTCTTCCGATCTTTGTGGAAAGGACGAAACACCG
P5 5016 1nt stagger	AATGATACGGCGACCACCGAGATCTACAC**TATTGG**ACACTCTTTCCCTACACGACGCTCTTCCGATCTCTTGTGGAAAGGACGAAACACCG
P5 5016 2nt stagger	AATGATACGGCGACCACCGAGATCTACAC**TATTGG**ACACTCTTTCCCTACACGACGCTCTTCCGATCTGCTTGTGGAAAGGACGAAACACCG
P5 5016 3nt stagger	AATGATACGGCGACCACCGAGATCTACAC**TATTGG**ACACTCTTTCCCTACACGACGCTCTTCCGATCTAGCTTGTGGAAAGGACGAAACACCG
P5 5016 4nt stagger	AATGATACGGCGACCACCGAGATCTACAC**TATTGG**ACACTCTTTCCCTACACGACGCTCTTCCGATCTCAACTTGTGGAAAGGACGAAACACCG
P5 5016 6nt stagger	AATGATACGGCGACCACCGAGATCTACAC**TATTGG**ACACTCTTTCCCTACACGACGCTCTTCCGATCTTGCACCTTGTGGAAAGGACGAAACACCG
P5 5016 7nt stagger	AATGATACGGCGACCACCGAGATCTACAC**TATTGG**ACACTCTTTCCCTACACGACGCTCTTCCGATCTACGCAACTTGTGGAAAGGACGAAACACCG
P5 5016 8nt stagger	AATGATACGGCGACCACCGAGATCTACAC**TATTGG**ACACTCTTTCCCTACACGACGCTCTTCCGATCTGAAGACCCTTGTGGAAAGGACGAAACACCG

Table 3. P7 7001-7096 PCR primer sequences. i7 Index sequences indicated in bold.

Name	Sequence
P7_A01_7001	CAAGCAGAAGACGGCATACGAGAT**ACGCTG**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCCTGCACTGT
P7_B01_7002	CAAGCAGAAGACGGCATACGAGAT**GTGATC**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCCTGCACTGT
P7_C01_7003	CAAGCAGAAGACGGCATACGAGAT**ACTGGT**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCCTGCACTGT
P7_D01_7004	CAAGCAGAAGACGGCATACGAGAT**CGTGCA**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCCTGCACTGT
P7_E01_7005	CAAGCAGAAGACGGCATACGAGAT**TAATGT**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCCTGCACTGT
P7_F01_7006	CAAGCAGAAGACGGCATACGAGAT**CTACAC**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCCTGCACTGT
P7_G01_7007	CAAGCAGAAGACGGCATACGAGAT**GACTAG**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCCTGCACTGT
P7_H01_7008	CAAGCAGAAGACGGCATACGAGAT**TGCACA**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCCTGCACTGT
P7_A02_7009	CAAGCAGAAGACGGCATACGAGAT**TCCTGA**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCCTGCACTGT
P7_B02_7010	CAAGCAGAAGACGGCATACGAGAT**TAACCG**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCCTGCACTGT
P7_C02_7011	CAAGCAGAAGACGGCATACGAGAT**AGTTAA**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCCTGCACTGT
P7_D02_7012	CAAGCAGAAGACGGCATACGAGAT**GTTCAT**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCCTGCACTGT
P7_E02_7013	CAAGCAGAAGACGGCATACGAGAT**CTTAGG**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCCTGCACTGT
P7_F02_7014	CAAGCAGAAGACGGCATACGAGAT**CGGATT**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCCTGCACTGT
P7_G02_7015	CAAGCAGAAGACGGCATACGAGAT**GCAGCC**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCCTGCACTGT
P7_H02_7016	CAAGCAGAAGACGGCATACGAGAT**AAGGTA**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCCTGCACTGT
P7_A03_7017	CAAGCAGAAGACGGCATACGAGAT**TTAAGA**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCCTGCACTGT

P7_B03_7018	CAAGCAGAAGACGGCATACGAGAT**CCTGAC**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCCTGCACTGT
P7_C03_7019	CAAGCAGAAGACGGCATACGAGAT**CAGTAT**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCCTGCACTGT
P7_D03_7020	CAAGCAGAAGACGGCATACGAGAT**GACATA**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCCTGCACTGT
P7_E03_7021	CAAGCAGAAGACGGCATACGAGAT**GGACTC**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCCTGCACTGT
P7_F03_7022	CAAGCAGAAGACGGCATACGAGAT**ACCTCC**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCCTGCACTGT
P7_G03_7023	CAAGCAGAAGACGGCATACGAGAT**AGTGTG**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCCTGCACTGT
P7_H03_7024	CAAGCAGAAGACGGCATACGAGAT**TTGCGG**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCCTGCACTGT
P7_A04_7025	CAAGCAGAAGACGGCATACGAGAT**CATAAA**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCCTGCACTGT
P7_B04_7026	CAAGCAGAAGACGGCATACGAGAT**GGCGTT**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCCTGCACTGT
P7_C04_7027	CAAGCAGAAGACGGCATACGAGAT**TGCTTG**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCCTGCACTGT
P7_D04_7028	CAAGCAGAAGACGGCATACGAGAT**TCGAGC**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCCTGCACTGT
P7_E04_7029	CAAGCAGAAGACGGCATACGAGAT**ATTCGC**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCCTGCACTGT
P7_F04_7030	CAAGCAGAAGACGGCATACGAGAT**AATCCA**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCCTGCACTGT
P7_G04_7031	CAAGCAGAAGACGGCATACGAGAT**GTAGGT**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCCTGCACTGT
P7_H04_7032	CAAGCAGAAGACGGCATACGAGAT**CCTTCG**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCCTGCACTGT
P7_A05_7033	CAAGCAGAAGACGGCATACGAGAT**CTATCT**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCCTGCACTGT
P7_B05_7034	CAAGCAGAAGACGGCATACGAGAT**TACCAA**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCCTGCACTGT
P7_C05_7035	CAAGCAGAAGACGGCATACGAGAT**GGTAAC**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCCTGCACTGT
P7_D05_7036	CAAGCAGAAGACGGCATACGAGAT**GTTGCG**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCCTGCACTGT
P7_E05_7037	CAAGCAGAAGACGGCATACGAGAT**TCGCCA**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCCTGCACTGT
P7_F05_7038	CAAGCAGAAGACGGCATACGAGAT**CACGGT**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCCTGCACTGT
P7_G05_7039	CAAGCAGAAGACGGCATACGAGAT**CTGTTG**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCCTGCACTGT
P7_H05_7040	CAAGCAGAAGACGGCATACGAGAT**ACAATC**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCCTGCACTGT
P7_A06_7041	CAAGCAGAAGACGGCATACGAGAT**CGAGAG**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCCTGCACTGT
P7_B06_7042	CAAGCAGAAGACGGCATACGAGAT**GTGTCA**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCCTGCACTGT
P7_C06_7043	CAAGCAGAAGACGGCATACGAGAT**TGTCTT**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCCTGCACTGT
P7_D06_7044	CAAGCAGAAGACGGCATACGAGAT**ATCTGT**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCCTGCACTGT
P7_E06_7045	CAAGCAGAAGACGGCATACGAGAT**AGCCTA**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCCTGCACTGT
P7_F06_7046	CAAGCAGAAGACGGCATACGAGAT**ATGGAG**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCCTGCACTGT
P7_G06_7047	CAAGCAGAAGACGGCATACGAGAT**CCATGC**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCCTGCACTGT
P7_H06_7048	CAAGCAGAAGACGGCATACGAGAT**GCTATT**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCCTGCACTGT
P7_A07_7049	CAAGCAGAAGACGGCATACGAGAT**TGGCAC**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCCTGCACTGT
P7_B07_7050	CAAGCAGAAGACGGCATACGAGAT**CCTCGA**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCCTGCACTGT
P7_C07_7051	CAAGCAGAAGACGGCATACGAGAT**TTAGTG**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCCTGCACTGT
P7_D07_7052	CAAGCAGAAGACGGCATACGAGAT**TATACC**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCCTGCACTGT
P7_E07_7053	CAAGCAGAAGACGGCATACGAGAT**CAGGCG**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCCTGCACTGT
P7_F07_7054	CAAGCAGAAGACGGCATACGAGAT**TTCATT**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCCTGCACTGT
P7_G07_7055	CAAGCAGAAGACGGCATACGAGAT**CGTTGT**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCCTGCACTGT
P7_H07_7056	CAAGCAGAAGACGGCATACGAGAT**GGATAT**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCCTGCACTGT
P7_A08_7057	CAAGCAGAAGACGGCATACGAGAT**AGTACT**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCCTGCACTGT
P7_B08_7058	CAAGCAGAAGACGGCATACGAGAT**TCTTAT**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCCTGCACTGT
P7_C08_7059	CAAGCAGAAGACGGCATACGAGAT**CTCACC**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCCTGCACTGT
P7_D08_7060	CAAGCAGAAGACGGCATACGAGAT**GCGGAA**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCCTGCACTGT
P7_E08_7061	CAAGCAGAAGACGGCATACGAGAT**CACTTC**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCCTGCACTGT
P7_F08_7062	CAAGCAGAAGACGGCATACGAGAT**GCATTG**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCCTGCACTGT
P7_G08_7063	CAAGCAGAAGACGGCATACGAGAT**AGACGT**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCCTGCACTGT
P7_H08_7064	CAAGCAGAAGACGGCATACGAGAT**GTCCTG**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCCTGCACTGT
P7_A09_7065	CAAGCAGAAGACGGCATACGAGAT**GAGCTT**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCCTGCACTGT
P7_B09_7066	CAAGCAGAAGACGGCATACGAGAT**ATCGTC**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCCTGCACTGT
P7_C09_7067	CAAGCAGAAGACGGCATACGAGAT**GAACGA**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCCTGCACTGT
P7_D09_7068	CAAGCAGAAGACGGCATACGAGAT**AATTGG**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCCTGCACTGT
P7_E09_7069	CAAGCAGAAGACGGCATACGAGAT**TCAACT**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCCTGCACTGT
P7_F09_7070	CAAGCAGAAGACGGCATACGAGAT**TCGGTT**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCCTGCACTGT
P7_G09_7071	CAAGCAGAAGACGGCATACGAGAT**CATCTG**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCCTGCACTGT
P7_H09_7072	CAAGCAGAAGACGGCATACGAGAT**TTCTAC**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCCTGCACTGT
P7_A10_7073	CAAGCAGAAGACGGCATACGAGAT**GATGTC**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCCTGCACTGT
P7_B10_7074	CAAGCAGAAGACGGCATACGAGAT**AGATCG**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCCTGCACTGT
P7_C10_7075	CAAGCAGAAGACGGCATACGAGAT**GCGACG**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCCTGCACTGT
P7_D10_7076	CAAGCAGAAGACGGCATACGAGAT**CGCCAT**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCCTGCACTGT
P7_E10_7077	CAAGCAGAAGACGGCATACGAGAT**ACCAAT**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCCTGCACTGT
P7_F10_7078	CAAGCAGAAGACGGCATACGAGAT**TGTGGC**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCCTGCACTGT
P7_G10_7079	CAAGCAGAAGACGGCATACGAGAT**TAGATG**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCCTGCACTGT
P7_H10_7080	CAAGCAGAAGACGGCATACGAGAT**CTTGTT**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCCTGCACTGT
P7_A11_7081	CAAGCAGAAGACGGCATACGAGAT**GGCTGC**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCCTGCACTGT
P7_B11_7082	CAAGCAGAAGACGGCATACGAGAT**AAGAGT**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCCTGCACTGT
P7_C11_7083	CAAGCAGAAGACGGCATACGAGAT**ATAGCA**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCCTGCACTGT
P7_D11_7084	CAAGCAGAAGACGGCATACGAGAT**CCACTT**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCCTGCACTGT
P7_E11_7085	CAAGCAGAAGACGGCATACGAGAT**TATGAG**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCCTGCACTGT
P7_F11_7086	CAAGCAGAAGACGGCATACGAGAT**GGTCGG**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCCTGCACTGT
P7_G11_7087	CAAGCAGAAGACGGCATACGAGAT**TTGGCC**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCCTGCACTGT
P7_H11_7088	CAAGCAGAAGACGGCATACGAGAT**TGGTCT**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCCTGCACTGT
P7_A12_7089	CAAGCAGAAGACGGCATACGAGAT**TCCGCG**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCCTGCACTGT
P7_B12_7090	CAAGCAGAAGACGGCATACGAGAT**AGCAGG**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCCTGCACTGT
P7_C12_7091	CAAGCAGAAGACGGCATACGAGAT**CAGCGC**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCCTGCACTGT
P7_D12_7092	CAAGCAGAAGACGGCATACGAGAT**AGGTTC**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCCTGCACTGT
P7_E12_7093	CAAGCAGAAGACGGCATACGAGAT**CTCGAA**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCCTGCACTGT
P7_F12_7094	CAAGCAGAAGACGGCATACGAGAT**GATTCT**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCCTGCACTGT
P7_G12_7095	CAAGCAGAAGACGGCATACGAGAT**ATGCCT**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCCTGCACTGT

P7_H12_7096	CAAGCAGAAGACGGCATACGAGAT**GCGTGT**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCCTGCACTGT

To simplify the processing of large numbers of samples, we built a multistep QMAP-Seq bioinformatic analysis pipeline **(Figure 5)**. First, the pipeline demultiplexed the 768 individual samples according to i5 and i7 index sequences. Second, it extracted the cell line barcode **(Table 4)** and sgRNA barcode **(Table 5)** from each read and counted the number of reads for each cell line-sgRNA pair. Third, the pipeline used the cell spike-in standards to generate a sample-specific standard curve and used the standard curve to interpolate cell number from sequencing reads. Finally, it calculated the number of cells for each cell line-sgRNA pair in the presence of compound relative to DMSO control.

Table 4. Cell line barcode read sequences.

Cell Line	Barcode Number	Read Sequence
ZR-75-1	B11	ATCACTGG
SKBR3	A09	CGGTGACC
HCC-38	A10	TACAGAGG
MDA-MB-231	A04	TAACAGTT
BT-20	A11	ATTGTCAA
293T	A02	GCTGGATT

Table 5. sgRNA barcode read sequences.

sgRNA	Read Sequence
sgNT_0	ATCGTTTCCGCTTAACGGCG
sgNT_1	GTTCGACTCGCGTGACCGTA
sgNT_2	GTCGTCATACAACGGCAACG
sgNT_3	GCCGACCAACGTCAGCGGTA
sgNT_4	GCGTTAGTCACTTAGCTCGA
sgNT_5	GCGCACGTATAATCCTGCGT
sgNT_6	GAGGTGATCGATTAAGTACA
sgNT_7	GGTCCTGGCGAGGAGCCTTG
sgNT_8	GCGGAGAGCATTGACCTCAT
sgNT_9	GAATCGGAATAGTTGATTCG
sgHSF1_3	AGCTTCCACGTGTTCGACCA
sgHSF2_3	ACACTAACGAGTTCATCACC
sgERN1_2	TCTCCCTCAATGGTACACGA
sgXBP1_2	CCTCCCAGGGGAATGAAGTG
sgATF3_4	CCAGCGCAGAGGACATCCGG
sgATF4_2	AGATGACCTTCTGACCACGT
sgATF6_4	TGAGCCCTGCAAATCAAAGG
sgNFE2L2_2	CACATCCAGTCAGAAACCAG
sgKEAP1_1	CAGCACCGTTCATGACGTGG
sgATG7_3	CTTGAAAGACTCGAGTGTGT
sgSLC35F2_3	CATGACAGAGCCATCAACAC

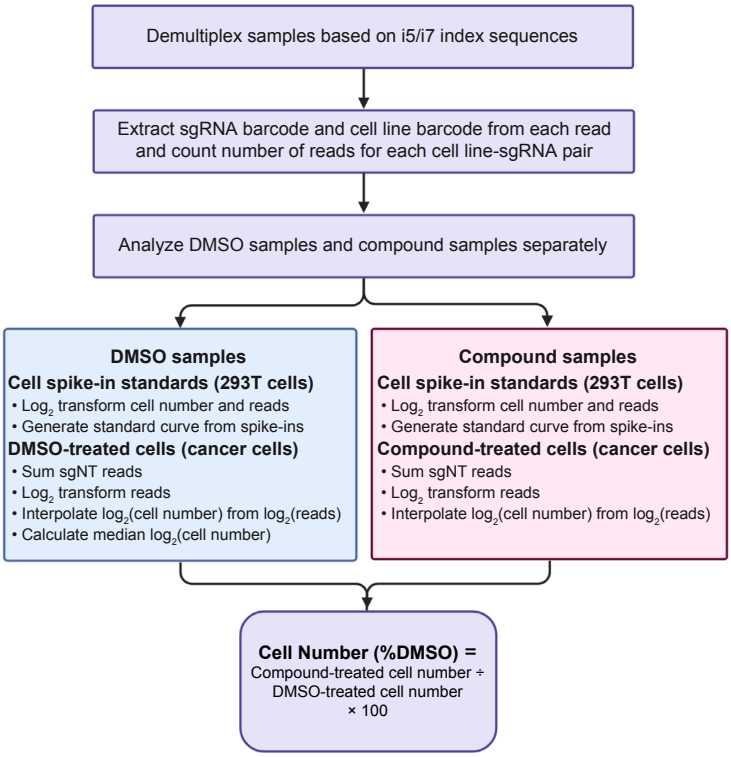


Figure 5. QMAP-Seq bioinformatic analysis pipeline. Pipeline for demultiplexing samples, counting sequencing reads for cell line-sgRNA pairs, interpolating cell number from sequencing reads, and calculating cell number in the presence of compound relative to DMSO.

We next performed a series of analyses to assess the quality of the data generated from QMAP-Seq. We first asked whether we could resolve differences in cell number by next-generation sequencing. Indeed, analysis of the cell spike-in standards revealed the expected increase in sequencing reads with increased input cell number **(Figure 6a)**. To assess the precision of QMAP-Seq, we compared the interpolated cell number between the two replicates for every compound-dose pair. Importantly, QMAP-Seq replicates were highly correlated (r = 0.9714) **(Figure 6b)**. To compare QMAP-Seq with an established method of measuring cellular response to compounds, we assessed the percent confluence of the population of cells in each compound-treated well using live cell imaging immediately prior to cell lysis. As expected, live cell imaging revealed a reduction in cell confluence with increasing concentrations of YM155 **(Figure 6c, Figure 6d, Top Left)**. The area under the dose-response curve (AUC) for the population of cells as determined using QMAP-Seq was similar to live cell imaging analysis (live cell imaging AUC = 74.12, QMAP-Seq AUC = 85.95). Importantly, QMAP-Seq was further able to resolve differences in drug response between the knockouts, such as knockout of SLC35F2 conferring resistance to YM155 **(Figure 6d, Bottom Left)**. As another example, live cell imaging revealed a dose-dependent reduction in cell confluence with the HSP90 inhibitor, Ganetespib **(Figure 6c, Figure 6d, Top Right)**. Once again, the AUC for the population of cells calculated using QMAP-Seq was similar to live cell imaging analysis (live cell imaging AUC = 120.90, QMAP-Seq AUC = 104.80). Furthermore, QMAP-Seq detected that knockout of HSF1 sensitized cells to Ganetespib, confirming another previously established chemical-genetic interaction (86) **(Figure 6d, Bottom Right)**. Notably, the nearly identical results obtained from the two assays was not limited to these two compounds. The AUCs for the population of cells calculated using QMAP-Seq were remarkably concordant with those calculated using live cell imaging across all 89 compounds (r = 0.9228), demonstrating the high degree of accuracy of QMAP-Seq **(Figure 6e)**. Taken together, these results indicate that QMAP-Seq generates precise, accurate, and sensitive quantitative measures of pharmacologic response in pooled format.

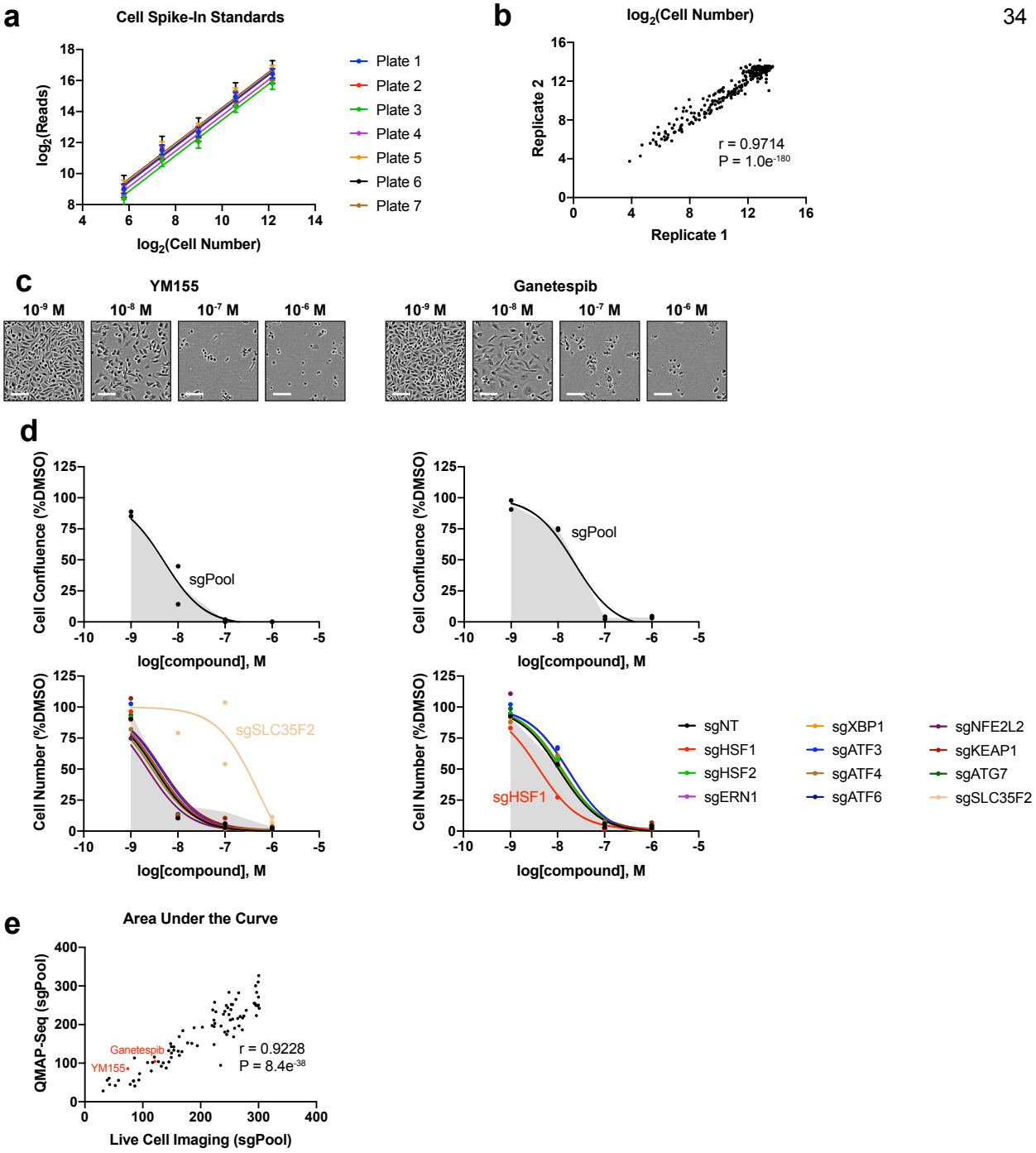


Figure 6. QMAP-Seq generates precise and accurate quantitative measures of drug response. a Standard curves generated from five uniquely barcoded 293T cell spike-in standards introduced at known cell numbers. Data are represented as mean number of sequencing reads across the six or eight DMSO samples on a plate ± standard deviation. **b** Scatterplot of interpolated cell number for two biologically independent replicates. Statistical significance of Pearson correlation was determined using a two-tailed test (n = 288 compound-dose combinations). **c** Live cell imaging of MDA-MB-231 sgPool cells 72 hours after treatment with YM155 or Ganetespib. Images are representative of two biologically independent replicates. Scale bar = 100 µm. **d** Top: Dose-response curves for MDA-MB-231 sgPool cells as measured using live cell imaging 72 hours after treatment with YM155 or Ganetespib. Bottom: Dose-response curves for 12 genetic perturbations of MDA-MB-231 cells as measured using QMAP-Seq 72 hours after treatment with YM155 or Ganetespib. Each data point represents one of two biologically independent replicates. The shaded region indicates the area under the curve (AUC) for sgPool. **e** Scatterplot of the dose-response curve AUC for sgPool as determined using live cell imaging versus QMAP-Seq. Statistical significance of Pearson correlation was determined using a two-tailed test (n = 89 compounds).

2.3 Expanding QMAP-Seq to multiple cell lines.

Because the genetic and epigenetic background of a cell line can impact therapeutic response, we next expanded QMAP-Seq to multiple cell lines **(Figure 7)**. We selected a panel of breast cancer cell lines comprising three major subtypes of breast cancer: ER+ (ZR-75-1), HER2+ (SKBR3), and triple-negative (HCC-38, MDA-MB-231, BT-20). We assembled a 180 compound collection for QMAP-Seq **(Table 6)** in a manner that allowed us to achieve two primary objectives for this assay. First, we selected FDA-approved drugs or compounds in clinical trials to facilitate the discovery of clinically-relevant chemical-genetic interactions and to enable drug repurposing **(Figure 8a)**. This included chemotherapeutics and targeted therapies that are either standard-of-care or currently being investigated in the context of breast cancer. Second, we selected compounds that target biological processes from 19 diverse pathways because proteostasis factors broadly impact cell biology **(Figure 8b)**.

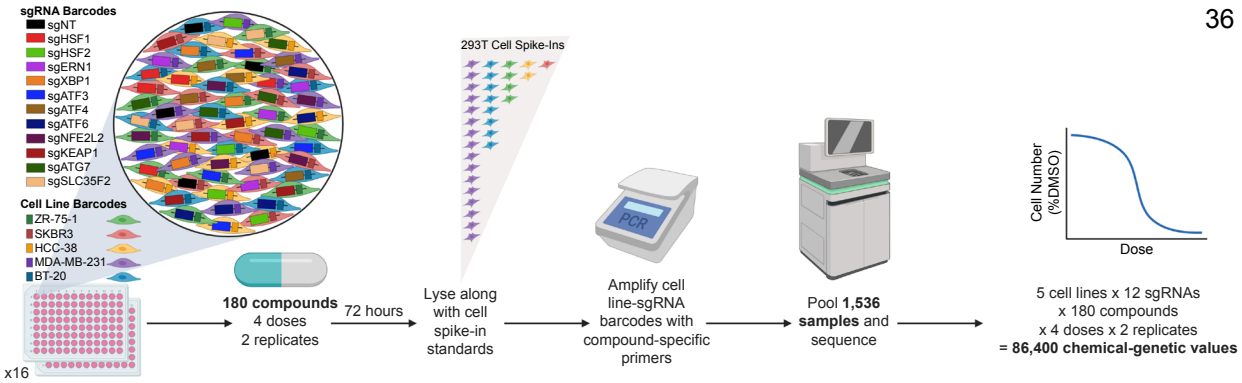


Figure 7. Experimental workflow for QMAP-Seq with five cell lines.

Table 6. Compounds, compound targets, and compound pathways for QMAP-Seq.

Compound	Compound Target	Compound Pathway
Abiraterone	Cytochrome P450	Proteasome/Metabolic Enzyme
ABT199	Bcl-2 Family; Autophagy	Apoptosis; Autophagy
ABT737	Bcl-2 Family; Autophagy; Mitophagy	Apoptosis; Autophagy
Afatinib	EGFR; Autophagy	JAK/STAT Signaling; Protein Tyrosine Kinase/RTK; Autophagy
Alisertib	Aurora Kinase; Autophagy; Apoptosis	Cell Cycle/DNA Damage; Epigenetics; Autophagy; Apoptosis
Anastrozole	Aromatase	Others
APY29	IRE1	Cell Cycle/DNA Damage
AT7867	Akt; PKA; Ribosomal S6 Kinase (RSK)	PI3K/Akt/mTOR; Protein Tyrosine Kinase/RTK; Stem Cell/Wnt; MAPK/ERK Pathway
AZD4547	FGFR	Protein Tyrosine Kinase/RTK
AZD5363	Akt; Autophagy	PI3K/Akt/mTOR; Autophagy
AZD7545	PDHK	Proteasome/Metabolic Enzyme
AZD8186	PI3K	PI3K/Akt/mTOR
BardoxoloneMethyl	Keap1-Nrf2; Autophagy; Apoptosis	NF-κB; Autophagy; Apoptosis
Belinostat	HDAC; Autophagy	Cell Cycle/DNA Damage; Epigenetics; Autophagy
Bendamustine	DNA Alkylator/Crosslinker; Apoptosis	Cell Cycle/DNA Damage; Apoptosis
BetulinicAcid	Apoptosis; Topoisomerase; HIV; Autophagy; Mitophagy	Apoptosis; Cell Cycle/DNA Damage; Anti-infection; Autophagy
Bexarotene	RAR/RXR; Autophagy	Proteasome/Metabolic Enzyme; Autophagy
BGJ398	FGFR	Protein Tyrosine Kinase/RTK
BIX01294	Histone Methyltransferase; Autophagy	Epigenetics; Autophagy
Blebbistatin	Myosin	Cytoskeleton
BMS345541	IKK	NF-κB
BMS536924	IGF-1R; Insulin Receptor	Protein Tyrosine Kinase/RTK
BMS754807	IGF-1R; Insulin Receptor	Protein Tyrosine Kinase/RTK
Bortezomib	Proteasome; Apoptosis; Autophagy	Proteasome/Metabolic Enzyme; Apoptosis; Autophagy
BrefeldinA	Autophagy; CRISPR/Cas9; Mitophagy; HSV	Autophagy; Cell Cycle/DNA Damage; Anti-infection
Brivanib	VEGFR; Autophagy	Protein Tyrosine Kinase/RTK; Autophagy
Cabozantinib	VEGFR; c-Met/HGFR; c-Kit; TAM Receptor; FLT3; Apoptosis	Protein Tyrosine Kinase/RTK; Apoptosis
CAL101	PI3K; Autophagy	PI3K/Akt/mTOR; Autophagy
Canertinib	EGFR	JAK/STAT Signaling; Protein Tyrosine Kinase/RTK
Carfilzomib	Proteasome; Autophagy; Apoptosis	Proteasome/Metabolic Enzyme; Autophagy; Apoptosis
Cediranib	VEGFR; Autophagy; PDGFR	Protein Tyrosine Kinase/RTK; Autophagy
CHIR99021	GSK-3; Autophagy	PI3K/Akt/mTOR; Stem Cell/Wnt; Autophagy
Chlorambucil	DNA Alkylator/Crosslinker	Cell Cycle/DNA Damage
Ciclopirox	Fungal; Autophagy	Anti-infection; Autophagy
Cimetidine	Histamine Receptor	GPCR/G Protein; Immunology/Inflammation; Neuronal Signaling
Cisplatin	DNA Alkylator/Crosslinker	Cell Cycle/DNA Damage
Clofarabine	Nucleoside Antimetabolite/Analog; Autophagy; Apoptosis	Cell Cycle/DNA Damage; Autophagy; Apoptosis
Crizotinib	ALK; c-Met/HGFR; ROS; Autophagy	Protein Tyrosine Kinase/RTK; Autophagy
Curcumin	Keap1-Nrf2; Autophagy; Histone Acetyltransferase; Epigenetic Reader Domain; Mitophagy; Influenza Virus	NF-κB; Autophagy; Epigenetics; Anti-infection
Cyclophosphamide	DNA Alkylator/Crosslinker	Cell Cycle/DNA Damage
Dabrafenib	Raf	MAPK/ERK Pathway
Dacarbazine	Nucleoside Antimetabolite/Analog; Apoptosis	Cell Cycle/DNA Damage; Apoptosis
Dasatinib	Bcr-Abl; Src; Autophagy; Apoptosis	Protein Tyrosine Kinase/RTK; Autophagy; Apoptosis
DBeQ	p97; Autophagy; Apoptosis	Cell Cycle/DNA Damage; Autophagy; Apoptosis
Decitabine	DNA Methyltransferase; Apoptosis	Epigenetics; Apoptosis
Dexamethasone	Glucocorticoid Receptor; Autophagy; Mitophagy	GPCR/G Protein; Autophagy
DimethylFumarate	Keap1-Nrf2	NF-κB
Doxorubicin	Topoisomerase; ADC Cytotoxin; Autophagy; Mitophagy; AMPK	Cell Cycle/DNA Damage; Antibody-drug Conjugate/ADC Related; Autophagy; Epigenetics; PI3K/Akt/mTOR

Entinostat	HDAC; Autophagy; Apoptosis	Cell Cycle/DNA Damage; Epigenetics; Autophagy; Apoptosis
Entrectinib	ROS; Trk Receptor; ALK; Autophagy	Protein Tyrosine Kinase/RTK; Neuronal Signaling; Autophagy
Erastin	Ferroptosis	Apoptosis
Erlotinib	EGFR; Autophagy	JAK/STAT Signaling; Protein Tyrosine Kinase/RTK; Autophagy
Etomoxir	Others	Others
Etoposide	Topoisomerase; Autophagy; Mitophagy; Apoptosis	Cell Cycle/DNA Damage; Autophagy; Apoptosis
ETP46464	mTOR; ATM/ATR	PI3K/Akt/mTOR; Cell Cycle/DNA Damage
Everolimus	mTOR; FKBP; Autophagy	PI3K/Akt/mTOR; Apoptosis; Autophagy; Immunology/Inflammation
EX527	Sirtuin	Cell Cycle/DNA Damage; Epigenetics
Exemestane	Aromatase	Others
Fingolimod	LPL Receptor; PAK	GPCR/G Protein; Cell Cycle/DNA Damage; Cytoskeleton
FiveFluorouracil	Nucleoside Antimetabolite/Analog; HIV	Cell Cycle/DNA Damage; Anti-infection
Fluvastatin	HMG-CoA Reductase (HMGCR); Autophagy	Proteasome/Metabolic Enzyme; Autophagy
FourHydroxytamoxifen	Estrogen Receptor/ERR	Others
Fulvestrant	Estrogen Receptor/ERR; Autophagy	Others; Autophagy
Ganetespib	HSP; Apoptosis	Cell Cycle/DNA Damage; Proteasome/Metabolic Enzyme; Apoptosis
GDC0879	Raf	MAPK/ERK Pathway
GDC0941	PI3K; Autophagy; Apoptosis	PI3K/Akt/mTOR; Autophagy; Apoptosis
Gefitinib	EGFR; Autophagy	JAK/STAT Signaling; Protein Tyrosine Kinase/RTK; Autophagy
Gemcitabine	DNA/RNA Synthesis; Nucleoside Antimetabolite/Analog; Autophagy	Cell Cycle/DNA Damage; Autophagy
Gossypol	Bcl-2 Family	Apoptosis
GSK1059615	PI3K; mTOR; Apoptosis	PI3K/Akt/mTOR; Apoptosis
GSK2636771	PI3K	PI3K/Akt/mTOR
GSK4112	Autophagy	Autophagy
GSK461364	Polo-like Kinase (PLK)	Cell Cycle/DNA Damage
GSKJ4	Histone Demethylase	Epigenetics
IBET151	Epigenetic Reader Domain	Epigenetics
IC87114	PI3K	PI3K/Akt/mTOR
Ifosfamide	DNA Alkylator/Crosslinker	Cell Cycle/DNA Damage
Isoliquiritigenin	Aldose Reductase; Autophagy; Apoptosis	Proteasome/Metabolic Enzyme; Autophagy; Apoptosis
Istradefylline	Adenosine Receptor	GPCR/G Protein
JQ1	Epigenetic Reader Domain; Autophagy; Ligand for Target Protein for PROTAC	Epigenetics; Autophagy; PROTAC
Ki8751	VEGFR	Protein Tyrosine Kinase/RTK
KU0063794	mTOR	PI3K/Akt/mTOR
KU55933	ATM/ATR; Autophagy	Cell Cycle/DNA Damage; PI3K/Akt/mTOR; Autophagy
KU60019	ATM/ATR	Cell Cycle/DNA Damage; PI3K/Akt/mTOR
KW2449	FLT3; FGFR; Bcr-Abl; Aurora Kinase; Apoptosis	Protein Tyrosine Kinase/RTK; Cell Cycle/DNA Damage; Epigenetics; Apoptosis
KX2391	Src; Microtubule/Tubulin	Protein Tyrosine Kinase/RTK; Cell Cycle/DNA Damage; Cytoskeleton
Lapatinib	EGFR; Autophagy	JAK/STAT Signaling; Protein Tyrosine Kinase/RTK; Autophagy
Lenvatinib	VEGFR; FGFR; PDGFR; c-Kit; RET	Protein Tyrosine Kinase/RTK
Letrozole	Aromatase; Autophagy	Others; Autophagy
Linifanib	PDGFR; VEGFR; FLT3; c-Fms; c-Kit; Apoptosis; Autophagy	Protein Tyrosine Kinase/RTK; Apoptosis; Autophagy
Linsitinib	Insulin Receptor; IGF-1R	Protein Tyrosine Kinase/RTK
Lomeguatrib	DNA Methyltransferase	Epigenetics
Lovastatin	HMG-CoA Reductase (HMGCR); Autophagy	Proteasome/Metabolic Enzyme; Autophagy
LRRK2IN1	LRRK2; Apoptosis	Autophagy; Apoptosis
Masitinib	c-Kit; PDGFR; Src; FGFR; FAK; Apoptosis	Protein Tyrosine Kinase/RTK; Apoptosis
Mdivi1	Dynamin; Mitophagy; Autophagy; Apoptosis	Cytoskeleton; Autophagy; Apoptosis
Mebendazole	Parasite; Apoptosis	Anti-infection; Apoptosis

Methotrexate	Antifolate; ADC Cytotoxin; Apoptosis	Cell Cycle/DNA Damage; Antibody-drug Conjugate/ADC Related; Apoptosis
MG132	Proteasome; Autophagy	Proteasome/Metabolic Enzyme; Autophagy
Mitoxantrone	Topoisomerase; PKC	Cell Cycle/DNA Damage; Epigenetics; TGF-beta/Smad
MK0752	γ-secretase	Neuronal Signaling; Stem Cell/Wnt
MK1775	Wee1	Cell Cycle/DNA Damage
MLN2238	Proteasome; Autophagy	Proteasome/Metabolic Enzyme; Autophagy
Navitoclax	Bcl-2 Family	Apoptosis
Necrostatin1	RIP kinase; Autophagy; Indoleamine 2,3-Dioxygenase (IDO	Apoptosis; Autophagy; Proteasome/Metabolic Enzyme
Nelarabine	Nucleoside Antimetabolite/Analog; Apoptosis	Cell Cycle/DNA Damage; Apoptosis
Neratinib	EGFR	JAK/STAT Signaling; Protein Tyrosine Kinase/RTK
Niclosamide	STAT; Parasite	JAK/STAT Signaling; Stem Cell/Wnt; Anti-infection
Nilotinib	Bcr-Abl; Autophagy	Protein Tyrosine Kinase/RTK; Autophagy
Nitroxoline	Bacterial; Autophagy	Anti-infection; Autophagy
NSC74859	STAT	JAK/STAT Signaling; Stem Cell/Wnt
NVPAUY922	HSP; Autophagy; Apoptosis	Cell Cycle/DNA Damage; Proteasome/Metabolic Enzyme; Autophagy; Apoptosis
Olaparib	PARP; Autophagy; Mitophagy	Cell Cycle/DNA Damage; Epigenetics; Autophagy
Ornidazole	Bacterial; Parasite	Anti-infection
OSI027	mTOR; Autophagy	PI3K/Akt/mTOR; Autophagy
PAC1	Caspase; Autophagy; Apoptosis	Apoptosis; Autophagy
Paclitaxel	Microtubule/Tubulin; ADC Cytotoxin; Autophagy	Cell Cycle/DNA Damage; Cytoskeleton; Antibody-drug Conjugate/ADC Related; Autophagy
Palbociclib	CDK	Cell Cycle/DNA Damage
Parthenolide	NF-κB; Autophagy; Mitophagy; Apoptosis	NF-κB; Autophagy; Apoptosis
Pazopanib	VEGFR; c-Kit; PDGFR; Autophagy; FGFR	Protein Tyrosine Kinase/RTK; Autophagy
PD318088	MEK	MAPK/ERK Pathway
PF3758309	PAK; Apoptosis	Cell Cycle/DNA Damage; Cytoskeleton; Apoptosis
PF573228	FAK; Apoptosis	Protein Tyrosine Kinase/RTK; Apoptosis
PHA793887	CDK; Apoptosis	Cell Cycle/DNA Damage; Apoptosis
Phloretin	SGLT	Membrane Transporter/Ion Channel
PI103	PI3K; mTOR; DNA-PK; Autophagy; Apoptosis	PI3K/Akt/mTOR; Cell Cycle/DNA Damage; Autophagy; Apoptosis
PifithrinAlpha	MDM-2/p53; Aryl Hydrocarbon Receptor	Apoptosis; Immunology/Inflammation
PIK93	PI4K; PI3K; Virus Protease	PI3K/Akt/mTOR
PLX4720	Raf	MAPK/ERK Pathway
Purmorphamine	Smo; Autophagy	Stem Cell/Wnt; Autophagy
PX12	Others	Others
PYR41	E1/E2/E3 Enzyme; Apoptosis	Proteasome/Metabolic Enzyme; Apoptosis
RacRotigotine	Dopamine Receptor	GPCR/G Protein; Neuronal Signaling
Regorafenib	VEGFR; Autophagy; PDGFR; Raf; RET	Protein Tyrosine Kinase/RTK; Autophagy; MAPK/ERK Pathway
RG108	DNA Methyltransferase	Epigenetics
Rigosertib	Polo-like Kinase (PLK); PI3K; Apoptosis	Cell Cycle/DNA Damage; PI3K/Akt/mTOR; Apoptosis
RITA	MDM-2/p53; Autophagy; DNA Alkylator/Crosslinker	Apoptosis; Autophagy; Cell Cycle/DNA Damage
RO4929097	γ-secretase; Notch	Neuronal Signaling; Stem Cell/Wnt
Ruxolitinib	JAK; Autophagy; Mitophagy	Epigenetics; JAK/STAT Signaling; Stem Cell/Wnt; Autophagy
Saracatinib	Src; Autophagy	Protein Tyrosine Kinase/RTK; Autophagy
SB743921	Kinesin	Cell Cycle/DNA Damage; Cytoskeleton
Selumetinib	MEK; Apoptosis	MAPK/ERK Pathway; Apoptosis
Semagacestat	γ-secretase; Amyloid-β; Notch	Neuronal Signaling; Stem Cell/Wnt
Sildenafil	Phosphodiesterase (PDE); Autophagy	Proteasome/Metabolic Enzyme; Autophagy
Simvastatin	HMG-CoA Reductase (HMGCR); Autophagy; Apoptosis; Mitophagy	Proteasome/Metabolic Enzyme; Autophagy; Apoptosis
Sitagliptin	Dipeptidyl Peptidase; Autophagy	Proteasome/Metabolic Enzyme; Autophagy
SixMercaptopurine	Nucleoside Antimetabolite/Analog; Autophagy	Cell Cycle/DNA Damage; Autophagy
SJ172550	MDM-2/p53	Apoptosis
SN38	Topoisomerase; ADC Cytotoxin; Autophagy	Cell Cycle/DNA Damage; Antibody-drug Conjugate/ADC Related; Autophagy
SNS032	CDK; Apoptosis	Cell Cycle/DNA Damage; Apoptosis

Sorafenib	Raf; VEGFR; FLT3; Autophagy; Apoptosis	MAPK/ERK Pathway; Protein Tyrosine Kinase/RTK; Autophagy; Apoptosis
Spautin1	Autophagy; Apoptosis	Autophagy; Apoptosis
SR1001	ROR	Proteasome/Metabolic Enzyme
StemRegenin1	Aryl Hydrocarbon Receptor	Immunology/Inflammation
STF083010	IRE1	Cell Cycle/DNA Damage
STF31	Autophagy	Autophagy
SU11274	c-Met/HGFR; Autophagy; Apoptosis	Protein Tyrosine Kinase/RTK; Autophagy; Apoptosis
Tacrolimus	Phosphatase; FKBP; Autophagy	Proteasome/Metabolic Enzyme; Apoptosis; Autophagy; Immunology/Inflammation
Temozolomide	Autophagy; DNA Alkylator/Crosslinker; Apoptosis	Autophagy; Cell Cycle/DNA Damage; Apoptosis
Teniposide	Topoisomerase	Cell Cycle/DNA Damage
TG101348	JAK; Apoptosis	Epigenetics; JAK/STAT Signaling; Stem Cell/Wnt; Apoptosis
Thalidomide	Ligand for E3 Ligase; Autophagy; Apoptosis	PROTAC; Autophagy; Apoptosis
Tigecycline	Bacterial; Autophagy	Anti-infection; Autophagy
Tivantinib	c-Met/HGFR; Apoptosis	Protein Tyrosine Kinase/RTK; Apoptosis
Tivozanib	VEGFR	Protein Tyrosine Kinase/RTK
Topotecan	Topoisomerase; Autophagy	Cell Cycle/DNA Damage; Autophagy
TPCA1	IKK; STAT; Apoptosis	NF-κB; JAK/STAT Signaling; Stem Cell/Wnt; Apoptosis
Trametinib	MEK; Autophagy; Apoptosis	MAPK/ERK Pathway; Autophagy; Apoptosis
Trifluoperazine	Dopamine Receptor; Autophagy; Influenza Virus	GPCR/G Protein; Neuronal Signaling; Autophagy; Anti-infection
Triptolide	NF-κB; Apoptosis	NF-κB; Apoptosis
TubastatinA	HDAC; Autophagy; Apoptosis	Cell Cycle/DNA Damage; Epigenetics; Autophagy; Apoptosis
Valdecoxib	COX	Immunology/Inflammation
Veliparib	PARP; Autophagy	Cell Cycle/DNA Damage; Epigenetics; Autophagy
VER155008	HSP; Autophagy	Cell Cycle/DNA Damage; Proteasome/Metabolic Enzyme; Autophagy
Vincristine	Microtubule/Tubulin	Cell Cycle/DNA Damage; Cytoskeleton
WAY362450	FXR; Autophagy	Proteasome/Metabolic Enzyme; Autophagy
WP1130	Deubiquitinase; Bcr-Abl; Autophagy; Apoptosis	Cell Cycle/DNA Damage; Protein Tyrosine Kinase/RTK; Autophagy; Apoptosis
YM155	Survivin; Autophagy	Apoptosis; Autophagy
Zebularine	DNA Methyltransferase; Autophagy	Epigenetics; Autophagy
ZSTK474	PI3K; Autophagy; Autophagy	PI3K/Akt/mTOR; Autophagy

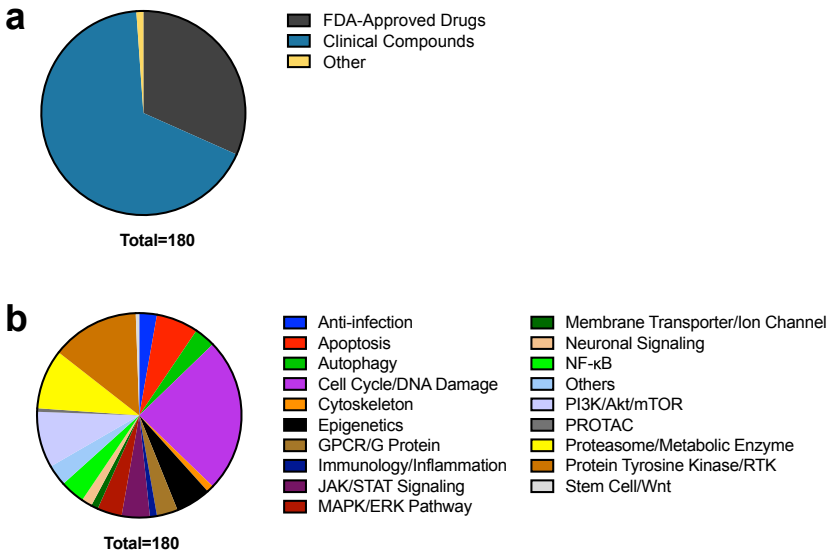


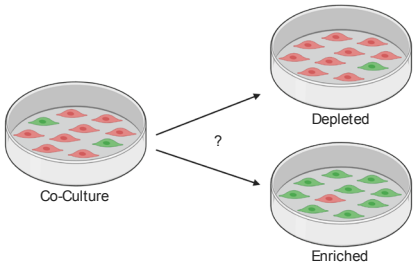
Figure 8. Selecting compounds for QMAP-Seq. a Pie chart showing breakdown of 180 compounds by stage of development. **b** Pie chart showing breakdown of 180 compounds by pathway.

