Identification of Genetic Vulnerabilities in Glioblastoma and Other Cancers Using CRISPR-Cas9-based Functional Genomic Approaches

Pia F. Hoellerbauer

A dissertation

submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

University of Washington

2020

Reading Committee:

Patrick J. Paddison, Chair

Jonathan A. Cooper

Eric C. Holland

Program Authorized to Offer Degree:

Molecular and Cellular Biology

University of Washington

Abstract

Identification of Genetic Vulnerabilities in Glioblastoma and Other Cancers Using CRISPR-Cas9-based Functional Genomic Approaches

Pia F. Hoellerbauer

Chair of the Supervisory Committee:

Patrick J. Paddison

Department of Genome Sciences

Glioblastoma (GBM) is the most common and aggressive form of brain cancer. It remains one of the deadliest cancers, with 90% of adult patients dying within two years of diagnosis with standard-of-care treatments. Although directly targeting oncogenic driver mutations is a popular strategy in developing new therapeutics, it has been met with limited success. Thus, we take the alternative, unbiased approach of using functional genomic screens to identify molecular vulnerabilities that arise in cancer cells due to the oncogenic state. In order to identify genes required for the proliferation and survival of GBM cells, the Paddison Lab previously performed lentiviral genome-scale pooled RNAi and CRISPR-Cas9 outgrowth screens in patient-derived GBM stem-like cells (GSCs) and human nontransformed neural stem cells (NSCs).

Retesting screen hits using single CRISPR sgRNAs can prove challenging with lentivirus due to protracted windows of indel formation (editing over the course of ~5-12 days), resulting in phenotypically mixed populations. Here, to address this challenge, I optimized a method for direct nucleofection of ribonucleoprotein complexes (RNPs) composed of chemically synthesized 2'-O-methyl 3'phosphorothioate-modified sgRNA and purified Cas9 protein. With this technique, we can routinely achieve >90% indel formation as well as targeted genomic deletions in only 3 days, even in hyperdiploid cells (such as GSCs), with no need to create clonal lines for simple loss-of-function experiments.

Additionally in this body of work, to further characterize candidate hits specific to GSCs (and not NSCs), we performed comprehensive pooled outgrowth retests of all putative screen hits that scored preferentially in GSC isolates in our whole-genome CRISPR-Cas9 screens. We then used these results to identify both GBM gene dependency groups and context-specific vulnerabilities, which was facilitated by incorporating published functional genomic screening data (DepMap database) into the analyses. By creating co-dependency networks, we identified GSC-specific gene vulnerability groups related to mitochondrial protein processing and turnover and membrane trafficking as well as metabolic enzymes and regulators. Furthermore, we also identified predicted context-specific vulnerabilities and further investigated the dsRNA-editing enzyme ADAR, which is required in cancer cells that have an interferon-stimulated gene expression signature, and the adapter protein EFR3A, which is required in cancer cells that have low expression of its paralog EFR3B. In addition, we investigated the particularly strongly GSC-specific screen hit FBXO42. We found this F-box protein to be essential in a subset of GSCs *in vitro* and *in vivo* but nonessential in NSCs *in vitro*, in addition to being required in subsets of cells of various other cancer types, suggesting the potential for both a large therapeutic window and

broad applicability. Furthermore, we demonstrated that the ubiquitin ligase role of FBXO42 is responsible for the viability phenotype, but likely not through p53 as a previous study might suggest. In searching for possible interactors, we determined that *FBXO42* and the gene *CCDC6* (Coiled-Coil Domain Containing 6, a common translocation partner for receptor kinase oncogenes) are both necessary to promote viability in FBXO42 loss-sensitive cells, likely working together rather than redundantly. Lastly, we demonstrated that the GSC-specific viability loss is due to an extended metaphase arrest upon FBXO42 knockout due to prolonged spindle assembly checkpoint activation. Altogether, this body of work both improves upon existing CRISPR-Cas9 technologies and uses this technology as part of an effort to identify potential novel therapeutic opportunities in GBM and other cancers.

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGEMENTS

My thesis work wouldn't have been possible without the support of so many wonderful people. I am truly grateful to all these people for helping me during the journey of my graduate career.

I am extremely thankful to my thesis advisor, Patrick Paddison, for taking me on as his graduate student and for always supporting me throughout this journey. It's rare to find a mentor who is both so excited and enthusiastic and so willing to support a mentee in whatever way he or she requires. Patrick always encouraged me to pursue my own ideas but gave valuable suggestions when I needed it. I enjoyed all of our stimulating scientific conversations, through which I learned so much, including how to expand my scope of scientific thinking. He also taught me how to put results into perspective and glean as much information as possible from data, especially when that data was not what I had expected. I am so thankful to Patrick for always taking the time to make my questions and needs a priority and for being so magically available whenever I needed his advice. Above all, I appreciate that Patrick has always made me feel like a truly indispensable, valued part of the lab.

All of my labmates have been a tremendous source of support. These people have been more like an extended family to me than just labmates, and I know that I can always count on them to help me with anything I need. Their incredible kindness and patience still often astounds me. They all provided me not only with priceless help with experiments, but also with the encouragement I needed throughout the roller coaster that is graduate school. In particular, I am indebted to Dan Kuppers for his continual guidance and advice and for being a wonderful bay mate for these 6 years. Thanks to Tobi Bonifert for teaching me about other aspects of Biology, listening to me rant about all kinds of things, and speaking German with me every day. Thanks to Heather Feldman for her kindness, good humor, and constant encouragement, and for many valuable discussions about both science itself and life as a scientist. Thanks to Jake Herman for being an exemplary scientist, often helping me to increase the scope of my scientific ideas, and providing incredible encouragement and support. Thanks to Megan Kufeld for her tremendous help with experiments for my thesis work, for many insightful life conversations, and for helping me become a better mentor. Thanks to Chad Toledo for mentoring me when I first joined the lab and being a wonderful role model for a productive and good-natured graduate student. Thanks to Sonali Arora for all her help with Bioinformatics and for her patience in juggling so many projects, and for such fun happy hours. Thanks to Phil Corrin for always making cool gadgets and contraptions in the lab for us and for always checking in with me about my thoughts on Star Wars and other matters. Thanks to Lucas Carter for all his help with cloning and screening and for his humorous comments. Last but not least, I want to thank Pam Lindberg and An Tyrrell for all their help with paperwork, budgets, ordering, etc. and for many wonderful chats about every aspect of life. I want to thank all of these people for being all-around amazing and fun to work with. They have supported me throughout graduate school and I definitely couldn't have done it without them.

Thank you to my thesis advisory committee members – Eric Holland, Jon Cooper, Nina Salama, and Pete Nelson – for all of their time, scientific insight, useful advice, and

encouragement. A special thanks goes out to Jon Cooper and Eric Holland for reading my thesis, and to Nina Salama and Jon Cooper for serving in the role of my GSR. It's remarkable that all of these faculty members with extremely busy schedules took the time to help a graduate student like me, and for that I am very grateful.

Graduate school has many ups and downs, and I am indebted to my friends for all their support, advice, encouragement, and companionship. They celebrated the highs with me and commiserated during the lows, and they made sure that I took some time away from work as well, which enriched my entire experience.

I truly cannot thank my parents enough. It is an extreme understatement to say that I wouldn't be here without them, not just because I owe my physical existence to them, but also because of everything they've done to help me become the person I am today. They never pushed any of their children to pursue any one particular thing. Instead, from a very young age, they raised us to be always curious and have a love for learning. They motivated me to figure out what I was interested in and to work hard to pursue that, and they taught me that I could do anything I set my mind to. I also want to thank my incredible siblings – Clara, Simon, and Jakob – for being such wonderful companions to have around my entire life. It has been an amazing experience to grow up with three people who are all so unique and have shared their valuable and varied insights with me over the years. From the bottom of my heart, I want to thank my whole family for all the love and support that they have given me for my whole life, including during graduate school, and for the lessons they have taught me, which helped me become the person I am today.

Finally, thank you to my husband Jason for the wonderful life that we've had so far together and for helping me to succeed in graduate school. He has always made me laugh when I needed it, made sure I stayed well-fed in graduate school by often cooking me food, listened to me excitedly tell him about experiments even though he is not a scientist, and provided me with encouragement to help me move projects forward. He is an amazing life buddy, and I will always appreciate all his love and support.

Chapter 1. INTRODUCTION

1.1 GLIOBLASTOMA OVERVIEW

1.1.1 *Disease Features and Glioblastoma Stem-like Cells*

Gliomas, named for the glial cell type that they most closely resemble histologically, account for approximately 30% of all central nervous system tumors[1]. They are classified into four grades of increasing aggressiveness, with the highest grade being glioblastoma (GBM)[2]. GBM is overall the most aggressive and common form of brain cancer[3,4], and there are currently no highly effective therapies against GBM. With standard-of-care treatments, ~90% of adult patients die within 2 years of diagnosis[5,6]. The median survival for GBM patients overall ranges from 14-17 months, with rare exceptions of long-term survival[4,7].

GBM is characterized by a complex vasculature and diffuse infiltration of tumor cells into the surrounding brain tissue. Current standard-of-care is surgery, radiation, and the alkylating agent temozolomide (TMZ)[4,7,8]. While effective at debulking the tumor mass, this regimen only provides a modest survival benefit, as tumors essentially always recur[7]. The prevailing rationale is that tumors harbor invasive, slow-dividing GBM stem-like cells (GSCs) that are missed by surgery and which are resistant to standard-of-care. This concept was elegantly demonstrated in a mouse model of glioma where a quiescent subset of endogenous glioma cells were shown to be responsible for tumor regrowth after TMZ treatment[9].

GBM tumors exhibit a remarkable degree of intratumoral heterogeneity, which may also contribute to their resistance to standard-of-care. In the tumor milieu, GSCs have been hypothesized to be responsible for giving rise to this cellular heterogeneity[10-16], exhibiting some

aspects of neurogenic developmental programs[10-12,14-17]. GCSs may also promote resistance to chemoradiotherapy by engendering a higher capacity for DNA damage repair[10,11].

There is strong evidence to support the notion that adult neural stem cells (NSCs) can serve as the cell-of-origin for GBMs, particularly since various mouse models that inactivate the tumor suppressors *PTEN*, *TP53*, and/or *NF1* or express GBM oncogenic drivers specifically in NSCs cause aggressive GBMs[18-24]. However, it is also possible that GBMs are derived from neural progenitor cells, or relatively mature astrocytes that dedifferentiate into a more stem-like state as they become transformed. A number of approaches have provided experimental evidence to support this[25-31], and it is clear that some but not all mutations are capable of initiating gliomas in mature astrocytes. The cell-of-origin of GBM may not have a clear answer and it is possible that different glioma tumors have multiple origins. Regardless, the important point is that GBM-initiating cells are functionally stem cell-like and have similar expression profiles and developmental potential as NSCs[17,32].

1.1.2 *Common Genomic Alterations and Glioblastoma Subtypes*

Mathematical modeling to predict the temporal sequence of driver events during tumorigenesis has shown that loss of chromosome 10 and gain of one or more copies of chromosome 7 occur very early in all primary GBMs[18]. *PTEN* is most likely the gene that provides the strongest selective advantage to losing a copy of chromosome 10, while *PDGFA* is most likely the gene that provides the strongest selective advantage to gaining extra copies of chromosome 7[18]. Therefore, hemizygosity of *PTEN* and increased PDGFA levels could represent important early drivers of primary GBM tumorigenesis.

Furthermore, analysis of hundreds of GBM patient samples in The Cancer Genome Atlas (TCGA) has revealed that the PI3K/MAPK pathway, the p53 pathway, and the Rb pathway are

altered in the vast majority of cases[33]. Considering RTK genes, PI3-kinase genes, and *PTEN*, 89.6% of GBMs have at least one alteration in the PI3K/MAPK pathway. The p53 pathway is dysregulated in 85.3% of tumors, through mutation/deletion of *TP53* (27.9%), amplification of *MDM1/2/4* (15.1%), and/or deletion of *CDKN2A* (57.8%). Lastly, 78.9% of tumors have one or more alteration affecting Rb function: 7.6% by direct *RB1* mutation/deletion, 15.5% by amplification of *CDK4/6*, and the remainder via *CDKN2A* deletion[33]. Additionally, a transgenic mouse model of deletion of *PTEN*, *TP53,* and *RB1* in GFAP-expressing cells created highly penetrant, high-grade gliomas[34]. These observations indicate that alterations in these three pathways are required events in the tumorigenesis of GBM.

Gene expression profiling of primary GBMs has divided them into four molecular signatures/subtypes: proneural, mesenchymal, classical, and neural[35-37]. Interestingly, recent work suggests that primary GBMs of all subtypes actually evolve from a common proneural-like precursor glioma[18]. Moreover, GBM tumors also display intratumoral heterogeneity, with multiple developmental subtypes co-existing in the same tumor[38-41]. Single cell RNA-seq studies have yielded critical insight into this intratumoral heterogeneity and developmental gene expression patterns for primary gliomas[38-40,42-44]. One key conclusion is that each tumor represents a complex neuro-developmental ecosystem, harboring diverse cell types, which presumably contribute to tumor growth and homeostasis in specific ways (e.g. vascular mimicry, immune evasion, recreating neural stem cell niches, neural injury responses, etc.).

Of note, a number of studies have sought to define responses to chemoradiation (e.g. NFkB, p53, STAT3, YAP/TAZ)[45-52]. Radiation treatment in particular induces or accentuates a proneural to mesenchymal transition via upregulation of signaling pathways (e.g. Jak/Stat) and mesenchymal transcription factors (e.g. CEBPB), which in turn may impart radiation resistance[46,48,53].

Importantly, it has been shown that the proneural to mesenchymal transition is a cell-intrinsic property of tumor cells that occurs within 6 hours of radiation[46]. However, it remains unclear from these and other studies which subpopulations of tumor cells mount these responses.

1.2 GLIOBLASTOMA RESEARCH MODELS

1.2.1 *Mouse Models of Glioma*

Animal models have the advantage of incorporating complex processes such as angiogenesis and metastasis and the tumor microenvironment into the study of cancer. Mice share extensive molecular and physiological similarities to humans, and they allow for studying tumorigenesis in a temporally and genetically controlled system. They can also be used to measure response to treatments, making them a powerful tool. The three main types of GBM mouse models are patient-derived xenograft (PDX) models, genetically engineered mouse models (GEMMs), and viral systems such as the replication-competent avian sarcoma-leukosis virus long terminal repeat with splice acceptor/tumor virus a (RCAS/tv-a) system.

PDX models in immune-compromised mice have the advantage that they represent actual human GBM cells with all of their many genomic alterations and complex biology, which immune-competent models cannot do. PDX models show an invasive growth pattern and can recapitulate the histology and gene expression profiles of their primary tumors, as well as display cellular hierarchies[12,17,54]. It is preferable to inject patient cells into mice as soon as possible after tumor biopsy or to use low-passage GSC cultures, and to transplant them orthotopically into the brain. However, tumor biopsies have also first been transplanted into and expanded in mouse flanks before subsequently being harvested and then injected into the brain, and this can also give rise to tumors that resemble human GBM[55]. Of course, since mice must be immune-compromised in order

for human cells to engraft, PDX models do not recapitulate the *complete* microenvironment with all types of immune cells.

In order to model a complete microenvironment, GEMMs can be used. GEMMs for GBM were initially rather unsuccessful until it became clear that all three main pathways that are commonly altered in GBM – the PI3K/MAPK pathway, the p53 pathway, and the Rb pathway – had to be altered in the mouse models as well[56]. There are now various GEMMs that can be used for gain-of-function, loss-of-function, and chromosome engineering studies[57]. Remarkably, tumors in these models can undergo additional gene copy number changes that have an astonishing similarity to the molecular subtypes found in human GBMs[56]. As in human gliomas, additional loss of the tumor suppressor *PTEN* causes higher-grade malignancy and reduced survival in mouse glioma models[20].

A disadvantage of GEMMs of glioma is that most achieve their effects through the widespread expression of or knockout of a gene across the entire brain, meaning alterations in these models are not restricted geographically. Therefore, multifocal lesions are often observed[58,59], which is more consistent with a cancer predisposition syndrome rather than a sporadic brain tumor. An alternative approach that allows for more localized delivery of tumorigenic genetic events is the use of viruses. This method can also be much faster since various combinations of alterations of interest can be induced simultaneously postnatally, without the need to both create new transgenic lines and then cross multiple transgenic lines in order to create combinations of genetic events[60]. One highly successful viral system that allows for localized, postnatal cell type-specific gene transfer in mice is the RCAS/tv-a system. In this system, virus-producing packaging cells transfected using an RCAS vector containing an oncogenic driver and/or tumor suppressor knockdown construct are injected into the brains of mice engineered to

express the RCAS receptor (tv-a) in Nestin+ or GFAP+ cells[61-66]. This allows for the targeting of glioneuronal progenitor cells or astrocytes, respectively. RCAS-mediated tumors are replete with expression patterns and also histological features common to human GBMs, including pseudopalisading necrosis, hypervascularity, and microvascular formation throughout the tumor[61-66].

1.2.2 *In Vitro Models*

While *in vitro* serum-cultured cell lines have yielded valuable insights into GBM cell biology, they are not ideal for faithfully representing the primary tumor from which they were derived. Principal component analysis of gene expression profiles of patient-derived GBM cells grown in defined serum-free media, the same cells grown in serum culture, and the primary tumors from which they were derived revealed that GBM cells grown in serum-free media cluster much more closely with their primary tumors[17]. Over time, serum-cultured GBM lines tend to adopt a more mesenchymal-like gene expression profile, regardless of the subtype(s) of their primary tumor[12,17]. Thus, patient-derived isolates are far preferable to serum-cultured lines for recapitulating GBM biology.

The tumor-initiating GSCs that have been isolated from patients and cultured under serum-free monolayer conditions retain the developmental potential and specific genetic alterations found in the original tumor[12,17,67,68]. They can also retain tumor-initiating potential and tumor-specific genetic and epigenetic signatures over extended outgrowth periods[12,17]. In addition, they have been shown to recreate tumor cellular hierarchies when implanted into the cortex of immune-compromised mice[12,17]. Furthermore, GSC isolates also retain expression of neural progenitor molecular networks, which may contribute to the aggressive behavior of GBM tumors through enhancing self-renewal or developmental programs[14-16], DNA repair pathways[11], angiogenesis[69,70], and/or invasiveness[71]. GSC isolates and GSC-derived xenograft tumor models cannot recreate the

complete cellular heterogeneity of the original patient tumor, since they likely represent a tumor GSC subpopulation which responds well to the defined *in vitro* outgrowth conditions. However, they represent useful models for studying subsets of GSCs with specific genetic and molecular profiles.

GBM biology can also be studied *in vitro* by creating transformation models. One recent paradigm achieves this by genetically engineering human induced pluripotent stem cells (hiPSCs) to reflect tumor-associated genetic driver mutations[72]. Interestingly, these models are shown to harbor intratumoral heterogeneity and extrachromosomal DNA amplifications seen in GBM[72]. Another option besides using hiPSCs is to similarly transform NSCs. We have created such models with fetal-derived NSCs by using a variety of combinations of GBM driver alterations, such as expression of *EGFRvIII* and constitutively active *AKT*[73], or CRISPR knockout of *CDKN2A*, *TP53*, *PTEN*, and *NF1* (see Chapter 2).

Other *in vitro* GBM models include GBM 3D organoids. These can be created by genetically manipulating brain organoids to develop oncogenic properties[74,75] or by co-culturing tumor spheres with brain organoids to model tumor cell invasion[76,77]. Alternatively, patient-derived GBM cells can be cultured in organoid format using Matrigel and exogenous EGF/bFGF. In this system, the cells can maintain stem cell heterogeneity and a hypoxic gradient for several weeks[78]. Furthermore, patient-derived organoids have recently also been cultured without cell dissociation (i.e. as "tumor chunks") under defined media conditions, without added extracellular matrix or EGF/bFGF[79]. These organoids can recapitulate the histological features, cellular diversity, gene expression, and mutational profiles of their corresponding parental tumors[79].

1.3 Functional Genomic Approaches in Cancer Research

1.3.1 *Functional Genomic Screening*

Even though large-scale genomic studies have given us extensive knowledge of the tumor suppressors and oncogenes that are commonly altered in various cancers, directly targeting such genomic alterations therapeutically has often been met with limited success. For instance, in GBM recent attempts to improve standard-of-care using targeted therapeutics for BRAF (sorafenib[80-82]), c-MET (onartuzumab[83]), EGFR (erlotinib[84] and lapatinib[85]), mTOR (everolimus[86] and temsirolimus[87]), and PI3K (PX-866[88]) have failed to produce significant survival benefits for GBM patients. An alternative approach is to identify vulnerabilities in cancer cells that arise due to the oncogenic state.

Cancer cells must tolerate various stresses on cellular processes, such as increased mitotic stress, which further promotes aneuploidy, which leads to proteotoxic stress. As another example, cancer cells often contain elevated levels of reactive oxygen species, which results in increased levels of DNA damage[89]. Additionally, the genomic instability that accompanies cancer causes passenger mutations, which can still contribute to the cell's phenotype, even if they are not oncogenic drivers[90]. Altogether, this oncogenic state leads to genetic vulnerabilities unique to tumor cells, which we may be able to exploit in order to selectively kill tumor cells compared to non-transformed cells, a concept first proposed by Hartwell et al. in 1997[91]. Since then, functional genomic screens have emerged as the top tool of choice for *unbiased* and *high-throughput* identification of such cancer vulnerabilities[73,92-105].

Functional genomic screens in cells of interest can be performed in either an arrayed or pooled format. Arrayed screens are performed by eliciting one perturbation at a time in individual wells or plates, which has the advantages of allowing for complex phenotypic readouts and

eliminating noise that can be seen when assaying multiple perturbations simultaneously in a pool. However, arrayed screens are constrained by the number of perturbations that can be performed at once due to time, lab infrastructure, and cost. Pooled functional genomic screens, on the other hand, allow for simultaneous assaying of tens to hundreds of thousands of perturbations, even genome-scale analyses. These screens are usually performed by infecting target cells with a lentiviral library where individual viral particles represent the desired perturbations, which most often utilize RNAi via short-hairpin RNAs (shRNAs) or the CRISPR-Cas9 system. ShRNAs function through the Dicer/Argonaute system to decrease levels of messenger RNA (mRNA) transcripts of interest, promoting a decrease in their corresponding protein products. In the case of CRISPR-Cas9, a single guide RNA (sgRNA) confers target recognition to the DNA endonuclease Cas9, which creates a double-strand break at the genomic target locus. This often results in protein knockout due to repair scars left behind when the break is repaired by the error-prone non-homologous end joining (NHEJ) pathway[106,107]. After cells are transduced with a library for a pooled screen and the desired selection (e.g. outgrowth, reporter activity, enrichment for a cell surface marker or some other phenotype) is allowed to occur, the genomically integrated shRNAs/sgRNAs are sequenced using next-generation sequencing. The change in shRNA/sgRNA abundance before and after the applied selection pressure can then be evaluated to determine the effect that these individual perturbations had on the phenotype being investigated.

The Paddison Lab previously performed multiple pooled functional genomic *outgrowth* screens using RNAi or CRISPR-Cas9 libraries in multiple GSC isolates. Here the phenotype being investigated is the degree of requirement of various genes for cell proliferation, survival, or growth. As screening controls, we have employed fetal neural stem cells (NSCs), which have similar expression profiles and developmental potential but are not transformed[12,17,108]. This paradigm has

allowed us to identify genes that when inhibited show differential lethality in GSCs. Hits from these screens have revealed GSC-specific dependency on: BUB1B/BUBR1 to suppress lethal consequences of altered kinetochore function[96]; PHF5A to maintain proper exon recognition[97]; BuGZ to facilitate Bub3 activity and chromosome congression[109]; PKMYT1 to block premature entry into mitosis[73]; and ZNF131 to promote expression of HAUS5 to maintain the integrity of centrosome function and viability[110].

1.3.2 *Precision Medicine*

The goal of precision medicine is the idea that when a patient presents with cancer, a biopsy of his or her tumor is genomically profiled and compared to adjacent normal tissue, and this information is then used to make informed decisions about the best possible treatment regimen. Sometimes genetic mutations in oncogenic drivers represent cancer dependencies that can be targeted, such as the *BCR-ABL* fusion in chronic myelogenous leukemia[111-113] and mutant *KRAS* in non-small cell lung cancer[114]. However, as discussed in the previous section, this is often not the case. The enormous advancements in high-throughput next-generation sequencing technologies have ushered in an exciting increase in cancer omics characterization. Combining such tumor characterization with functional genomic screening allows for the identification of cancer synthetic lethal interactions, which brings us closer to the promise of precision medicine.

Two genes are considered to be synthetic lethal if inactivation of one alone is tolerable but their combined inactivation leads to cell death[115,116]. Likely one of the most famous synthetic lethal relationships in cancer is that between BRCA1/2 deficiency and PARP inhibition[117,118]. Other examples include PTEN loss combined with mTOR inhibition[119]; ARID1A loss (the most frequently mutated subunit of SWI/SNF chromatin remodeling complexes) combined with ARID1B inhibition[120,121]; and SMARCA4 loss combined with SMARCA2 inhibition[122-124].

Furthermore, cancer patient-specific vulnerabilities may be identified not only in tumor cells that carry inactivation of a gene, but also in cells that display a particular pathway activation, gene or protein expression profile, gene copy number change, gain-of-function mutation, metabolic profile, etc. Therefore, any genes that are required due to such a unique cancer cellular context can be more aptly termed "context-specific vulnerabilities". For instance, in previous CRISPR-Cas9 genome-scale screens, we identified the Wee1-like kinase PKMYT1/Myt1 as being required in the context of combined hyperactivation of the RTK/Ras and PI3K/AKT signaling pathways[73].

Interestingly, context-specific vulnerabilities are sometimes also the result of the copy number "passenger loss" of a gene that is proximal to a tumor suppressor gene that is the tumorigenic "driver loss". When such a passenger loss is an essential gene (often in the ribosome, proteasome, or spliceosome), its hemizygosity can create a preferential vulnerability for further depletion of that same gene or related pathway members[90]. An intriguing example is that the metabolic enzyme *MTAP* is frequently co-deleted with *CDKN2A* due to its genomic proximity to this tumor suppressor, conferring dependency on the arginine methyltransferase *PRMT5* because of the accumulation of an inhibiting metabolite[125]. Clearly, we as a scientific community are making enormous advancements in the identification of cancer-specific vulnerabilities, and we must continue this pursuit in order to develop improved therapies for cancer patients.

1.4 Research Aims

In order to identify genes required for the proliferation and survival of GBM cells, we previously performed genome-scale pooled outgrowth screens in human GSCs and NSCs using both RNAi and CRISPR-Cas9 platforms[73,96,97,109,110]. Partial retests of these screens led to the identification of genetic vulnerabilities associated with the oncogenic activity of RTK/Ras and PI3K/Akt pathways or MYC. In order to further characterize candidate GSC-specific hits from our CRISPR-Cas9

whole genome screens, here we performed comprehensive pooled retests of all putative screen hits that scored preferentially in any of three GSC isolates over two NSCs isolates.

While using lentiviral pooled libraries to perform these screens is immensely valuable, retesting screen hits using single sgRNAs can prove challenging when using lentivirus. In GSCs and NSCs, we have observed that the timeline of CRISPR editing is not very uniform with lentivirus and occurs over the course of ~5-12 days post infection in individual cells. This is problematic when studying essential genes since protracted indel formation results in mixtures of wild-type and knockout cells and thereby mixtures of live and dead cells. This represents a critical experimental limitation of the use of CRISPR-Cas9 platforms in primary cells. Therefore, I sought to create an optimized protocol that would allow maximal targeted indel formation over the shortest possible experimental window.

In chapter 2, I describe a method I optimized that employs direct nucleofection of ribonucleoprotein complexes (RNPs) composed of chemically synthesized 2'-O-methyl 3’phosphorothioate-modified sgRNA and purified Cas9 protein. With this optimized technique, we can routinely achieve >90% indel formation in only 3 days, even in hyperdiploid cells (such as GSCs). There is no need to create clonal lines for simple loss-of-function experiments, as I observed near-total protein loss of target genes in cell pools. Additionally, we found that this approach allows for the creation of targeted genomic deletions. This represents an extremely useful tool for the research community as a whole when wishing to inactivate not only coding genes, but also non-coding RNAs, UTRs, enhancers, and promoters. I extensively employed this technique in chapter 3 for high-efficiency gene knockout of GSC-specific functional screen hits.

In chapter 3, I describe the comprehensive pooled retest screens that we performed for all potential whole-genome CRISPR-Cas9 screen hits. Here the main aims were to identify GBM

gene dependency groups and to identify context-specific vulnerabilities. We observed that a subset of top GSC-specific retest hits display a strong and consistent increased dependency in GSCs compared to NSCs. To identify groups of such genes, particularly related genetic vulnerabilities that can be expected to be shared in the same GBM tumor, we integrated our results with data from the Broad Institute's DepMap[102-105], which contains functional genomic screening data for over 500 serum-cultured cell lines from various tumor types. By creating co-dependency networks using CRISPR gene dependency correlations, we identified GSC-specific gene vulnerability groups related to mitochondrial protein processing and turnover and membrane trafficking as well as metabolic enzymes and regulators.

In addition to identifying gene vulnerability groups that may be applied to GBM as a whole, we also wanted to use our retest screen results to identify context-specific vulnerabilities, or genes that are required in subsets of tumor cells with a particular cellular context (e.g. increased/decreased expression of a particular gene or group of genes). In order to generate such hypotheses, we integrated our functional screening data with descriptive genomic data for our GSC lines, as well as DepMap data. We specifically investigated two examples, ADAR and EFR3A. We observed that ADAR, the primary enzyme responsible for editing dsRNA by catalyzing the hydrolytic deamination of adenosine to inosine[126], is preferentially required in the context of an interferon-stimulated gene expression signature. We found that EFR3A, an adapter protein required to localize PI4KIIIα to the plasma membrane[127,128], is differentially required in cells that have low expression of the paralog EFR3B.

Furthermore, we investigated the particularly strongly GSC-specific screen hit FBXO42, a thus-far poorly studied F-box protein. FBXO42 is a unique screen hit that we found to be essential in a subset of GSCs *in vitro* and *in vivo* but nonessential in NSCs *in vitro*. This indicates that

FBXO42 could be particularly promising as a therapeutic target since it may have a very large therapeutic window. Furthermore, we determined that FBXO42 loss-sensitive tumors actually likely occur in various cancer types of different tissue lineages. A previous study described a role for FBXO42 in the destabilization of p53[129], and we thus wondered if *TP53* mutation status might explain requirement for FBXO42, which seemed plausible based on our highly *FBXO42* loss-sensitive GSC isolate being *TP53* wild-type. We demonstrated that the ubiquitin ligase role of FBXO42 is responsible for the viability phenotype, but that in fact, this does not appear to be through p53. In searching for possible interactors, we determined that FBXO42 and the co-dependency correlating gene CCDC6 are both necessary to promote viability in FBXO42 loss-sensitive cells, likely working together rather than redundantly. Lastly, we demonstrated that the GSC-specific viability loss is due to an extended metaphase arrest upon FBXO42 knockout due to prolonged spindle assembly checkpoint activation.

Altogether, this body of work both improves upon existing CRISPR-Cas9 technology and uses this technology as part of an effort to identify potential GBM novel therapeutic opportunities, which may represent therapeutic avenues in other cancer types as well.

Chapter 2. A SIMPLE AND EFFICIENT METHOD FOR MULTI-ALLELIC CRISPR-CAS9 EDITING IN GSCS/NSCS AND OTHER PRIMARY CELL CULTURES

2.1 ABSTRACT

Background: CRISPR-Cas9-based technologies have revolutionized experimental manipulation of mammalian genomes. None-the-less, limitations of the delivery and efficacy of these technologies restrict their application in primary cells.

Aims: To create an optimized protocol for penetrant, reproducible, and fast targeted insertion-deletion mutation (indel) formation in cell cultures derived from primary cells, using patient-derived glioblastoma (GBM) stem-like cells (GSCs) and human neural stem/progenitor cells (NSCs) for proof-of-concept experiments.

Methods: We employed transient nucleofection of Cas9:sgRNA ribonucleoprotein complexes using chemically synthesized 2'-O-methyl 3'phosphorothioate-modified sgRNAs and purified Cas9 protein. Indel frequency and size distribution were measured via computational deconvolution of Sanger sequencing trace data. Western blotting was used to evaluate protein loss. RNA-seq in edited NSCs was used to assess gene expression changes resulting from knockout of tumor suppressors commonly altered in GBM.

Results: We found that with this optimized technique, we can routinely achieve >90% indel formation in only 3 days, without the need to create clonal lines for simple loss-of-function experiments. We observed near-total protein loss of target genes in cell pools. Additionally, we found that this approach allows for the creation of targeted genomic deletions. We also

demonstrated the utility of this method for quickly creating a series of gene knockouts that allow for the study of oncogenic activities.

Conclusion: Our data suggest that this relatively simple method can be used for highly efficient and fast gene knockout, as well as for targeted genomic deletions, even in hyperdiploid cells (such as GSCs). This represents an extremely useful tool for the cancer research community when wishing to inactivate not only coding genes, but also non-coding RNAs, UTRs, enhancers, and promoters. This method can be readily applied to diverse cell types by varying the nucleofection conditions.

2.2 BACKGROUND

In bacteria and archaea, the CRISPR-Cas (Clustered, Regularly Interspaced, Short Palindromic Repeats (CRISPR)–CRISPR-associated (Cas)) pathway acts as an adaptive immune system, conferring resistance to genetic parasites and bacteriophage[130,131]. CRISPR-Cas systems are able to target and degrade DNA[130,131], and this property has been harnessed for directed genome editing in prokaryotes and (more recently) eukaryotes, including human cells[132-135], using the type II CRISPR-Cas system from *Streptococcus pyogenes*. In its simplest form, this system consists of a complex of two components, the Cas9 protein and a single guide RNA (sgRNA). Cas9 is an RNA-guided DNA endonuclease. The sgRNA is a chimeric guide RNA composed of a ~20nt 'protospacer' sequence, which is used for target recognition, and a structural RNA required for Cas9:sgRNA complex formation (i.e. tracrRNA). In addition, DNA cleavage by Cas9 occurs only in the presence of an appropriate protospacer adjacent motif (PAM) at the 3' end of the protospacer sequence in the target genomic locus (for Cas9 this is "NGG", where N is any nucleotide[131]).

When Cas9 and an sgRNA are expressed together, a double-strand DNA (dsDNA) break is created about 3 bp upstream of the PAM site[136,137]. This break is then repaired by the cell either

via the high-fidelity homology-directed repair (HDR) pathway, or much more commonly, via the error-prone non-homologous end joining (NHEJ) pathway, which leaves behind repair scars in the form of small insertion-deletion (indel) mutations[106,107]. When these indels occur in an exon, they can cause frameshifts and premature stop codons in the target gene, effectively ablating protein function[133,135,138].

Cas9:sgRNA targeting efficiency in human cells varies considerably depending on the methods, reagents, and cell types used. In general, successful generation of indels using transient DNA transfection occurs in a range of ~1-30%[137]. However, it was shown that lentiviral-based stable expression of Cas9:sgRNA greatly improves targeting efficiency to >90%[131,139,140]. As a result, we and others have successfully performed pooled lentiviral-based sgRNA screens in various human cell types[73,139,140]. However, retesting single sgRNAs from these screens, especially those targeting essential genes, can prove challenging. For example, in human neural stem/progenitor cells (NSCs) and patient-derived glioblastoma stem-like cells (GSCs), we have observed that "all-in-one" lentiviral-based CRISPR systems can result in protracted windows of indel formation and phenotypically mixed populations, requiring incubation of up to 12 days to achieve complete indel formation[73]. As a result, it can be difficult to set up rigorous experiments analyzing a particular gene knockout (KO) when cell populations contain variable mixtures of wild-type (wt) and indel-containing alleles and, if the target gene is essential, thereby contain mixtures of alive and dead cells. This represents a critical experimental limitation of the use of CRISPR-Cas9 platforms in primary cells. As a result, we wished to create an optimized protocol that would allow maximal targeted indel formation over the shortest possible experimental window.

We found a robust method that utilizes transient nucleofection of *in vitro*-formed Cas9:sgRNA ribonucleoprotein (RNP) complexes using chemically synthesized 2'-O-methyl 3'phosphorothioate-modified sgRNAs and purified Cas9 protein (Figure 3.1A). With this optimized approach, we are able to achieve >90% indel formation in multiple human GSCs and NSCs in only 3 days, without the need to create clonal lines for simple loss-of-function experiments. In addition, we find that this approach allows for the creation of targeted deletions in cell pools or cell clones, depending on the size of the desired deletion. Here we present these results illustrating the utility of this method and a detailed step-by-step protocol.

2.3 RESULTS

2.3.1 *Nucleofection of Cas9:sgRNA (chemically synthesized, modified) RNPs results in highly penetrant generation of small indels in human GSCs and NSCs*

Due to the previously-discussed challenges with indel formation using lentiviral-based Cas9:sgRNA delivery, we explored alternative approaches, including the use of Cas9:sgRNA ribonucleoprotein (RNP) complexes composed of purified Cas9 protein and purified gRNA. Such RNPs have recently been used effectively for several applications, including gene loss-of-function in human cell lines[141-143] and ES cells[141], editing of CXCR4 in human T cells[144], HDR tests in human cells via insertion of restriction sites[141,145], epitope tagging in mouse NSCs[146], and studying effects of sgRNA sequence on editing efficiency[147]. The efficiencies reported for *in vitro* editing in these contexts are most often in the range of ~15-60%, with most studies reporting *maximum* (not routine) efficiencies ≤80%. We wanted to further improve upon these RNP methods so we could routinely achieve high efficiency, multi-allelic editing in GSCs and NSCs (Figure 2.1A).

Since most studies using RNPs utilize *in vitro*-transcribed gRNA, we instead tested chemically synthesized sgRNAs with 2'-O-methyl 3' phosphorothioate modifications in the first

and last 3 nucleotides, which are more nuclease resistant than unmodified sgRNAs and therefore likely increase RNP half-life[148,149]. We formed RNPs by combining these sgRNAs with purified sNLS-SpCas9-sNLS nuclease and then delivered them via nucleofection, a modified electroporation technique developed by Amaxa (now Lonza) that allows direct transfer of nucleic acids into the nucleus of mammalian cells in culture.

