Constructing and deconstructing cancer using CRISPR-Cas9

by

Francisco J. Sánchez-Rivera

B.S., Microbiology
University of Puerto Rico-Mayagüez, 2008

Submitted to the Department of Biology
in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

February 2016

Signature of Author _____
Francisco J. Sánchez-Rivera
Department of Biology
December 16th, 2015

Certified by_____
Tyler Jacks
Professor of Biology
Thesis Supervisor

Certified by_____
Michael T. Hemann
Professor of Biology
Co-Chair, Biology Graduate Committee

Constructing and deconstructing cancer using CRISPR-Cas9

by

Francisco J. Sánchez-Rivera

Submitted to the Department of Biology on December 16th, 2015 in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy in Biology

ABSTRACT

Cancer is a genetic disease that arises through the sequential acquisition of genetic and epigenetic alterations in oncogenes and tumor suppressor genes. Large-scale efforts to re-sequence protein-coding genes from human cancer cell lines and tumor biopsies have begun to catalog the spectrum of mutations existing in human cancers. One major limitation of these studies has been the inability to rapidly and systematically determine which of these mutations are causally related to tumorigenesis, particularly in the context of *in vivo* models of the disease. Although existing genetically engineered mouse models (GEMMs) have led to critical insights into the initiation and progression of human cancer, their use for rapid functional characterization of new cancer genes has been historically limited, partly due to the cost and time required to generate appropriate murine models. In the first part of this thesis, I describe a novel CRISPR-Cas9-based approach for rapid functional investigation of candidate genes *in vivo* using well established autochthonous mouse models of cancer. By employing this platform in a GEMM of lung adenocarcinoma *in vivo*, I have functionally validated both known and novel tumor suppressor genes - all of which promote one or more aspects of lung cancer. These findings underscore the power and versatility of this platform for the rapid and functional interrogation of the cancer genome. In the second part of this thesis, I describe the development of a novel CRISPR-Cas9-based genetic screening approach for systematically uncovering genotype-specific vulnerabilities that could be exploited for the therapeutic benefit of specific lung cancer patient subpopulations. I demonstrate the successful application of this system for the discovery of novel *Keap1* mutant-specific genetic dependencies, many of which could potentially be pursued for the clinical benefit of patients whose tumors harbor loss of function mutations in *KEAP1* (~17% of lung adenocarcinoma patients). These results demonstrate the power of CRISPR-based genetic screens for uncovering novel genetic dependencies in the context of clinically relevant cancer-associated genotypes.

Thesis Advisor: Tyler Jacks

Title: Professor of Biology

CURRICULUM VITAE

Francisco J. Sánchez-Rivera

David H. Koch Institute for Integrative Cancer Research at MIT

500 Main Street, 76-421

Cambridge, MA 02139

Tel: 939-292-5555

E-mail: fsr19720@mit.edu

EDUCATION

Fall 2008 - Present	**Massachusetts Institute of Technology** Department of Biology PhD Candidate
Fall 2003 - Spring 2008	**University of Puerto Rico, Mayagüez, PR** Department of Biology Bachelor of Science in Microbiology

RESEARCH

June 2009 - Present	**Department of Biology at MIT and Koch Institute for Integrative Cancer Research** Graduate student in the laboratory of Dr. Tyler Jacks Main Projects: 1) Modeling cancer *in vivo* using the CRISPR-Cas9 system. 2) CRISPR screens for uncovering genotype-specific cancer drug targets. 3) Molecular mechanisms dictating tumor-specific responses to p53 restoration.
June 2008 - Aug 2008	**Whitehead Institute for Biomedical Research** Summer student researcher in the laboratory of Dr. Harvey F. Lodish Project: The role of histone deacetylases in enucleation of mouse fetal erythroblasts.
June 2007 - Aug 2007	**Department of Biology at MIT** Summer student researcher in the laboratory of Dr. Stephen P. Bell Project: Identification of dominant negative helicase mutants by insertional mutagenesis of the yeast replicative helicase Mcm2-7.
Jan 2006 - May 2008	**University of Puerto Rico, Mayagüez Campus** Department of Biology Undergraduate student researcher in the laboratory of Dr. Carlos Ríos-Velázquez Main project: Development and analysis of gene libraries from tropical hypersaline microbial mats.

ACADEMIC HONORS

2012	AACR Minority Scholar in Cancer Research Award to attend the 2012 AACR Meeting
2011	Repligen KIICR Graduate Fellowship
2006 – 2008	PR - LSAMP Scholar
2006 – 2008	MARC Scholar

PRESENTATIONS

2015 **Invited speaker**, HMS Graduate Course - *Cell Biology: Current Topics and Techniques in Cancer Biology Research*

2014 **Invited speaker**, Koch Institute Crossfire Seminar - *Applications of the CRISPR-Cas9 System in Cancer Biology*

2014 **Speaker**, Colrain Meeting, Harvard, MA

2014 **Poster presenter**, Mechanisms and Models of Cancer Meeting, Cold Spring Harbor Laboratory, NY

2014 **Speaker**, AACR Annual Meeting, San Diego, CA

2013 **Poster presenter**, Frontiers in Tumour Heterogeneity and Plasticity, Nature/CNIO, Madrid, Spain

2012 **Speaker***,* AACR Annual Meeting, Chicago, IL

2011 **Poster presenter**, Mechanisms and Models of Cancer Meeting, Salk Institute, La Jolla, CA

2008 **Poster presenter**, ASM General Meeting, Boston, MA

2007 **Poster presenter**, ASCB Annual Meeting*,* Washington, DC

2007 **Poster presenter**, ACS 234th National Meeting and Exposition*,* Boston, MA

2007 **Poster presenter**, ASM General Meeting, Toronto, Canada

TEACHING EXPERIENCE

Fall 2015	**Biological and Biomedical Sciences Graduate Program at Harvard Medical School**
	Invited lecturer - Cell Biology: Current Topics and Techniques in Cancer Biology Research
Spring 2013	**Department of Biology at Harvard University**
	Teaching assistant for LS1b - Genetics, Genomics and Evolution
Spring 2012	**Department of Biology at MIT**
	Teaching assistant for 7.16 - Biotechnology II
Oct 2010, 2011, 2014, 2015	**Massachusetts Junior Academy of Sciences**
	Scientific judge for the MJAS Symposium
Fall 2009	**Department of Biology at MIT**
	Teaching assistant for 7.012 - Introductory Biology
Summer 2007 & 2008	**University of Puerto Rico, Mayagüez Campus**
	Workshop assistant: Genetic Engineering Applied to Molecular Microbiology

PUBLICATIONS

1. Soto-Feliciano Y.M., Bartlebaugh, J.M.E.*, Liu, Y.*, **Sánchez-Rivera, F.J.**, Weintraub, A.S., Bhutkar, A., Jacks, T., Young, R.A., Hemann, M.T. PHF6 controls B-cell identity in acute lymphoblastic leukemia. *Submitted.* ***Co-second author**

2. **Sánchez-Rivera, F.J.**, Ryan, J., Beytagh, M.C., Soto-Feliciano, Y.M., Xuan, L., Feldser, D.M., Hemann, M.T., Letai, A., Jacks, T. The level of mitochondrial apoptotic priming is a critical determinant of cell fate upon p53 restoration. *In preparation*.

3. Tammela T., **Sánchez-Rivera, F.J.**, Helenius K., Joshi N., Azimi R., Kerper N., Park Y., Wesselhoeft A., Schmidt L., Jacks T. Wnt-secreting niche controls intratumoral hierarchy in lung adenocarcinoma. *In preparation.*

4. Roper, J.*, Tammela, T.*, Cetinbas, N., **Sánchez-Rivera, F.J.**, Almeqdadi, M., Azimi, R., Park, Y., Papagiannakopoulos, T., Kedrin, D., Jacks, T.#, Yilmaz, O.# Rapid modeling of colon cancer by organoid transplantation and mucosal-directed somatic genome editing. *In preparation.* ***Co-first author, #Co-corresponding author.**

5. Mazur, P.K.*#, Herner, A.*, Mello, S.S., Wirth, M., Hausmann, S., Kuschma, T., **Sánchez-Rivera, F.J.**, Lofgren, S., Hahn, S.A., Vangala, D., Trajkovic-Arsic, M., Gupta, A., Heid, I., Noël, P.B., Braren, R., Erkan, M., Kleeff, J., Sipos, B., Sayles, L.C., Heikenwälder, M., Heßmann, E., Ellenrieder, V., Esposito, I., Jacks, T., Bradner, J.E., Khatri, P., Sweet-Cordero, E.A., Attardi, L.D., Schmid, R.M., Schneider, G., Sage, J.#, Siveke, J.T.# (2015) Combined inhibition of BET family proteins and histone deacetylases as a potential epigenetics-based therapy for pancreatic ductal adenocarcinoma. *Nature Medicine*. **21,** 1163-1171. ***Co-first author, #Co-corresponding author.**

6. **Sánchez-Rivera, F.J.**, Jacks, T. (2015) Applications of the CRISPR-Cas9 system in cancer biology. *Nature Reviews Cancer.* **15,** 387-395.

7. **Sánchez-Rivera, F.J.***, Papagiannakopoulos, T.*, Romero, R., Tammela, T., Bauer, M.R., Bhutkar, A., Joshi, N.S., Subbaraj, L., Bronson, R.T., Xue, W., Jacks, T. (2014) Rapid modelling of cooperating genetic events in cancer through somatic genome editing. *Nature.* **516,** 428–431. ***Co-first author.**

8. Feldser, D.M., Kostova, K.K., Winslow, M.M., Taylor, S.E., Cashman, C., Whittaker, C.A., **Sánchez-Rivera, F.J.**, Resnick, R., Bronson, R.T., Hemann, M.T., Jacks, T. (2010) Stage-specific sensitivity to p53 restoration during lung cancer progression. *Nature.* **468,** 572-575.

REFERENCES

Dr. Tyler Jacks, PhD
Koch Institute for Integrative Cancer Research at MIT, 76-453
77 Massachusetts Avenue
Cambridge, MA 02139
Tel: 617-253-0263
E-mail: tjacks@mit.edu

Dr. Michael T. Hemann, PhD
Koch Institute for Integrative Cancer Research at MIT, 76-361
77 Massachusetts Avenue
Cambridge, MA 02139
Tel: 617-324-1964
E-mail: hemann@mit.edu

Dr. Robert A. Weinberg, PhD
Whitehead Institute for Biomedical Research, WI-361
Nine Cambridge Center
Cambridge, MA 02142-1479
Tel: 617-258-5159
E-mail: weinberg@wi.mit.edu

ACKNOWLEDGEMENTS

I am very thankful to have been able to form part of the family that is the Jacks Lab Family. I can honestly say that the last 6 1/2 years have been the best and most transformative years in my life, both personally, as well as scientifically. It is hard to believe that my defense date is finally coming, and it is even harder to believe that I will have to leave this place the day after, a place that I consider to be my home, and leave many people behind, people who I consider to be my family. Thank you all - past and present lab members - for making the Jacks Lab feel like home. I will truly miss all of you - every single one of you.

I would like to thank Tyler for letting me join his group and for somehow believing that I could make it some day in science. I have many fond memories of my interactions with Tyler, even from way back in the day. I remember that one of the first times I actually saw Tyler in real life was when the Koch Institute groundbreaking took place. And then a few days after, I remember talking to him during the interview weekend and, as soon as I got back to Puerto Rico, I emailed him asking to see if he could send me a paper he had mentioned during his talk (since I had no access to these articles back home) and one of the things that impressed me the most about Tyler was that he actually emailed me back, and even attached the PDF to the email (Kirsch et al., 2007)… I know it sounds simple but the fact that Tyler took time out of his busy schedule to email me back and attach a paper to the email meant a lot to me. To me it meant that he actually cared about people. And that is one of the qualities I admire the most about him. He truly cares about his people, and this is something that I will emulate when I have a chance to get my own lab. Tyler - thank you for making me the scientist that I am today. Thank you for letting me fail when I needed to fail, and thank you for letting me be successful when I needed to be successful. Thank you for believing in me, and for trusting me, and for giving me scientific freedom, and honestly just for taking the chance with me. You are a role model to follow in every sense of the word, and I am very lucky to have had the chance to be a student in your lab. I will truly miss you.