To ensure that each cell line would be similarly represented in our assay despite different doubling times, we measured the relative abundance of each of the five cell lines grown in heterogeneous pools. We prepared five pools, each containing 20% of one cell line expressing ZsGreen and 20% of each of the other four cell lines expressing dTomato **(Figure 9a)**. We co-cultured and analyzed the pools by flow cytometry to quantify the percentage of GFP positive cells over time. We found that SKBR3 cells were the most depleted, whereas MDA-MB-231 cells were the most enriched **(Figure 9b, Left)**. We used the relative cell abundances from this competition experiment to mathematically model an optimized pooling ratio to ensure adequate representation of all cell lines at seven days post-pooling, the timepoint when the cell lines are exposed to compounds during QMAP-Seq **(Figure 9b, Right, 0 Days)**. Optimized pools predicted to contain 20% of each cell line after seven days exhibited similar representation of the cell lines at this timepoint **(Figure 9b, Right, 7 Days)**.

After pooling the five cell lines each possessing our panel of proteostasis factor knockouts at the optimized pooling ratio, we performed the QMAP-Seq workflow **(Figure 7)**. Seven days after thawing, we treated the pooled cells with compounds or DMSO. After three days of treatment, we measured the representation of the cell line-sgRNA pairs in the DMSO samples. Each of the cell lines was covered by at least 10% of the total sequencing reads, indicating adequate representation of the five cell lines **(Figure 10a)**. We also observed sufficient representation of most cell line-sgRNA pairs **(Figure 10b)**. Unsurprisingly, the two cell line-sgRNA pairs with the lowest representation in the pool (ZR-75-1 sgHSF1 and ZR-75-1 sgKEAP1) exhibited the greatest variation between DMSO samples (standard deviation > 0.7) and were thus excluded from downstream analysis **(Figure 10c)**.

a

	ZR-75-1 Pool	SKBR3 Pool	HCC-38 Pool	MDA-MB-231 Pool	BT-20 Pool
20%	ZR-75-1 ZsGreen	ZR-75-1 dTomato	ZR-75-1 dTomato	ZR-75-1 dTomato	ZR-75-1 dTomato
20%	SKBR3 dTomato	SKBR3 ZsGreen	SKBR3 dTomato	SKBR3 dTomato	SKBR3 dTomato
20%	HCC-38 dTomato	HCC-38 dTomato	HCC-38 ZsGreen	HCC-38 dTomato	HCC-38 dTomato
20%	MDA-MB-231 dTomato	MDA-MB-231 dTomato	MDA-MB-231 dTomato	MDA-MB-231 ZsGreen	MDA-MB-231 dTomato
20%	BT-20 dTomato	BT-20 dTomato	BT-20 dTomato	BT-20 dTomato	BT-20 ZsGreen



b

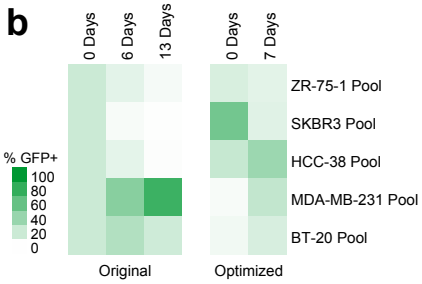


Figure 9. Optimizing conditions for pooling cell lines with different doubling times. a Schematic of competition experiment to optimize the starting representation of the five co-cultured breast cancer cell lines. Original cell line pools were prepared by mixing equal numbers of one cell line expressing ZsGreen with each of the other four cell lines expressing dTomato. Flow cytometry analysis measured the percentage of GFP positive cells in each pool over time. **b** Heat maps displaying the percentage of GFP positive cells at various time points as measured using flow cytometry. Left: Original cell line pools that started with 20% of each cell line on Day 0. Right: Optimized cell line pools predicted to contain 20% of each cell line on Day 7.

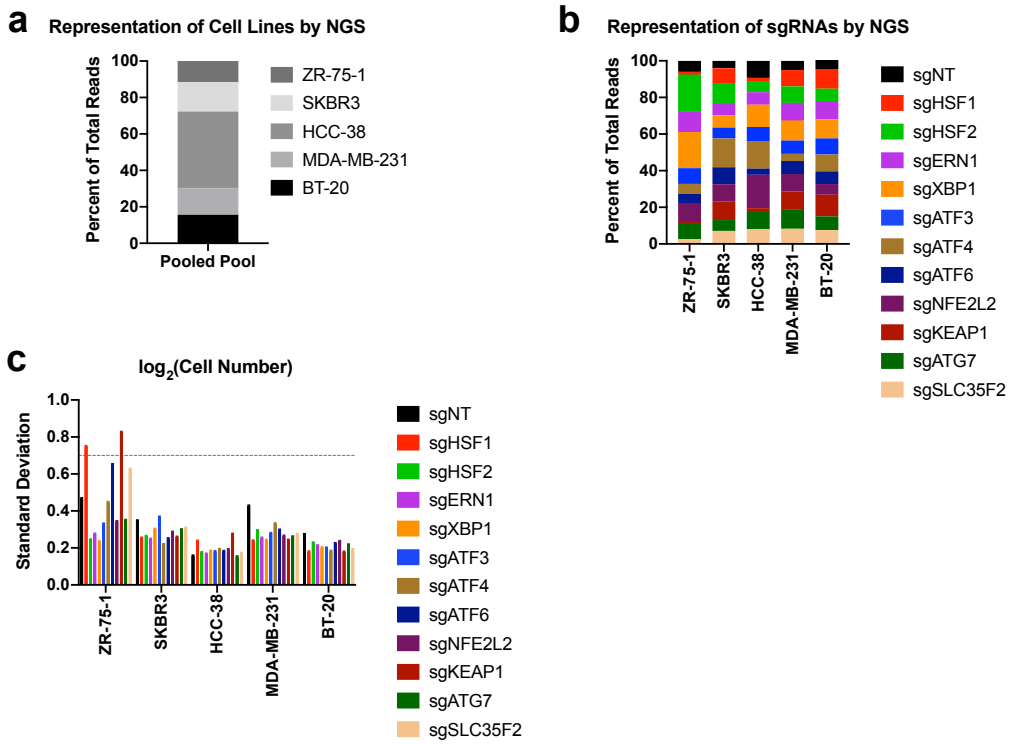


Figure 10. Representation of cell lines and sgRNAs. a Representation of five breast cancer cell lines as measured by counting the number of sequencing reads for each cell line barcode across the 96 DMSO samples. **b** Representation of sgRNAs as measured by counting the number of sequencing reads for each sgRNA barcode relative to the total number of sequencing reads for that cell line across the 96 DMSO samples. **c** Standard deviation of the interpolated cell number for each cell line-sgRNA pair across the 96 DMSO samples. Dotted line indicates threshold for excluding cell line-sgRNA pairs with high variability.

Next, we asked whether cell spike-in standards could reduce the technical variation between plates that we observed in the raw sequencing reads across the DMSO control samples **(Figure 11b, Top)**. For example, Plate 6 had markedly fewer sequencing reads compared to the other plates, but this plate also had a standard curve with a lower y-intercept, reflecting proportionally lower total reads compared to the other plates **(Figure 11a)**. Thus, utilizing sample-specific standard curves to interpolate cell number from sequencing reads improved this and other variation between plates **(Figure 11b, Bottom)**. For each of the five cell lines, we observed a statistically significant decrease in variation between plates upon interpolation (ZR-75-1: SD, -53% compared with raw reads, F test: $P = 6.3e^{-3}$) (SKBR3: SD, -66% compared with raw reads, F test: $P = 1.4e^{-4}$) (HCC-38: SD, -88% compared with raw reads, F test: $P = 5.9e^{-11}$) (MDA-MB-231: SD, -71% compared with raw reads, F test: $P = 2.4e^{-5}$) (BT-20: SD, -67% compared with raw reads, F test: $P = 1.1e^{-4}$) **(Figure 11c)**. Thus, cell spike-in standards serve as an internal control for normalizing data for each sample, thereby reducing technical variation and improving overall data quality.

2.4 Identification and validation of cell line-gene-drug interactions.

We next examined whether we could detect known cell line-specific and gene-specific drug vulnerabilities within complex mixtures of cells (five cell lines x 12 genetic perturbations) using QMAP-Seq. As expected, the ER+ cell line ZR-75-1 was preferentially sensitive to 4-Hydroxytamoxifen, and the HER2+ cell line SKBR3 was preferentially sensitive to Lapatinib **(Figure 12)**. In addition, knockout of SLC35F2 conferred resistance to YM155, most prominently in the triple-negative breast cancer cell lines that were most sensitive to this compound **(Figure 12)**. It is worth mentioning that the ZR-75-1 cells, which exhibited the greatest sample-to-sample variability of the five cell lines across the DMSO samples **(Figure 10c)**, also displayed the noisiest relative cell number data for the compound-treated cells **(Figure 12)**, suggesting that validation is even more crucial for this cell line.

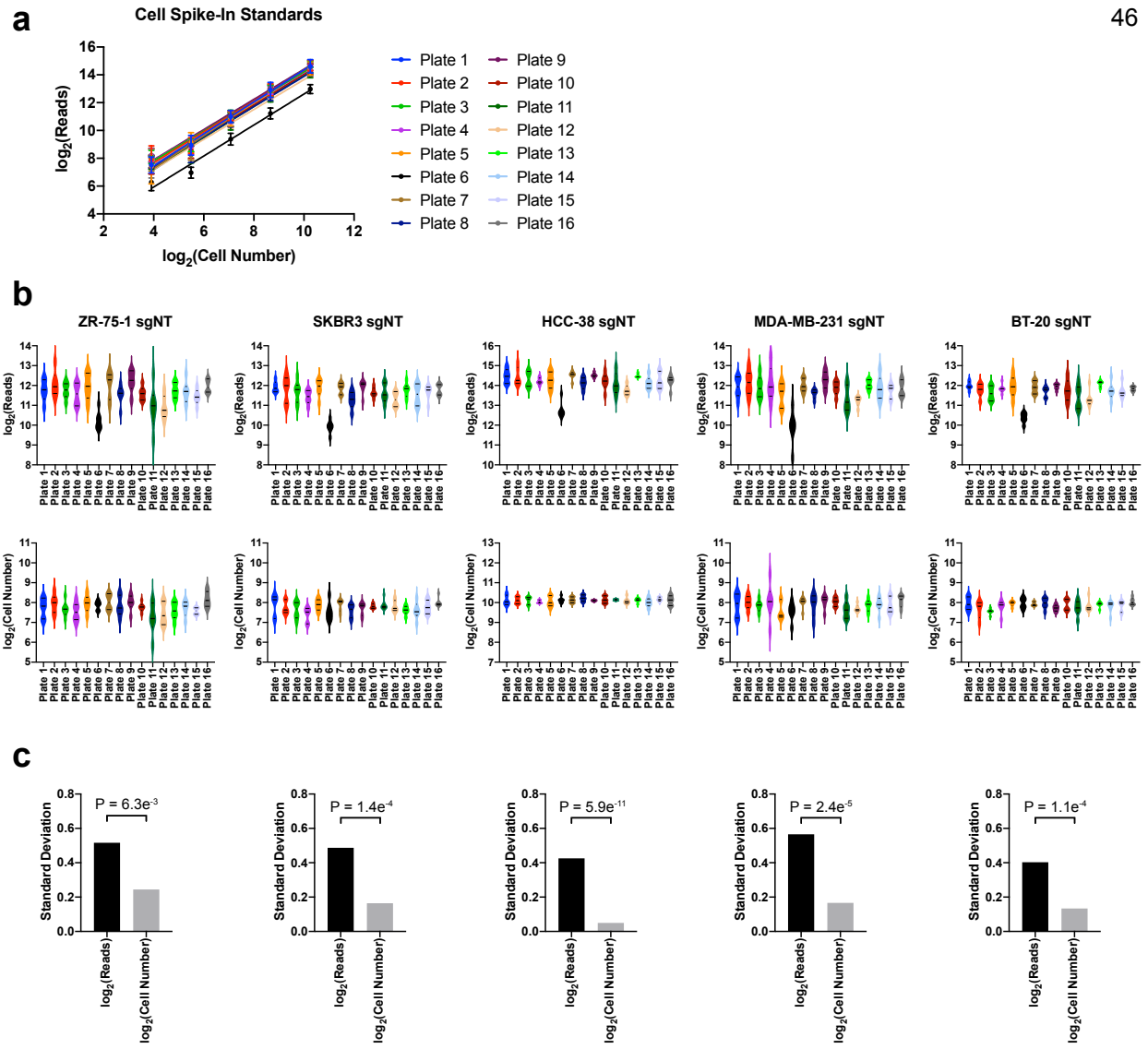


Figure 11. Cell spike-in standards reduce technical variation between plates. a Standard curves generated from five uniquely barcoded 293T cell spike-in standards introduced at known cell numbers. Data are represented as the mean number of sequencing reads across the six DMSO samples on a plate ± standard deviation. **b** Top: Violin plots indicating the distribution of sgNT sequencing reads for each cell line across the six DMSO samples on a plate. Bottom: Violin plots indicating the distribution of sgNT interpolated cell number for each cell line across the six DMSO samples on a plate. Thick black line denotes median. Thin black lines denote quartiles. **c** Standard deviation of sgNT sequencing reads between plates versus standard deviation of sgNT interpolated cell number between plates. Statistical significance was determined using an unpaired, two-tailed F test to compare variances (n = 16 plates).

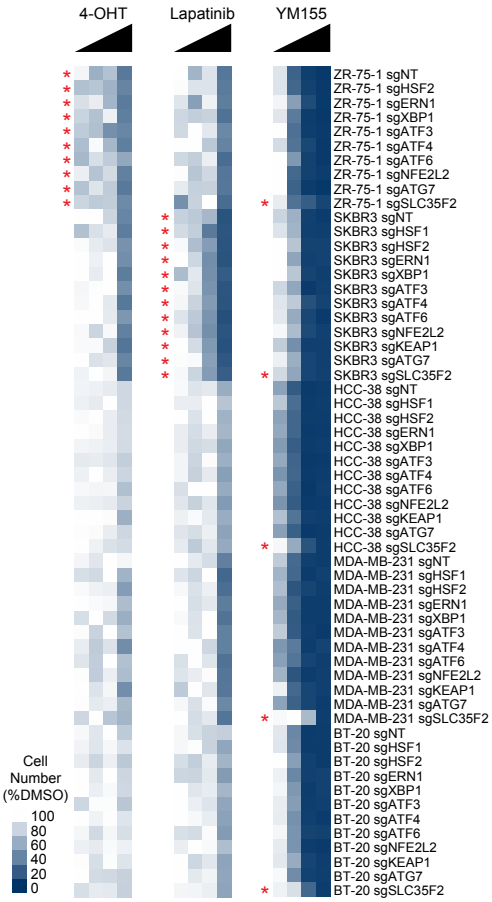


Figure 12. QMAP-Seq detects known cell line-specific and gene-specific drug vulnerabilities within complex mixtures of cells. Heat maps displaying the relative cell number for each cell line-sgRNA pair 72 hours after treatment with 4-Hydroxytamoxifen (4-OHT), Lapatinib, or YM155 as measured using QMAP-Seq. Data are represented as mean of two biologically independent replicates. Asterisks denote positive controls.

In total, a single next-generation sequencing run consisting of 1.3 billion reads provided the capacity to make 86,400 relative cell number measurements, plot 10,800 dose-response curves **(Figure 13)**, and calculate 21,600 AUCs, demonstrating the remarkable power of this assay. This dataset will hopefully serve as a valuable resource for the proteostasis community. Among the many chemical-genetic interactions we observed, we uncovered a synthetic lethal chemical-genetic interaction between the HSP90 inhibitor, NVPAUY922, and knockout of HSF1 in all cell lines **(Figure 13)**. Conversely, we identified a synthetic rescue chemical-genetic interaction between NVPAUY922 and knockout of KEAP1, a negative regulator of the proteostasis network, in all cell lines **(Figure 13)**. The magnitude of the AUC difference between sgHSF1 and sgNT varied across cell lines (HCC-38: -44.30, MDA-MB-231: -33.05, SKBR3: -16.15, BT-20: -2.00), but the cell lines with the largest effect size for sgHSF1 also had the largest effect size for sgKEAP1, but in the opposite direction (HCC-38: 49.85, MDA-MB-231: 7.55, SKBR3: 4.00, BT-20: 2.85), pointing to an anticorrelation between HSF1 and KEAP1 in the context of HSP90 inhibition that is consistent across all cell lines.

Profiling multiple cell lines enabled the identification of cell line-specific vulnerabilities. For example, while the proteasome inhibitor, Carfilzomib, was synthetic lethal with knockout of individual proteostasis factors in MDA-MB-231 cells, it had a modest synthetic rescue interaction with knockout of individual proteostasis factors in SKBR3 cells **(Figure 13)**. As expected, SKBR3 cells were preferentially sensitive to multiple EGFR inhibitors (Afatinib, Lapatinib, Neratinib) across all genetic perturbations **(Figure 13)**. As another example, HCC-38 and MDA-MB-231 cells were more resistant to the effects of RITA, a compound that inhibits the MDM2-p53 interaction, compared to the other three cell lines **(Figure 13)**, but the functional significance of the differential cell line sensitivity remains to be determined.

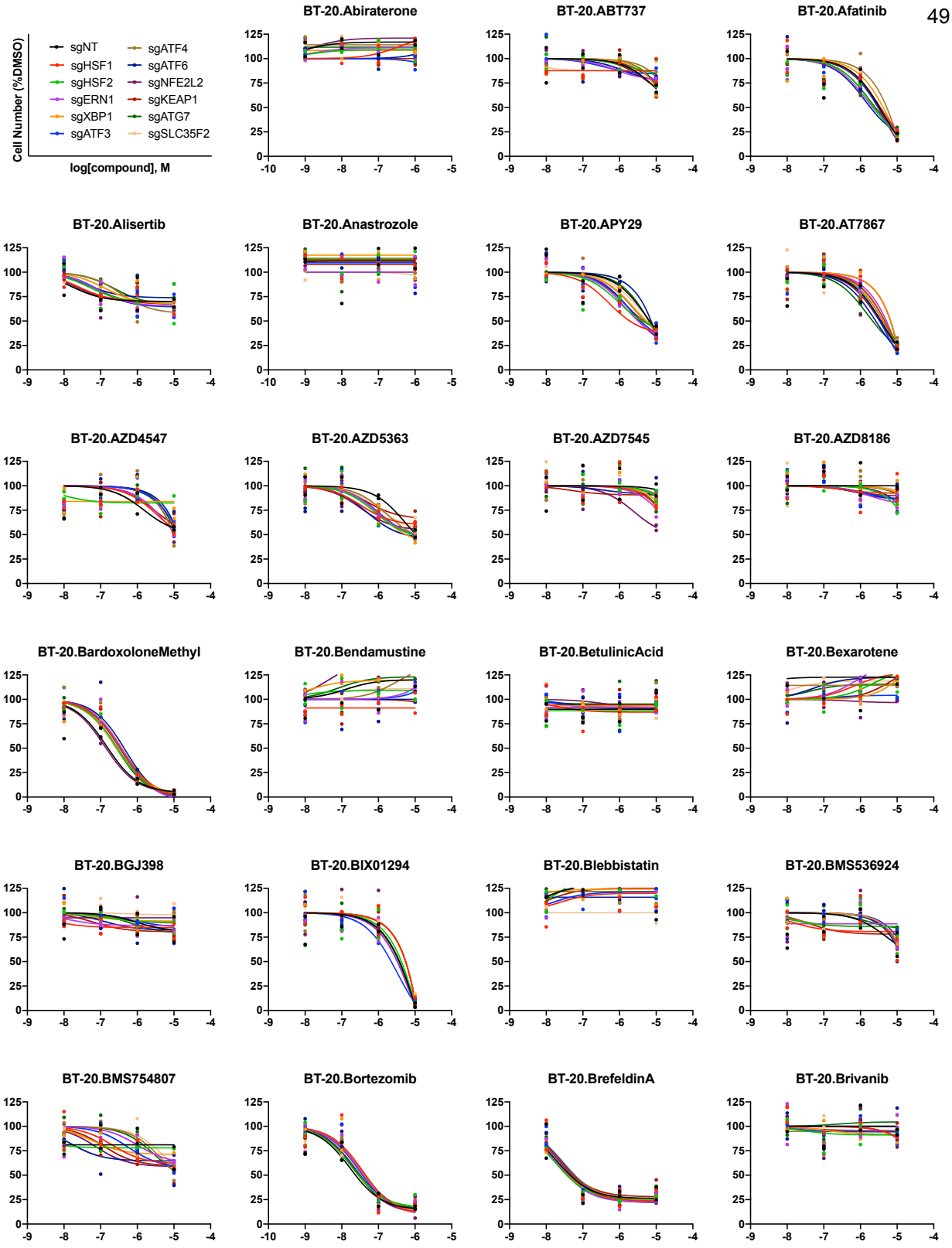


Figure 13. Dose-response curves generated using QMAP-Seq. Each plot represents a unique cell line-compound combination. Each data point represents one of two biologically independent replicates for the indicated sgRNA.

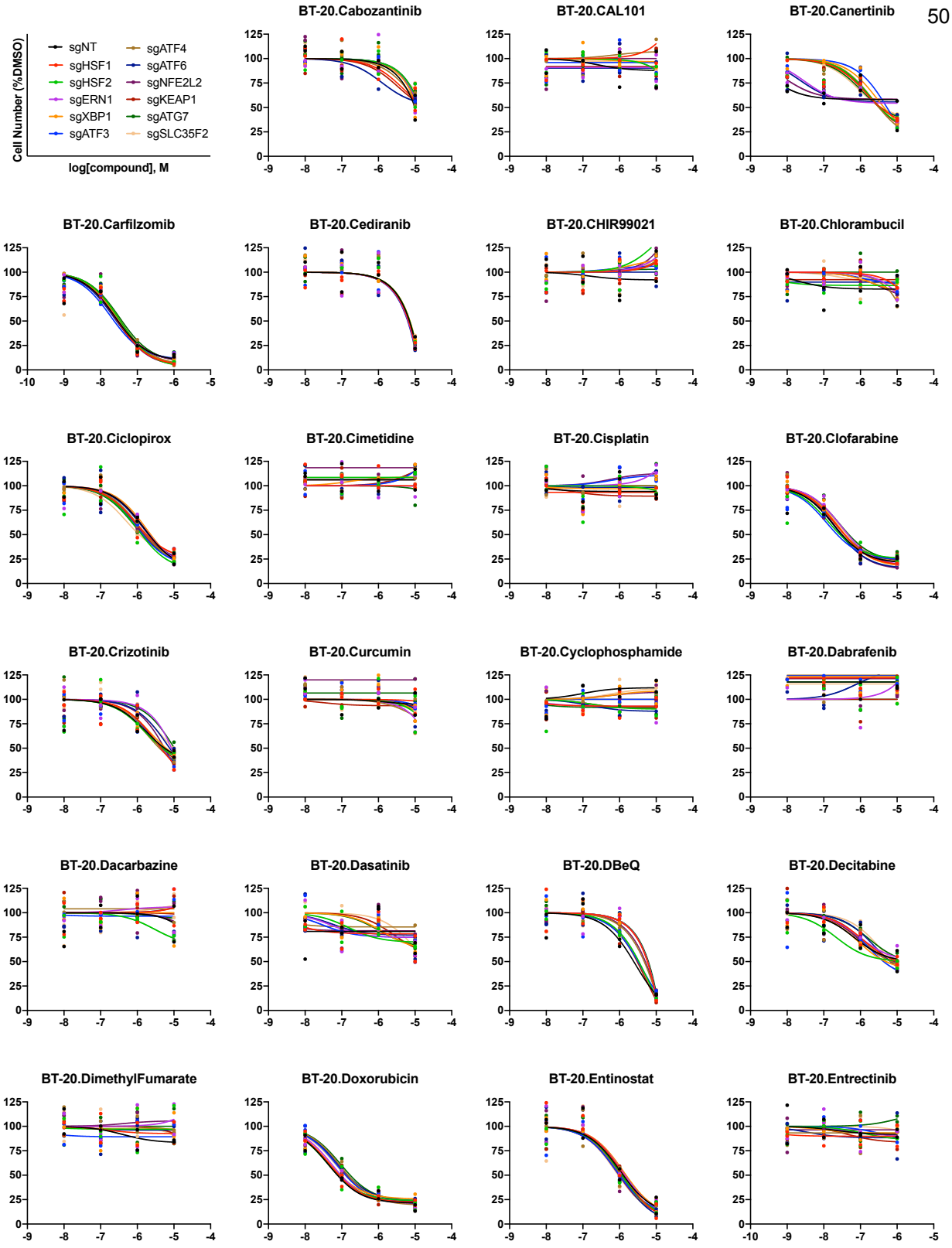


Figure 13 (continued). Dose-response curves generated using QMAP-Seq. Each plot represents a unique cell line-compound combination. Each data point represents one of two biologically independent replicates for the indicated sgRNA.

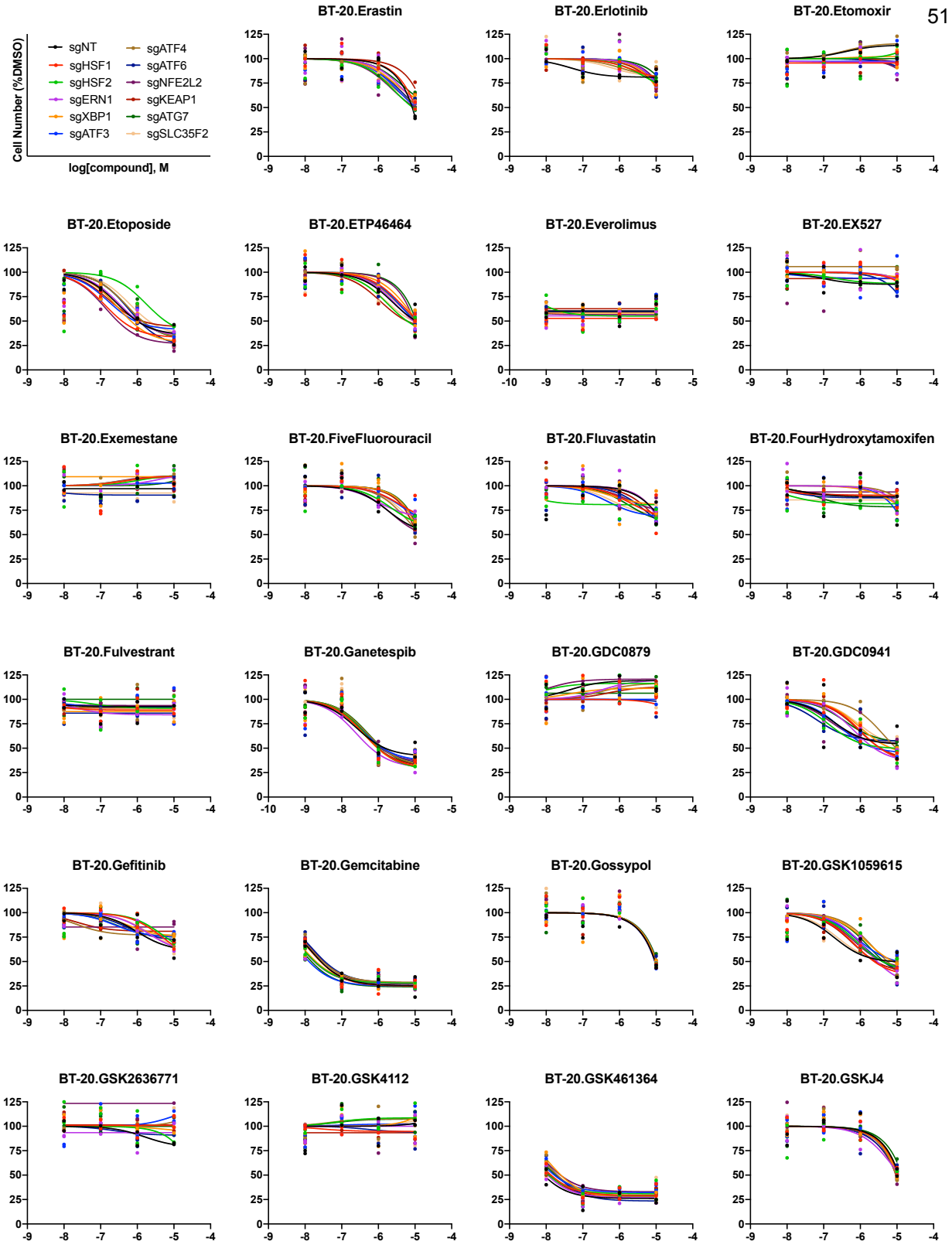


Figure 13 (continued). Dose-response curves generated using QMAP-Seq. Each plot represents a unique cell line-compound combination. Each data point represents one of two biologically independent replicates for the indicated sgRNA.

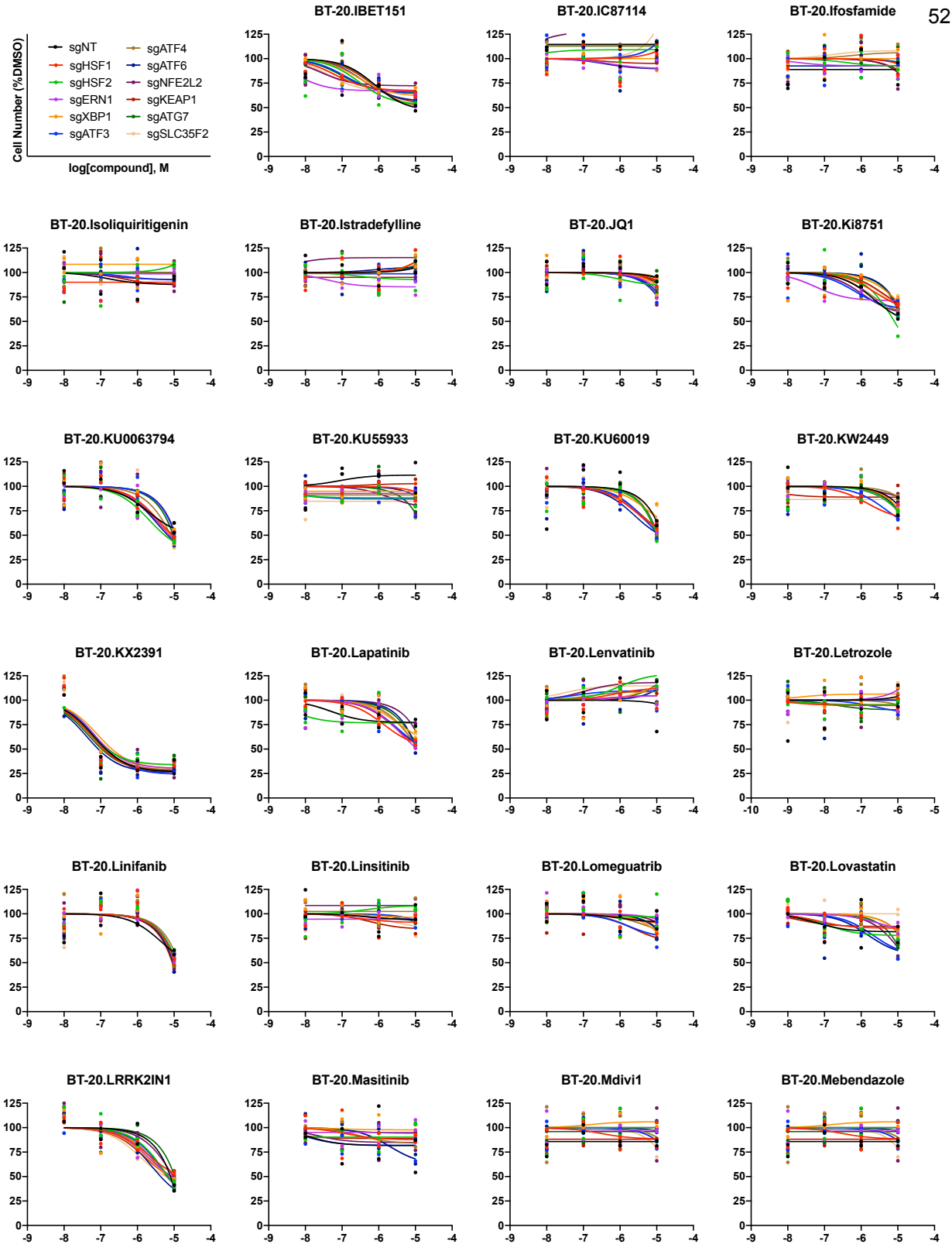


Figure 13 (continued). Dose-response curves generated using QMAP-Seq. Each plot represents a unique cell line-compound combination. Each data point represents one of two biologically independent replicates for the indicated sgRNA.

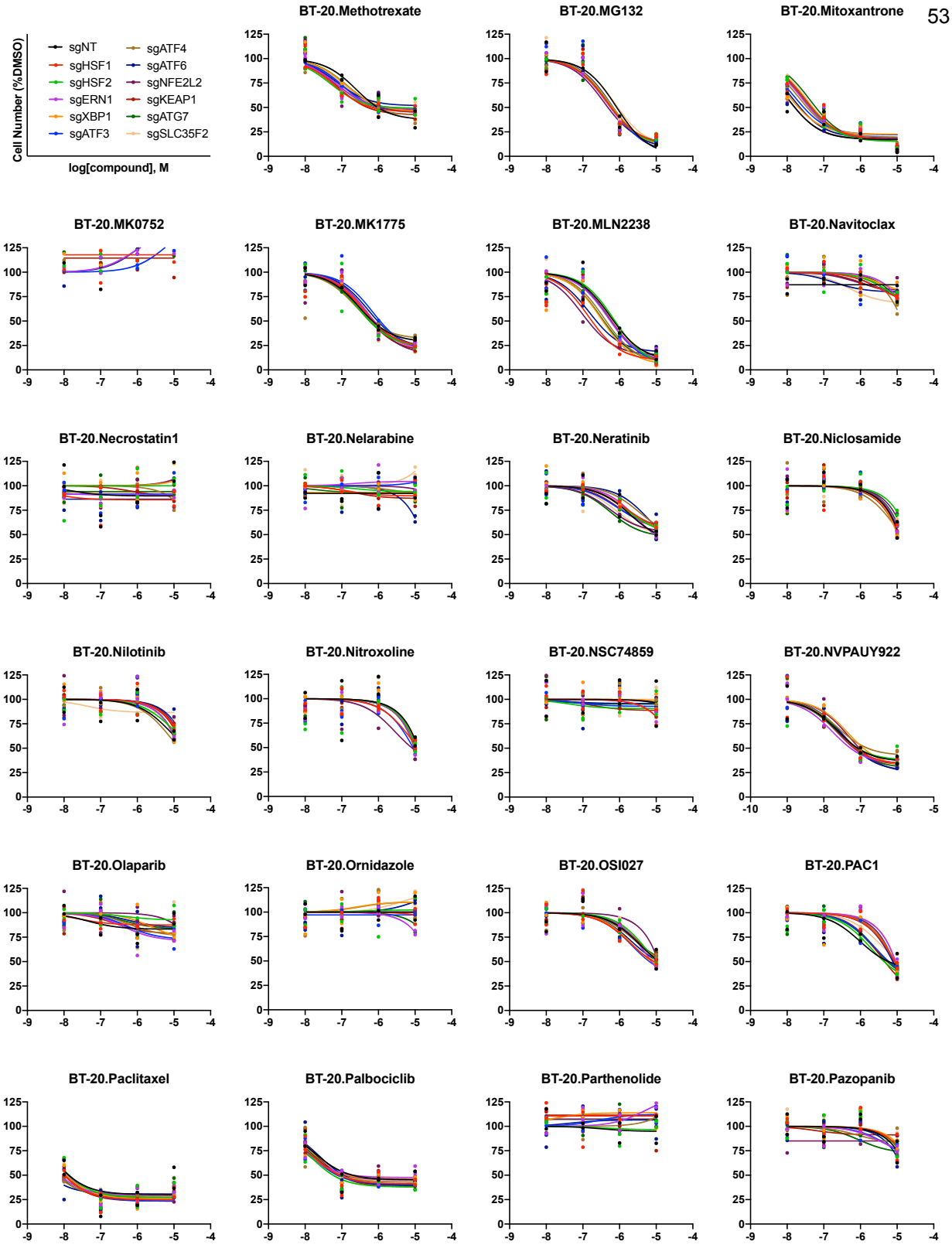


Figure 13 (continued). Dose-response curves generated using QMAP-Seq. Each plot represents a unique cell line-compound combination. Each data point represents one of two biologically independent replicates for the indicated sgRNA.