To follow indel formation in cell pools, we employed a method that uses Sanger sequencing of sgRNA target site-spanning PCR amplicons followed by computational trace decomposition of the control and experimental traces to predict indel frequency and KO frequency[150,151], where KO frequency is the percent of predicted sequences that result in a frameshift or an indel ≥21 bp in length[151] (Figure 2.1B). In order to determine the efficiency and timing of indel formation, we measured indel and KO frequency for RNP doses ranging from ~2-60 pmol (Figure 2.1C) and at 24, 48, and 72 hours post-nucleofection (Figure 2.1D) in diploid NSC-CB660 cells, near-diploid GSC-0131 cells, and hypertriploid GSC-0827 cells, for single sgRNAs targeting *TP53* exon 7 (sgTP53-1) and *NF1* exon 2 (sgNF1-1). *TP53* and *NF1* are both located on chromosome 17, of which GSC-0131 have 2 copies and GSC-0827 have 3 copies. These experiments revealed that high (>90%) multi-allelic indel efficiencies can be achieved starting at RNP doses of 7.5-15 pmol in GSCs and NSCs (Figure 2.1C). Furthermore, we observed that 50-70% of editing has occurred by 24 hours post-nucleofection, reaching its maximum by 72 hours (Figure 2.1D).

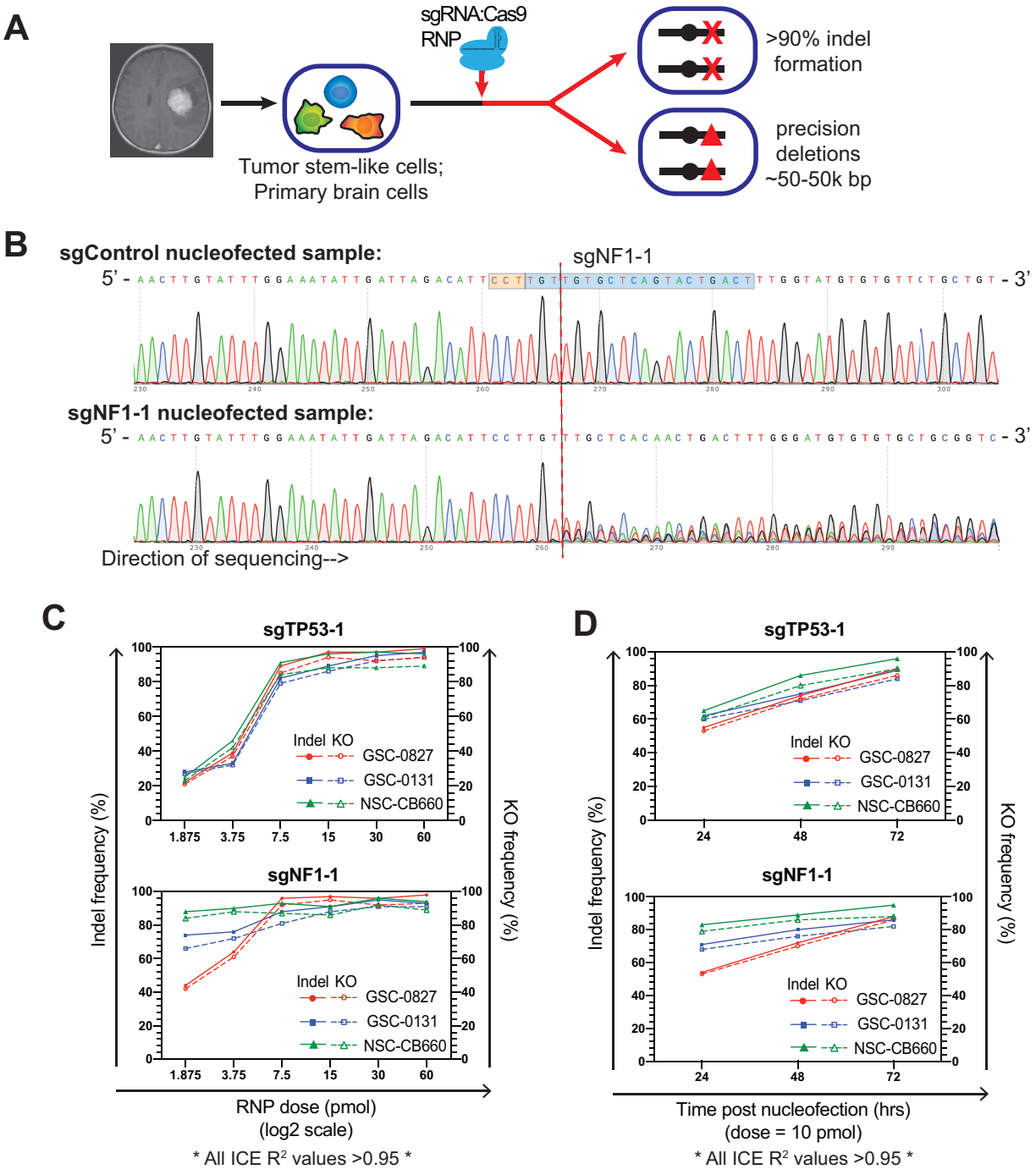


Figure 2.1. Highly efficient and fast indel formation using RNPs composed of purified Cas9 and chemically synthesized 2'-O-methyl 3'phosphorothioate-modified sgRNA. **(A)** Overview of CRISPR RNP targeting strategy. **(B)** Sanger trace example for a pool of NSC-CB660 nucleofected with 15 pmol RNPs using either a non-targeting control sgRNA or sgNF1-1 (72 hours post-nucleofection). Sequence was created using a reverse sequencing primer. Blue shaded box denotes sgRNA, orange shaded box denotes PAM sequence, and dotted red line represents sgRNA cut site. The sgNF1-1 treated cells produce a clean sequence trace that mirrors the control up until the sgRNA cut site, at which point the trace begins to represent the compounding effect of multiple

overlapping traces due to various indel mutations. Sanger trace data like this was used in conjunction with a freely-available bioinformatics tool (ICE) in order to predict CRISPR editing sequence distribution in cell pools. **(C)** CRISPR editing efficiency as a function of RNP dose for 2 different sgRNAs. Solid data lines denote indel frequency while dotted lines denote predicted KO frequency (% of predicted sequences that result in a frameshift or an indel ≥21 bp in length). All samples were harvested 72 hours post-nucleofection. **(D)** CRISPR indel frequency and KO frequency as a function of time post-nucleofection for 2 different sgRNAs. A dose of 10 pmol was used.

Similar indel penetrance for sgRNAs targeting other genes was also observed in multiple GSC isolates and NSCs which had been immortalized and oncogenically transformed (Figure 2.2). Because many of our sgRNA sequences had been prevalidated through our lentiviral screens[73] (e.g. sgTP53-1; sgNF1-1; 13 of 23 sgRNAs in Figure 2.2), these experiments illustrate representative results for active RNPs. Of note, several of the genes targeted are present at hyperdiploid levels in GSCs; for instance, GSC-G166 contain 3 copies of *SCAP*, *FBXO42*, and *GMPPB*. Furthermore, targets included top scoring essential genes for both GSCs and NSCs, which cause profound viability loss[73], indicating that the high efficiencies we observe are *not* simply due to outgrowth of edited cells.

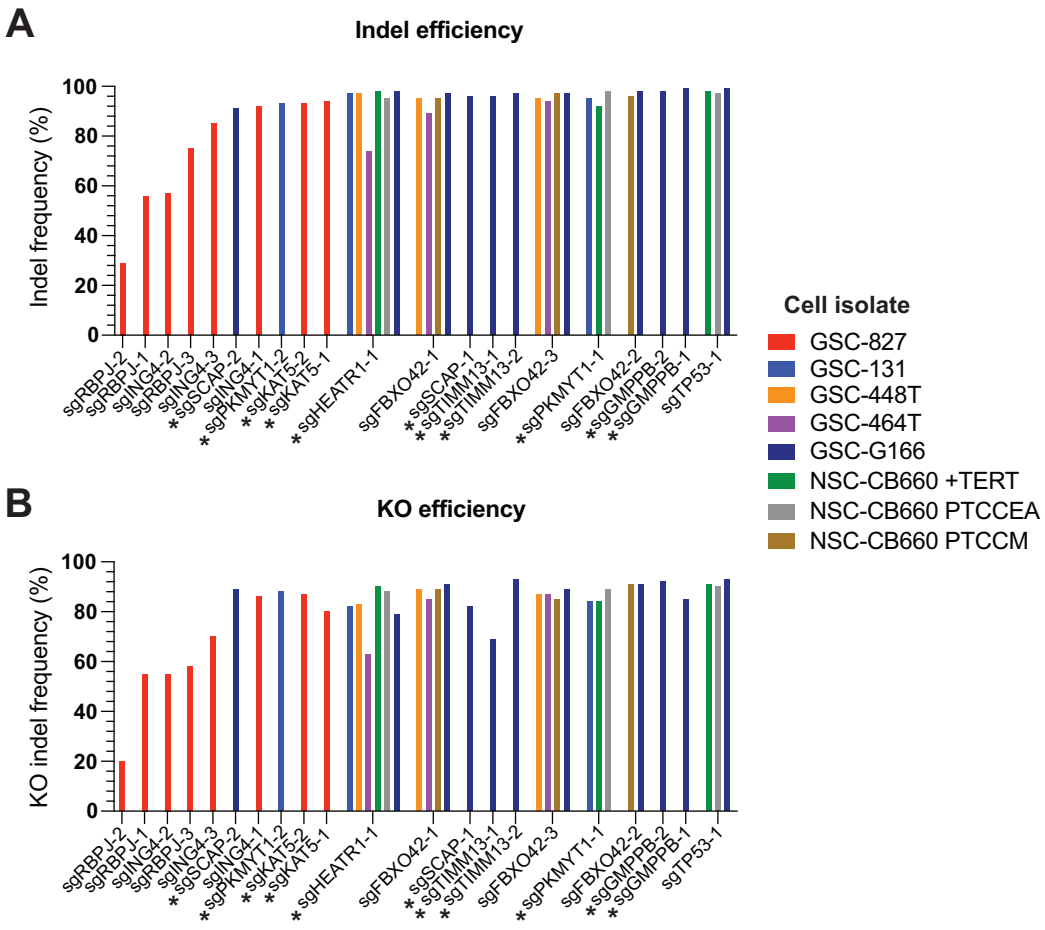


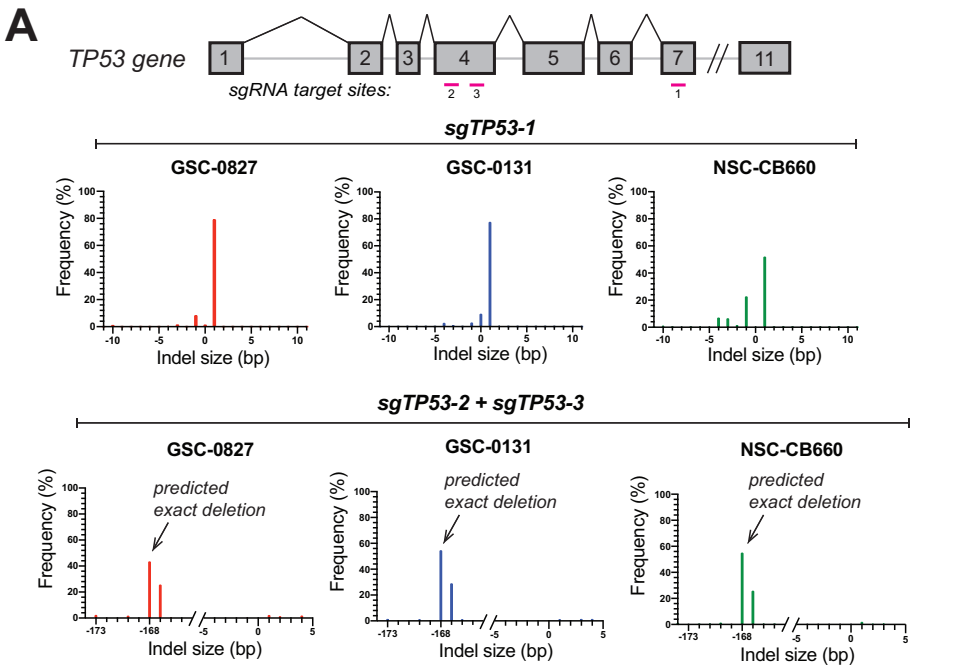
Figure 2.2. Additional sgRNA targeting for different loci and different cell isolates. **(A)** Predicted indel frequencies for 21 different sgRNAs tested in various GSC and NSC cell isolates. Gene knockout using sgRNAs labeled with * caused reduced viability and/or reduced proliferation in the cell isolate(s) shown. +TERT means cells were transduced with a retroviral vector expressing TERT. +PTCCEA means cells were transduced with dominant-negative TP53DD, CCND1 + CDK4R24C (p16 binding deficient), EGFRvIII, and myristoylated AKT. +PTCCM means cells were transduced with dominant-negative TP53DD, CCND1 + CDK4R24C (p16 binding deficient), and c-Myc. **(B)** Corresponding predicted knockout frequencies for the sgRNAs shown in (A).

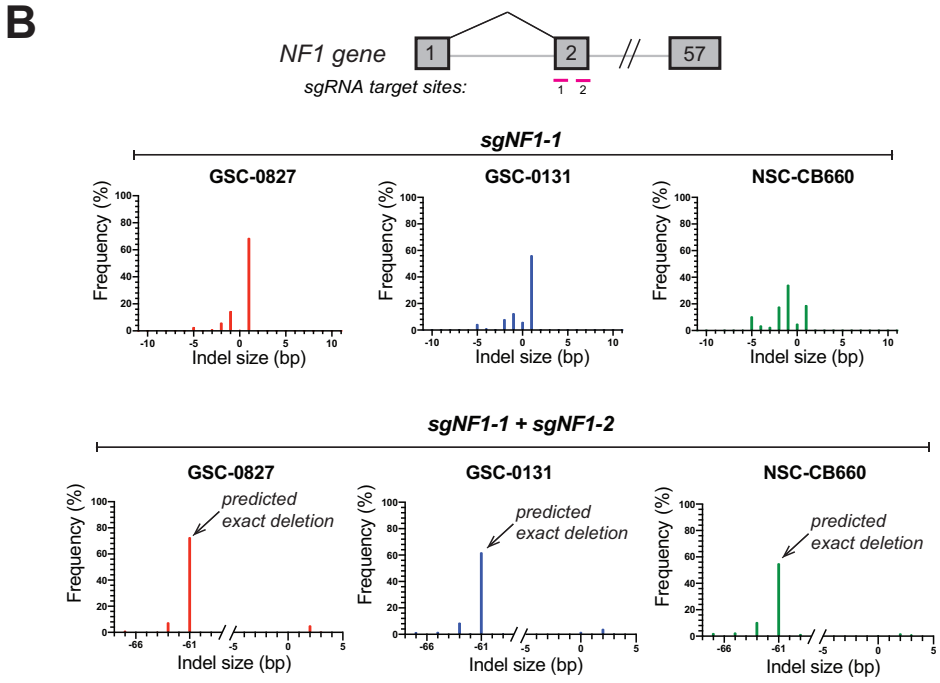
2.3.2 *RNP nucleofection allows for deletion of several hundred bp genomic regions*

In our efforts to assess RNP efficiency and dosing, we also observed that for sgTP53-1 and sgNF1-1, the KO efficiency closely mirrored the indel efficiency (Figure 2.1C-D). In these cases, this is due to repair bias at the target sites, where each sgRNA produced a high percentage of 1bp insertions in GSCs and a mixture of frameshifting 1bp insertions/deletions and 2 bp deletions in

NSCs (Figure 2.3A-B, top panels). This is consistent with a recent analysis of sgRNA targeting repair events in human cells, which found that frameshift frequencies are higher than expected (81% versus the expected 67% for random NHEJ-mediated repair) due to unexpectedly high 1 bp insertion/deletion events[138].

Given the precise and reproducible nature of the indels created for sgNF1-1 and sgTP53-1, we wondered if using two sgRNAs in close proximity (e.g. 50-1000 bp) would favor the production of precise deletions using our protocol. This was indeed the case when we simultaneously nucleofected with two sgRNAs targeting *TP53* or *NF1* (Figure 2.3A-B, bottom panels). We observed remarkably high predicted exact deletion frequencies between 44-73% for GSCs and NSCs (Figure 2.3C). Allowing +/-2bp for the deletion size window further increased the predicted "near-precise" deletion efficiencies to 68-86% (Figure 2.3C). It is also important to note that in this multiple sgRNAs scenario, the bioinformatic predictions for total indel frequencies were somewhat reduced due to adjustment for slightly lower regression fit R^2 values. (These R^2 values are a measure of how well the prediction correlates with the actual Sanger trace data). However, essentially 0% wt sequences were predicted in the trace data (see indel size of "0" in bottom panels of Figure 2.3A-B), suggesting that the total indel frequencies – and thereby also the deletion frequencies – may actually have been even higher. We observed similar near-precise deletion results when nucleofecting with two other sets of double sgRNAs (Figure 2.4).





C

	sgTP53-2 + sgTP53-3			sgNF1-1 + sgNF1-2		
	GSC-827	*GSC-131*	*NSC-CB660*	*GSC-827*	*GSC-131*	*NSC-CB660*
Total indel frequency (%)	83	94	94	96	89	87
ICE R^2 value	0.83	0.94	0.94	0.96	0.91	0.88
Exact deletion frequency (%)	43.58	54.66	55.1	72.91	62.11	55.19
Exact deletion +/- 2bp (%)	71.21	86.35	83.21	80.78	71.87	67.53

* All other indel sizes each present at <6% (most <2%) *

Figure 2.3. Cas9:sgRNA RNPs can be used to create small, near-precise targeted deletions. **(A)** Indel size distribution of predicted indel sequences for cell isolates nucleofected with individual sgTP53s (top panel) or

two simultaneous sgTP53s (bottom panel) (168 bp apart). **(B)** Indel size distribution of predicted indel sequences for cell isolates nucleofected with individual sgNF1s (top panel) or two simultaneous sgNF1s (bottom panel) (61 bp apart). **(C)** Summary of dual sgRNA indel profiles shown in (A) and (B).

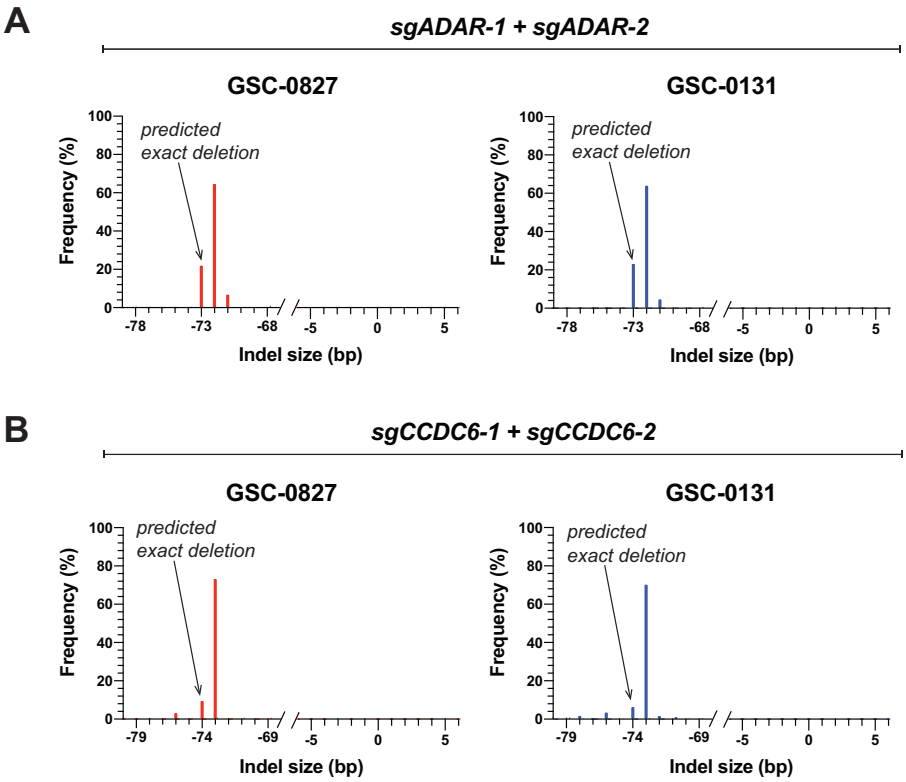


Figure 2.4. Further evidence that CRISPR RNPs can be used to create near-precise targeted deletions. **(A)** Indel size distribution of predicted indel sequences for 2 GSC cell isolates nucleofected simultaneously with two sgRNAs targeting ADAR. The dual sgRNA cut sites were 73 bp apart. **(B)** Indel size distribution of predicted indel sequences for 2 GSC cell isolates nucleofected simultaneously with two sgRNAs targeting CCDC6. The dual sgRNA cut sites were 74 bp apart.

To further investigate these results, we designed sets of 3 sgRNAs to target 13 different non-coding genomic loci on chromosomes X, 2, 5, 13, 15, and 21 in GSC-0827 cells, which contain 4, 3, 3, 2, 3, and 3 copies of these chromosomes, respectively. These sgRNA sequences were designed using the Broad GPP Web Portal[152] and were used without prevalidation. Each target locus was defined by two outer/flanking sgRNA cut sites (176-981 bp apart) and a third sgRNA

targeting roughly the midpoint (Figure 2.5A). Five days after nucleofection using these sgRNA pools (compared to a non-targeting control sgRNA), deletion production was visualized via PCR using primers flanking the outermost cut sites (Figure 2.5A). The results showed that deletions, spanning either the flanking cut sites or a flank-to-mid cut, were dramatically favored over simple small indels (which are contained in the ~wt-sized band due to lack of separation for size differences of only a few bp) (Figure 2.5B). These results suggest that deletions that are near-precise lengths of <1000 bp can be readily generated using this protocol.

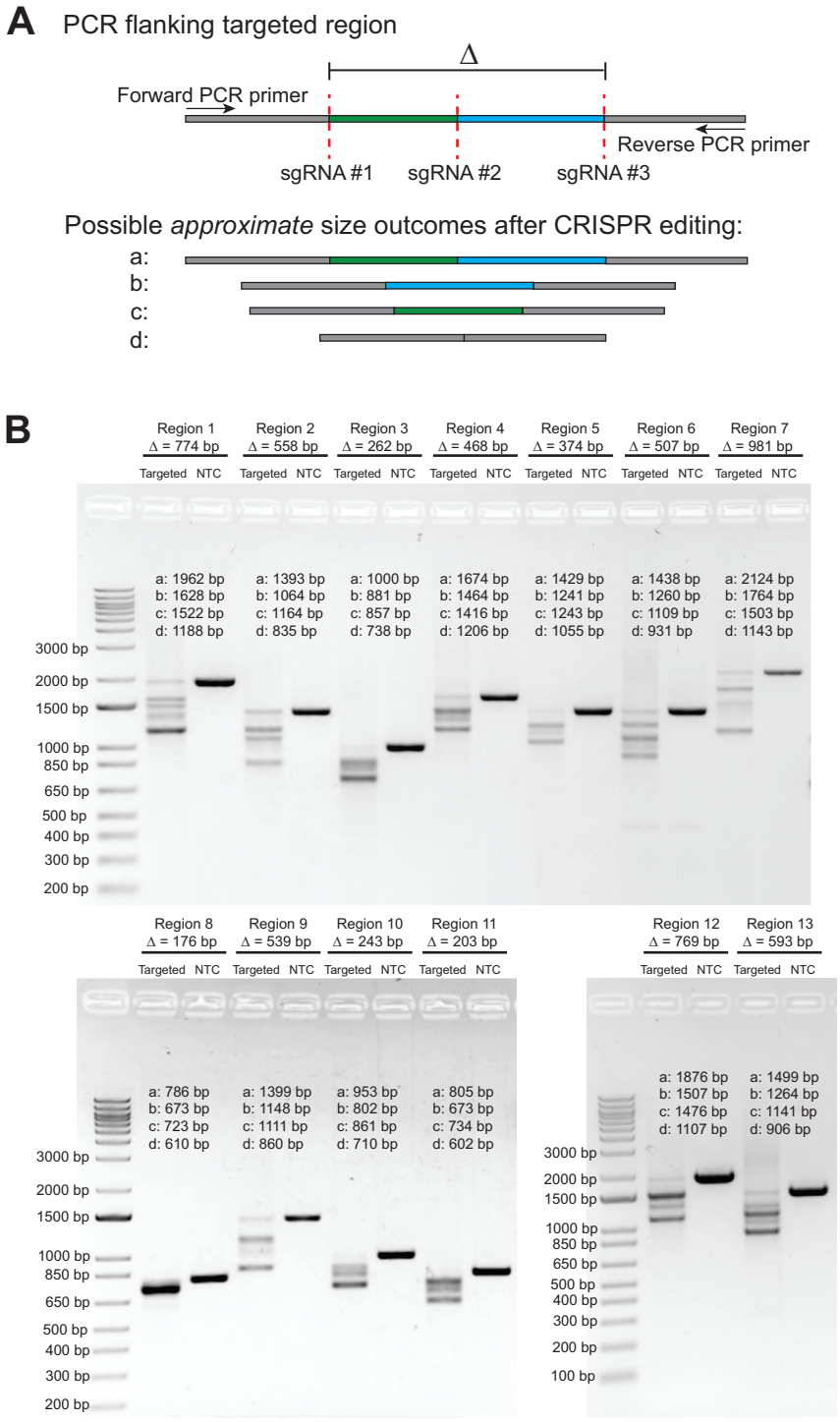


Figure 2.5. CRISPR RNPs allow for deletion of several hundred bp genomic regions. **(A)** Schematic showing targeting strategy for deletion of various genomic regions (not to scale). Three sgRNAs were designed for each region, and sgRNA cut sites are denoted by red dotted lines. Δ denotes largest deletion, between outermost sgRNAs. As shown, PCR amplicons of varying sizes are predicted to result from deletion of the region between all 3 sgRNAs (outcome "d"), deletions between only two of the 3 sgRNAs (outcomes "b" and "c"), and no large deletion (wt allele or only small indels; outcome "a"). **(B)** PCRs for scheme shown in (A) for 13 different

genomic regions, which were each targeted in GSC-0827 cells using a total of 60 pmoles RNPs, with a Cas9:sgRNA ratio of 1:3. Cells were harvested for gDNA extraction 5 days post nucleofection. Each region's PCR is shown for a targeted sample compared to a sample nucleofected with a non-targeting control sgRNA RNP (60 pmoles). Predicted sizes of the various PCR amplicons a-d described in (A), as well as the predicted size of the outermost deletion Δ, are noted for each region.

2.3.3 *RNP nucleofection generates dramatic protein loss in cell pools*

Given the potential of our approach to create highly penetrant multi-allelic KOs in pools, we wanted to further demonstrate its utility by creating a series of KO mutants in human NSCs. Previously, we and others have used ectopic expression of human oncogenes (e.g. *EGFRvIII*, *RasV12, myr-Akt1*, *CCND1*, *CDK4*^{R24C}, dominant-negative *TP53*) to partially or completely transform NSCs[73]. Our current method afforded us the opportunity to affect the same pathways by instead creating loss-of-function mutations in tumor suppressors. We chose to successively target four genes commonly mutated or deleted in adult GBM tumors: *TP53*, *CDKN2A*, *PTEN*, and *NF1*, which affect distinct pathways required for glioma progression, including the p53-pathway, the Rb-axis[153], the PI3-k/Akt pathway[154], and the RTK-Ras-MAPK pathway[155].

We nucleofected TERT-immortalized NSC-CB660 cells with pools of 3 sgRNAs for each gene, in four successive rounds of nucleofection spaced 1 week apart (to allow cells time to recover from electroporation) (Figure 2.6A). In this case, we chose to spread the sgRNAs across each gene to favor individual indels rather than deletions, reasoning that the cumulative effect of 3 sgRNA sites for each gene would lead to a high percentage of cells that contained at least one out-of-frame edit in each allele. We examined the effect on target protein expression via western blotting on the pool at each stage in the process, as well as on eight subclones of the final pool. Remarkably, we observed dramatic protein loss for each gene in the targeted pools (Figure 2.6B). Examination of the eight clones of the final *TP53+CDKN2A+PTEN+NF1* targeted KO pool revealed a similar

result, where all proteins showed similar reduction in individual clones, except for one clone which still showed protein expression of NF1 (Figure 2.6C).

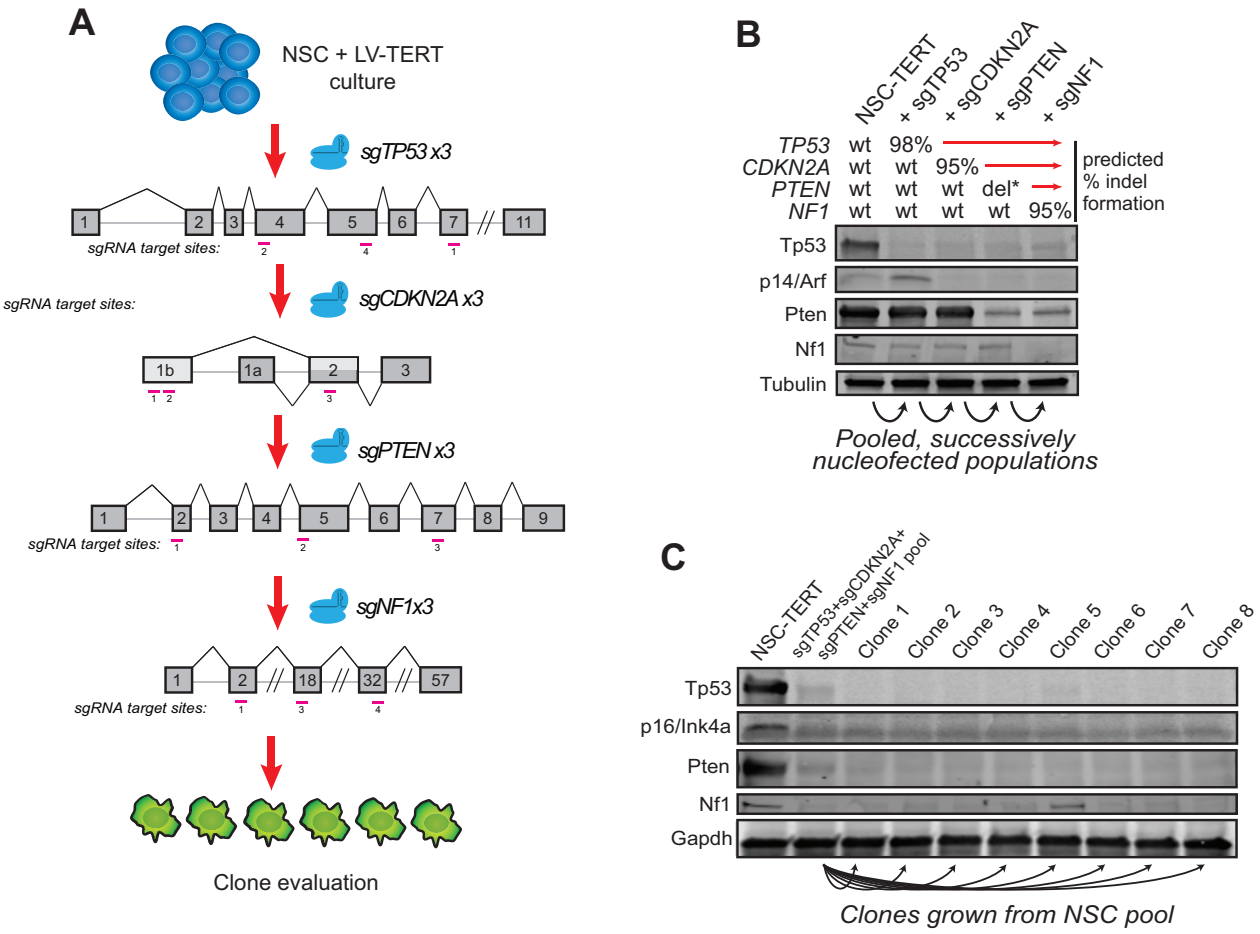


Figure 2.6. Use of Cas9:sgRNA RNPs to generate multi-gene knockout cells to model oncogenic transformation reveals dramatic protein loss. **(A)** Overview of strategy for successive Cas9:sgRNA RNP nucleofections targeting various tumor suppressors in NSCs. **(B)** Western blots for targeted genes in the cell pool at each successive stage. Predicted % indel formation for each gene from Sanger sequencing results is shown above blot. del* indicates that large deletions were identified in addition to indels (detailed in Figure 4). For detection of p53, cells were treated with doxorubin to stabilize the protein. **(C)** Western blots for 8 clones derived from the final sgTP53+sgCDKN2A+ sgPTEN+sgNF1 cell pool.

2.3.4 *RNP nucleofection allows for targeted biallelic deletion of multi-kb genomic regions in cell clones*

We also assessed indel efficiencies in the tumor suppressor KO cell pools across one test sgRNA for each gene. For *TP53*, *CDKN2A*, and *NF1* we observed high predicted indel frequencies of 98%,

95%, and 95%, respectively. For *PTEN*, however, we noted a discrepancy between the Western blot results, which showed near total ablation of protein expression, and the indel analysis for all 3 sgRNAs, which each revealed only ~33% efficiency. Since probability suggests that even the cumulative editing effect of these 3 sgRNAs should not quite account for near total protein loss, we suspected that the *PTEN* sgRNA pool may have allowed for the generation of a large deletion. To investigate this possibility, we performed PCR with primers flanking the outer sgRNA cut sites (Figure 2.7A). In this case, an allele that did *not* harbor deletion of the entire ~64kbp region would not amplify properly in our PCR conditions since the product would be too large. We observed that 2 of 8 clones tested contained a deletion allele, and one of these actually produced two products of similar but distinct sizes, indicating a biallelic deletion with slightly different editing results (Figure 2.7B, left panel). As a positive control for gDNA integrity, we used a second PCR spanning a small *PTEN* region outside the deletion region, and we observed the correct product for all 8 clones in this case (Figure 2.7B, right panel).

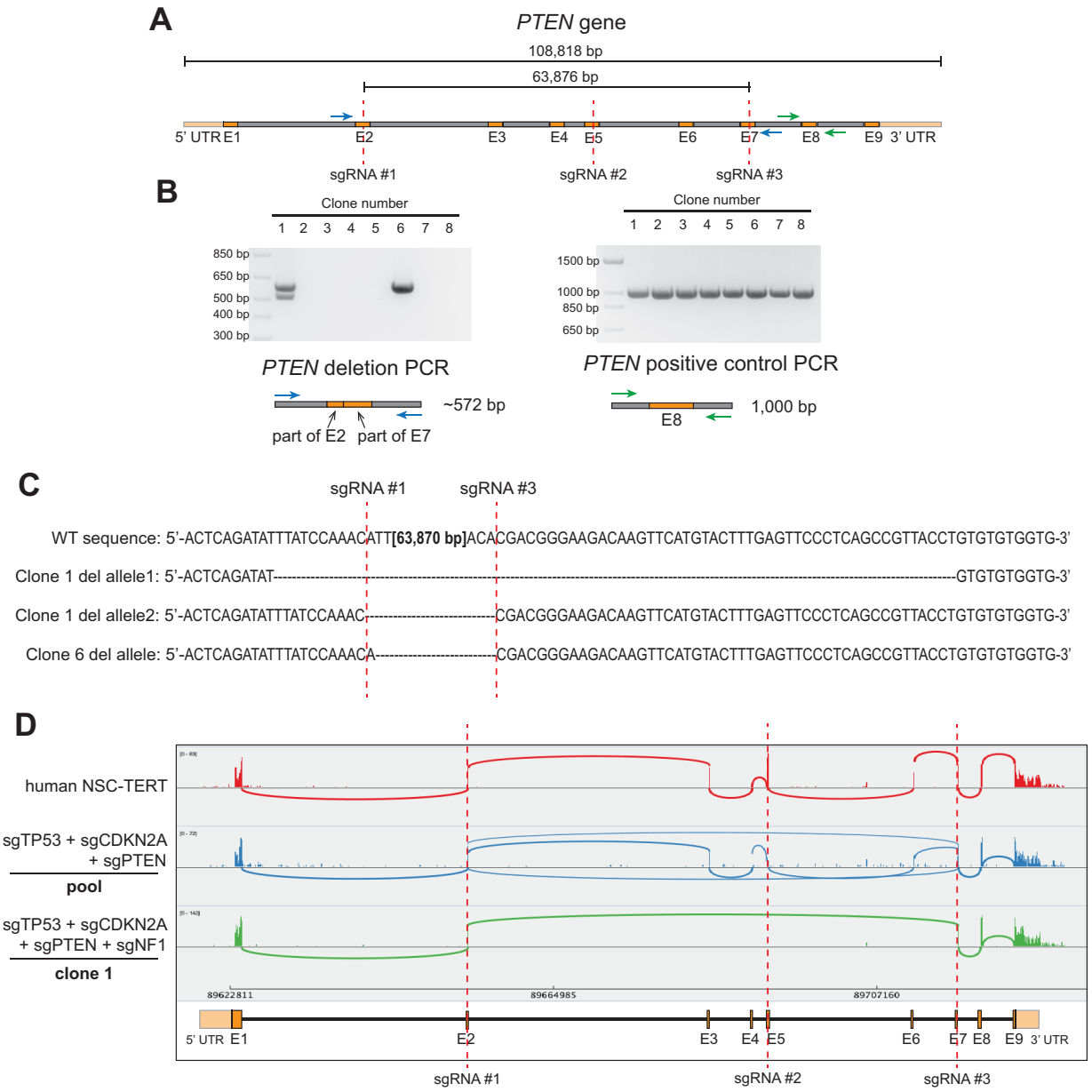


Figure 2.7. Cas9:sgRNA RNPs can be used to generate multi-kb genomic deletions in cell clones. **(A)** The targeting strategy for the *PTEN* gene. Three sgRNAs were designed targeting exons 2, 5, and 7. To check for the potential deletion of the ~64kb segment between the outermost sgRNAs, deletion PCR primers (in blue) were devised to amplify a product only if the entire region had been deleted. Amplification of the region around the non-targeted exon 8 (primers in green) served as a positive control for gDNA integrity. **(B)** Left: Deletion PCR as described in (A) for 8 different clones of transformed NSC-CB660 that had been nucleofected with the 3 sgRNAs targeting *PTEN*. Right: Positive control PCR as described in (A) for the 8 clones. **(C)** Genomic sequences of the 3 deletion alleles identified in (B). Red dotted lines denote sgRNA cut sites. **(D)** A sashimi plot of RNA-seq reads covering the *PTEN* gene for parental NSC cells, targeted pool, and clone 1. "Transcripts" with a minimum junction coverage of 5 reads are shown.