I would like to sincerely thank Mike Hemann and Bob Weinberg for all the years of support and guidance as members of my thesis committee. I am deeply honored that I have had the chance to have such an amazing duo of brilliant minds helping me grow as a scientist. The guidance that I have received from each one of you has had a tremendous impact in my development as a scientist. I would also like to extend my most sincere thanks to Steve Elledge for serving as my external thesis committee member. It is really an honor that you agreed, since your science has motivated me to do much of the research that I am trying to do now and will do in the future.

I would like to thank my friends in the lab, past and present, but most importantly, I would like to thank David McFadden, Tuomas Tammela, and Thales Papagiannakopoulos. I think it will be very hard to come across such good friends in the future, friends that I consider to be my brothers. Hopefully we can all reunite and have the symposiums we always talk about. I would like to give a special thanks to Thales, Matt, and Rodrigo, with whom I shared some of the best scientific moments of my career. I will always think back to the mythical CRISPR bay days, while listening to some of the main songs we used to listen to while putting 14-16hr work days, 7 days a week. Its hard to imagine that I will not be sitting at the bay anymore. It is a very special bay…

I would like to thank Mary Clare for being the most amazing UROP in every sense of the word, for working extremely hard (which in turn made me work harder) and for having a ton of patience to deal with me. I know it wasn't easy! I'm looking forward to hearing about all of your YGG awards in the future - you truly deserve them!

I would like to thank Etienne Meylan and David Feldser for allowing me to work with them during my rotation, and especially Feldser for giving me a lot of guidance and mentoring at the beginning. I would like to thank Nadya for being a great baymate, and a very critical one when I needed feedback. I would like to thank Wen for being a great friend and colleague, and for being one of the nicest scientists I know. I would like to thank AJ Bhutkar - an extremely talented bioinformatician without whom I would not have been able to achieve many of the experiments presented in this thesis. I will greatly miss working with you. I would like to thank Judy for being the best in every sense of the word. I will really miss your enthusiasm, smile, work ethic, and all the jokes! I will truly miss you. And Anne Deconinck for welcoming me to the lab the first day and being my friend for all these years, and for all the help she gave me when she was still around. I still miss you! I would also like to thank Ines and Karen for running the lab and also for helping me a lot with grant applications, letters, and other things. I really appreciate the hard work you've done. I would like to thank everyone else in the lab, past and present, for all their help and support in one way or another.

I would also like to thank my collaborators: Chris Fillmore, Joo-Hyeon Lee, Carla Kim, Isaac Harris, Jose M. Amich, Joanne Brugge, Peter Choi, Matt Meyerson, Jatin Roper, Naniye Cetinbas, Ömer Yilmaz, Jeremy Ryan, and Tony Letai, among others. These collaborations have been invaluable to me.

I would also like to thank the entire staff in the MIT Biology Department, particularly Betsey Walsh, Mandana Sassanfar, Steve Bell, Harvey Lodish, and Frank Solomon. I would like to particularly thank Mandana and Steve for giving me my first exposure to MIT through their MSRP summer program, as well as Harvey for allowing me to work in his lab the summer right before starting graduate school.

Beyond MIT, I would like to give my most sincere thanks to my undergraduate mentor, Carlos Ríos-Velázquez, who was one of the professors that inspired me to become a scientist and pursue graduate studies in Biology. His invaluable mentoring was critical in my decision to carry out research in Biology, particularly at the MIT Biology Department. Carlos made sure that I was able to carry out solid biological research and present my work in multiple local and US scientific forums, even when the funding was scarce. Beyond research, Carlos instilled in me the importance of acknowledging those who help you along the way, and this is something that I greatly value and will always value. I would also like to thank my two graduate student mentors in Carlos' lab - Josué Malavé-Orengo and Roger Valle-Molinares. They contributed immensely to my development as a young scientist. Special thanks to Josué for all of our technical conversations about biology, for constantly motivating me, and for supporting me in my decision to apply to MIT, even when I thought the chances of me getting in were pretty slim. I would also like to thank Franklin Carrero-Martínez, another great mentor and friend who inspired me to pursue science, and who constantly challenged (and still challenges) my ideas about science and life. Our discussions have been invaluable in my growth as a person and as a scientist. I would also like to thank my genetics professor, Juan Martínez-Cruzado, who was another professor that greatly inspired me to become a scientist. His lectures are still the major highlights of my undergraduate education, and they served as a source of motivation to pursue graduate studies in cancer genetics.

So now the difficult ones…

I would first like to thank my mother because she's always been there for me and believed in me no matter what, even when I messed up. And when I messed up, she put me back on the right path… more than once… If it wasn't for her, I wouldn't be here getting a PhD. In fact, I don't know what I would be doing right now… So I owe a big part of who I am to her. Gracias mami!

I would also like to thank my father because he's always supported me and has always stimulated me to try to be the best I can, and to excel in what I do. He's a role model to me because he worked extremely hard to get what he wanted, and he stimulated me to also work hard to get what I wanted. Gracias Paco!

I would also like to thank my brother for constantly supporting me in everything I've done. He's always believed in me, and has always stimulated me to achieve the things that I want to achieve. He's never told me that something is impossible, and I thank him for that. He's always listened to me when I've had problems, and has always supported me. Even though we don't live close to each other, I feel your strong support. Gracias Tony!

I would also like to thank my other family in Puerto Rico - Steven, Kristina, Valentina, Neggy, Eugenio - for also supporting me for 13 years now.

Then there's my "boo"… 10 days from now, it will be 13 years together. Hard to imagine that the first time I met you, that was going to be it! I love you with all my heart, and I am looking forward to the next time we get to sit right next to each other, for 12 hours a day or so, for ~ 2 months, writing a thesis haha… thanks for keeping me on the right track, for believing in me, and for never giving up on me… and for being in this 13-year trip with me… I can't imagine life without you… honestly, and I can't wait to start the next chapter in NYC so that we can walk Silvia around the city and look at her play with other bulldogs… I have a feeling we'll be very happy :)

And lastly… the toughest one for me. My aunt, Blanca Rivera Sanabria, to whom I owe everything that I am today… Without her, I wouldn't be anywhere near where I am right now… Even though you passed away last year, I still would like to think that you would be proud to witness what I have accomplished now, which is all thanks to you. It is to you that I dedicate all of my present and future work, and it is to you that I dedicate this thesis. I would like to think that you would be proud of me right now, just as you always were…

Sincerely,

Francisco

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1: Introduction

CHAPTER 2: Rapid modeling of cooperating genetic events in cancer through somatic genome editing

CHAPTER 3: Systematic identification of *Keap1* mutant-specific vulnerabilities in lung adenocarcinoma through CRISPR-Cas9 genetic screening

CHAPTER 4: Discussion

APPENDIX 1: The level of mitochondrial apoptotic priming is a critical determinant of cell fate upon p53 restoration

APPENDIX 2: *Brg1* suppresses several aspects of lung cancer and its inactivation creates unique genetic dependencies

Chapter 1 – Introduction

Part I – Modeling cancer in the mouse

I - Historical Perspective

As Douglas Hanahan and Robert Weinberg stated in their seminal "*The Hallmarks of Cancer*" essay, tumorigenesis is a multistep process that is characterized by genetic (and epigenetic) alterations that drive the progressive transformation of normal cells into highly malignant derivatives (Hanahan and Weinberg 2000). This highly complex process, however, takes place in the context of an entire organism. Implicit in such statement is the fact that whole-organism physiology - from the presence of an intact immune system to an intricate set of whole-body multi-organ metabolic reactions - cannot be fully recapitulated using *in vitro* cell culture-based models. Therefore, despite the fact that the foundations and some of the most significant breakthroughs of modern cancer research can be unequivocally attributed to discoveries made using *in vitro* culture systems, these methods are unable to faithfully recapitulate the endogenous *in vivo* environment in which such sequential alterations take place. Moreover, dissecting the mechanisms driving several of the hallmarks of cancer, such as invasion and metastasis, necessitates the use of appropriate *in vivo* systems in which each of the sequential tumorigenic steps can be properly interrogated in the context of an entire organism.

As it often happens in science, the technologies needed to solve a given problem or answer a particular question co-evolve with the problem itself. Therefore, while G. Steven Martin was discovering the v-*src* oncogene (Martin 1970), and while the groups of Michael Bishop and Harold Varmus where discovering that v-*src* was in fact a virally hijacked version of the endogenous *Src* proto-oncogene (Stehelin et al., 1976), and while the groups of Robert Weinberg, Michael Wigler, and Mariano Barbacid where identifying and cloning the first *bona fide* human oncogene (Tabin et al., 1982, Parada et al., 1982, Reddy et al., 1982, and Taparowsky et al., 1982), and the Weinberg group was identifying and cloning the first human tumor suppressor gene (Friend et al., 1986), Rudolf Jaenisch and Beatrice Mintz were microinjecting SV40 viral DNA or Moloney leukemia retroviruses into early embryos to generate the very first transgenic mice in history (Jaenisch and Mintz 1974, Jaenisch 1976). That foreign DNA molecules, including retroviruses and recombinant DNA, could be utilized to generate transgenic animals meant that the biological role of the recently discovered oncogenes and tumor suppressor genes could be dissected in the *in vivo* context. Outstanding questions in the field, such as how oncogene activation and tumor suppressor inactivation can lead to the malignant transformation of otherwise normal cells, or how multiple oncogenes cooperate in the context of tumorigenesis, could now be tackled using physiological *in vivo* models. In the next few sections, I briefly discuss the pioneering studies that opened the door for modeling cancer in mice.

A. Transgenic technologies

As mentioned above, the pioneering work of Rudolf Jaenisch and Beatrice Mintz firmly established the feasibility of generating transgenic mice using foreign viral DNA or retroviruses (Jaenisch and Mintz 1974, Jaenisch 1976). A few years later, Ralph Brinster and Richard Palmiter generated the first set of transgenic mouse models of cancer by demonstrating that widespread expression of the SV40 T antigen by itself or fused to metallothionein regulatory elements resulted in the development of choroid plexus tumors with high penetrance (Brinster et al., 1984), as well as hepatocellular carcinomas and pancreatic adenomas (Messing et al., 1985). Subsequent studies by Douglas Hanahan, as well as the Brinster and Palmiter groups, demonstrated that the SV40 T antigen could drive pancreatic tumorigenesis when expressed under the control of the insulin or elastase regulatory regions, respectively (Hanahan 1985, Ornitz et al., 1987). These landmark studies using the SV40 T antigen greatly influenced the way cancer research was pursued, both from a biological standpoint, as well as a methodological standpoint. The fact that large cohorts of transgenic mice could now be generated, each of which could serve as a physiological model in which to study different types of human cancers, fundamentally changed the way researchers approached the study of cancer.