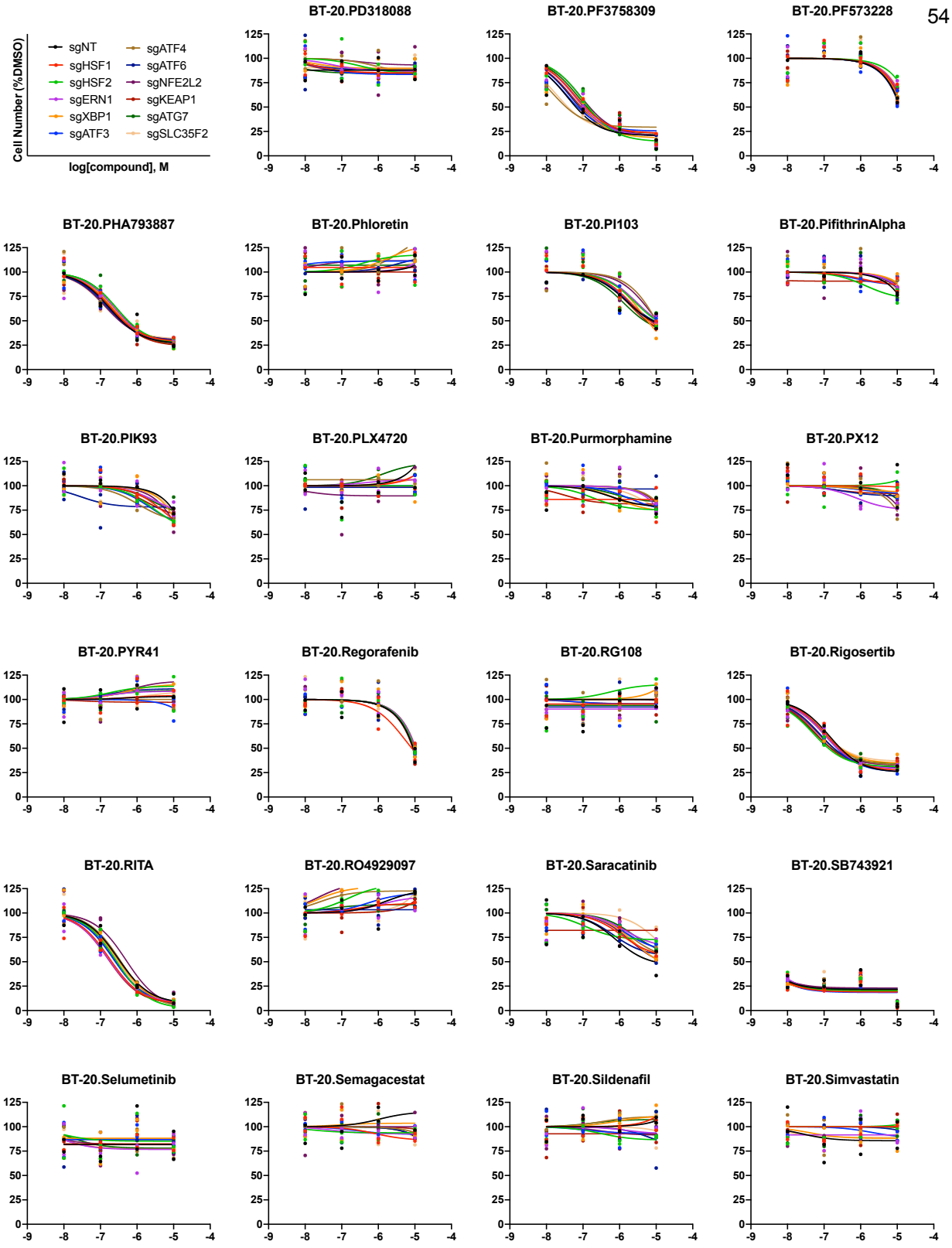


Figure 13 (continued). Dose-response curves generated using QMAP-Seq. Each plot represents a unique cell line-compound combination. Each data point represents one of two biologically independent replicates for the indicated sgRNA.

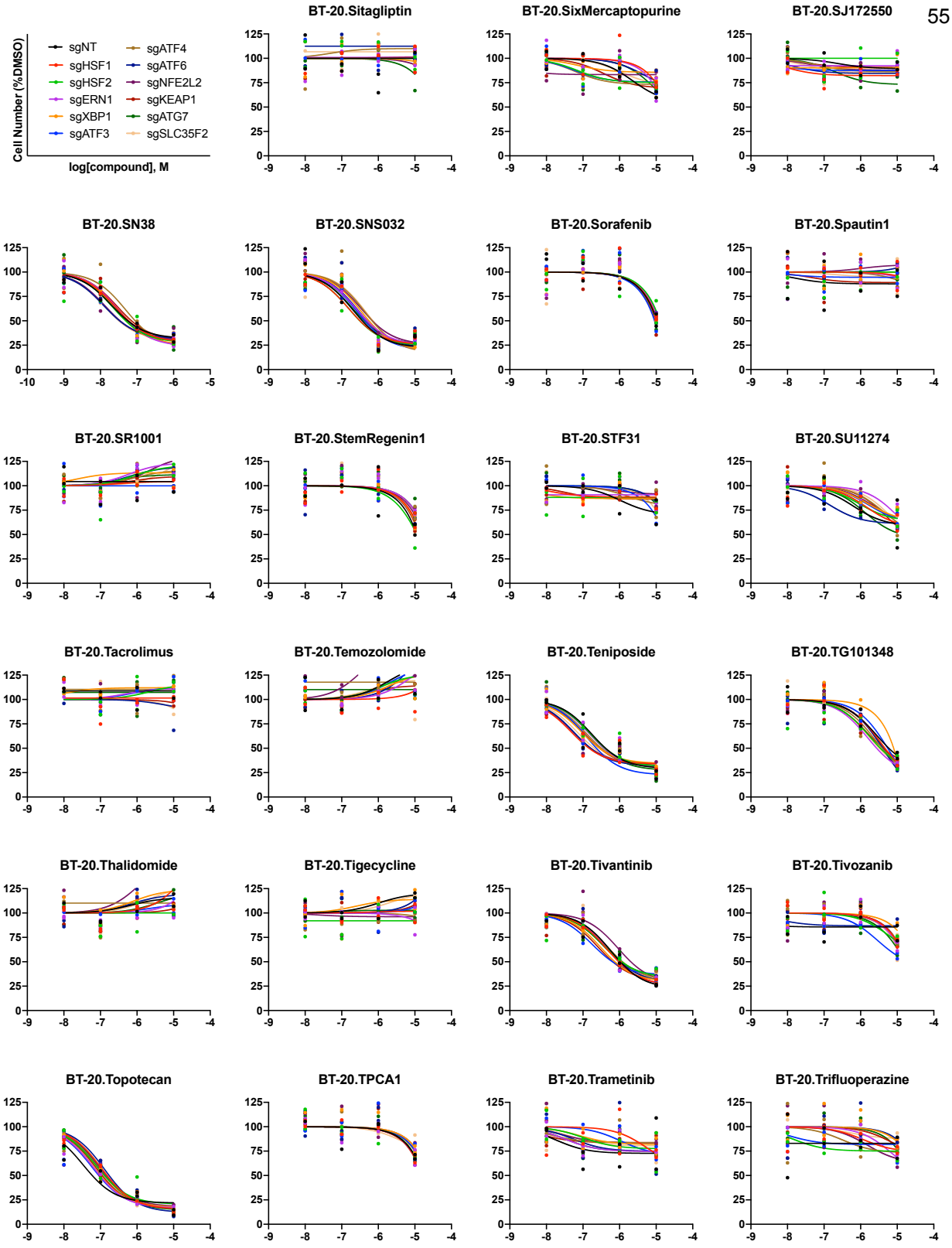


Figure 13 (continued). Dose-response curves generated using QMAP-Seq. Each plot represents a unique cell line-compound combination. Each data point represents one of two biologically independent replicates for the indicated sgRNA.

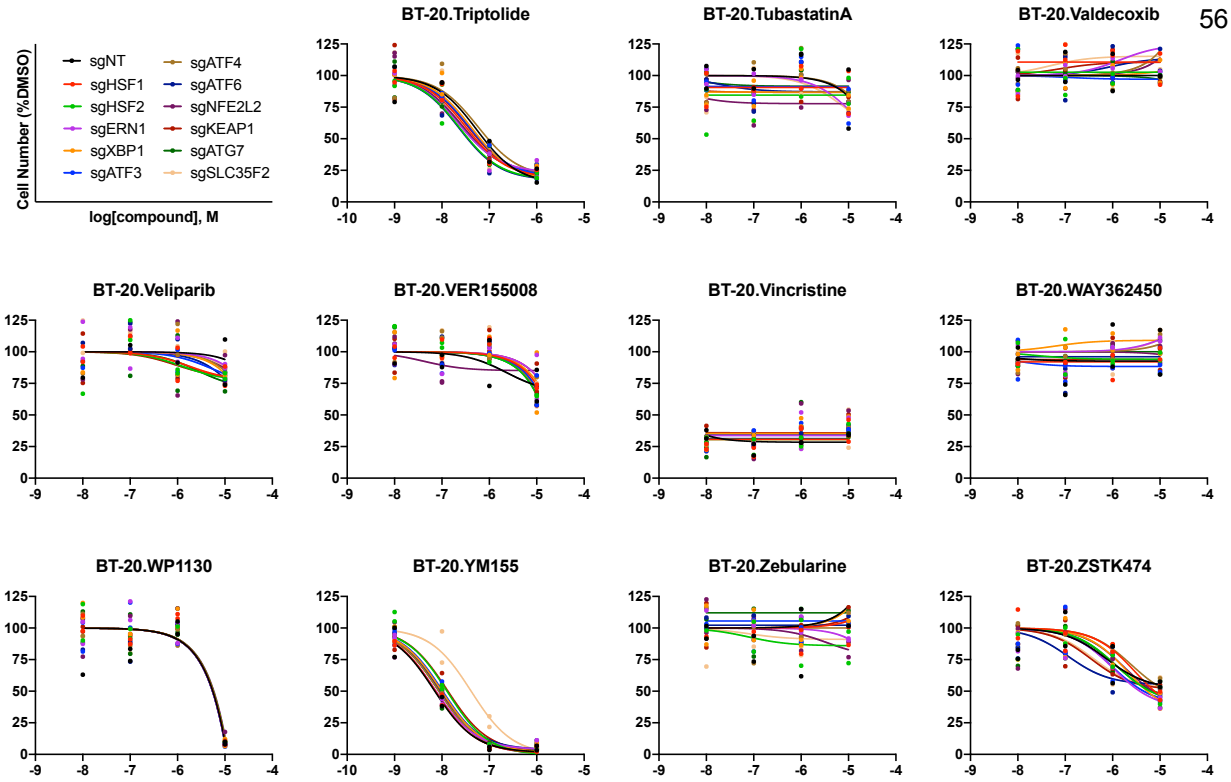


Figure 13 (continued). Dose-response curves generated using QMAP-Seq. Each plot represents a unique cell line-compound combination. Each data point represents one of two biologically independent replicates for the indicated sgRNA.

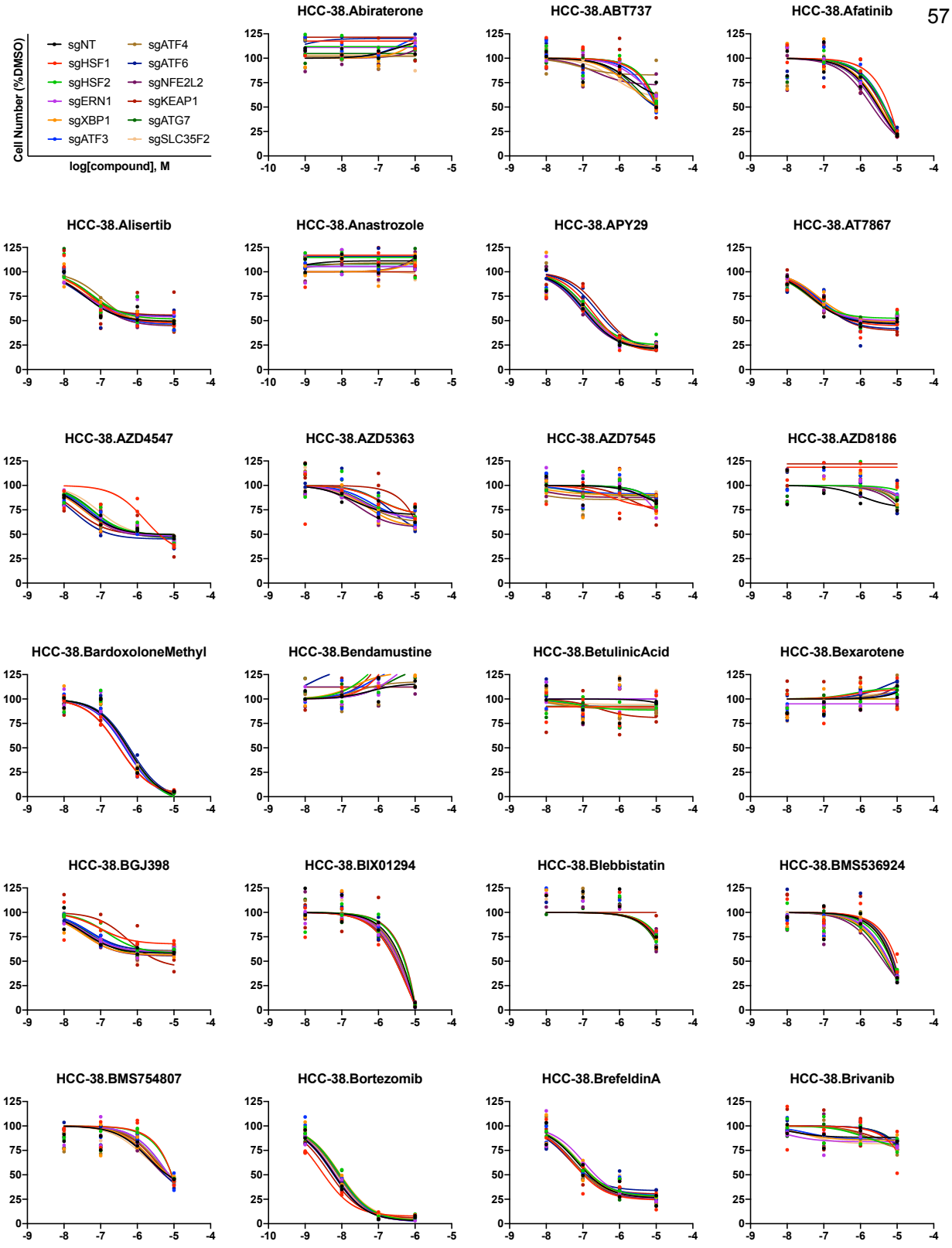


Figure 13 (continued). Dose-response curves generated using QMAP-Seq. Each plot represents a unique cell line-compound combination. Each data point represents one of two biologically independent replicates for the indicated sgRNA.

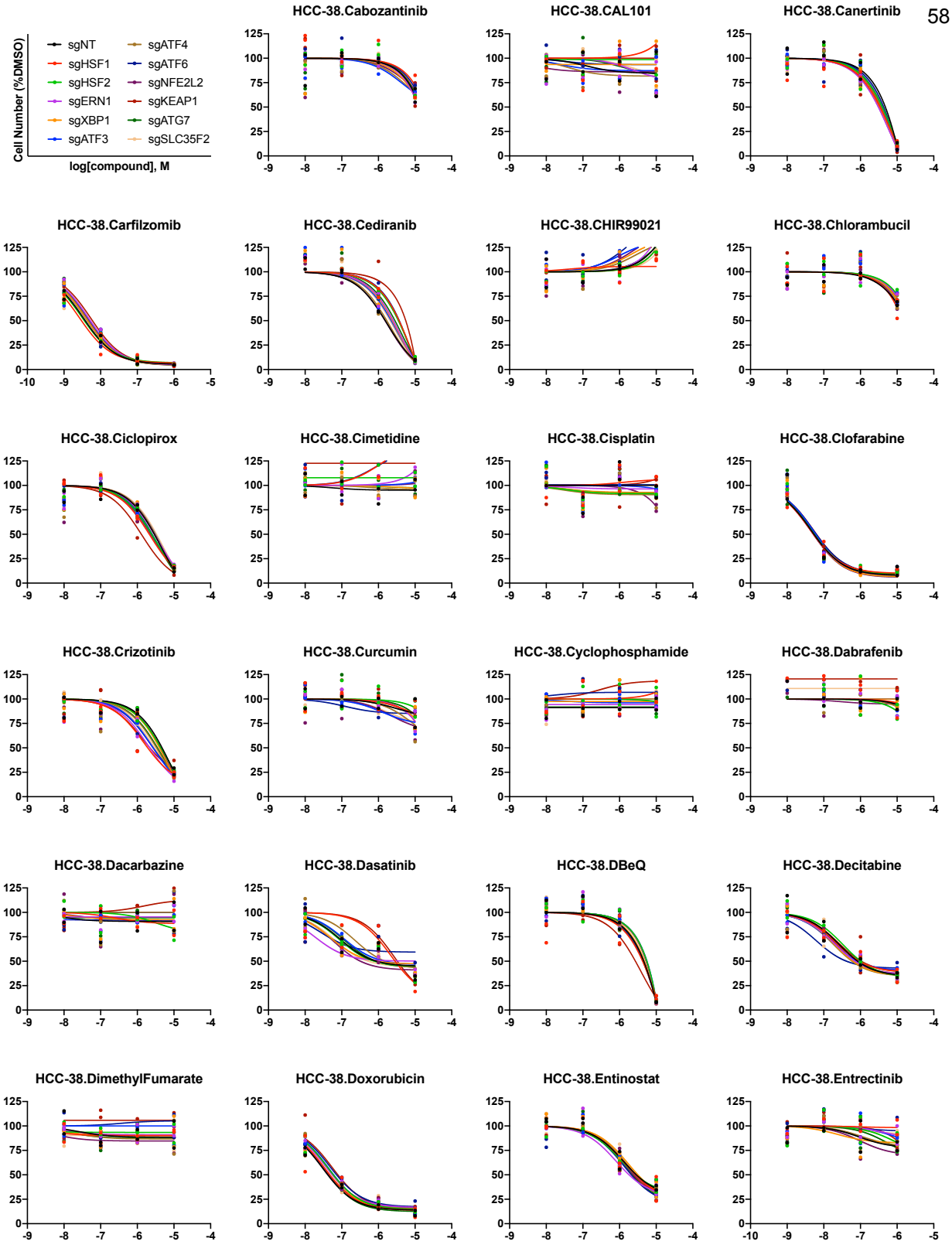


Figure 13 (continued). Dose-response curves generated using QMAP-Seq. Each plot represents a unique cell line-compound combination. Each data point represents one of two biologically independent replicates for the indicated sgRNA.

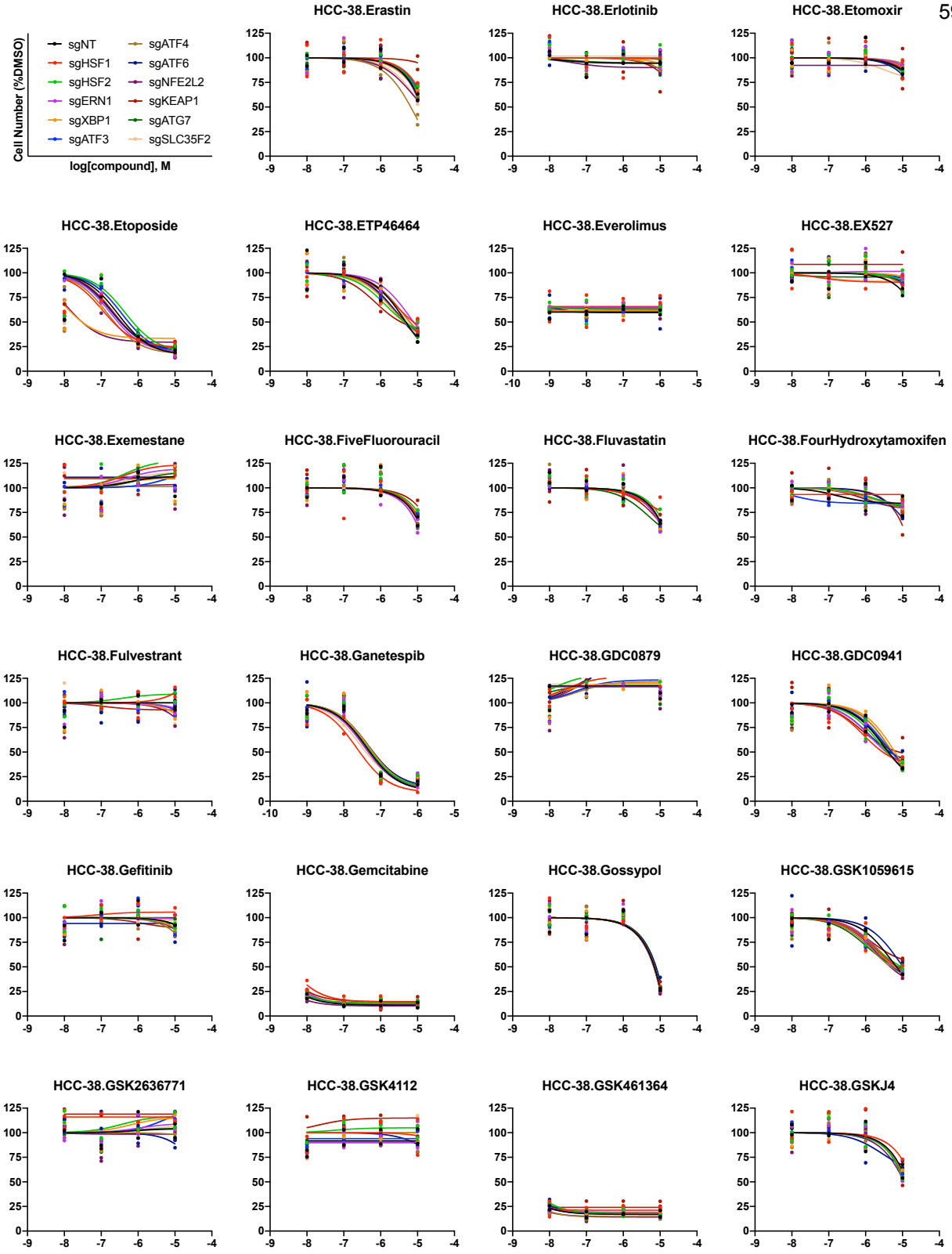


Figure 13 (continued). Dose-response curves generated using QMAP-Seq. Each plot represents a unique cell line-compound combination. Each data point represents one of two biologically independent replicates for the indicated sgRNA.

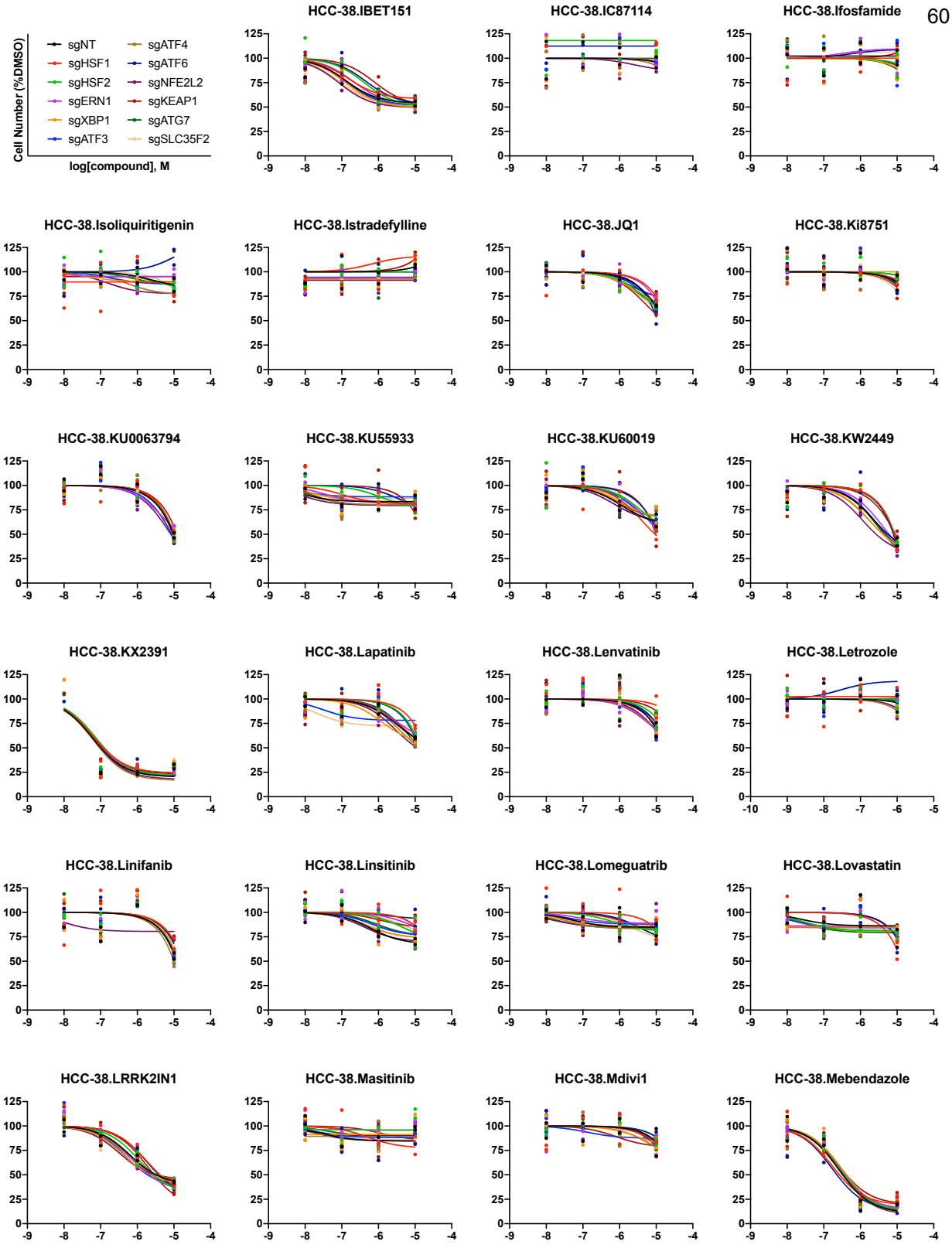


Figure 13 (continued). Dose-response curves generated using QMAP-Seq. Each plot represents a unique cell line-compound combination. Each data point represents one of two biologically independent replicates for the indicated sgRNA.

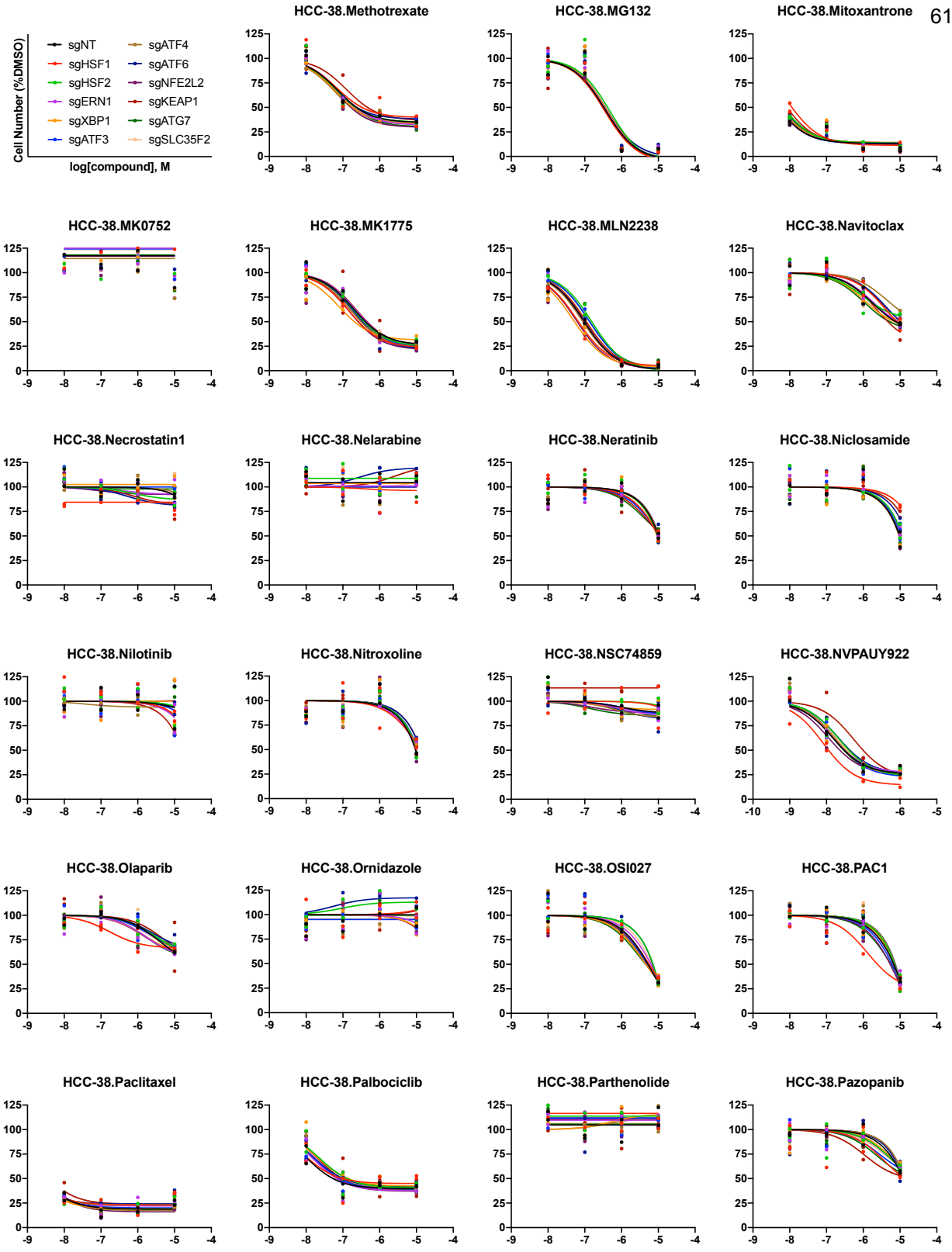


Figure 13 (continued). Dose-response curves generated using QMAP-Seq. Each plot represents a unique cell line-compound combination. Each data point represents one of two biologically independent replicates for the indicated sgRNA.

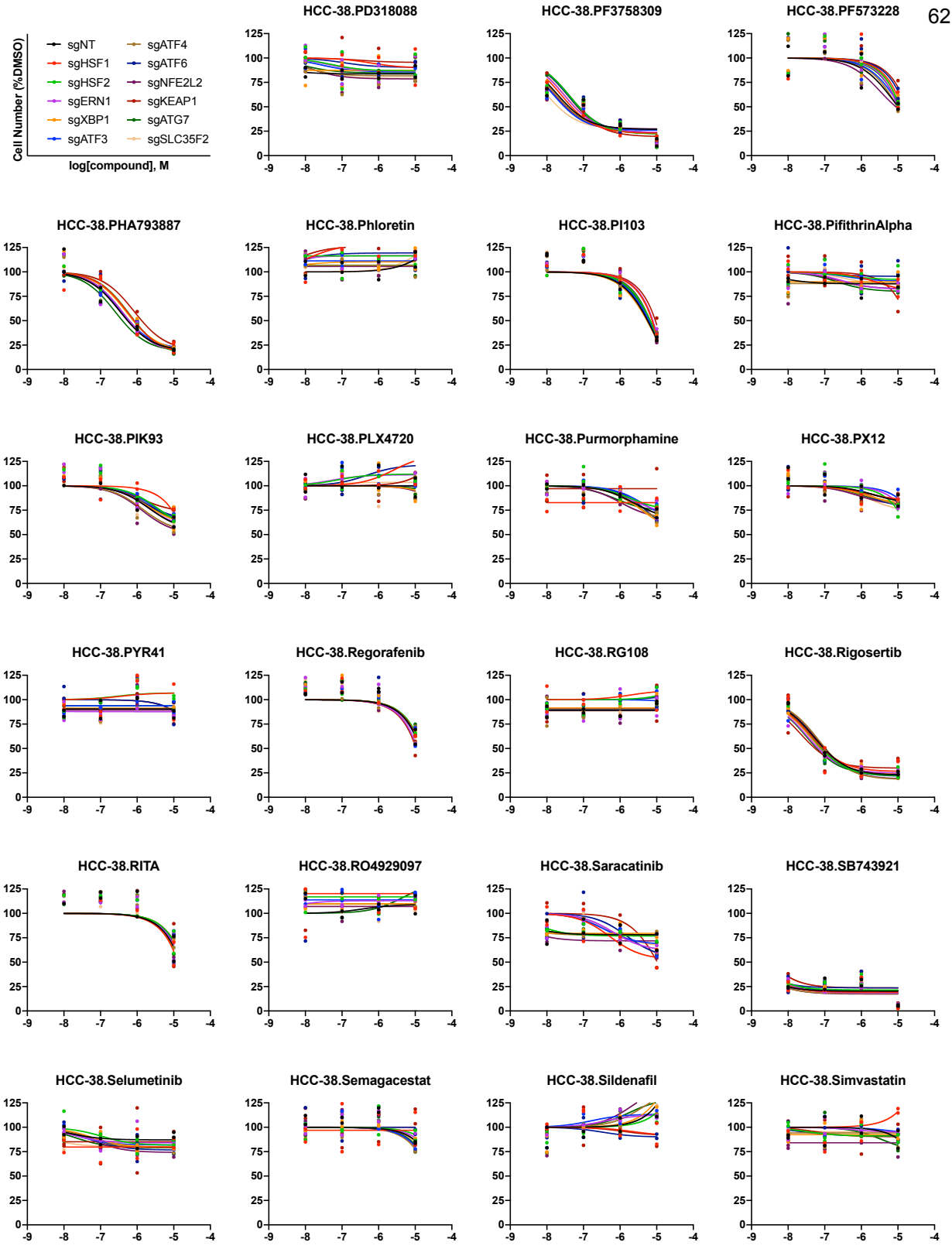


Figure 13 (continued). Dose-response curves generated using QMAP-Seq. Each plot represents a unique cell line-compound combination. Each data point represents one of two biologically independent replicates for the indicated sgRNA.

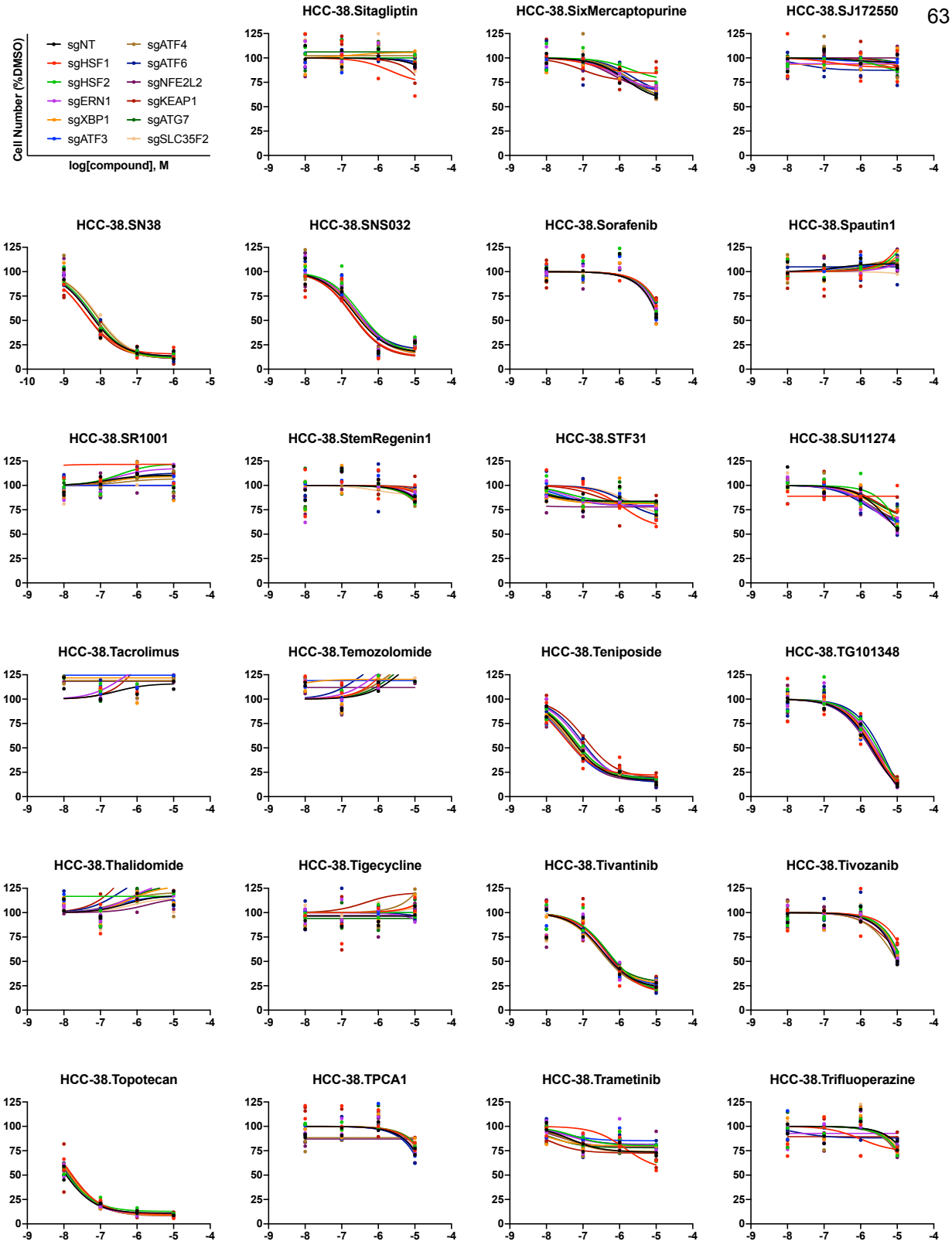


Figure 13 (continued). Dose-response curves generated using QMAP-Seq. Each plot represents a unique cell line-compound combination. Each data point represents one of two biologically independent replicates for the indicated sgRNA.

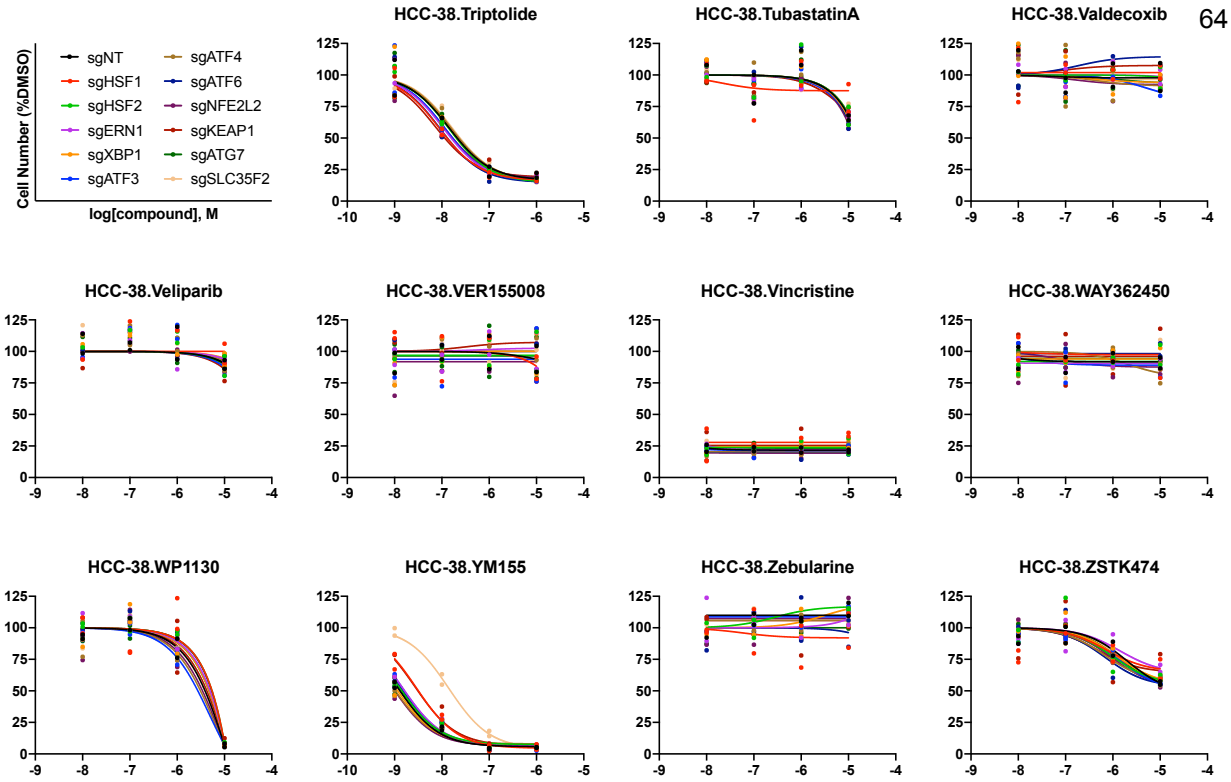


Figure 13 (continued). Dose-response curves generated using QMAP-Seq. Each plot represents a unique cell line-compound combination. Each data point represents one of two biologically independent replicates for the indicated sgRNA.

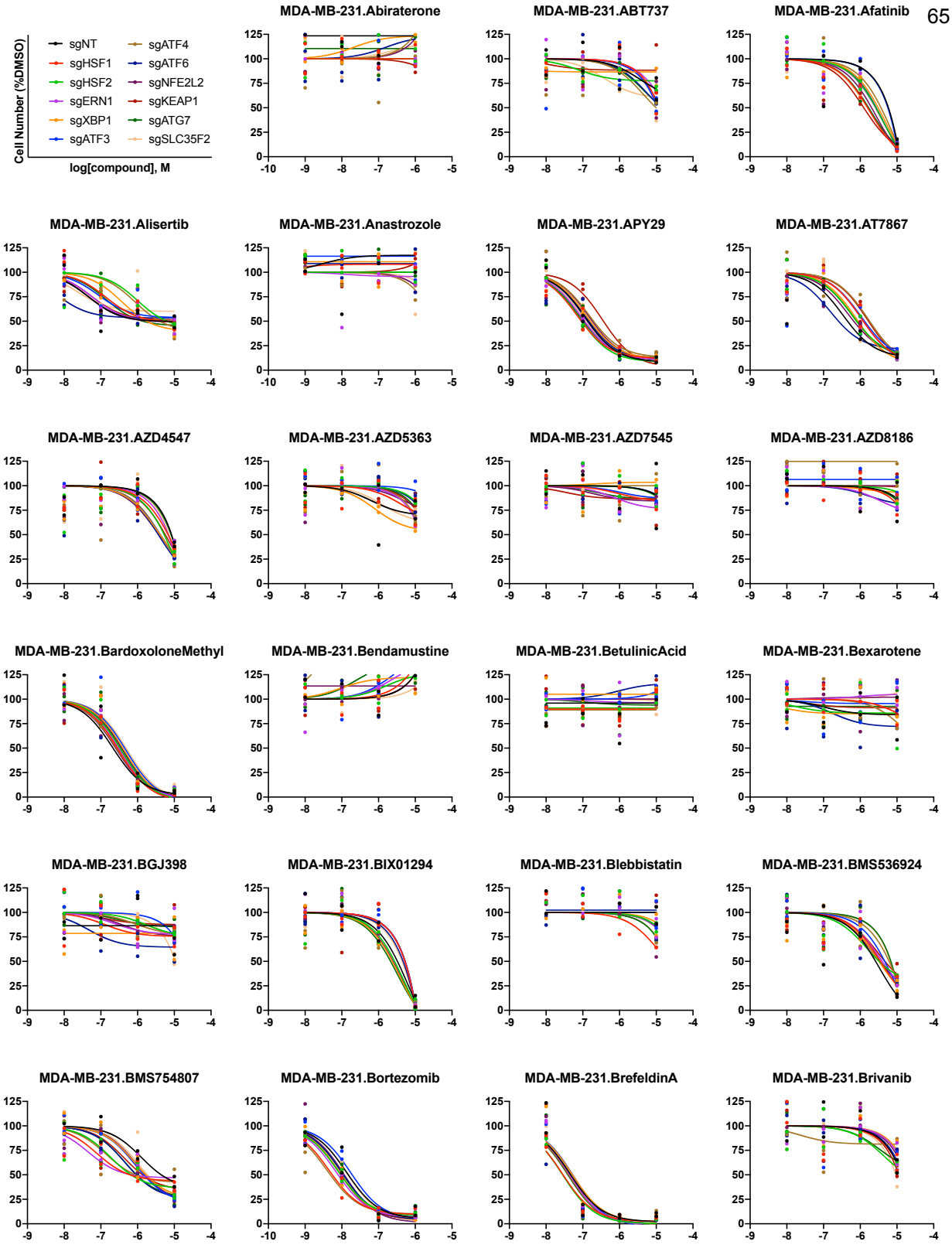


Figure 13 (continued). Dose-response curves generated using QMAP-Seq. Each plot represents a unique cell line-compound combination. Each data point represents one of two biologically independent replicates for the indicated sgRNA.