Gel-purification of the deletion PCR bands followed by Sanger sequencing confirmed that each deletion allele observed in the clones was a result of cutting near the predicted outermost sgRNA cut sites, with one allele tested being an exact deletion, one containing an additional 1bp insertion, and one containing an additional 10bp 5'-deletion and 48bp 3'-deletion (Figure 2.7C). To further investigate this, we performed RNA-seq and examined reads mapping to the *PTEN* locus in the KO pool and in clone 1. Analysis of predicted "splice junctions" based on RNA-seq reads showed that mRNA containing the exact deletions observed at the gDNA level could be identified in clone 1, and no properly spliced reads were present, corroborating the fact that this clone did not contain any non-deletion allele (Figure 2.7D). Furthermore, reads corresponding to large deletions could be observed in the cell pool as well, in addition to the expected normally spliced reads. These results suggest that using multiple sgRNAs with our method has the potential to create large (>50kbp) deletions, which may be monoallelic or biallelic in subsets of clones.

2.3.5 *Analysis of gene expression changes induced by tumor suppressor knockout targeting events*

To further assess the fidelity of gene KOs via CRISPR RNP nucleofection, we examined the progressive changes in gene expression after each successive targeting event (sg*TP53*, sg*CDKN2A*, sg*PTEN*, sg*NF1*) in NSC-CB660-TERT cells, by performing RNA-seq on the parental cells compared to the targeted cells at each stage. We used edgeR[156] for differential gene expression analysis, and Figure 2.8A shows the overall relationship of these data in cluster analysis and the gene expression changes after each targeting. The greatest number of changes were produced by *TP53* KO (269 genes with log2 fold-change (log2FC)>0.5 and 682 genes log2FC<-0.5 (FDR<0.05)) and *CDKN2A* KO (1340 genes log2FC>0.5 and 1733 genes log2FC<-0.5 (FDR<.05)). Importantly, ≥80% of *TP53* KO-induced expression changes were maintained (at

FDR<0.05) even after further *CDKN2A* KO, *PTEN* KO, and *NF1* KO, and ≥78% of *CDKN2A* KO induced expression changes were maintained even after further *PTEN* KO and *NF1* KO.

The gene expression changes associated with *TP53* KO cells were consistent with p53's known transcriptional function. We observed downregulation of key p53 transcriptional targets, including 27 of 132 found in Pathway Interaction Database (p=8.2E-20) and 51 of 116 literature-curated p53 targets, including: *BAX*, *BBC3/PUMA*, *BTG2*, *CDKN1A/p21, RRM2B,* and *ZMAT3*[157] (Figure 2.8B-C; Tables S2 & S3). *CDKN2A* KO cells most prominently revealed upregulation of genes associated with the cell cycle (285 genes; GO:0007049; p=6E-116) and specifically *E2F* targets (52 genes; p=2.4E-42), including *E2F1* and *E2F2* themselves (Figure 2.8B-C; Tables S2 & S3). This is consistent with loss of *CDKN2A*'s p16 protein, which inhibits Cyclin D/CDK activity in G1, preventing de-repression of E2F[153]. Other prominent cell cycle regulated genes included those associated with: *CDK1* interactions (67 genes, p=1.7E-28), *MCM6* interactions (33 genes, p=6.1E-19), *PCNA* interactions (64 genes, p=6.2E-28), and *PLK1* interactions (50 genes, p=6.4E-20). Thus, loss of *CDKN2A* leads to profound reprogramming of the transcription of critical cell cycle-regulated genes in human NSCs, consistent with loss of p16 function.

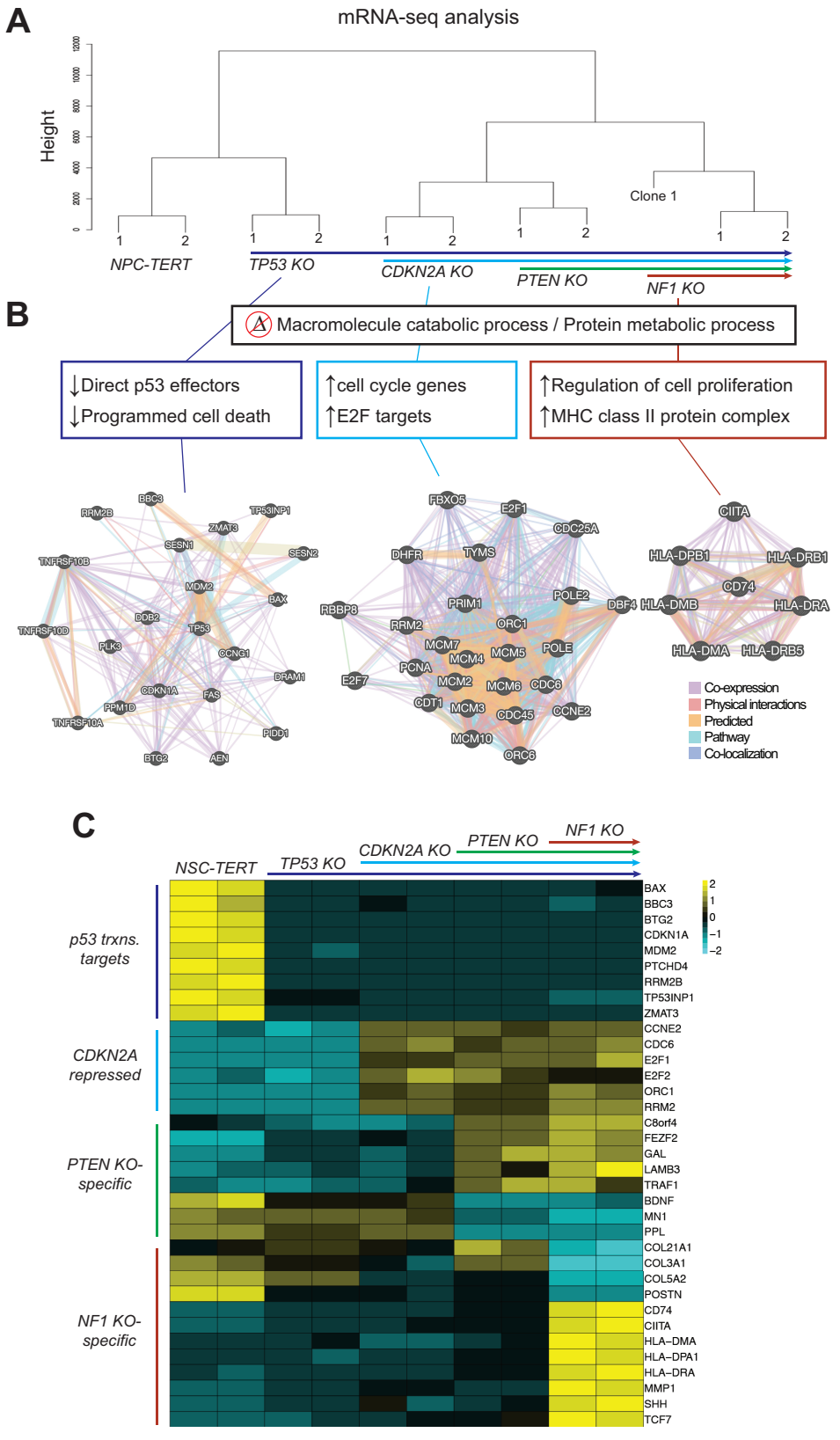


Figure 2.8. Gene expression changes induced by tumor suppressor knockout targeting events in human NSCs. **(A)** Cluster analysis showing relative gene expression profile relationships based on RNA-seq for parental NSC-TERT

cells and NSC-TERT successively nucleofected with sgTP53, then sgCDKN2A, then sgPTEN, and then sgNF1, as well as for clone 1 derived from the final pool where all 4 genes had been targeted. Duplicates were sequenced for all samples except clone 1. **(B)** Networks showing gene relationships for genes altered upon TP53, CDKN2A, or NF1 knockout. **(C)** A heatmap displaying examples of sgTP53, sgCDKN2A, sgPTEN, and sgNF1-specific gene expression changes observed in RNA-seq data. Relative gene expression values were normalized for all samples within each gene. Each RNA-seq replicate is shown as a separate column.

In contrast to *TP53* and *CDKNA*, *PTEN* KO produced the fewest changes in gene expression in our scheme (upregulation of 169 genes and downregulation of 216 genes (+/-0.5 log2FC, FDR<0.05)). This may be due to epistasis with gene expression changes caused by *CDKN2A* KO. Nonetheless, manual curation of these genes revealed possible connections to the PI-3 kinase pathway itself, suggestive of feedback regulation. For example, within the downregulated genes, *PPL/Plakin* binds Akt directly[158], and brain-derived neurotrophic factor (*BDNF*) activates Akt[159] and Pten itself[160]. Within the upregulated genes, *GAL/galanin* codes for neuroendocrine peptide that exhibits an autocrine mitogenic effect through Erk and Akt activity[161] (Figure 2.8C). In addition, there were many novel genes affected by *PTEN* loss, including, among others: *C8orf4/TCIM*, a positive regulator of the Wnt/beta-catenin pathway[161]; *FEZF2*, a marker and transcription factor associated with NSC-dependent patterning of the cerebral cortex[162]; and *TRAF1*, which mediates the anti-apoptotic signals from TNF receptors[163] (Figure 2.8C).

With the further addition of *NF1* KO in *TP53+CDKN2A+PTEN* KO cells, we observed changes in expression of an additional 1022 genes (321 up- and 701 downregulated; log2FC+/-0.5; FDR<0.05). The downregulated genes were most significantly enriched for extracellular matrix organization genes (43 genes; GO:0030198; p=1.8E-15), which included many collagen encoding genes and Periostin (Figure 2.8C). Upregulated genes included genes involved in regulation of cell proliferation (43 genes, p=5.2E-6), including members of the WNT signaling pathway (*WNT7B*, *FZD4*, *LEF1*, and *TCF7*). Interestingly, also upregulated were eight genes

coding for major histocompatibility complex (MHC) class II proteins, including *CIITA*, the master transcriptional activator controlling expression of the MHC class II genes[164] (Figure 2.8B-C; Tables S2 & S3). *CIITA* expression was upregulated ~10-fold in *NF1 KO* cells relative to NSC-CB660-TERT cells. MHC class II genes are primarily expressed by professional antigen presenting cells, such as dendritic cells, macrophages, and B cells[165]. However, remarkably, upregulation of MHC class II protein complex is a hallmark of *NF1*$^{-/-}$ human neurofibroma tumor cells, which *CIITA* activity is required to maintain[164,166]. It has also been observed that gliomas and other cancers have a high proportion of MHC class II-expressing tumor cells[167], which may promote tolerance to tumor-associated antigens[168]. Our results show that *NF1* loss is one route to de-repression of MHC class II machinery in human NSCs.

In addition to gene expression changes in response to loss of *TP53, CDKN2A, PTEN,* and *NF1*, we also identified 806 genes whose expression remained largely unaltered in all conditions (FDR≥0.2 across all comparisons with ≥10 CPM counts across all samples). Gene set enrichment analysis revealed that 487 of these were involved in a "cellular metabolic process" (GO:0044237; FDR=1.5E-15), with 129 encoding mitochondrial proteins (including 19 involved in oxidative phosphorylation) and 29 encoding ribosomal proteins (Tables S2 & S3). Our previous CRISPR-Cas9 lethality screens in NSC-CB660 cells demonstrate that at least 208 of these genes score as essential, including, for example, 34 genes associated with ribosome biogenesis, 11 genes coding for respiratory electron transport machinery, and 23 overlapping with "housekeeping" genes[169] (GSEA: 111197; p= 7.12E-13). These results suggest that a subset of genes expressed in human NSCs are transcriptionally regulated and/or maintain their mRNA levels independently of p53, Rb-axis, PI-3 kinase, and Nf1 pathways.

Altogether, our gene expression results highlight the utility of our CRISPR/Cas9 RNP nucleofection method for quickly creating a series of KOs that allow for the study of gene activities. Importantly, due to the high targeting efficiency using our method, we were able to confirm many known as well as identify novel transcriptional changes associated with loss of the genes we targeted, using nucleofected cell *pools* rather than cell clones.

2.4 DISCUSSION

Here, we present a method for creating bi- and multi-allelic loss-of-function indel mutations, using *in vitro* assembled Cas9:sgRNA RNPs composed of chemically synthesized, modified sgRNA and purified Cas9 protein. By this method, indel efficiencies of >90% can routinely be achieved in populations of cells, obviating the need for clonal selection of edited cells or chemical selection of DNA-based sgRNA expression systems. Moreover, because gene editing is complete within three days of RNP introduction, this approach offers better experimental tractability over current approaches, which can suffer from lack of indel penetrance and protracted windows of indel formation, as well as the potential for increased off-target effects that can occur with constitutive expression of Cas9[170].

We illustrate the utility of our method in primary human NSCs by knocking out four tumor suppressor loci in quick succession: *TP53*, *CDKN2A*, *PTEN*, and *NF1*. This combination of tumor suppressor gene KOs mimics common loss-of-function alterations in the p53-pathway, the Rb-axis, the PI3K/Akt pathway, and the RTK-Ras-MAPK pathway, which are observed in the mesenchymal glioblastoma subtype[37]. Remarkably, our results demonstrate that multiple tumor suppressor KOs can quickly be created with >95% efficacy in cell pools when using combinations of 3 sgRNAs. Thus, this approach offers the potential to oncogenically transform normal diploid

cell types without the need for ectopic expression of oncogenes, which may better capture cancer-specific biology.

This method also improves upon existing methods for the creation of precise or near-precise deletions. In mammals, single dsDNA breaks, including those generated by Cas9:sgRNA cutting, produce NHEJ-dependent small insertions or deletions at break sites (i.e., error-prone repair)[138,171]. However, adding a second dsDNA break in close proximity to the first can cause rejoining without error via "accurate" NHEJ[172,173]. Our results support this notion. We observe a high frequency and penetrance of conversion of single indels to precise and near-precise deletions. It has been previously shown that using 2-3 proximal sgRNAs can create deletions of ~10bp to 1Mb in mouse embryos and cultured cells[133,172,174,175]. However, these approaches produce deletion formation efficiencies ranging from ~2% to ~40%. By contrast, we observe near-precise deletion frequencies as high as >90% using sgRNAs spaced up to 1000bp apart. This suggests that the use of 2 sgRNAs using our approach can have the added benefit of triggering accurate NHEJ and being able to specify a high frequency of precise or near-precise deletions.

This technique does have a few limitations. First, achieving high multi-allelic indel efficiencies may require pre-validation of sgRNAs. However, we have had good success designing sgRNAs using the Broad GPP Web Portal[152] design tool (where ~60-70% of sgRNAs that we choose via manual curation produce >80% indel formation) or, alternatively, choosing sgRNA sequences from positively scoring CRISPR-Cas9 lentiviral-based screen hits. Second, although we can specify deletions using this approach, the repair of the dsDNA break(s) can at times be unpredictable, as illustrated with the case of *PTEN* where using a pool of widely spaced sgRNAs unexpectedly produced ~64kb deletions. Third, reliance on chemically synthesized sgRNAs can be cost-prohibitive for large-scale projects. An alternative option is to generate *in vitro* transcribed

sgRNA using T7, T3, or SP6 RNA polymerase in the presence of ribonucleoside triphosphates and a DNA template[132]. However, this requires additional steps, namely the initial creation of an accurate DNA template and the purification of the transcribed sgRNA to remove unincorporated triphosphates, enzyme, and template DNA. In addition, *in vitro* transcription can result in errors toward the 5' end of the RNA[176]. Also, it is not possible to easily generate modifications, meaning *in vitro* transcribed sgRNA does not possess increased protection from nucleases once it has entered the cell, resulting in decreased editing efficiency compared to chemically synthesized, modified sgRNA. Nonetheless, it is a viable alternative to consider when cost is a concern. Another alternative is to employ a commercially sourced dual gRNA system (crRNA:tracrRNA), which may represent only a slight reduction in efficiency. The two chemically synthesized RNAs can still be modified to enhance nuclease resistance, but they are usually available at a lower cost since the accurate synthesis of these shorter RNAs is less complex compared to a longer chimeric sgRNA.

Our data suggest that the relatively simple method described here can be used for highly efficient (>90%) and fast (72 hours) gene knockout, as well as for targeted genomic deletions, even in hyperdiploid cells (such as many tumor cells). This represents an extremely useful tool for inactivating not only coding genes, but also non-coding elements such as non-coding RNAs, UTRs, enhancers, and promoters. The gain in efficiency that we observe can allow for systematic well-by-well screening (similar to small interfering RNA screens), but provides the flexibility of targeting any small element in the genome. This method can be readily applied to diverse mammalian cell types by varying the nucleofection buffer and program (Lonza can provide appropriate conditions for many cell types). Thus, it represents an important step forward in the

ability to manipulate the genomes of cell cultures derived from primary cells, such as patient-derived tumor cells and human stem/progenitor cells.

2.5 METHODS

2.5.1 *CRISPR sgRNA Design*

CRISPR sgRNAs were designed via manual curation of all possible sgRNA sequences for a given region as identified by the Broad Institute's GPP Web Portal[152]. See Table S1 for a list of sgRNA sequences used.

2.5.2 *Cas9:sgRNA RNP Nucleofection*

See the detailed protocol at the end of this chapter. Briefly, to prepare RNP complexes, reconstituted sgRNA (Synthego) and then sNLS-SpCas9-sNLS (Aldevron) were added to complete SG Cell Line Nucleofector Solution (Lonza), to a final volume of 20 μL. The mixture was incubated at room temperature for 15 minutes to allow RNP complexes to form. A Cas9:sgRNA molar ratio of 1:2 was used, unless otherwise noted. Total RNP doses described refer to the amount of the limiting complex member (Cas9). To nucleofect, 1.5 x 10^5 cells were harvested, washed with PBS, resuspended in 20 μL of RNPs, and electroporated using the Amaxa 96-well Shuttle System or 4D X Unit (Lonza) and program EN-138.

2.5.3 *CRISPR Editing Analysis*

Nucleofected cells were harvested at indicated timepoints and genomic DNA was extracted (MicroElute Genomic DNA Kit, Omega Bio-Tek). Genomic regions around CRISPR target sites were PCR amplified using Phusion polymerase (Thermo Fisher) and primers located (whenever possible) at least 250bp outside cut sites. After size verification by agaorse gel electrophoresis,

PCR products were column-purified (Monarch PCR & DNA Clean-up Kit, New England BioLabs) and submitted for Sanger sequencing (Genewiz) using unique sequencing primers. The resulting trace files for edited cells versus control cells (nucleofected with non-targeting Cas9:sgRNA) were analyzed for predicted indel composition using the Inference of CRISPR Edits (ICE) web tool[151]. See Table S1 for a list of all PCR and sequencing primers used.

2.5.4 *Cell Culture*

Patient tumor-derived GSCs and fetal tissue-derived NSCs were provided by Drs. Do-Hyun Nam, Jeongwu Lee, and Steven M. Pollard, were obtained via informed consent, and have been previously published[13,32,177]. Isolates were cultured in NeuroCult NS-A basal medium (StemCell Technologies) supplemented with B27 (Thermo Fisher), N2 (homemade 2x stock in Advanced DMEM/F-12 (Thermo Fisher))[32], EGF and FGF-2 (20 ng/ml) (PeproTech), glutamax (Thermo Fisher), and antibiotic-antimycotic (Thermo Fisher). Cells were cultured on laminin (Trevigen or in-house-purified) -coated polystyrene plates and passaged as previously described[32], using Accutase (EMD Millipore) to detach cells.

2.5.5 *Western Blotting*

Cells were harvested, washed with PBS, and lysed with modified RIPA buffer (150mM NaCl, 25mM Tris-HCl (pH 8.0), 1mM EDTA, 1.0% Igepal CA-630 (NP-40), 0.5% sodium deoxycholate, 0.1% SDS, 1X protease inhibitor cocktail (complete Mini EDTA-free, Roche)). Lysates were sonicated (Bioruptor, Diagenode) and then quantified using Pierce BCA assay (Thermo Fisher). Identical amounts of proteins (20-40μg) were electrophoresed on 4–15% Mini-PROTEAN TGX precast protein gels (Bio-Rad). For transfer, the Trans-Blot Turbo transfer system (Bio-Rad) with nitrocellulose membranes was used according to the manufacturer's instructions. TBS (137mM

NaCl, 20mM Tris, pH 7.6) +5% nonfat milk was used for blocking, and TBS+0.1%Tween-20+5% milk was used for antibody incubations. The following commercial primary antibodies were used: Tp53 (Cell Signaling #48818, 1:500), p14/Arf (Bethyl Laboratories #A300-340A-T, 1:500), p16/Ink4a (Abcam #ab16123, 1:200), Pten (Cell Signaling #9559S, 1:1,000), Nf1 (Santa Cruz #sc-67, 1:50), αTubulin (Sigma #T9026, 1:1,000), Gapdh (Sigma #SAB2500450, 1:100). The following secondary antibodies were used (LI-COR): #926-68073, #926-32212, #926-32214, #926-68074. An Odyssey infrared imaging system (LI-COR) was used to visualize blots.

2.5.6 *RNA-seq and Analysis*

Cells were lysed with Trizol (Thermo Fisher). Total RNA was isolated (Direct-zol RNA kit, Zymo Research) and quality validated on the Agilent 2200 TapeStation. Illumina sequencing libraries were generated with the KAPA Biosystems Stranded RNA-Seq Kit[178] and sequenced using HiSeq 2000 (Illumina) with 100bp paired-end reads. RNA-seq reads were aligned to the UCSC hg19 assembly using STAR2 (v 2.6.1)[179] and counted for gene associations against the UCSC genes database with HTSeq[180]. Normalized gene count data was used for subsequent hierarchical clustering (R package ggplot2[181]) and differential gene expression analysis (R/Bioconductor package edgeR[156]). Heatmaps were made using R package pheatmap[182].

2.6 In-depth Protocol for Cas9:sgRNA Ribonucleoprotein Nucleofection Using Amaxa 4D Nucleofector X Unit or 96-well Shuttle System

The goal of this protocol is to produce bi- or multi-allelic loss of gene function in cell populations via insertion-deletion (indel) or near-precise deletion mutations triggered by CRISPR-Cas9 targeting. This protocol represents our current preferred method, which utilizes *in vitro*-formed

Cas9:sgRNA ribonucleoprotein (RNP) complexes composed of one to three chemically synthesized 2'-O-methyl 3'phosphorothioate-modified sgRNAs and purified Cas9 protein. We currently purchase our sgRNAs from Synthego (https://www.synthego.com/). These sgRNAs are a single RNA molecule that contains both the custom-designed short crRNA sequence fused to the scaffold tracrRNA sequence. Synthego's 2'-O-methyl 3'phosphorothioate-modifications occur in the first and last 3 nucleotides. Other companies also provide sgRNA synthesis services, including IDT, Dharmacon, Thermo Fisher Scientific, and Sigma-Aldrich. In addition, sgRNAs can be synthesized *in vitro*, e.g. using *in vitro* transcription reactions and short DNA templates (although this usually results in inferior editing efficiency).

The Amaxa Nucleofector System uses proprietary electrical parameters and buffer formulations. This protocol is optimized for glioblastoma stem-like cells (GSCs) and human neural stem cells (NSCs) (with nucleofection program based on conditions described for mouse NSCs in Bressan et al. 2017). However, we and collaborators have successfully employed CRISPR in other cell types as well using this method. *Nucleofection conditions (i.e. nucleofection buffer and program) should be optimized for other cell types.* Lonza can provide guidance for many cell lines/types (https://knowledge.lonza.com/).

2.6.1 *Designing sgRNAs*

a. There are many sgRNA design resources available. We typically use the Broad Institute's GPP Web Portal[152] (https://portals.broadinstitute.org/gpp/public/analysis-tools/sgrna-design) in combination with Synthego's design tool (https://design.synthego.com/#/) to identify potential top sgRNAs for a particular gene. We then manually curate these lists by evaluating on-target score versus predicted off-target effects in order to identify ~3-4 top sgRNAs per gene for experiments. In our experience, when using nucleofection of RNPs,

about 60-70% of such sgRNAs result in high editing efficiency (>80% indel efficiency). Total editing efficiency can usually also be increased by simultaneously nucleofecting with multiple sgRNAs targeting the same gene.

b. If a near-precise deletion is desired, we typically design 2 sgRNAs spaced ~50-300bp apart. We have tested up to 1000bp spacing and this works well, although efficiencies for deletions <300bp can be more easily assessed (compared to larger deletions), via simple Sanger sequencing. When the goal is a gene KO, we usually design both sgRNAs to be in exons (or the same exon), so that even alleles that do not receive a deletion still have a good chance of resulting in gene KO via small indels.

c. By necessity, available sgRNA design tools list possible sgRNA sequences based on a reference genome for the species you are using. It is a good idea to check the genomic sequence of your particular cells for polymorphisms (and mutations in the case of cancer cell lines) that may exist and may cause some sgRNA sequences created by design tools to have mismatches in your particular cells. In other words, it is best to ensure that the exact sgRNA sequence you decide to use is actually intact in your cells' genome.

2.6.2 *CRISPR RNP Nucleofection*

Note: Before starting, wipe down workspace and equipment with RNaseZap (Thermo Fisher) to inhibit RNases that may degrade your sgRNA. Wear clean gloves and use pipet tips designated for RNA work only.

a. Reconstitute your synthetic RNA (and dilute it if desired):

   i) Briefly centrifuge your tubes or plates containing synthetic modified single guide RNA (sgRNA) oligos to ensure that the dried RNA pellet is collected at the bottom.

   ii) Carefully dissolve sgRNA in the provided nuclease-free 1X TE Buffer (Tris-EDTA,

pH 8.0) by adding 15µl buffer to 1.5nmol of dried sgRNA. (This will result in a final concentration of 100µM (100pmol/µl)). Flick the bottom of the tube gently so the liquid covers the bottom, then pulse vortex the tube for ~30 seconds and quick spin reconstituted RNA down. Dissolved RNA should be stored at -20 °C when not in use. Under these conditions, RNA will be stable for at least one year.

iii) If desired, to create a working stock, make a 30µM sgRNA dilution (e.g. 14µl nuclease-free water + 6µl of reconstituted 100µM sgRNA, creating a total of 20µl of 30µM (30pmol/µl) sgRNA). Use diluted sgRNA immediately or store at -20 °C for up to three months. Keep RNA on ice while in use. (If you will make larger master mixes of RNPs using the same sgRNA(s), it may not be necessary to dilute your sgRNA. sgRNAs are diluted primarily to allow for easier pipetting volumes.)

⇒ *Tip: If reduced editing efficiencies are experienced over time, this may be due to improper handling and/or storage of RNA. It is best to store RNA in 1X TE Buffer long-term and to dilute with water only the amount required in the short-term. Freeze-thawing of RNA should also be kept to a minimum.*

b. Dilute Cas9:

i) Briefly centrifuge the stock Cas9 (Aldevron sNLS-SpCas9-sNLS, 61 µM) to ensure that all liquid is at the bottom. To make a working stock, make a 1:5 Cas9:PBS dilution (e.g. 10µl PBS (pH 7.4) + 2µl Cas9, creating 12µl of 10.17µM (10.17pmol/µl) Cas9). Keep diluted/undiluted Cas9 on ice. Only make the amount of diluted Cas9 that you will need at any given time since it is best not to freeze the diluted Cas9.

⇒ *Tip: It is best to dilute Cas9 only in PBS (pH 7.4), not directly in Nucleofector solution. We have found that concentrated Cas9 can precipitate when added directly to some buffers.*

⇒ *Tip: We have tested several commercial sources of purified Cas9 protein and have found that Aldevron's product provides the highest editing efficiency at the lowest dose. Other sources or in-lab purified Cas9 can also be used but may require a higher molar dose for equivalent editing efficiency. Note also that we have compared purified Cas9 proteins with three slightly different versions of nuclear localization signal (NLS) location: sNLS-SpCas9-sNLS, 2XNLS-SpCas9-sNLS, and sNLS-SpCas9-2XNLS. We found that these three varieties performed identically in our cells.*

c. Assemble RNP complexes:

   i) Create complete Nucleofector Solution by mixing SG Cell Line Solution (or other solution specific to your cell type) with the provided Supplement at a 4.5:1 ratio (e.g. for one nucleofection, combine 16.4µl Cell Line Solution + 3.6µl Supplement). Make a master mix for multiple nucleofections. (Once combined, the complete Nucleofector Solution is stable for three months.)

   ii) Assemble RNP complexes in complete Nucleofector solution, adding reagents in the order shown below. The amounts below will allow for 15 pmol total RNPs to be nucleofected into one cell sample (with 10% extra volume), at an sgRNA:Cas9 ratio of 2:1. You may make a master mix for multiple nucleofections using identical RNPs.

ul complete Nucleofector Solution	19.28 µl
ul Synthego gRNA (**30 pmol/ul**)	1.10 µl
ul Cas9 (**10.17 pmol/ul**)	1.62 µl
Total	22 µl

⇒ *Tip: We have also tested higher sgRNA:Cas9 molar ratios up to 9:1 and have not found an increased ratio to result in significantly higher editing efficiency in NSCs and GSCs. However, it is possible that other cell types may benefit from an increased ratio, and this should be empirically determined. Similarly, the optimal total RNP dose should also be empirically determined for other*

cell types.

iii) Mix by pipetting and briefly centrifuge the sample so all the liquid is at the bottom.

iv) Incubate RNPs for 10-20 minutes at room temperature and then keep on ice (or at 4°C) until shortly before use. Immediately before using complexed RNPs to resuspend cells (see below), make sure to re-equilibrate RNPs back to room temperature.

d. Collect and prepare cells:

i) Prepare a 12-well or 6-well dish (pre-coated if necessary, following your standard culturing procedure) so it contains appropriate culture media and pre-warm the plate by placing in a 37°C incubator until you are ready to plate your nucleofected cells. (The size dish that the cells will be plated into after nucleofection should be empirically determined so that your cells are ~30-50% confluent the day after nucleofection. This will vary depending on the size of your particular cells and how many die due to the nucleofection process.)

ii) Harvest cells normally and count. For each sample to be nucleofected, aliquot 1.5-2 x 10^5 cells into a separate 1.5 mL microcentrifuge tube and centrifuge at 300g for 5 minutes.

iii) Resuspend cells in PBS to wash, and centrifuge at 300g for 7 minutes.

iv) Aspirate the liquid as completely as possible. Resuspend the cell pellet in 20μL of RNP complexes. Work quickly, but carefully, and avoid leaving cells in Nucleofector Solution for longer than 30 minutes total (from the time you resuspend them to the time nucleofection is complete). Avoid bubble formation.

v) Transfer all of the cell-RNP solution to one well of a 16-well Nucleocuvette strip (placed inside the supplied "plate skeleton" if using the 96-well Shuttle System), and

cover with the provided lid (either strip lid or plate lid if using the Amaxa 4D Nucleofector X unit or the 96-well Shuttle Unit, respectively). Make sure there are no bubbles in your Nucleocuvette.

e. Nucleofect cells:

i) If using the Amaxa 4D Nucleofector X unit, turn on the core unit and then use the touch screen to select the appropriate options for nucleofecting a 16-well strip. Alternatively, if using the Amaxa 96-well Shuttle Unit, turn on the Amaxa core unit, then the 96-well Shuttle Unit, and then the laptop computer attached to the 96-well Shuttle Unit. Then open the 96-well Shuttle software.

ii) Visually inspect the Nucleocuvette Vessel to make sure that the sample covers the bottom of the cuvette and that there are no bubbles in the cuvette. If you notice problems, gently tap the Nucleocuvette Vessel against your hand and/or use a thin sterile pipet tip to pop any bubbles.

iii) If using a 16-well Nucleocuvette strip, insert the strip into the open Nucleofector 4D X unit. Make sure the larger gap in the strip lid is at the top of the strip rather than at the bottom, so that the yellow indicator in the X unit fits through the large gap at the top of the lid. Alternatively, if using the 96-well Shuttle Unit, place the Nucleocuvette Vessel with plate lid into the retainer of the 96-well Shuttle Unit. Check for proper orientation of the plate (A1 should be at the top left).

iv) Run nucleofection program EN-138 (or other cell-type specific program). After run completion, the screen should display a "+" over the wells that were successfully electroporated. Remove the cuvette strips/plate from the machine.

v) Carefully resuspend the cells in each well of the Nucleocuvette with 80µl of pre-

warmed growth media, and mix very **slowly and gently** by pipetting up and down 3-4 times.

⇒ *Tip: Excessive pipetting can greatly increase cell death, so keep pipetting to a minimum. Furthermore, for some cell types Lonza recommends incubating the Nucleocuvettes at room temperature or 37℃ for ~20 minutes immediately after nucleofection (before resuspending with media), as this can greatly increase cell survival. We have not found this to be necessary for NSCs and GSCs, but it may be important for other cell types.*

vi) Transfer all 100µl to one well of the pre-warmed, media-containing cell culture plate and gently pipet up and down a few times.

vii) Shake the plate to distribute the cells. Leave the plate at room temperature for ~20-25 minutes and then place in incubator. (This helps distribute cells evenly.)

⇒ *Tip: Cell survival is usually highest if you keep the cells somewhat dense after nucleofection and do not plate them too sparse, i.e. plate them into an appropriate well/plate so they are ~30-50% confluent the day after nucleofection. Exact cell numbers to use and size well they should be plated into should be empirically determined for your particular cell type.*

f. Change the media 12-24 hours after nucleofection to remove dead cells and debris.

g. Three days after nucleofection, cells may be collected for indel analysis and/or phenotypic analysis, or cells may be expanded if creating a stable line. If studying an essential gene, phenotypic analysis may need to be done sooner as the majority of editing happens within 24 hours and the stability of the particular remaining protein will determine how quickly a phenotype can be observed.

⇒ *Tip: If absolutely necessary, cells can be passaged only 24 hours post nucleofection. However, this should be avoided if possible, as we have found that it results in reduced cell health and proliferation.*

2.6.3 *Editing Efficiency Analysis*

a. Harvest nucleofected cells *and control cells* (nucleofected with RNPs containing a non-targeting sgRNA or mock nucleofected) according to normal culturing procedure. Cell pellet can either be used immediately or stored at -20°C or -80°C for several weeks before gDNA extraction.
b. Extract gDNA using a column purification kit, according to the manufacturer's protocol. We typically use the MicroElute Genomic DNA Kit (Omega Bio-Tek).
c. Perform PCR amplification of the region around the sgRNA cut site(s):
    i) Whenever possible, design PCR primers to be ~250-350bp outside of the sgRNA cut site. If multiple sgRNAs are being used, design primers to be ~250bp outside of the outermost cut sites. We usually devise a ~1000bp (wt) amplicon. Use webtools such as Primer3 and NCBI Primer-Blast to check primer parameters and check for off-target amplification.
    ii) Design one or multiple sequencing primers that are ~20-80bp inside either end of the PCR amplicon.
    iii) Perform PCR amplification (50µl reaction) on the edited and control samples using a high-fidelity polymerase such as Phusion (New England BioLabs), following manufacturer's suggestions for PCR conditions for the particular enzyme. We typically include a final concentration of 3% DMSO (which aids with template denaturation) in

our PCR reactions with gDNA templates. (See Table S1 for details on primer sequences, annealing temperatures, and extension times that we used for the sgRNAs described in our study.)

iv) Run 10µl of the PCR reactions on a 1% agarose gel and visualize with ethidium bromide staining. Check that there is only one product of the correct size in the control sample (the edited sample may have multiple bands if your RNPs created one or more deletions). If you have multiple products in the control sample, design new PCR primers and repeat the amplification.

⇒ *Tip: Sometimes it is simply not possible to identify PCR primers that only produce one band. If you have tried multiple primer sets and this is the case, you may be able to use a primer set that creates multiple bands. We have found that as long as the correctly-sized band is by far the most intense band, the following steps using simple column purification of the PCR product can still work properly, as long as a unique internal sequencing primer is used (a sequencing primer that does not overlap significantly with either PCR primer). Alternatively, if your best PCR primers produce multiple bands, you can also employ gel purification and only sequence a specific band. However, we do not recommend this since you may not sequence certain indels that are present in your edited sample if you only select a specific portion of the gel to sequence (rather than sequencing everything present in your PCR reaction).*

d. Use a column PCR purification kit to purify the remainder of the edited and control sample PCR products (~40µl), according to the manufacturer's protocol. We typically use the Monarch PCR & DNA Clean-up Kit (New England BioLabs).

e. Perform Sanger sequencing on the PCR amplicons for your edited and control samples.

f. Perform editing analysis using Sanger trace files:

i) Using a freely-available webtool such as Inference of CRISPR edits (ICE, https://ice.synthego.com/#/), enter the trace files (.ab1 files) for your edited and control samples, as well as the sequence(s) of the sgRNA(s) that you used. ICE can be used for samples where up to 3 sgRNAs were used simultaneously.

ii) ICE results will display information including your predicted total editing efficiency (called "ICE score"), predicted KO score (percent of predicted sequences that result in a frameshift or an indel ≥21 bp in length), and an R^2 value that describes "goodness of fit". Furthermore, the output allows visualization of indel size distribution, a list of the sequences predicted in your cell pool (in order of frequency of occurrence), and a discordance plot showing where (or if) your edited trace deviates from your control trace. Find more detailed explanation of outputs at https://www.synthego.com/guide/how-to-use-crispr/ice-analysis-guide.

2.6.4 *Supplies Required for Nucleofection*

The following list does not include cell culture reagents. Also note that SG Cell Line Solution should be used for GSCs and NSCs (and most likely other neural cell types as well), but a different solution may be optimal for different cell types.

Item	*Company*	*Catalog number*
2'-O-methyl 3'phosphorothioate-modified, chemically synthesized sgRNA	Synthego	N/A ***
1X TE Buffer (Tris-EDTA, pH 8.0)	Synthego	Provided with sgRNA kit
Nuclease-free water	Synthego	Provided with sgRNA kit
sNLS-SpCas9-sNLS nuclease	Aldevron	9212-0.25MG
Phosphate buffered saline, pH 7.4	homemade or e.g. Fisher	e.g. 10010-023
SG Cell Line 4D-Nucleofector™ Kit (20µl format)	Lonza	V4XC-3032 (32 reactions) V4SC-3096 (96 reactions)

Note that for this protocol we necessarily used Lonza's instruction manual for the Amaxa 4D Nucleofector as a reference. See also Synthego's CRISPR RNP nucleofection protocol.