Building upon the aforementioned discoveries of cellular oncogenes, as well as the recent development of transgenic technologies, several groups developed transgenic mice expressing the *Ras* and *Myc* oncogenes, in addition to several other novel oncogenes discovered around that time (for comprehensive and contemporaneous

reviews, see Cory and Adams 1988, Hanahan 1988 and Hanahan 1989). Brinster and colleagues developed the first set of transgenic mice expressing either wild type *HRAS* (referred to as c-Ha-*ras* during this time) or mutant *HRAS* (containing a Glycine to Valine mutation at position 12 of the protein, abbreviated as $HRAS^{G12V}$) under the control of elastase regulatory regions (Quaife et al., 1987). Strikingly, expression of wild type *HRAS* was well tolerated throughout embryogenesis and adulthood, while expression of $HRAS^{G12V}$ was incompatible with life. Transgenic mice expressing $HRAS^{G12V}$ were characterized by having an abnormal pancreas, with clear signs of dysplasia, marked disruption of pancreatic architecture, and histological features reminiscent of malignant carcinomas. In contemporaneous studies, Gerlinger and colleagues developed one of the first transgenic models that expressed the *HRAS* oncogene in the mammary glands by employing a transgene expressing mutant *HRAS* under the control of murine whey acidic protein (Wap) regulatory elements (Andres et al., 1987). Similar to the pancreas study described above, expression of mutant *HRAS* resulted in adenocarcinomas, albeit with the caveat that these animals actually developed salivary gland adenocarcinomas due to inappropriate expression of the transgene (Andres et al., 1987). In 1984, Philip Leder's group published a seminal study describing a set of 13 transgenic strains of mice overexpressing the c-*myc* proto-oncogene under the control of the mouse mammary tumor virus (MMTV) regulatory sequences (Stewart et al., 1984). Strikingly, overexpression of c-*myc* resulted in the development of mammary adenocarcinomas in a subset of the transgenic founders. Notably, the progeny of this subset of transgenic founders also developed mammary adenocarcinomas with very high penetrance (Stewart et al., 1984). Three years later,

the same group reported the generation of transgenic mice expressing v-Ha-*ras* (the viral oncogene) under the control of MMTV regulatory sequences (Sinn et al., 1987). Similar to their previous study, transgenic mice expressing v-Ha-*ras* in the mammary glands developed mammary adenocarcinomas with high penetrance (as well as other tumor types, such as salivary gland adenocarcinomas and, with lower frequency, lymphomas). Notably, when they generated bi-transgenic mice co-expressing v-Ha-*ras* and c-*myc,* the mice succumbed to a dramatically accelerated disease, demonstrating that these oncogenes act cooperatively *in vivo* in the context of tumorigenesis (Sinn et al., 1987). Another notable and contemporaneous transgenic mouse model of cancer was reported by the Palmiter and Brinster groups in 1985 - the Eμ-Myc model (Adams et al., 1985). This mouse model was developed by expressing c-*myc* under the control of the immunoglobulin heavy chain enhancer (Eμ). These mice, which develop B-cell lymphomas with close to 100% penetrance, have been an invaluable model for studying several fundamental mechanisms of cancer biology, such as oncogene cooperation (Vaux et al., 1988, Strasser et al., 1990) and the role of the $p19^{Arf}$-p53 axis in mediating tumor suppression (Eischen et al., 1999, Schmitt et al., 1999, Jacobs et al., 1999).

Despite the fact that these pioneering transgenic mouse models have played an invaluable role in uncovering several fundamental aspects of cancer biology, they have several limitations. Due to the use of strong enhancer / regulatory regions, transgenes are almost always expressed at supraphysiological levels, which can lead to confounding results due to the fact that the expression levels do not mimic those of the endogenous proto-oncogenes and oncogenes. Another limitation tightly linked to supraphysiological expression is the fact that transgenes are often randomly integrated

into the genome. Therefore, the gene that is embedded within the transgene is not under the control of its native regulatory elements. In addition, depending on the chromatin landscape, the transgene can either be highly expressed (if it integrates in a euchromatic region) or even become silenced (if it integrates in a heterochromatic region) (Garrick et al., 1996, Henikoff et al., 1998, Garrick et al., 1998). Beyond these limitations, the recessive nature of most tumor suppressor genes made the use of transgenic technologies virtually impossible to study this class of cancer genes due to the fact that their study required their loss of function. As an attempt to overcome these and other shortcomings, the mouse modeling community turned its attention to the emerging area of gene targeting via homologous recombination.

B. Gene targeting technologies

i. Homologous recombination in yeast

Just as transgenic technologies were being developed at the same time as viral and human oncogenes were being discovered, the technologies for precise gene targeting via homologous recombination were also being developed contemporaneously. The first example of gene replacement via homologous recombination was published by the group of Gerald Fink in 1978 (Hinnen et al., 1978). In this landmark study, Fink and colleagues demonstrated that *Saccharomyces cerevisiae* leucine auxotrophs (*LEU2* mutant) could be rescued via the incorporation of a wild type copy of the *LEU2* gene carried in a hybrid bacterial-yeast plasmid. In addition to pioneering the technique of yeast transformation, Fink and colleagues demonstrated that the exogenous *LEU2* gene was incorporated into the endogenous *LEU2* locus, proving - for the first time - that

exogenous DNA sequences could replace endogenous DNA sequences provided that there was sufficient sequence homology between the two (Hinnen et al., 1978). Subsequent studies by Terry Orr-Weaver, Jack Szostak and Rodney Rothstein confirmed Fink's work and further demonstrated that the linear ends of a DNA molecule were highly recombinogenic with endogenous homologous sequences and that the experimental introduction of DNA double-strand breaks into exogenous templates increased the frequency of homologous recombination by several orders of magnitude (Orr-Weaver et al., 1981 and reviewed in Orr-Weaver et al., 1983 and in Szostak 1983). These fundamental molecular biology experiments in yeast unequivocally paved the way towards the adaptation of homologous recombination technologies for genome engineering in mammalian cells.

ii. Homologous recombination in mammalian cells

Pioneering work by the groups of Mario Capecchi and Oliver Smithies in the 1980s demonstrated that gene targeting via homologous recombination in mammalian cells was experimentally possible. Similar to the genetic correction of the mutant *LEU2* gene by the Fink group, Capecchi and colleagues demonstrated that an integrated/ exogenous gene encoding a mutant neomycin-resistance cassette (Neo-R) could be replaced with an exogenous corrected Neo-R cassette via homologous recombination in mammalian cells (Thomas et al., 1986). Notably, they demonstrated the correction of both point mutations and deletions in the mutant Neo-R cassette via homologous recombination. A contemporaneous study by the group of Oliver Smithies demonstrated efficient replacement of the endogenous human *β*-globin locus via homologous

recombination with exogenous plasmids containing significant sequence homology (Smithies et al., 1985). Fortuitously, Martin Evans had just pioneered the derivation and culturing of karyotypically-stable mouse embryonic stem (ES) cells (Evans and Kaufman 1981), which where utilized by both the Capecchi and Smithies groups to demonstrate that endogenous genes could be modified via homologous recombination (Thomas et al., 1987, Doetschman et al., 1987, Koller and Smithies 1989). Specifically, they utilized the X-linked hypoxanthine phosphoribosyltransferase gene (HPRT) and the β_2-globin loci as prototypical examples to demonstrate efficient mutation and correction of endogenous genes (Thomas et al., 1987, Doetschman et al., 1987, Koller and Smithies 1989). Mario Capecchi's group took it a step further by developing the powerful and widely-used positive-negative selection strategy based on the incorporation of Neo-R and thymidine kinase (TK) cassettes into targeting vectors in order to enrich for correct homologous recombination events in ES cells (Mansour et al., 1988). These groundbreaking advancements in the field of genome engineering in mammalian cells revolutionized the way cancer biologists started to approach the study of oncogenes and tumor suppressor genes due to the fact that this technology had the potential to overcome many of the limitations associated with transgenic mouse models. The fact that endogenous loci could be engineered with almost surgical precision offered several advantages (**Figure 1**). First, it meant that endogenous promoters (as opposed to exogenous strong enhancers and regulatory regions) could be utilized to drive physiological expression of proto-oncogenes and oncogenes (as opposed to the supraphysiological and incontrollable expression levels generally achieved with transgenes). It also meant that heterozygous events could be properly modeled, as

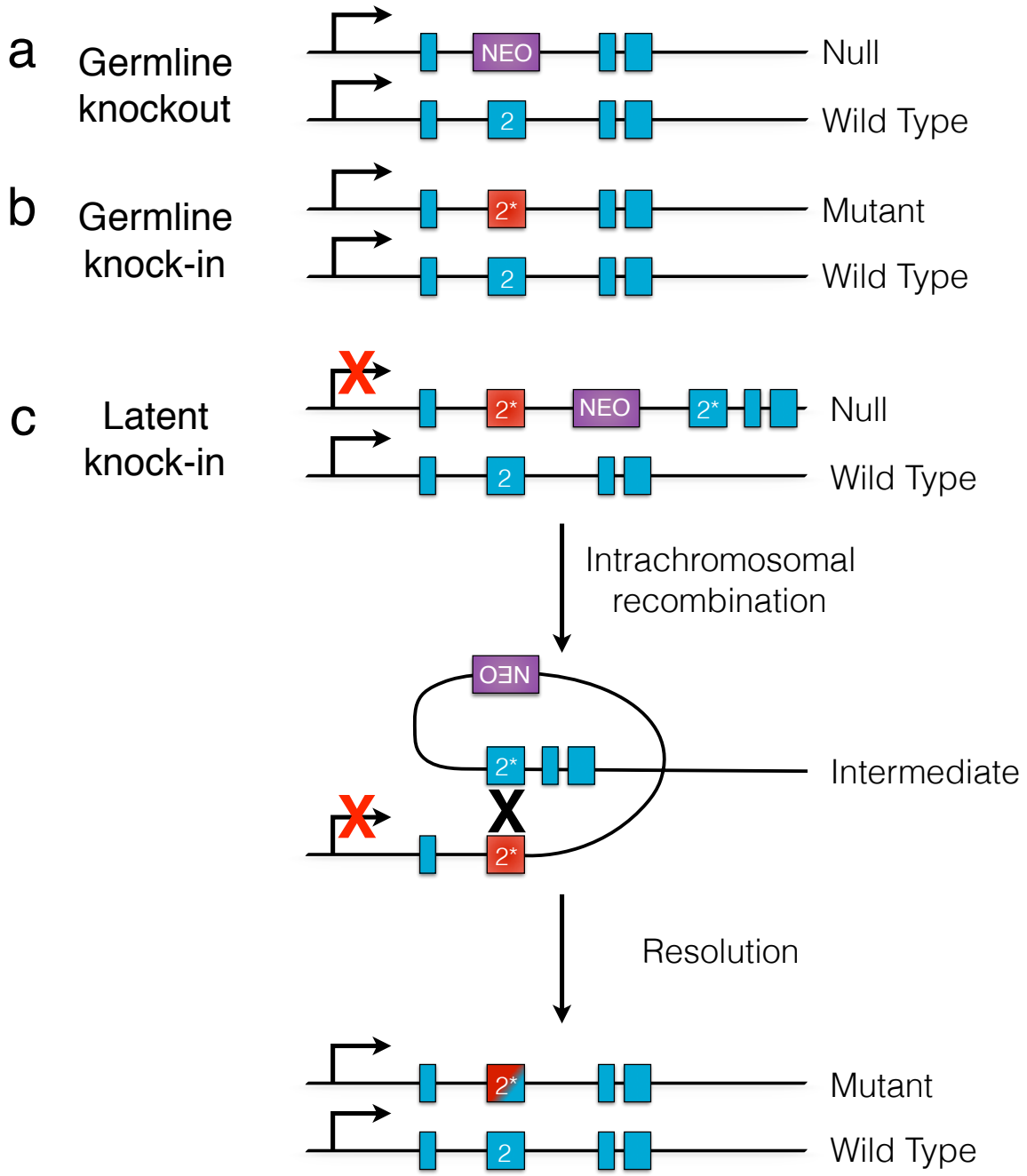


Figure 1: Germline endogenous alleles. (a) ***Knockout allele*** - designed to constitutively knockout a gene of interest by replacing functionally important coding exons with a selectable marker. **(b)** ***Knock-in allele*** - designed to constitutively express a mutation of interest by engineering the mutation into the relevant exon. **(c)** ***Latent alleles*** - designed to trigger the spontaneous somatic activation of a mutation via the introduction of a tandem duplication of the mutant exon, which promotes an intrachromosomal recombination event *in vivo*. Figures (a) and (b) were adapted from Frese & Tuveson (2007).

mice harboring single heterozygous mutations in either tumor suppressor genes or oncogenes could be generated via intercrossing. Not surprisingly, this led to an

explosion of mouse models harboring mutations in cancer-associated genes, particularly tumor suppressor genes, that were generated via gene targeting through homologous recombination in ES cells.