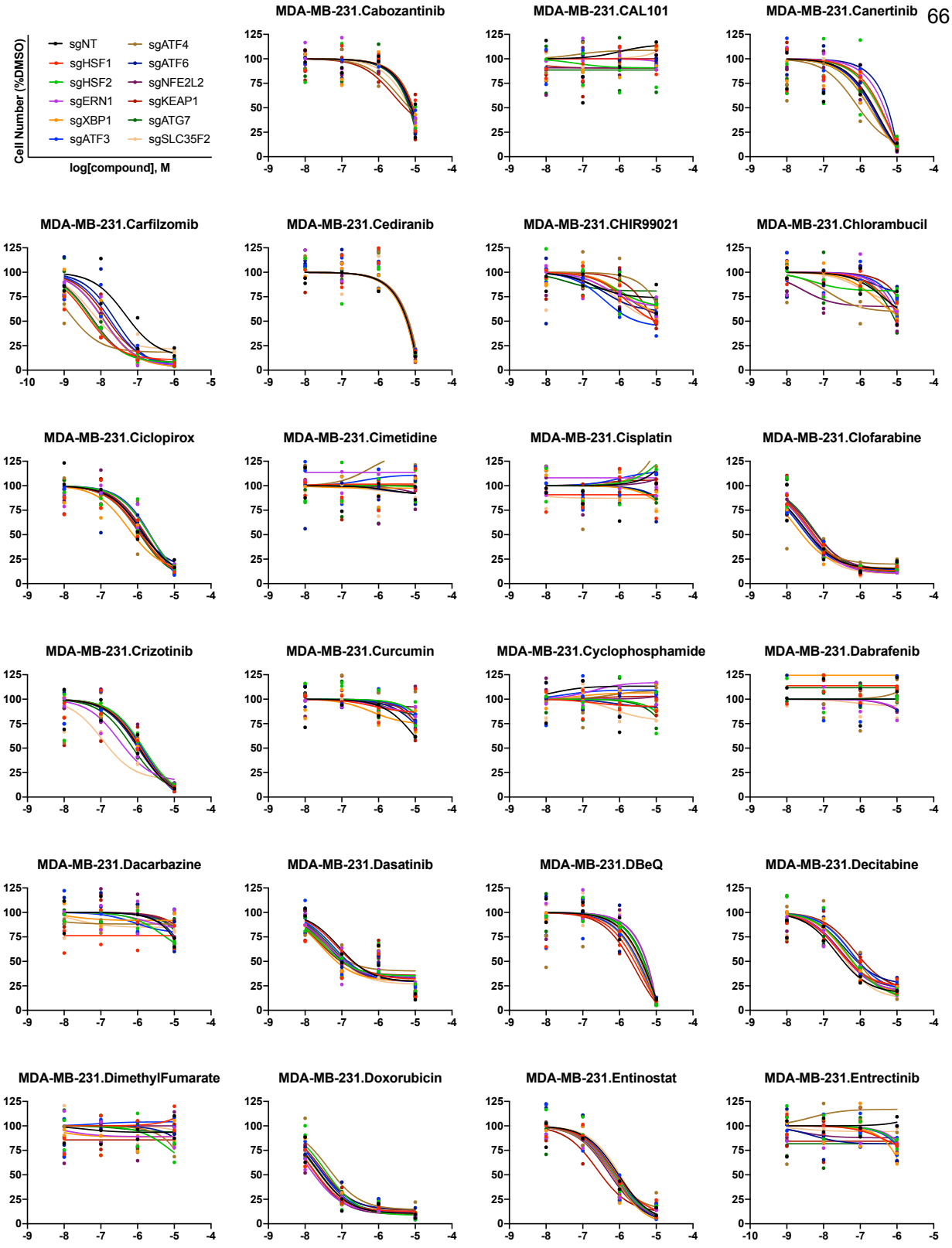


Figure 13 (continued). Dose-response curves generated using QMAP-Seq. Each plot represents a unique cell line-compound combination. Each data point represents one of two biologically independent replicates for the indicated sgRNA.

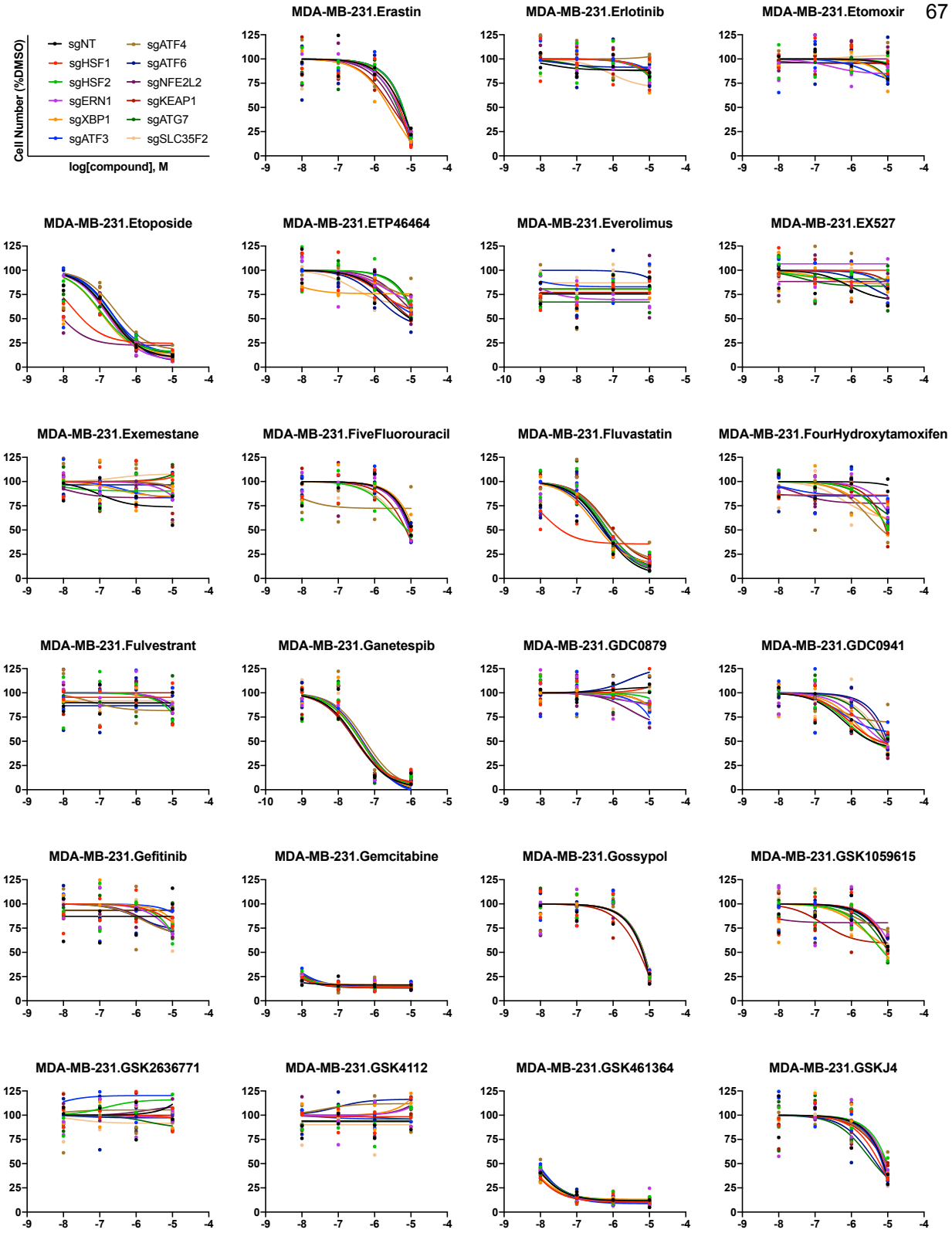


Figure 13 (continued). Dose-response curves generated using QMAP-Seq. Each plot represents a unique cell line-compound combination. Each data point represents one of two biologically independent replicates for the indicated sgRNA.

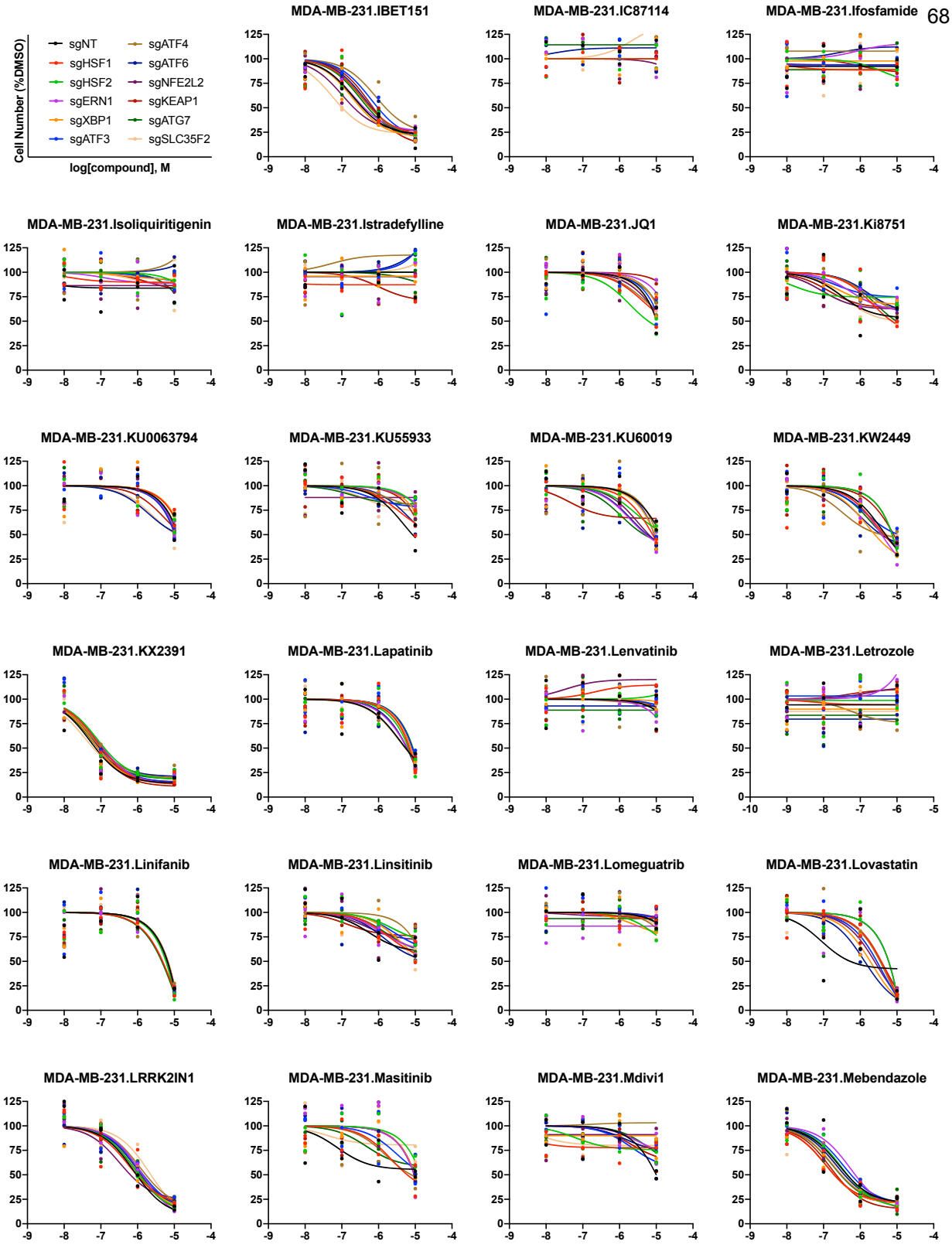


Figure 13 (continued). Dose-response curves generated using QMAP-Seq. Each plot represents a unique cell line-compound combination. Each data point represents one of two biologically independent replicates for the indicated sgRNA.

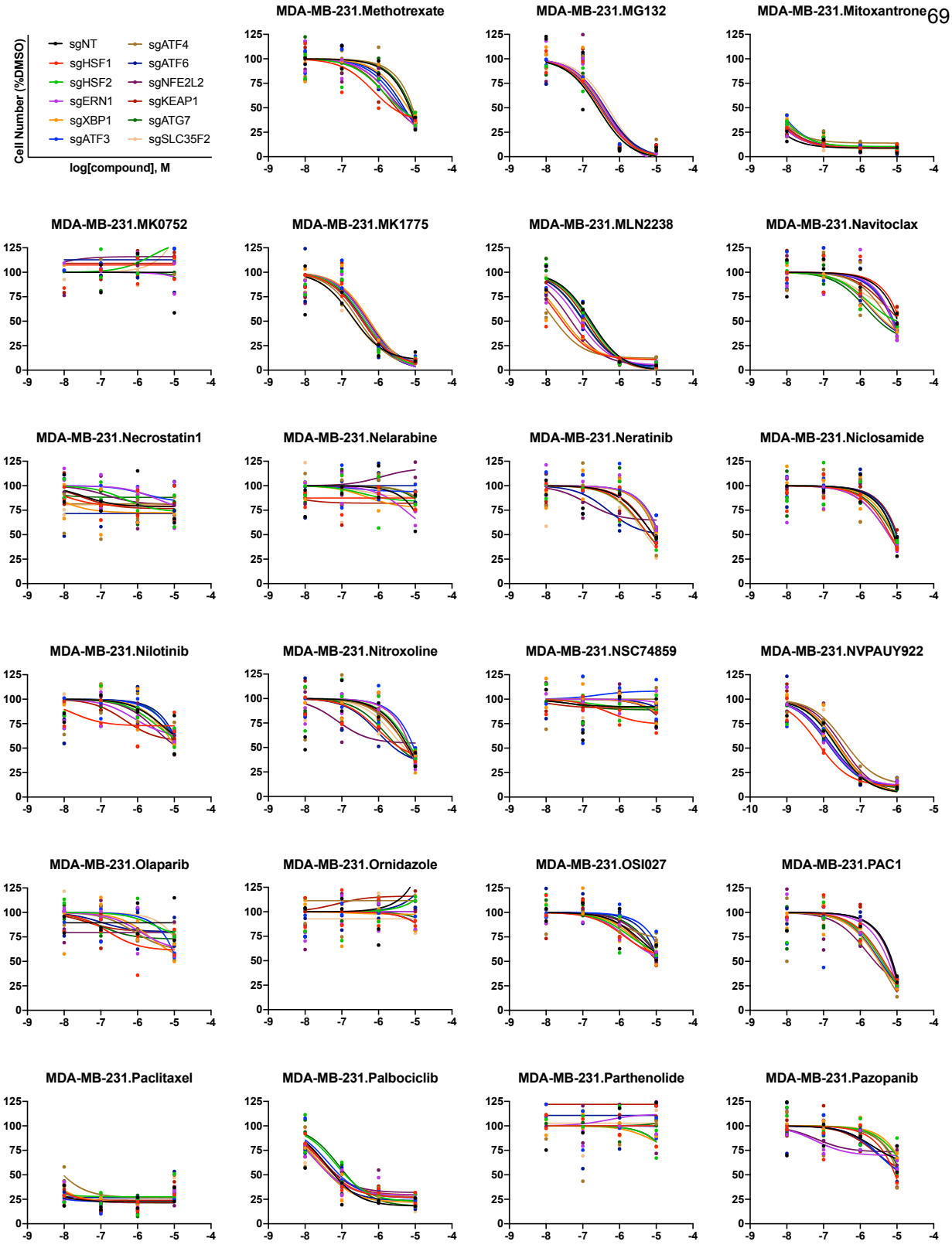


Figure 13 (continued). Dose-response curves generated using QMAP-Seq. Each plot represents a unique cell line-compound combination. Each data point represents one of two biologically independent replicates for the indicated sgRNA.

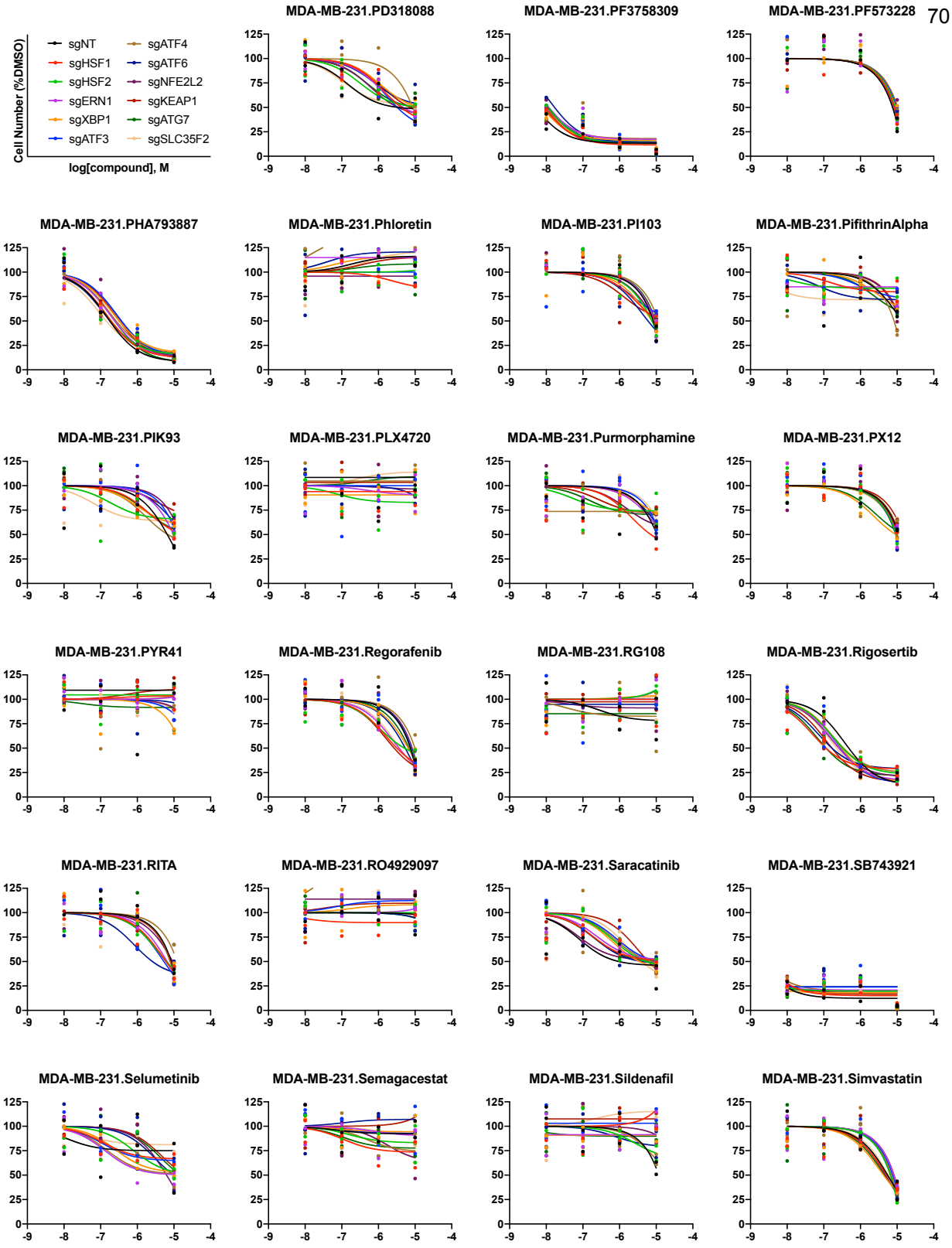


Figure 13 (continued). Dose-response curves generated using QMAP-Seq. Each plot represents a unique cell line-compound combination. Each data point represents one of two biologically independent replicates for the indicated sgRNA.

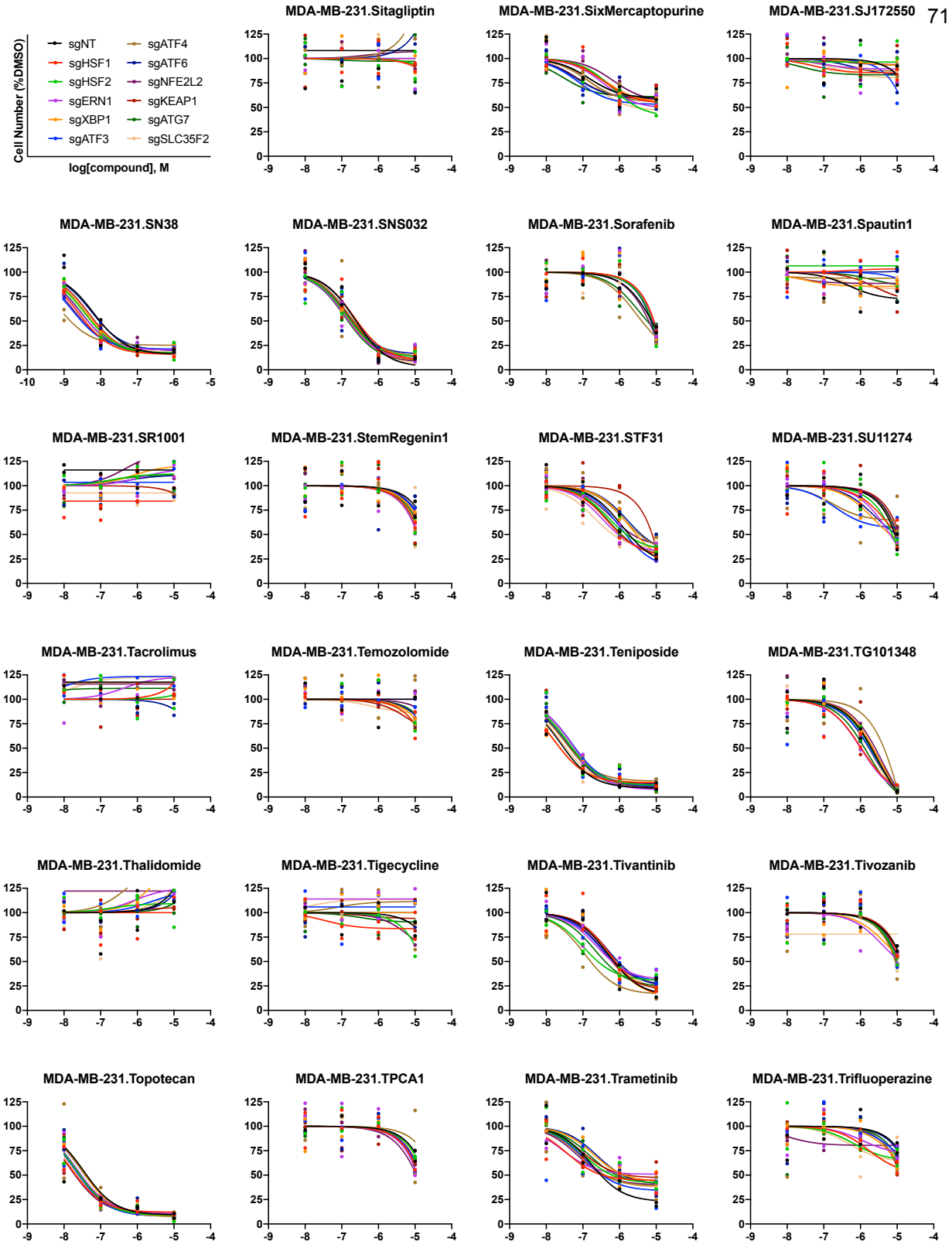


Figure 13 (continued). Dose-response curves generated using QMAP-Seq. Each plot represents a unique cell line-compound combination. Each data point represents one of two biologically independent replicates for the indicated sgRNA.

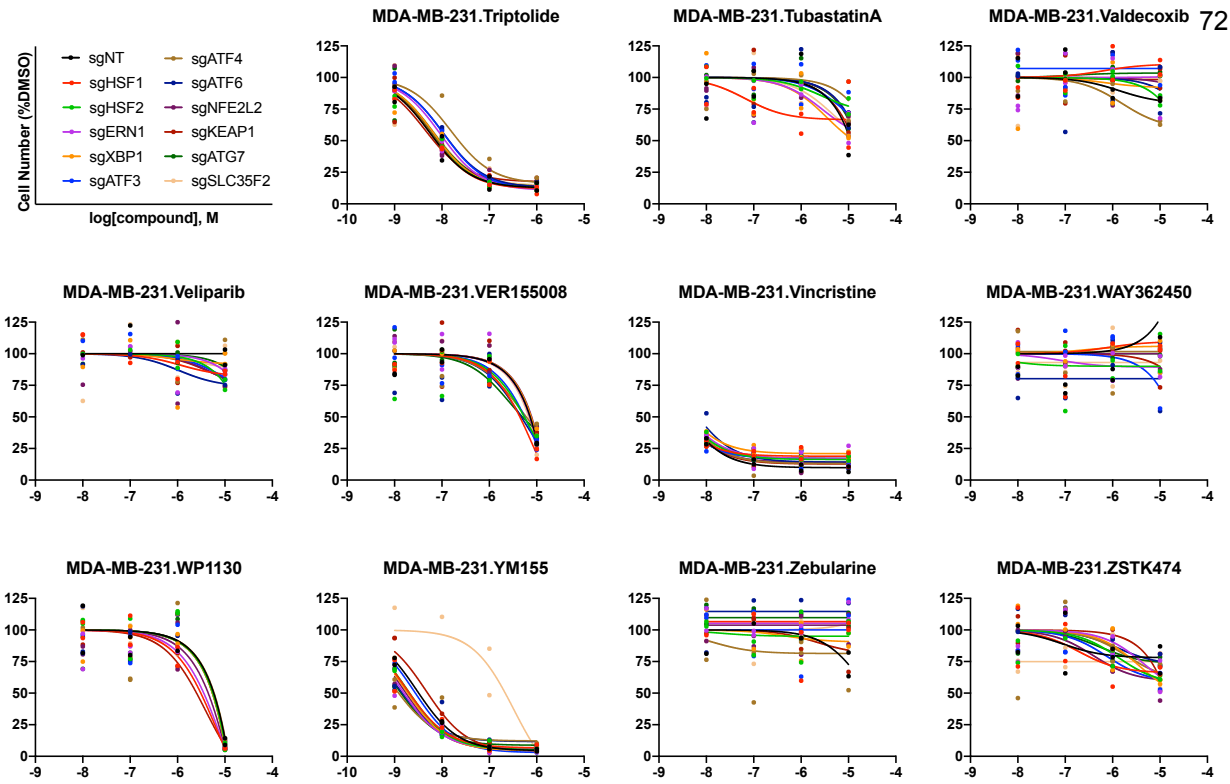


Figure 13 (continued). Dose-response curves generated using QMAP-Seq. Each plot represents a unique cell line-compound combination. Each data point represents one of two biologically independent replicates for the indicated sgRNA.

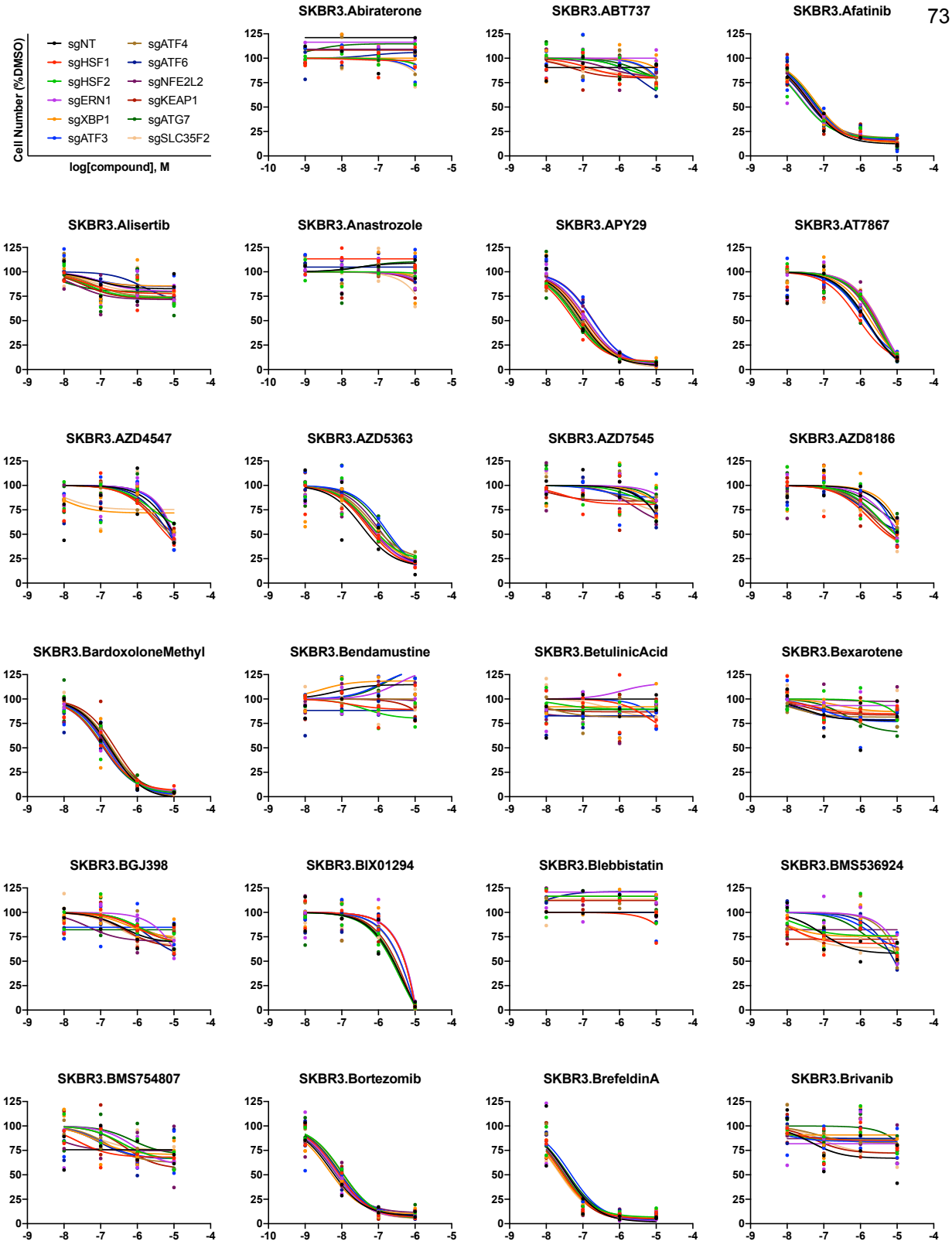


Figure 13 (continued). Dose-response curves generated using QMAP-Seq. Each plot represents a unique cell line-compound combination. Each data point represents one of two biologically independent replicates for the indicated sgRNA.

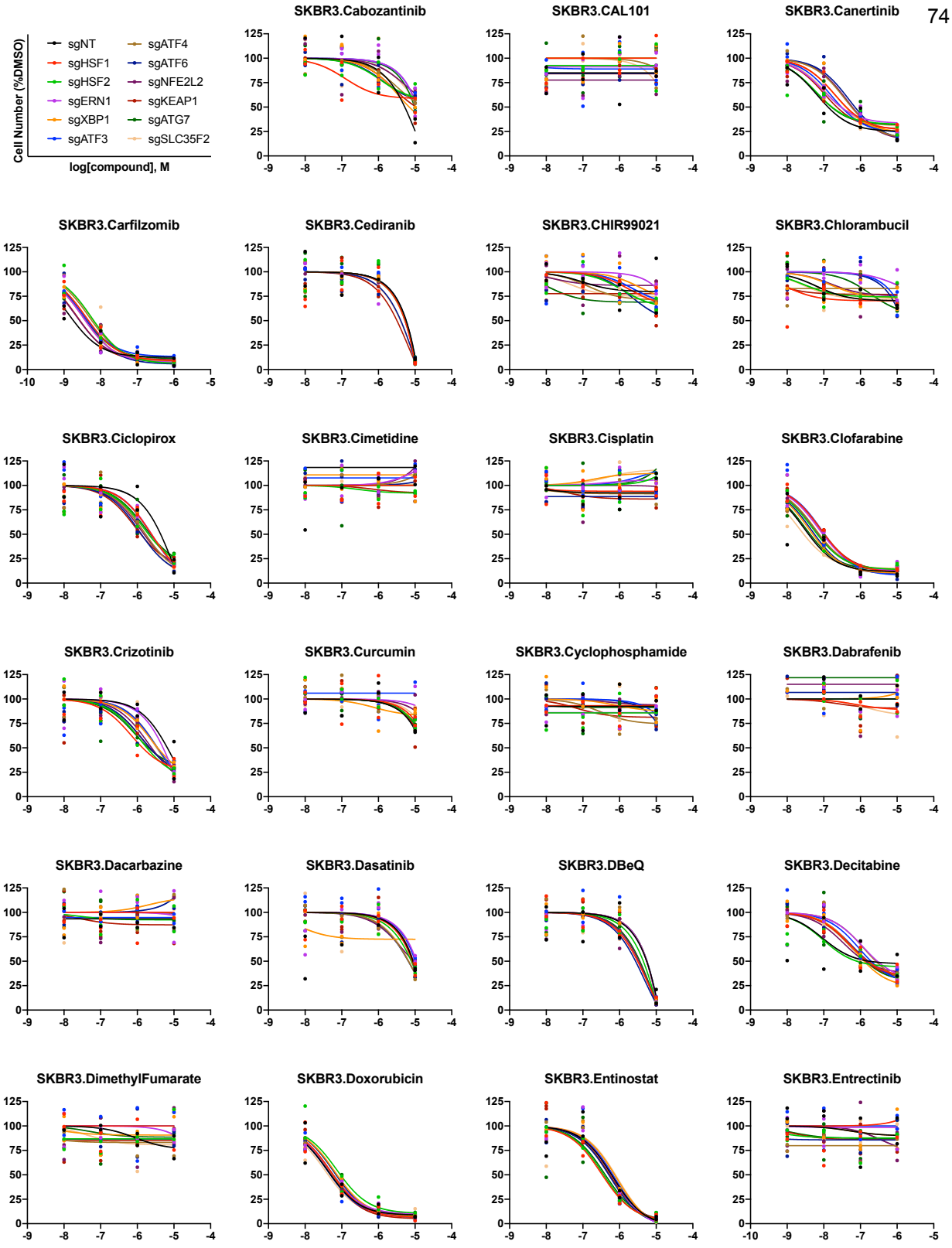


Figure 13 (continued). Dose-response curves generated using QMAP-Seq. Each plot represents a unique cell line-compound combination. Each data point represents one of two biologically independent replicates for the indicated sgRNA.

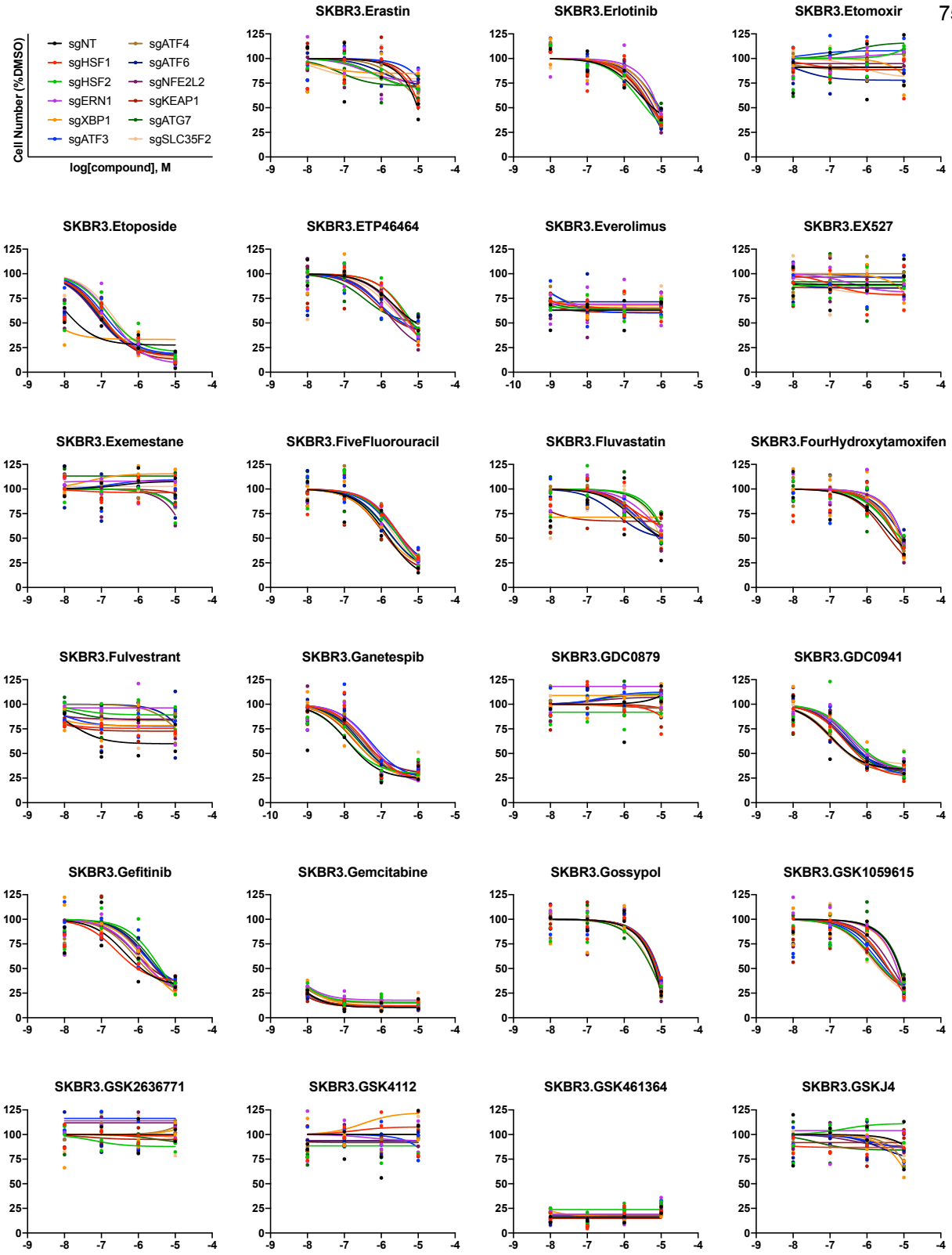


Figure 13 (continued). Dose-response curves generated using QMAP-Seq. Each plot represents a unique cell line-compound combination. Each data point represents one of two biologically independent replicates for the indicated sgRNA.

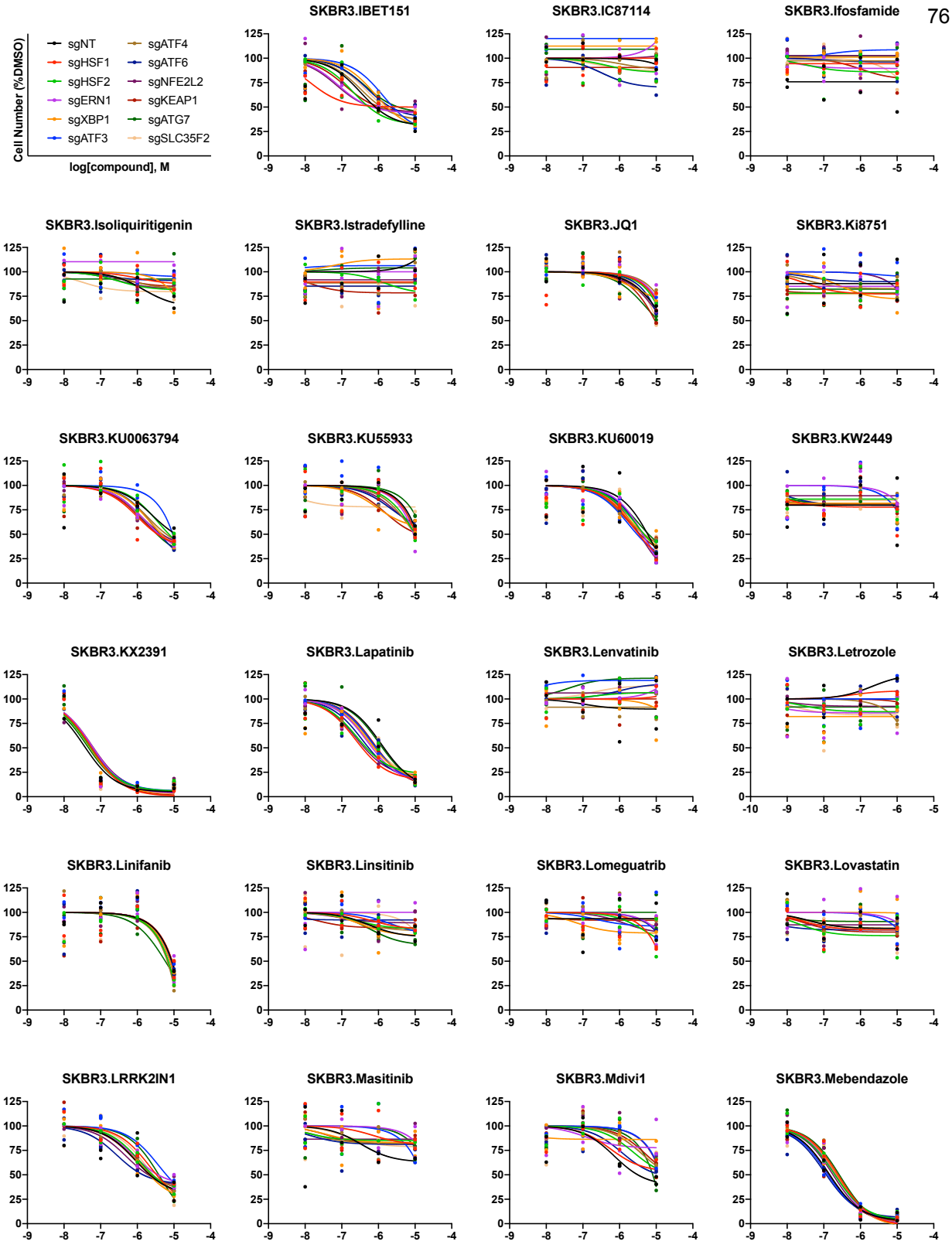


Figure 13 (continued). Dose-response curves generated using QMAP-Seq. Each plot represents a unique cell line-compound combination. Each data point represents one of two biologically independent replicates for the indicated sgRNA.