Chapter 3. IDENTIFICATION OF GENE DEPENDENCY GROUPS AND CONTEXT-SPECIFIC GENETIC VULNERABILITIES IN GLIOBLASTOMA AND OTHER CANCERS USING CRISPR-CAS9 SCREENS

3.1 ABSTRACT

Glioblastoma (GBM) is the most common and aggressive form of brain cancer and remains one of the deadliest cancers. Although directly targeting oncogenic driver mutations is a popular strategy in developing new therapeutics, it has been met with limited success. Thus, we take the alternative, unbiased approach of using functional genomic screens to identify molecular vulnerabilities that arise in cancer cells due to the oncogenic state. We previously performed genome-scale RNAi and CRISPR-Cas9 pooled outgrowth screens in human nontransformed neural stem cells (NSCs) and patient-derived GBM stem-like cells (GSCs). Here, in order to further characterize candidate GSC-specific hits, we performed comprehensive pooled retests of all putative screen hits that scored preferentially in GSC isolates in our whole-genome CRISPR-Cas9 screens. We then used this data to identify both GBM gene dependency groups and context-specific vulnerabilities, which was facilitated by incorporating published functional genomic screening data (DepMap database) into the analyses. By creating co-dependency networks, we identified GSC-specific gene vulnerability groups related to mitochondrial protein processing and turnover and membrane trafficking as well as metabolic enzymes and regulators. Furthermore, we identified predicted context-specific vulnerabilities and further investigated the dsRNA-editing enzyme ADAR and the adapter protein EFR3A. In addition, we explored the particularly strongly GSC-specific screen hit FBXO42, which we found to be essential in a subset of GSCs and cell

lines of various cancer types, but nonessential in NSCs, indicating this target's potential for a large therapeutic window and broad applicability. We demonstrated that the ubiquitin ligase role of FBXO42 is responsible for the viability phenotype, but likely not through p53 as a previous study might suggest. In searching for possible interactors, we determined that FBXO42 and CCDC6 likely work together rather than redundantly to promote viability in FBXO42 loss-sensitive cells. Lastly, we demonstrated that the viability loss results from an extended metaphase arrest upon FBXO42 knockout due to prolonged spindle assembly checkpoint activation. Altogether, this work identifies therapeutic avenues that may be exploited not only in GBM but also in other cancer types.

3.2 BACKGROUND

Glioblastoma (GBM) is the most aggressive and common form of brain cancer[3,4]. There are currently no highly effective therapies against GBM, and this cancer is one of the few tumor types with both a poor outcome and minimal improvement in survival in the past decades[183]. With standard of care treatments, including surgery, radiation, and chemotherapy, ~90% of adult patients die within 2 years of diagnosis[5,6]. Thus, improved treatments are desperately needed.

GBM stem-like cells (GSCs) have been isolated from patient tumors and retain the developmental potential and specific genetic alterations found in the original tumor[12,17,67,68]. When cultured under serum-free monolayer conditions, GSCs can retain tumor-initiating potential and tumor-specific genetic and epigenetic signatures over extended outgrowth periods[12,17]. In addition, they have been shown to recreate tumor cellular hierarchies when implanted into the cortex of immune-compromised mice[12,17]. While they cannot recreate the *entire* complexity of a GBM tumor, *in vitro*-cultured GSCs represent extremely useful models for studying subsets of GSCs with specific genetic and molecular profiles.

Even though large-scale genomic studies have given us extensive knowledge of the tumor suppressors and oncogenes that are commonly altered in cancer, directly targeting such genomic alterations therapeutically has often met with limited success. For instance, frequent genetic alterations in GBMs stimulate common signal transduction pathways involving RTK/Ras and PI3K/AKT signaling[184,185]. Although neither Ras nor AKT activation alone is sufficient to drive GBM, their combined activation is sufficient to induce GBM formation in mice[186]. However, directly targeting these pathways in clinical trials has led to limited patient responses[187,188], calling into question their use as "actionable" therapeutic targets for GBM. An alternative approach is to identify novel therapeutic targets in an unbiased and high-throughput manner by using functional genomic screens. We have performed multiple such screens using both RNAi and CRISPR-Cas9 platforms. This has allowed us to successfully identify genes selectively required for the self-renewal and expansion of GSCs, but not fetal-derived neural stem cells (NSCs)[73,96,97,109,110], which have similar expression profiles and developmental potential but are not transformed[12,108]. Our recent whole-genome CRISPR-Cas9 screens led to the identification of the Wee1-like kinase PKMYT1/Myt1 as a vulnerability associated with the oncogenic activity of RTK/Ras and PI3K/AKT pathways. This demonstrates the utility of our functional genomic screening approach[73].

Besides general oncogenic pathway activations, specific oncogenic activities such as loss-of-function and gain-of-function mutations, gene expression changes, copy number changes, etc. can also lead to vulnerabilities unique to subsets of tumor cells, which we may be able to exploit in order to treat a defined patient population[91]. Here, we wanted to further characterize candidate GSC-specific hits from our CRISPR-Cas9 whole genome screens in order to: 1) identify GBM gene dependency groups, or multiple genetic vulnerabilities that can be expected to be shared in

the same GBM tumor; and 2) identify context-specific vulnerabilities, or genes that are required in a certain cellular context (e.g. specific oncogenic activities listed above). We therefore performed comprehensive pooled retests of all putative screen hits that scored preferentially in GSC isolates over NSCs isolates in our whole-genome CRISPR-Cas9 screens. We then incorporated published functional genomic screening data (DepMap database) into our data in order to facilitate the desired analyses.

By creating co-dependency networks, we identified GSC-specific gene vulnerability groups related to mitochondrial protein processing and turnover and membrane trafficking as well as metabolic enzymes and regulators. Furthermore, we identified predicted context-specific vulnerabilities and further investigated two of these – the dsRNA-editing enzyme ADAR, which we demonstrate is required in the context of an interferon-stimulated gene expression signature, and the adapter protein EFR3A, which we determine is required in cells that have low expression of its paralog EFR3B.

Additionally, we explored the intriguing GSC-specific hit *FBXO42*, which encodes a poorly-studied F-box protein that scored particularly differentially in one of the three GSC isolates screened compared to NSCs. We found that FBXO42 is essential in a subset of GSCs and cell lines of various cancer types, but nonessential in NSCs, indicating this target's potential for a large therapeutic window and broad applicability. A previous study suggested that p53 is a substrate of FBXO42 and that stabilization of p53 upon FBXO42 loss causes reduced viability in cancer cells[129]. We demonstrated that the ubiquitin ligase role of FBXO42 is indeed responsible for the viability phenotype, but very likely not through p53. We identified a possible interactor of FBXO42, the phosphoprotein CCDC6, and we showed that these two proteins likely work together rather than redundantly to promote viability in FBXO42 loss-sensitive cells. Lastly, we

demonstrated that the observed viability loss upon FBXO42 knockout results from an extended metaphase arrest due to prolonged spindle assembly checkpoint activation. Our work illustrates the utility of using CRISPR-Cas9 functional genomic screens to identify novel therapeutic opportunities in patient tumor samples.

3.3 RESULTS

3.3.1 *Successive rounds of CRISPR-Cas9 functional genomic screening to identify strongest in vitro growth-promoting genes in GBM*

We previously performed genome-scale pooled outgrowth screens in human GSCs and NSCs using both RNAi and CRISPR-Cas9 platforms in order to identify genes required for the proliferation and survival of GBM cells[73,96,97,109,110]. Partial retests of these screens led to the identification of genetic vulnerabilities associated with the oncogenic activity of RTK/Ras and PI3K/Akt pathways or MYC. In order to further characterize candidate GSC-specific hits from our CRISPR-Cas9 screens, here we performed comprehensive retests of all putative strong, medium, and weak screen hits that scored preferentially in any of three GSC isolates (mesenchymal GSC-0131, proneural GSC-0827, and mesenchymal GSC-1502) over two NSCs isolates. Our goal was to facilitate identification of GBM gene dependency groups and context-specific vulnerabilities. We therefore took our whole-genome CRISPR-Cas9 screen results and selected all genes that fit the relaxed criteria of false discovery rate (FDR) <0.05 and $\log_2$ fold-change (logFC) <-1 for at least one sgRNA in any GSC line but in neither of the NSC lines (Figure 3.1B). We then took this list of 907 genes and mined the CRISPR-Cas9 GeCKO v2 library[189] for improved sgRNA sequences, which we then used to create a new library that we termed the comprehensive retest library (Figure 3.1B). We also included sgRNAs targeting essential and growth-limiting control genes (for a total library size of 1,079 genes), as well as 110 non-targeting control sgRNAs.

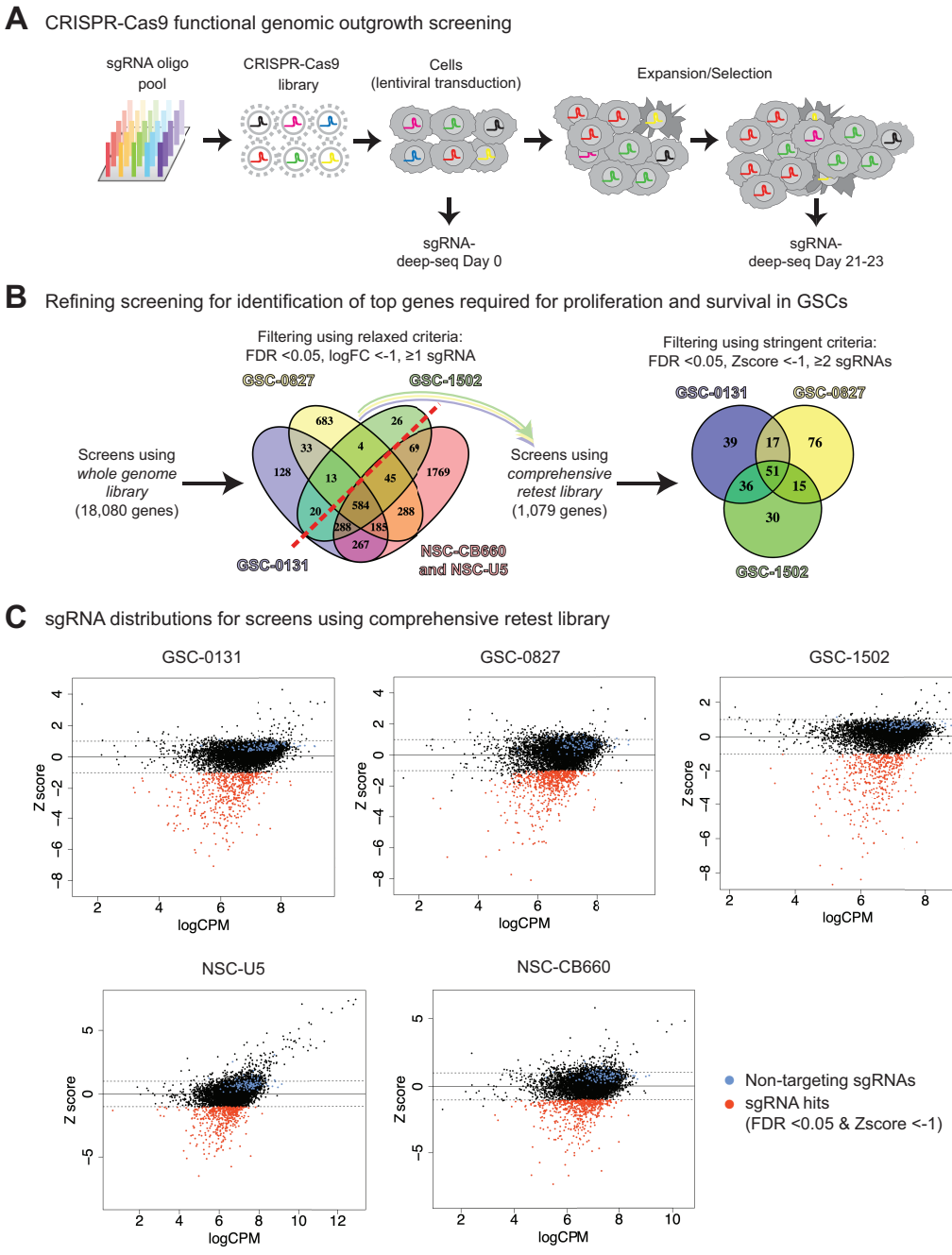


Figure 3.1. Successive rounds of CRISPR-Cas9 functional genomic screening to identify strongest *in vitro* growth-promoting genes in GBM. **(A)** Overview of CRISPR-Cas9 *in vitro* functional genomic outgrowth screening procedure. **(B)** Strategy for initial whole genome screens followed by refining retest screening using a comprehensive retest library targeting genes based on the whole genome screen results. **(C)** Distribution of logCPM (counts per million) versus Zscores for all sgRNAs in the comprehensive retest library, shown for the 3 GSC and 2 NSC lines that were screened. Each dot represents an sgRNA. Non-targeting sgRNAs are marked

in blue, while sgRNAs meeting the scoring criteria for top growth-promoting genes (FDR <0.05 & Zscore <-1) are marked in red.

We performed outgrowth screens using our comprehensive retest library in the same GSCs and NSCs in which we had performed the whole genome screens, in a manner consistent with our previous outgrowth screen procedure (Figure 3.1A). Briefly, after cloning the library from an oligo pool, we generated lentivirus and transduced cells at a low MOI to ensure a high probability that each cell would only receive one sgRNA integration. Two days after transduction, we treated the cells with puromycin for 3 days in order to select for successfully transduced cells. We then harvested the initial timepoint before culturing the cells for ~21 days to allow for competition between the integrated sgRNAs. After collecting the final timepoint, we used deep sequencing to measure relative abundance of the integrated sgRNAs before and after outgrowth. Using this system, an sgRNA that is overrepresented after outgrowth (positive logFC) indicates a potentially growth-limiting gene, while an sgRNA that is underrepresented after outgrowth (negative logFC) indicates a potentially growth-promoting gene that may be required for proliferation, survival, or growth.

We analyzed the results of our comprehensive retest library screens by normalizing the logFC data via conversion to Zscores and by using more stringent gene selection criteria. In order to identify growth-promoting genes, we filtered for sgRNAs with FDR <0.05 and Zscore <-1, and we termed "gene hits" any gene which had at least 2 sgRNAs scoring in this manner. This produced a list of 264 top potentially growth-promoting gene hits in GSCs, including 51 genes that were shared between all 3 GSC isolates (Figure 3.1B). We observed that the overall distribution of all sgRNAs in the library was very similar for each of the lines screened, and in each case the majority of non-targeting control sgRNAs clustered together and fell slightly above zero, but below the positive growth-limiting scoring cutoff of Zscore >1 (Figure 3.1C). (The observation that the non-

targeting controls appeared slightly positive is likely due to the fact that the library was designed to contain many sgRNAs that would score negatively, and since all sgRNA effects are relative, this would cause a concomitant apparently slightly positive effect for the non-targeting sgRNAs.)

In order to identify sgRNAs that scored more negatively in GSCs compared to NSCs, we created differential Zscores for each GSC vs. each NSC. SgRNAs that have negative GSC-NSC differential Zscores would correspond to genes that are preferentially required in GSCs compared to NSCs. Figure 3.2 shows heatmaps of GSC – NSC differential Zscores for all sgRNAs that scored in all 3 GSC lines, corresponding to 51 genes (left two heatmaps), as well as for all non-targeting sgRNAs (right two heatmaps). It is clear that a subset of these top GSC-shared hits display a strong and consistent increased dependency in GSCs compared to each NSC line (e.g. *CAB39*, *GMPPB*, *KIAA1432/RIC1*, etc.). Importantly, the spread of differential Zscores for the non-targeting controls was much lower than for gene hits and the values were much closer to zero, indicating that our screen results are not somehow biased to be overall more negative in GSCs (right two heatmaps).

We wanted to further characterize screen hits in order to: 1) identify GBM gene dependency groups, or multiple genetic vulnerabilities that can be expected to be shared in the same GBM tumor; and 2) identify context-specific vulnerabilities, or genes that are required in a certain cellular context, such as a gene expression profile, inactivation/deficiency of a gene, gain-of-function mutation, copy number alteration, etc. We therefore took advantage of the Broad Institute's DepMap database, which represents recent efforts to comprehensively identify vulnerabilities in hundreds of serum-cultured human cancer cell lines from various tumor types. While such cell lines are not ideal for modeling cancer biology[190,191], they have led to numerous

key discoveries[102-105]. By incorporating both functional screening and descriptive genomic profiling data from DepMap into our results, we were able to address both of our aims.

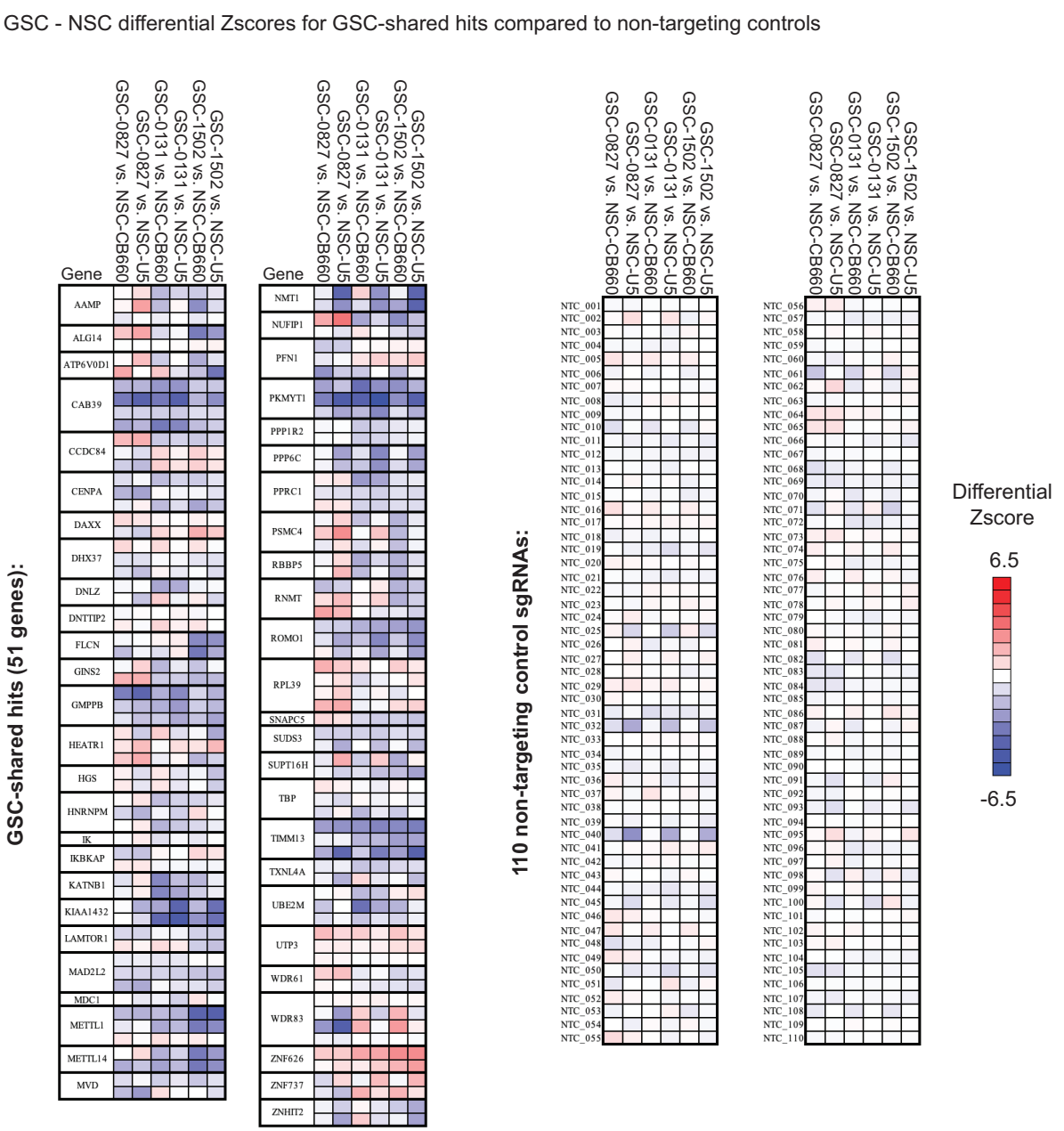


Figure 3.2. A subset of top GSC hits display increased dependency in GSCs compared to NSCs. **(A)** Heatmap of GSC – NSC differential Zscores for all sgRNAs that scored (FDR <0.05 & Zscore <-1) in all 3 GSC lines in the comprehensive retest screens (corresponding to 51 genes, displayed in two sets). The columns in each set correspond to all six combinations of Zscore(GSC) – Zscore(NSC). Each row corresponds to an sgRNA, and multiple sgRNAs targeting the same gene are grouped together. A heatmap of the differential Zscores for the 110 non-targeting sgRNAs in the library is shown for comparison.

3.3.2 *Mitochondrial-related and metabolism-related genetic vulnerability groups in GBM identified through co-dependency network analysis*

First, we wanted to take advantage of our results to identify groups of genes that are preferentially required in GBM compared to normal NSCs, particularly related genetic vulnerabilities that can be expected to be shared in the same GBM tumor. However, the 3 GSC lines that we used for screening do not provide a large enough sample size to make predictions about co-dependencies, or relationships where dependency (i.e. degree of requirement for proliferation, survival, or growth) on one gene correlates with dependency on another gene. Therefore, we mined DepMap[102-105] for CRISPR gene dependency correlations to determine which of our hits were connected by co-dependency relationships. DepMap CRISPR screening data is presented as one "CERES" gene dependency score per gene per cell line, normalized for copy-number specific effects and for the median of common essential genes[103]. The interpretation of this score is similar to a Zscore from our screens, where a positive value corresponds to a gene whose loss caused an increase in cell fitness (i.e. growth-limiting gene), while a negative value corresponds to a gene whose loss caused a decrease in cell fitness (i.e. growth-promoting gene).

We used DepMap correlations for our top screen hits that scored in 1, 2, or all 3 GSC lines and filtered for a correlation coefficient >0.2. We then used this data to create a large co-dependency network, where each node represents a gene, and each edge represents a correlation between dependency on one gene and dependency on the connected gene. Genes scoring in 1, 2, or all 3 GSCs in our screens are marked in yellow, blue, or green, respectively (Figure 3.3A). Interestingly, we noticed that the vast majority of our hits were connected indirectly by co-dependency relationships, with very few genes falling outside the large network. We also observed that this large network contained two especially concentrated clusters (red boxes in Figure 3.3A and Figure 3.4A). We pulled out the gene hits from these clusters and created subnetworks (Figure

3.3B and Figure 3.4B) by incorporating DepMap co-dependency correlations with data regarding physical interactions, predicted interactions, and shared pathways (gathered from GeneMania[192]). Edge widths were created to correspond to the value of the correlation coefficient for co-dependency correlations and to the weight of the association for GeneMania-derived data. Since we wanted to determine if the clusters we had identified were in fact preferentially required in GSCs compared to NSCs, we then incorporated our differential Zscore data into these subnetworks to identity gene hits with a GSC-NSC Zscore differential <-0.5 for all GSCs in which a particular gene scored, versus each NSC (see node borders in Figure 3.3B and Figure 3.4B).

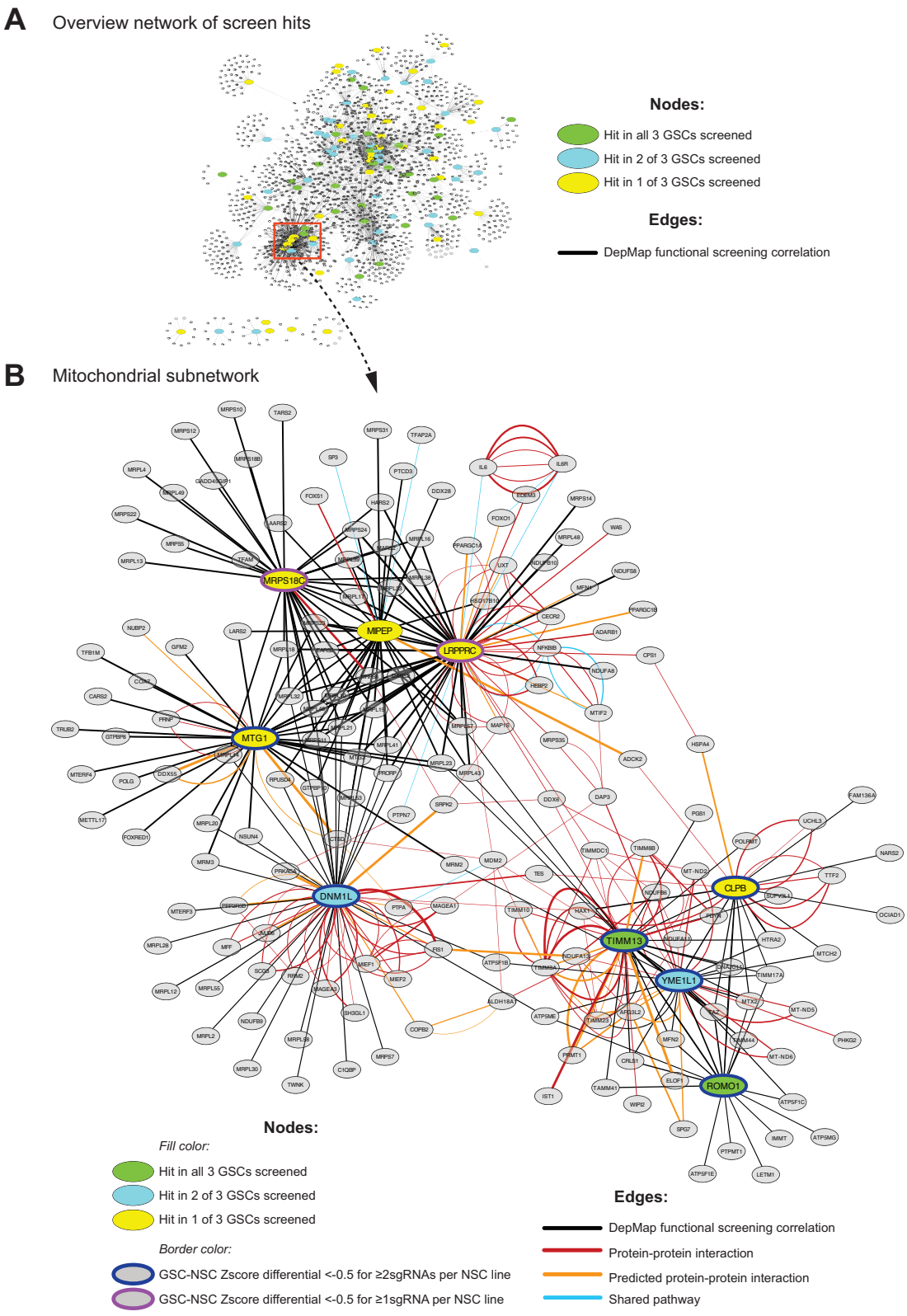


Figure 3.3. Identification of a group of mitochondrial-related genetic vulnerabilities through co-dependency network analysis. **(A)** Larger overview network of functional screening correlations for our comprehensive retest

screen hits. Each node represents a gene, and genes scoring in 1, 2, or all 3 GSCs in our screens are shown larger and marked in yellow, blue, or green, respectively. Each black edge is based on DepMap's publicly-available functional genomic CRISPR-Cas9 screening data and represents a correlation between dependency on one gene and dependency on the connected gene among hundreds of cancer cell lines. Red box denotes the location of a mitochondrial subnetwork within this larger network. **(B)** The mitochondrial subnetwork that can be identified in the larger network shown in (A). Gene hit node fills are colored as in (A), and gene hits with a GSC-NSC Zscore differential <-0.5 (for any scoring GSC) are also marked by border colors. Edges display DepMap correlations as well as protein-protein interactions and pathway relationships (GeneMania). Edge widths correspond to the value of the correlation coefficient for DepMap-derived correlations and to the weight of the association for GeneMania-derived data.

The first gene cluster we had identified in this manner and for which we thus created a subnetwork contained genes involved in mitochondrial protein processing and turnover and membrane trafficking (Figure 3.3B). Specifically, there was a close association between the mitochondrial ribosomal protein *MRPS18C*[193]; the GTPase *MTG1*, which plays a role in the regulation of mitochondrial ribosome assembly and translational activity[194,195]; *LRPPRC*, which binds to poly(A) mRNA in mitochondria and works to regulate certain mitochondrially encoded genes and gluconeogenic genes[196]; and *MIPEP*, which cleaves a specific class of nuclear-encoded proteins targeted to the mitochondrial matrix or inner membrane to their mature size, particularly oxidative phosphorylation-related proteins[197,198]. This group of genes was connected via *DNM1L*, which plays an important role in mitochondrial division by mediating membrane fission[199-201], to another group of genes involved in membrane processes. Specifically, this group consisted of the mitochondrial membrane protein *ROMO1,* which is responsible for increasing the level of reactive oxygen species[202-204]; *YME1L1*, an ATP-dependent metalloprotease that catalyzes the degradation of folded and unfolded proteins in the mitochondrial intermembrane region[205,206]; *TIMM13,* a mitochondrial intermembrane chaperone that participates in the import and insertion of multi-pass transmembrane proteins into the mitochondrial inner membrane[207,208]; and *CLPB*, a regulatory ATPase that contributes to protein trafficking and disaggregation of protein aggregates[209,210].

Remarkably, 8 of the 9 total gene hits in our mitochondrial subnetwork displayed an increased requirement in GSCs compared to NSCs via differential Zscores. Altogether, this subnetwork demonstrates that GBM cells are more likely to be sensitive to inhibition of genes involved in mitochondrial protein processing and turnover and membrane trafficking. Interestingly, while some screen hits that displayed co-dependency were also connected by known or predicted protein-protein interactions or pathway annotations, most of them were not, indicating that such data is not sufficient for the identification of gene vulnerability groups, and highlighting the importance of incorporating functional genomic screening data.

The second gene cluster we had identified in our larger co-dependency network corresponded to genes related to metabolism, especially metabolic enzymes, so we created a separate subnetwork for these genes (Figure 3.4B). The most dense group of genes within this cluster included *GMPPB*, a GDP-mannose pyrophosphorylase[211]; *PGP*, a glycerol-3-phosphate phosphatase that regulates the cellular levels of glycerol-3-phosphate[212], a metabolic intermediate of glucose, lipid, and energy metabolism; *MAT2A*, which catalyzes the formation of S-adenosylmethionine from methionine and ATP[213]; *MVD*, which catalyzes the conversion of mevalonate pyrophosphate into isopentenyl pyrophosphate in one of the early steps in cholesterol biosynthesis[214]; *PGD*, which is the second dehydrogenase in the pentose phosphate shunt[215]; *UMPS*, a uridine 5'-monophosphate synthase that catalyzes the final two steps of the *de novo* pyrimidine biosynthetic pathway[216]; as well as two enzymes involved in the *de novo* synthesis of guanine nucleotides, *GMPS*[217] and *IMPDH2*[218]. Furthermore, these gene hits displayed co-dependency with the catalytic subunits of ATPases *ATP1A1*, which is responsible for establishing and maintaining the electrochemical gradients of sodium and potassium ions across the plasma

membrane[219], and *ATP6V1G1*, part of the peripheral V1 complex of vacuolar ATPase that is responsible for acidifying a variety of intracellular compartments[220].

Interestingly, the above-described group of enzymes was connected to genes involved in regulating AMPK signaling: the serine/threonine-protein kinase *STK11*, which acts as a key upstream regulator of AMPK by mediating phosphorylation and activation of AMPK catalytic subunits[221], as well as two members of a complex that binds and activates *STK11 – CAB39* and *STRADA*[222]. This connection is supported by a recent publication showing that *PGD* contributes to AMPK activation through inhibition of PP2A[223]. Traditionally, AMPK has been regarded as a tumor suppressor since its activation promotes ATP conservation by inhibiting proliferation and growth-promoting pathways including synthesis of proteins, sterols, and fatty acids[224]. However, in recent years this generalization has been challenged since oncogenic signaling causes rapid cellular proliferation, which comes with high metabolic demand that creates a metabolically stressed state. AMPK could serve to alleviate this stress by reducing energy demand through inhibition of mTORC1-regulated anabolic growth and by promoting energy generation through increased mitochondrial complex I and catabolic metabolism[225]. Our results are consistent with this notion that AMPK signaling is actually required for some tumor cells, and point toward this signaling pathway as being important for survival of GBM cells.

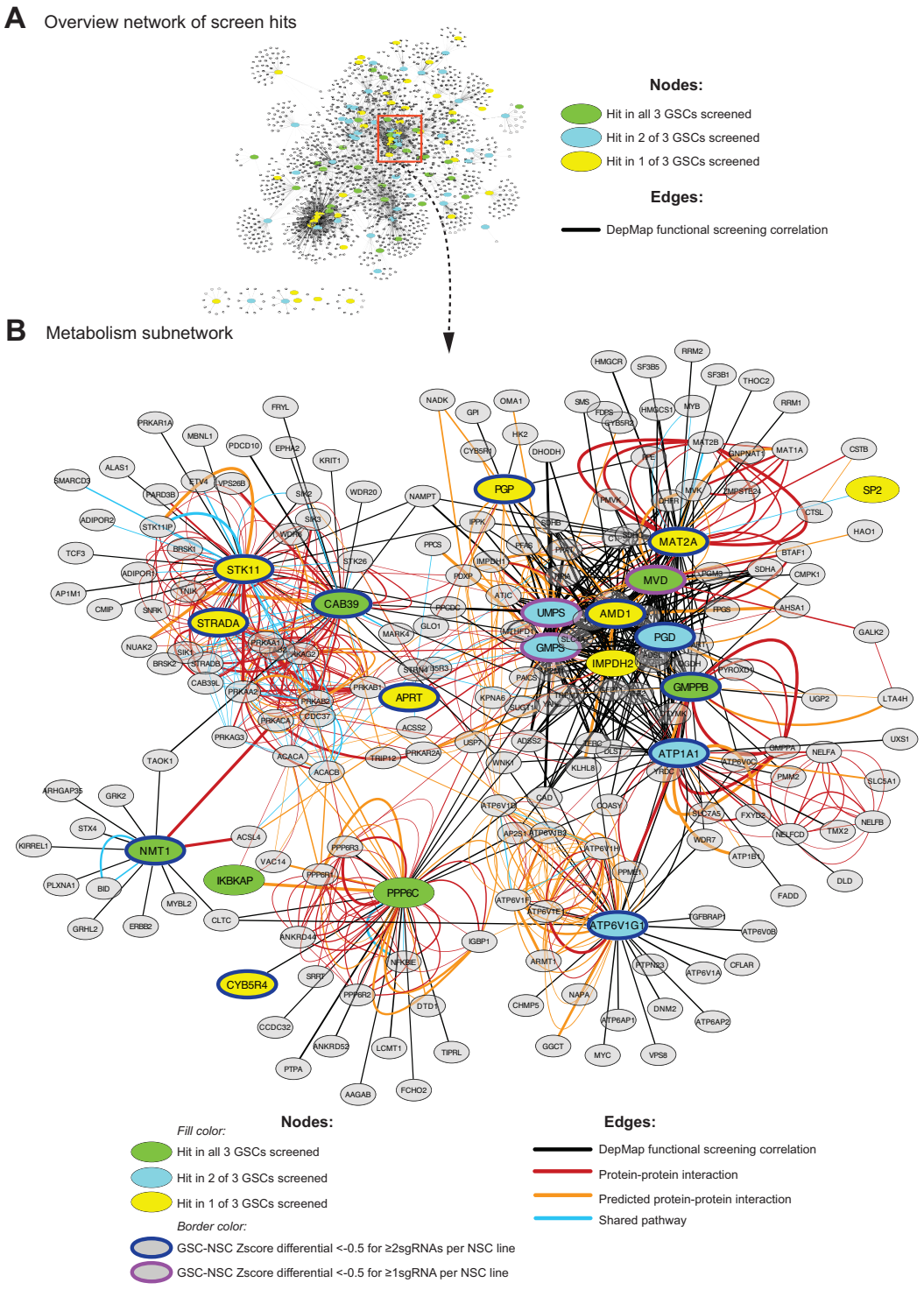


Figure 3.4. Identification of a group of metabolism-related genetic vulnerabilities through co-dependency network analysis. **(A)** Larger overview network of functional screening correlations for our comprehensive retest screen hits. Each node represents a gene, and genes scoring in 1, 2, or all 3 GSCs in our screens are shown larger and marked in yellow, blue, or green, respectively. Each black edge is based on DepMap's publicly-available

functional genomic CRISPR-Cas9 screening data and represents a correlation between dependency on one gene and dependency on the connected gene among hundreds of cancer cell lines. Red box denotes the location of a metabolism subnetwork within this larger network. **(B)** The metabolism subnetwork that can be identified in the larger network shown in (A). Gene hit node fills are colored as in (A), and gene hits with a GSC-NSC Zscore differential <-0.5 (for any scoring GSC) are also marked by border colors. Edges display DepMap correlations as well as protein-protein interactions and pathway relationships (GeneMania). Edge widths correspond to the value of the correlation coefficient for DepMap-derived correlations and to the weight of the association for GeneMania-derived data.

Overall, 16 of 20 total gene hits in our metabolism-related subnetwork displayed an increased requirement in GSCs compared to NSCs via differential Zscores, highlighting the increased dependency of GBM cells on these specific metabolic enzymes and regulators. As one might expect, our mitochondrial-related and metabolism-related subnetworks were connected within the initial larger network we had created, although they still represent distinct gene vulnerability groups, with one emphasizing mitochondrial protein processing and turnover and membrane trafficking and the other highlighting metabolic enzymes and regulators.

3.3.3 *Integration of functional and descriptive genomics for the identification of context-specific cancer vulnerabilities*

In addition to using co-dependency network analysis to identify gene vulnerability groups that may be applied to GBM in general, we also wanted to use our screen results to identify context-specific vulnerabilities, or genes that are required in subsets of tumor cells with a particular cellular context (e.g. increased/decreased expression of a particular gene or group of genes). To this end, we integrated our functional screening data with descriptive genomic data for our GSC lines, namely RNA-seq, exome-seq, and CNV data, in order to generate hypotheses regarding context-specific vulnerabilities. We then looked for additional evidence for our hypotheses, once again turning to DepMap data, and created a list of top predictions (Figure 3.5A). The contexts for these predictions could be broken up into 3 categories: tumor subtype / expression profile, where certain genes are

required in subsets of tumors with a particular gene expression profile; specific oncogenic activities, e.g. where vulnerabilities might occur in the context of RTK and PI3K pathway activation[73] or in the context of high expression of an oncogene itself (which represents the concept of oncogene addiction); or deficiency in a gene's paralog or family member, which represents the classic concept of synthetic lethality (Figure 3.5A). We wanted to explore two of these context-specific vulnerability predictions further – one for which evidence already exists (*ADAR*) and one for which it does not (*EFR3A*).