C. Tumor suppressor knockout mice

Armed with this new technology, several cancer researchers successfully generated multiple mice harboring knockout mutations in some of the most commonly mutated tumor suppressor genes (reviewed in Jacks 1996). In 1992, the groups of Robert Weinberg, Allan Bradley and Martin Hooper simultaneously reported the generation of the first tumor suppressor knockout mouse - the *Rb* knockout mouse (Lee et al., 1992, Jacks et al., 1992, Clarke et al., 1992). Interestingly, heterozygous *Rb* knockout mice were predisposed to the development of pituitary tumors and such tumors invariably underwent loss of heterozygosity of the 2nd wild type allele of *Rb* (Jacks et al., 1992), supporting Alfred Knudson's famous two-hit hypothesis (Knudson 1971). Strikingly, homozygous mutation of *Rb* was found to be embryonic lethal. That knockout of a tumor suppressor gene was incompatible with life would prove to be a common feature for the majority of tumor suppressor genes inactivated via gene targeting in the mouse during the next several years, such as *Wt1* (Kreidberg et al., 1993), *Apc* (Fodde et al., 1994, Oshima et al., 1995, Moser et al., 1995), *Nf1* (Brannan et al., 1994, Jacks et al., 1994,), *Nf2* (McClatchey et al., 1997), *Brca1* (Hakem et al., 1996, Gowen et al., 1996), *Vhl* (Gnarra et al., 1997), and *Pten* (Di Cristofano et al., 1998, Stambolic et al., 1998),

among others. The second tumor suppressor knockout mouse - the *p53* knockout mouse - was reported by a number of groups, including the group of Allan Bradley and the group of Robert Weinberg (Donehower et al., 1992, Jacks et al., 1994). Interestingly, homozygous *p53* mutation was not embryonic lethal; however, both heterozygous and homozygous mutant animals rapidly succumbed to cancer, particularly T-cell lymphomas and soft tissue sarcomas (Donehower et al., 1992, Jacks et al., 1994).

The fact that aberrant expression/overexpression of an oncogene in the context of a transgenic mouse was found to promote tumorigenesis, and the fact that homozygous or heterozygous genetic inactivation of a tumor suppressor gene via gene targeting was also found to promote tumorigenesis unequivocally demonstrated that mutation of either class of genes was a *causal* event in human cancer. This was a very important conclusion at the time since mutation in either class of genes appeared to be sufficient to transform an otherwise normal cell into a malignant cancer cell. With these relatively efficient genome engineering technologies at hand, the field of cancer biology was determined to move towards the generation of faithful, physiologically-relevant genetically engineered mouse models (GEMMs) of cancer harboring mutations in genes relevant to the human disease. Despite the fact that these technologies fundamentally changed the way cancer researchers started to tackle the study of oncogenes and tumor suppressor genes *in vivo*, there were still a number of technological barriers that needed to be overcome in order to fulfill the promise that GEMMs of cancer were supposed to deliver. For example, how could researchers limit malignant transformation to a small number of cells in the mouse to closely mimic sporadic tumorigenesis (while also avoiding the lethality associated with germline mutation of several oncogenes and

tumor suppressor genes)? How could researchers spatiotemporally (and perhaps reversibly) modulate the expression of oncogenes and tumor suppressor genes *in vivo*. Some of these and other technological hurdles were solved by the introduction of sophisticated conditional mouse modeling technologies, which are the focus of the next section.

II - Sophisticated GEMMs of cancer

Two of the most transformative events in the history of mouse modeling were (1) the introduction of inducible-reversible systems based on tetracycline-regulatable genetic elements and hormone-regulatable estrogen receptor gene fusions, and (2) the introduction of site-specific recombinase (SSR) systems for inducible manipulation of the genome. These tools allowed cancer researchers to spatiotemporally regulate the expression of oncogenes and tumor suppressor genes *in vivo*, overcoming some of the major technical issues mentioned at the end of the last section.

A. Tetracycline-regulatable systems

The development and application of tetracycline-regulatable systems represented a major milestone for the rapidly evolving field of mouse modeling (**Figure 2A-B**). The first demonstration of the power of this system for inducible and reversible gene expression was published in 1992 by Wendenburg and colleagues, who demonstrated that juxtaposing three *tet* operator (tetO) sequences to the TATA-box of the normally constitutive cauliflower mosaic virus (CaMV) 35S promoter was sufficient to render expression of the β-glucuronidase (GUS) reporter gene inducible and reversible in transgenic tobacco plants (Gatz et al., 1992). In this variant of what was later referred to

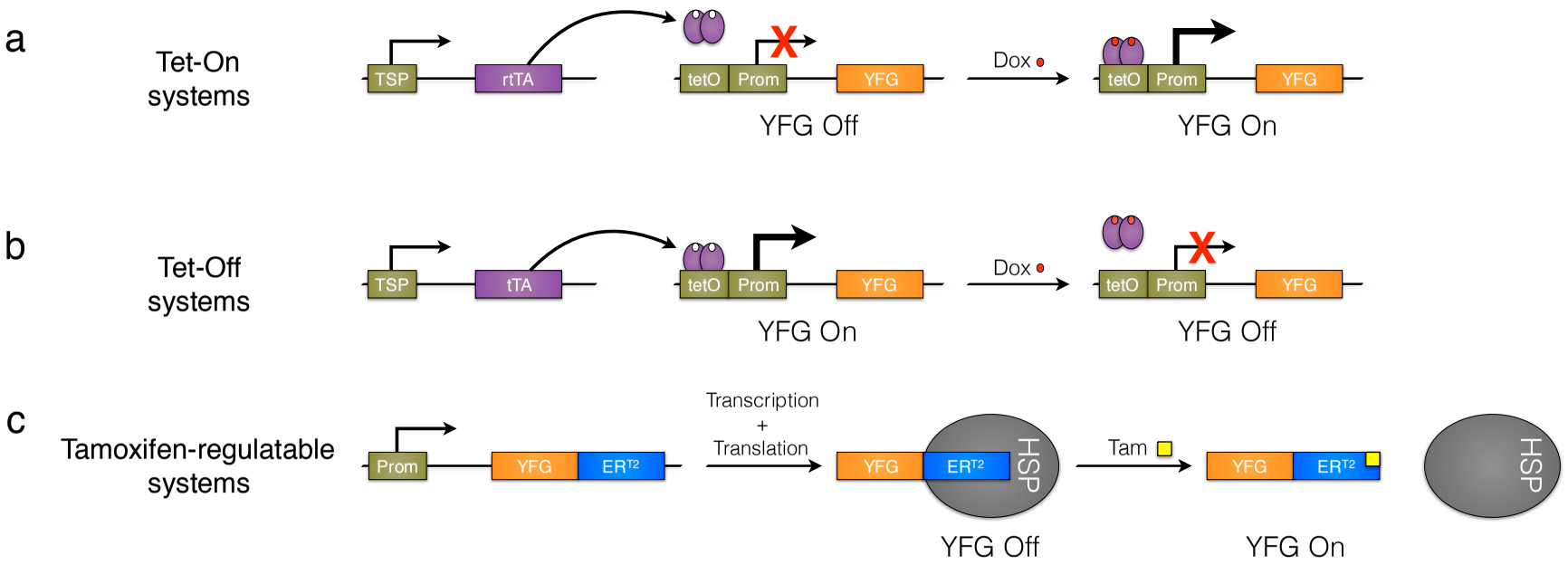


Figure 2: Tetracycline- and hormone-regulatable systems. (a-b) *Tet-On* and *Tet-Off* systems - Juxtaposing tetO regulatory sequences to the coding sequence of your favorite gene (YFG) allows for tetracycline-regulatable induction - *Tet-On* (a) - or repression - *Tet-Off* (b) - of expression of YFG upon administration of the tetracycline analog doxycycline (Dox). **(c) *Hormone-regulatable systems*** - Fusing YFG to a fragment of the estrogen receptor (ER) allows for hormone-regulatable activity of YFG at the protein level. In the absence of tamoxifen, the fusion protein is sequestered in the cytoplasm by heat shock proteins (HSPs). Tamoxifen disrupts this interaction, which liberates the protein of interest. TSP: tissue-specific promoter. Prom: promoter. Adapted from Frese & Tuveson (2007).

as the “tet-on” configuration (**Figure 2A**), the GUS gene was only expressed when plants were exposed to tetracycline. Remarkably, this induction of gene expression was shown to be reversible upon withdrawal of tetracycline. In a contemporaneous study, Gossen and Bujard described the generation of the “tet-off” system (**Figure 2B**), in which the *tet* repressor was fused to the virion protein 16 (VP16) of the herpes simplex virus to generate the tetracycline-controlled transactivator (tTA) protein (Gossen and Bujard 1992). Using a fusion gene encoding *Photinus pyralis* luciferase under the control of 1-7 tetO sequences as a reporter in HeLa cells stably expressing tTA, they demonstrated strong (and reversible) gene repression and induction in the absence and presence of tetracycline, respectively (Gossen and Bujard 1992). Three years later, Bujard and colleagues reported the generation of the reverse tetracycline-controlled transactivator (rtTA) protein, which forms the basis of one of the most widely used “tet on” systems (**Figure 2A**) (Gossen et al., 1995). In this system, potent induction or repression of gene expression is achieved upon administration or withdrawal of tetracycline (or the tetracycline analog doxycycline), respectively (Gossen et al., 1995). These groundbreaking studies were followed by the generation of transgenic mice expressing tTA or rtTA under the control of the strong CMV promoter (Kistner et al., 1996).