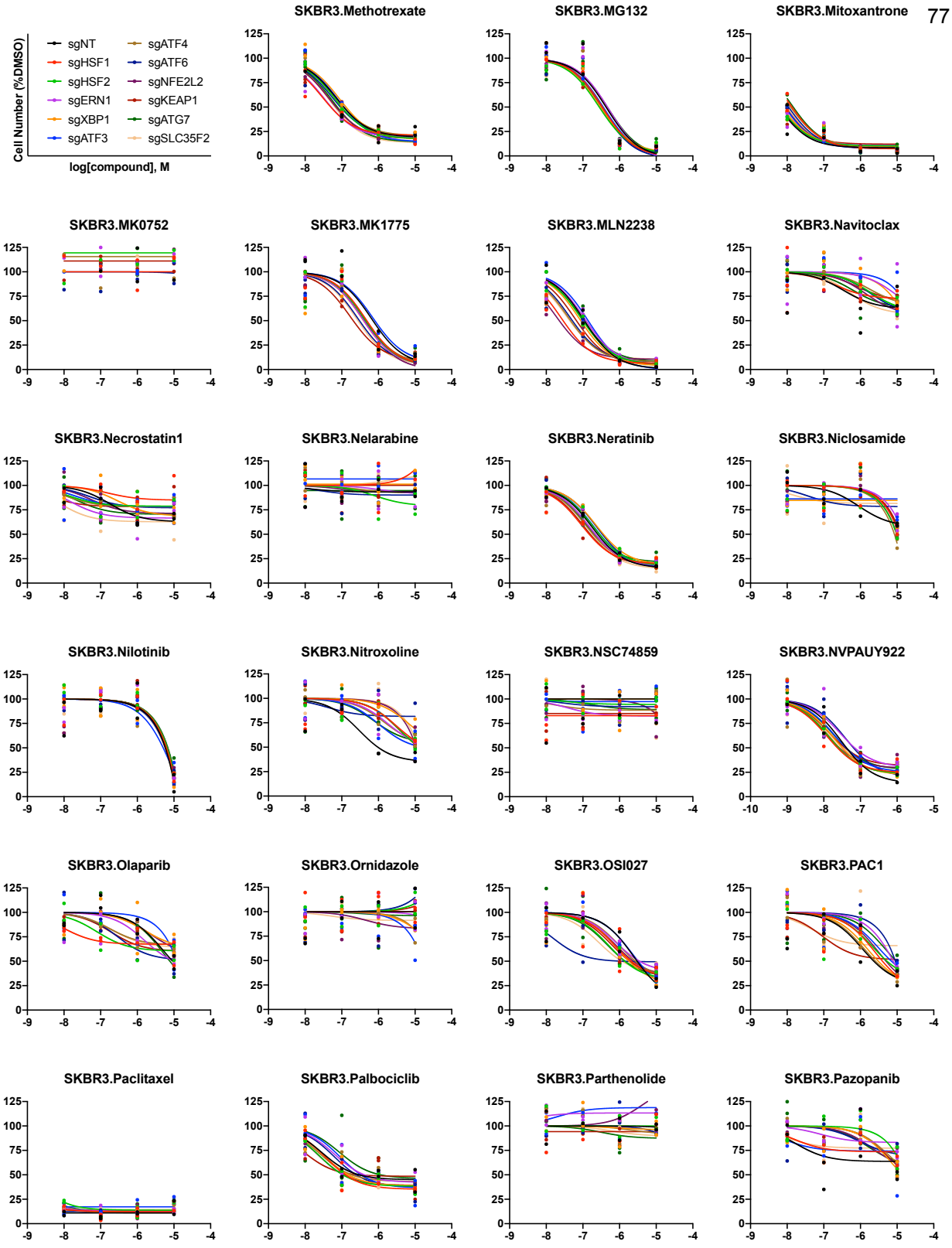


Figure 13 (continued). Dose-response curves generated using QMAP-Seq. Each plot represents a unique cell line-compound combination. Each data point represents one of two biologically independent replicates for the indicated sgRNA.

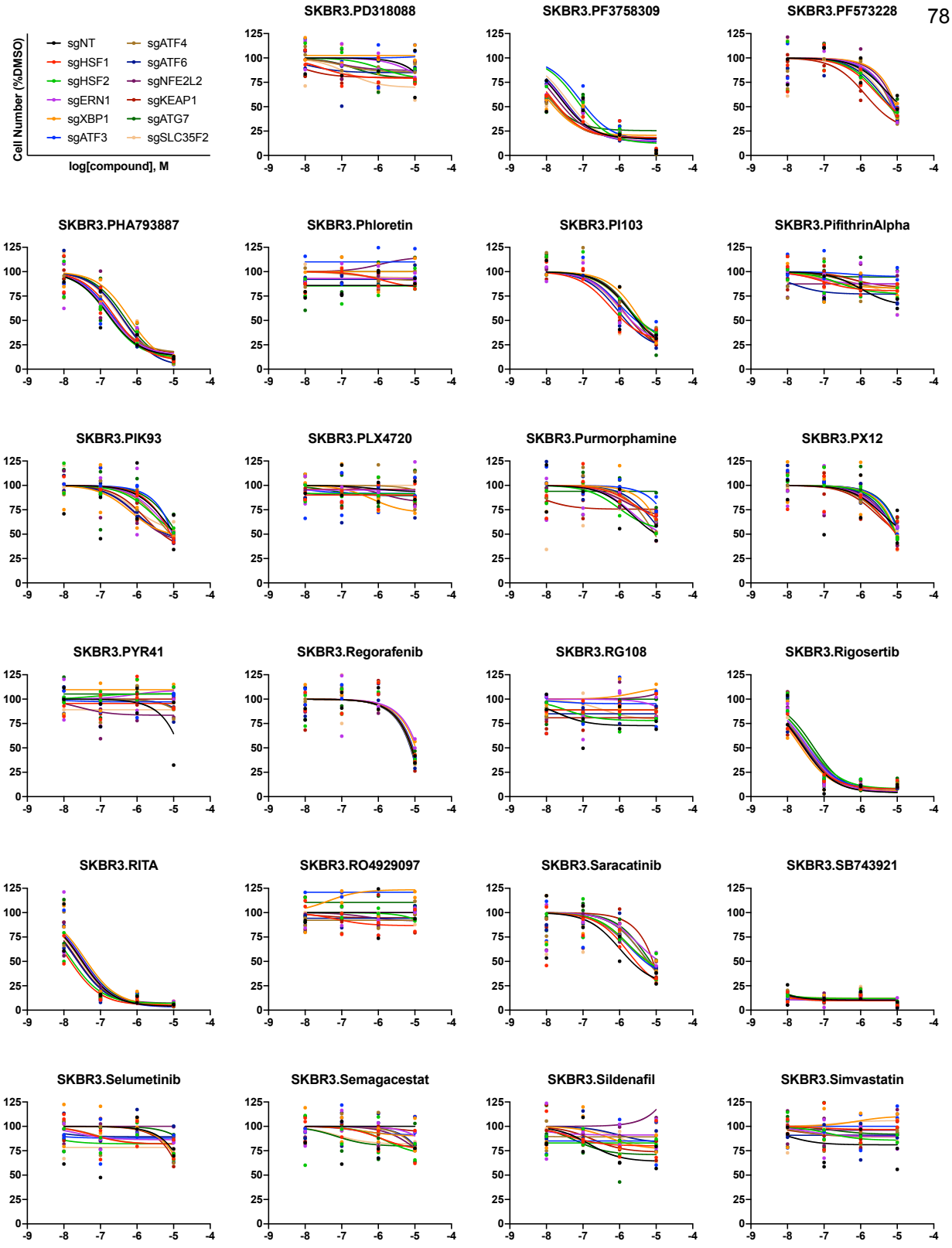


Figure 13 (continued). Dose-response curves generated using QMAP-Seq. Each plot represents a unique cell line-compound combination. Each data point represents one of two biologically independent replicates for the indicated sgRNA.

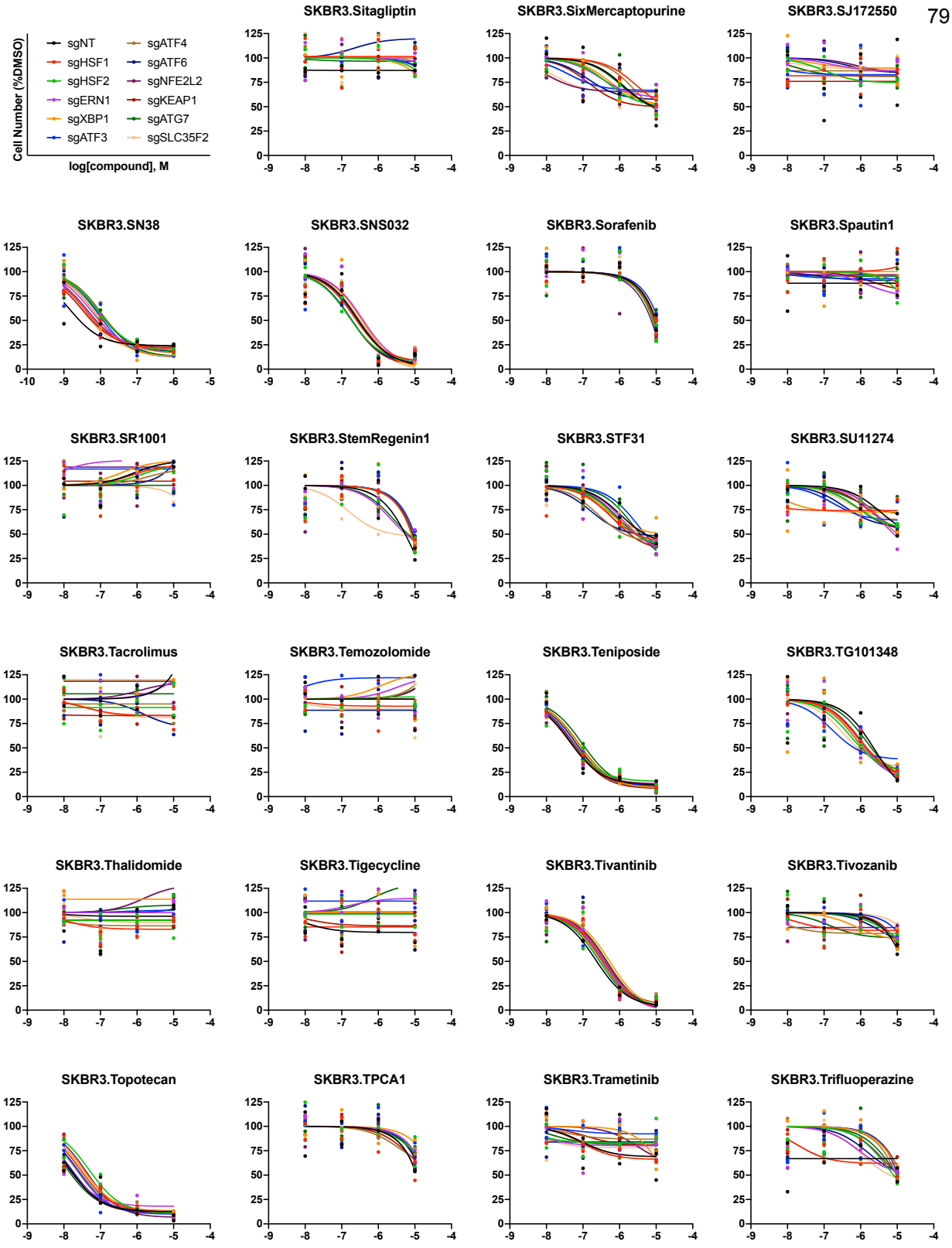


Figure 13 (continued). Dose-response curves generated using QMAP-Seq. Each plot represents a unique cell line-compound combination. Each data point represents one of two biologically independent replicates for the indicated sgRNA.

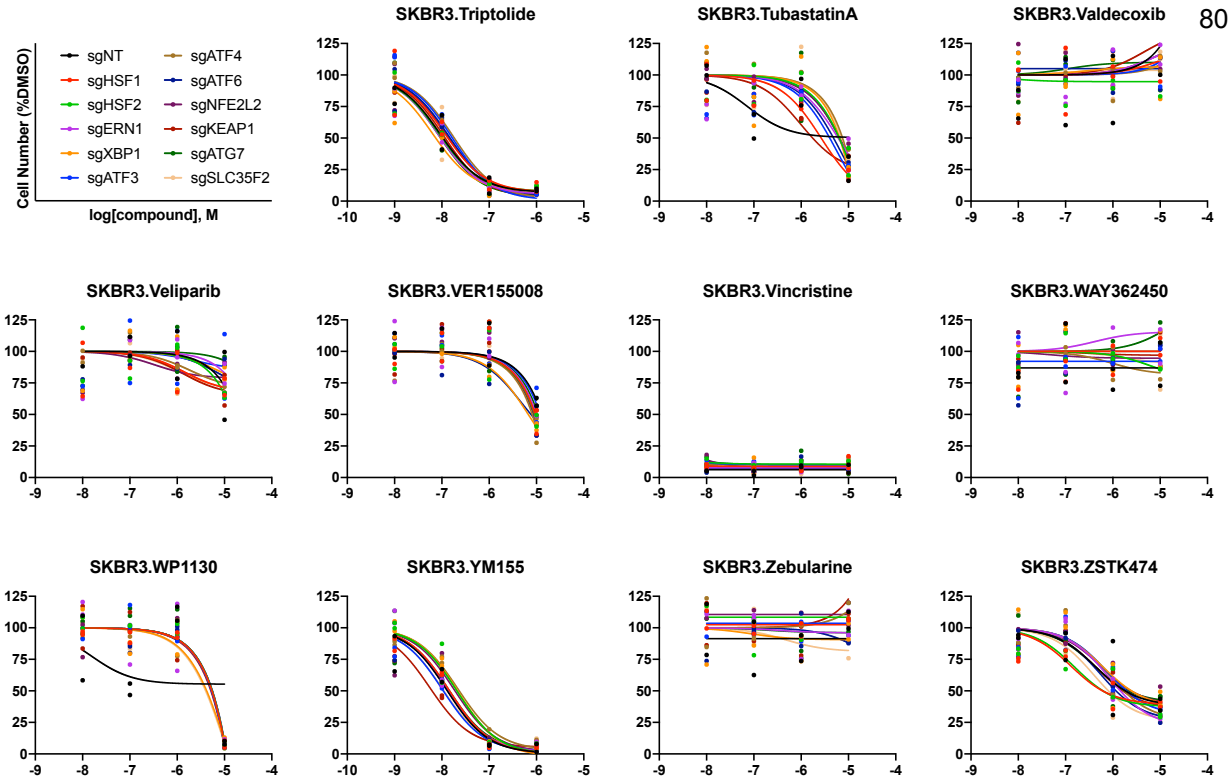


Figure 13 (continued). Dose-response curves generated using QMAP-Seq. Each plot represents a unique cell line-compound combination. Each data point represents one of two biologically independent replicates for the indicated sgRNA.

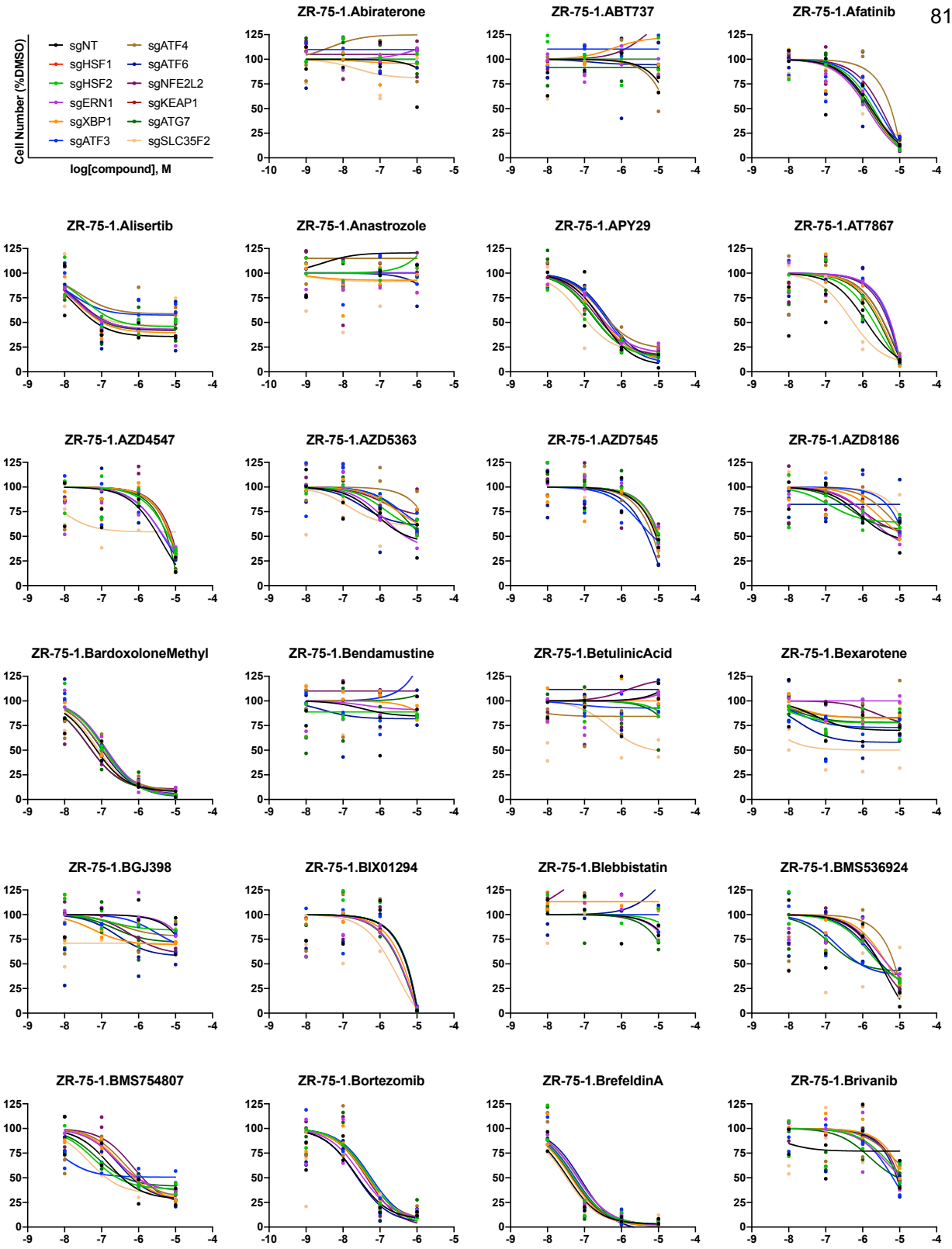


Figure 13 (continued). Dose-response curves generated using QMAP-Seq. Each plot represents a unique cell line-compound combination. Each data point represents one of two biologically independent replicates for the indicated sgRNA.

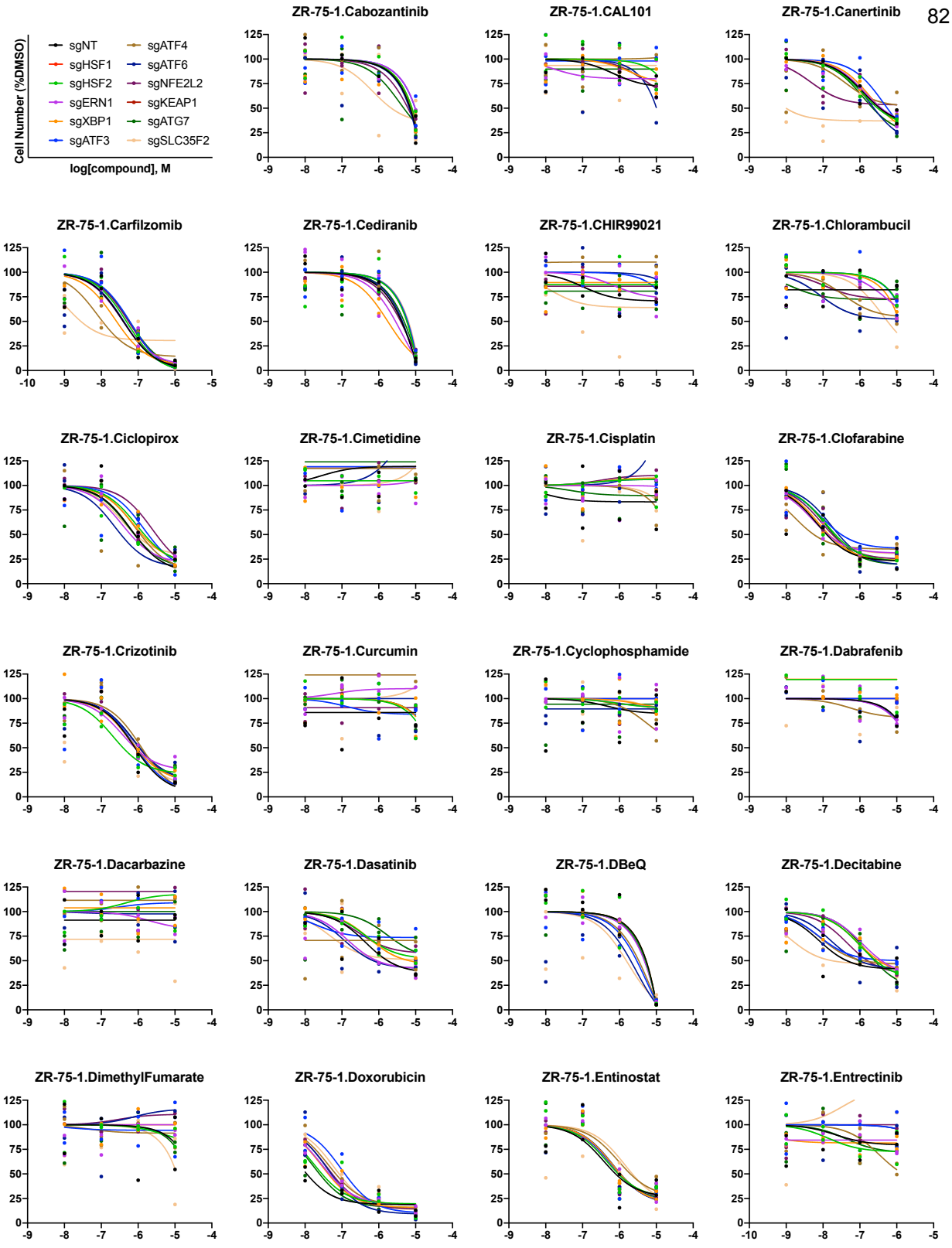


Figure 13 (continued). Dose-response curves generated using QMAP-Seq. Each plot represents a unique cell line-compound combination. Each data point represents one of two biologically independent replicates for the indicated sgRNA.

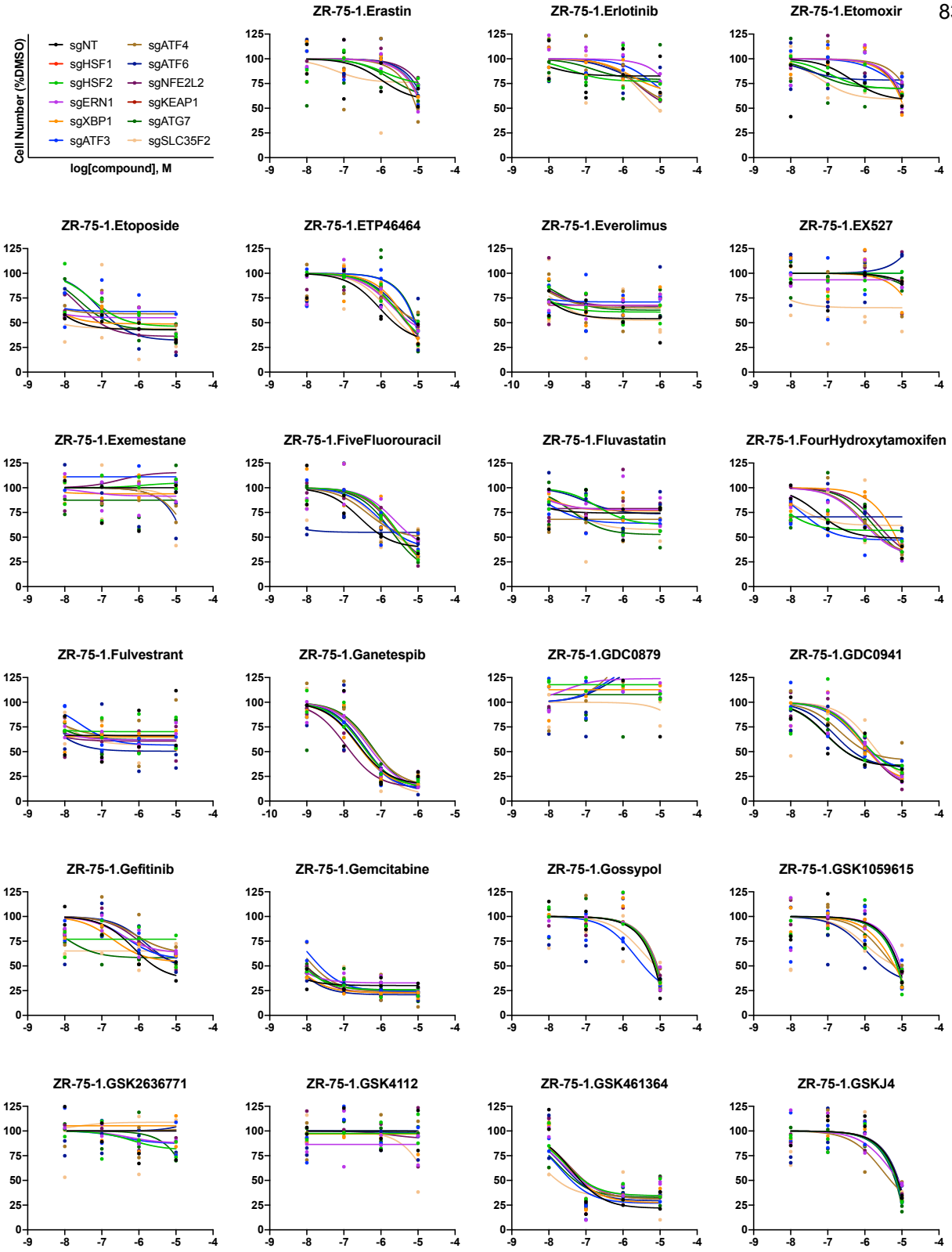


Figure 13 (continued). Dose-response curves generated using QMAP-Seq. Each plot represents a unique cell line-compound combination. Each data point represents one of two biologically independent replicates for the indicated sgRNA.

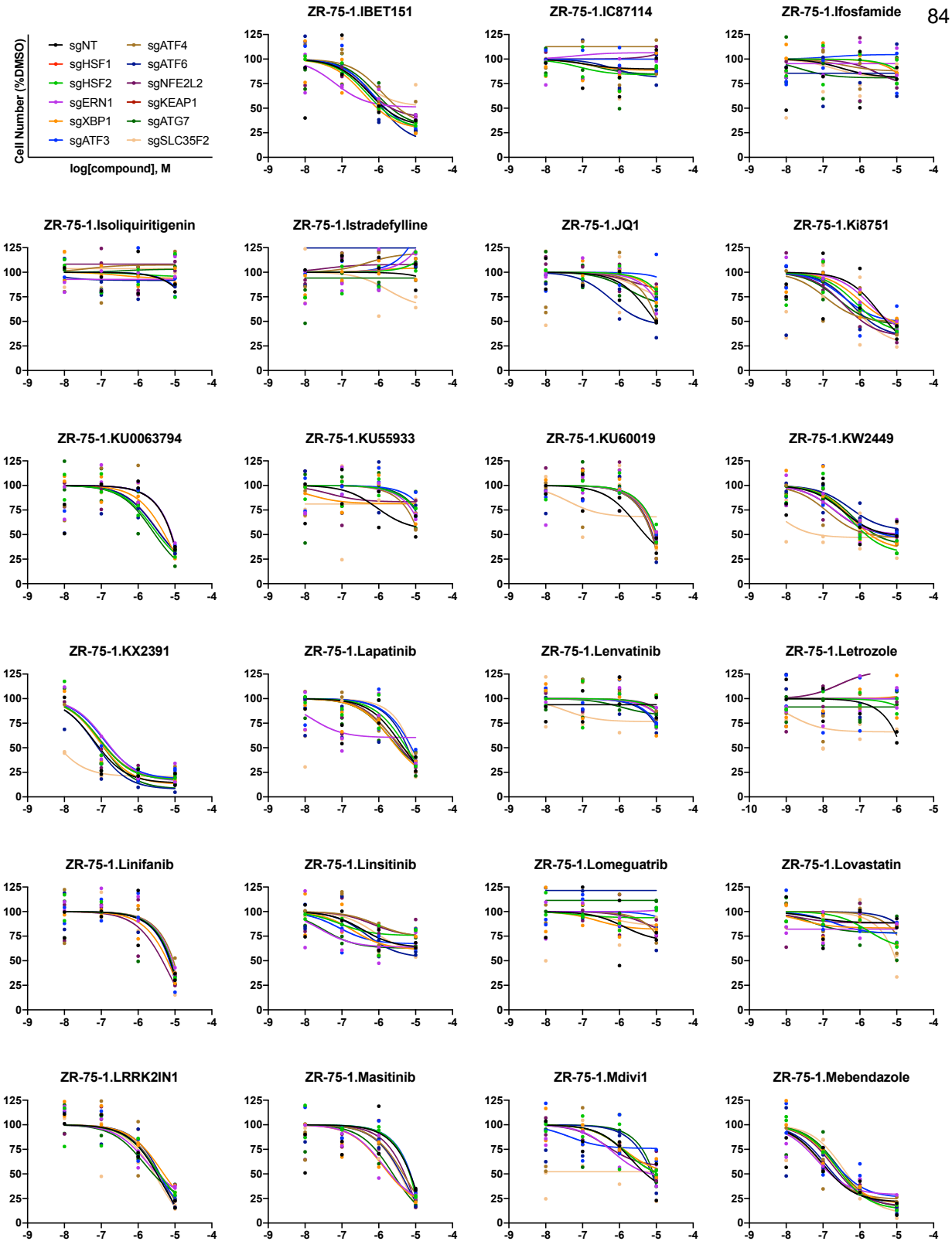


Figure 13 (continued). Dose-response curves generated using QMAP-Seq. Each plot represents a unique cell line-compound combination. Each data point represents one of two biologically independent replicates for the indicated sgRNA.

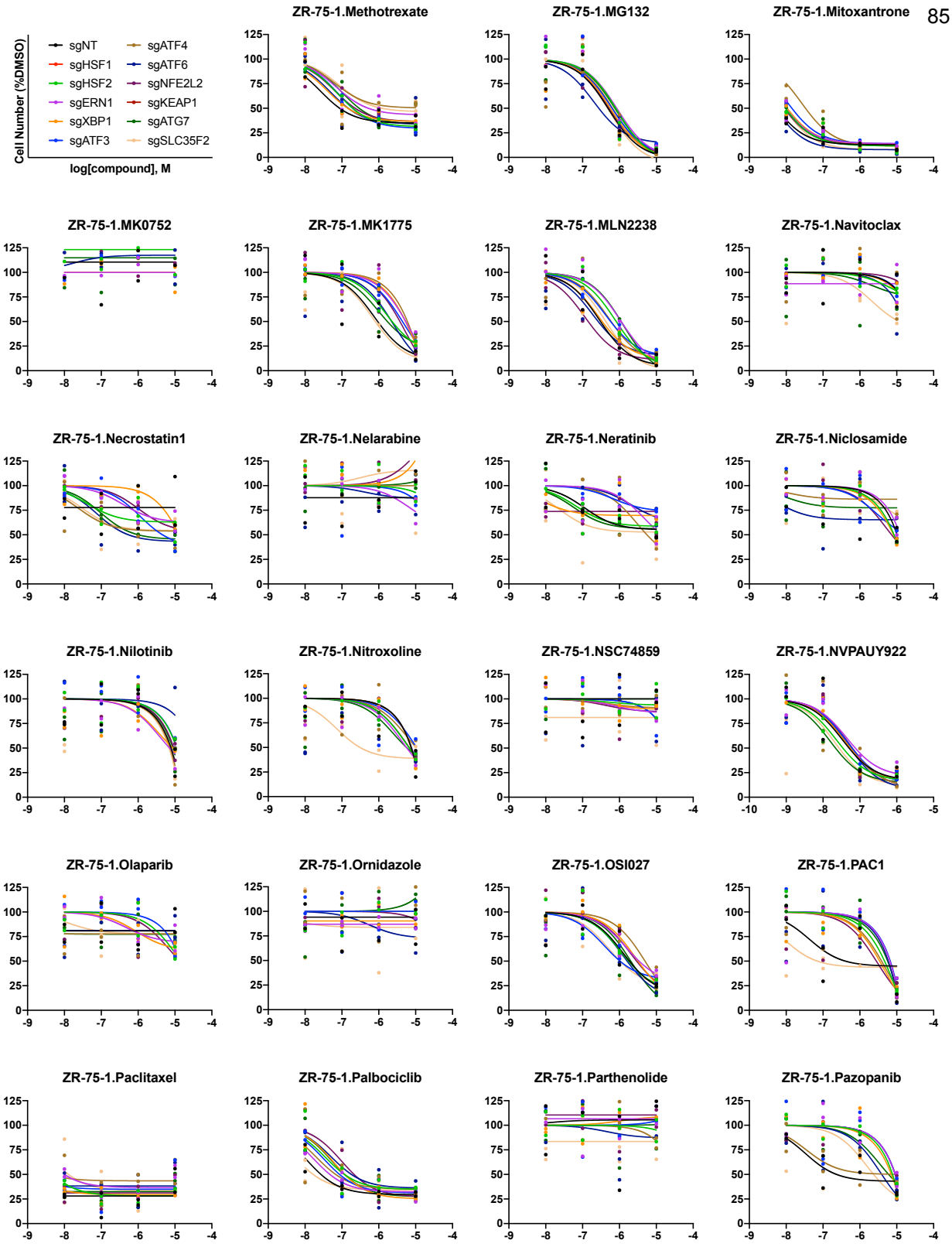


Figure 13 (continued). Dose-response curves generated using QMAP-Seq. Each plot represents a unique cell line-compound combination. Each data point represents one of two biologically independent replicates for the indicated sgRNA.

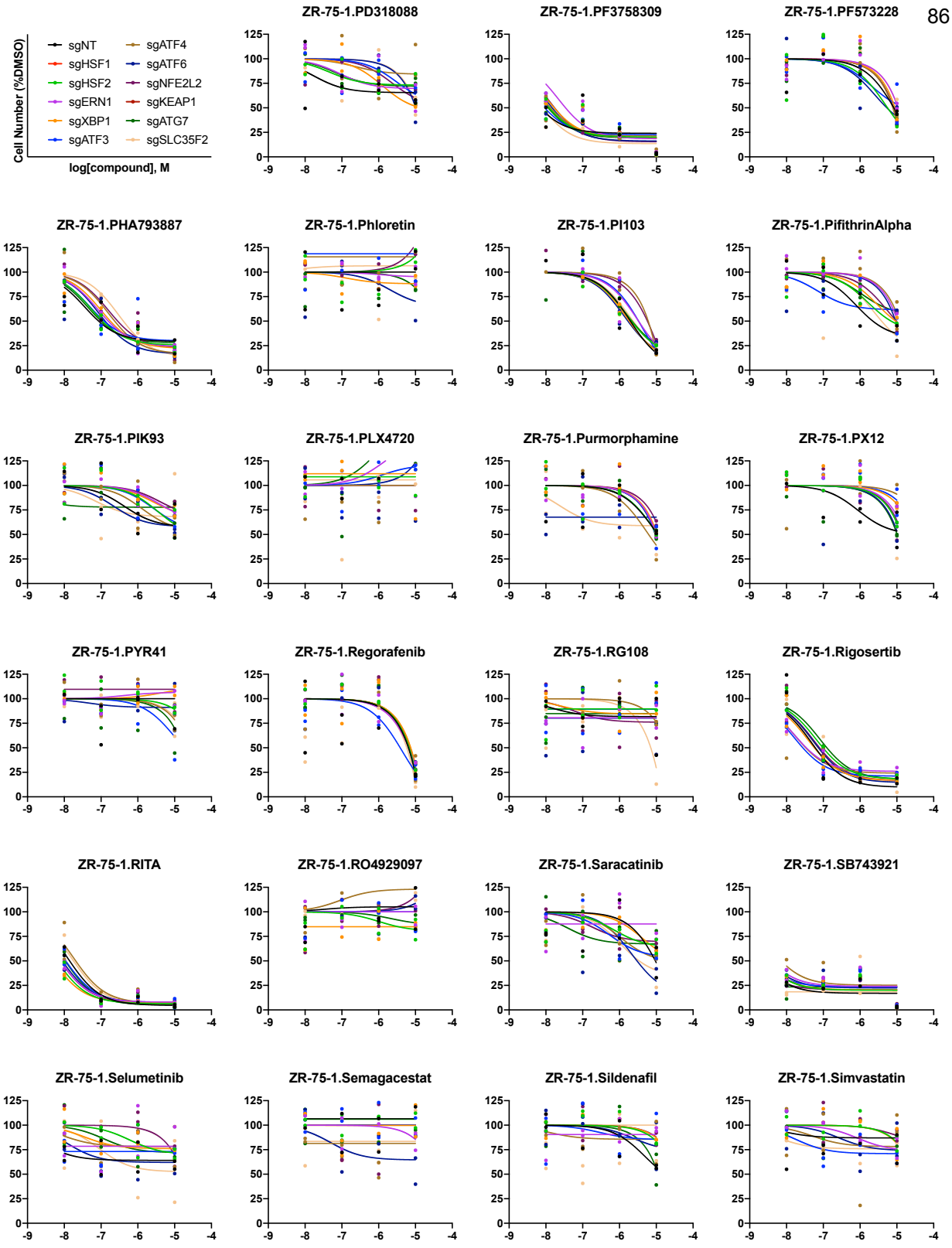


Figure 13 (continued). Dose-response curves generated using QMAP-Seq. Each plot represents a unique cell line-compound combination. Each data point represents one of two biologically independent replicates for the indicated sgRNA.

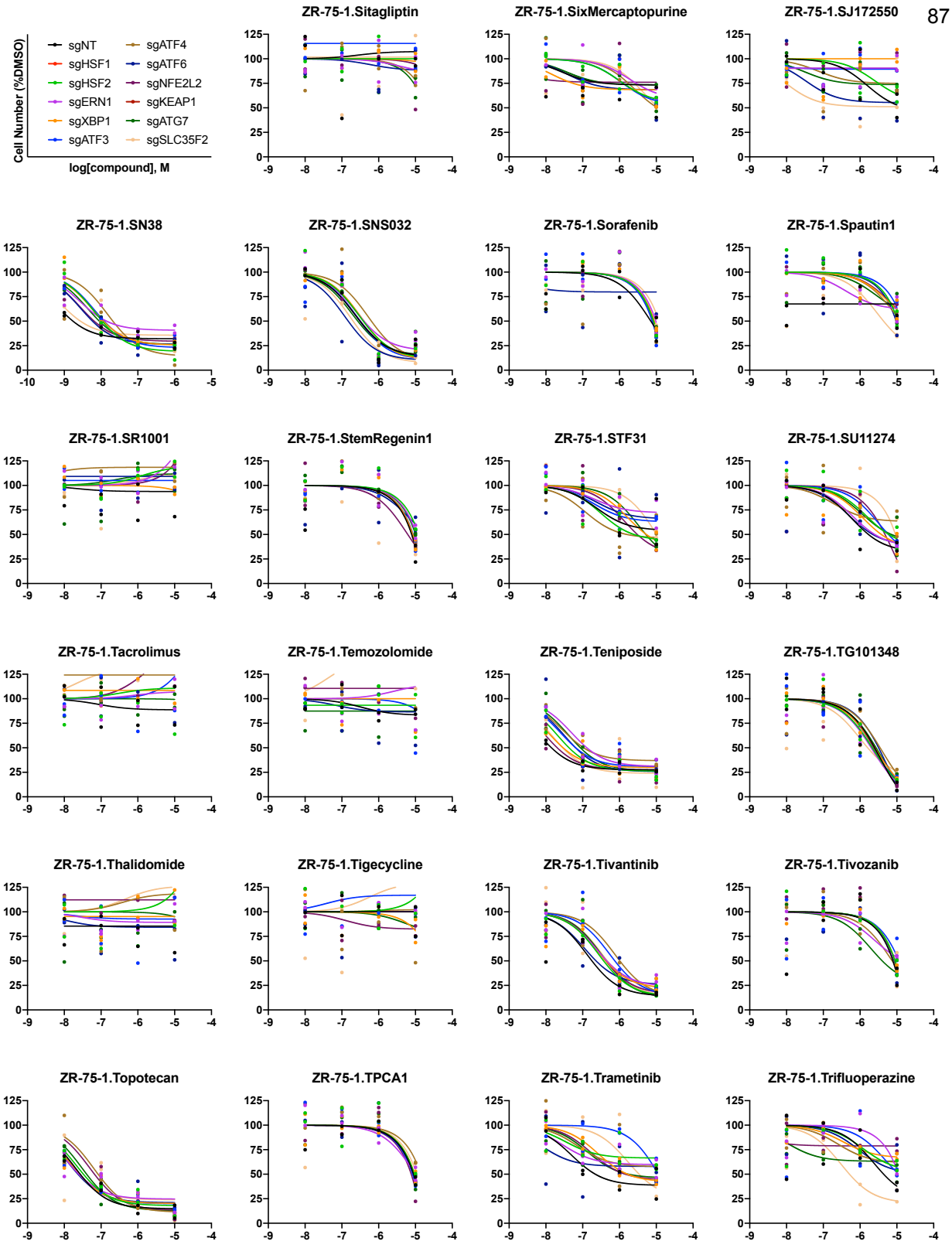


Figure 13 (continued). Dose-response curves generated using QMAP-Seq. Each plot represents a unique cell line-compound combination. Each data point represents one of two biologically independent replicates for the indicated sgRNA.