ADAR (or more specifically the isoform *ADAR1*) is the primary enzyme responsible for editing dsRNA by catalyzing the hydrolytic deamination of adenosine to inosine[126]. We became interested in this gene because *ADAR* was a patient-specific hit in GSC-0131 cells in our screens. *ADAR* serves to neutralize immunogenic dsRNAs, and it is known that loss of *ADAR* leads to cell lethality under interferon (IFN) treatment[226]. It was recently shown that tumor cells themselves can produce IFN, creating an IFN-stimulated gene expression (ISG) signature, which makes them more vulnerable to *ADAR* loss compared to normal cells due to accumulation of dsRNAs[227,228]. Indeed, when we looked at top DepMap correlations with *ADAR* dependency, we saw that they included mainly negative correlations for expression of ISG genes, which suggests that there is an increased requirement for *ADAR* when expression of these genes is higher (Figure 3.5B). Comparing RNAseq data for our various GSC lines, we observed that GSC-0131 cells do display some upregulation of ISG genes (Figure 3.5C, top panel; GSC-0131 highlighted by red box). We also analyzed a panel of pediatric brain tumor isolates and saw that a subset of these had an even higher upregulation of ISG genes (Figure 3.5C, bottom panel; left side of heatmap).

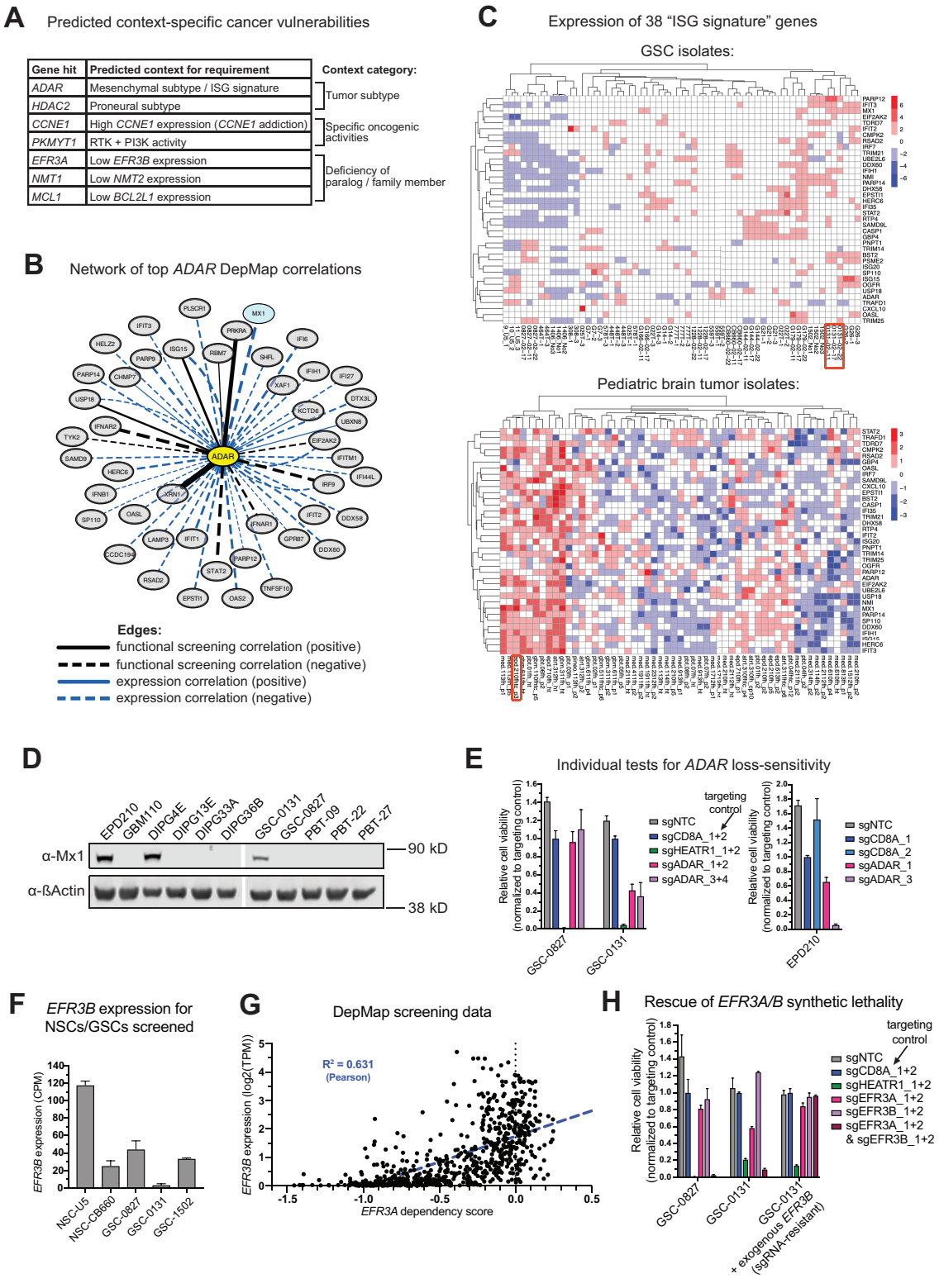


Figure 3.5. Integration of functional and descriptive genomics allows for the identification of context-specific cancer vulnerabilities. **(A)** Table of predicted context-specific cancer vulnerabilities, identified by incorporating DepMap correlation data with patient-specific gene hit data from our CRISPR screens. **(B)** Mini-network of top

DepMap gene co-dependency and gene expression correlations for the GSC-0131-specific screen hit *ADAR*. Blue expression edges indicate that expression of the gene to which *ADAR* is connected correlates with *ADAR* dependency score. **(C)** Clustered heatmaps showing relative expression of genes in a 38-gene ISG signature for a set of GSCs/NSCs and a set of various pediatric brain tumor isolates. **(D)** Western blot showing MX1 protein expression in a subset of lines shown in (C). **(E)** Relative cell viability (normalized to targeting control sgCD8A) for lines nucleofected with CRISPR RNPs targeting *ADAR*. *HEATR1* is an essential control gene. sgNTC = non-targeting control sgRNA. Measured at 10 days post nucleofection. **(F)** Relative *EFR3B* expression, based on RNAseq data, for GSCs and NSCs that were used for CRISPR screens with comprehensive retest library. **(G)** Cell line functional genomic screening data from DepMap showing *EFR3A* dependency score (CERES score from CRISPR Avana 19Q4 dataset) vs. *EFR3B* expression. Each dot represents a cell line. Dotted blue line shows linear regression fit. **(H)** Relative cell viability (normalized to targeting control sgCD8A) for cells nucleofected with CRISPR RNPs targeting *EFR3A*, *EFR3B*, or both. GSC-0131 are compared to GSC-0131 that were transduced with a lentiviral construct expressing *EFR3B*. *HEATR1* is an essential control gene. sgNTC = non-targeting control sgRNA. Measured at 9 days post nucleofection.

Since *MX1*, an IFN-induced dynamin-like GTPase with antiviral activity against a wide range of RNA viruses[229], was the strongest expression correlation with *ADAR* dependency in DepMap (Figure 3.5B), we assessed MX1 protein levels in some of the isolates for which we had RNA expression data (Figure 3.5D). We observed that GSC-0131 as well as the pediatric ependymoma line EPD210, which was among the ISG-positive pediatric lines (shown in red boxes in Figure 3.5C), did in fact have much higher MX1 protein expression than other lines, for which MX1 protein was actually undetectable (Figure 3.5D). (MX1 was also higher in a diffuse intrinsic pontine glioma line, DIPG4E, for which we did not have RNA expression data.) We therefore decided to individually test *ADAR* loss in GSC-0131 and EPD210 using CRISPR Cas9:sgRNA ribonucleoprotein (RNP) nucleofection and subsequent measurement of relative cell viability. We compared the effect on viability in these two lines to the ISG gene expression signature-negative line GSC-0827 (Figure 3.5C), which had no detectable expression of MX1 protein (Figure 3.5D) and did not score as *ADAR* loss-sensitive in our screens. Since it has been suggested that the DNA damage associated with CRISPR cutting can lead to negative viability effects independently of

whether a gene is truly required or not[103,230-232], we chose to use sgRNAs targeting the unrelated gene *CD8A*, which is not expressed in GSCs and NSCs, as a targeting control for normalization in all of our viability experiments, rather than using only a non-targeting control. When we nucleofected GSC-0131 and GSC-0827 with CRISPR RNPs targeting the essential control gene *HEATR1*, we observed a similar reduction in viability between these lines (Figure 3.5E). However, targeting of *ADAR* using two different combinations of dual sgRNAs significantly reduced viability in GSC-0131 but not GSC-0827. In addition, we observed that like GSC-0131, EPD210 cells were also sensitive to *ADAR* loss (Figure 3.5E). These results are in line with the recent publications suggesting that *ADAR* is specifically required in the context of increased expression of ISG genes. Furthermore, our results suggest that *MX1* expression (particularly MX1 protein expression) may be the most useful individual predictive biomarker for *ADAR* loss-sensitivity.

The second context-specific vulnerability hypothesized from our screens that we wanted to explore further was *EFR3A*. EFR3A and EFR3B are recently-characterized adapter proteins that are required to localize the enzyme PI4KIIIα to the plasma membrane[127,128], where it generates PI4P, which is important for plasma membrane identity[233]. *EFR3A* scored as a gene hit specific to GSC-0131 in our screens, which we noticed had very low expression of the paralog *EFR3B* compared to the other GSCs and NSCs that we had screened (Figure 3.5F), suggesting a potential synthetic lethal relationship between these two paralogs. When we examined DepMap data, we also observed a trend toward increased requirement for *EFR3A* in tumor cells with low *EFR3B* expression (Figure 3.5G). We therefore decided to use CRISPR RNPs to individually knock out *EFR3A* and *EFR3B* in the *EFR3B*-higher line GSC-0827, the *EFR3B*-lower line GSC-0131, and GSC-0131 which we had transduced with lentivirus overexpressing an sgEFR3B-resistant version of exogenous *EFR3B* (Figure 3.5H). We observed that indeed, the *EFR3B*-lower line GSC-0131

was more sensitive to *EFR3A* loss than the *EFR3B*-higher line GSC-0827, and this effect was fully rescued by exogenous expression of *EFR3B* in GSC-0131. When we knocked out *EFR3B* in GSC-0131 and GSC-0827, we did not observe a loss of viability in either line, which we expected since these lines still express high levels of *EFR3A*. However, we observed that simultaneous targeting of *EFR3A* and *EFR3B* was extremely lethal to both GSC lines (Figure 3.5H). Taken together, our results indicate that *EFR3A* and *EFR3B* are truly synthetic lethal, which creates a unique therapeutic opportunity for targeting *EFR3A* in tumor cells that endogenously express low levels of *EFR3B*.

3.3.4 *FBXO42 is a unique screen hit that is essential in a subset of GSCs in vitro and in vivo but is nonessential in NSCs in vitro*

Among the strongest GSC-specific hits in our screen retests was *FBXO42*. It scored in a manner where it appears essential in GSC-0827, while being seemingly completely dispensable in GSC-0131, GSC-1502, and NSCs. *FBXO42* encodes an F-box protein that serves as the substrate-recognition component of an SCF (SKP1-CUL1-F-box protein)-type E3 ubiquitin ligase complex, and it is thus far very poorly studied, with only a few publications exploring the roles of the protein[129,234-236]. One previous study described a role for FBXO42 in the destabilization of p53, similar to MDM2[129]. Because GSC-0827 cells are *TP53*-wt and GSC-0131 are *TP53*-mutant, we wondered if *TP53* status might explain FBXO42's GSC-0827-specific requirement. We therefore decided to explore this screen hit further.

First, we wanted to individually retest *FBXO42* loss to determine if the patient-specific effect in GSC-0827 compared to NSCs really was as stark as it appeared from our screening data. When we nucleofected GSC-0827 and NSC-U5 with CRISPR RNPs targeting *FBXO42* and subsequently measured relative cell viability, we found that loss of this gene did indeed have a

profound effect on GSC-0827, with almost no cells surviving, but had not even a partial negative viability effect on NSC-U5 (Figure 3.6A). Importantly, we did see a strong reduction in viability in NSC-U5 when targeting the essential gene *HEATR1* as a control, indicating that our differential results were not due to reduced nucleofection efficiency in NSC-U5s. We also tested additional GSC lines and found that similar to NSC-U5, GSC-0131 cells were unaffected by loss of *FBXO42*, while GSC-G166 cells did display a reduction in viability, though to a lesser extent than GSC-0827s (Figure 3.6A). This confirmed our initial hypothesis, that a subset of patient-derived GSC lines are *FBXO42* loss-sensitive (which we termed F42L-sensitive), while normal non-transformed NSCs are not sensitive (which we termed F42L-resistant).

We wanted to determine if the sensitivity to *FBXO42* loss holds true *in vivo*, and we wanted to model the therapeutic potential of targeting *FBXO42* in established tumors rather than simply testing engraftment of *FBXO42* knockout tumor cells. Therefore, we cloned an sgFBXO42_1-resistant version of the *FBXO42* ORF into the retroviral Tet-On vector pTURN-tight and transduced GSC-0827 cells with virus for this construct. Keeping the cells on doxycycline to maintain exogenous *FBXO42* expression, we then nucleofected them with CRISPR RNPs for sgFBXO42_1 in order to knock out the endogenous *FBXO42* alleles. In order to obtain cells with a uniform and well-controllable induction of exogenous expression, we then derived clones from this cell pool. We screened the clones for proper ability to turn off the construct by taking a subset of cells of each clone and testing for viability loss upon doxycycline removal. Based on these results, we took our top two clones (clones 6 and 15), PCR amplified the endogenous FBXO42 locus, and sequenced individual alleles using a TA cloning-like method (Figure 3.6B). We found that, as expected, both clones 6 and 15 had been successfully CRISPR-edited at this site. Clone 15 contained a net frameshift mutation in each allele, while clone 6 contained one frameshift allele

and one allele with a net 33 bp deletion, which would result in an in-frame protein that nonetheless contains an 11 aa deletion in the F-box domain. As we expected based on this result, Western blotting for FBXO42 protein revealed that after 4 days of doxycycline removal, clone 6 still displayed a faint band while clone 15 displayed no visible FBXO42 protein expression (Figure 3.6C). The fact that clone 6 responded to doxycycline removal with the expected dramatic loss in viability suggests that the remaining protein may not be entirely functional. However, we decided to move ahead with clone 15 for our *in vivo* experiment due to the observed total protein loss after doxycycline removal.

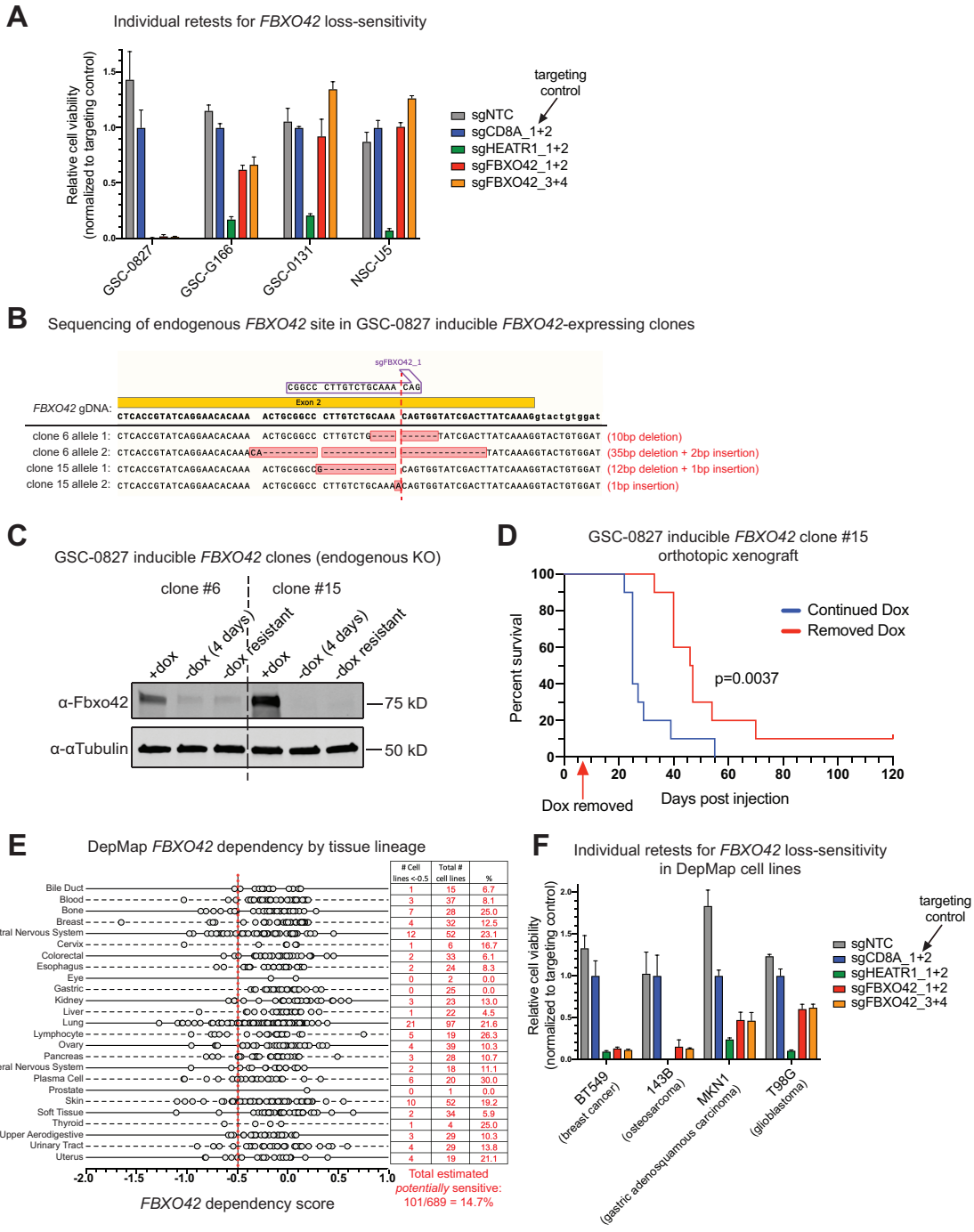


Figure 3.6. *FBXO42* is a unique screen hit that is essential in a subset of GSCs *in vitro* and *in vivo* but is nonessential in NSCs *in vitro*. *FBXO42* loss-sensitive tumors occur in various cancer types. **(A)** Relative cell viability (normalized to targeting control sgCD8A) for GSCs and NSCs nucleofected with CRISPR RNPs targeting *FBXO42*. *HEATR1* is an essential control gene. sgNTC = non-targeting control sgRNA. Measured at 9 days post nucleofection for GSCs and 11 days for NSCs (due to doubling time differences). **(B)** Genomic sequences of the *FBXO42* exon 2 region surrounding the sgFBXO42_1 target site, shown for individual alleles of 2 clones of GSC-0827 cells that had been transduced with a doxycycline-inducible *FBXO42* construct

(resistant to sgFBXO42_1) and then nucleofected with CRISPR RNPs for sgFBXO42_1 to knock out the endogenous *FBXO42*. PCR primers used were specific to the endogenous locus and could not amplify the exogenous construct. Red dotted line denotes sgRNA cut site. **(C)** Western blot for FBXO42 for clones described in (B) with either continuous doxycycline exposure, 4 days doxycycline removal, or 3-4 weeks of doxycycline removal (which allowed a very small percentage of doxycycline removal-resistant cells to expand). **(D)** Survival curve for immunocompromised mice which were orthotopically xenografted with GSC-0827 doxycycline-inducible *FBXO42* clone 15 described in (B) and (C). All mice were initially kept on doxycycline, and then doxycycline was removed for half the mice on Day 7 post injection. n=10 mice per group. **(E)** Breakdown by tissue lineage of *FBXO42* dependency in cell lines screened in DepMap (CERES score from CRISPR Avana 19Q4 dataset). Each circle corresponds to a cell line. Summary of proportion of potentially sensitive cell lines (score <-0.5) by tissue lineage is shown to the right. **(F)** Relative cell viability (normalized to targeting control sgCD8A) for 4 cell lines that were nucleofected with CRISPR RNPs targeting *FBXO42*. Lines had been predicted *FBXO42* loss-sensitive based on DepMap data (using CRISPR Avana dataset for all except MKN1, which were predicted based on Combined RNAi dataset.) *HEATR1* is an essential control gene. sgNTC = non-targeting control sgRNA. Measured at 8-12 days post nucleofection (depending on doubling time).

We implanted cells from GSC-0827 inducible pTT-FBXO42(sg#1-resistant) clone 15 orthotopically into the brains of immunocompromised mice. Before implantation and for the first 7 days after implantation, the mice were maintained on doxycycline water in order to prevent death of tumor cells before tumor initiation. On Day 7 after injection, we removed doxycycline from half the mice, normalizing the tumor burden in each cohort based on bioluminescent imaging (the parental GSC-0827 cells had been transduced with a luciferase construct). We monitored survival and found a significant increase in survival in the group which experienced doxycycline removal compared to the group that was kept on doxycycline, indicating that loss of *FBXO42* in established tumors *in vivo* resulted in improved outcome (Figure 3.6D). Nine of the 10 mice that experienced doxycycline removal did eventually succumb due to tumor burden, but the remaining mouse was actually still alive at the study endpoint of 120 days and had a small stable tumor that was not progressing. Interestingly, the observation that most of the tumors in mice not treated with doxycycline eventually continued proliferating is consistent with our *in vitro* results, where we observed that the extremely small percentage of cells that survived doxycycline removal (<3%) in

both clones 6 and 15 were eventually able to proliferate after ~3 weeks. When we assessed FBXO42 protein levels in these *in vitro* doxycycline independent cells, we found that their survival was *not* due to doxycycline-independent reactivation of the viral construct, as FBXO42 levels in both clones were identical for long-term doxycycline removal-resistant cells and cells which had been under doxycycline removal for 4 days and were thus just beginning to experience the effects of *FBXO42* loss (Figure 3.6C). This suggests that a very small subpopulation of cells had likely undergone some sort of genomic change that allowed them to become F42L-resistant. Nevertheless, the survival difference we observed is quite dramatic for glioblastoma, a tumor type with an extremely poor outcome and minimal improvement in survival in the past decades[183]. Altogether, these results illustrate that the requirement for FBXO42 holds true *in vivo* in F42L-sensitive GSCs and present FBXO42 as a viable potential therapeutic target in a subset of glioblastomas.

3.3.5 *FBXO42 loss-sensitive tumors occur in various cancer types*

We wondered if *FBXO42* loss-sensitivity is restricted to glioblastomas (or perhaps various types of brain tumors), or if other cancer types might also contain subsets of sensitive tumors. We thus turned to DepMap[102-105] to assess *FBXO42* dependency across cell lines derived from tumors of various tissue lineages (Figure 3.6E). Using the DepMap-suggested cutoff of a dependency score of -0.5, we quantified the percentage of cell lines within each tissue lineage that scored below this cutoff (Figure 3.6E, right side). We saw that most tissues appear to contain a subset of potentially F42L-sensitive cell lines, with a range of 0-30%, a mean of 13.2%, and a median of 11.1%. Overall across all tissues, almost 15% of DepMap cell lines were predicted to be potentially F42L-sensitive.

We wanted to confirm the predicted sensitivity in some of these cell lines to determine if DepMap screening results were indeed representative of sensitivity to *FBXO42* loss. Therefore, we took 3 cell lines that were predicted sensitive based on the CRISPR Avana dataset (Figure 3.6E) – the breast cancer ductal carcinoma line BT549, the osteosarcoma line 143B, and the glioblastoma line T98G – as well as one line that was predicted sensitive based on the Combined RNAi dataset (not shown) – the gastric adenosquamous carcinoma line MKN1 – and nucleofected them with CRISPR RNPs targeting *FBXO42* (Figure 3.6F). Compared to our targeting control sgCD8A, we observed that these 4 lines were indeed F42L-sensitive, with BT549 and 143B being exquisitely sensitive and MKN1 and T98G being moderately sensitive. These results indicate that tumors of completely different tissue origins contain subsets of F42L-sensitive cells, which suggests that *FBXO42* is involved in fundamental and universal biological processes and further highlights the need to study this gene as a potentially widely-applicable therapeutic target in multiple cancer types.

3.3.6 *The ubiquitin ligase role of FBXO42 is responsible for the observed phenotype through a novel, unknown substrate*

Since *FBXO42* is an Fbox protein that serves as the substrate-recognition component of an SCF (SKP1-CUL1-F-box protein)-type E3 ubiquitin ligase complex, we hypothesized that there are one or more substrates of *FBXO42* that must be degraded in sensitive cells, and that if these proteins are not degraded, their accumulation has a negative effect on proliferation and survival of the cells. We first wanted to determine if it was in fact its participation in an E3 ligase complex that causes the sensitivity we observe. F-box proteins interact with SKP1 through the F-box domain, and they interact with ubiquitination targets through other protein interaction domains. Besides an F-box domain, FBXO42 additionally contains solely a Kelch repeat domain[237], which serves as its

substrate-binding domain[129,234,236,237]. Therefore, we created lentiviral constructs containing 3XFLAG-tagged ΔF-box domain (deletion of aa 39-93) and ΔKelch domain (deletion of aa 107-354) mutant versions of *FBXO42*, along with a 3XFLAG-tagged full-length *FBXO42* control construct. We then transduced GSC-0827 cells with these constructs, which also conferred puromycin resistance, and selected them with puromycin. In order to test whether exogenous expression of the full-length, ΔF-box, and ΔKelch versions was able to rescue loss of endogenous *FBXO42*, we nucleofected the cells with CRISPR RNPs for two *FBXO42* sgRNAs to which the lentiviral vectors had been designed to be resistant (via synonymous mutation of multiple bases and use of an sgRNA spanning an intron-exon boundary), along with non-targeting and targeting control sgRNAs, and subsequently measured relative cell viability (Figure 3.7A). We observed that relative to the sgCD8A targeting controls, expression of full-length *FBXO42* was able to almost fully rescue the viability loss observed upon endogenous *FBXO42* knockout. However, with expression of the ΔF-box and ΔKelch versions, we observed a similar viability loss upon *FBXO42* knockout as we observed in untransduced control GSC-0827 cells. Importantly, the protein expression levels of the ΔF-box and ΔKelch versions were similar to expression of the full-length version (Figure 3.7B), indicating that the difference was not simply due to insufficient expression of the ΔF-box and ΔKelch proteins. These data suggest that both the F-box domain and the Kelch domain are required for survival of F42L-sensitive cells, which supports the idea that FBXO42's role in an E3 ubiquitin ligase complex is responsible for the observed phenotype.

Building upon this idea, we next wanted to know if a published target of FBXO42 is primarily responsible for the viability phenotype. If there is any one target that is primarily responsible, then it follows that accumulation of this target is what causes the negative effect in F42L-sensitive cells, since loss of FBXO42 would cause a decrease in degradation of the target.

We thus reasoned that in this scenario, knocking out the target would rescue the viability loss seen with FBXO42 knockout. To this end, we nucleofected GSC-0827 cells with CRISPR RNPs targeting the published targets/interactors *TP53*[129,234], *ING4*[236] (a chromatin reader), or *RBPJ*[238] (a transcriptional regulator important in Notch signaling) in combination with a non-targeting control sgRNA or sgFBXO42. Using microscopy to assess relative cell numbers, we did not observe any rescue for *FBXO42* loss with knockout of these 3 genes (Figure 3.7C). While this does not imply that none of these genes are targets of *FBXO42* per se, it does suggest that none of them are primarily responsible for the *FBXO42* sensitivity phenotype.

p53 has many important roles in the cell cycle and in cancer, and in particular, one group reported that FBXO42 directly interacts with p53 to cause its ubiquitination and degradation, arguing that loss of FBXO42 in cancer cells causes p53 stabilization, which leads to cell cycle arrest in the G1 phase and promotes apoptosis[129,234]. We thus wanted to further explore the possibility of this particular gene being a substrate of FBXO42 and causing the lethality phenotype we were observing. We therefore assessed p53 protein levels in F42L-non-sensitive GSCs and NSCs which we had nucleofected with CRISPR RNPs targeting *FBXO42*, as well as in our doxycycline-inducible clones 6 and 15 of F42L-sensitive GSC-0827 cells (Figure 3.7D). Across these non-sensitive and sensitive cell lines, we did not observe the stabilization of p53 protein that one would expect upon loss of FBXO42 if p53 were a vital substrate. It is also noteworthy that we did not observe a difference in p53 protein levels in the long-term doxycycline removal-resistant versions of GSC-0827 clones 6 and 15 (Figure 3.7D).

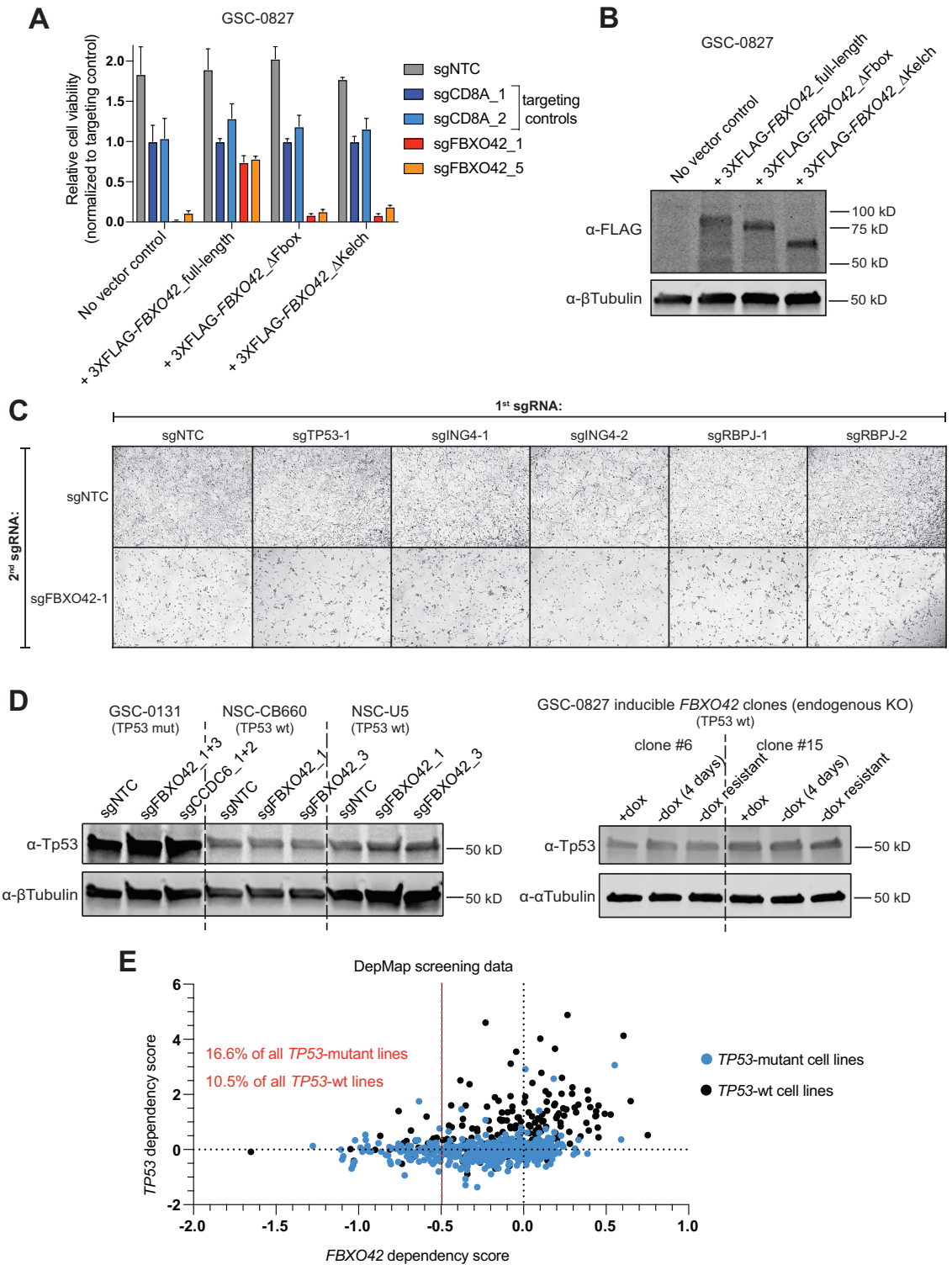


Figure 3.7. The ubiquitin ligase role of *FBXO42* is responsible for the observed phenotype through a novel, unknown substrate. **(A)** Relative cell viability (normalized to targeting control sgCD8A) for GSC-0827 transduced with full-length, F-box domain deletion mutant, or Kelch domain deletion mutant versions of 3XFLAG-*FBXO42* and then nucleofected with CRISPR RNPs targeting *FBXO42*, measured at 8 days post

nucleofection. The lentiviral expression constructs are resistant to the *FBXO42* sgRNAs used here. Untransduced GSC-0827 are shown for comparison. sgNTC = non-targeting control sgRNA. **(B)** Western blot for FLAG tag in transduced GSC-0827 used for viability assay in (A). **(C)** Representative images of GSC-0827 nucleofected with CRISPR RNPs targeting the published FBXO42 targets/interactors *TP53*, *ING4*, or *RBPJ* in combination with sgNTC or sgFBXO42, taken at 5 days post nucleofection. **(D)** Western blot for p53 levels in various GSCs and NSCs nucleofected with CRISPR RNPs targeting *FBXO42* or *CCDC6*. **(E)** Cell line functional genomic screening data from DepMap showing *FBXO42* dependency score vs. *TP53* dependency score (CERES scores from CRISPR Avana 19Q4 dataset). *TP53*-mutant cell lines are marked in blue.

Furthermore, we reasoned that if the negative effect on viability observed with *FBXO42* loss were indeed due to stabilization of wild-type p53 as reported by Sun et al.[129,234], then *TP53*-mutant cancer cell lines should be less sensitive to *FBXO42* loss than *TP53*-wt cell lines. In order to determine if this is true across a large set of cell lines, we once again turned to DepMap[102-105] data and assessed the relationship between *FBXO42* dependency, *TP53* dependency, and *TP53* mutation status (Figure 3.7E). As expected, the vast majority of *TP53*-mutant cell lines do not display significant positive *TP53* dependency scores (>0.5), indicating that loss of *TP53* in mutant lines does not result in an increase in proliferation (since mutant *TP53* is already mimicking the effect of loss of *TP53*). This serves as a validation of the data in making these kinds of observations. When in turn comparing *FBXO42* dependency between *TP53*-mutant and *TP53*-wt cell lines, it is clear that *TP53*-mutant lines do not have a lower proportion of F42L-sensitive lines (dependency score <-0.5) than do *TP53*-wt lines (79 of 477 = 16.6% for mutant lines and 22 of 210 = 10.5% for wt lines) (Figure 3.7E). Taken together, our data suggest that while FBXO42's role in an E3 ubiquitin ligase complex is responsible for the phenotype we observe in F42L-sensitive cells, the effect occurs primarily through a novel, unknown substrate. In particular, the evidence suggests that p53 is likely not a primary target responsible for the phenotype.

3.3.7 *FBXO42 and CCDC6 are both necessary to promote viability in FBXO42 loss-sensitive cells, likely working together rather than redundantly*

Since incorporating our observations with DepMap[102-105] had been an extremely useful resource for us thus far, we turned to DepMap to see if we could generate a hypothesis about potential interactors of FBXO42 that contribute to viability loss in F42L-sensitive cells. By far the top DepMap correlation with *FBXO42* dependency was dependency on the gene *CCDC6*, with a correlation coefficient of 0.746 (Figure 3.8A) (compared to the next-highest correlation value of 0.315). *CCDC6* is a coiled-coil domain-containing protein that is considered a tumor suppressor and was first identified due to its involvement in chromosomal rearrangements with the RET proto-oncogene in thyroid papillary carcinomas[239]. It is a pro-apoptotic protein substrate of ATM that has been shown to be involved in the DNA damage response[240-242]. Interestingly, the biophysical interactions of ORFeome-based complexes (BioPlex) network database, which is the result of mass spectrometry analysis done on cell lines that each express a tagged version of a protein from the ORFeome collection, actually suggests an interaction between FBXO42 and CCDC6[243,244]. The fact that there is a positive DepMap correlation between dependencies on these genes suggests that CCDC6 is likely not a ubiquitination target of FBXO42. If CCDC6 were a target, then loss of FBXO42 would result in increased CCDC6 levels, meaning that in F42L-sensitive cells, an increase rather than a decrease in CCDC6 should result in a negative viability effect. Nevertheless, we wanted to test whether CCDC6 protein levels change upon loss of FBXO42 or vice versa. In Western blots of two NSC lines and one GSC line that had been nucleofected with CRISPR RNPs targeting *FBXO42*, we did not observe a significant change in CCDC6 levels (Figure 3.8B). We also did not observe a change in FBXO42 levels in GSC-0131 cells in which we had knocked out *CCDC6* (Figure 3.8B). These results support the notion that CCDC6 is likely not an E3 ligase target of FBXO42 and also suggest that CCDC6 does not serve to affect FBXO42 levels.