These novel technologies offered cancer researchers the ability - for the first time - to turn on and turn off oncogenes and tumor suppressor genes *in vivo* to tackle some of the most pressing questions in the field at the time. For example, did established tumors depend on continuous expression of the oncogene? In other words, where tumors *addicted* to the oncogenic lesions that they harbored in their genomes? The first group

to tackle this was Lothar Hennighausen's group, who generated bi-transgenic mice expressing the SV40 T antigen under the control of a tetO promoter as well as tTA under the control of MMTV regulatory sequences (Ewald et al., 1996). These mice developed extensive salivary gland hyperplasias by 16 weeks of age. They made three important observations that are worth pointing out. First, they demonstrated that inactivation of the SV40 T antigen at 16 weeks of age (through the administration of tetracycline) led to the regression of these hyperplastic lesions, demonstrating that these lesions require continuous expression of the oncogene. Secondly, they demonstrated that re-expression of the SV40 T antigen three weeks after it had been inactivated let to the re-emergence of hyperplastic lesions, demonstrating that the residual transformed cells could serve as a seedbed for the growth of future hyperplastic lesions. Lastly, they demonstrated that prolonged expression of the SV40 T antigen for over 28 weeks led to a point of no return, whereby hyperplastic lesions persisted even after inactivating the expression of the SV40 T antigen (Ewald et al., 1996). That malignant lesions progressed to a point where they no longer required sustained expression of the driving oncogene for growth and survival would not prove to be the norm. Instead, a series of groundbreaking studies led by the groups of Ronald DePinho, Michael Bishop and Dean Felsher experimentally demonstrated the concept of oncogene addiction *in vivo*, whereby tumor cells become addicted to sustained expression of a particular oncogene and become exquisitely sensitive to removal of this oncogene (Weinstein et al., 2002, Sharma and Settleman 2007). DePinho and colleagues employed bi-transgenic mice expressing oncogenic $HRAS^{G12V}$ under the control of a tetO promoter as well as rtTA under the control of tyrosinase regulatory

elements to restrict inducible expression of the oncogene to melanocytes (Chin et al., 1999). In this setting, expression of the *HRAS*^{G12V} oncogene was induced upon doxycycline administration, which, in the genetic background of homozygous *Ink4a* mutation, led to the development of melanomas. Remarkably, inactivation of the *HRAS*^{G12V} oncogene in established melanomas led to their rapid regression, thereby validating the model of oncogene addiction in the context of RAS-driven melanomas. Similar to the study from Hennighausen's group, re-activation of *HRAS*^{G12V} in microscopic melanoma lesions was sufficient to induce the re-emergence of *bona fide* melanoma tumors. In a contemporaneous study utilizing bi-transgenic mice expressing the *MYC* proto-oncogene under the control of a tetO promoter as well as tTA under the control of the immunoglobulin heavy chain enhancer (EμSR-tTA) to restrict expression of *MYC* to cells of the hematopoietic lineage, Dean Felsher and Michael Bishop demonstrated that established *MYC*-driven T cell lymphomas and acute myeloid leukemias were also exquisitely sensitive to oncogene inactivation (Felsher and Bishop 1999). Interestingly, disease regression was accompanied by a combination of cell cycle arrest, differentiation, and apoptosis. In a subsequent study using the same mouse model, the Felsher group demonstrated that brief inactivation of *MYC* in established osteosarcomas also led to sustained tumor regression and cellular differentiation (Jain et al., 2002). Strikingly, re-activation of *MYC* in differentiated, mature osteocytes led to massive cell death, which led the authors to postulate that *MYC* inactivation in established malignancies leads to a series of irreversible epigenetic changes that prevent subsequent malignant transformation by *MYC* (Jain et al., 2002).

In addition to proving the concept of oncogene addiction, these landmark studies unequivocally cemented the power of tetracycline-regulatable systems for dissecting cancer *in vivo*. Not surprisingly, multiple researchers over the years have developed over 100 transgenic strains for studying many different tumor types based on tetracycline-regulatable systems (for a recent comprehensive review by Lewis Chodosh and colleagues, see Yeh et al., 2014).

B. Hormone-regulatable systems

Another powerful approach for inducibly and reversibly controlling the function of a protein of interest is through the use of hormone-regulatable systems, particularly those that are based on estrogen receptor gene fusions (**Figure 2C**) (for a contemporaneous review discussing some of the first applications of this technology, see Picard 1993). Initially described by the group of Keith Yamamoto in 1988 (Picard et al., 1988), estrogen receptor gene fusion technology exploits the capacity of the steroid binding domain of this receptor to act independently of the DNA binding domain of the entire intact estrogen receptor. Therefore, a fusion between a gene of interest and the steroid binding domain of the estrogen receptor places the protein of interest under the control of exogenously provided hormone (Picard et al., 1988). In the absence of hormone, the protein is kept sequestered in the cytoplasm via direct physical interaction of the steroid binding domain with heat shock proteins, particularly heat shock protein 90 (Hsp90). In the presence of hormone, this physical interaction is abolished and the protein of interest is therefore “induced” (Picard et al., 1988). After Picard’s initial demonstration of a functional estrogen receptor fusion with the adenovirus E1A gene product, several

groups went on to develop hormone-regulatable fusions of many oncogenes, such as *Myc* (Eilers et al., 1989, Littlewood et al., 1995), *Myb* (Burk et al., 1991), *Rel* (Boehmelt et al., 1992), *Abl* (Jackson et al., 1993), and *Raf1* (Samuels et al., 1993), among many others. Notably, Gerard Evan's group went on to generate an estrogen receptor fusion gene using *MYC* (Littlewood et al., 1995) and subsequently used it to develop a transgenic mouse that expressed the Myc-ER^{TM} fusion gene under the control of involucrin regulatory sequences, thereby limiting its expression to the mouse epidermis (specifically to the suprabasal keratinocytes) (Pelengaris et al., 1999). Using this inducible and reversible hormone-regulatable transgenic mouse model, they went on to demonstrate that *MYC* activation in the mouse epidermis led to epidermal hyperplasias that closely resembled pre-cancerous lesions associated with squamous cell carcinomas. Strikingly, inactivation of the Myc-ER^{TM} transgene through withdrawal of the estrogen analog, 4-hydroxy-tamoxifen (4-OHT), completely reversed these malignant phenotypes, demonstrating that these hyperplasias are addicted to the *MYC* oncogene (Pelengaris et al., 1999).

While both tetracycline-regulatable and estrogen receptor-regulatable technologies were incredibly powerful for studying oncogenes, they were not readily applicable to the study of tumor suppressor genes, at least during this time period. The introduction of genetic technologies that allowed for conditional deletion of tumor suppressor genes, and the combination of these technologies with tetracycline-regulatable and estrogen receptor-regulatable systems marked another major milestone in the field of mouse modeling. These conditional systems are described below.

C. Conditional systems based on site-specific recombinases

The groundbreaking work of Nat Sternberg and Daniel Hamilton in the early 1980s on P1 bacteriophage genetics established the foundation for the development of conditional systems based on SSRs (Sternberg and Hamilton 1981a, Sternberg et al., 1981b, Sternberg et al., 1981c). In a series of classical experiments aimed at tackling the scientific conundrum of why the genetic map of P1 was linear, even though it was well known at the time that its genome was circular (Ikeda and Tomizawa 1968), Sternberg and colleagues discovered a site-specific recombination system that allowed P1 to integrate into the genome of the bacterial host as part of its life cycle. This SSR system - called the Cre/loxP system - was found to be composed of at least two parts: a 34bp locus of crossing over [x] in P1 (loxP) sequence and a locus that causes recombination (Cre). In this system, any sequence of DNA that is flanked by loxP sites positioned in direct orientation (➡ sequence ➡) or in an inverted orientation (➡ sequence ⬅) in *cis* can be excised or inverted via the action of the SSR Cre, respectively (see **Figure 3** for multiple examples of these configurations). Building upon the pioneering work of Sternberg and colleagues, Brian Sauer and Nancy Henderson adapted the phage Cre/loxP system to catalyze Cre-mediated DNA excision or inversion of a series of genetic elements introduced via transfection into mammalian cells stably expressing Cre recombinase under the control of metallothionein regulatory elements (Sauer and Henderson 1988). Subsequently, two studies reported the generation of transgenic mice stably expressing Cre recombinase and Cre-responsive transgenes. In one of these studies, Marth and colleagues demonstrated that a transgene containing the *β*-galactosidase gene flanked by a pair of loxP sites positioned in direct orientation

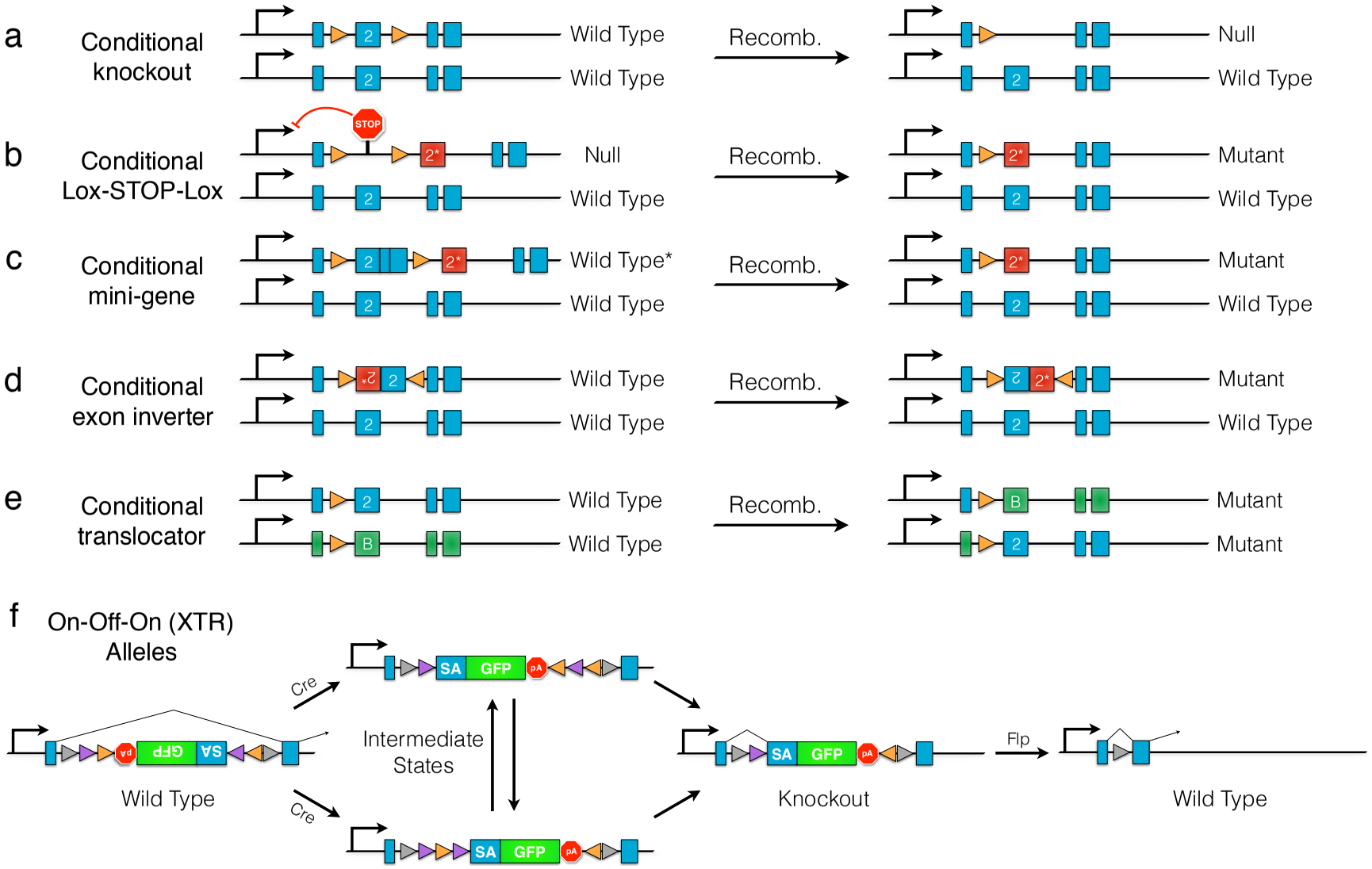

a
Conditional knockout
Wild Type
Wild Type
Recomb.
Null
Wild Type
b
Conditional Lox-STOP-Lox
STOP
Null
Wild Type
Recomb.
Mutant
Wild Type
c
Conditional mini-gene
Wild Type*
Wild Type
Recomb.
Mutant
Wild Type
d
Conditional exon inverter
Wild Type
Wild Type
Recomb.
Mutant
Wild Type
e
Conditional translocator
Wild Type
Wild Type
Recomb.
Mutant
Mutant
f
On-Off-On (XTR) Alleles
Cre
Cre
SA
GFP
pA
Intermediate States
Wild Type
Knockout
Flp
Wild Type