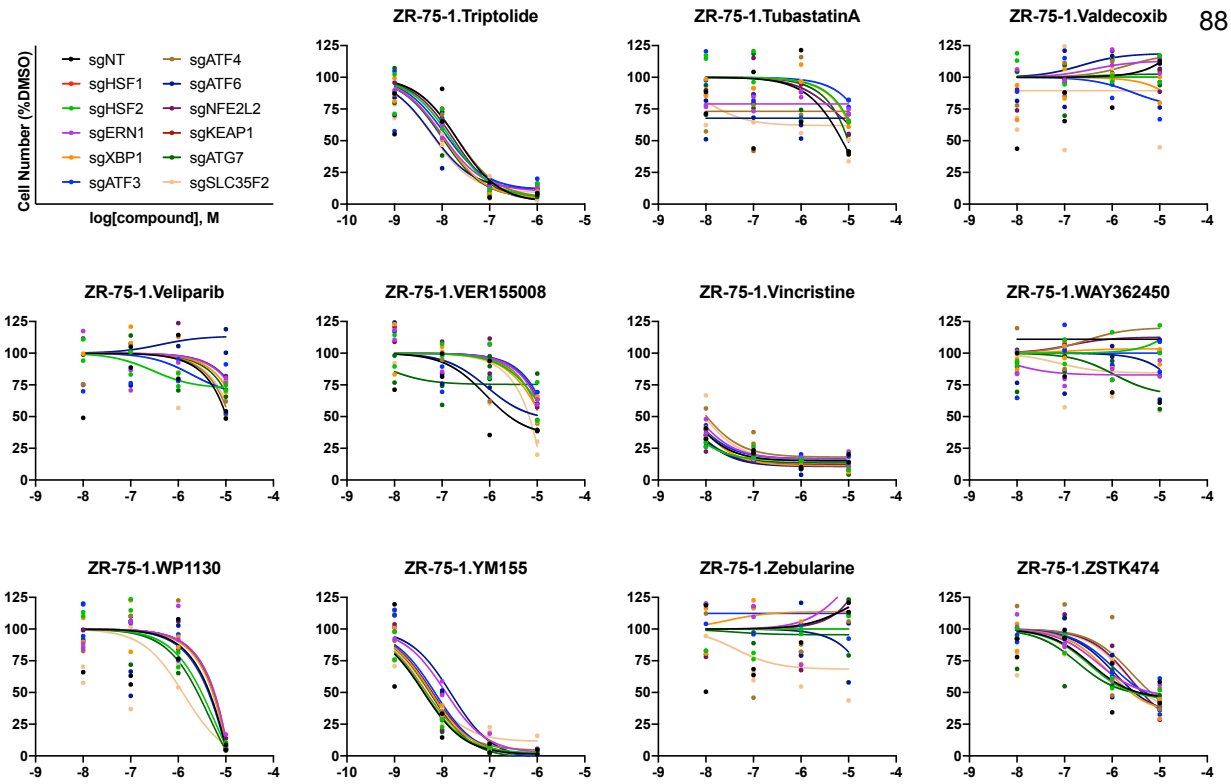
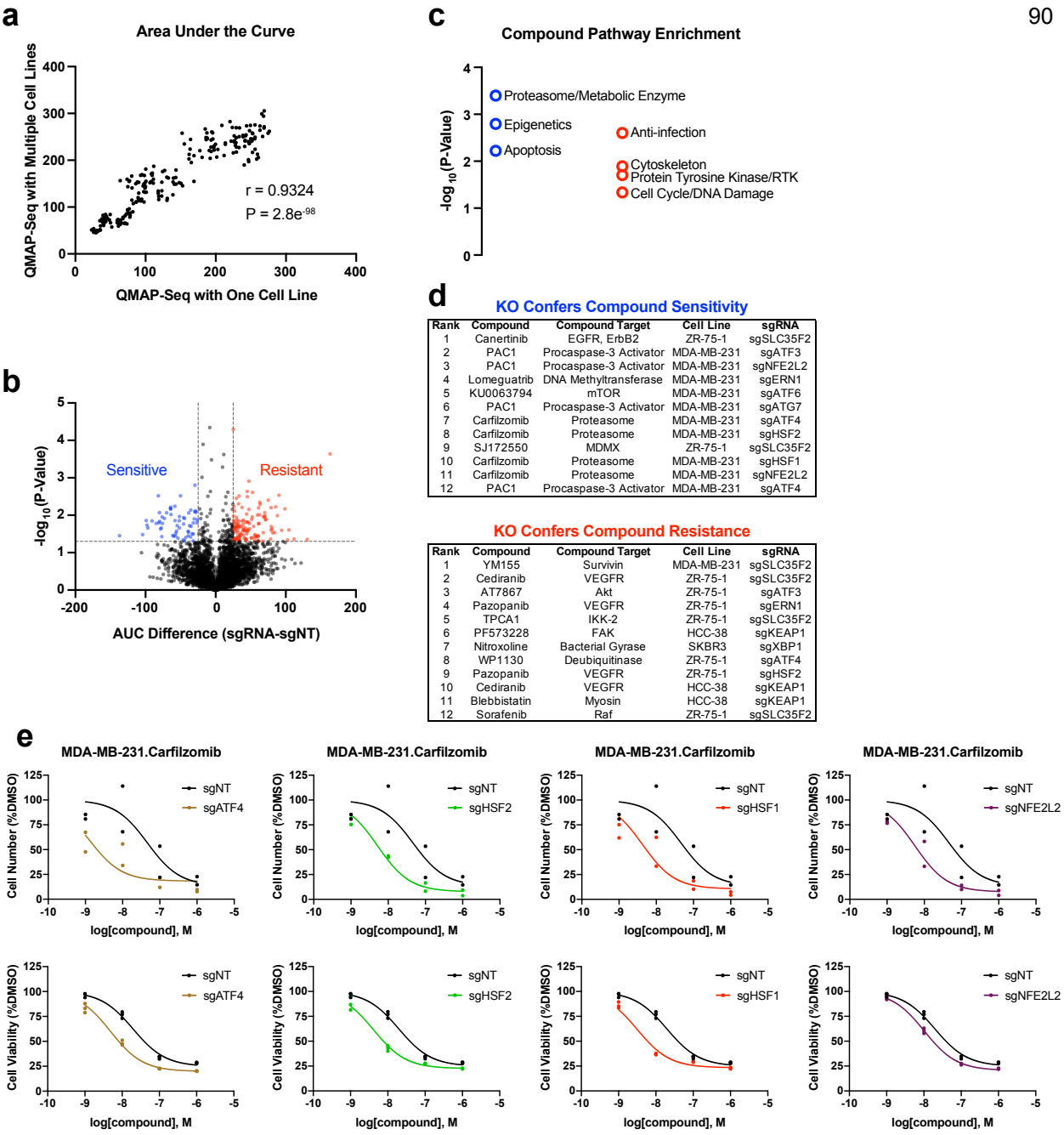


Figure 13 (continued). Dose-response curves generated using QMAP-Seq. Each plot represents a unique cell line-compound combination. Each data point represents one of two biologically independent replicates for the indicated sgRNA.

Despite profiling five times as many cell lines for this QMAP-Seq experiment compared to our pilot QMAP-Seq experiment with one cell line, the AUC measurements from the common cell line-sgRNA-compound combinations were highly correlated between these two independent experiments ($r = 0.9324$) **(Figure 14a)**. Thus, neither mixing different parental cell lines nor increasing the total number of cell types analyzed altered the performance of QMAP-Seq.

Statistical analysis revealed 60 cell line-gene-drug interactions that conferred compound sensitivity (AUC Difference < -25 and $P < 0.05$) and 124 cell line-gene-drug interactions that conferred compound resistance (AUC Difference > 25 and $P < 0.05$) compared to sgNT (see "Materials and Methods" for details) **(Figure 14b)**. To characterize the type of pathways enriched among the compounds involved in the top sensitivity and resistance interactions, we compared the expected distribution of the 19 pathways in our compound collection **(Figure 8b)** with the observed distribution of these pathways among the 60 sensitivity interactions and the 124 resistance interactions. Notably, the most significantly enriched pathway among the sensitivity interactions was proteasome and metabolic enzyme compounds, particularly proteasome inhibitors **(Figure 14c)**, suggesting that proteasome inhibition is especially lethal under conditions of proteostasis factor depletion. Epigenetics and apoptosis were also significantly enriched among the sensitivity interactions, whereas anti-infection, cytoskeleton, protein tyrosine kinase, cell cycle, and DNA damage were the most enriched pathways among the resistance interactions **(Figure 14c)**.

Among the top chemical-genetic interactions identified using QMAP-Seq **(Figure 14d)** was loss of four proteostasis factors (ATF4, HSF2, HSF1, NFE2L2) further sensitizing MDA-MB-231 cells to the proteasome inhibitor, Carfilzomib **(Figure 14e, Top)**, corroborating previous findings that loss of these factors enhances sensitivity to proteasome inhibitors in cancer cells (87-90). To validate these pooled screening hits in traditional arrayed format, we treated MDA-MB-231 cells possessing individual gene knockouts with Carfilzomib for 72 hours and measured their intracellular reducing potential as a proxy of cell viability. As we observed using QMAP-Seq, knockout of these proteostasis factors further sensitized MDA-MB-231 cells to Carfilzomib **(Figure 14e, Bottom)**. We conclude that proteostasis factor depletion is synthetic lethal with proteasome inhibition and more broadly, chemical-genetic interactions uncovered using heterogenous mixtures of cell lines are reproducible using homogenous cell lines.



Rank	Compound	Compound Target	Cell Line	sgRNA
1	Canertinib	EGFR, ErbB2	ZR-75-1	sgSLC35F2
2	PAC1	Procaspase-3 Activator	MDA-MB-231	sgATF3
3	PAC1	Procaspase-3 Activator	MDA-MB-231	sgNFE2L2
4	Lomeguatrib	DNA Methyltransferase	MDA-MB-231	sgERN1
5	KU0063794	mTOR	MDA-MB-231	sgATF6
6	PAC1	Procaspase-3 Activator	MDA-MB-231	sgATG7
7	Carfilzomib	Proteasome	MDA-MB-231	sgATF4
8	Carfilzomib	Proteasome	MDA-MB-231	sgHSF2
9	SJ172550	MDMX	ZR-75-1	sgSLC35F2
10	Carfilzomib	Proteasome	MDA-MB-231	sgHSF1
11	Carfilzomib	Proteasome	MDA-MB-231	sgNFE2L2
12	PAC1	Procaspase-3 Activator	MDA-MB-231	sgATF4

Rank	Compound	Compound Target	Cell Line	sgRNA
1	YM155	Survivin	MDA-MB-231	sgSLC35F2
2	Cediranib	VEGFR	ZR-75-1	sgSLC35F2
3	AT7867	Akt	ZR-75-1	sgATF3
4	Pazopanib	VEGFR	ZR-75-1	sgERN1
5	TPCA1	IKK-2	ZR-75-1	sgSLC35F2
6	PF573228	FAK	HCC-38	sgKEAP1
7	Nitroxoline	Bacterial Gyrase	SKBR3	sgXBP1
8	WP1130	Deubiquitinase	ZR-75-1	sgATF4
9	Pazopanib	VEGFR	ZR-75-1	sgHSF2
10	Cediranib	VEGFR	HCC-38	sgKEAP1
11	Blebbistatin	Myosin	HCC-38	sgKEAP1
12	Sorafenib	Raf	ZR-75-1	sgSLC35F2

Figure 14. Identification and validation of cell line-gene-drug interactions. a Scatterplot of the dose-response curve AUC as determined using QMAP-Seq with one cell line versus QMAP-Seq with multiple cell lines. Common compounds for MDA-MB-231 cells are shown. Statistical significance of Pearson correlation was determined using a two-tailed test (n = 220 compound-sgRNA combinations). **b** Volcano plot depicting cell line-gene-drug interactions. Magnitude was determined by calculating the difference in mean AUC between sgRNA and sgNT for every cell line-compound combination. Statistical significance was determined using an unpaired, two-tailed t test (n = 2 biologically independent replicates). Blue dots indicate interactions where the knockout confers greater sensitivity than sgNT. Red dots indicate interactions where the knockout confers greater resistance than sgNT. **c** Pathways targeted by compounds involved in the top 60 sensitivity interactions (blue circles) or top 124 resistance interactions (red circles) that were significantly enriched compared to expected pathway representation. Statistical significance of pathway enrichment was determined using a one-tailed binomial test to compare observed distribution with expected distribution (n = 180 compounds). **d** List of the top 12 conditions that confer compound sensitivity or resistance. **e** Top: Dose-response curves for four of the top chemical-genetic interactions as measured using QMAP-Seq. Each data point represents one of two biologically independent replicates. Bottom: Dose-response curves as measured using Resazurin Cell Viability Kit. Each data point represents one of three biologically independent replicates.

2.5 QMAP-Seq enables proteostasis network mapping in breast cancer.

To further investigate the strongest chemical-genetic interactions in our dataset, we assembled a chemical-genetic interaction map of the highest confidence interactions (Absolute AUC Difference > 60 and P < 0.05) **(Figure 15a)**. This network incorporated both synthetic lethal gene-drug interactions and synthetic rescue gene-drug interactions. The map revealed hub compounds that synergize with loss of multiple proteostasis factors. For example, the proteasome inhibitor, Carfilzomib displayed five synthetic lethal chemical-genetic interactions with proteostasis factors from distinct branches of the proteostasis network. We also observed hub characteristics for compounds not previously connected to proteostasis, including the procaspase-3 activator, PAC1, and the DNA methyltransferase inhibitor, Lomeguatrib.

To model the overall structure of the proteostasis network, we assessed the functional similarity of each proteostasis factor's chemical-genetic interaction profiles across all cell lines and compounds profiled. Specifically, we quantified the Spearman correlation between all gene-gene pairs based on AUC difference **(Figure 15b)**. Using this approach, genes with similar compound sensitivity and resistance profiles clustered together, revealing several known genetic relationships **(Figure 15c)**. For example, because KEAP1 negatively regulates NRF2 (*NFE2L2*), the master transcriptional regulator of the oxidative stress response (23), we expected a low degree of correlation between these two factors. Indeed, NFE2L2 and KEAP1 displayed the third lowest correlation. In addition, we observed a high correlation between ERN1 and XBP1, which was expected given that the ER stress sensor, IRE1 (*ERN1*), activates XBP1 through mRNA splicing (17).

This approach also revealed previously unknown genetic relationships **(Figure 15c)**. For example, the highest correlation was observed between NFE2L2 and XBP1. While the relationship between these two factors in the context of breast cancer is unknown, XBP1 has been shown to activate NRF2 in atherosclerosis (91) and retinal pigment epithelium cells (92). Furthermore, the paralogs HSF1 and HSF2 have been reported to function cooperatively (10-12, 93), antagonistically (13), and in a context-dependent manner (14). Interestingly, HSF1 and HSF2 displayed the second highest correlation, providing evidence in support of a cooperative interaction in breast cancer. The lowest correlation genetic interaction was

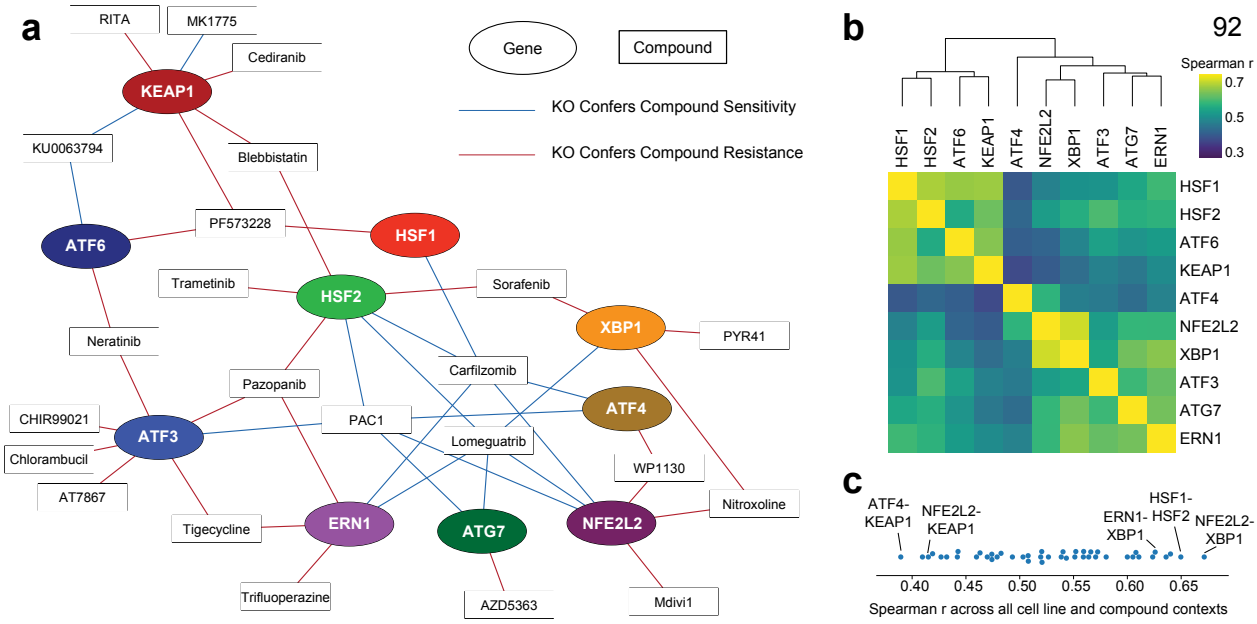


Figure 15. QMAP-Seq enables proteostasis network mapping in breast cancer. a Chemical-genetic interaction map displaying the highest confidence interactions identified by QMAP-Seq (Absolute AUC Difference > 60 and P < 0.05). Genes are represented as oval nodes, and compounds are represented as rectangular nodes. Chemical-genetic interactions where the knockout confers compound sensitivity (i.e., synthetic lethal interactions) are represented as blue edges. Chemical-genetic interactions where the knockout confers compound resistance (i.e., synthetic rescue interactions) are represented as red edges. **b** Spearman correlation matrix of gene-gene correlations based on AUC differences across all cell line and compound contexts (n = 488 cell line-compound combinations). **c** Distribution of gene-gene Spearman correlations across all cell line and compound contexts. Known and previously unknown genetic relationships are labeled.

observed between ATF4 and KEAP1, which is not entirely surprising considering KEAP1 has been reported to negatively regulate ATF4 expression in other cancer models (53). Altogether, we conclude that chemical-genetic profiling using QMAP-Seq provides insight into the organization of the proteostasis network in breast cancer cells and has the power to reveal genetic relationships.

CHAPTER 3: DISCUSSION

Here, we develop QMAP-Seq to identify drug vulnerabilities dependent on the activation status of individual proteostasis factors and to map the organization of the proteostasis network in breast cancer. QMAP-Seq is a highly multiplexed chemical-genetic profiling strategy that enables systematic phenotyping of dozens of cell lines with defined genetic perturbations across thousands of individual compound treatments. We label individual cell lines, sgRNAs, and compound treatments with unique DNA barcodes and utilize next-generation sequencing to monitor cell viability in pooled format. We also present a bioinformatic analysis pipeline that simplifies the processing of thousands of fastq files into over 10,000 dose-response curves.

Using complex mixture of cells, we uncover a total of 60 sensitivity interactions and 124 resistance interactions, including both known and novel interactions involving critical proteostasis factors. We uncover a small number of compounds (e.g., Carfilzomib, PAC1, Lomeguatrib) that act as synthetic lethal hubs connecting multiple nodes of the proteostasis network. Finally, we integrate our large-scale chemical-genetic interaction dataset generated across a diverse landscape of compounds to infer functional relationships between individual members of the proteostasis network in the context of breast cancer. We identify both known and previously unknown functional relationships, including cooperative relationships (e.g., NRF2-XBP1, HSF1-HSF2, IRE1-XBP1) and antagonistic relationships (e.g., ATF4-KEAP1, NRF2-KEAP1). KEAP1 is the only negative regulator of the proteostasis network that we included in our gene panel, so it is not surprising that it acts antagonistically to the positive regulators of the proteostasis network. It is worth noting that the antagonistic relationships displayed low positive gene-gene Spearman correlations instead of negative correlations, which might be expected for antagonistic relationships. This finding potentially reflects the multitude of compounds that exhibited similar compound sensitivity across the various genetic perturbations. The compounds with similar compound sensitivity likely masked the effect size of the compounds with differential compound sensitivity.

3.1 Discussion of unexpected results

We incorporated SLC35F2 as a positive control for QMAP-Seq based on a previous study demonstrating that knockout of SLC35F2 conferred resistance to YM155 in the near-haploid chronic myeloid leukemia cell line, KBM7 (80). In our hands, we observed that knockout of SLC35F2 in the triple-negative breast cancer cell lines conferred resistance to YM155 **(Figure 12)**. However, knockout of SLC35F2 in the non-triple-negative cell lines conferred resistance to YM155 to a much lesser extent **(Figure 12)**. A possible explanation for the smaller effect in the non-triple-negative cell lines is that they are already resistant to YM155 due to lower SLC35F2 mRNA expression in the non-triple-negative cell lines compared to the triple-negative cell lines (70).

Although YM155 has shown promise as an anticancer agent in preclinical studies, its performance in multiple Phase I and Phase II clinical trials has been largely disappointing due to its limited antitumor efficacy when used alone (94-99) or in combination with other cytotoxic therapeutic agents (100-103). It has been proposed that the failure of YM155 in clinical trials could be due to the chemical instability of the compound (104). For example, one study showed a rapid decrease in YM155 levels in serum and tumors after completion of YM155 treatment (105). Another potential explanation for the disappointing clinical trial efficacy is that the patients included in the trials were not the appropriate patient populations. As others have suggested previously (80), we see value in stratifying patients based on SLC35F2 expression for future clinical trials. We predict that patients with high expression of SLC35F2 are likely to respond better than those with low expression. Patient stratification may help distinguish between potential “responders” and “non-responders” to YM155, a compound that has been well-tolerated in patients (95).

3.2 Advantages of QMAP-Seq

QMAP-Seq has several major advantages over existing methods for the identification of chemical-genetic interactions in mammalian cells. For one, our pooled approach provides substantial gains in throughput, facilitating over 50 times as many cell viability measurements per sample as established arrayed assays and enabling all cells in a well to be exposed to identical culture conditions and compound concentrations. By introducing sample-specific indexes, thousands of compound treatment samples are

further pooled and sequenced together in a single next-generation sequencing run. This distinguishes QMAP-Seq from PRISM (77), an approach that uses microsphere bead technology for multiplexing cell lines, but not compound treatments. Another important feature of QMAP-Seq is the use of genetically engineered isogenic cell line pairs, which enables direct assessment of the contribution of a single genetic perturbation to compound sensitivity. It is worth noting, however, that other pooled assays could in principle profile isogenic cell line pairs, while QMAP-Seq could be performed with pools of non-isogenic cell lines. QMAP-Seq leverages readily available next-generation sequencing, which itself continues to improve in throughput, and is compatible with common sequencing libraries, such as RNA-Seq. As a result, the costs associated with QMAP-Seq are $1/10^{th}$ the cost of an ATP-based cell viability assay of the same magnitude performed in 384-well format. The compound requirements are also lower, which is particularly useful when profiling scarce compounds or natural products.

Another advantage of QMAP-Seq is that it is quantitative and accurate. The use of cell spike-in standards allows quantitative determination of cell number from sequencing reads. The spike-ins also provide a standard curve for each sample, which normalizes for differences in PCR amplification efficiency across samples and reduces technical variability. Moreover, QMAP-Seq features direct cell counting, which provides a more accurate measure of cell viability than metabolic-based readouts that use metabolic activity as a proxy of cell viability. Metabolic-based cell viability assays do not accurately reflect cell number when studying compounds that alter cell size or ATP metabolism because the expected relationship between ATP level and viable cell number is distorted (106-108).

Importantly, the short-term nature of QMAP-Seq (i.e., 72 hour compound treatment) enables profiling of thousands of compounds, which would be technically challenging with existing genome-scale chemical-genetic methods that require several weeks of passaging and treating cells prior to readout. The short-term format also minimizes potential secondary effects that may arise due to several weeks of exposure to compounds, and thus identifies the most relevant chemical-genetic interactions. Collectively, these attributes of QMAP-Seq enable accurate mapping of biological network structure, as we demonstrate for the proteostasis network.

3.3 Limitations of QMAP-Seq

Nevertheless, QMAP-Seq has certain limitations. For one, pooling multiple cell lines requires optimization of cell pooling ratios to ensure adequate representation. However, this could be bypassed by pooling cell lines with similar doubling times (78) or by profiling one cell line at a time. Additionally, factors secreted by one cell line could conceivably alter the compound sensitivity of another cell line in the pool. Although such paracrine-mediated effects would certainly warrant further studies, we **(Figure 14a, Figure 14e)** and others (77, 78) have yet to observe any major difference in drug response measurements from pooled versus arrayed formats. Moreover, pooling heterogeneous mixtures of cells more closely mimics the complexity and heterogeneity of the tumor microenvironment. Finally, while QMAP-Seq is compatible with standard cell viability readouts, it is less suitable for assessing non-cell viability phenotypes. It could, however, be adapted to non-standard readouts by isolating cells that possess the desired phenotype prior to sequencing.

3.4 Potential applications

We envision several potential applications of QMAP-Seq. For one, networks of genes identified using genome-scale chemical-genetic studies could be investigated further using QMAP-Seq with a more focused set of genes, but a broader range of compounds. Notably, the QMAP-Seq experiments described in this dissertation were performed without significant automation. By employing automation, this assay could readily be scaled to encompass an even larger number of compounds. In addition, QMAP-Seq could be used to directly characterize the impact of individual cancer-specific alterations identified through large-scale cancer sequencing and molecular characterization efforts by The Cancer Genome Atlas and the Cancer Genome Project (109, 110) on therapeutic response, which may improve patient stratification and treatment outcomes. For this application, it would be beneficial to prioritize clinically-relevant compounds over a diverse compound collection. QMAP-Seq could also be modified to accommodate gene regulation using CRISPR interference (CRISPRi) or CRISPR activation (CRISPRa). To better recapitulate the effect of physiologically-relevant gene expression levels on drug response, gene dosage could be precisely tuned using established methods (111).

QMAP-Seq could be adapted to enable more sophisticated assays. For example, to better model the contribution of the tumor microenvironment to drug response, it could be used for chemical-genetic profiling of barcoded cancer cells grown in the presence of stromal cells. These tumor-stroma co-culture experiments would not be possible with traditional cell viability assays that are unable to distinguish between cancer cells and stromal cells. Moreover, other types of perturbations besides chemical treatments could be applied, such as nutrient conditions, to assess the consequence of diverse perturbations on cell viability across multiple genetic contexts.

QMAP-Seq could be modified to accommodate pooled *in vivo* chemical-genetic profiling experiments by injecting immunocompromised mice with a pool of barcoded cells and sequencing tumors for barcode enrichment or depletion in compound-treated mice relative to vehicle-control treated mice. Recent reports have demonstrated the feasibility of applying similar *in vivo* barcode sequencing approaches. In one example, multiple barcoded cells pre-treated with compounds were simultaneously tested for their capacity to reduce *in vivo* metastatic seeding ability (112). In another example, multiple sgRNAs targeting known and candidate tumor suppressor genes were concurrently tested for their effect on *in vivo* lung tumor growth (113). As more QMAP-Seq data is collected over time from these and other potential applications, we see value in building searchable databases that could serve as a resource for the broader scientific community.

3.5 Future directions

The results from this project present several new possibilities for future investigation. To enhance our understanding of the biology behind chemical-genetic interactions uncovered using QMAP-Seq, transcriptional profiling of cells in response to specific genetic and chemical perturbations could be performed at the whole population level (114-116) or single-cell level (117, 118). This data could be layered with QMAP-Seq data to enable a richer annotation of connections between genes and compounds based on gene expression signatures. Further studies will be required to determine the precise molecular mechanism by which loss of an individual proteostasis factor sensitizes breast cancer cells to the proteasome inhibitor, Carfilzomib. We predict that proteasome inhibition increases proteotoxic stress, and

the absence of critical proteostasis factors reduces the capacity to cope with proteotoxic stress. Carfilzomib is already FDA approved for use as a single agent or as part of a combination therapy for the treatment of adult patients with relapsed or refractory multiple myeloma, so it could potentially be repurposed for use in breast cancer patients with perturbed proteostasis networks.

The barcoded cells generated in this study could also be used to investigate the role of individual proteostasis factors in promoting breast cancer invasion and metastasis. To determine which proteostasis factors are required for the invasive potential of certain breast cancer cells, pooled transwell invasion assays could be performed by seeding pooled cells in the upper chamber of a Boyden chamber in low serum growth medium. Cells that invade through the collagen extracellular matrix (ECM) layer to the lower chamber containing high serum growth medium could be collected and sequenced to quantify barcode abundance.

To establish which proteostasis factors are required for the metastatic potential of specific breast cancer cells, various *in vivo* metastasis assays could be performed. Metastasis is a multi-step process that involves cancer cell invasion through the ECM, migration through the blood or lymph vessels, and colonization at distant sites. To recapitulate all of the metastatic steps, pooled cells could be orthotopically injected into the mammary fat pad of female immunocompromised mice. Knockout could be induced by feeding mice doxycycline + glucose water. After metastatic lesions have formed, the metastases could be harvested and sequenced. Alternatively, to recapitulate just the later steps of metastasis, a similar experiment could be performed using tail vein injection instead of orthotopic injection. It would be interesting to determine whether individual genetic perturbations have different abilities to colonize specific organs.

3.6 Concluding remarks

Overall, we report the development of the QMAP-Seq experimental and bioinformatic pipeline and the application of this chemical-genetic profiling approach to the proteostasis network. QMAP-Seq addresses the need for a direct, non-correlative method for assessing compound selectivity across multiple cellular and genetic contexts in a high-throughput and scalable manner. It provides a path towards cancer precision medicine by predicting clinically actionable synthetic lethal and synthetic rescue interactions and

potential combinatorial therapies. QMAP-Seq represents the first application of chemical-genetic profiling to map biological networks and reveal functional genetic relationships in mammalian cells. Above all, the QMAP-Seq platform is well-suited to answer a broad range of clinical and biological questions and can be readily adopted by standard laboratories without the need for highly specialized equipment.

CHAPTER 4: MATERIALS AND METHODS

4.1 Plasmid construction

For cloning doxycycline-inducible Cas9, 3xFLAG-Cas9-EGFP was amplified from pSpCas9(BB)-2A-GFP (Addgene, Plasmid #48138) (119) with Cas9 EcoRI F and EGFP BamHI R **(Table 7)** and inserted within the EcoRI and BamHI restriction sites of the pLVX-TetOne vector (Clontech, #631846) using the In-Fusion HD Cloning Kit (Clontech, #638909), according to manufacturer's instructions.

For inserting cell line barcodes downstream of the gRNA scaffold and upstream of cPPT/CTS within the lentiGuide-Puro vector (Addgene, Plasmid #52963) (120), EcoRI and BamHI restriction sites were introduced flanking the future cell line barcode insertion site by amplifying two overlapping fragments with Barcode Frag1 F/R and Barcode Frag2 F/R **(Table 7)**. The two fragments were inserted within the Mph1103I and SmaI restriction sites of lentiGuide-Puro using the In-Fusion HD Cloning Kit (Clontech, #638909), according to manufacturer's instructions. Six unique 8 bp cell line barcode sequences were inserted within the EcoRI and BamHI restriction sites of the modified lentiGuide-Puro vector by resuspending the top and bottom strands of the cell line barcode oligos **(Table 7)** to a final concentration of 100 μM. Oligo pairs were phosphorylated and annealed by combining 1 μL 100 μM cell line barcode top, 1 μL 100 μM cell line barcode bottom, 5 μL 2x Quick Ligation Buffer, 1 μL T4 PNK, and 2 μL ddH_2O and incubating in thermocycler as follows: 37°C for 30 minutes, 95°C for 5 minutes, ramp down to 25°C (5°C per minute). Cell line barcode oligos were inserted within the EcoRI and BamHI restriction sites of the modified lentiGuide-Puro vector.

Gene-targeting sgRNAs sourced from the Brunello Human CRISPR Knockout Pooled Library (121) and non-targeting (NT) sgRNAs sourced from Ref. (120) and Ref. (62) were inserted into the cell line barcoded lentiGuide-Puro vectors by resuspending the top and bottom strands of the sgRNA oligos **(Table 7)** to a final concentration of 100 μM. Oligo pairs were phosphorylated and annealed by combining 1 μL 100 μM sgRNA top, 1 μL 100 μM sgRNA bottom, 5 μL 2x Quick Ligation Buffer, 1 μL T4 PNK, and 2 μL ddH_2O and incubating in thermocycler as follows: 37°C for 30 minutes, 95°C for 5 minutes, ramp down to 25°C (5°C per minute). sgRNA oligos were cloned into the barcoded lentiGuide-Puro vectors by combining

100 ng vector, 2 μL 1:200 diluted oligo duplex, 2 μL 10x FastDigest Buffer, 1 μL 10 mM DTT, 1 μL 10 mM ATP, 1 μL BsmBI, 0.5 μL Quick Ligase, and ddH_2O to a final volume of 20 μL and incubating in thermocycler as follows: 37°C for 5 minutes, 21°C for 5 minutes, repeat for a total of 6 cycles. Ligation reactions were treated with PlasmidSafe exonuclease by combining 11 μL ligation reaction, 1.5 μL 10x PlasmidSafe Buffer, 1.5 μL 10 mM ATP, and 1 μL PlasmidSafe exonuclease and incubating in thermocycler as follows: 37°C for 30 minutes, 70°C for 30 minutes. Ligation reactions were transformed into Stbl3 cells. All plasmid sequences were verified by Sanger sequencing.

Table 7. Cloning primer sequences.

Primer Name	Primer Sequence
Cas9 EcoRI F	CCCTCGTAAAGAATTCATGGACTATAAGGACCACGACGGAGA
EGFP BamHI R	GAGGTGGTCTGGATCCTTAGAATTCCTTGTACAGCTCGTCCAT
Barcode Frag1 F	AAATGTTTCATCTTCAAATGCATGCTCTT
Barcode Frag1 R	GGATCCTCCCACGCGACACGAACCTAGCCTGAATTCAAGCTTAAAAAAGCACCGACTCG
Barcode Frag2 F	GAATTCAGGCTAGGTTCGTGTCGCGTGGGAGGATCCTCCACAATTTTAAAAGAAAAGGGGGGATTGGG
Barcode Frag2 R	TATCCATCTTTGCACCCGGGCC
Barcode B11 F	aattcAGGCTAGGATCACTGGGCGTGGGAg
Barcode B11 R	gatccTCCCACGCCCAGTGATCCTAGCCTg
Barcode A09 F	aattcAGGCTAGGCGGTGACCGCGTGGGAg
Barcode A09 R	gatccTCCCACGCGGTCACCGCCTAGCCTg
Barcode A10 F	aattcAGGCTAGGTACAGAGGGCGTGGGAg
Barcode A10 R	gatccTCCCACGCCCTCTGTACCTAGCCTg
Barcode A04 F	aattcAGGCTAGGTAACAGTTGCGTGGGAg
Barcode A04 R	gatccTCCCACGCAACTGTTACCTAGCCTg
Barcode A11 F	aattcAGGCTAGGATTGTCAAGCGTGGGAg
Barcode A11 R	gatccTCCCACGCTTGACAATCCTAGCCTg
Barcode A02 F	aattcAGGCTAGGGCTGGATTGCGTGGGAg
Barcode A02 R	gatccTCCCACGCAATCCAGCCCTAGCCTg
sgNT_0 F	caccgATCGTTTCCGCTTAACGGCG
sgNT_0 R	aaacCGCCGTTAAGCGGAAACGATc
sgNT_1 F	caccgGTTCGACTCGCGTGACCGTA
sgNT_1 R	aaacTACGGTCACGCGAGTCGAACc
sgNT_2 F	caccgGTCGTCATACAACGGCAACG
sgNT_2 R	aaacCGTTGCCGTTGTATGACGACc
sgNT_3 F	caccgGCCGACCAACGTCAGCGGTA
sgNT_3 R	aaacTACCGCTGACGTTGGTCGGCc
sgNT_4 F	caccgGCGTTAGTCACTTAGCTCGA
sgNT_4 R	aaacTCGAGCTAAGTGACTAACGCc
sgNT_5 F	caccgGCGCACGTATAATCCTGCGT
sgNT_5 R	aaacACGCAGGATTATACGTGCGCc
sgNT_6 F	caccgGAGGTGATCGATTAAGTACA
sgNT_6 R	aaacTGTACTTAATCGATCACCTCc
sgNT_7 F	caccgGGTCCTGGCGAGGAGCCTTG
sgNT_7 R	aaacCAAGGCTCCTCGCCAGGACCc
sgNT_8 F	caccgGCGGAGAGCATTGACCTCAT
sgNT_8 R	aaacATGAGGTCAATGCTCTCCGCc
sgNT_9 F	caccgGAATCGGAATAGTTGATTCG
sgNT_9 R	aaacCGAATCAACTATTCCGATTCc
sgHSF1_3 F	caccgAGCTTCCACGTGTTCGACCA
sgHSF1_3 R	aaacTGGTCGAACACGTGGAAGCTc
sgHSF2_3 F	caccgACACTAACGAGTTCATCACC

sgHSF2_3 R	aaacGGTGATGAACTCGTTAGTGTc
sgERN1_2 F	caccgTCTCCCTCAATGGTACACGA
sgERN1_2 R	aaacTCGTGTACCATTGAGGGAGAc
sgXBP1_2 F	caccgCCTCCCAGGGGAATGAAGTG
sgXBP1_2 R	aaacCACTTCATTCCCCTGGGAGGc
sgATF3_4 F	caccgCCAGCGCAGAGGACATCCGG
sgATF3_4 R	aaacCCGGATGTCCTCTGCGCTGGc
sgATF4_2 F	caccgAGATGACCTTCTGACCACGT
sgATF4_2 R	aaacACGTGGTCAGAAGGTCATCTc
sgATF6_4 F	caccgTGAGCCCTGCAAATCAAAGG
sgATF6_4 R	aaacCCTTTGATTTGCAGGGCTCAc
sgNFE2L2_2 F	caccgCACATCCAGTCAGAAACCAG
sgNFE2L2_2 R	aaacCTGGTTTCTGACTGGATGTGc
sgKEAP1_1 F	caccgCAGCACCGTTCATGACGTGG
sgKEAP1_1 R	aaacCCACGTCATGAACGGTGCTGc
sgATG7_3 F	caccgCTTGAAAGACTCGAGTGTGT
sgATG7_3 R	aaacACACACTCGAGTCTTTCAAGc
sgSLC35F2_3 F	caccgCATGACAGAGCCATCAACAC
sgSLC35F2_3 R	aaacGTGTTGATGGCTCTGTCATGc

4.2 Cell culture

293T cells were cultured in DMEM medium (Gibco, #11995073) supplemented with 10% Tet System Approved Fetal Bovine Serum (Clontech, #631106) and 1% Penicillin/Streptomycin (Gibco, #15140122). To facilitate co-culturing cell lines in mixtures, ZR-75-1, SKBR3, HCC-38, MDA-MB-231, and BT-20 cells were cultured individually or collectively in a common growth medium: RPMI-1640 medium (Gibco, #11875119) supplemented with 10% Tet System Approved Fetal Bovine Serum (Clontech, #631106) and 1% Penicillin/Streptomycin (Gibco, #15140122). All cell lines were authenticated at the University of Arizona Genetics Core and tested negative for mycoplasma contamination.

4.3 Cell engineering

For engineering Cas9-expressing breast cancer cell lines, virus was produced from pLVX-TetOne Cas9 lentiviral transfer plasmid as specified in the application note for Lipofectamine 3000 Transfection Reagent (Invitrogen, #L3000015), but substituting pMD2.G envelope plasmid and psPAX2 packaging plasmid for ViraPower Lentiviral Packaging Mix. ZR-75-1, SKBR3, HCC-38, MDA-MB-231, and BT-20 cells were transduced with virus and induced with either 100 ng/mL doxycycline (ZR-75-1, SKBR3, BT-20) or 10 ng/mL doxycycline (HCC-38, MDA-MB-231) for seven days (refreshing doxycycline every three days). Induced cells were gated into three equal-size bins based on GFP brightness and sorted for medium GFP+ cells.

For introducing sgRNA libraries, pooled sgRNA plasmid libraries possessing appropriate cell line barcodes were generated by combining equal amounts of 11 gene-targeting sgRNA plasmids and a pool of five NT sgRNA plasmids (sgNT_0-4). Virus was produced with Lipofectamine 3000 Transfection Reagent (Invitrogen, #L3000015) as described above and was functionally titered on its respective cell line using serially-diluted virus. ZR-75-1, SKBR3, HCC-38, MDA-MB-231, and BT-20 Cas9 cells were transduced with appropriate sgRNA libraries at MOI = 0.3, maintaining coverage of at least 7,500 cells per sgRNA. Transduced cells were selected using 2 μg/mL puromycin (InvivoGen, #ant-pr-1) for three days.

For generating cell spike-in standards, virus for five NT sgRNA plasmids (sgNT_5-9) possessing the 293T cell line barcode was produced in arrayed format. 293T cells were transduced with the viruses individually. Transduced cells were selected using 2 μg/mL puromycin (InvivoGen, #ant-pr-1) for three days. To pool, cell lines were detached, resuspended in PBS, counted, and pooled at the following ratio: 1x sgNT_5, 3x sgNT_6, 9x sgNT_7, 27x sgNT_8, 81x sgNT_9. Pooled cells were aliquoted, pelleted at 425 x *g* for 5 minutes, and stored at -20°C.

4.4 Western blot analysis

For assessing induction of FLAG-Cas9, MDA-MB-231 pLVX-TetOne Cas9 cells were treated with 0, 0.5, 1, 2, 5, or 10 ng/mL doxycycline (Clontech, #631311) for 48 hours. Cells were harvested and lysed in buffer containing 50 mM Tris, pH 7.5, 1 mM EDTA, 150 mM NaCl, 1% Triton X-100, 0.1% SDS. Protein concentration was measured using the BCA Protein Assay Kit (Pierce, #23225). 5 μg total protein per lane was electrophoresed and transferred using an iBlot 2 Dry Blotting System (Thermo Fisher Scientific). Membrane was probed with 1:1000 Anti-FLAG primary antibody (Sigma-Aldrich, #F3165) followed by 1:10,000 Anti-Mouse IgG-Peroxidase secondary antibody (Sigma-Aldrich, #A9044), developed with Immobilon Western Chemiluminescent HRP Substrate (Millipore, #WBKLS0500), visualized using a ChemiDoc Touch Imaging System (Bio-Rad), and analyzed using Image Lab 5.2.1 (Bio-Rad). Membrane was stripped with ReBlot Plus Mild Antibody Stripping Solution (Millipore, #2502) and reprobed with 1:10,000 Anti-Alpha Tubulin primary antibody (abcam, #ab80779) followed by 1:10,000 Anti-Mouse IgG-Peroxidase secondary antibody (Sigma-Aldrich, #A9044).

For confirming whole population knockout of the proteostasis factors, MDA-MB-231 pLVX-TetOne Cas9 cells transduced with appropriate sgRNAs were treated with 10 ng/mL doxycycline (Clontech, #631311) for 96 hours (refreshing doxycycline every two days) to induce Cas9 expression prior to harvesting. Western blot analysis was performed as described above using the following antibodies: 1:1000 Anti-HSF1 (Santa Cruz Biotechnology, #sc-9144), 1:1000 Anti-HSF2 (Santa Cruz Biotechnology, #sc-13517), 1:1000 Anti-IRE1 (Cell Signaling Technology, #3294), 1:1000 Anti-XBP1 (Cell Signaling Technology, #12782), 1:1000 Anti-ATF3 (abcam, #ab207434), 1:1000 Anti-ATF4 (Cell Signaling Technology, #11815), 1:1000 Anti-ATF6 (Cell Signaling Technology, #65880), 1:1000 Anti-ATG7 (Cell Signaling Technology, #8558), 1:1000 Anti-NRF2 (Cell Signaling Technology, #12721), 1:1000 Anti-KEAP1 (Cell Signaling Technology, #4617), 1:10,000 Anti-Alpha Tubulin (abcam, #ab80779), 1:10,000 Anti-Beta Actin (Thermo Fisher Scientific, #MA5-15739).

4.5 Relative cell abundance competition experiment

ZR-75-1, SKBR3, HCC-38, MDA-MB-231, and BT-20 cells were transduced with pHIV-Luc-ZsGreen (Addgene, Plasmid #39196) or pUltra-Chili-Luc (Addgene, Plasmid #48688) and sorted for GFP+ or RFP+ cells, as appropriate.

For preparing the five original pools, fluorescently labeled cell lines were counted, pooled, and frozen in liquid nitrogen. Pools were thawed on Day 0 and cultured normally. Six days and 13 days after thawing, the percentages of GFP+ and RFP+ cells were quantified by flow cytometry analysis using a LSRFortessa Cell Analyzer (BD Biosciences).

For estimating the growth rate (r) of each of the five cell lines, the equation for exponential cell growth was used:

$$x_t = x_0(1+r)^t,$$

where x_t = percentage at day t, x_0 = percentage at day 0, r = growth rate, and t = time [days]. Growth rates (r) were then used to calculate the optimal starting percentage (x_0) of each cell line to achieve 20% representation at t = 7 days.