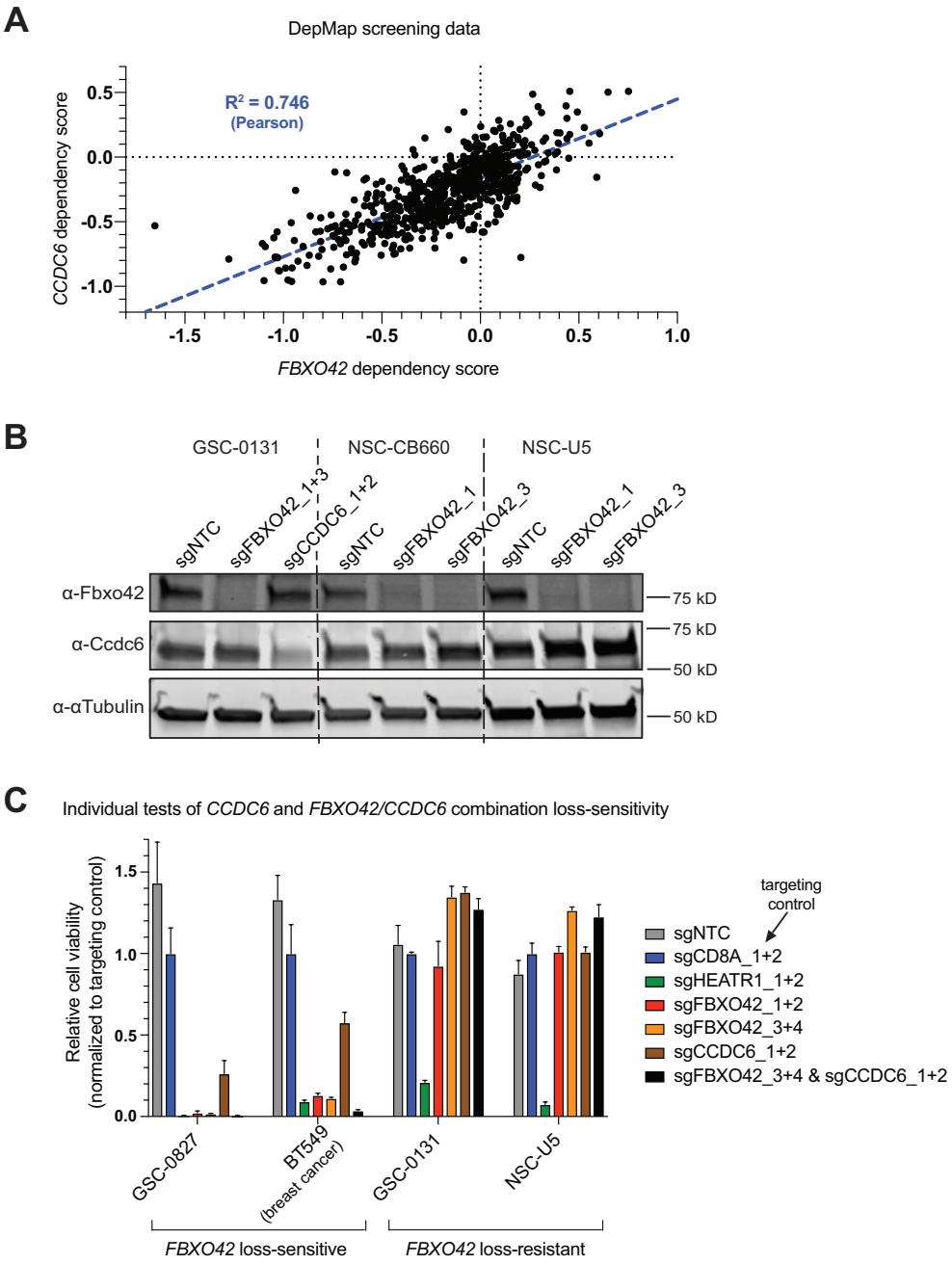


Figure 3.8. *FBXO42* and *CCDC6* are both necessary to promote viability in *FBXO42* loss-sensitive cells, likely working together rather than redundantly. **(A)** Cell line functional genomic screening data from DepMap showing *FBXO42* dependency score vs. *CCDC6* dependency score (CERES scores from CRISPR Avana 19Q4 dataset). Each dot represents a cell line. Dotted blue line shows linear regression fit. **(B)** Relative cell viability (normalized to targeting control sgCD8A) for cells nucleofected with CRISPR RNPs targeting *FBXO42*, *CCDC6*, or both. *HEATR1* is an essential control gene. sgNTC = non-targeting control sgRNA. Measured at 8-11 days post nucleofection (depending on doubling time). **(C)** Western blot for FBXO42 and CCDC6 levels in GSCs and NSCs nucleofected with CRISPR RNPs targeting *FBXO42* or *CCDC6*.

However, the observed positive gene dependency correlation could still indicate that these two proteins either work together in some way to achieve a particular function or that they act independently and redundantly to produce the same net outcome. In order to explore the relationship between *FBXO42* dependency and *CCDC6* dependency in our system, we assessed relative cell viability after nucleofecting the two exquisitely F42L-sensitive lines GSC-0827 and BT549 and the two F42L-resistant lines GSC-0131 and NSC-U5 with CRISPR RNPs targeting *FBXO42* or *CCDC6* individually or both in combination (Figure 3.8C). The results confirmed that *CCDC6* dependency does correlate with *FBXO42* dependency in our system as well, since *CCDC6* knockout reduced viability in GSC-0827 and BT549 but had no effect in GSC-0131 or NSC-U5. Furthermore, combined loss of *FBXO42* and *CCDC6* had an even greater effect on viability in F42L-sensitive cells than loss of either gene alone. However, interestingly, combined loss of both genes did not have a negative effect in either F42L-resistant line (Figure 3.8C). These results indicate that F42L-sensitive cells require both genes for survival, but that *FBXO42* and *CCDC6* are *not* synthetic lethal. An observed synthetic lethality would be more consistent with two proteins working redundantly to contribute to the same net outcome, so our results suggest that this is probably not the case. Instead, they suggest that FBXO42 and CCDC6 likely work together to achieve a particular function, a function which is required in a subset of cancer cells but not normal cells. This further highlights that therapeutically, combined targeting of FBXO42 and CCDC6 could have a dramatic effect in F42L-sensitive tumors while sparing normal tissues.

3.3.8 *FBXO42 loss-sensitive cells suffer an extended metaphase arrest upon FBXO42 knockout due to prolonged spindle assembly checkpoint activation*

We wanted to gain insight into what could be causing the reduced viability in F42L-sensitive cells. Since many E3 ubiquitin ligases are involved in cell cycle control, we assessed the cell cycle

profiles in *FBXO42* knockout GSCs and NSCs compared to controls using flow cytometry DNA content analysis via DAPI staining (Figure 3.9A-B). Using a Dean Jett Fox model to predict the relative proportions of cells in the G1, S, and G2/M stages of the cell cycle that would result in the observed flow profile, we saw a marked increase in the percent of cells in the G2/M stage in F42L-sensitive lines that had been nucleofected with sgFBXO42, with a concomitant decrease in the percent in G1, while the relative percent in S phase remained stable (Figure 3.9B). The degree of F42L-sensitivity correlated with the increase in percent G2/M cells, as the exquisitely-sensitive lines GSC-0827 and BT549 displayed the largest increase while the moderately-sensitive line GSC-G166 displayed a lesser increase. Importantly, there was no difference in cell cycle profile upon *FBXO42* loss in the F42L-resistant line GSC-0131 (Figure 3.9A-B).

Since it is not possible to distinguish between cells in G2 and M phase when using a DAPI DNA content profile to assess cell cycle, we used time-lapse microscopy in GSCs that had been transduced with a PGK-H2B-EGFP construct to assay mitosis in *FBXO42* knockout cells (Figure 3.9C). We imaged cells (phase and GFP) every 5 minutes to compile time-lapse videos and then assessed the mitotic transit time for individual cells. We found that F42L-sensitive cells with *FBXO42* knockout spent a significantly longer time in mitosis. Once again, the difference correlated with degree of F42L-sensitivity, with GSC-0827 sgFBXO42 cells spending an average of ~3.1 times as long in mitosis as control cells (up to a maximum of ~8.9 times) and GSC-G166 sgFBXO42 cells spending an average of ~1.50 times as long in mitosis as control cells (up to a maximum of ~4.9 times) (Figure 3.9C). We also observed that GSC-0827 cells were not able to overcome their mitotic arrest and eventually suffered cell death during mitosis, while GSC-G166 cells were eventually able to overcome the arrest and exit mitosis. As expected based on the flow cytometry results, there was no significant difference in mitotic transit time between sgFBXO42

and control cells for the F42L-resistant line GSC-0131 (Figure 3.9C). Altogether, these results indicate that loss of *FBXO42* in F42L-sensitive cells leads to a prolonged mitosis, which causes the loss of viability in these cells.

We observed from our time-lapse microscopy that the mitotic arrest seemed to be specifically a metaphase arrest, since we could see a formed metaphase plate in the arrested cells (visualized by H2B-EGFP). We wanted to explore this further and thus fixed sgFBXO42 and control GSC-0827 cells, stained with DAPI and α-αtubulin and α-centromere antibodies, and performed Z-stack ultra-high resolution microscopy to create 3D projections of mitotic cells (Figure 3.9D). This confirmed that the vast majority of arrested cells were indeed at metaphase, and it showed that sgFBXO42 cells suffer from both a distorted spindle and a dramatic increase in chromosome alignment defects, with many misaligned chromosomes present at the spindle poles (Figure 3.9D).

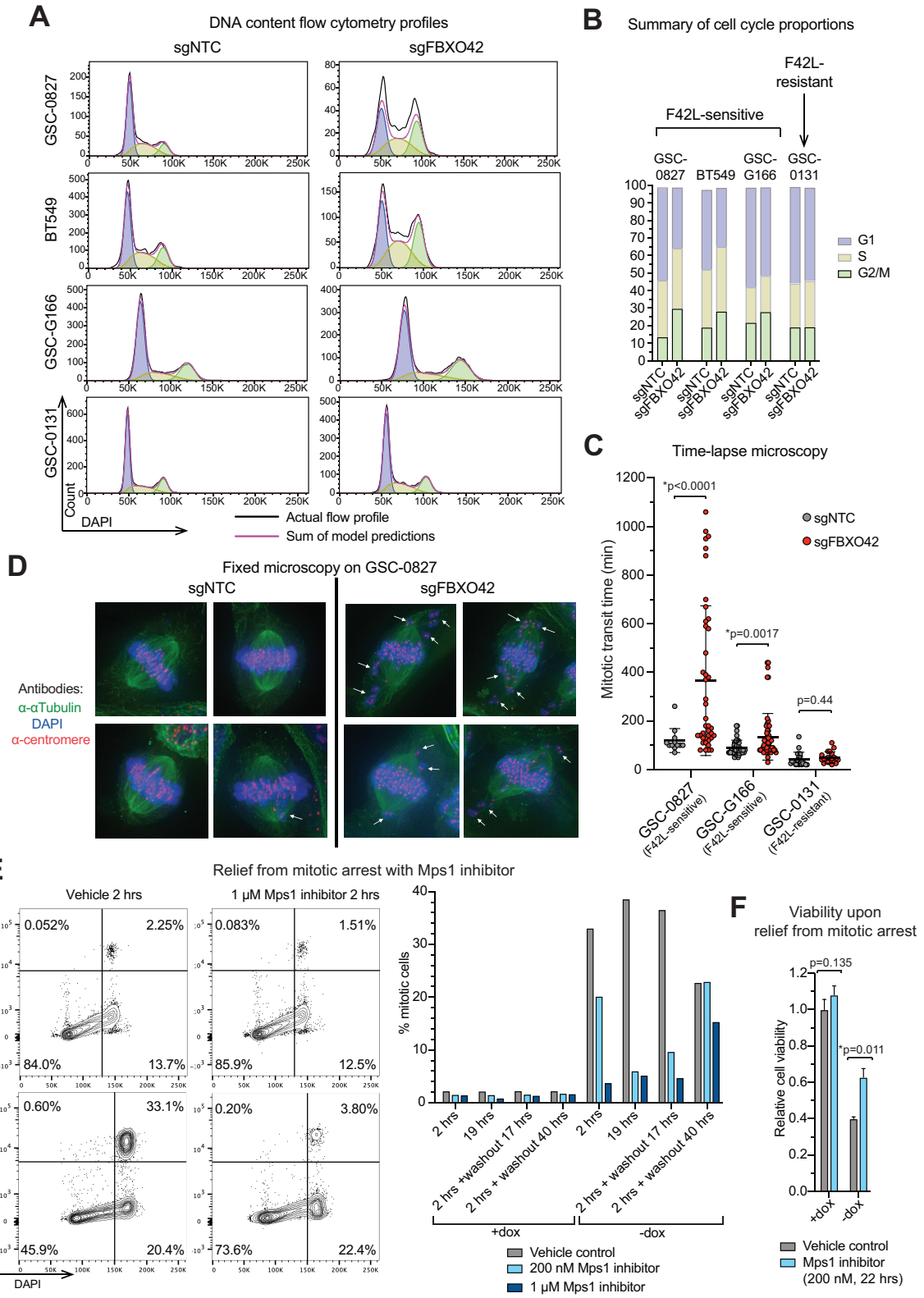


Figure 3.9. *FBXO42* loss-sensitive cells suffer an extended metaphase arrest upon *FBXO42* knockout due to prolonged spindle assembly checkpoint activation. **(A)** DNA content (DAPI) flow cytometry profiles for cells nucleofected with CRISPR RNPs targeting *FBXO42* compared to a non-targeting sgRNA (sgNTC) (4 days post

nucleofection). Dean Jett Fox model for cell cycle distribution (FlowJo software) is shown under the histogram for each sample, with predictions for G1, S, and G2/M in violet, yellow, and green, respectively. **(B)** Summary of cell cycle proportion predictions for samples shown in (A). **(C)** Mitotic transit time, measured using analysis of time-lapse microscopy, for individual H2B-EGFP-expressing cells with or without knockout of *FBXO42*. Bars show mean and standard deviation. **(D)** Representative images of visualization of mitotic spindle structure and chromosome alignment at metaphase upon *FBXO42* knockout using fixed F42L-sensitive GSC-0827 cells stained with DAPI and α-αtubulin and α-centromere antibodies. Images are a result of 3D projections of multiple Z-stacks taken using a DeltaVision Ultra High Resolution Microscope. **(E)** Left panel: Examples of DAPI vs. p-H3 (Ser10) flow cytometry profiles for GSC-0827 doxycycline-inducible *FBXO42* clone 15 kept in +/- doxycycline for 4 days and then treated with vehicle or an Mps1 inhibitor (NMS-P715) for 2 hours. Right panel: Summary of % mitotic cells from flow cytometry analysis (as shown in left panel) for cells that were kept in +/- doxycycline for 4 days and then treated with vehicle/drug for 2 hrs or 19 hrs, or for 2 hrs followed by a 17 hrs/40 hrs washout of the drug. **(F)** Relative viability for cells that were kept in +/- doxycycline for 4 days and then treated with vehicle or 200 nM Mps1 inhibitor for 22 hours.

A metaphase arrest with misaligned chromosomes is consistent with activation of the spindle assembly checkpoint (SAC), which is a feedback-control system in eukaryotic cells that monitors the attachment of kinetochores to the microtubule fibers of the mitotic spindle. Sister chromatids that do not have proper bi-oriented attachments at kinetochores cause SAC signaling, activating the SAC effector, the mitotic checkpoint complex (MCC). The MCC binds and inhibits APC/CCdc20, which is required for the metaphase–anaphase transition, thus preventing entry into anaphase. In this manner, the SAC serves to prevent premature chromosome segregation in the presence of chromosomes that are not properly attached to the spindle, thereby preserving the genome from the disastrous consequences that aneuploidy can bring[245]. The Mps1 kinase serves to phosphorylate the kinetochore protein Knl1[246-248], creating docking sites for the recruitment of additional SAC proteins[249-255].

We wanted to determine if SAC activation was indeed the cause of the prolonged metaphase arrest in F42L-sensitive cells. Since Mps1 promotes assembly of the SAC, we hypothesized that if SAC activation were the cause of the arrest, then F42L-sensitive cells could

overcome their arrest if Mps1 were inhibited and the SAC were thereby relieved. For these experiments, we wanted to model *FBXO42* loss as uniformly time-wise as possible and thus turned to our system of F42L-sensitive GSC-0827 that express doxycycline-inducible exogenous *FBXO42* and have knockout of the endogenous locus. We cultured cells with or without doxycycline for 4 days, a timepoint which corresponded to complete loss of FBXO42 protein (Figure 3.6C) and the beginning of the mitotic arrest phenotype. We then treated these cells with either a high-dose Mps1 inhibitor or vehicle control for 2 hours and performed flow cytometry for DAPI and p-H3 (Ser10) to mark mitotic cells (Figure 3.9E, left panel). From the vehicle controls, we could see that at this timepoint, 33.1% of cells in the -doxycycline condition (*FBXO42* loss) were in mitosis (cells that were p-H3 (Ser10) positive and had a DAPI staining level consistent with the G2/M peak), compared to 2.25% in mitosis in the +doxycycline control condition. Remarkably, -doxycycline cells that were treated with the Mps1 inhibitor for 2 hours only had 3.8% of cells in mitosis, indicating that inhibition of Mps1 did alleviate the metaphase arrest to very near normal mitotic levels.

We expanded upon this experiment in order to test if cells that had been relieved of their initial mitotic arrest would then re-arrest when they entered their next mitosis. To this end, we took cells that had been cultured with or without doxycycline for 4 days and treated them with a high dose Mps1 inhibitor, low dose Mps1 inhibitor, or vehicle control for either 2 hours or 19 hours. We also took cells and treated them for 2 hours, washed out the drug/vehicle, and then allowed the cells to grow for another 17 or 40 hours. After all of these treatments and timepoints, we collected cells and again performed flow cytometry for DAPI and p-H3 (Ser10) (Figure 3.9E, right panel). As expected, the vehicle control showed that in -doxycycline cells, the percent mitotic cells remains fairly consistent between 2 and 19 hours, as some cells die from their arrest but are

replaced by additional newly-arrested cells. By 40 hours, the percent mitotic cells slightly drops off as there are fewer newly-arrested cells remaining to replace the ones that have already died from their arrest. We observed that in the -doxycycline cells treated with both the higher and lower doses of Mps1 inhibitor, the vast majority of cells were at first relieved of their arrest, which took 2 hours for the higher dose inhibitor and up to 19 hours for the lower dose inhibitor. Then, after the drug washout, -doxycycline cells that had previously been treated with the higher dose inhibitor did display a re-arrest after 40 hours. Cells that had been previously treated with the lower dose inhibitor displayed an even greater re-arrest after 40 hours, with a percent mitotic cells equivalent to -doxycycline cells at this timepoint that had never been treated with the Mps1 inhibitor at all (Figure 3.9E, right panel). All of these results clearly demonstrate that SAC activation is indeed the cause of the prolonged metaphase arrest in F42L-sensitive cells.

While we had shown that relieving the SAC would relieve the mitotic arrest in F42L-sensitive cells and that a population of cells could re-arrest, our data did not specifically show that cells that had been relieved would not immediately die upon exiting mitosis. We thus assessed relative cell viability in cells cultured with or without doxycycline for 4 days and then treated with a low dose of Mps1 inhibitor for 22 hours (Figure 3.9F). We observed a significant increase in viability in -doxycycline cells that had been treated with the inhibitor compared to vehicle control -doxycycline cells. This suggests that arrested cells did not die immediately upon being artificially relieved of their arrest, again lending support to the idea that the SAC-triggered metaphase arrest is primarily responsible for the viability phenotype.

Altogether, our data indicate that upon *FBXO42* loss, F42L-sensitive cells suffer reduced viability due to an extended metaphase arrest triggered by prolonged SAC activation, which suggests that FBXO42 activity contributes to inhibition of the SAC.

3.4 DISCUSSION

One popular concept in cancer research is the notion that genomic and molecular profiling of patient samples will enable the discovery of patient-tailored therapeutic strategies. However, it remains unclear whether analytic or computational approaches based solely on descriptive data sets are powerful enough to predict successful therapies. An alternative approach is to directly identify molecular vulnerabilities in patient samples using functional genetic experimentation. Here, we perform functional genomic screens in patient isolates and integrate this data with other publicly-available functional data in order to create co-dependency networks that identify GBM gene vulnerability groups. Furthermore, we demonstrate the utility of incorporating our functional data with descriptive genomics in order to identify cancer cell context-specific vulnerabilities. Lastly, we further investigate the patient-specific screen hit *FBXO42*, a thus far poorly-studied F-box protein that we find is likely required for a subset of tumors of various cancer types, and thus presents a potentially widely applicable therapeutic target that should be explored further.

We performed successive rounds of CRISPR-Cas9 functional genomic screening in order to identify the strongest *in vitro* growth-promoting genes in GBM. Co-dependency networks based on integrating DepMap screening data with our own screen hits then revealed that GBMs preferentially require a group of genes related to mitochondrial function and another group related to metabolic enzymes. The dependency of tumors on mitochondrial function is now well-appreciated, since it is understood that cancer cells require both glycolysis and mitochondrial metabolism to provide the necessary building blocks for macromolecule synthesis as well as ATP and NADPH production[256]. Strategies for targeting mitochondrial metabolism in cancer have mainly focused on: redox capacity[257,258]; bioenergetics, including the drug metformin[259] and electron transport chain inhibitors[260]; and biosynthetic function, including targeting glutamine

catabolism[261] and autophagy[262]. Our results demonstrate that inhibiting mitochondrial protein processing and membrane trafficking specifically could be another promising path for targeting mitochondrial metabolism in GBM.

The therapeutic potential of targeting metabolic enzymes in cancer stem-like cells is also being thoroughly researched, with several studies investigating various genes that we identified in our networks, such as *PGD*[263,264], *ATP1A1*[265,266], *IMPDH2*[267,268], and *MAT2A*[269-271]. However, our results highlight the functional interconnectedness of these genes, and also suggest an interesting potential connection between this group of enzymes and AMPK signaling. Overall, we show that if GSCs in a particular tumor require one of these genes for growth and survival, it is likely that those GSCs also require other genes identified in our networks. Therefore, these co-dependency networks identify groups of gene vulnerabilities that represent potential therapeutic avenues in GBM.

In addition, the identification of context-specific vulnerabilities has long been the pinnacle of the precision medicine movement, particularly since the synthetic lethal interaction between PARP inhibition and *BRCA1* or *BRCA2* mutation was first described in 2005[117,118]. Here, we demonstrate that unbiased functional genomic screening combined with descriptive genomics allows for the identification of context-specific cancer vulnerabilities, focusing specifically on *ADAR*, the primary enzyme responsible for editing dsRNA[126], and *EFR3A*, one of two EFR3 proteins that are required to localize the enzyme PI4KIIIα to the plasma membrane[127,128]. *ADAR* and *EFR3A* serve as examples of vulnerabilities that occur in the context of a specific expression profile / tumor subtype and low expression of a paralog, respectively. The intriguing implication of *ADAR* vulnerability is that this gene could be a potential target for increasing or restoring tumor sensitivity to immunotherapies, as demonstrated by a study which identified *ADAR* as boosting

immunotherapy in an *in vivo* CRISPR knockout screen of melanoma tumors in mice treated with anti-PD1 therapy[272]. Further findings suggest that inflammation caused by *ADAR* deletion can overcome resistance to checkpoint blockade caused by loss of a tumor-specific CD8+ T cell response[228]. These results are likely a product of *ADAR* loss leading to accumulation of endogenous immunogenic dsRNAs, which triggers IFN responses via nucleic acid sensors. *EFR3A*, on the other hand, is an attractive example of a classic synthetic lethal interaction. The enzyme PI4KIIIα has been suggested as a potential cancer target[273-275], but as with many cancer therapies, concerns about possible toxicities on normal cells remain. Since *EFR3A* and *EFR3B* are paralogs that both function to localize PI4KIIIα to the plasma membrane[127,128], where it can exert its activity, targeting *EFR3A* in *EFR3B*-low cancer cells represents a unique way to inhibit PI4KIIIα activity in these tumors while sparing normal cells that express *EFR3B*.

We also became interested in the context-specific vulnerability *FBXO42* due to the uniquely extreme degree with which this gene scored as specific to one particular GSC line in our screens compared to NSC controls. We demonstrate that even in individual retests, *FBXO42* is essential in a subset of GSCs *in vitro* and *in vivo* but is completely nonessential in NSCs *in vitro*. We further show evidence suggesting that *FBXO42* loss-sensitive tumors occur in various cancer types derived from completely different tissue lineages, which implies not only that there is important fundamental biology underlying this gene, but also that it could represent an important therapeutic target in many different cancers. This has never been documented before, and both of these implications demand further study of this F-box protein.

In an effort to explore whether FBXO42's role in an E3 ubiquitin ligase complex is actually responsible for the patient-specific viability phenotype upon *FBXO42* loss, we show that only full-length *FBXO42* and not ΔF-box or ΔKelch versions can rescue the phenotype. However, we also

demonstrate that the published targets p53[129,234] and ING4[236] and the published interactor RBPJ[238] are not mainly responsible for the phenotype. In particular, the evidence suggests that p53 is likely not a primary target responsible for the phenotype we observe. In order to learn more about potential interactors that affect the phenotype, we thus turned once again to functional screening data, which led us to investigate a potential connection with *CCDC6*, which has never been explored thus far. We find that *FBXO42* and *CCDC6* are each necessary to promote viability in *FBXO42* loss-sensitive cells, but not in *FBXO42* loss-resistant cells. We also show that while CCDC6 is likely not a ubiquitination target of FBXO42, these two genes do appear to function together. *FBXO42* and *CCDC6* are not synthetic lethal in resistant cells, but loss of both genes simultaneously does result in even more dramatic viability loss in sensitive cells than *FBXO42* knockout alone. Therefore, our results suggest that *FBXO42* and *CCDC6* likely work together to achieve a particular function, a function which is required in a subset of cancer cells but not normal cells. This further highlights that therapeutically, combined targeting of FBXO42 and CCDC6 could have a dramatic effect in sensitive tumors while sparing normal tissues.

Lastly, through exploring the nature of the viability phenotype in *FBXO42* loss-sensitive cells, we identify a thus-far completely unknown role for FBXO42 in affecting the spindle assembly checkpoint (SAC). We find that upon *FBXO42* knockout, *FBXO42* loss-sensitive cells suffer an extended metaphase arrest due to prolonged SAC activation, which causes a reduction in viability. This suggests that FBXO42 activity contributes to inhibition of the SAC, although it is thus far unclear how exactly it does so. We did observe chromosome alignment defects in sensitive GSC-0827 cells, which is the classic cause of SAC activation (reviewed in[276]). However, the majority of chromosomes were aligned at the metaphase plate, and the number that were unaligned should in theory not be sufficient to cause the extremely prolonged SAC activation that we saw in

GSC-0827 cells[277], which was up to over 17 hours. It is therefore possible that the chromosome misalignment observation is a by-product of additional mitotic functions of FBXO42, but that the true culprit of *prolonged* SAC activation is that FBXO42 directly regulates the SAC proteins themselves, perhaps even directly affecting Mps1 levels. Alternatively, we also observed that while the mitotic spindle in *FBXO42* knockout cells was still able to form and was properly bipolar, it was noticeably twisted and distorted. Therefore, it is also possible that FBXO42 substrate(s) act on the mitotic spindle and thereby cause prolonged activation of the SAC[278]. The exact cellular context of *FBXO42* sensitivity is thus far unclear, but our data suggests that in either scenario, it is possible that it relates to a cell's innate ability to regulate the SAC independently of FBXO42 function, i.e. that the total signaling in some cells is more inclined to push through SAC activation than in other cells[279].

It is also unclear how CCDC6 interacts with FBXO42 to contribute to the above-described processes. However, it has been shown that CCDC6 is important for intra-S-phase checkpoint control and that loss of CCDC6 shortens the S phase transition and causes cells to skip through the DNA damage checkpoint[280]. Another study also showed that following genotoxic stress, loss of *CCDC6* in cancers that carry the *CCDC6-RET* fusion accelerates the dephosphorylation of pH2AX (S139), resulting in premature mitotic entry, potentially without proper DNA damage repair[242]. This concept has thus far been mainly discussed as a role for *CCDC6* as a tumor suppressor, where loss of the gene causes genome instability and thus leads to cancer initiation and/or progression. However, it could also contribute to the cancer-specific vulnerability that we observe here. It has been shown that during mitosis, a DNA damage checkpoint also feeds into the SAC[281]. Therefore, we speculate that it is possible that cells with higher intrinsic levels of DNA damage would be sensitive to loss of *CCDC6/FBXO42* since having skipped the S phase DNA damage checkpoint

would then trigger prolonged SAC activation during mitosis. Additional studies will need to explore these ideas further.

Overall, all of our results highlight the importance of integrating functional genomic approaches into precision oncology paradigms to identify new therapeutic strategies for GBM and other cancers.

3.5 METHODS

3.5.1 *Cell Culture*

Patient tumor-derived GSCs and fetal tissue-derived NSCs were provided by Drs. Do-Hyun Nam, Jeongwu Lee, and Steven M. Pollard, were obtained via informed consent, and have been previously published[17,32,177]. Isolates were cultured in NeuroCult NS-A basal medium (StemCell Technologies) supplemented with B27 (Thermo Fisher), N2 (homemade 2x stock in Advanced DMEM/F-12 (Thermo Fisher)), EGF and FGF-2 (20 ng/ml) (PeproTech), glutamax (Thermo Fisher), and antibiotic-antimycotic (Thermo Fisher). Cells were cultured on laminin (Trevigen or in-house-purified) -coated polystyrene plates and passaged as previously described[32], using Accutase (EMD Millipore) to detach cells.

3.5.2 *Lentivirus Production*

The day before transfection, 293T cells were seeded by splitting a confluent plate at a 1:5 surface area ratio onto plates that had been coated with poly-L-lysine (Sigma) for 20 minutes and then washed with PBS. To prepare transfection mixture for a 10cm plate, 12μg of lentiviral transfer vector of interest was combined with 8μg of psPAX2 and 3μg of pMD2.G (Addgene) in 1mL of serum-free DMEM. 34.5μg of polyethylenimine (Polysciences Inc.) was then added to the DNA mixture, and the solution was vortexed briefly, incubated at room temperature for 20 minutes, and

then added drop-wise to cells with gentle shaking. 24 hours after transfection, media on the cells was changed to NSC/GSC culture media for generation of unconcentrated lentivirus or fresh DMEM containing 10% FBS for generation of concentrated lentivirus. 48 hours after transfection, the supernatant containing the virus was harvested and filtered through a 0.45μm pore-size filter. For individual experiments using unconcentrated lentivirus, this filtered supernatant was then stored at -80°C and subsequently directly used to infect cells. For concentrated lentivirus (e.g. for use in CRISPR-Cas9 screens), this filtered supernatant was centrifuged for 20-24 hours at 6,000g and then resuspended in serum-free DMEM to create ~100X concentrated virus, which was stored at -80°C.

3.5.3 *CRISPR-Cas9 Functional Genomic Screening*

Our genome-wide screens in GSCs and NSCs have been previously published and described[73]. For generating our comprehensive retest library, sgRNAs for chosen genes and controls were mined from the human GeCKO v2 library[189] as well as our whole genome screen results, and an oligo pool representing these sgRNAs and PCR adapters was obtained (Twist Bioscience). The oligo pool was PCR amplified using Herculase II Fusion DNA Polymerase (Agilent) and cloned into lentiCRISPRv2 puro vector (Addgene) using Gibson Assembly Master Mix (New England Biolabs). The assembled pool was then transformed into Stellar Competent Cells (Clontech) and plated onto LB agar plates (liquid culture was avoided in order to minimize competition between clones containing different sgRNAs). The resulting colonies were scraped from the plates and the finished lentiCRISPRv2 plasmid comprehensive retest library was extracted using a NucleoBond Xtra Midi Endotoxin-Free Kit (Macherey–Nagel). This lentiviral library was then used to generate virus and infect cells for outgrowth screening. For screening, cells were transduced to achieve ~750X representation of the library (at ~30% infection efficiency to ensure a high proportion of

single integrants). 2 days after transduction, media was replaced with media containing 2 μg/mL puromycin. After 3 days of selection, portions of cells representing 500-750X coverage of the library were collected as the "Day_0" samples. The remaining cells were cultured and consistently maintained at 500-750X representation for 21-23 days, after which time the "Day_final" samples were collected. Screening was carried out in triplicate. To read out screen results, genomic DNA was extracted using the QIAamp DNA Blood Mini Kit (QIAGEN), and a two-step PCR procedure was used to first amplify the genomically integrated sgRNA sequences and then to incorporate Illumina deep sequencing adapters and barcodes onto the sgRNA amplicons. For the first round of PCR, a sufficient number of PCR reactions were carried out to use all gDNA from the 500-750X coverage sample of cells at 2μg genomic DNA per PCR reaction, using MagniTaq Multiplex PCR Master Mix (Affymetrix) and 12 cycles. For the second round of PCR, 5μL of the first round product was used as a template in combination with primers that would add the deep sequencing adapters and barcodes, using Herculase II Fusion DNA Polymerase (Agilent) and 16 cycles. Amplicons from the second round PCR were then column purified using the PureLink Quick PCR Purification Kit (Invitrogen). Purified PCR products were sequenced using HiSeq 2500 (Illumina). Bowtie[282] was used to align the sequenced reads to the sgRNA library, allowing for 1 mismatch. The R/Bioconductor package edgeR[156] was used to assess changes across groups.

3.5.4 *Network Generation*

Networks were created using Cytoscape software. For the initial larger network, the top 100 co-dependency correlations per gene were downloaded from the DepMap[102-105] database (CRISPR Avana Public 19Q4 dataset) for top hits from our comprehensive retest screens (genes scoring in 1, 2, or all 3 GSC lines). This data was then filtered to keep only correlations with a correlation coefficient >0.2, and the resulting information was used to create the network. Based on the results

of this larger network, two subnetworks (mitochondrial-related and metabolism-related) were then created using screen hits that clustered together in the larger network. The DepMap[102-105] co-dependency correlations for these genes were filtered for a correlation coefficient >0.25, and to avoid an overly complicated network, only a maximum of the top 30 correlations per gene were kept. Furthermore, negative correlations were filtered out since their meaning is more difficult to interpret. In addition to DepMap[102-105] co-dependency correlations, data regarding physical interactions, predicted interactions, and shared pathways were downloaded from GeneMania[192] for the gene hits in the subnetworks, using Gene Ontology Biological Process-based network weighting. Data with a weight >0.05 were kept and incorporated into the networks. In this manner, subnetworks for our screen hits were created by combining DepMap[102-105] and GeneMania[192] data. Edge widths were created to correspond to the value of the correlation coefficient for DepMap-derived correlations and to the weight of the association for GeneMania-derived data. After creation of the networks, genes were marked by different border colors to identity gene hits with a GSC-NSC Zscore differential <-0.5 for all GSCs in which a particular gene scored.

3.5.5 *CRISPR sgRNA Design for Nucleofection*

CRISPR sgRNA sequences that scored in our screens were also used in chemically synthesized form for RNP nucleofections. Additional sgRNA sequences were designed via manual curation of all possible sgRNA sequences for a given gene as identified by the Broad Institute's GPP Web Portal[152]. See Table S1 for a list of sgRNA sequences used.

3.5.6 *Cas9:sgRNA RNP Nucleofection*

See detailed protocol previously published. Briefly, to prepare RNP complexes, reconstituted chemically synthesized and 2'-O-methyl 3'phosphorothioate-modified sgRNA (Synthego) and

then sNLS-SpCas9-sNLS (Aldevron) were added to complete SG Cell Line Nucleofector Solution (Lonza), to a final volume of 20 μL. The mixture was incubated at room temperature for 15 minutes to allow RNP complexes to form. A Cas9:sgRNA molar ratio of 1:2 was used, unless otherwise noted. A total RNP dose of 15-40 pmol was used (depending on the cell line), and total dose refers to the amount of the limiting complex member (Cas9). To nucleofect, 1.5-2.5 x 10^5 cells (depending on the cell line) were harvested, washed with PBS, resuspended in 20 μL of RNPs, and electroporated using the Amaxa 96-well Shuttle System or 4D X Unit (Lonza) and program EN-138. After nucleofection, cells were plated onto 24-well, 12-well, or 6-well plates (depending on the cell line), and media was changed ~24 hours later.

3.5.7 *CRISPR Editing Analysis*

Nucleofected cells were harvested at indicated timepoints and genomic DNA was extracted (MicroElute Genomic DNA Kit, Omega Bio-Tek). Genomic regions around CRISPR target sites were PCR amplified using Phusion polymerase (Thermo Fisher) and primers located (whenever possible) at least 250bp outside cut sites. After size verification by agaorse gel electrophoresis, PCR products were column-purified (Monarch PCR & DNA Clean-up Kit, New England BioLabs) and submitted for Sanger sequencing (Genewiz) using unique sequencing primers. The resulting trace files for edited cells versus control cells (nucleofected with non-targeting Cas9:sgRNA) were analyzed for predicted indel composition using the Inference of CRISPR Edits (ICE) web tool[151]. See Table S1 for a list of all PCR and sequencing primers used.

3.5.8 *Western Blotting*

Cells were harvested, washed with PBS, and lysed with modified RIPA buffer (150mM NaCl, 25mM Tris-HCl (pH 8.0), 1mM EDTA, 1.0% Igepal CA-630 (NP-40), 0.5% sodium deoxycholate,

0.1% SDS, 1X protease inhibitor cocktail (complete Mini EDTA-free, Roche)). Lysates were sonicated (Bioruptor, Diagenode) and then quantified using Pierce BCA assay (Thermo Fisher). Identical amounts of proteins (20-40μg) were electrophoresed on 4–15% Mini-PROTEAN TGX precast protein gels (Bio-Rad). For transfer, the Trans-Blot Turbo transfer system (Bio-Rad) with nitrocellulose membranes was used according to the manufacturer's instructions. TBS (137mM NaCl, 20mM Tris, pH 7.6) +5% nonfat milk was used for blocking, and TBS+0.1%Tween-20+5% milk was used for antibody incubations. The following commercial primary antibodies were used: MX1 (Cell Signaling #37849S, 1:500), Tp53 (Cell Signaling #48818, 1:500), αTubulin (Sigma #T9026, 1:1,000), βActin (Cell Signaling #3700S, 1:2,000). The following secondary antibodies were used (LI-COR, all 1:10,000): #926-68073, #926-32212, #926-32214, #926-68074, #925-32212, #925-68071. An Odyssey infrared imaging system (LI-COR) was used to visualize blots.

3.5.9 *Viability Assays*

For viability assays when targeting individual genes, cells were given 2-3 days post nucleofection to recover. Cells were then counted (NucleoCounter, NBS) and equal numbers of all samples for a particular cell line were plated in triplicate into 96-well plates at various dilutions to allow for multiple timepoints to be assessed. Cells were fed with fresh medium every 2-3 days. After 6-12 days under standard growth conditions (timepoint depending on cell doubling time), relative viable cell numbers were measured using CellTiter-Glo Luminescent Cell Viability Assay (Promega) according to manufacturer's instructions.

3.5.10 *Time-lapse Microscopy*

Cells were first infected with lentivirus for PGK-H2B-EGFP (Addgene) to create stable cell lines in which the nuclei could be visualized easily. Cells were then nucleofected with CRISPR RNPs

as needed and plated into 24-well or 12-well plates. 1-3 days post nucleofection (depending on cell line and onset of gene knockout phenotype), cells were placed into an IncuCyte S3 (Sartorius) instrument, which was in an incubator set to normal culturing conditions. For the mitotic transit time analysis, phase and fluorescence (GFP) images were taken every 5 minutes for 48-72 hours. Videos were compiled using the IncuCyte S3 software, and mitotic transit time was then analyzed for individual cells. A cell was considered to enter mitosis when nuclear envelope breakdown was evident based on EGFP visualization and when a visible morphology change from flat to round was observed. Following successful cytokinesis (proper cytoplasmic division resulting in two daughter cells), a cell was categorized as having successfully completed mitosis. A cell was classified as cytokinesis failure if the cell failed to divide following mitotic entry due to an abrupt mitotic exit while in metaphase or anaphase, or failure to complete cytokinesis. If a cell seemed to experience cytokinesis failure, it was followed for additional time to ensure that this was indeed the case. A cell was categorized as cell death in mitosis if a cell erupted and died during mitosis (between nuclear envelope breakdown and cytokinesis).