Figure 3: Conditional alleles based on SSRs. (a) ***Conditional knockout / floxed alleles*** - in this design, the sequence of the entire gene of interest or a critical region of the gene of interest is flanked by loxP sites positioned in direct orientation (➡ Gene ➡), such that Cre-mediated recombination leads to deletion of the floxed sequence. **(b)** ***Lox-STOP-Lox alleles*** - in this design, one or both alleles are kept silenced via the incorporation of a Lox-STOP-Lox cassette in a region upstream of the gene of interest, typically between the promoter and coding sequence, such that Cre-mediated recombination leads to the activation of the allele (s). **(c)** ***Conditional mini-gene alleles*** - in this design, one or both alleles are modified such that they contain a Lox-mini-gene-Lox cassette that expresses a cDNA encoding the wild type version of the allele (s) before Cre-mediated recombination. This cassette gets deleted upon Cre-mediated recombination, which results in the expression of the mutant allele (s) of interest. **(d)** ***Conditional exon inverter alleles*** - in this design, a cassette that contains an inverted mutant exon upstream of the wild type exon is flanked by a pair of heterospecific loxP sites that are positioned in a head-to-head orientation (➡ cassette ⬅), such that Cre-mediated recombination leads to the inversion of the cassette and subsequent expression of the mutant exon accompanied by loss of expression of the wild type exon. **(e)** ***Conditional translocator alleles*** - in this design, individual loxP sites are inserted into different chromosomes, which results in a translocation event upon Cre-mediated recombination between the loxP sites. **(f)** ***On-Off-On alleles*** - in this clever design, a Cre-mediated recombination event (between purple loxP sites *or* orange loxP sites, but not between purple *and* orange loxP sites) induces a stable inversion event that traps the endogenous gene via the use of a gene-trap GFP reporter cassette, leading to a stable (but reversible) gene knockout. Subsequently, Flp-mediated recombination between Frt sites (colored in grey) leads to the excision of the gene-trap GFP reporter cassette, thereby restoring the expression of the endogenous gene. SA: splice acceptor. pA: polyadenylation sequence. Figures (a) through (e) were adapted from Frese & Tuveson (2007). Figure (f) was adapted from Robles-Oteiza et al., 2015.

(➡ Gene ➡) (**Figure 3A**) could be efficiently excised *in vivo* upon expression of Cre recombinase from a second transgene (Orban et al., 1992). In a contemporaneous study, Westphal and colleagues introduced the important technological concept of a "Lox-STOP-Lox" (LSL) cassette, in which a sequence corresponding to (1) the 3' portion of the *HIS3* gene from yeast, (2) an SV40 polyadenylation sequence and (3) a false translation initiation codon followed by a 5' splice donor site, is placed between two loxP sites positioned in direct (➡ STOP ➡) orientation (Lakso et al., 1992). Placement of a LSL cassette upstream of a gene potently blocks transcription and translation of the gene (**Figure 3B**). Upon exposure to Cre recombinase, recombination between these direct loxP sites leads to excision of the STOP cassette and thus gene activation (**Figure 3B**) (Lakso et al., 1992). In this study, the authors demonstrated that Cre-inducible activation of the SV40 T antigen in the lens of mice invariably led to the development of lens tumors, a finding that opened the door for the generation of mice expressing Cre-inducible oncogenes (Lakso et al., 1992). A subsequent study by Hua Gu, Klaus Rajewsky and colleagues took this technology a step further by combining gene targeting via homologous recombination with Cre/loxP technology to generate *endogenous* alleles flanked by loxP sites (commonly referred to as "floxed alleles"), which would get efficiently deleted upon exposure to Cre recombinase (Gu et al., 1994). In this landmark study, the authors demonstrated that a floxed allele of the essential gene encoding for the DNA polymerase β (*pol* β) could be selectively deleted in the T-cell compartment via the incorporation of a transgene encoding Cre recombinase under the control of the T-cell-specific *Lck* promoter. Therefore, this study achieved two very important goals for the mouse modeling community: (1) floxed alleles of endogenous

genes could be efficiently deleted by Cre recombinase and (2) this could be performed in a cell type-specific or tissue-specific manner by restricting expression of the recombinase to a given cell lineage (Gu et al., 1994).

Just like the development of transgenic technologies triggered an explosion in the generation of several dozens of transgenic mouse models of cancer, the development of conditional systems based on SSRs led to a rapid expansion of the catalog of sophisticated conditional GEMMs of cancer (for a series of excellent contemporaneous reviews that synthesize the thinking during these years, see Rossant and Nagy 1995, Lobe and Nagy 1998, Cohen-Tannoudji and Babinet 1998, Wu and Pandolfi 2001, Wong and Witte 2001, Jonkers and Berns 2002, Van Dyke and Jacks 2002 and Bernardi et al., 2002). Cancer researchers could now tackle some of the most crippling limitations of pre-existing technologies. For example, they could finally get around the embryonic lethality associated with germline deletion of some of the most relevant cancer genes, including the *Apc* (Fodde et al., 1994, Oshima et al., 1995, Moser et al., 1995), *Pten* (Di Cristofano et al., 1998, Stambolic et al., 1998), and *Rb* (Lee et al., 1992, Jacks et al., 1992, Clarke et al., 1992) tumor suppressor genes, as well as the *Egfr* (Threadgill et al., 1995), *Kras* (Johnson et al., 1997), *Mdm2* (Montes de Oca Luna et al., 1995, Jones et al., 1995), *Myc* (Davis et al., 1993), and *Pik3ca* (Bi et al., 1999) oncogenes, among many others. In addition to circumventing the embryonic lethality problem, the availability of a suite of tissue-specific promoters and multiple inducible technologies, and the ability to simultaneously combine all of these technologies allowed researchers to generate novel alleles that restricted the expression of the SSR to specific cell types in a tightly controlled inducible (and reversible) manner. In the next

section, I will briefly describe some of the most innovating mouse models of cancer that have been developed during the years using conditional systems mainly based on the use of SSRs, such as Cre and Flp (Andrews et al., 1985) recombinases.

D. GEMMs of cancer based on SSR technologies

Many of the most innovating GEMMs of cancer that have been developed so far have employed a combination of multiple experimental strategies based on the simultaneous application of tetracycline-regulatable systems (**Figure 1**), hormone-regulatable systems (**Figure 1**) and single or multiple SSRs (**Figure 3**), sometimes even combining all of the above with stable or inducible RNAi technologies (**Figure 5**) (RNAi-based technologies for modeling cancer *in vivo* are discussed in part E of this section of the introduction), as well as emerging genome engineering technologies that will be discussed in **Part II** of the introduction. Most of the strategies that employ increasingly sophisticated designs based on the use of single or multiple SSRs are summarized in **Figure 3** and will be briefly described below exclusively in the context of GEMMs of cancer.

i. Lox-STOP-Lox alleles and floxed alleles

Several dozens of GEMMs harboring conditional alleles of tumor suppressor genes and oncogenes have been generated throughout the years, and the list keeps expanding. In fact, the scientific desire and demand for generating conditional alleles for almost every gene in the mouse genome has been so high that it has led to the creation of multiple international consortia whose sole purpose is the generation of gene-targeting vectors as well as pre-targeted ES cells and GEMM-derived ES cells harboring constitutive or

conditional alterations in almost every coding gene (including microRNAs) in the mouse genome (Adams et al., 2004, Skarnes et al., 2011, Dow and Lowe 2012). In its most basic design, conditional alleles can refer to either (1) floxed alleles, in which the sequence of the entire gene of interest or a critical region of the gene of interest is flanked by loxP sites positioned in direct orientation (➡ gene ➡) (**Figure 3A**) or (2) LSL alleles, in which the gene is kept silenced via the incorporation of a Lox-STOP-Lox cassette in a region upstream of the gene of interest, typically between the promoter and coding sequence (**Figure 3B**). Several well described GEMMs of cancer that are directly relevant to this thesis have been generated using these two strategies. For example, the group of Anton Berns generated a conditional knockout allele of the *p53* tumor suppressor gene via the incorporation of loxP sites in introns 1 and 10 of the gene, thereby generating a *p53* floxed allele ($p53^{fl/fl}$) (Marino et al., 2000, Jonkers et al., 2001). In this scenario, Cre-mediated recombination between these two loxP sites leads to the efficient deletion of exons 2-10 from the *p53* gene, thereby producing a full knockout allele in a conditional manner. These mice have played an invaluable role in modern cancer research, as demonstrated by numerous publications over the last 14 years, ranging from fundamental studies in lung cancer initiation, progression, metastasis and therapeutic resistance (Meuwissen et al., 2003, Meylan et al., 2009, Oliver et al., 2010, Winslow et al., 2011, Xue et al., 2011, Sutherland et al., 2014, among many others), as well as multiple other studies in pancreatic cancer and breast cancer, among several other cancer types (Gidekel Friedlander et al., 2009, Chiou et al., 2015, among many others). Another central GEMM in this thesis is the Lox-STOP-Lox-Kras allele (hereafter referred to as the $Kras^{LSL\text{-}G12D/+}$ allele) which was developed by the

Jacks group in 2001 (Jackson et al., 2001). This model harbors a LSL cassette upstream of exon 1 of *Kras,* which was simultaneously engineered to contain the oncogenic glycine (G) to aspartic acid (D) mutation at position 12 (G12D). In this scenario, Cre-mediated recombination leads to excision of the STOP cassette and subsequent induction of expression of the oncogenic $Kras^{G12D}$ allele. Notably, this allele offers several advantages over transgenic oncogenic alleles (such as the previously described tetO-driven $HRAS^{G12V}$ allele) or latent oncogenic alleles (**Figure 1C**). First, the researcher is able to control the time of initiation and the number of initiating events by limiting the dose of Cre recombinase administered (either inducible Cre being expressed from a transgene or an endogenous locus or exogenously-delivered Cre recombinase embedded within the backbone of an adenovirus or lentivirus) (Jackson et al., 2001, DuPage et al., 2009, Chiou et al., 2015). Secondly, mice harboring this allele can be readily crossed to any of the pre-existing strains of mice expressing Cre recombinase in a constitutive or inducible tissue-specific fashion, thereby allowing for the interrogation of the consequences of expression of oncogenic *Kras* in a multitude of different tissue and tumor types. Thirdly, these mice can be readily crossed to any of the numerous strains of mice harboring additional floxed or Cre-inducible alleles to tackle an innumerable number of pressing questions in cancer biology. By employing the latter strategy, the group of Tyler Jacks developed the $Kras^{LSL\text{-}G12D/+;}$ $p53^{flox/flox}$ (KP) mouse model of lung adenocarcinoma (Jackson et al., 2005) (**Figure 4**), in which lung tumors are initiated via intranasal or intratracheal administration of viral vectors expressing Cre recombinase (Jackson et al., 2005, DuPage et al., 2009), which concomitantly activates the oncogenic $Kras^{G12D}$ allele and deletes both copies of the *p53* tumor suppressor gene