For preparing the five optimized pools, fluorescently labeled cell lines were counted, pooled according to our model, and frozen in liquid nitrogen. Pools were thawed on Day 0 and cultured normally. Seven days after thawing, the percentages of GFP+ and RFP+ cells were quantified by flow cytometry analysis using a LSRFortessa Cell Analyzer (BD Biosciences). Heat maps were generated using Cluster 3.0 and Java TreeView 1.1.6r4.

4.6 Cell pooling

For preparing cell pools for QMAP-Seq with multiple cell lines (five cell lines x 12 genetic perturbations), individual cell line pools consisting of 12 sgRNAs (11 gene-targeting sgRNAs + 5 pooled NT sgRNAs) were prepared as described in “Cell engineering”. Each of the five individual cell line pools was counted, pooled with the other four cell line pools according to ratios derived from the relative cell abundance competition experiment, and frozen in liquid nitrogen (2,000,000 total cells per vial).

4.7 Cell viability assays

For confirming the efficacy of the sgRNAs that target SLC35F2, MDA-MB-231 pLVX-TetOne Cas9 cells engineered with either sgNT or one of the four sgRNAs targeting SLC35F2 were induced with 100 ng/mL doxycycline for six days (refreshing doxycycline every three days). In total, 1,000 cells were seeded in a volume of 50 μL in opaque 384-well plates. The next day, YM155 was added over the indicated nine-point concentration range in quadruplicate using a D300e Digital Dispenser (Tecan). 72 hours later, cell viability was measured using the CellTiter-Glo Luminescent Cell Viability Assay (Promega, #G7572). Luminescence was read using an Infinite M1000 PRO (Tecan) with an integration time of 500 ms.

For live cell imaging, pooled MDA-MB-231 pLVX-TetOne Cas9 cells engineered with 12 sgRNAs were induced with 100 ng/mL doxycycline for six days (refreshing doxycycline every three days). In total, 5,000 pooled cells were seeded in a volume of 100 μL in 96-well plates. The next day, 89 compounds were added over the indicated four-point concentration range in duplicate from custom compound plates prepared at the High Throughput Analysis Laboratory (Northwestern University). 72 hours later, the percent confluence of the population of cells in each well was measured (4X objective, whole-well imaging, phase

channel) using an IncuCyte ZOOM Live-Cell Analysis System GUI version 2015A (Essen BioScience). To calculate relative cell confluence, the percent confluence of a compound-treated well was normalized to the median percent confluence of DMSO-treated wells.

For validating QMAP-Seq hits using Resazurin, MDA-MB-231 pLVX-TetOne Cas9 cells engineered with appropriate sgRNAs were induced with 100 ng/mL doxycycline for six days (refreshing doxycycline every three days). In total, 5,000 cells were seeded in a volume of 100 μL in 96-well plates. The next day, Carfilzomib was added over the indicated four-point concentration range in triplicate using a D300e Digital Dispenser (Tecan). 72 hours later, cell viability was measured using the Resazurin Cell Viability Kit (Cell Signaling Technology, #11884). Relative fluorescent units were read (Excitation = 550 nm, Emission = 605 nm) using an Infinite M1000 PRO (Tecan).

For all cell viability assays, dose-response curves were fit in GraphPad Prism 8 using the log(inhibitor) vs. response model (three parameters) with the top constrained to 100%.

4.8 Compounds

YM155 was obtained from Selleckchem (#S1130). Compounds for QMAP-Seq were obtained from the FDA-Approved Drug Library (MedChemExpress, #HY-L022), the Clinical Compound Library (MedChemExpress, #HY-L026), or from the following vendors: 4-Hydroxytamoxifen was obtained from Sigma-Aldrich (#H7904), and Bortezomib was obtained from Cayman Chemical (#10008822). For validation experiments, Carfilzomib was obtained from a different vendor (Cayman Chemical, #17554) than where it was sourced for QMAP-Seq.

4.9 Selection of compounds and doses for QMAP-Seq

Compounds for QMAP-Seq with multiple cell lines spanned six categories: positive controls, breast cancer chemotherapeutics, targeted therapies (NCI), targeted therapies (OncoKB), diverse compounds (Informer Set), and proteostasis-modulating compounds. Relevant positive controls (4-Hydroxytamoxifen, Fulvestrant, Lapatinib, and YM155) were included. For all remaining categories, compounds that were not part of the FDA-Approved Drug Library (MedChemExpress, #HY-L022) or the Clinical Compound Library

(MedChemExpress, #HY-L026) were filtered out. One compound per compound class was selected from the breast cancer chemotherapeutics. Compounds that were classified as OncoKB Levels 1-4 for the indications of breast cancer, all solid tumors, or all tumors were selected from the OncoKB targeted therapies (122). The existence of specific genetic alterations in breast cancer was confirmed using cBioPortal (123, 124). Up to two proteostasis-modulating compounds per target were incorporated.

Compounds were applied over a four-point concentration range (10-fold dilutions) in duplicate using one of two dose ranges. The standard dose range encompassed doses from 10 nM to 10 μM, and the low dose range covered doses from 1 nM to 1 μM. Dose ranges were selected by referencing doses previously used for treating cancer cell lines, including the Informer Set (75), Cancer Therapeutics Response Portal v2 (74-76), Genomics of Drug Sensitivity in Cancer (71-73), and Selleckchem.

4.10 QMAP-Seq

4.10.1 Induction, seeding, and compound treatment

A step-by-step protocol describing the QMAP-Seq protocol can be found at Protocol Exchange (125). For QMAP-Seq with multiple cell lines, pooled cell lines (five cell lines x 12 genetic perturbations) were thawed in media containing 100 ng/mL doxycycline (Clontech, #631311). Three days later, cells were expanded and doxycycline was refreshed. Three more days later, 5,000 pooled cells were seeded in a volume of 100 μL in 96-well plates using an EL406 Microplate Washer Dispenser (BioTek Instruments), while maintaining doxycycline induction. The next day, 180 compounds were added over a four-point concentration range in duplicate from custom compound plates prepared at the High Throughput Analysis Laboratory (Northwestern University). Briefly, compound plates containing 500 nL of compound at 1,000x were resuspended in 250 μL media per well to achieve 2x compound concentration, and 100 μL of 2x compound was distributed to replicate wells using multichannel pipet. Cells were treated for 72 hours.

4.10.2 Lysis of cell spike-in standards

Cell spike-in standards were thawed at room temperature for 5 minutes and resuspended in Lysis Buffer (10% 10x Taq DNA Polymerase Buffer (Invitrogen, #18067017), 0.45% IGEPAL CA-630 (Sigma-

Aldrich, #I8896), 0.45% TWEEN 20 (Sigma-Aldrich, P9416), 10% Proteinase K (Qiagen, #19133), 79.1% Nuclease-Free Water (Qiagen, #129115)) to achieve a concentration of 36.3 total cell spike-in standards/μL. Cells were homogenized using a 5 mL syringe and a 21G x 1" needle three times followed by a 27G x ½" needle three times. Cells were incubated in 60°C water bath for 1 hour, pipetting up and down every 20 minutes.

4.10.3 Lysis of compound-treated cells

After 72 hours of compound treatment, compound-treated cells were washed with 100 μL PBS. 50 μL Lysis Buffer containing 36.3 total cell spike-in standards/μL (1,815 total spike-in cells/well) was added using multichannel pipet. Plates were covered with foil adhesive and incubated in 60°C oven for 1 hour. Cell lysates were transferred to PCR plates using multichannel pipet. Proteinase K from Lysis Buffer was inactivated at 95 ºC for 15 minutes using thermocycler.

4.10.4 Sequencing library preparation

A 359-367 bp fragment (depending on stagger length) containing the sgRNA and cell line barcode was amplified with a unique set of P5 and P7 primers for each well by combining the following reagents: 6 μL 10x Taq DNA Polymerase Buffer, 1.2 μL 10 mM dNTP Mix, 1.8 μL 50 mM $MgCl_2$, 0.6 μL DMSO (Sigma-Aldrich, #D8418), 6 μL 1 μM P5 Primer Mix (mix of 0-8nt staggered primers), 6 μL 1 μM P7 Primer, 8 μL Lysate (Proteinase K inactivated again just prior to PCR), 30.16 μL Water, 0.24 μL Platinum Taq DNA Polymerase (Invitrogen, #10966034). To reduce the likelihood of PCR jackpot effects, half of the PCR volume from each reaction was transferred to a second set of PCR plates. PCRs were run in thermocycler as follows: 94 ºC for 4 minutes, (94 ºC for 30 seconds, 60 ºC for 30 seconds, 72 ºC for 30 seconds, repeat for a total of 29 cycles), 72 ºC for 15 minutes, 4 ºC hold. Technical duplicates were recombined.

Library size was verified by running 5 μL of PCR products on E-Gel 96 2% Agarose gels (Invitrogen, #G7008-02). An equal volume (6 μL) of each PCR product was pooled together. The concentration of the pooled PCR products was measured using the Qubit dsDNA HS Assay Kit (Invitrogen, #Q32854). Pooled PCR products were purified using the QIAquick PCR Purification Kit (Qiagen, #28106) with sufficient PCR

purification columns to avoid exceeding the maximum binding capacity of each column. The concentration of the purified sample was measured using the Qubit dsDNA HS Assay Kit (Invitrogen, #Q32854). The purified sample was purified a second time using the Purification Module with Magnetic Beads (Lexogen, #022.96). The final library concentration was measured using the Qubit dsDNA HS Assay Kit (Invitrogen, #Q32854).

4.10.5 Next-generation sequencing

The sequencing library was diluted to 2.5 nM, combined with PhiX (Illumina, #FC-110-3001) to achieve 25% PhiX (to increase nucleotide diversity), and denatured. The sequencing library was loaded on a NovaSeq 6000 (Illumina) using a NovaSeq 6000 S1 Reagent Kit, 200 Cycles, 1.3B Reads (Illumina, #20012864). Single-end sequencing was performed using the following run parameters: Illumina Read 1 Primer: 164 cycles (to sequence sgRNA, cell line barcode), Illumina Index 1 Primer: 6 cycles (to sequence i7 index), Illumina Index 2 Primer: 6 cycles (to sequence i5 index).

4.10.6 Sequencing data processing and data analysis

Individual samples were demultiplexed based on i5 and i7 index sequences by running bcl2fastq2 Conversion Software v2.20 (Illumina). sgRNA and cell line barcode sequences were extracted from Read 1 sequences as follows: [ACACCG][sgRNA:20][Interval:102][CellLineBarcode:8]. The number of reads for each cell line-sgRNA pair was counted, allowing up to one nucleotide mismatch per barcode. Next, cell number was interpolated from sequencing reads using sample-specific standard curves. The number of cells for a compound-treated sample was normalized to the median number of cells for the DMSO-treated samples.

Relative cell numbers were plotted as heat maps and dose-response curves. Heat maps were generated using Cluster 3.0 and Java TreeView 1.1.6r4. Dose-response curves were fit in GraphPad Prism 8 using the log(inhibitor) vs. response model (three parameters) with the top constrained to 100%. The area under the curve (AUC) was calculated using the equation: AUC = 0.5(Dose 1 Cell Number) + Dose 2 Cell Number + Dose 3 Cell Number + 0.5(Dose 4 Cell Number).

Data was filtered using the following exclusion criteria. First, samples with low read counts were excluded, which included these compounds: ABT199, Belinostat, BMS345541, Dexamethasone, Fingolimod, and RacRotigotine. Second, cell line-sgRNA pairs with high variation (standard deviation of $\log_2$(cell number) > 0.7) were excluded, which included these cell line-sgRNA pairs: ZR-75-1 sgHSF1 and ZR-75-1 sgKEAP1. Third, compounds with an AUC > 550 for any cell line-sgRNA pair were excluded, which included these compounds: Ruxolitinib and STF083010 (both in well B7, suggesting a likely technical problem with that well). Fourth, noncytotoxic cell line-compound pairs (relative cell number of sgNT with highest dose ≥ 90%) were excluded, which included 225 cell line-compound pairs. Fifth, cell line-compound pairs without a significant dose-dependent reduction in cell viability (difference in relative cell number of sgNT between lowest and highest dose < 25%) were excluded, which included 366 cell line-compound pairs (union between fourth and fifth exclusion criteria = 380 cell line-compound pairs).

4.10.7 QMAP-Seq with one cell line

Experimental and analysis workflows were performed as described for QMAP-Seq with multiple cell lines with the following modifications. MDA-MB-231 cells with 12 genetic perturbations were used instead of five cell lines with 12 genetic perturbations each. Because each sample had five times as many cells per perturbation in the one cell line experiment (5,000 cells ÷ 12 perturbations ≈ 417 cells per perturbation) compared to in the five cell line experiment (5,000 cells ÷ 60 perturbations ≈ 83 cells per perturbation), five times as many cells for each of the cell spike-in standards were added per sample for the one cell line experiment. The data from Plate 8 was excluded from analysis due to a technical problem with the addition of cell spike-in standards for that plate. The sequencing library was loaded on a NextSeq 500 (Illumina) using a NextSeq 500/550 High Output Reagent Kit, 400M Reads (Illumina, #20024906).

4.11 Network analysis

For constructing a chemical-genetic interaction network, data was filtered as described above. All interactions were then filtered for those considered significant ($P < 0.05$) and that had a large effect size

(absolute magnitude of AUC change > 60). These interactions were visualized as an unweighted network using a standard force-directed layout in Cytoscape v3.7.2 (cytoscape.org).

For assessing functional similarity between proteostasis genes targeted in our sgRNA library, data was filtered as described above. The AUC difference was quantified for each genetic perturbation across all 488 cell line-compound contexts. A Spearman correlation was then calculated for all gene pairs based on overall similarity of their chemical-genetic interaction profiles.

4.12 Statistical analysis

Statistical analysis was performed with GraphPad Prism 8 statistical software. Replicate measurements were taken from distinct biological samples.

For correlation analysis of replicates, Pearson r was reported, and statistical significance of Pearson correlation was determined using a two-tailed test (n = 288 compound-dose combinations after excluding Plate 8 data). For correlation analysis of AUCs between live cell imaging and QMAP-Seq, Pearson r was reported, and statistical significance of Pearson correlation was determined using a two-tailed test (n = 89 compounds). For correlation analysis of AUCs between two independent QMAP-Seq experiments, Pearson r was reported, and statistical significance of Pearson correlation was determined using a two-tailed test (n = 220 compound-sgRNA combinations). For comparing the variation between plates for reads versus cell number, statistical significance was determined using an unpaired, two-tailed F test to compare variances (n = 16 plates).

For identifying the chemical-genetic interactions with the greatest effect, a volcano plot was produced after data filtering as described above. The magnitude was determined by calculating the difference in mean AUC between sgRNA and sgNT for every cell line-compound combination. The statistical significance was determined using an unpaired, two-tailed t test (n = 2 biologically independent replicates). The chemical-genetic interactions with an Absolute AUC Difference > 25 and $P < 0.05$ were designated as hits.

For determining compound pathway enrichment, statistical significance was determined using a one-tailed binomial test to compare observed distribution to expected distribution (n = 180 compounds).

4.13 Data availability

The data discussed in this dissertation, including raw fastq files, read counts, and relative cell numbers, have been deposited in NCBI's Gene Expression Omnibus and are accessible through GEO Series accession number GSE155855 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE155855).

4.14 Code availability

Custom code is available at GitHub (https://github.com/mendillolab/QMAP-Seq) (126) and Code Ocean (https://codeocean.com/capsule/3022355/tree/v1) (127).

REFERENCES

1. Balch WE, Morimoto RI, Dillin A, Kelly JW. Adapting proteostasis for disease intervention. Science. 2008;319(5865):916-9. Epub 2008/02/16. doi: 10.1126/science.1141448. PubMed PMID: 18276881.
2. Himanen SV, Sistonen L. New insights into transcriptional reprogramming during cellular stress. J Cell Sci. 2019;132(21). Epub 2019/11/05. doi: 10.1242/jcs.238402. PubMed PMID: 31676663.
3. Galluzzi L, Yamazaki T, Kroemer G. Linking cellular stress responses to systemic homeostasis. Nat Rev Mol Cell Biol. 2018;19(11):731-45. Epub 2018/10/12. doi: 10.1038/s41580-018-0068-0. PubMed PMID: 30305710.
4. Rabindran SK, Haroun RI, Clos J, Wisniewski J, Wu C. Regulation of heat shock factor trimer formation: role of a conserved leucine zipper. Science. 1993;259(5092):230-4. Epub 1993/01/08. doi: 10.1126/science.8421783. PubMed PMID: 8421783.
5. Xu YM, Huang DY, Chiu JF, Lau AT. Post-translational modification of human heat shock factors and their functions: a recent update by proteomic approach. J Proteome Res. 2012;11(5):2625-34. Epub 2012/04/13. doi: 10.1021/pr201151a. PubMed PMID: 22494029.
6. Jaeger AM, Makley LN, Gestwicki JE, Thiele DJ. Genomic heat shock element sequences drive cooperative human heat shock factor 1 DNA binding and selectivity. J Biol Chem. 2014;289(44):30459-69. Epub 2014/09/11. doi: 10.1074/jbc.M114.591578. PubMed PMID: 25204655; PMCID: PMC4215228.
7. Pelham HR. A regulatory upstream promoter element in the Drosophila hsp 70 heat-shock gene. Cell. 1982;30(2):517-28. Epub 1982/09/01. doi: 10.1016/0092-8674(82)90249-5. PubMed PMID: 6814763.
8. Rallu M, Loones M, Lallemand Y, Morimoto R, Morange M, Mezger V. Function and regulation of heat shock factor 2 during mouse embryogenesis. Proc Natl Acad Sci U S A. 1997;94(6):2392-7. Epub 1997/03/18. doi: 10.1073/pnas.94.6.2392. PubMed PMID: 9122205; PMCID: PMC20098.

9. Mathew A, Mathur SK, Morimoto RI. Heat shock response and protein degradation: regulation of HSF2 by the ubiquitin-proteasome pathway. Mol Cell Biol. 1998;18(9):5091-8. Epub 1998/08/26. doi: 10.1128/mcb.18.9.5091. PubMed PMID: 9710593; PMCID: PMC109094.

10. Ostling P, Bjork JK, Roos-Mattjus P, Mezger V, Sistonen L. Heat shock factor 2 (HSF2) contributes to inducible expression of hsp genes through interplay with HSF1. J Biol Chem. 2007;282(10):7077-86. Epub 2007/01/11. doi: 10.1074/jbc.M607556200. PubMed PMID: 17213196.

11. Loison F, Debure L, Nizard P, le Goff P, Michel D, le Drean Y. Up-regulation of the clusterin gene after proteotoxic stress: implication of HSF1-HSF2 heterocomplexes. Biochem J. 2006;395(1):223-31. Epub 2005/12/13. doi: 10.1042/bj20051190. PubMed PMID: 16336210; PMCID: PMC1409688.

12. Sandqvist A, Bjork JK, Akerfelt M, Chitikova Z, Grichine A, Vourc'h C, Jolly C, Salminen TA, Nymalm Y, Sistonen L. Heterotrimerization of heat-shock factors 1 and 2 provides a transcriptional switch in response to distinct stimuli. Mol Biol Cell. 2009;20(5):1340-7. Epub 2009/01/09. doi: 10.1091/mbc.E08-08-0864. PubMed PMID: 19129477; PMCID: PMC2649261.

13. Rossi A, Riccio A, Coccia M, Trotta E, La Frazia S, Santoro MG. The proteasome inhibitor bortezomib is a potent inducer of zinc finger AN1-type domain 2a gene expression: role of heat shock factor 1 (HSF1)-heat shock factor 2 (HSF2) heterocomplexes. J Biol Chem. 2014;289(18):12705-15. Epub 2014/03/13. doi: 10.1074/jbc.M113.513242. PubMed PMID: 24619424; PMCID: PMC4007460.

14. Korfanty J, Stokowy T, Widlak P, Gogler-Piglowska A, Handschuh L, Podkowinski J, Vydra N, Naumowicz A, Toma-Jonik A, Widlak W. Crosstalk between HSF1 and HSF2 during the heat shock response in mouse testes. Int J Biochem Cell Biol. 2014;57:76-83. Epub 2014/12/03. doi: 10.1016/j.biocel.2014.10.006. PubMed PMID: 25450459.

15. Hetz C. The unfolded protein response: controlling cell fate decisions under ER stress and beyond. Nat Rev Mol Cell Biol. 2012;13(2):89-102. Epub 2012/01/19. doi: 10.1038/nrm3270. PubMed PMID: 22251901.

16. Senft D, Ronai ZA. UPR, autophagy, and mitochondria crosstalk underlies the ER stress response. Trends Biochem Sci. 2015;40(3):141-8. Epub 2015/02/07. doi: 10.1016/j.tibs.2015.01.002. PubMed PMID: 25656104; PMCID: PMC4340752.

17. Yoshida H, Matsui T, Yamamoto A, Okada T, Mori K. XBP1 mRNA is induced by ATF6 and spliced by IRE1 in response to ER stress to produce a highly active transcription factor. Cell. 2001;107(7):881-91. Epub 2002/01/10. doi: 10.1016/s0092-8674(01)00611-0. PubMed PMID: 11779464.

18. Harding HP, Zhang Y, Ron D. Protein translation and folding are coupled by an endoplasmic-reticulum-resident kinase. Nature. 1999;397(6716):271-4. Epub 1999/02/04. doi: 10.1038/16729. PubMed PMID: 9930704.

19. Harding HP, Novoa I, Zhang Y, Zeng H, Wek R, Schapira M, Ron D. Regulated translation initiation controls stress-induced gene expression in mammalian cells. Mol Cell. 2000;6(5):1099-108. Epub 2000/12/07. doi: 10.1016/s1097-2765(00)00108-8. PubMed PMID: 11106749.

20. Ye J, Rawson RB, Komuro R, Chen X, Davé UP, Prywes R, Brown MS, Goldstein JL. ER stress induces cleavage of membrane-bound ATF6 by the same proteases that process SREBPs. Mol Cell. 2000;6(6):1355-64. Epub 2001/02/13. doi: 10.1016/s1097-2765(00)00133-7. PubMed PMID: 11163209.

21. Shoulders MD, Ryno LM, Genereux JC, Moresco JJ, Tu PG, Wu C, Yates JR, 3rd, Su AI, Kelly JW, Wiseman RL. Stress-independent activation of XBP1s and/or ATF6 reveals three functionally diverse ER proteostasis environments. Cell Rep. 2013;3(4):1279-92. Epub 2013/04/16. doi: 10.1016/j.celrep.2013.03.024. PubMed PMID: 23583182; PMCID: PMC3754422.

22. Adamson B, Norman TM, Jost M, Cho MY, Nunez JK, Chen Y, Villalta JE, Gilbert LA, Horlbeck MA, Hein MY, Pak RA, Gray AN, Gross CA, Dixit A, Parnas O, Regev A, Weissman JS. A Multiplexed Single-Cell CRISPR Screening Platform Enables Systematic Dissection of the Unfolded Protein Response. Cell. 2016;167(7):1867-82.e21. Epub 2016/12/17. doi: 10.1016/j.cell.2016.11.048. PubMed PMID: 27984733; PMCID: PMC5315571.

23. Zhang DD, Lo SC, Cross JV, Templeton DJ, Hannink M. Keap1 is a redox-regulated substrate adaptor protein for a Cul3-dependent ubiquitin ligase complex. Mol Cell Biol. 2004;24(24):10941-53. Epub 2004/12/02. doi: 10.1128/mcb.24.24.10941-10953.2004. PubMed PMID: 15572695; PMCID: PMC533977.

24. Itoh K, Wakabayashi N, Katoh Y, Ishii T, Igarashi K, Engel JD, Yamamoto M. Keap1 represses nuclear activation of antioxidant responsive elements by Nrf2 through binding to the amino-terminal Neh2 domain. Genes Dev. 1999;13(1):76-86. Epub 1999/01/14. doi: 10.1101/gad.13.1.76. PubMed PMID: 9887101; PMCID: PMC316370.

25. Cullinan SB, Gordan JD, Jin J, Harper JW, Diehl JA. The Keap1-BTB protein is an adaptor that bridges Nrf2 to a Cul3-based E3 ligase: oxidative stress sensing by a Cul3-Keap1 ligase. Mol Cell Biol. 2004;24(19):8477-86. Epub 2004/09/16. doi: 10.1128/mcb.24.19.8477-8486.2004. PubMed PMID: 15367669; PMCID: PMC516753.

26. Kobayashi A, Kang MI, Okawa H, Ohtsuji M, Zenke Y, Chiba T, Igarashi K, Yamamoto M. Oxidative stress sensor Keap1 functions as an adaptor for Cul3-based E3 ligase to regulate proteasomal degradation of Nrf2. Mol Cell Biol. 2004;24(16):7130-9. Epub 2004/07/30. doi: 10.1128/mcb.24.16.7130-7139.2004. PubMed PMID: 15282312; PMCID: PMC479737.

27. Baird L, Swift S, Llères D, Dinkova-Kostova AT. Monitoring Keap1-Nrf2 interactions in single live cells. Biotechnol Adv. 2014;32(6):1133-44. Epub 2014/04/01. doi: 10.1016/j.biotechadv.2014.03.004. PubMed PMID: 24681086; PMCID: PMC4165437.

28. Itoh K, Chiba T, Takahashi S, Ishii T, Igarashi K, Katoh Y, Oyake T, Hayashi N, Satoh K, Hatayama I, Yamamoto M, Nabeshima Y. An Nrf2/small Maf heterodimer mediates the induction of phase II detoxifying enzyme genes through antioxidant response elements. Biochem Biophys Res Commun. 1997;236(2):313-22. Epub 1997/07/18. doi: 10.1006/bbrc.1997.6943. PubMed PMID: 9240432.

29. Lee JM, Calkins MJ, Chan K, Kan YW, Johnson JA. Identification of the NF-E2-related factor-2-dependent genes conferring protection against oxidative stress in primary cortical astrocytes using

oligonucleotide microarray analysis. J Biol Chem. 2003;278(14):12029-38. Epub 2003/01/31. doi: 10.1074/jbc.M211558200. PubMed PMID: 12556532.

30. Malhotra D, Portales-Casamar E, Singh A, Srivastava S, Arenillas D, Happel C, Shyr C, Wakabayashi N, Kensler TW, Wasserman WW, Biswal S. Global mapping of binding sites for Nrf2 identifies novel targets in cell survival response through ChIP-Seq profiling and network analysis. Nucleic Acids Res. 2010;38(17):5718-34. Epub 2010/05/13. doi: 10.1093/nar/gkq212. PubMed PMID: 20460467; PMCID: PMC2943601.

31. Clague MJ, Urbé S. Ubiquitin: same molecule, different degradation pathways. Cell. 2010;143(5):682-5. Epub 2010/11/30. doi: 10.1016/j.cell.2010.11.012. PubMed PMID: 21111229.

32. White E. Deconvoluting the context-dependent role for autophagy in cancer. Nat Rev Cancer. 2012;12(6):401-10. Epub 2012/04/27. doi: 10.1038/nrc3262. PubMed PMID: 22534666; PMCID: PMC3664381.

33. Levine B, Kroemer G. Autophagy in the pathogenesis of disease. Cell. 2008;132(1):27-42. Epub 2008/01/15. doi: 10.1016/j.cell.2007.12.018. PubMed PMID: 18191218; PMCID: PMC2696814.

34. Rabinowitz JD, White E. Autophagy and metabolism. Science. 2010;330(6009):1344-8. Epub 2010/12/04. doi: 10.1126/science.1193497. PubMed PMID: 21127245; PMCID: PMC3010857.

35. Green DR, Levine B. To be or not to be? How selective autophagy and cell death govern cell fate. Cell. 2014;157(1):65-75. Epub 2014/04/01. doi: 10.1016/j.cell.2014.02.049. PubMed PMID: 24679527; PMCID: PMC4020175.

36. Dai C, Whitesell L, Rogers AB, Lindquist S. Heat shock factor 1 is a powerful multifaceted modifier of carcinogenesis. Cell. 2007;130(6):1005-18. Epub 2007/09/25. doi: 10.1016/j.cell.2007.07.020. PubMed PMID: 17889646; PMCID: PMC2586609.

37. Santagata S, Hu R, Lin NU, Mendillo ML, Collins LC, Hankinson SE, Schnitt SJ, Whitesell L, Tamimi RM, Lindquist S, Ince TA. High levels of nuclear heat-shock factor 1 (HSF1) are associated with poor prognosis in breast cancer. Proc Natl Acad Sci U S A. 2011;108(45):18378-83. Epub 2011/11/02. doi: 10.1073/pnas.1115031108. PubMed PMID: 22042860; PMCID: PMC3215027.

38. Mendillo ML, Santagata S, Koeva M, Bell GW, Hu R, Tamimi RM, Fraenkel E, Ince TA, Whitesell L, Lindquist S. HSF1 drives a transcriptional program distinct from heat shock to support highly malignant human cancers. Cell. 2012;150(3):549-62. Epub 2012/08/07. doi: 10.1016/j.cell.2012.06.031. PubMed PMID: 22863008; PMCID: PMC3438889.

39. Chen X, Iliopoulos D, Zhang Q, Tang Q, Greenblatt MB, Hatziapostolou M, Lim E, Tam WL, Ni M, Chen Y, Mai J, Shen H, Hu DZ, Adoro S, Hu B, Song M, Tan C, Landis MD, Ferrari M, Shin SJ, Brown M, Chang JC, Liu XS, Glimcher LH. XBP1 promotes triple-negative breast cancer by controlling the HIF1alpha pathway. Nature. 2014;508(7494):103-7. Epub 2014/03/29. doi: 10.1038/nature13119. PubMed PMID: 24670641; PMCID: PMC4105133.

40. Zhang C, Wang HJ, Bao QC, Wang L, Guo TK, Chen WL, Xu LL, Zhou HS, Bian JL, Yang YR, Sun HP, Xu XL, You QD. NRF2 promotes breast cancer cell proliferation and metastasis by increasing RhoA/ROCK pathway signal transduction. Oncotarget. 2016;7(45):73593-606. Epub 2016/10/08. doi: 10.18632/oncotarget.12435. PubMed PMID: 27713154; PMCID: PMC5342001.

41. Desai S, Liu Z, Yao J, Patel N, Chen J, Wu Y, Ahn EE, Fodstad O, Tan M. Heat shock factor 1 (HSF1) controls chemoresistance and autophagy through transcriptional regulation of autophagy-related protein 7 (ATG7). J Biol Chem. 2013;288(13):9165-76. Epub 2013/02/07. doi: 10.1074/jbc.M112.422071. PubMed PMID: 23386620; PMCID: PMC3610989.

42. Beltran H, Prandi D, Mosquera JM, Benelli M, Puca L, Cyrta J, Marotz C, Giannopoulou E, Chakravarthi BV, Varambally S, Tomlins SA, Nanus DM, Tagawa ST, Van Allen EM, Elemento O, Sboner A, Garraway LA, Rubin MA, Demichelis F. Divergent clonal evolution of castration-resistant neuroendocrine prostate cancer. Nat Med. 2016;22(3):298-305. Epub 2016/02/09. doi: 10.1038/nm.4045. PubMed PMID: 26855148; PMCID: PMC4777652.

43. Witkiewicz AK, McMillan EA, Balaji U, Baek G, Lin WC, Mansour J, Mollaee M, Wagner KU, Koduru P, Yopp A, Choti MA, Yeo CJ, McCue P, White MA, Knudsen ES. Whole-exome sequencing of pancreatic cancer defines genetic diversity and therapeutic targets. Nat Commun. 2015;6:6744. Epub 2015/04/10. doi: 10.1038/ncomms7744. PubMed PMID: 25855536; PMCID: PMC4403382.

44. Kumar A, Coleman I, Morrissey C, Zhang X, True LD, Gulati R, Etzioni R, Bolouri H, Montgomery B, White T, Lucas JM, Brown LG, Dumpit RF, DeSarkar N, Higano C, Yu EY, Coleman R, Schultz N, Fang M, Lange PH, Shendure J, Vessella RL, Nelson PS. Substantial interindividual and limited intraindividual genomic diversity among tumors from men with metastatic prostate cancer. Nat Med. 2016;22(4):369-78. Epub 2016/03/02. doi: 10.1038/nm.4053. PubMed PMID: 26928463; PMCID: PMC5045679.

45. Comprehensive genomic characterization of squamous cell lung cancers. Nature. 2012;489(7417):519-25. Epub 2012/09/11. doi: 10.1038/nature11404. PubMed PMID: 22960745; PMCID: PMC3466113.

46. Comprehensive genomic characterization of head and neck squamous cell carcinomas. Nature. 2015;517(7536):576-82. Epub 2015/01/30. doi: 10.1038/nature14129. PubMed PMID: 25631445; PMCID: PMC4311405.

47. Comprehensive molecular profiling of lung adenocarcinoma. Nature. 2014;511(7511):543-50. Epub 2014/08/01. doi: 10.1038/nature13385. PubMed PMID: 25079552; PMCID: PMC4231481.

48. Jordan EJ, Kim HR, Arcila ME, Barron D, Chakravarty D, Gao J, Chang MT, Ni A, Kundra R, Jonsson P, Jayakumaran G, Gao SP, Johnsen HC, Hanrahan AJ, Zehir A, Rekhtman N, Ginsberg MS, Li BT, Yu HA, Paik PK, Drilon A, Hellmann MD, Reales DN, Benayed R, Rusch VW, Kris MG, Chaft JE, Baselga J, Taylor BS, Schultz N, Rudin CM, Hyman DM, Berger MF, Solit DB, Ladanyi M, Riely GJ. Prospective Comprehensive Molecular Characterization of Lung Adenocarcinomas for Efficient Patient Matching to Approved and Emerging Therapies. Cancer Discov. 2017;7(6):596-609. Epub 2017/03/25. doi: 10.1158/2159-8290.Cd-16-1337. PubMed PMID: 28336552; PMCID: PMC5482929.

49. Dang CV, Reddy EP, Shokat KM, Soucek L. Drugging the 'undruggable' cancer targets. Nat Rev Cancer. 2017;17(8):502-8. Epub 2017/06/24. doi: 10.1038/nrc.2017.36. PubMed PMID: 28643779; PMCID: PMC5945194.

50. B'Chir W, Maurin AC, Carraro V, Averous J, Jousse C, Muranishi Y, Parry L, Stepien G, Fafournoux P, Bruhat A. The eIF2alpha/ATF4 pathway is essential for stress-induced autophagy gene

expression. Nucleic Acids Res. 2013;41(16):7683-99. Epub 2013/06/28. doi: 10.1093/nar/gkt563. PubMed PMID: 23804767; PMCID: PMC3763548.

51. Del Vecchio CA, Feng Y, Sokol ES, Tillman EJ, Sanduja S, Reinhardt F, Gupta PB. De-differentiation confers multidrug resistance via noncanonical PERK-Nrf2 signaling. PLoS Biol. 2014;12(9):e1001945. Epub 2014/09/10. doi: 10.1371/journal.pbio.1001945. PubMed PMID: 25203443; PMCID: PMC4159113.

52. Dayalan Naidu S, Dikovskaya D, Gaurilcikaite E, Knatko EV, Healy ZR, Mohan H, Koh G, Laurell A, Ball G, Olagnier D, de la Vega L, Ganley IG, Talalay P, Dinkova-Kostova AT. Transcription factors NRF2 and HSF1 have opposing functions in autophagy. Sci Rep. 2017;7(1):11023. Epub 2017/09/10. doi: 10.1038/s41598-017-11262-5. PubMed PMID: 28887499; PMCID: PMC5591275.

53. Gwinn DM, Lee AG, Briones-Martin-Del-Campo M, Conn CS, Simpson DR, Scott AI, Le A, Cowan TM, Ruggero D, Sweet-Cordero EA. Oncogenic KRAS Regulates Amino Acid Homeostasis and Asparagine Biosynthesis via ATF4 and Alters Sensitivity to L-Asparaginase. Cancer Cell. 2018;33(1):91-107.e6. Epub 2018/01/10. doi: 10.1016/j.ccell.2017.12.003. PubMed PMID: 29316436; PMCID: PMC5761662.

54. Giaever G, Shoemaker DD, Jones TW, Liang H, Winzeler EA, Astromoff A, Davis RW. Genomic profiling of drug sensitivities via induced haploinsufficiency. Nat Genet. 1999;21(3):278-83. Epub 1999/03/18. doi: 10.1038/6791. PubMed PMID: 10080179.

55. Giaever G, Flaherty P, Kumm J, Proctor M, Nislow C, Jaramillo DF, Chu AM, Jordan MI, Arkin AP, Davis RW. Chemogenomic profiling: identifying the functional interactions of small molecules in yeast. Proc Natl Acad Sci U S A. 2004;101(3):793-8. Epub 2004/01/14. doi: 10.1073/pnas.0307490100. PubMed PMID: 14718668; PMCID: PMC321760.

56. Parsons AB, Brost RL, Ding H, Li Z, Zhang C, Sheikh B, Brown GW, Kane PM, Hughes TR, Boone C. Integration of chemical-genetic and genetic interaction data links bioactive compounds to cellular target pathways. Nat Biotechnol. 2004;22(1):62-9. Epub 2003/12/09. doi: 10.1038/nbt919. PubMed PMID: 14661025.

57. Piotrowski JS, Li SC, Deshpande R, Simpkins SW, Nelson J, Yashiroda Y, Barber JM, Safizadeh H, Wilson E, Okada H, Gebre AA, Kubo K, Torres NP, LeBlanc MA, Andrusiak K, Okamoto R, Yoshimura M, DeRango-Adem E, van Leeuwen J, Shirahige K, Baryshnikova A, Brown GW, Hirano H, Costanzo M, Andrews B, Ohya Y, Osada H, Yoshida M, Myers CL, Boone C. Functional annotation of chemical libraries across diverse biological processes. Nat Chem Biol. 2017;13(9):982-93. Epub 2017/08/02. doi: 10.1038/nchembio.2436. PubMed PMID: 28759014; PMCID: PMC6056180.

58. Parsons AB, Lopez A, Givoni IE, Williams DE, Gray CA, Porter J, Chua G, Sopko R, Brost RL, Ho CH, Wang J, Ketela T, Brenner C, Brill JA, Fernandez GE, Lorenz TC, Payne GS, Ishihara S, Ohya Y, Andrews B, Hughes TR, Frey BJ, Graham TR, Andersen RJ, Boone C. Exploring the mode-of-action of bioactive compounds by chemical-genetic profiling in yeast. Cell. 2006;126(3):611-25. Epub 2006/08/12. doi: 10.1016/j.cell.2006.06.040. PubMed PMID: 16901791.

59. Hillenmeyer ME, Fung E, Wildenhain J, Pierce SE, Hoon S, Lee W, Proctor M, St Onge RP, Tyers M, Koller D, Altman RB, Davis RW, Nislow C, Giaever G. The chemical genomic portrait of yeast: uncovering a phenotype for all genes. Science. 2008;320(5874):362-5. Epub 2008/04/19. doi: 10.1126/science.1150021. PubMed PMID: 18420932; PMCID: PMC2794835.

60. Jost M, Chen Y, Gilbert LA, Horlbeck MA, Krenning L, Menchon G, Rai A, Cho MY, Stern JJ, Prota AE, Kampmann M, Akhmanova A, Steinmetz MO, Tanenbaum ME, Weissman JS. Combined CRISPRi/a-Based Chemical Genetic Screens Reveal that Rigosertib Is a Microtubule-Destabilizing Agent. Mol Cell. 2017;68(1):210-23.e6. Epub 2017/10/07. doi: 10.1016/j.molcel.2017.09.012. PubMed PMID: 28985505; PMCID: PMC5640507.

61. Shalem O, Sanjana NE, Hartenian E, Shi X, Scott DA, Mikkelson T, Heckl D, Ebert BL, Root DE, Doench JG, Zhang F. Genome-scale CRISPR-Cas9 knockout screening in human cells. Science. 2014;343(6166):84-7. Epub 2013/12/18. doi: 10.1126/science.1247005. PubMed PMID: 24336571; PMCID: PMC4089965.

62. Wang T, Wei JJ, Sabatini DM, Lander ES. Genetic screens in human cells using the CRISPR-Cas9 system. Science. 2014;343(6166):80-4. Epub 2013/12/18. doi: 10.1126/science.1246981. PubMed PMID: 24336569; PMCID: PMC3972032.

63. Lin A, Giuliano CJ, Palladino A, John KM, Abramowicz C, Yuan ML, Sausville EL, Lukow DA, Liu L, Chait AR, Galluzzo ZC, Tucker C, Sheltzer JM. Off-target toxicity is a common mechanism of action of cancer drugs undergoing clinical trials. Sci Transl Med. 2019;11(509). Epub 2019/09/13. doi: 10.1126/scitranslmed.aaw8412. PubMed PMID: 31511426.

64. Dobzhansky T. Genetics of natural populations; recombination and variability in populations of Drosophila pseudoobscura. Genetics. 1946;31(3):269-90. Epub 1946/05/01. PubMed PMID: 20985721; PMCID: PMC1209328.

65. Bryant HE, Schultz N, Thomas HD, Parker KM, Flower D, Lopez E, Kyle S, Meuth M, Curtin NJ, Helleday T. Specific killing of BRCA2-deficient tumours with inhibitors of poly(ADP-ribose) polymerase. Nature. 2005;434(7035):913-7. Epub 2005/04/15. doi: 10.1038/nature03443. PubMed PMID: 15829966.

66. Farmer H, McCabe N, Lord CJ, Tutt AN, Johnson DA, Richardson TB, Santarosa M, Dillon KJ, Hickson I, Knights C, Martin NM, Jackson SP, Smith GC, Ashworth A. Targeting the DNA repair defect in BRCA mutant cells as a therapeutic strategy. Nature. 2005;434(7035):917-21. Epub 2005/04/15. doi: 10.1038/nature03445. PubMed PMID: 15829967.

67. Sahu AD, J SL, Wang Z, Zhang G, Iglesias-Bartolome R, Tian T, Wei Z, Miao B, Nair NU, Ponomarova O, Friedman AA, Amzallag A, Moll T, Kasumova G, Greninger P, Egan RK, Damon LJ, Frederick DT, Jerby-Arnon L, Wagner A, Cheng K, Park SG, Robinson W, Gardner K, Boland G, Hannenhalli S, Herlyn M, Benes C, Flaherty K, Luo J, Gutkind JS, Ruppin E. Genome-wide prediction of synthetic rescue mediators of resistance to targeted and immunotherapy. Mol Syst Biol. 2019;15(3):e8323. Epub 2019/03/13. doi: 10.15252/msb.20188323. PubMed PMID: 30858180; PMCID: PMC6413886.

68. Srivas R, Shen JP, Yang CC, Sun SM, Li J, Gross AM, Jensen J, Licon K, Bojorquez-Gomez A, Klepper K, Huang J, Pekin D, Xu JL, Yeerna H, Sivaganesh V, Kollenstart L, van Attikum H, Aza-

Blanc P, Sobol RW, Ideker T. A Network of Conserved Synthetic Lethal Interactions for Exploration of Precision Cancer Therapy. Mol Cell. 2016;63(3):514-25. Epub 2016/07/28. doi: 10.1016/j.molcel.2016.06.022. PubMed PMID: 27453043; PMCID: PMC5209245.