3.5.11 *RNA-seq and Analysis*

Cells were lysed with Trizol (Thermo Fisher). Total RNA was isolated (Direct-zol RNA Kit, Zymo Research) and quality validated on the Agilent 2200 TapeStation. Illumina sequencing libraries were generated with the KAPA Biosystems Stranded RNA-Seq Kit[178] and sequenced using HiSeq 2000 (Illumina) with 100bp paired-end reads. RNA-seq reads were aligned to the UCSC hg19 assembly using STAR2 (v 2.6.1)[179] and counted for gene associations against the UCSC genes database with HTSeq[180]. Normalized gene count data was used for subsequent hierarchical clustering (R package ggplot2[181]) and differential gene expression analysis (R/Bioconductor package edgeR[156]). Heatmaps were made using R package pheatmap[182].

3.5.12 *Statistics*

All Student's t-tests were performed as unpaired 2-tailed tests, assuming unequal variance.

3.5.13 *In Vivo Experiments*

All *in vivo* experiments were conducted in accordance with the NIH Guide for the Care and Use of Experimental Animals and with approval from the Fred Hutchinson Cancer Research Center, Institutional Animal Care and Use Committee (Protocol 1457). Two days prior to tumor cell implant, 2mg/ml doxycycline with 5% sucrose (w/v) was added to the mouse drinking water of adult, female, HSD:athymic nude Foxn1nu mice (#069, Envigo). A cell clone was derived from GSC-0827 that had been transduced with doxycycline-inducible pTT-FBXO42(sgRNA#1-resistant)-V5 and nucleofected with CRISPR RNPs to knock out the endogenous *FBXO42* gene. Cells from this clone were implanted at 100,000 cells per mouse on Day 0 as previously described[283]. Mice underwent bioluminescent imaging for luciferase expression of the tumor beginning on Day 1 and repeated twice weekly. On Day 7 after tumor implant, mice were separated into experimental cohorts, normalizing the bioluminescent tumor burden in each group. One group was maintained on doxycycline water while the second group was given untreated water. Fresh water with or without doxycycline and sucrose was provided weekly. A subset of mice was euthanized at planned time points on study days 14 and 17 (7 and 10 days after doxycycline removal). For the survival curve, mice were euthanized when they became moribund due to tumor burden.

Chapter 4. DISCUSSION

4.1 CONSIDERATIONS FOR APPLYING CRISPR TECHNOLOGY

As I have demonstrated in this body of work, CRISPR technology is an incredibly useful tool for exploring gene function due to its efficiency and simplicity. However, as with any relatively new technique, we are still exploring the "rules" that govern it. For instance, even though the targeting specificity of Cas9 is controlled by the sgRNA sequence as well as the presence of a PAM adjacent to the target sequence in the genome, reports have shown that off-target cleavage can still occur with even 3-5 bp mismatches in the PAM-distal portion of the sgRNA sequence[284-288]. Some studies, such as a study that corrected mutations causing beta-thalassemia using iPSCs[289], do find that off-target effects are rare, but the salient point is that with the plethora of genes being targeted and applications being explored, restricting potential off-target effects needs to be a top priority. This area of current investigation is being facilitated by the development of additional methods of detecting CRISPR off-target effects, such as integrase-defective lentiviral vectors[290], Digenome-seq (*in vitro* digestion of genomic DNA with Cas9 and gRNA(s) of interest, followed by fragmentation and whole genome next-generation sequencing)[291], and GUIDE-seq (integration of double-stranded oligodeoxynucleotides into CRISPR-induced double-strand breaks, followed by amplification and next-generation sequencing)[292].

While the method of Cas9:sgRNA ribonucleoprotein (RNP) nucleofection that I describe in Chapter 2 does have the potential to result in reduced off-target effects compared to lentiviral methods, due to the short-term activity of RNP complexes, it is still important to further limit potential off-target effects. This can be done by using the lowest effective doses of both Cas9 and sgRNA, since it has been shown that high complex concentrations can cleave additional, more

promiscuous off-target sites that are not cleaved when enzyme concentrations are limiting[293]. The appropriate doses for a particular cell type at hand should thus be empirically determined. Furthermore, in any scenario, sgRNA sequences should be carefully designed to avoid potential off-target effects, and there are many freely-available bioinformatics tools for this purpose[152,294-297]. Whenever possible, care should be taken to consider potential off-target sites by analyzing not only mismatches but also the influence of gaps/bulges between any given gRNA and target DNA[287,298], and by accounting for the fact that PAM-proximal mismatches are generally less tolerated than PAM-distal mismatches[287,299-302]. Furthermore, off-target cleavage can be further reduced by using high-fidelity Cas9 mutants[303,304] or nickase versions of Cas9 that produce single-strand breaks (and employ a dual sgRNA system)[305-307].

An additional consideration is that in some cases the repair outcome of CRISPR-induced double-strand breaks can be predicted based on sgRNA sequence characteristics[308], and taking this into account can be useful when the goal is to create out-of-frame indels using individual sgRNAs (rather than targeted deletions using multiple sgRNAs). Another vital factor that should be taken into account when employing CRISPR using lentiviral methods is that stretches of a few thymines can cause early termination from Pol III promoters that are commonly used for sgRNA expression (U6, 7SK, and H1)[309], meaning that sgRNA sequences with multiple consecutive thymines should be avoided.

Moreover, it has been suggested that the DNA damage associated with CRISPR cutting can lead to negative viability effects independently of whether a gene is truly required or not[103,230-232], indicating that it is important to compare the effects of targeting a gene of interest to those of a non-expressed gene or an intergenic region. My work shows that indeed, some cell lines are highly sensitive to CRISPR cutting, with a reduction in viability of up to almost 50%. This is a

consideration that the vast majority of studies are not taking into account, with most published viability experiments still only using non-targeting sgRNAs. This also again highlights the importance of designing sgRNAs that have limited off-target effects so that cutting at additional, unintended target sites does not cause confounding viability loss. Furthermore, it demonstrates the importance of not overinterpreting viability results when targeting highly amplified regions in cancer cells.

Studies involving CRISPR systems are being released at an astonishing rate, and our understanding of this valuable and widely-applicable technology will only continue to grow.

4.2 Limitations of Functional Genomic Screening for Therapeutic Target Identification

While using functional genomic screens to identify prospective therapeutic targets clearly has immense potential, there are some limitations. CRISPR-Cas9 technology in particular has both the advantage and disadvantage of the ability to completely knock out a gene rather than knocking it down. The advantage is that genes which are identified as extremely strongly scoring differentially in cancer cells versus normal cells (such as my gene hit *FBXO42*) may have much larger therapeutic windows than genes that are similarly identified using siRNA or shRNA technology. However, this is also a disadvantage since CRISPR-Cas9 screening in this manner may not be able to identify some genes that are essential in all cells but nonetheless display a differential cancer sensitivity upon partial inhibition. Also, partial inhibition via siRNAs or shRNAs is more akin to modeling the effect that a future drug for a particular target might have. Still, my CRISPR-Cas9 screening data suggests that some such targets can still be identified by using differential cutoffs rather than absolute cutoffs.

A general limitation of using functional genomic screening to identify potential therapeutic targets is that not all targets identified will be druggable. In this case, having an understanding of the biology involved in the genetic vulnerability can shed light onto cooperating pathway members that may represent druggable proteins. Still, even if a protein is druggable, the process of target discovery to drug discovery and development, pre-clinical testing, and clinical testing is extremely lengthy, taking about 17 years on average[310,311]. Furthermore, even if a drug can be developed, the chance of it demonstrating clinical efficacy is low, with only 3.4% of oncology drugs currently succeeding in clinical trials[312]. Many drugs fail due to a lack of target specificity, lack of penetration into solid tumors (and ability to cross the blood-brain barrier in the case of brain tumors), and unforeseen and intolerable side effects. The alternative approach of performing well-by-well style drug screens to identify novel indications for drugs that are already approved for clinical use greatly reduces the amount of time it takes for new treatment regiments to reach patients, and is therefore also a valuable pursuit. However, this obviously in no way precludes the need to perform functional studies to identify additional novel and improved therapeutic targets, as our understanding of the intricacies of cancer is constantly evolving.

Furthermore, functional genomic screens are mainly performed in cell lines[102-105,313,314], and the lingering question remains of how well established cell lines reflect the cancer biology seen in patient tumors. Comparing cell lines characterized as part of the DepMap effort with patient tumors from The Cancer Genome Atlas (TCGA) has shown that some cell line models vary significantly from primary tumors of the same tissue lineage with regards to copy-number changes, mutations, and mRNA expression profiles[190]. Over time, different cell lines grown in culture tend to resemble each other more than they resemble primary tumors, suggesting that growth on plastic allows for selection of a tumor subclone and/or evolution of the cells due to *in vitro* growth

pressures that are very unique from *in vivo* growth pressures[191]. Patient isolates (such as GSCs) that are cultured under serum-free conditions are known to more faithfully represent their primary tumors[108] and thus represent an improved model for functional genomic screening that seeks to identify potential therapeutic targets that will be applicable to patients. However, currently we as a scientific community probably do not have enough cell lines available to us for functional screening – and even far fewer patient-derived isolates – to even partially recapitulate the tremendous diversity of combinations of cell lineages and genomic alterations that is seen in patient tumors. Therefore, the derivation of additional patient isolates (cultured under serum-free conditions) and the integration of these isolates into routine lab use in functional genomic screening should be a top priority.

Another related limitation is that performing screens in cell lines does not assay the entire complexity and heterogeneity of the cancer cells within a patient tumor. Even GSC isolates and GSC-derived PDX tumor models do not recreate the complete cellular repertoire of the original patient tumor, likely representing a tumor subpopulation which responds well to the defined *in vitro* outgrowth conditions. This means that the therapeutic targets identified using such isolates or cell lines is limited to that particular subpopulation of the tumor. This challenge is similarly seen when using many drugs to treat patient tumors, since patients that initially respond often relapse, which can occur due to clonal selection of a rare treatment-resistant subclone. However, this scenario precisely demonstrates the need for incorporating both descriptive genomics and functional genomics into the clinic. If we can accurately identify and characterize (by mutation, CNV, and expression analysis) the diverse subpopulations of cells that exist within a patient's tumor, we can then apply combination therapies targeting various context-specific vulnerabilities that have been identified for those particular subpopulations.

Furthermore, the impact of the tumor microenvironment on tumor initiation and progression and treatment response cannot be underestimated. It is now well-recognized that tumors consist not only of stem-like cells and their progeny but also a network of various cell types such as endothelial cells, pericytes, immune cells, fibroblasts, and more[315]. It has been demonstrated that simultaneously targeting tumor cells and their microenvironment leads to improved treatment response. For instance, this strategy resulted in significantly increased survival in a GBM preclinical mouse model where tumor cells were targeted with a PI3K inhibitor while macrophages and microglia were targeted with an inhibitor of Colony-stimulating factor 1 receptor (*CSF1R*)[316]. *In vitro* functional genomic screens cannot properly assay the contribution of the microenvironment or the complex signaling that happens *in vivo*. However, in recent years CRISPR-Cas9 screens have been performed in *in vivo* models as well, demonstrating the power and simplicity of this technology. The first *in vivo* CRISPR-Cas9 screen sought to identify genes that modulate tumor metastasis by transducing a non-metastatic cell line *in vitro*, transplanting it subcutaneously into nude mice, monitoring for metastasis to the lung, and sequencing the resulting lung metastases for the integrated sgRNAs[317]. Additional *in vivo* CRISPR-Cas9 screens have since been performed to identify tumor suppressors[318-320], oncogenes[321], synthetically lethal genes[322], and regulators of cancer immunotherapy[272,323]. Another particularly compelling example of the power of *in vivo* screening in identifying modulators of immune cell contribution is a study that performed genome-scale CRISPR-Cas9 screens in CD8+ T cells under cancer immunotherapy settings to identify regulators of tumor infiltration and degranulation[324]. This screen re-identified canonical immunotherapy targets such as PD-1 and Tim-3, along with new modulators such as the RNA helicase *DHX37*, whose knockout in CD8+ T cells enhanced the efficacy of antigen-specific CD8+ T cells against triple-negative breast cancer[324]. These studies demonstrate that functional

genomic screening can be applied *in vivo* as well to glean a variety of useful information about both tumor biology and potential therapeutic avenues, and more complex systems are constantly being developed for this purpose.

Lastly, another limitation of standard pooled functional genomic screening is that it cannot assay the effect of individual (or combinatorial) gene perturbations on complex cellular responses such as changes in transcriptional profiles. Rather, when such complex phenotypes are being assayed, it is usually necessary to perform perturbations in separate plates/wells and read out the results individually as well, which can be labor intensive, time-consuming, and cost ineffective. However, as a recent method termed Perturb-seq[325,326] demonstrates, it is possible to develop techniques that integrate pooled functional genomic screening with complex readouts. Perturb-seq combines single-cell RNA sequencing (scRNA-seq) and CRISPR-based perturbations to perform many such assays in a large-scale pooled format[325,326]. In this technique, the viral vector used for pooled transduction encodes expression of an sgRNA as well as a separate guide barcode that relates the identity of the sgRNA. After cells are infected, they can be grown, differentiated, stimulated, selected for some marker, etc. ScRNA-seq is then performed in a manner that tags each cell's mRNAs, including the guide barcode, with a unique cell barcode and a unique molecular identifier. Therefore, a cell's delivered genetic perturbation can be associated with its transcriptional profile. By infecting cells at a variety of MOIs, the combined epistatic effect of multiple perturbations can also be assessed, since each individual perturbation relates a different guide barcode while one cell contains the same cell barcode[325,326]. This method is an example of techniques that can be developed to dramatically increase the scope of pooled functional genomic screens.

The last two decades have been defined by massive advancements in high-throughput next-generation sequencing technologies, which has corresponded with enormous advancements in cancer omics characterization. The importance of also incorporating functional genomic screens with this data is now being recognized and is steadily gaining popularity, especially since the advent of CRISPR-Cas9 technology. The use of CRISPR–Cas9 screens to supplement the data already produced by RNAi techniques will continue to ascertain many valuable cancer vulnerabilities. I envision that we are in an era where the amalgamation of many such studies produced by the science community, including the body of work presented here, will ultimately lead to improved therapeutic strategies for patients battling cancer.

REFERENCES

1. Goodenberger, M.L. & Jenkins, R.B. Genetics of adult glioma. *Cancer Genetics* **205**, 613-621 (2012).
2. Louis, D.N. et al. The 2007 WHO classification of tumours of the central nervous system. *Acta Neuropathologica* **114**, 97-109 (2007).
3. American Cancer Society *American Cancer Society: Cancer Facts and Figures 2010*.
4. Stupp, R. et al. Radiotherapy plus concomitant and adjuvant temozolomide for glioblastoma. *N Engl J Med* **352**, 987-996 (2005).
5. CBTRUS, Vol. 2012 (www.cbtrus.org, 2011).
6. Latera, J. & Brem, H. in Diseases of the Nervous System: Clinical Neuroscience and Theraputic Principals
, Vol. 2, Edn. Third Edition. (ed. A. Asbury, McKhann, McDonald, Goadsby, McArthur) 1431-1466 (Cambridge University Press, Cambridge, UK; 2002).
7. O'Reilly, S.M. et al. Temozolomide: a new oral cytotoxic chemotherapeutic agent with promising activity against primary brain tumours. *Eur J Cancer* **29A**, 940-942 (1993).
8. Hegi, M.E. et al. MGMT gene silencing and benefit from temozolomide in glioblastoma. *N Engl J Med* **352**, 997-1003 (2005).
9. Chen, J. et al. A restricted cell population propagates glioblastoma growth after chemotherapy. *Nature* **488**, 522-526 (2012).
10. Lan, X. et al. Fate mapping of human glioblastoma reveals an invariant stem cell hierarchy. *Nature* **549**, 227-232 (2017).
11. Bao, S. et al. Glioma stem cells promote radioresistance by preferential activation of the DNA damage response. *Nature* **444**, 756-760 (2006).
12. Pollard, S.M. et al. Glioma stem cell lines expanded in adherent culture have tumor-specific phenotypes and are suitable for chemical and genetic screens. *Cell stem cell* **4**, 568-580 (2009).
13. Lee, J. et al. Tumor stem cells derived from glioblastomas cultured in bFGF and EGF more closely mirror the phenotype and genotype of primary tumors than do serum-cultured cell lines. *Cancer Cell*, 391-403 (2006).
14. Stiles, C.D. & Rowitch, D.H. Glioma stem cells: a midterm exam. *Neuron* **58**, 832-846 (2008).
15. Mangiola, A. et al. Stem cell marker nestin and c-Jun NH2-terminal kinases in tumor and peritumor areas of glioblastoma multiforme: possible prognostic implications. *Clinical cancer*

research : an official journal of the American Association for Cancer Research **13**, 6970-6977 (2007).

16. Gangemi, R.M. et al. SOX2 silencing in glioblastoma tumor-initiating cells causes stop of proliferation and loss of tumorigenicity. *Stem cells (Dayton, Ohio)* **27**, 40-48 (2009).
17. Lee, J. et al. Tumor stem cells derived from glioblastomas cultured in bFGF and EGF more closely mirror the phenotype and genotype of primary tumors than do serum-cultured cell lines. *Cancer Cell* **9**, 391-403 (2006).
18. Ozawa, T. et al. Most Human Non-GCIMP Glioblastoma Subtypes Evolve from a Common Proneural-like Precursor Glioma. *Cancer Cell* **26**, 288-300 (2014).
19. Alcantara Llaguno, S. et al. Malignant Astrocytomas Originate from Neural Stem/Progenitor Cells in a Somatic Tumor Suppressor Mouse Model. *Cancer Cell* **15**, 45-56 (2009).
20. Kwon, C.-H. et al. Pten Haploinsufficiency Accelerates Formation of High-Grade Astrocytomas. *Cancer Research* **68**, 3286 (2008).
21. Llaguno, S.A., Chen, J., Kwon, C.H. & Parada, L.F. Neural and Cancer Stem Cells in Tumor Suppressor Mouse Models of Malignant Astrocytoma. *Cold Spring Harbor Symposia on Quantitative Biology* **73**, 421-426 (2008).
22. Zheng, H. et al. p53 and Pten control neural and glioma stem/progenitor cell renewal and differentiation. *Nature* **455**, 1129-1133 (2008).
23. Zhu, Y. et al. Early inactivation of p53 tumor suppressor gene cooperating with NF1 loss induces malignant astrocytoma. *Cancer Cell* **8**, 119-130 (2005).
24. Marumoto, T. et al. Development of a novel mouse glioma model using lentiviral vectors. *Nature Medicine* **15**, 110-116 (2009).
25. Bachoo, R.M. et al. Epidermal growth factor receptor and Ink4a/Arf: Convergent mechanisms governing terminal differentiation and transformation along the neural stem cell to astrocyte axis. *Cancer Cell* **1**, 269-277 (2002).
26. Endersby, R., Zhu, X., Hay, N., Ellison, D.W. & Baker, S.J. Nonredundant Functions for Akt Isoforms in Astrocyte Growth and Gliomagenesis in an Orthotopic Transplantation Model. *Cancer Research* **71**, 4106 (2011).
27. Liu, K.-W. et al. SHP-2/PTPN11 mediates gliomagenesis driven by PDGFRA and INK4A/ARF aberrations in mice and humans. *The Journal of Clinical Investigation* **121**, 905-917 (2011).
28. Ghazi, S.O. et al. Cell of Origin Determines Tumor Phenotype in an Oncogenic Ras/p53 Knockout Transgenic Model of High-Grade Glioma. *Journal of Neuropathology & Experimental Neurology* **71**, 729-740 (2012).

29. Muñoz, D.M. et al. Differential transformation capacity of neuro-glial progenitors during development. *Proceedings of the National Academy of Sciences* **110**, 14378 (2013).
30. Muñoz, D.M. et al. Loss of p53 cooperates with K-ras activation to induce glioma formation in a region-independent manner. *Glia* **61**, 1862-1872 (2013).
31. Radke, J., Bortolussi, G. & Pagenstecher, A. Akt and c-Myc Induce Stem-Cell Markers in Mature Primary p53−/− Astrocytes and Render These Cells Gliomagenic in the Brain of Immunocompetent Mice. *PLOS ONE* **8**, e56691 (2013).
32. Pollard, S.M. et al. Glioma stem cell lines expanded in adherent culture have tumor-specific phenotypes and are suitable for chemical and genetic screens. *Cell Stem Cell* **4**, 568-580 (2009).
33. Brennan, C.W. et al. The somatic genomic landscape of glioblastoma. *Cell* **155**, 462-477 (2013).
34. Chow, L.M.L. et al. Cooperativity within and among Pten, p53, and Rb Pathways Induces High-Grade Astrocytoma in Adult Brain. *Cancer Cell* **19**, 305-316 (2011).
35. Brennan, C. et al. Glioblastoma subclasses can be defined by activity among signal transduction pathways and associated genomic alterations. *PLoS ONE* **4** (2009).
36. Phillips, H.S. et al. Molecular subclasses of high-grade glioma predict prognosis, delineate a pattern of disease progression, and resemble stages in neurogenesis. *Cancer Cell* **9**, 157-173 (2006).
37. Verhaak, R.G. et al. Integrated genomic analysis identifies clinically relevant subtypes of glioblastoma characterized by abnormalities in PDGFRA, IDH1, EGFR, and NF1. *Cancer Cell* **17**, 98-110 (2010).
38. Neftel, C. et al. An Integrative Model of Cellular States, Plasticity, and Genetics for Glioblastoma. *Cell* **178**, 835-849 e821 (2019).
39. Patel, A.P. et al. Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma. *Science* **344**, 1396-1401 (2014).
40. Darmanis, S. et al. Single-Cell RNA-Seq Analysis of Infiltrating Neoplastic Cells at the Migrating Front of Human Glioblastoma. *Cell Rep* **21**, 1399-1410 (2017).
41. Wang, L. et al. The Phenotypes of Proliferating Glioblastoma Cells Reside on a Single Axis of Variation. *Cancer discovery* **9**, 1708-1719 (2019).
42. Tirosh, I. et al. Single-cell RNA-seq supports a developmental hierarchy in human oligodendroglioma. *Nature* **539**, 309-313 (2016).
43. Venteicher, A.S. et al. Decoupling genetics, lineages, and microenvironment in IDH-mutant gliomas by single-cell RNA-seq. *Science* **355**, eaai8478 (2017).
44. Filbin, M.G. et al. Developmental and oncogenic programs in H3K27M gliomas dissected by single-cell RNA-seq. *Science* **360**, 331-335 (2018).

45. Squatrito, M. et al. Loss of ATM/Chk2/p53 pathway components accelerates tumor development and contributes to radiation resistance in gliomas. *Cancer Cell* **18**, 619-629 (2010).
46. Halliday, J. et al. In vivo radiation response of proneural glioma characterized by protective p53 transcriptional program and proneural-mesenchymal shift. *Proc Natl Acad Sci U S A* **111**, 5248-5253 (2014).
47. Minata, M. et al. Phenotypic Plasticity of Invasive Edge Glioma Stem-like Cells in Response to Ionizing Radiation. *Cell Rep* **26**, 1893-1905 e1897 (2019).
48. Bhat, K.P. et al. Mesenchymal differentiation mediated by NF-kappaB promotes radiation resistance in glioblastoma. *Cancer Cell* **24**, 331-346 (2013).
49. Rabe, M. et al. Identification of a transient state during the acquisition of temozolomide resistance in glioblastoma. *Cell Death Dis* **11**, 19 (2020).
50. Lee, H.H. et al. Histone 2A Family Member J Drives Mesenchymal Transition and Temozolomide Resistance in Glioblastoma Multiforme. *Cancers (Basel)* **12** (2019).
51. Sachdeva, R. et al. BMP signaling mediates glioma stem cell quiescence and confers treatment resistance in glioblastoma. *Sci Rep* **9**, 14569 (2019).
52. Zhao, S.G. et al. Xenograft-based platform-independent gene signatures to predict response to alkylating chemotherapy, radiation, and combination therapy for glioblastoma. *Neuro Oncol* (2019).
53. Segerman, A. et al. Clonal Variation in Drug and Radiation Response among Glioma-Initiating Cells Is Linked to Proneural-Mesenchymal Transition. *Cell Rep* **17**, 2994-3009 (2016).
54. Olav, E., Geir Olav, H., Henry, H. & Øystein, F. Growth of precultured human glioma biopsy specimens in nude rat brain. *Journal of Neurosurgery* **90**, 125-132 (1999).
55. Antunes, L. et al. Analysis of Tissue Chimerism in Nude Mouse Brain and Abdominal Xenograft Models of Human Glioblastoma Multiforme: What Does It Tell Us About the Models and About Glioblastoma Biology and Therapy? *Journal of Histochemistry & Cytochemistry* **48**, 847-858 (2000).
56. Chow, Lionel M.L. et al. Cooperativity within and among Pten, p53, and Rb Pathways Induces High-Grade Astrocytoma in Adult Brain. *Cancer Cell* **19**, 305-316 (2011).
57. Fomchenko, E.I. & Holland, E.C. Mouse Models of Brain Tumors and Their Applications in Preclinical Trials. *Clinical Cancer Research* **12**, 5288 (2006).
58. Ding, H. et al. Astrocyte-specific Expression of Activated p21-ras Results in Malignant Astrocytoma Formation in a Transgenic Mouse Model of Human Gliomas. *Cancer Research* **61**, 3826 (2001).

59. Xiao, A., Wu, H., Pandolfi, P.P., Louis, D.N. & Van Dyke, T. Astrocyte inactivation of the pRb pathway predisposes mice to malignant astrocytoma development that is accelerated by PTEN mutation. *Cancer Cell* **1**, 157-168 (2002).
60. Huse, J.T. & Holland, E.C. Genetically Engineered Mouse Models of Brain Cancer and the Promise of Preclinical Testing. *Brain Pathology* **19**, 132-143 (2009).
61. Holland, E.C. & Varmus, H.E. Basic fibroblast growth factor induces cell migration and proliferation after glia-specific gene transfer in mice. *Proceedings of the National Academy of Sciences* **95**, 1218-1223 (1998).
62. Pao, W., Klimstra, D.S., Fisher, G.H. & Varmus, H.E. Use of avian retroviral vectors to introduce transcriptional regulators into mammalian cells for analyses of tumor maintenance. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 8764-8769 (2003).
63. Oldrini, B. et al. Somatic genome editing with the RCAS/TVA-CRISPR/Cas9 system for precision tumor modeling. *bioRxiv*, 162669 (2017).
64. Holland, E.C., Hively, W.P., DePinho, R.A. & Varmus, H.E. A constitutively active epidermal growth factor receptor cooperates with disruption of G1 cell-cycle arrest pathways to induce glioma-like lesions in mice. *Genes Dev* **12**, 3675-3685 (1998).
65. Ozawa, T. et al. Most Human Non-GCIMP Glioblastoma Subtypes Evolve from a Common Proneural-like Precursor Glioma. *Cancer Cell* **26**, 288-300 (2014).
66. Bleau, A.M. et al. PTEN/PI3K/Akt Pathway Regulates the Side Population Phenotype and ABCG2 Activity in Glioma Tumor Stem-like Cells. *Cell Stem Cell* **4**, 226-235 (2009).
67. Hemmati, H.D. et al. Cancerous stem cells can arise from pediatric brain tumors. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 15178-15183 (2003).
68. Singh, S.K. et al. Identification of a cancer stem cell in human brain tumors. *Cancer Res* **63**, 5821-5828 (2003).
69. Bao, S. et al. Stem cell-like glioma cells promote tumor angiogenesis through vascular endothelial growth factor. *Cancer Res* **66**, 7843-7848 (2006).
70. Cheng, L. et al. Glioblastoma stem cells generate vascular pericytes to support vessel function and tumor growth. *Cell* **153**, 139-152 (2013).
71. Liu, Q. et al. Molecular properties of CD133+ glioblastoma stem cells derived from treatment-refractory recurrent brain tumors. *J Neurooncol* **94**, 1-19 (2009).
72. Koga, T. et al. Longitudinal assessment of tumor development using cancer avatars derived from genetically engineered pluripotent stem cells. *Nat Commun* **11**, 550-550 (2020).
73. Toledo, C.M. et al. Genome-wide CRISPR-Cas9 Screens Reveal Loss of Redundancy between PKMYT1 and WEE1 in Glioblastoma Stem-like Cells. *Cell Rep* **13**, 2425-2439 (2015).

74. Bian, S. et al. Genetically engineered cerebral organoids model brain tumor formation. *Nature Methods* **15**, 631-639 (2018).
75. Ogawa, J., Pao, G.M., Shokhirev, M.N. & Verma, I.M. Glioblastoma Model Using Human Cerebral Organoids. *Cell Reports* **23**, 1220-1229 (2018).
76. da Silva, B., Mathew, R.K., Polson, E.S., Williams, J. & Wurdak, H. Spontaneous Glioblastoma Spheroid Infiltration of Early-Stage Cerebral Organoids Models Brain Tumor Invasion. *SLAS DISCOVERY: Advancing the Science of Drug Discovery* **23**, 862-868 (2018).
77. Linkous, A. et al. Modeling Patient-Derived Glioblastoma with Cerebral Organoids. *Cell Reports* **26**, 3203-3211.e3205 (2019).
78. Hubert, C.G. et al. A Three-Dimensional Organoid Culture System Derived from Human Glioblastomas Recapitulates the Hypoxic Gradients and Cancer Stem Cell Heterogeneity of Tumors Found In Vivo. *Cancer Research* **76**, 2465 (2016).
79. Jacob, F. et al. A Patient-Derived Glioblastoma Organoid Model and Biobank Recapitulates Inter- and Intra-tumoral Heterogeneity. *Cell* **180**, 188-204.e122 (2020).
80. Hainsworth, J.D. et al. Concurrent radiotherapy and temozolomide followed by temozolomide and sorafenib in the first-line treatment of patients with glioblastoma multiforme. *Cancer* **116**, 3663-3669 (2010).
81. Galanis, E. et al. Phase II study of bevacizumab in combination with sorafenib in recurrent glioblastoma (N0776): a north central cancer treatment group trial. *Clinical cancer research : an official journal of the American Association for Cancer Research* **19**, 4816-4823 (2013).
82. Lee, E.Q. et al. Phase I/II study of sorafenib in combination with temsirolimus for recurrent glioblastoma or gliosarcoma: North American Brain Tumor Consortium study 05-02. *Neuro Oncol* **14**, 1511-1518 (2012).
83. Cloughesy, T. et al. Randomized, Double-Blind, Placebo-Controlled, Multicenter Phase II Study of Onartuzumab Plus Bevacizumab Versus Placebo Plus Bevacizumab in Patients With Recurrent Glioblastoma: Efficacy, Safety, and Hepatocyte Growth Factor and O6-Methylguanine-DNA Methyltransferase Biomarker Analyses. *J Clin Oncol* **35**, 343-351 (2017).
84. Reardon, D.A. et al. Phase 2 trial of erlotinib plus sirolimus in adults with recurrent glioblastoma. *J Neurooncol* **96**, 219-230 (2010).
85. Thiessen, B. et al. A phase I/II trial of GW572016 (lapatinib) in recurrent glioblastoma multiforme: clinical outcomes, pharmacokinetics and molecular correlation. *Cancer Chemother Pharmacol* **65**, 353-361 (2010).
86. Ma, D.J. et al. A phase II trial of everolimus, temozolomide, and radiotherapy in patients with newly diagnosed glioblastoma: NCCTG N057K. *Neuro Oncol* **17**, 1261-1269 (2015).

87. Lassen, U., Sorensen, M., Gaziel, T.B., Hasselbalch, B. & Poulsen, H.S. Phase II study of bevacizumab and temsirolimus combination therapy for recurrent glioblastoma multiforme. *Anticancer Res* **33**, 1657-1660 (2013).
88. Pitz, M.W. et al. Phase II study of PX-866 in recurrent glioblastoma. *Neuro Oncol* **17**, 1270-1274 (2015).
89. Luo, J., Solimini, N.L. & Elledge, S.J. Principles of Cancer Therapy: Oncogene and Non-oncogene Addiction. *Cell* **136**, 823-837 (2009).
90. Nijhawan, D. et al. Cancer Vulnerabilities Unveiled by Genomic Loss. *Cell* **150**, 842-854 (2012).
91. Hartwell, L.H., Szankasi, P., Roberts, C.J., Murray, A.W. & Friend, S.H. Integrating Genetic Approaches into the Discovery of Anticancer Drugs. *Science* **278**, 1064 (1997).
92. Chudnovsky, Y. et al. ZFHX4 interacts with the NuRD core member CHD4 and regulates the glioblastoma tumor-initiating cell state. *Cell Rep* **6**, 313-324 (2014).
93. Gargiulo, G. et al. In vivo RNAi screen for BMI1 targets identifies TGF-beta/BMP-ER stress pathways as key regulators of neural- and malignant glioma-stem cell homeostasis. *Cancer Cell* **23**, 660-676 (2013).
94. Goidts, V. et al. RNAi screening in glioma stem-like cells identifies PFKFB4 as a key molecule important for cancer cell survival. *Oncogene* **31**, 3235-3243 (2012).
95. Wurdak, H. et al. An RNAi screen identifies TRRAP as a regulator of brain tumor-initiating cell differentiation. *Cell Stem Cell* **6**, 37-47 (2010).
96. Ding, Y. et al. Cancer-Specific requirement for BUB1B/BUBR1 in human brain tumor isolates and genetically transformed cells. *Cancer discovery* **3**, 198-211 (2013).
97. Hubert, C.G. et al. Genome-wide RNAi screens in human brain tumor isolates reveal a novel viability requirement for PHF5A. *Genes Dev* **27**, 1032-1045 (2013).
98. Toledo, C.M. et al. BuGZ is required for Bub3 stability, Bub1 kinetochore function, and chromosome alignment. *Dev Cell* **28**, 282-294 (2014).
99. Hart, T. et al. High-Resolution CRISPR Screens Reveal Fitness Genes and Genotype-Specific Cancer Liabilities. *Cell* **163**, 1515-1526 (2015).
100. MacLeod, G. et al. Genome-Wide CRISPR-Cas9 Screens Expose Genetic Vulnerabilities and Mechanisms of Temozolomide Sensitivity in Glioblastoma Stem Cells. *Cell Rep* **27**, 971-986 e979 (2019).
101. Miller, T.E. et al. Transcription elongation factors represent in vivo cancer dependencies in glioblastoma. *Nature* **547**, 355-359 (2017).
102. Broad, D. DepMap 19Q4 Public. (2020).

103. Meyers, R.M. et al. Computational correction of copy number effect improves specificity of CRISPR-Cas9 essentiality screens in cancer cells. *Nature Genetics* **49**, 1779-1784 (2017).
104. Dempster, J.M. et al. Extracting Biological Insights from the Project Achilles Genome-Scale CRISPR Screens in Cancer Cell Lines. *bioRxiv* (2019).
105. Tsherniak, A. et al. Defining a Cancer Dependency Map. *Cell* **170**, 564-576.e516 (2017).
106. Hartlerode, A.J. & Scully, R. Mechanisms of double-strand break repair in somatic mammalian cells. *The Biochemical journal* **423**, 157-168 (2009).
107. Lieber, M.R., Ma, Y., Pannicke, U. & Schwarz, K. Mechanism and regulation of human non-homologous DNA end-joining. *Nat Rev Mol Cell Biol* **4**, 712-720 (2003).
108. Lee, J. et al. Tumor stem cells derived from glioblastomas cultured in bFGF and EGF more closely mirror the phenotype and genotype of primary tumors than do serum-cultured cell lines. *Cancer Cell* **9**, 391-403 (2006).
109. Toledo, C.M. et al. BuGZ is required for Bub3 stability, Bub1 kinetochore function, and chromosome alignment. *Developmental cell* **28**, 282-294 (2014).
110. Ding, Y. et al. ZNF131 suppresses centrosome fragmentation in glioblastoma stem-like cells through regulation of HAUS5. *Oncotarget* (2017).
111. Druker, B.J. et al. Effects of a selective inhibitor of the Abl tyrosine kinase on the growth of Bcr–Abl positive cells. *Nature Medicine* **2**, 561-566 (1996).
112. Zimmermann, J., Buchdunger, E., Mett, H., Meyer, T. & Lydon, N.B. Potent and selective inhibitors of the Abl-kinase: phenylamino-pyrimidine (PAP) derivatives. *Bioorganic & Medicinal Chemistry Letters* **7**, 187-192 (1997).
113. O'Brien, S.G. et al. Imatinib Compared with Interferon and Low-Dose Cytarabine for Newly Diagnosed Chronic-Phase Chronic Myeloid Leukemia. *New England Journal of Medicine* **348**, 994-1004 (2003).
114. Sunaga, N. et al. Knockdown of oncogenic KRAS in non-small cell lung cancers suppresses tumor growth and sensitizes tumor cells to targeted therapy. *Molecular cancer therapeutics* **10**, 336-346 (2011).
115. Bridges, C.B. The origin of variations in sexual and sex-limited characters. *Amer Nat* **1922**, 51–63 (1922).
116. Kaelin, W.G. The Concept of Synthetic Lethality in the Context of Anticancer Therapy. *Nature Reviews Cancer* **5**, 689-698 (2005).
117. Bryant, H.E. et al. Specific killing of BRCA2-deficient tumours with inhibitors of poly(ADP-ribose) polymerase. *Nature* **434**, 913-917 (2005).

118. Farmer, H. et al. Targeting the DNA repair defect in BRCA mutant cells as a therapeutic strategy. *Nature* **434**, 917-921 (2005).
119. Neshat, M.S. et al. Enhanced sensitivity of PTEN-deficient tumors to inhibition of FRAP/mTOR. *Proceedings of the National Academy of Sciences of the United States of America* **98**, 10314-10319 (2001).
120. Helming, K.C. et al. ARID1B is a specific vulnerability in ARID1A-mutant cancers. *Nature Medicine* **20**, 251-254 (2014).
121. Helming, Katherine C., Wang, X. & Roberts, Charles W.M. Vulnerabilities of Mutant SWI/SNF Complexes in Cancer. *Cancer Cell* **26**, 309-317 (2014).
122. Wilson, B.G. et al. Residual Complexes Containing SMARCA2 (BRM) Underlie the Oncogenic Drive of SMARCA4 (BRG1) Mutation. *Molecular and Cellular Biology* **34**, 1136 (2014).
123. Oike, T. et al. A Synthetic Lethality–Based Strategy to Treat Cancers Harboring a Genetic Deficiency in the Chromatin Remodeling Factor BRG1. *Cancer Research* **73**, 5508 (2013).
124. Hoffman, G.R. et al. Functional epigenetics approach identifies BRM/SMARCA2 as a critical synthetic lethal target in BRG1-deficient cancers. *Proceedings of the National Academy of Sciences* **111**, 3128 (2014).
125. Kryukov, G.V. et al. MTAP deletion confers enhanced dependency on the PRMT5 arginine methyltransferase in cancer cells. *Science* **351**, 1214-1218 (2016).
126. Tan, M.H. et al. Dynamic landscape and regulation of RNA editing in mammals. *Nature* **550**, 249-254 (2017).
127. Baskin, J.M. et al. The leukodystrophy protein FAM126A (hyccin) regulates PtdIns(4)P synthesis at the plasma membrane. *Nature cell biology* **18**, 132-138 (2016).
128. Bojjireddy, N., Guzman-Hernandez, M.L., Reinhard, N.R., Jovic, M. & Balla, T. EFR3s are palmitoylated plasma membrane proteins that control responsiveness to G-protein-coupled receptors. *Journal of Cell Science* **128**, 118 (2015).
129. Sun, L. et al. JFK, a Kelch domain-containing F-box protein, links the SCF complex to p53 regulation. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 10195-10200 (2009).
130. Wiedenheft, B., Sternberg, S.H. & Doudna, J.A. RNA-guided genetic silencing systems in bacteria and archaea. *Nature* **482**, 331-338 (2012).
131. Mali, P., Esvelt, K.M. & Church, G.M. Cas9 as a versatile tool for engineering biology. *Nat Methods* **10**, 957-963 (2013).
132. Cho, S.W., Kim, S., Kim, J.M. & Kim, J.S. Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease. *Nat Biotechnol* **31**, 230-232 (2013).