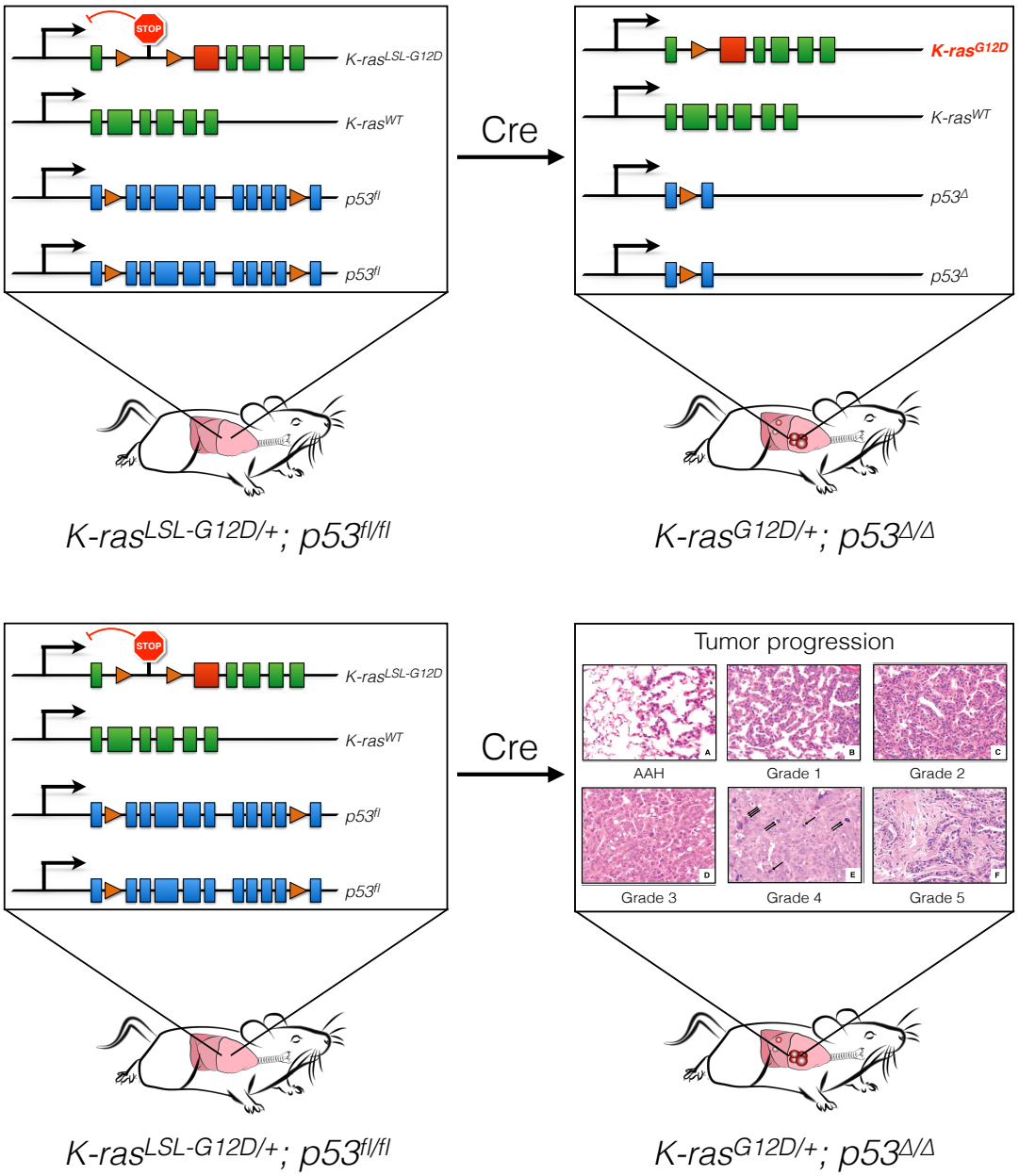


Figure 4: The KP mouse model of lung adenocarcinoma. In this model, a LSL cassette inserted in the first intron of the endogenous *Kras* gene prevents gene transcription. Intratracheal delivery of an adenovirus or lentivirus expressing Cre recombinase triggers activation of oncogenic *Kras* and concomitant deletion of both alleles of the *p53* tumor suppressor gene. Importantly, this mouse model closely recapitulates the histopathological progression of human lung adenocarcinoma. Histology pictures shown in the bottom right were adapted from Jackson et al., 2005.

in the murine lung epithelium. Notably, this model closely recapitulates the human counterpart, showing initial development of lung hyperplasias followed by the formation of adenomas, which subsequently progresses to lung adenocarcinomas and (in a subset of the animals) local and distant metastases (Jackson et al., 2005) (**Figure 4B**). The impact that the *Kras*$^{LSL\text{-}G12D/+}$ (K) and *Kras*$^{LSL\text{-}G12D/+}$*; p53*$^{flox/flox}$ (KP) mouse models of cancer have had in modern cancer research is enormous - from discovering fundamental regulators metastasis (Winslow et al., 2011) to informing current and future clinical trials through their use in co-clinical trials (Engelman et al., 2008, Singh et al., 2010, Chen et al., 2012).

ii. Conditional mini-gene alleles

One potential disadvantage of LSL alleles is the fact that GEMMs harboring such alleles are inherently heterozygous for the gene of interest, which could lead to confounding results due to haploinsufficiency. A notable example involves *Kras* itself - at least four groups have demonstrated that wild type *Kras* can exhibit tumor suppressive properties in the context of lung adenocarcinomas, teratomas, colorectal adenomas and T-cell leukemias (Zhang et al., 2001, James et al., 2003, Luo et al., 2014, Staffas et al., 2015). To address this potential shortcoming, several groups have adopted a conditional mini-gene strategy (**Figure 3C**), which features a Lox-mini-gene-Lox cassette that expresses a cDNA encoding the wild type version of the allele before Cre-mediated recombination (Mercer et al., 2005, Forster et al., 2005, Bayascas et al., 2006, Dankort et al., 2007, Dhomen et al., 2009), which gets deleted upon Cre-mediated recombination of the Lox-mini-gene-Lox cassette. Two of the best examples that demonstrate the power of this

approach were reported by the groups of Richard Marais in 2005 and Martin McMahon in 2007, who generated the widely used *LSL-Braf*^{V600E} and *BRaf*^{CA} Cre-inducible alleles, respectively (Dankort et al., 2007, Mercer et al., 2005).

iii. Conditional exon inverter alleles

Another elegant strategy that addresses this potential shortcoming makes use of a conditional exon inverter strategy (**Figure 3D**), which features a cassette that contains an inverted mutant exon upstream of the wild type exon (Bayascas et al., 2006). Importantly, this cassette is flanked by a pair of engineered heterospecific loxP sites (which are mutant loxP sites that favor recombination in one direction) (Langer et al., 2002) that are positioned in a head-to-head orientation (➡ cassette ⬅), such that Cre-mediated recombination leads to the inversion of the cassette and subsequent expression of the mutant exon accompanied by loss of expression of the wild type exon (**Figure 3D**) (Bayascas et al., 2006).

iv. Conditional translocator and inverter alleles

A number of cancers are characterized by the presence of large chromosomal aberrations, including translocations, inversions and large chromosomal deletions (Mitelman et al., 2007, Solimini et al., 2012, Davoli et al., 2013). Several groups have reported the successful adaptation of the Cre/loxP system to model these large chromosomal rearrangements in an inducible manner (see **Figure 3E** for an example of a strategy for modeling translocations using this system). Albeit technically challenging, this approach was successfully used by the group of Terence Rabbitts to model the *EWS-ERG* fusion (Forster et al., 2005). Notably, they employed a sophisticated

combination of genetic tricks to achieve a chromosomal inversion involving *ERG* by placing loxP sites in a head-to-head orientation (➡ sequence ⬅) surrounding an intronic, inverted *ERG* cassette that, upon Cre-mediated inversion, got fused to endogenous mouse *Ews* regulatory sequences (Forster et al., 2005). Other groups have also elegantly demonstrated the general applicability of Cre/loxP-based approaches for modeling chromosomal translocations and other large chromosomal rearrangements (Smith et al., 1995, Van Deursen et al., 1995, Buchholz et al., 2000, Collins et al., 2000, Forster et al., 2003, Bagchi et al., 2007, Barlow et al., 2009, Wong et al., 2015) (**Figure 3E**). As I will extensively discuss in **Part II** of the introduction, the recent emergence of several genome engineering technologies, such as the CRISPR-Cas9 system, have greatly facilitated the technical aspects associated with generating these and other complex genetic events.

v. Lox-STOP-Lox restorable alleles and On-Off-On alleles

LSL alleles (**Figure 3B**) have also been cleverly employed for studying the consequences of reactivation of the *p53* tumor suppressor gene in established tumors. Tyler Jacks and colleagues first demonstrated the power of this approach by generating a mouse model expressing a *LSL-p53* allele in the germline (Ventura et al., 2007). Homozygous mutant mice ($p53^{LSL/LSL}$) are phenotypically identical to the homozygous knockout mice ($p53^{-/-}$) described by Lawrence Donehower and Tyler Jacks in the early 1990s. However, upon induction of Cre recombinase (which in this case was expressed from the ubiquitously expressed *Rosa26* locus (Zambrowicz et al., 1997) as an estrogen receptor fusion protein called Cre-ER^{T2}) and excision of the LSL cassette, p53

transcription is restored. Utilizing this mouse model, Tyler Jacks and colleagues demonstrated that established tumors that spontaneously arise in the context of p53 deficiency (mostly soft tissue sarcomas and lymphomas) (Donehower et al., 1992, Jacks et al., 1994) remain exquisitely sensitive to the reactivation of p53 (Ventura et al., 2007). Notably, two contemporaneous studies by the groups of Gerard Evan and Scott Lowe reached the same conclusion using estrogen receptor-regulatable systems and inducible-reversible RNAi technologies, respectively (Martins et al., 2006, Xue et al., 2007). In a subsequent study, Tyler Jacks and colleagues crossed the p53 restorable mouse to the *Kras^{LA2}* mouse model (**Figure 1C**), which bears a latent allele of oncogenic *Kras^{G12D}* that gets spontaneously activated *in vivo* (Johnson et al., 2001), for the purposes of studying the effects of p53 reactivation in established lung tumors (Feldser et al., 2010). Surprisingly, they found that the tumor suppressive effects of p53 reactivation in established lung tumors are stage-specific, whereby p53 triggers the elimination of highly advanced lesions but spares low grade lesions. Mechanistically, they discovered that high grade lesions harbor hyperactive MAPK signaling, which in turn leads to potent induction of the p19^{Arf} tumor suppressor gene, leading to stabilization and activation of the p53 tumor suppressor (Feldser et al., 2010). Importantly, a contemporaneous study by Gerard Evan and colleagues using an estrogen receptor-regulatable allele of *p53* reached the same conclusion (Junttila et al., 2010). Albeit undoubtedly powerful, the use of LSL alleles for studies involving tumor suppressor gene restoration are limited to a minority of cases where germline knockout of the gene does not lead to embryonic lethality, such as *p53* (Donehower et al., 1992, Jacks et al., 1994), *Ink4a/Arf* (Serrano et al., 1996), *p19^{Arf}* (Kamijo et al., 1997), *Ink4a*

(Sharpless et al., 2001), and *Atm* (Barlow et al., 1996, Xu et al., 1996, Elson et al., 1996). Moreover, mice harboring germline non-lethal knockout alleles of these tumor suppressor genes are invariably highly tumor prone, which severely limits the utility of these models for tackling questions related to advanced stages of tumor progression, such as metastasis. To overcome these and other limitations, the group of David Feldser recently reported the generation of a novel "On-Off-On" system that allows for SSR-based conditional and reversible gene regulation *in vivo* (Robles-Oteiza et al., 2015) (**Figure 3F**). This system, called *XTR* for expressed (XTR), Trapped (TR) and Restored (R), allows for simultaneous conditional inactivation of an endogenous gene and fluorescent readout of transcription of the endogenous locus in a reversible manner. Specifically, Cre-mediated recombination induces a stable inversion event that traps the endogenous gene via the use of a gene-trap GFP reporter cassette, leading to a stable (but reversible) gene knockout. Subsequently, Flp-mediated recombination leads to the excision of the gene-trap GFP reporter cassette, thereby restoring the expression of the endogenous gene (Robles-Oteiza et al., 2015) (**Figure 3F**). Notably, this approach will enable cancer researchers to study the effects of tumor suppressor restoration in established tumors *in vivo* by effectively circumventing the embryonic lethal effects caused by germline knockout of the vast majority of tumor suppressor genes, such as *Rb* (Lee et al., 1992, Jacks et al., 1992, Clarke et al., 1992), *Pten* (Di Cristofano et al., 1998, Stambolic et al., 1998), and *Apc* (Fodde et al., 1994, Oshima et al., 1995, Moser et al., 1995).