69. Guo E, Ishii Y, Mueller J, Srivatsan A, Gahman T, Putnam CD, Wang JYJ, Kolodner RD. FEN1 endonuclease as a therapeutic target for human cancers with defects in homologous recombination. Proc Natl Acad Sci U S A. 2020. Epub 2020/07/29. doi: 10.1073/pnas.2009237117. PubMed PMID: 32719125.

70. Barretina J, Caponigro G, Stransky N, Venkatesan K, Margolin AA, Kim S, Wilson CJ, Lehar J, Kryukov GV, Sonkin D, Reddy A, Liu M, Murray L, Berger MF, Monahan JE, Morais P, Meltzer J, Korejwa A, Jane-Valbuena J, Mapa FA, Thibault J, Bric-Furlong E, Raman P, Shipway A, Engels IH, Cheng J, Yu GK, Yu J, Aspesi P, Jr., de Silva M, Jagtap K, Jones MD, Wang L, Hatton C, Palescandolo E, Gupta S, Mahan S, Sougnez C, Onofrio RC, Liefeld T, MacConaill L, Winckler W, Reich M, Li N, Mesirov JP, Gabriel SB, Getz G, Ardlie K, Chan V, Myer VE, Weber BL, Porter J, Warmuth M, Finan P, Harris JL, Meyerson M, Golub TR, Morrissey MP, Sellers WR, Schlegel R, Garraway LA. The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. Nature. 2012;483(7391):603-7. Epub 2012/03/31. doi: 10.1038/nature11003. PubMed PMID: 22460905; PMCID: PMC3320027.

71. Garnett MJ, Edelman EJ, Heidorn SJ, Greenman CD, Dastur A, Lau KW, Greninger P, Thompson IR, Luo X, Soares J, Liu Q, Iorio F, Surdez D, Chen L, Milano RJ, Bignell GR, Tam AT, Davies H, Stevenson JA, Barthorpe S, Lutz SR, Kogera F, Lawrence K, McLaren-Douglas A, Mitropoulos X, Mironenko T, Thi H, Richardson L, Zhou W, Jewitt F, Zhang T, O'Brien P, Boisvert JL, Price S, Hur W, Yang W, Deng X, Butler A, Choi HG, Chang JW, Baselga J, Stamenkovic I, Engelman JA, Sharma SV, Delattre O, Saez-Rodriguez J, Gray NS, Settleman J, Futreal PA, Haber DA, Stratton MR, Ramaswamy S, McDermott U, Benes CH. Systematic identification of genomic markers of drug sensitivity in cancer cells. Nature. 2012;483(7391):570-5. Epub 2012/03/31. doi: 10.1038/nature11005. PubMed PMID: 22460902; PMCID: PMC3349233.

72. Yang W, Soares J, Greninger P, Edelman EJ, Lightfoot H, Forbes S, Bindal N, Beare D, Smith JA, Thompson IR, Ramaswamy S, Futreal PA, Haber DA, Stratton MR, Benes C, McDermott U, Garnett MJ. Genomics of Drug Sensitivity in Cancer (GDSC): a resource for therapeutic biomarker discovery in cancer cells. Nucleic Acids Res. 2013;41(Database issue):D955-61. Epub 2012/11/28. doi: 10.1093/nar/gks1111. PubMed PMID: 23180760; PMCID: PMC3531057.

73. Iorio F, Knijnenburg TA, Vis DJ, Bignell GR, Menden MP, Schubert M, Aben N, Goncalves E, Barthorpe S, Lightfoot H, Cokelaer T, Greninger P, van Dyk E, Chang H, de Silva H, Heyn H, Deng X, Egan RK, Liu Q, Mironenko T, Mitropoulos X, Richardson L, Wang J, Zhang T, Moran S, Sayols S, Soleimani M, Tamborero D, Lopez-Bigas N, Ross-Macdonald P, Esteller M, Gray NS, Haber DA, Stratton MR, Benes CH, Wessels LFA, Saez-Rodriguez J, McDermott U, Garnett MJ. A Landscape of Pharmacogenomic Interactions in Cancer. Cell. 2016;166(3):740-54. Epub 2016/07/12. doi: 10.1016/j.cell.2016.06.017. PubMed PMID: 27397505; PMCID: PMC4967469.

74. Basu A, Bodycombe NE, Cheah JH, Price EV, Liu K, Schaefer GI, Ebright RY, Stewart ML, Ito D, Wang S, Bracha AL, Liefeld T, Wawer M, Gilbert JC, Wilson AJ, Stransky N, Kryukov GV, Dancik V, Barretina J, Garraway LA, Hon CS, Munoz B, Bittker JA, Stockwell BR, Khabele D, Stern AM, Clemons PA, Shamji AF, Schreiber SL. An interactive resource to identify cancer genetic and lineage dependencies targeted by small molecules. Cell. 2013;154(5):1151-61. Epub 2013/09/03. doi: 10.1016/j.cell.2013.08.003. PubMed PMID: 23993102; PMCID: PMC3954635.

75. Seashore-Ludlow B, Rees MG, Cheah JH, Cokol M, Price EV, Coletti ME, Jones V, Bodycombe NE, Soule CK, Gould J, Alexander B, Li A, Montgomery P, Wawer MJ, Kuru N, Kotz JD, Hon CS, Munoz B, Liefeld T, Dancik V, Bittker JA, Palmer M, Bradner JE, Shamji AF, Clemons PA, Schreiber SL. Harnessing Connectivity in a Large-Scale Small-Molecule Sensitivity Dataset. Cancer Discov. 2015;5(11):1210-23. Epub 2015/10/21. doi: 10.1158/2159-8290.Cd-15-0235. PubMed PMID: 26482930; PMCID: PMC4631646.

76. Rees MG, Seashore-Ludlow B, Cheah JH, Adams DJ, Price EV, Gill S, Javaid S, Coletti ME, Jones VL, Bodycombe NE, Soule CK, Alexander B, Li A, Montgomery P, Kotz JD, Hon CS, Munoz B, Liefeld T, Dancik V, Haber DA, Clish CB, Bittker JA, Palmer M, Wagner BK, Clemons PA, Shamji

AF, Schreiber SL. Correlating chemical sensitivity and basal gene expression reveals mechanism of action. Nat Chem Biol. 2016;12(2):109-16. Epub 2015/12/15. doi: 10.1038/nchembio.1986. PubMed PMID: 26656090; PMCID: PMC4718762.

77. Yu C, Mannan AM, Yvone GM, Ross KN, Zhang YL, Marton MA, Taylor BR, Crenshaw A, Gould JZ, Tamayo P, Weir BA, Tsherniak A, Wong B, Garraway LA, Shamji AF, Palmer MA, Foley MA, Winckler W, Schreiber SL, Kung AL, Golub TR. High-throughput identification of genotype-specific cancer vulnerabilities in mixtures of barcoded tumor cell lines. Nat Biotechnol. 2016;34(4):419-23. Epub 2016/03/02. doi: 10.1038/nbt.3460. PubMed PMID: 26928769; PMCID: PMC5508574.

78. Corsello SM, Nagari RT, Spangler RD, Rossen J, Kocak M, Bryan JG, Humeidi R, Peck D, Wu X, Tang AA, Wang VM, Bender SA, Lemire E, Narayan R, Montgomery P, Ben-David U, Garvie CW, Chen Y, Rees MG, Lyons NJ, McFarland JM, Wong BT, Wang L, Dumont N, O'Hearn PJ, Stefan E, Doench JG, Harrington CN, Greulich H, Meyerson M, Vazquez F, Subramanian A, Roth JA, Bittker JA, Boehm JS, Mader CC, Tsherniak A, Golub TR. Discovering the anticancer potential of non-oncology drugs by systematic viability profiling. Nature Cancer. 2020;1(2):235-48. doi: 10.1038/s43018-019-0018-6.

79. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2020. CA Cancer J Clin. 2020;70(1):7-30. Epub 2020/01/09. doi: 10.3322/caac.21590. PubMed PMID: 31912902.

80. Winter GE, Radic B, Mayor-Ruiz C, Blomen VA, Trefzer C, Kandasamy RK, Huber KVM, Gridling M, Chen D, Klampfl T, Kralovics R, Kubicek S, Fernandez-Capetillo O, Brummelkamp TR, Superti-Furga G. The solute carrier SLC35F2 enables YM155-mediated DNA damage toxicity. Nat Chem Biol. 2014;10(9):768-73. Epub 2014/07/30. doi: 10.1038/nchembio.1590. PubMed PMID: 25064833; PMCID: PMC4913867.

81. Hsu PD, Scott DA, Weinstein JA, Ran FA, Konermann S, Agarwala V, Li Y, Fine EJ, Wu X, Shalem O, Cradick TJ, Marraffini LA, Bao G, Zhang F. DNA targeting specificity of RNA-guided Cas9 nucleases. Nat Biotechnol. 2013;31(9):827-32. Epub 2013/07/23. doi: 10.1038/nbt.2647. PubMed PMID: 23873081; PMCID: PMC3969858.

82. Morgens DW, Wainberg M, Boyle EA, Ursu O, Araya CL, Tsui CK, Haney MS, Hess GT, Han K, Jeng EE, Li A, Snyder MP, Greenleaf WJ, Kundaje A, Bassik MC. Genome-scale measurement of off-target activity using Cas9 toxicity in high-throughput screens. Nat Commun. 2017;8:15178. Epub 2017/05/06. doi: 10.1038/ncomms15178. PubMed PMID: 28474669; PMCID: PMC5424143.

83. Jiang L, Schlesinger F, Davis CA, Zhang Y, Li R, Salit M, Gingeras TR, Oliver B. Synthetic spike-in standards for RNA-seq experiments. Genome Res. 2011;21(9):1543-51. Epub 2011/08/06. doi: 10.1101/gr.121095.111. PubMed PMID: 21816910; PMCID: PMC3166838.

84. Orlando DA, Chen MW, Brown VE, Solanki S, Choi YJ, Olson ER, Fritz CC, Bradner JE, Guenther MG. Quantitative ChIP-Seq normalization reveals global modulation of the epigenome. Cell Rep. 2014;9(3):1163-70. Epub 2014/12/02. doi: 10.1016/j.celrep.2014.10.018. PubMed PMID: 25437568.

85. Grzybowski AT, Chen Z, Ruthenburg AJ. Calibrating ChIP-Seq with Nucleosomal Internal Standards to Measure Histone Modification Density Genome Wide. Mol Cell. 2015;58(5):886-99. Epub 2015/05/26. doi: 10.1016/j.molcel.2015.04.022. PubMed PMID: 26004229; PMCID: PMC4458216.

86. Chen Y, Chen J, Loo A, Jaeger S, Bagdasarian L, Yu J, Chung F, Korn J, Ruddy D, Guo R, McLaughlin ME, Feng F, Zhu P, Stegmeier F, Pagliarini R, Porter D, Zhou W. Targeting HSF1 sensitizes cancer cells to HSP90 inhibition. Oncotarget. 2013;4(6):816-29. Epub 2013/04/26. doi: 10.18632/oncotarget.991. PubMed PMID: 23615731; PMCID: PMC3757240.

87. Acosta-Alvear D, Cho MY, Wild T, Buchholz TJ, Lerner AG, Simakova O, Hahn J, Korde N, Landgren O, Maric I, Choudhary C, Walter P, Weissman JS, Kampmann M. Paradoxical resistance of multiple myeloma to proteasome inhibitors by decreased levels of 19S proteasomal subunits. Elife. 2015;4:e08153. Epub 2015/09/04. doi: 10.7554/eLife.08153. PubMed PMID: 26327694; PMCID: PMC4602331.

88. Joutsen J, Da Silva AJ, Luoto JC, Budzynski MA, Nylund AS, de Thonel A, Concordet JP, Mezger V, Saberan-Djoneidi D, Henriksson E, Sistonen L. Heat Shock Factor 2 Protects against

Proteotoxicity by Maintaining Cell-Cell Adhesion. Cell Rep. 2020;30(2):583-97.e6. Epub 2020/01/16. doi: 10.1016/j.celrep.2019.12.037. PubMed PMID: 31940498.

89. Ye P, Mimura J, Okada T, Sato H, Liu T, Maruyama A, Ohyama C, Itoh K. Nrf2- and ATF4-dependent upregulation of xCT modulates the sensitivity of T24 bladder carcinoma cells to proteasome inhibition. Mol Cell Biol. 2014;34(18):3421-34. Epub 2014/07/09. doi: 10.1128/mcb.00221-14. PubMed PMID: 25002527; PMCID: PMC4135628.

90. Riz I, Hawley TS, Marsal JW, Hawley RG. Noncanonical SQSTM1/p62-Nrf2 pathway activation mediates proteasome inhibitor resistance in multiple myeloma cells via redox, metabolic and translational reprogramming. Oncotarget. 2016;7(41):66360-85. Epub 2016/09/15. doi: 10.18632/oncotarget.11960. PubMed PMID: 27626179; PMCID: PMC5340085.

91. Martin D, Li Y, Yang J, Wang G, Margariti A, Jiang Z, Yu H, Zampetaki A, Hu Y, Xu Q, Zeng L. Unspliced X-box-binding protein 1 (XBP1) protects endothelial cells from oxidative stress through interaction with histone deacetylase 3. J Biol Chem. 2014;289(44):30625-34. Epub 2014/09/06. doi: 10.1074/jbc.M114.571984. PubMed PMID: 25190803; PMCID: PMC4215241.

92. Chen C, Zhong Y, Wang JJ, Yu Q, Plafker K, Plafker S, Zhang SX. Regulation of Nrf2 by X Box-Binding Protein 1 in Retinal Pigment Epithelium. Front Genet. 2018;9:658. Epub 2019/01/09. doi: 10.3389/fgene.2018.00658. PubMed PMID: 30619478; PMCID: PMC6306429.

93. Santopolo S, Riccio A, Rossi A, Santoro MG. The proteostasis guardian HSF1 directs the transcription of its paralog and interactor HSF2 during proteasome dysfunction. Cell Mol Life Sci. 2020. Epub 2020/07/02. doi: 10.1007/s00018-020-03568-x. PubMed PMID: 32607595.

94. Tolcher AW, Mita A, Lewis LD, Garrett CR, Till E, Daud AI, Patnaik A, Papadopoulos K, Takimoto C, Bartels P, Keating A, Antonia S. Phase I and pharmacokinetic study of YM155, a small-molecule inhibitor of survivin. J Clin Oncol. 2008;26(32):5198-203. Epub 2008/10/01. doi: 10.1200/jco.2008.17.2064. PubMed PMID: 18824702; PMCID: PMC4879696.

95. Giaccone G, Zatloukal P, Roubec J, Floor K, Musil J, Kuta M, van Klaveren RJ, Chaudhary S, Gunther A, Shamsili S. Multicenter phase II trial of YM155, a small-molecule suppressor of survivin,

in patients with advanced, refractory, non-small-cell lung cancer. J Clin Oncol. 2009;27(27):4481-6. Epub 2009/08/19. doi: 10.1200/jco.2008.21.1862. PubMed PMID: 19687333.

96. Satoh T, Okamoto I, Miyazaki M, Morinaga R, Tsuya A, Hasegawa Y, Terashima M, Ueda S, Fukuoka M, Ariyoshi Y, Saito T, Masuda N, Watanabe H, Taguchi T, Kakihara T, Aoyama Y, Hashimoto Y, Nakagawa K. Phase I study of YM155, a novel survivin suppressant, in patients with advanced solid tumors. Clin Cancer Res. 2009;15(11):3872-80. Epub 2009/05/28. doi: 10.1158/1078-0432.Ccr-08-1946. PubMed PMID: 19470738.

97. Lewis KD, Samlowski W, Ward J, Catlett J, Cranmer L, Kirkwood J, Lawson D, Whitman E, Gonzalez R. A multi-center phase II evaluation of the small molecule survivin suppressor YM155 in patients with unresectable stage III or IV melanoma. Invest New Drugs. 2011;29(1):161-6. Epub 2009/10/16. doi: 10.1007/s10637-009-9333-6. PubMed PMID: 19830389.

98. Cheson BD, Bartlett NL, Vose JM, Lopez-Hernandez A, Seiz AL, Keating AT, Shamsili S, Papadopoulos KP. A phase II study of the survivin suppressant YM155 in patients with refractory diffuse large B-cell lymphoma. Cancer. 2012;118(12):3128-34. Epub 2011/10/19. doi: 10.1002/cncr.26510. PubMed PMID: 22006123.

99. Tolcher AW, Quinn DI, Ferrari A, Ahmann F, Giaccone G, Drake T, Keating A, de Bono JS. A phase II study of YM155, a novel small-molecule suppressor of survivin, in castration-resistant taxane-pretreated prostate cancer. Ann Oncol. 2012;23(4):968-73. Epub 2011/08/24. doi: 10.1093/annonc/mdr353. PubMed PMID: 21859898.

100. Kelly RJ, Thomas A, Rajan A, Chun G, Lopez-Chavez A, Szabo E, Spencer S, Carter CA, Guha U, Khozin S, Poondru S, Van Sant C, Keating A, Steinberg SM, Figg W, Giaccone G. A phase I/II study of sepantronium bromide (YM155, survivin suppressor) with paclitaxel and carboplatin in patients with advanced non-small-cell lung cancer. Ann Oncol. 2013;24(10):2601-6. Epub 2013/07/17. doi: 10.1093/annonc/mdt249. PubMed PMID: 23857959; PMCID: PMC3784336.

101. Clemens MR, Gladkov OA, Gartner E, Vladimirov V, Crown J, Steinberg J, Jie F, Keating A. Phase II, multicenter, open-label, randomized study of YM155 plus docetaxel as first-line treatment in patients with HER2-negative metastatic breast cancer. Breast Cancer Res Treat. 2015;149(1):171-

9. Epub 2014/12/31. doi: 10.1007/s10549-014-3238-6. PubMed PMID: 25547219; PMCID: PMC4298663.

102. Kudchadkar R, Ernst S, Chmielowski B, Redman BG, Steinberg J, Keating A, Jie F, Chen C, Gonzalez R, Weber J. A phase 2, multicenter, open-label study of sepantronium bromide (YM155) plus docetaxel in patients with stage III (unresectable) or stage IV melanoma. Cancer Med. 2015;4(5):643-50. Epub 2014/12/24. doi: 10.1002/cam4.363. PubMed PMID: 25533314; PMCID: PMC4430257.

103. Papadopoulos KP, Lopez-Jimenez J, Smith SE, Steinberg J, Keating A, Sasse C, Jie F, Thyss A. A multicenter phase II study of sepantronium bromide (YM155) plus rituximab in patients with relapsed aggressive B-cell Non-Hodgkin lymphoma. Leuk Lymphoma. 2016;57(8):1848-55. Epub 2016/02/10. doi: 10.3109/10428194.2015.1113275. PubMed PMID: 26857688.

104. Li F, Aljahdali I, Ling X. Cancer therapeutics using survivin BIRC5 as a target: what can we do after over two decades of study? J Exp Clin Cancer Res. 2019;38(1):368. Epub 2019/08/24. doi: 10.1186/s13046-019-1362-1. PubMed PMID: 31439015; PMCID: PMC6704566.

105. Nakahara T, Kita A, Yamanaka K, Mori M, Amino N, Takeuchi M, Tominaga F, Hatakeyama S, Kinoyama I, Matsuhisa A, Kudoh M, Sasamata M. YM155, a novel small-molecule survivin suppressant, induces regression of established human hormone-refractory prostate tumor xenografts. Cancer Res. 2007;67(17):8014-21. Epub 2007/09/07. doi: 10.1158/0008-5472.Can-07-1343. PubMed PMID: 17804712.

106. Salani B, Marini C, Rio AD, Ravera S, Massollo M, Orengo AM, Amaro A, Passalacqua M, Maffioli S, Pfeffer U, Cordera R, Maggi D, Sambuceti G. Metformin impairs glucose consumption and survival in Calu-1 cells by direct inhibition of hexokinase-II. Sci Rep. 2013;3:2070. Epub 2013/06/26. doi: 10.1038/srep02070. PubMed PMID: 23797762; PMCID: PMC3691576.

107. Harris EA, Koh EJ, Moffat J, McMillen DR. Automated inference procedure for the determination of cell growth parameters. Phys Rev E. 2016;93(1):012402. Epub 2016/02/13. doi: 10.1103/PhysRevE.93.012402. PubMed PMID: 26871096.

108. Niepel M, Hafner M, Mills CE, Subramanian K, Williams EH, Chung M, Gaudio B, Barrette AM, Stern AD, Hu B, Korkola JE, Gray JW, Birtwistle MR, Heiser LM, Sorger PK. A Multi-center Study on the Reproducibility of Drug-Response Assays in Mammalian Cell Lines. Cell Syst. 2019;9(1):35-48.e5. Epub 2019/07/16. doi: 10.1016/j.cels.2019.06.005. PubMed PMID: 31302153; PMCID: PMC6700527.

109. Alexandrov LB, Nik-Zainal S, Wedge DC, Aparicio SA, Behjati S, Biankin AV, Bignell GR, Bolli N, Borg A, Borresen-Dale AL, Boyault S, Burkhardt B, Butler AP, Caldas C, Davies HR, Desmedt C, Eils R, Eyfjord JE, Foekens JA, Greaves M, Hosoda F, Hutter B, Ilicic T, Imbeaud S, Imielinski M, Jager N, Jones DT, Jones D, Knappskog S, Kool M, Lakhani SR, Lopez-Otin C, Martin S, Munshi NC, Nakamura H, Northcott PA, Pajic M, Papaemmanuil E, Paradiso A, Pearson JV, Puente XS, Raine K, Ramakrishna M, Richardson AL, Richter J, Rosenstiel P, Schlesner M, Schumacher TN, Span PN, Teague JW, Totoki Y, Tutt AN, Valdes-Mas R, van Buuren MM, van 't Veer L, Vincent-Salomon A, Waddell N, Yates LR, Zucman-Rossi J, Futreal PA, McDermott U, Lichter P, Meyerson M, Grimmond SM, Siebert R, Campo E, Shibata T, Pfister SM, Campbell PJ, Stratton MR. Signatures of mutational processes in human cancer. Nature. 2013;500(7463):415-21. Epub 2013/08/16. doi: 10.1038/nature12477. PubMed PMID: 23945592; PMCID: PMC3776390.

110. Weinstein JN, Collisson EA, Mills GB, Shaw KR, Ozenberger BA, Ellrott K, Shmulevich I, Sander C, Stuart JM. The Cancer Genome Atlas Pan-Cancer analysis project. Nat Genet. 2013;45(10):1113-20. Epub 2013/09/28. doi: 10.1038/ng.2764. PubMed PMID: 24071849; PMCID: PMC3919969.

111. Maji B, Moore CL, Zetsche B, Volz SE, Zhang F, Shoulders MD, Choudhary A. Multidimensional chemical control of CRISPR-Cas9. Nat Chem Biol. 2017;13(1):9-11. Epub 2016/11/08. doi: 10.1038/nchembio.2224. PubMed PMID: 27820801; PMCID: PMC5531067.

112. Gruner BM, Schulze CJ, Yang D, Ogasawara D, Dix MM, Rogers ZN, Chuang CH, McFarland CD, Chiou SH, Brown JM, Cravatt BF, Bogyo M, Winslow MM. An in vivo multiplexed small-molecule screening platform. Nat Methods. 2016;13(10):883-9. Epub 2016/09/13. doi: 10.1038/nmeth.3992. PubMed PMID: 27617390; PMCID: PMC5088491.

113. Rogers ZN, McFarland CD, Winters IP, Naranjo S, Chuang CH, Petrov D, Winslow MM. A quantitative and multiplexed approach to uncover the fitness landscape of tumor suppression in vivo. Nat Methods. 2017;14(7):737-42. Epub 2017/05/23. doi: 10.1038/nmeth.4297. PubMed PMID: 28530655; PMCID: PMC5495136.

114. Lamb J, Crawford ED, Peck D, Modell JW, Blat IC, Wrobel MJ, Lerner J, Brunet JP, Subramanian A, Ross KN, Reich M, Hieronymus H, Wei G, Armstrong SA, Haggarty SJ, Clemons PA, Wei R, Carr SA, Lander ES, Golub TR. The Connectivity Map: using gene-expression signatures to connect small molecules, genes, and disease. Science. 2006;313(5795):1929-35. Epub 2006/09/30. doi: 10.1126/science.1132939. PubMed PMID: 17008526.

115. Subramanian A, Narayan R, Corsello SM, Peck DD, Natoli TE, Lu X, Gould J, Davis JF, Tubelli AA, Asiedu JK, Lahr DL, Hirschman JE, Liu Z, Donahue M, Julian B, Khan M, Wadden D, Smith IC, Lam D, Liberzon A, Toder C, Bagul M, Orzechowski M, Enache OM, Piccioni F, Johnson SA, Lyons NJ, Berger AH, Shamji AF, Brooks AN, Vrcic A, Flynn C, Rosains J, Takeda DY, Hu R, Davison D, Lamb J, Ardlie K, Hogstrom L, Greenside P, Gray NS, Clemons PA, Silver S, Wu X, Zhao WN, Read-Button W, Wu X, Haggarty SJ, Ronco LV, Boehm JS, Schreiber SL, Doench JG, Bittker JA, Root DE, Wong B, Golub TR. A Next Generation Connectivity Map: L1000 Platform and the First 1,000,000 Profiles. Cell. 2017;171(6):1437-52.e17. Epub 2017/12/02. doi: 10.1016/j.cell.2017.10.049. PubMed PMID: 29195078; PMCID: PMC5990023.

116. Ye C, Ho DJ, Neri M, Yang C, Kulkarni T, Randhawa R, Henault M, Mostacci N, Farmer P, Renner S, Ihry R, Mansur L, Keller CG, McAllister G, Hild M, Jenkins J, Kaykas A. DRUG-seq for miniaturized high-throughput transcriptome profiling in drug discovery. Nat Commun. 2018;9(1):4307. Epub 2018/10/20. doi: 10.1038/s41467-018-06500-x. PubMed PMID: 30333485; PMCID: PMC6192987.

117. McFarland JM, Paolella BR, Warren A, Geiger-Schuller K, Shibue T, Rothberg M, Kuksenko O, Colgan WN, Jones A, Chambers E, Dionne D, Bender S, Wolpin BM, Ghandi M, Tirosh I, Rozenblatt-Rosen O, Roth JA, Golub TR, Regev A, Aguirre AJ, Vazquez F, Tsherniak A. Multiplexed single-cell transcriptional response profiling to define cancer vulnerabilities and

therapeutic mechanism of action. Nat Commun. 2020;11(1):4296. Epub 2020/08/29. doi: 10.1038/s41467-020-17440-w. PubMed PMID: 32855387; PMCID: PMC7453022 Immunitas, and an SAB Member of ThermoFisher Scientific, Neogene Therapeutics, Asimov, and Syros Pharmaceuticals. A.J.A. has consulted for Oncorus, Inc., Arrakis Therapeutics, and Merck & Co., Inc. and has research funding from Mirati Therapeutics and Deerfield Management that are unrelated to this work. A.T. is a consultant for Tango Therapeutics. T.R.G. is a consultant to GlaxoSmithKline, a founder of Sherlock Biosciences, and was formerly a consultant and equity holder in Foundation Medicine, acquired by Roche. T.R.G. also receives research funding unrelated to this project from Bayer Healthcare. B.M.W. receives research funding unrelated to this project from Celgene and Lilly and is a consultant for BioLineRx, Celgene, G1 Therapeutics, and GRAIL. The Broad Institute filed a United States Provisional Patent Application directed to work described in this article. The pending Application is entitled “Rapid Prediction of Drug Responsiveness” and was filed on September 2, 2019. The current inventor list includes A.T., A.J.A., F.V., B.R.P., and J.M.M. A.W., K.G.-S., T.S., M.R., O.K., W.N.C., A.J., E.C., D.D., S.B., M.G., I.T., O.R.-R., and J.A.R. declare no competing interests.

118. Srivatsan SR, McFaline-Figueroa JL, Ramani V, Saunders L, Cao J, Packer J, Pliner HA, Jackson DL, Daza RM, Christiansen L, Zhang F, Steemers F, Shendure J, Trapnell C. Massively multiplex chemical transcriptomics at single-cell resolution. Science. 2020;367(6473):45-51. Epub 2019/12/07. doi: 10.1126/science.aax6234. PubMed PMID: 31806696.

119. Ran FA, Hsu PD, Wright J, Agarwala V, Scott DA, Zhang F. Genome engineering using the CRISPR-Cas9 system. Nat Protoc. 2013;8(11):2281-308. Epub 2013/10/26. doi: 10.1038/nprot.2013.143. PubMed PMID: 24157548; PMCID: PMC3969860.

120. Sanjana NE, Shalem O, Zhang F. Improved vectors and genome-wide libraries for CRISPR screening. Nat Methods. 2014;11(8):783-4. Epub 2014/07/31. doi: 10.1038/nmeth.3047. PubMed PMID: 25075903; PMCID: PMC4486245.

121. Doench JG, Fusi N, Sullender M, Hegde M, Vaimberg EW, Donovan KF, Smith I, Tothova Z, Wilen C, Orchard R, Virgin HW, Listgarten J, Root DE. Optimized sgRNA design to maximize activity and

minimize off-target effects of CRISPR-Cas9. Nat Biotechnol. 2016;34(2):184-91. Epub 2016/01/19. doi: 10.1038/nbt.3437. PubMed PMID: 26780180; PMCID: PMC4744125.

122. Chakravarty D, Gao J, Phillips SM, Kundra R, Zhang H, Wang J, Rudolph JE, Yaeger R, Soumerai T, Nissan MH, Chang MT, Chandarlapaty S, Traina TA, Paik PK, Ho AL, Hantash FM, Grupe A, Baxi SS, Callahan MK, Snyder A, Chi P, Danila D, Gounder M, Harding JJ, Hellmann MD, Iyer G, Janjigian Y, Kaley T, Levine DA, Lowery M, Omuro A, Postow MA, Rathkopf D, Shoushtari AN, Shukla N, Voss M, Paraiso E, Zehir A, Berger MF, Taylor BS, Saltz LB, Riely GJ, Ladanyi M, Hyman DM, Baselga J, Sabbatini P, Solit DB, Schultz N. OncoKB: A Precision Oncology Knowledge Base. JCO Precis Oncol. 2017;2017. Epub 2017/09/12. doi: 10.1200/po.17.00011. PubMed PMID: 28890946; PMCID: PMC5586540.

123. Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, Jacobsen A, Byrne CJ, Heuer ML, Larsson E, Antipin Y, Reva B, Goldberg AP, Sander C, Schultz N. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. Cancer Discov. 2012;2(5):401-4. Epub 2012/05/17. doi: 10.1158/2159-8290.Cd-12-0095. PubMed PMID: 22588877; PMCID: PMC3956037.

124. Gao J, Aksoy BA, Dogrusoz U, Dresdner G, Gross B, Sumer SO, Sun Y, Jacobsen A, Sinha R, Larsson E, Cerami E, Sander C, Schultz N. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. Sci Signal. 2013;6(269):pl1. Epub 2013/04/04. doi: 10.1126/scisignal.2004088. PubMed PMID: 23550210; PMCID: PMC4160307.

125. Brockway S, Wang G, Jackson JM, Amici DR, Takagishi SR, Clutter MR, Bartom ET, Mendillo ML. QMAP-Seq: Quantitative and Multiplexed Analysis of Phenotype by Sequencing. Protoc Exch. 2020. doi: https://doi.org/10.21203/rs.3.pex-1196/v1.

126. Brockway S, Wang G, Jackson JM, Amici DR, Takagishi SR, Clutter MR, Bartom ET, Mendillo ML. QMAP-Seq code. GitHub2020. doi: https://doi.org/10.5281/zenodo.4067234.

127. Brockway S, Wang G, Jackson JM, Amici DR, Takagishi SR, Clutter MR, Bartom ET, Mendillo ML. QMAP-Seq compute capsule. Code Ocean2020. doi: https://doi.org/10.24433/CO.6760903.v1.

VITA

Sonia Olikara Brockway

Education

Northwestern University, Chicago, IL
Ph.D., Cancer Biology — Expected December 2020

Northwestern University, Kellogg School of Management, Evanston, IL
Pharmaceutical Strategy Course — April 2018–June 2018
Management for Scientists and Engineers Certificate — August 2014

Loyola University Chicago, Maywood, IL
M.S., Molecular Biology — December 2011

DePauw University, Greencastle, IN
B.A., Biochemistry and Spanish double major — May 2004

Universidad de Salamanca, Salamanca, Spain
Study Abroad Student — September 2002–December 2002

Research Experience

Northwestern University, Chicago, IL
Graduate Student — March 2016–Present
Advisor: Marc Mendillo, Ph.D.
Dissertation: *Quantitative and Multiplexed Chemical-Genetic Phenotyping Reveals the Architecture of the Proteostasis Network in Breast Cancer*

Northwestern University, Chicago, IL
Graduate Student — July 2012–March 2015
Advisor: Marcus Peter, Ph.D.
Project: *Investigating the role of the Fas receptor and ligand in cancer cell survival and nonapoptotic signaling*

Loyola University Chicago, Maywood, IL
Graduate Student — August 2009–August 2011
Advisor: Nancy Zeleznik-Le, Ph.D.
Thesis: *WEE1 is a biological target of the miR-17-92 cluster in leukemia*

University of Chicago, Chicago, IL
Research Technologist — July 2005–July 2009
Advisor: Steve Goldstein, M.D., Ph.D.
Projects: *One SUMO is sufficient to silence the dimeric potassium channel K2P1* and *Sumoylation of Kv2.1 potassium channels controls excitability of hippocampal neurons*

DePauw University, Greencastle, IN
Undergraduate Research Assistant — May 2003–December 2003
Advisor: Chester Fornari, Ph.D.
Project: *Discovering, sequencing, and analyzing new genes from rotifers by a rapid sequencing method and the BLAST algorithm*

DePauw University, Greencastle, IN
Undergraduate Research Assistant — May 2001–August 2001
Advisors: Jacqueline Roberts, Ph.D. and David Roberts, Ph.D.
Projects: *Crystallization studies of gamma-glutamylcysteine synthetase from E. coli* and *Towards the structure determination of long-chain acyl-CoA dehydrogenase: expression and initial purification of the cloned enzyme*

Peer-Reviewed Publications

Brockway S, Wang G, Jackson JM, Amici DR, Takagishi SR, Clutter MR, Bartom ET, Mendillo ML. Quantitative and multiplexed chemical-genetic phenotyping in mammalian cells with QMAP-Seq. *Nat Commun.* 2020 Nov 12;11(1):5722.

Murmann AE, McMahon KM, Haluck-Kangas A, Ravindran N, Patel M, Law CY, **Brockway S**, Wei JJ, Thaxton CS, Peter ME. Induction of DISE in ovarian cancer cells in vivo. *Oncotarget*. 2017 Oct 4;8(49):84643-58.

Qadir AS, Ceppi P, **Brockway S**, Law C, Mu L, Khodarev NN, Kim J, Zhao JC, Putzbach W, Murmann AE, Chen Z, Chen W, Liu X, Salomon AR, Liu H, Weichselbaum RR, Yu J, Peter ME. CD95/Fas increases stemness in cancer cells by inducing a STAT1-dependent type I interferon response. *Cell Rep*. 2017 Mar 7;18(10):2373-86.

Brockway S, Zeleznik-Le NJ. WEE1 is a validated target of the microRNA miR-17-92 cluster in leukemia. *Cancer Genet*. 2015 May;208(5):279-87.

Peter ME, Hadji A, Murmann AE, **Brockway S**, Putzbach W, Pattanayak A, Ceppi P. The role of CD95 and CD95 ligand in cancer. *Cell Death Differ*. 2015 Apr;22(4):549-59.

Hadji A, Ceppi P, Murmann AE, **Brockway S**, Pattanayak A, Bhinder B, Hau A, DeChant S, Parimi V, Kolesza P, Richards J, Chandel N, Djaballah H, Peter ME. Death induced by CD95 or CD95 ligand elimination. *Cell Rep*. 2014 Apr 10;7(1):208-22.

Plant LD, Dementieva IS, Kollewe A, **Olikara S**, Marks JD, Goldstein SA. One SUMO is sufficient to silence the dimeric potassium channel K2P1. *Proc Natl Acad Sci U S A*. 2010 Jun 8;107(23):10743-8.

Oral Presentations

Brockway S. Systematic chemical-genetic profiling reveals the architecture of the proteostasis network in breast cancer. Northwestern University Biochemistry and Molecular Genetics Retreat (Rapid Fire Talk), Chicago, IL, October 2019.

Brockway S. Systematic chemical-genetic profiling reveals the architecture of the proteostasis network in breast cancer. Gordon Research Conference on Stress Proteins in Growth, Development and Disease (Poster Preview Talk), Barga, Italy, June 2019.

Brockway S. Systematic chemical-genetic profiling reveals the architecture of the proteostasis network in breast cancer. Gordon Research Seminar on Stress Proteins in Growth, Development and Disease, Barga, Italy, June 2019.

Brockway S. Systematic chemical-genetic profiling reveals the architecture of the proteostasis network in breast cancer. Northwestern University Simpson Querrey Center for Epigenetics Forum on Biochemistry, Epigenetics, and Metabolism, Chicago, IL, October 2018.

Brockway S. Systematic characterization of the protein homeostasis network in breast cancer therapeutic response. Northwestern University Biochemistry and Molecular Genetics Data Group, Chicago, IL, April 2017.

Olikara S. The regulation of WEE1 by the miR-17-92 cluster in MLL fusion leukemias. Loyola University Chicago Molecular Biology Seminar, Maywood, IL, May 2011.

Poster Presentations

Brockway S, Wang G, Jackson, J, Takagishi S, Bartom E, Mendillo ML. Systematic chemical-genetic profiling reveals the architecture of the proteostasis network in breast cancer. Northwestern University Biochemistry and Molecular Genetics Retreat, Chicago, IL, October 2019.

Brockway S, Wang G, Takagishi S, Bartom E, Mendillo ML. Systematic chemical-genetic profiling reveals the architecture of the proteostasis network in breast cancer. Gordon Research Conference on Stress Proteins in Growth, Development and Disease, Barga, Italy, June 2019.

Brockway S, Wang G, Takagishi S, Bartom E, Mendillo ML. Systematic chemical-genetic profiling reveals the architecture of the proteostasis network in breast cancer. Gordon Research Seminar on Stress Proteins in Growth, Development and Disease, Barga, Italy, June 2019.

Brockway S, Wang G, Takagishi S, Bartom E, Mendillo ML. Systematic chemical-genetic profiling reveals the architecture of the proteostasis network in breast cancer. Midwest Stress Response and Molecular Chaperone Meeting, Evanston, IL, January 2019.

Brockway S, Wang G, Takagishi S, Mendillo ML. Systematic chemical-genetic profiling reveals the architecture of the proteostasis network in breast cancer. Northwestern University Biochemistry and Molecular Genetics Retreat, Evanston, IL, September 2018.

Brockway S, Wang G, Takagishi S, Mendillo ML. Systematic characterization of the proteostasis network in breast cancer therapeutic response. Midwest Stress Response and Molecular Chaperone Meeting, Evanston, IL, January 2018.

Brockway S, Wang G, Takagishi S, Mendillo ML. Systematic characterization of the protein homeostasis network in breast cancer therapeutic response. Northwestern University Biochemistry and Molecular Genetics Retreat, Evanston, IL, November 2017.

Brockway S, Wang G, Takagishi S, Mendillo ML. The protein homeostasis network in breast cancer therapeutic response. Northwestern University Annual Lurie Cancer Center Symposium and Scientific Poster Session, Chicago, IL, June 2017.

Brockway S, Wang G, Takagishi S, Mendillo ML. The protein homeostasis network in breast cancer therapeutic response. Northwestern University Lewis Landsberg Research Day, Chicago, IL, April 2017.

Brockway S, Wang G, Takagishi S, Mendillo ML. Systematic assessment of the role of the protein homeostasis network in modulating breast cancer therapeutic response. Midwest Stress Response and Molecular Chaperone Meeting, Evanston, IL, January 2017.

Brockway S, Wang G, Takagishi S, Mendillo ML. Systematic assessment of the role of the cellular stress response network in modulating breast cancer therapeutic response. Northwestern University Biochemistry and Molecular Genetics Retreat, Wilmette, IL, October 2016.

Brockway S, Hadji A, Ceppi P, Murmann AE, Pattanayak A, Bhinder B, Hau A, DeChant S, Parimi V, Kolesza P, Richards J, Chandel N, Djaballah H, Peter ME. Death induced by CD95 or CD95 ligand elimination (DICE). Northwestern University Nuclear Architecture and Chromatin Dynamics in Health and Disease Symposium, Chicago, IL, July 2014.

Brockway S, Hadji A, Ceppi P, Murmann AE, Djaballah H, Peter ME. Using RNA interference to kill cancer cells by eliminating CD95 or CD95 ligand. Keystone Symposia on RNA Silencing, Seattle, WA, February 2014.

Olikara S, Zeleznik-Le NJ. The regulation of WEE1 by the miR-17-92 cluster in MLL fusion leukemias. Loyola University Chicago St. Albert's Research Day, Maywood, IL, November 2010.

Olikara S, Zeleznik-Le NJ. The regulation of WEE1 by the miR-17-92 cluster in MLL fusion leukemias. Loyola University Chicago Molecular Biology Retreat, Brookfield, IL, September 2010.

Olikara S, Bales B, Evans P, Fornari C. Discovering, sequencing, and analyzing new genes from rotifers by a rapid sequencing method and the BLAST algorithm. National Conference on Undergraduate Research, Indianapolis, IN, April 2004.

Olikara S, Faulkner L, Huffer A, Abbott M, Moore D, Young C, Roberts D, Roberts J. Crystallization studies of gamma-glutamylcysteine synthetase from E. coli. American Chemical Society Meeting, Indianapolis, IN, October 2001.

Abbott M, Moore D, Young C, **Olikara S**, Faulkner L, Huffer A, Roberts D, Roberts J. Towards the structure determination of long-chain acyl-CoA dehydrogenase: expression and initial purification of the cloned enzyme. American Chemical Society Meeting, Indianapolis, IN, October 2001.

Grants, Fellowships, and Awards

Northwestern University, Chicago, IL

The Graduate School Conference Travel Grant	April 2019
Katten Muchin Rosenman Travel Scholarship	March 2019
Driskill Graduate Program in Life Sciences Travel Grant	February 2019
Biochemistry and Molecular Genetics Retreat Best Poster Presentation	September 2018
NIH/NCI Carcinogenesis Training Grant T32 CA 09560-28	September 2014–August 2015
NIH/NCI Carcinogenesis Training Grant T32 CA 09560-27	September 2013–August 2014
Integrated Graduate Program in the Life Sciences Fellowship	September 2011–March 2013

DePauw University, Greencastle, IN

Sigma Delta Pi National Collegiate Spanish Language Honor Society	April 2003–May 2004
Condit Science Award	August 2002–May 2004
Undergraduate Science Research Fellowship	May 2001–May 2004
Donald J. Cook Excellence in Chemistry Award	May 2001–May 2004
Academic Merit Award	August 2000–May 2004
Honor Scholar Grant	August 2000–May 2004
Multicultural Student Leadership Award	August 2000–May 2004
Science and Mathematics Scholarship	August 2000–May 2002

Teaching Experience

Northwestern University, Chicago, IL

Graduate Teaching Assistant, Molecular Mechanisms of Carcinogenesis	January 2014–March 2014

DePauw University, Greencastle, IN

Undergraduate Teaching Assistant, Molecular Biology	August 2003–December 2003