133. Cong, L. et al. Multiplex Genome Engineering Using CRISPR/Cas Systems. *Science* **339**, 819-823 (2013).
134. Jinek, M. et al. RNA-programmed genome editing in human cells. *Elife* **2**, e00471 (2013).
135. Mali, P. et al. RNA-guided human genome engineering via Cas9. *Science* **339**, 823-826 (2013).
136. Jinek, M. et al. A Programmable Dual-RNA–Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science* **337**, 816-821 (2012).
137. Ran, F.A. et al. Genome engineering using the CRISPR-Cas9 system. *Nature Protocols* **8**, 2281-2308 (2013).
138. Chakrabarti, A.M. et al. Target-Specific Precision of CRISPR-Mediated Genome Editing. *Mol Cell* **73**, 699-713 e696 (2019).
139. Shalem, O. et al. Genome-scale CRISPR-Cas9 knockout screening in human cells. *Science* **343**, 84-87 (2014).
140. Wang, T., Wei, J.J., Sabatini, D.M. & Lander, E.S. Genetic screens in human cells using the CRISPR-Cas9 system. *Science* **343**, 80-84 (2014).
141. Kim, S., Kim, D., Cho, S.W., Kim, J. & Kim, J.S. Highly efficient RNA-guided genome editing in human cells via delivery of purified Cas9 ribonucleoproteins. *Genome Research* **24**, 1012-1019 (2014).
142. Zuris, J.A. et al. Cationic lipid-mediated delivery of proteins enables efficient protein-based genome editing in vitro and in vivo. *Nature Biotechnology* **33**, 73 (2014).
143. DeWitt, M.A., Corn, J.E. & Carroll, D. Genome editing via delivery of Cas9 ribonucleoprotein. *Methods* **121-122**, 9-15 (2017).
144. Schumann, K. et al. Generation of knock-in primary human T cells using Cas9 ribonucleoproteins. *Proceedings of the National Academy of Sciences* **112**, 10437 (2015).
145. Lin, S., Staahl, B.T., Alla, R.K. & Doudna, J.A. Enhanced homology-directed human genome engineering by controlled timing of CRISPR/Cas9 delivery. *eLife* **3**, e04766 (2014).
146. Bressan, R.B. et al. Efficient CRISPR/Cas9-assisted gene targeting enables rapid and precise genetic manipulation of mammalian neural stem cells. *Development* **144**, 635-648 (2017).
147. Graf, R., Li, X., Chu, V.T. & Rajewsky, K. sgRNA Sequence Motifs Blocking Efficient CRISPR/Cas9-Mediated Gene Editing. *Cell Reports* **26**, 1098-1103.e1093 (2019).
148. Eckstein, F. Phosphorothioates, Essential Components of Therapeutic Oligonucleotides. *Nucleic Acid Therapeutics* **24**, 374-387 (2014).
149. Allerson, C.R. et al. Fully 2‘-Modified Oligonucleotide Duplexes with Improved in Vitro Potency and Stability Compared to Unmodified Small Interfering RNA. *Journal of Medicinal Chemistry* **48**, 901-904 (2005).

150. Brinkman, E.K., Chen, T., Amendola, M. & van Steensel, B. Easy quantitative assessment of genome editing by sequence trace decomposition. *Nucleic Acids Res* **42**, e168 (2014).
151. Hsiau, T. et al. Inference of CRISPR Edits from Sanger Trace Data. *bioRxiv* (2019).
152. Doench, J.G. et al. Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. *Nat Biotechnol* **34**, 184-191 (2016).
153. Sherr, C.J. & McCormick, F. The RB and p53 pathways in cancer. *Cancer Cell* **2**, 103-112 (2002).
154. Keniry, M. & Parsons, R. The role of PTEN signaling perturbations in cancer and in targeted therapy. *Oncogene* **27**, 5477-5485 (2008).
155. Brennan, C.W. et al. The somatic genomic landscape of glioblastoma. *Cell* **155**, 462-477 (2013).
156. Robinson, M.D., McCarthy, D.J. & Smyth, G.K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics (Oxford, England)* **26**, 139-140 (2010).
157. Fischer, M. Census and evaluation of p53 target genes. *Oncogene* **36**, 3943-3956 (2017).
158. van den Heuvel, A.P. et al. Binding of protein kinase B to the plakin family member periplakin. *J Cell Sci* **115**, 3957-3966 (2002).
159. Song, W., Volosin, M., Cragnolini, A.B., Hempstead, B.L. & Friedman, W.J. ProNGF induces PTEN via p75NTR to suppress Trk-mediated survival signaling in brain neurons. *J Neurosci* **30**, 15608-15615 (2010).
160. Plotkin, J.L. et al. Impaired TrkB receptor signaling underlies corticostriatal dysfunction in Huntington's disease. *Neuron* **83**, 178-188 (2014).
161. Banerjee, R., Henson, B.S., Russo, N., Tsodikov, A. & D'Silva, N.J. Rap1 mediates galanin receptor 2-induced proliferation and survival in squamous cell carcinoma. *Cell Signal* **23**, 1110-1118 (2011).
162. Dwyer, N.D. et al. Neural Stem Cells to Cerebral Cortex: Emerging Mechanisms Regulating Progenitor Behavior and Productivity. *J Neurosci* **36**, 11394-11401 (2016).
163. Shu, H.B., Takeuchi, M. & Goeddel, D.V. The tumor necrosis factor receptor 2 signal transducers TRAF2 and c-IAP1 are components of the tumor necrosis factor receptor 1 signaling complex. *Proceedings of the National Academy of Sciences of the United States of America* **93**, 13973-13978 (1996).
164. Choi, N.M., Majumder, P. & Boss, J.M. Regulation of major histocompatibility complex class II genes. *Curr Opin Immunol* **23**, 81-87 (2011).
165. Neefjes, J., Jongsma, M.L., Paul, P. & Bakke, O. Towards a systems understanding of MHC class I and MHC class II antigen presentation. *Nat Rev Immunol* **11**, 823-836 (2011).

166. Reuss, D.E. et al. Functional MHC class II is upregulated in neurofibromin-deficient Schwann cells. *J Invest Dermatol* **133**, 1372-1375 (2013).
167. Tran, C.T. et al. Differential expression of MHC class II molecules by microglia and neoplastic astroglia: relevance for the escape of astrocytoma cells from immune surveillance. *Neuropathol Appl Neurobiol* **24**, 293-301 (1998).
168. Thibodeau, J., Bourgeois-Daigneault, M.C. & Lapointe, R. Targeting the MHC Class II antigen presentation pathway in cancer immunotherapy. *Oncoimmunology* **1**, 908-916 (2012).
169. Hsiao, L.L. et al. A compendium of gene expression in normal human tissues. *Physiol Genomics* **7**, 97-104 (2001).
170. Vakulskas, C.A. & Behlke, M.A. Evaluation and Reduction of CRISPR Off-Target Cleavage Events. *Nucleic Acid Therapeutics* **29**, 167-174 (2019).
171. Symington, L.S. & Gautier, J. Double-strand break end resection and repair pathway choice. *Annu Rev Genet* **45**, 247-271 (2011).
172. Guo, T. et al. Harnessing accurate non-homologous end joining for efficient precise deletion in CRISPR/Cas9-mediated genome editing. *Genome Biol* **19**, 170 (2018).
173. Betermier, M., Bertrand, P. & Lopez, B.S. Is non-homologous end-joining really an inherently error-prone process? *PLoS Genet* **10**, e1004086 (2014).
174. Canver, M.C. et al. Characterization of genomic deletion efficiency mediated by clustered regularly interspaced short palindromic repeats (CRISPR)/Cas9 nuclease system in mammalian cells. *J Biol Chem* **289**, 21312-21324 (2014).
175. Wolfs, J.M. et al. Biasing genome-editing events toward precise length deletions with an RNA-guided TevCas9 dual nuclease. *Proceedings of the National Academy of Sciences of the United States of America* **113**, 14988-14993 (2016).
176. Helm, M., Brulé, H., Giegé, R. & Florentz, C. More mistakes by T7 RNA polymerase at the 5' ends of in vitro-transcribed RNAs. *RNA* **5**, 618–621 (1999).
177. Sun, Y. et al. Long-term tripotent differentiation capacity of human neural stem (NS) cells in adherent culture. *Molecular and cellular neurosciences* **38**, 245-258 (2008).
178. Hart, T. et al. High-Resolution CRISPR Screens Reveal Fitness Genes and Genotype-Specific Cancer Liabilities Screens Reveal Fitness Genes. *Cell*, 1-12 (2015).
179. Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics (Oxford, England)* **29**, 15-21 (2013).
180. Anders, S., Pyl, P.T. & Huber, W. HTSeq--a Python framework to work with high-throughput sequencing data. *Bioinformatics (Oxford, England)* **31**, 166-169 (2015).

181. Wickham, H., Sievert, C. & Springer International Publishing, A.G. ggplot2 : elegant graphics for data analysis. (2016).
182. Kolde, R. pheatmap: Pretty Heatmaps. R package version 1.0.10. *https://CRAN.R-project.org/package=pheatmap* (2018).
183. Ostrom, Q.T. et al. CBTRUS Statistical Report: Primary Brain and Central Nervous System Tumors Diagnosed in the United States in 2008-2012. *Neuro Oncol* **17 Suppl 4**, iv1-iv62 (2015).
184. Guha, A., Feldkamp, M.M., Lau, N., Boss, G. & Pawson, A. Proliferation of human malignant astrocytomas is dependent on Ras activation. *Oncogene* **15**, 2755-2765 (1997).
185. Bredel, M. et al. A network model of a cooperative genetic landscape in brain tumors. *JAMA* **302**, 261-275 (2009).
186. Holland, E.C. et al. Combined activation of Ras and Akt in neural progenitors induces glioblastoma formation in mice. *Nat Genet* **25**, 55-57 (2000).
187. Chautard, E., Ouedraogo, Z.G., Biau, J. & Verrelle, P. Role of Akt in human malignant glioma: from oncogenesis to tumor aggressiveness. *J Neurooncol* **117**, 205-215 (2014).
188. Reardon, D.A., Wen, P.Y. & Mellinghoff, I.K. Targeted molecular therapies against epidermal growth factor receptor: past experiences and challenges. *Neuro Oncol* **16 Suppl 8**, viii7-13 (2014).
189. Sanjana, N.E., Shalem, O. & Zhang, F. Improved vectors and genome-wide libraries for CRISPR screening. *Nature Methods* **11**, 783-784 (2014).
190. Domcke, S., Sinha, R., Levine, D.A., Sander, C. & Schultz, N. Evaluating cell lines as tumour models by comparison of genomic profiles. *Nat Commun* **4**, 2126 (2013).
191. Gillet, J.-P. et al. Redefining the relevance of established cancer cell lines to the study of mechanisms of clinical anti-cancer drug resistance. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 18708-18713 (2011).
192. Warde-Farley, D. et al. The GeneMANIA prediction server: biological network integration for gene prioritization and predicting gene function. *Nucleic Acids Research* **38**, W214-W220 (2010).
193. Mushtaq, M., Ali, R.H., Kashuba, V., Klein, G. & Kashuba, E. S18 family of mitochondrial ribosomal proteins: evolutionary history and Gly132 polymorphism in colon carcinoma. *Oncotarget; Vol 7, No 34* (2016).
194. Kim, H.-J. & Barrientos, A. MTG1 couples mitoribosome large subunit assembly with intersubunit bridge formation. *Nucleic Acids Research* **46**, 8435-8453 (2018).
195. Kotani, T., Akabane, S., Takeyasu, K., Ueda, T. & Takeuchi, N. Human G-proteins, ObgH1 and Mtg1, associate with the large mitochondrial ribosome subunit and are involved in translation and assembly of respiratory complexes. *Nucleic Acids Research* **41**, 3713-3722 (2013).

196. Siira, S.J. et al. LRPPRC-mediated folding of the mitochondrial transcriptome. *Nat Commun* **8**, 1532 (2017).
197. Coenen, M.J.H., Smeitink, J.A.M., Smeets, R., Trijbels, F.J.M. & van den Heuvel, L.P. Mutation detection in four candidate genes (OXA1L, MRS2L, YME1L and MIPEP) for combined deficiencies in the oxidative phosphorylation system. *Journal of Inherited Metabolic Disease* **28**, 1091-1097 (2005).
198. Chew, A. et al. Cloning, Expression, and Chromosomal Assignment of the Human Mitochondrial Intermediate Peptidase Gene (MIPEP). *Genomics* **40**, 493-496 (1997).
199. Sheffer, R. et al. Postnatal microcephaly and pain insensitivity due to a de novo heterozygous DNM1L mutation causing impaired mitochondrial fission and function. *American Journal of Medical Genetics Part A* **170**, 1603-1607 (2016).
200. Zaha, K. et al. DNM1L-related encephalopathy in infancy with Leigh syndrome-like phenotype and suppression-burst. *Clinical Genetics* **90**, 472-474 (2016).
201. Nasca, A. et al. Biallelic Mutations in DNM1L are Associated with a Slowly Progressive Infantile Encephalopathy. *Human Mutation* **37**, 898-903 (2016).
202. Norton, M. et al. ROMO1 Is an Essential Redox-Dependent Regulator of Mitochondrial Dynamics. *Science Signaling* **7**, ra10 (2014).
203. Yu, M.O. et al. Romo1 is associated with ROS production and cellular growth in human gliomas. *Journal of Neuro-Oncology* **121**, 73-81 (2015).
204. Richter, F. et al. ROMO1 is a constituent of the human presequence translocase required for YME1L protease import. *Journal of Cell Biology* **218**, 598-614 (2018).
205. Rainbolt, T.K., Lebeau, J., Puchades, C. & Wiseman, R.L. Reciprocal Degradation of YME1L and OMA1 Adapts Mitochondrial Proteolytic Activity during Stress. *Cell Reports* **14**, 2041-2049 (2016).
206. Shi, H., Rampello, A.J. & Glynn, S.E. Engineered AAA+ proteases reveal principles of proteolysis at the mitochondrial inner membrane. *Nat Commun* **7**, 13301 (2016).
207. Paschen, S.A. et al. The role of the TIM8–13 complex in the import of Tim23 into mitochondria. *The EMBO Journal* **19**, 6392-6400 (2000).
208. Curran, S.P., Leuenberger, D., Schmidt, E. & Koehler, C.M. The role of the Tim8p–Tim13p complex in a conserved import pathway for mitochondrial polytopic inner membrane proteins. *Journal of Cell Biology* **158**, 1017-1027 (2002).
209. Deville, C., Franke, K., Mogk, A., Bukau, B. & Saibil, H.R. Two-Step Activation Mechanism of the ClpB Disaggregase for Sequential Substrate Threading by the Main ATPase Motor. *Cell Reports* **27**, 3433-3446.e3434 (2019).

210. Rizo, A.N. et al. Structural basis for substrate gripping and translocation by the ClpB AAA+ disaggregase. *Nat Commun* **10**, 2393 (2019).
211. Ning, B. & Elbein, A.D. Cloning, expression and characterization of the pig liver GDP-mannose pyrophosphorylase. *European Journal of Biochemistry* **267**, 6866-6874 (2000).
212. Possik, E., Madiraju, S.R.M. & Prentki, M. Glycerol-3-phosphate phosphatase/PGP: Role in intermediary metabolism and target for cardiometabolic diseases. *Biochimie* **143**, 18-28 (2017).
213. Nordgren, K.K.S. et al. Methionine Adenosyltransferase 2A/2B and Methylation: Gene Sequence Variation and Functional Genomics. *Drug Metabolism and Disposition* **39**, 2135 (2011).
214. Voynova, N.E. et al. Human mevalonate diphosphate decarboxylase: Characterization, investigation of the mevalonate diphosphate binding site, and crystal structure. *Archives of Biochemistry and Biophysics* **480**, 58-67 (2008).
215. Stincone, A. et al. The return of metabolism: biochemistry and physiology of the pentose phosphate pathway. *Biological Reviews* **90**, 927-963 (2015).
216. McReynolds, M.R., Wang, W., Holleran, L.M. & Hanna-Rose, W. Uridine monophosphate synthetase enables eukaryotic de novo NAD+ biosynthesis from quinolinic acid. *Journal of Biological Chemistry* **292**, 11147-11153 (2017).
217. Welin, M. et al. Substrate Specificity and Oligomerization of Human GMP Synthetase. *Journal of Molecular Biology* **425**, 4323-4333 (2013).
218. Shah, C.P. & Kharkar, P.S. Discovery of novel human inosine 5′-monophosphate dehydrogenase 2 (hIMPDH2) inhibitors as potential anticancer agents. *European Journal of Medicinal Chemistry* **158**, 286-301 (2018).
219. Manoharan, P. et al. Chronic and selective inhibition of basolateral membrane Na-K-ATPase uniquely regulates brush border membrane Na absorption in intestinal epithelial cells. *American Journal of Physiology-Cell Physiology* **308**, C650-C656 (2015).
220. Palmgren, M. & Morsomme, P. The plasma membrane H+-ATPase, a simple polypeptide with a long history. *Yeast* **36**, 201-210 (2019).
221. Lizcano, J.M. et al. LKB1 is a master kinase that activates 13 kinases of the AMPK subfamily, including MARK/PAR-1. *The EMBO Journal* **23**, 833-843 (2004).
222. Boudeau, J. et al. MO25α/β interact with STRADα/β enhancing their ability to bind, activate and localize LKB1 in the cytoplasm. *The EMBO Journal* **22**, 5102-5114 (2003).
223. Gao, X. et al. γ-6-Phosphogluconolactone, a Byproduct of the Oxidative Pentose Phosphate Pathway, Contributes to AMPK Activation through Inhibition of PP2A. *Molecular Cell* **76**, 857-871.e859 (2019).

224. Fay, J.R., Steele, V. & Crowell, J.A. Energy Homeostasis and Cancer Prevention: The AMP-Activated Protein Kinase. *Cancer Prevention Research* **2**, 301 (2009).
225. Kishton, Rigel J. et al. AMPK Is Essential to Balance Glycolysis and Mitochondrial Metabolism to Control T-ALL Cell Stress and Survival. *Cell Metabolism* **23**, 649-662 (2016).
226. Chung, H. et al. Human ADAR1 Prevents Endogenous RNA from Triggering Translational Shutdown. *Cell* **172**, 811-824.e814 (2018).
227. Liu, H. et al. Tumor-derived IFN triggers chronic pathway agonism and sensitivity to ADAR loss. *Nature Medicine* **25**, 95-102 (2019).
228. Ishizuka, J.J. et al. Loss of ADAR1 in tumours overcomes resistance to immune checkpoint blockade. *Nature* **565**, 43-48 (2019).
229. Haller, O., Staeheli, P., Schwemmle, M. & Kochs, G. Mx GTPases: dynamin-like antiviral machines of innate immunity. *Trends in Microbiology* **23**, 154-163 (2015).
230. Munoz, D.M. et al. CRISPR Screens Provide a Comprehensive Assessment of Cancer Vulnerabilities but Generate False-Positive Hits for Highly Amplified Genomic Regions. *Cancer discovery* **6**, 900-913 (2016).
231. Aguirre, A.J. et al. Genomic Copy Number Dictates a Gene-Independent Cell Response to CRISPR/Cas9 Targeting. *Cancer discovery* **6**, 914-929 (2016).
232. Sheel, A. & Xue, W. Genomic Amplifications Cause False Positives in CRISPR Screens. *Cancer discovery* **6**, 824-826 (2016).
233. Nakatsu, F. et al. PtdIns4P synthesis by PI4KIIIα at the plasma membrane and its impact on plasma membrane identity. *Journal of Cell Biology* **199**, 1003-1016 (2012).
234. Sun, L. et al. Substrate phosphorylation and feedback regulation in JFK-promoted p53 destabilization. *Journal of Biological Chemistry* **286**, 4226-4235 (2011).
235. Nagler, A. et al. A genome-wide CRISPR screen identifies FBXO42 involvement in resistance toward MEK inhibition in NRAS-mutant melanoma. *Pigment Cell & Melanoma Research* **n/a** (2019).
236. Yan, R. et al. SCFJFK is a bona fide E3 ligase for ING4 and a potent promoter of the angiogenesis and metastasis of breast cancer. *Genes Dev* **29**, 672-685 (2015).
237. Jin, J. et al. Systematic analysis and nomenclature of mammalian F-box proteins. *Genes Dev* **18**, 2573-2580 (2004).
238. Li, X. et al. Proteomic analyses reveal distinct chromatin-associated and soluble transcription factor complexes. *Mol Syst Biol* **11**, 775-775 (2015).

239. Pierotti, M.A. et al. Characterization of an inversion on the long arm of chromosome 10 juxtaposing D10S170 and RET and creating the oncogenic sequence RET/PTC. *Proceedings of the National Academy of Sciences* **89**, 1616-1620 (1992).
240. Merolla, F. et al. Involvement of H4(D10S170) protein in ATM-dependent response to DNA damage. *Oncogene* **26**, 6167-6175 (2007).
241. JunGang, Z., Jun, T., WanFu, M. & KaiMing, R. FBXW7-mediated degradation of CCDC6 is impaired by ATM during DNA damage response in lung cancer cells. *FEBS Letters* **586**, 4257-4263 (2012).
242. Merolla, F. et al. Loss of CCDC6, the First Identified RET Partner Gene, Affects pH2AX S139 Levels and Accelerates Mitotic Entry upon DNA Damage. *PLOS ONE* **7**, e36177 (2012).
243. Schweppe, D.K., Huttlin, E.L., Harper, J.W. & Gygi, S.P. BioPlex Display: An Interactive Suite for Large-Scale AP–MS Protein–Protein Interaction Data. *Journal of Proteome Research* **17**, 722-726 (2018).
244. Huttlin, E.L. et al. Architecture of the human interactome defines protein communities and disease networks. *Nature* **545**, 505-509 (2017).
245. Santaguida, S. & Amon, A. Short- and long-term effects of chromosome mis-segregation and aneuploidy. *Nature Reviews Molecular Cell Biology* **16**, 473-485 (2015).
246. Yamagishi, Y., Yang, C.-H., Tanno, Y. & Watanabe, Y. MPS1/Mph1 phosphorylates the kinetochore protein KNL1/Spc7 to recruit SAC components. *Nature Cell Biology* **14**, 746-752 (2012).
247. London, N., Ceto, S., Ranish, Jeffrey A. & Biggins, S. Phosphoregulation of Spc105 by Mps1 and PP1 Regulates Bub1 Localization to Kinetochores. *Current Biology* **22**, 900-906 (2012).
248. Shepperd, Lindsey A. et al. Phosphodependent Recruitment of Bub1 and Bub3 to Spc7/KNL1 by Mph1 Kinase Maintains the Spindle Checkpoint. *Current Biology* **22**, 891-899 (2012).
249. Zhang, G., Lischetti, T. & Nilsson, J. A minimal number of MELT repeats supports all the functions of KNL1 in chromosome segregation. *Journal of Cell Science* **127**, 871 (2014).
250. Lischetti, T., Zhang, G., Sedgwick, G.G., Bolanos-Garcia, V.M. & Nilsson, J. The internal Cdc20 binding site in BubR1 facilitates both spindle assembly checkpoint signalling and silencing. *Nat Commun* **5**, 5563 (2014).
251. Primorac, I. et al. Bub3 reads phosphorylated MELT repeats to promote spindle assembly checkpoint signaling. *eLife* **2**, e01030 (2013).
252. Moyle, M.W. et al. A Bub1–Mad1 interaction targets the Mad1–Mad2 complex to unattached kinetochores to initiate the spindle checkpoint. *Journal of Cell Biology* **204**, 647-657 (2014).

253. Vleugel, M. et al. Sequential Multisite Phospho-Regulation of KNL1-BUB3 Interfaces at Mitotic Kinetochores. *Molecular Cell* **57**, 824-835 (2015).
254. Vleugel, M. et al. Arrayed BUB recruitment modules in the kinetochore scaffold KNL1 promote accurate chromosome segregation. *Journal of Cell Biology* **203**, 943-955 (2013).
255. Krenn, V., Overlack, K., Primorac, I., van Gerwen, S. & Musacchio, A. KI Motifs of Human Knl1 Enhance Assembly of Comprehensive Spindle Checkpoint Complexes around MELT Repeats. *Current Biology* **24**, 29-39 (2014).
256. Ward, Patrick S. & Thompson, Craig B. Metabolic Reprogramming: A Cancer Hallmark Even Warburg Did Not Anticipate. *Cancer Cell* **21**, 297-308 (2012).
257. Cheng, G. et al. Mitochondria-Targeted Drugs Synergize with 2-Deoxyglucose to Trigger Breast Cancer Cell Death. *Cancer Research* **72**, 2634 (2012).
258. Nazarewicz, R.R. et al. Does Scavenging of Mitochondrial Superoxide Attenuate Cancer Prosurvival Signaling Pathways? *Antioxidants & Redox Signaling* **19**, 344-349 (2013).
259. Buzzai, M. et al. Systemic Treatment with the Antidiabetic Drug Metformin Selectively Impairs p53-Deficient Tumor Cell Growth. *Cancer Research* **67**, 6745 (2007).
260. Zhang, X. et al. Induction of mitochondrial dysfunction as a strategy for targeting tumour cells in metabolically compromised microenvironments. *Nat Commun* **5**, 3295 (2014).
261. Hensley, C.T., Wasti, A.T. & DeBerardinis, R.J. Glutamine and cancer: cell biology, physiology, and clinical opportunities. *The Journal of Clinical Investigation* **123**, 3678-3684 (2013).
262. Guo, J.Y. et al. Activated Ras requires autophagy to maintain oxidative metabolism and tumorigenesis. *Genes Dev* **25**, 460-470 (2011).
263. Li, H. et al. 6-Phosphogluconate Dehydrogenase Links Cytosolic Carbohydrate Metabolism to Protein Secretion via Modulation of Glutathione Levels. *Cell Chemical Biology* **26**, 1306-1314.e1305 (2019).
264. Sarfraz, I. et al. 6-Phosphogluconate dehydrogenase fuels multiple aspects of cancer cells: From cancer initiation to metastasis and chemoresistance. *BioFactors* **n/a** (2020).
265. Li, S. et al. Targeting β2 subunit of Na(+)/K(+)-ATPase induces glioblastoma cell apoptosis through elevation of intracellular Ca(2). *Am J Cancer Res* **9**, 1293-1308 (2019).
266. Wu, I.-C. et al. Overexpression of ATPase Na+/K+ transporting alpha 1 polypeptide, ATP1A1, correlates with clinical diagnosis and progression of esophageal squamous cell carcinoma. *Oncotarget* **7** (2016).
267. Kofuji, S. et al. IMP dehydrogenase-2 drives aberrant nucleolar activity and promotes tumorigenesis in glioblastoma. *Nature Cell Biology* **21**, 1003-1014 (2019).

268. Wieczorek, P., Bałut-Wieczorek, M., Jasinski, M., Szabłoński, W. & Antczak, A. Inosine monophosphate dehydrogenase 2 as a marker of aggressive and advanced prostate cancer. *Cent European J Urol* **71**, 399-403 (2018).
269. Wang, Z. et al. Methionine is a metabolic dependency of tumor-initiating cells. *Nature Medicine* **25**, 825-837 (2019).
270. Strekalova, E. et al. S-adenosylmethionine biosynthesis is a targetable metabolic vulnerability of cancer stem cells. *Breast Cancer Res Treat* **175**, 39-50 (2019).
271. Quinlan, C.L. et al. Targeting S-adenosylmethionine biosynthesis with a novel allosteric inhibitor of Mat2A. *Nat Chem Biol* **13**, 785-792 (2017).
272. Manguso, R.T. et al. In vivo CRISPR screening identifies Ptpn2 as a cancer immunotherapy target. *Nature* **547**, 413-418 (2017).
273. Kwon, J. et al. Targeting Phosphatidylinositol 4-Kinase IIIα for Radiosensitization: A Potential Model of Drug Repositioning Using an Anti-Hepatitis C Viral Agent. *International Journal of Radiation Oncology*Biology*Physics* **96**, 867-876 (2016).
274. Kattan, W.E. et al. Targeting plasma membrane phosphatidylserine content to inhibit oncogenic KRAS function. *Life Science Alliance* **2**, e201900431 (2019).
275. Sbrissa, D. et al. A novel cross-talk between CXCR4 and PI4KIIIα in prostate cancer cells. *Oncogene* **38**, 332-344 (2019).
276. Musacchio, A. The Molecular Biology of Spindle Assembly Checkpoint Signaling Dynamics. *Current Biology* **25**, R1002-R1018 (2015).
277. Meraldi, P., Draviam, V.M. & Sorger, P.K. Timing and Checkpoints in the Regulation of Mitotic Progression. *Developmental Cell* **7**, 45-60 (2004).
278. Matson, D.R. & Stukenberg, P.T. Spindle poisons and cell fate: a tale of two pathways. *Mol Interv* **11**, 141-150 (2011).
279. Bing, C. & Karen, C. Consequences of mitotic slippage for antimicrotubule drug therapy. *Endocrine-Related Cancer* **24**, T97-T106 (2017).
280. Thanasopoulou, A., Stravopodis, D.J., Dimas, K.S., Schwaller, J. & Anastasiadou, E. Loss of CCDC6 affects cell cycle through impaired intra-S-phase checkpoint control. *PLoS ONE* **7** (2012).
281. Kim, E.M. & Burke, D.J. DNA Damage Activates the SAC in an ATM/ATR-Dependent Manner, Independently of the Kinetochore. *PLOS Genetics* **4**, e1000015 (2008).
282. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S.L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biology* **10**, R25 (2009).

283. Brabetz, S. et al. A biobank of patient-derived pediatric brain tumor models. *Nature Medicine* **24**, 1752-1761 (2018).
284. Fu, Y. et al. High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nature Biotechnology* **31**, 822-826 (2013).
285. Cong, L. et al. Multiplex Genome Engineering Using CRISPR/Cas Systems. *Science* **339**, 819 (2013).
286. Mali, P. et al. CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. *Nature Biotechnology* **31**, 833-838 (2013).
287. Hsu, P.D. et al. DNA targeting specificity of RNA-guided Cas9 nucleases. *Nature Biotechnology* **31**, 827-832 (2013).
288. Pattanayak, V. et al. High-throughput profiling of off-target DNA cleavage reveals RNA-programmed Cas9 nuclease specificity. *Nature Biotechnology* **31**, 839-843 (2013).
289. Song, B. et al. Improved Hematopoietic Differentiation Efficiency of Gene-Corrected Beta-Thalassemia Induced Pluripotent Stem Cells by CRISPR/Cas9 System. *Stem Cells and Development* **24**, 1053-1065 (2014).
290. Wang, X. et al. Unbiased detection of off-target cleavage by CRISPR-Cas9 and TALENs using integrase-defective lentiviral vectors. *Nature Biotechnology* **33**, 175-178 (2015).
291. Kim, D. et al. Digenome-seq: genome-wide profiling of CRISPR-Cas9 off-target effects in human cells. *Nature Methods* **12**, 237-243 (2015).
292. Tsai, S.Q. et al. GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases. *Nature Biotechnology* **33**, 187-197 (2015).
293. Pattanayak, V. et al. High-throughput profiling of off-target DNA cleavage reveals RNA-programmed Cas9 nuclease specificity. *Nat Biotechnol* **31**, 839-843 (2013).
294. Cradick, T.J., Qiu, P., Lee, C.M., Fine, E.J. & Bao, G. COSMID: A Web-based Tool for Identifying and Validating CRISPR/Cas Off-target Sites. *Mol Ther Nucleic Acids* **3**, e214 (2014).
295. Heigwer, F., Kerr, G. & Boutros, M. E-CRISP: fast CRISPR target site identification. *Nat Methods* **11**, 122-123 (2014).
296. Labun, K., Montague, T.G., Gagnon, J.A., Thyme, S.B. & Valen, E. CHOPCHOP v2: a web tool for the next generation of CRISPR genome engineering. *Nucleic Acids Res* **44**, W272-276 (2016).
297. Listgarten, J. et al. Prediction of off-target activities for the end-to-end design of CRISPR guide RNAs. *Nature Biomedical Engineering* **2**, 38-47 (2018).
298. Lin, Y. et al. CRISPR/Cas9 systems have off-target activity with insertions or deletions between target DNA and guide RNA sequences. *Nucleic Acids Res* **42**, 7473-7485 (2014).

299. Anderson, E.M. et al. Systematic analysis of CRISPR–Cas9 mismatch tolerance reveals low levels of off-target activity. *Journal of Biotechnology* **211**, 56-65 (2015).
300. Fu, B.X., St Onge, R.P., Fire, A.Z. & Smith, J.D. Distinct patterns of Cas9 mismatch tolerance in vitro and in vivo. *Nucleic Acids Res* **44**, 5365-5377 (2016).
301. Zheng, T. et al. Profiling single-guide RNA specificity reveals a mismatch sensitive core sequence. *Sci Rep* **7**, 40638 (2017).
302. Klein, M., Eslami-Mossallam, B., Arroyo, D.G. & Depken, M. Hybridization Kinetics Explains CRISPR-Cas Off-Targeting Rules. *Cell Rep* **22**, 1413-1423 (2018).
303. Kleinstiver, B.P. et al. High-fidelity CRISPR-Cas9 nucleases with no detectable genome-wide off-target effects. *Nature* **529**, 490-495 (2016).
304. Chen, J.S. et al. Enhanced proofreading governs CRISPR-Cas9 targeting accuracy. *Nature* **550**, 407-410 (2017).
305. Fujii, W., Onuma, A., Sugiura, K. & Naito, K. Efficient generation of genome-modified mice via offset-nicking by CRISPR/Cas system. *Biochemical and Biophysical Research Communications* **445**, 791-794 (2014).
306. Ran, F.A. et al. Double Nicking by RNA-Guided CRISPR Cas9 for Enhanced Genome Editing Specificity. *Cell* **154**, 1380-1389 (2013).
307. Cho, S.W. et al. Analysis of off-target effects of CRISPR/Cas-derived RNA-guided endonucleases and nickases. *Genome research* **24**, 132-141 (2014).
308. Chakrabarti, A.M. et al. Target-specific precision of CRISPR-mediated genome editing. *Molecular Cell*, 387027 (2018).
309. Gao, Z., Herrera-Carrillo, E. & Berkhout, B. Delineation of the Exact Transcription Termination Signal for Type 3 Polymerase III. *Molecular therapy. Nucleic acids* **10**, 36-44 (2018).
310. Green, L.W., Ottoson, J.M., García, C. & Hiatt, R.A. Diffusion Theory and Knowledge Dissemination, Utilization, and Integration in Public Health. *Annual Review of Public Health* **30**, 151-174 (2009).
311. Mohs, R.C. & Greig, N.H. Drug discovery and development: Role of basic biological research. *Alzheimers Dement (N Y)* **3**, 651-657 (2017).
312. Wong, C.H., Siah, K.W. & Lo, A.W. Estimation of clinical trial success rates and related parameters. *Biostatistics* **20**, 273-286 (2018).
313. Barbie, D.A. et al. Systematic RNA interference reveals that oncogenic KRAS-driven cancers require TBK1. *Nature* **462**, 108-112 (2009).
314. Luo, J. et al. A genome-wide RNAi screen identifies multiple synthetic lethal interactions with the Ras oncogene. *Cell* **137**, 835-848 (2009).

315. Hanahan, D. & Weinberg, Robert A. Hallmarks of Cancer: The Next Generation. *Cell* **144**, 646-674 (2011).
316. Quail, D.F. et al. The tumor microenvironment underlies acquired resistance to CSF-1R inhibition in gliomas. *Science* **352**, aad3018 (2016).
317. Chen, S. et al. Genome-wide CRISPR Screen in a Mouse Model of Tumor Growth and Metastasis. *Cell* **160**, 1246-1260 (2015).
318. Kodama, M. et al. In vivo loss-of-function screens identify KPNB1 as a new druggable oncogene in epithelial ovarian cancer. *Proceedings of the National Academy of Sciences* **114**, E7301 (2017).
319. Song, C.-Q. et al. Genome-Wide CRISPR Screen Identifies Regulators of Mitogen-Activated Protein Kinase as Suppressors of Liver Tumors in Mice. *Gastroenterology* **152**, 1161-1173.e1161 (2017).
320. Katigbak, A. et al. A CRISPR/Cas9 Functional Screen Identifies Rare Tumor Suppressors. *Scientific Reports* **6**, 38968 (2016).
321. Braun, C. et al. Versatile in vivo regulation of tumor phenotypes by dCas9-mediated transcriptional perturbation. *Proceedings of the National Academy of Sciences of the United States of America* **113** (2016).
322. Rana, T.M. et al. Genome-wide CRISPR screen for essential cell growth mediators in mutant KRAS colorectal cancers. *Cancer Research*, canres.2043.2017 (2017).
323. Patel, S.J. et al. Identification of essential genes for cancer immunotherapy. *Nature* **548**, 537-542 (2017).
324. Dong, M.B. et al. Systematic Immunotherapy Target Discovery Using Genome-Scale In Vivo CRISPR Screens in CD8 T Cells. *Cell* **178**, 1189-1204.e1123 (2019).
325. Dixit, A. et al. Perturb-Seq: Dissecting Molecular Circuits with Scalable Single-Cell RNA Profiling of Pooled Genetic Screens. *Cell* **167**, 1853-1866.e1817 (2016).
326. Adamson, B. et al. A Multiplexed Single-Cell CRISPR Screening Platform Enables Systematic Dissection of the Unfolded Protein Response. *Cell* **167**, 1867-1882.e1821 (2016).