E. GEMMs of cancer based on inducible *in vivo* RNAi

Another powerful alternative to the use of SSR-based conditional systems for modeling loss of function mutations *in vivo* is RNA interference (RNAi). The phenomenon of RNAi was originally described by Andrew Fire, Craig Mello, and colleagues in their seminal, nobel-prize winning 1998 letter to *Nature* (Fire et al., 1998). The serendipitous discovery that double stranded RNA (dsRNA) molecules could potently inhibit gene expression through a sequence-specific mechanism immediately triggered a revolution that transformed nearly every single aspect of biological research. Multiple groups over the next several years identified the core components of the RNAi pathway and systematically dissected the mechanisms mediating potent, sequence-specific genetic interference (for comprehensive and contemporaneous reviews, see Hammond et al., 2001, Hannon et al., 2002, McManus and Sharp 2002, Paddison and Hannon, 2002, Hannon and Rossi 2004, Mello and Conte 2004, and Meister and Tuschl 2004). The RNAi pathway functions through a highly conserved mechanism, in which dsRNA molecules trigger sequence-specific post-transcriptional gene suppression. The mechanics of the endogenous RNAi pathway in animals involve the initial RNA polymerase II-dependent transcription of a long polyadenylated precursor microRNA molecule (pri-miRNA) that gets cleaved into a mature, 20-25nt long miRNA molecule through a series of highly conserved steps (reviewed in Bartel 2004 and Bartel 2009). First, the Drosha/DGCR8 complex cleaves the nuclear pri-miRNA molecule, generating a pre-miRNA molecule that gets exported to the cytoplasm via the action of the Ran-GTP-dependent dsRNA -binding protein Exportin 5 (Lee et al., 2003, Gregory et al., 2004, Denli et al., 2004, Yi et al., 2003, Lund et al., 2004, Bohnsack et al., 2004). This

pre-miRNA molecule is subsequently cleaved by the endoribonuclease Dicer, which leads to the generation of the 20-25nt long mature miRNA molecule (Bernstein et al., 2001, Hutvágner et al., 2001, Hutvágner and Zamore 2002). This mature miRNA molecule is then unwound into each of its respective single strands and the guide strand subsequently associates with the Argonaute family of proteins to produce the RNA-induced silencing complex (RISC) (Liu et al., 2004, Rand et al., 2005). Using the guide strand to scan the transcriptome of the cell, the RISC then pairs with complementary messenger RNAs (mRNAs) via Watson-Crick base pairing to direct potent post-transcriptional silencing via transcript degradation or destabilization (Meister et al., 2004, Yekta et al., 2004, Guo et al., 2010) or translation inhibition (Pillai et al., 2005, Humphreys et al., 2005, Bhattacharyya et al., 2006, Kiriakidou et al., 2007) (also reviewed in Filipowicz et al., 2008 and in Huntzinger and Izaurralde 2011). Researchers over the last 10-15 years have exploited the capacity of other types of dsRNA molecules, such as synthetic short hairpin RNA molecules (shRNAs) that closely resemble pri-miRNAs, to undergo similar processing by Drosha/DGCR8 and subsequent cleavage by Dicer to generate mature ~22nt dsRNAs that are subsequently incorporated into RISC to mediate potent mRNA silencing (Hannon and Rossi et al., 2004).

A comprehensive discussion of the applications of RNAi technologies for modern cancer research is beyond the scope of the introduction of this thesis. For this reason, I would like to point the reader to multiple excellent reviews that have been published in the last 5 years, in which most of the applications up to date have been extensively discussed, particularly some of the ones that relate to high-throughput RNAi screens and RNAi-

based GEMMs of cancer: Boehm and Hahn 2011, Livshits and Lowe 2013, Mohr et al., 2014, Fellmann and Lowe 2014, Crotty and Pipkin 2015, and Rytlewski and Beronja 2015. Nevertheless, it is imperative for me to briefly discuss a few pioneering examples that illustrate how modern RNAi technologies have revolutionized the way cancer researchers can model and study cancer *in vivo*.

i. Using transplant-based approaches and transgenic mice

The group of Scott Lowe has pioneered some of the most sophisticated RNAi technologies for modeling cancer *in vivo*. Using the well established RNA polymerase II-driven miR30 system, in which potent synthetic RNAi trigger sequences are embedded within the context of the endogenous miR30 backbone (Zeng et al., 2002, Zeng et al., 2005, Stegmeier et al., 2005), they demonstrated that shRNA-mediated suppression of *p53* in hematopoietic stem and progenitor cells (HSPCs) derived from the Eμ-Myc mouse model dramatically accelerated the onset of lymphomagenesis *in vivo* upon transplantation into syngeneic recipient mice (Hemann et al., 2003). In subsequent studies, they demonstrated that these shRNA cassettes could be placed under the control of tetracycline response element (TRE) promoters (**Figure 2A-B**), thereby allowing for tetracycline (or doxycycline) regulatable expression of the shRNA molecule in both Tet-On and Tet-Off configurations (Dickins et al., 2005, Dickins et al., 2007, Xue et al., 2007, Zuber et al., 2011a, Zuber et al., 2011b, and others). One notable example that elegantly demonstrates the power of these approaches comes from a study published in 2007 by the Lowe group (Xue et al., 2007). In this study, they utilized a Tet-Off system to turn off the expression of an shRNA targeting p53 *in vivo* to demonstrate

that p53 reactivation in established liver carcinomas triggers a potent cellular senescence program accompanied by extensive tumor regression and immune clearance mediated by the innate immune system (Xue et al., 2007). In addition to the use of RNAi technologies for modeling loss of function of individual genes *in vivo*, the groups of Scott Lowe and Michael Hemann pioneered the adaptation of RNAi technologies for performing small- to large-scale pooled shRNA screens *in vivo* (Burgess et al., 2008, Zender et al., 2008, Meacham et al., 2009, Bric et al., 2009, Zuber et al., 2011b, Scuoppo et al., 2012, Meacham et al., 2015). These innovating approaches - which were all initially based on transplantable approaches - have had a tremendous impact in the field of functional cancer genomics, as exemplified by the implementation of similar strategies by other groups for performing *in vivo* screens in many other experimental settings (Possemato et al., 2011, Gargiulo et al., 2013, Miller et al., 2013, Järås et al., 2014, and others).

Beyond the development and implementation of sophisticated RNAi technologies for performing loss of function studies *in vivo* using transplantable approaches, several groups have generated various types of transgenic mice expressing constitutive, Cre/loxP-based conditional, or tetracycline-inducible shRNA cassettes. Multiple groups initially reported the generation of transgenic mice expressing constitutive shRNAs (Hasuwa et al., 2002, Carmell et al., 2003, Kunath et al., 2003, Tiscornia et al., 2003). Despite the fact that these proof-of-principle studies demonstrated that this approach was experimentally feasible, these constitutive strategies were not widely adopted by the scientific community due to the desire to manipulate shRNA expression at will. One notable example of a system that allowed for inducible shRNA expression was the

pSico and pSicoR pair of Cre-recombinase-regulatable lentiviral vectors developed by the group of Tyler Jacks (Ventura et al., 2004). The pSico and pSicoR vectors allowed for Cre-mediated induction or repression of shRNA expression, respectively. They demonstrated the utility of this technology *in vivo* by generating transgenic mice expressing a Cre-inducible shRNA targeting the T lymphocyte cell surface marker CD8 (Ventura et al., 2004). Another notable contemporaneous example was the generation of transgenic mice expressing tetracycline-regulatable shRNAs by the group of Scott Lowe (Dickins et al., 2007). In this study, Lowe and colleagues generated Tet-On and Tet-Off transgenic mice expressing a potent miR-30 based shRNA targeting the *p53* tumor suppressor gene. They further demonstrated potent and reversible induction and repression of shRNA expression in established malignancies *in vivo* by employing rtTA and tTA alleles, respectively (Dickins et al., 2007). Despite the fact that transgenic RNAi mice and cell lines derived from them are very powerful for studying the effects of loss of function of genes of interest using *in vivo* and *ex vivo* approaches, they invariably require the generation of a mouse and extensive mouse husbandry. To overcome some of these experimental hurdles, several groups adapted lentiviral- and transposon-based approaches to directly deliver shRNAs to tissues of interest *in vivo*. Below, I will describe a few key examples that illustrate the power of these approaches, focusing exclusively on lentiviral systems.

ii. Using direct delivery of lentiviruses to somatic tissues

Multiple groups have reported the successful adaptation of well established lentiviral approaches to deliver potent shRNAs directly to tissues of interest *in vivo*. One of the

first examples demonstrating the power of this approach utilized stereotactic injections to directly deliver shRNA-expressing lentiviruses to the adult mouse brain, achieving effective suppression of an exogenous reporter gene that was co-delivered using a second lentivirus (Van den Haute et al., 2003). Using a conceptually similar approach, Patrick Aebischer and colleagues demonstrated that potent suppression of a mutant allele of superoxide dismutase (*SOD1^{G93A}*) *in vivo* ameliorated the onset and progression of amyotrophic lateral sclerosis (ALS) in a *SOD1^{G93A}* transgenic mouse model of ALS (Raoul et al., 2005). Another pioneering study was published by the group of Elaine Fuchs in 2010, in which they demonstrated that ultrasound-guided delivery of bi-functional lentiviruses co-expressing shRNAs and cDNAs (encoding fluorescent reporters or Cre recombinase) *in utero* could be utilized to achieve potent loss of function in the murine epidermis (Beronja et al., 2010, Beronja and Fuchs 2013). Notably, Fuchs and colleagues extended this approach to perform large-scale RNAi screens *in vivo* to uncover novel regulators of tumor growth in epidermal hyperplasias and squamous cell carcinomas (Beronja et al., 2013, Schramek et al., 2014, Rytlewski and Beronja 2015). Similarly, by leveraging their previous expertise in lentiviral approaches for modeling cancer *in vivo* (Marumoto et al., 2009), Inder Verma and colleagues combined Cre/loxP technology with *in vivo* RNAi to rapidly generate novel mouse models of Kras-driven lung cancer via intratracheal delivery of bi-functional lentiviruses co-expressing Cre recombinase and shRNAs (Xia et al., 2012). In subsequent studies, Verma and colleagues adapted this approach for performing large-scale RNAi screens *in vivo* to uncover novel lung cancer tumor suppressor genes (Yeddula et al., 2015). Although powerful, these lentiviral-based approaches (as well as

plasmid-based systems encoding sleeping beauty transposons that co-express shRNAs and cDNAs of interest - for notable examples, see Heggestad et al., 2004, Fletcher et al., 2010, Wuestefeld et al., 2013, Rudalska et al., 2014, Tschaharganeh et al., 2014) are invariably limited by the capacity of these constructs to directly access the tissue of interest. Therefore, these approaches have been limited to readily accessible tissues, such as the liver (Rudalska et al., 2014), lung (Yeddula et al., 2015), brain (Marumoto et al., 2009), and pancreas (Taniguchi et al., 2003, Houbracken et al., 2012, Chiou et al., 2015), among others. To overcome some of these technical hurdles, multiple groups have turned their attention to sophisticated GEMMs and non-germline GEMMs (Heyer et al., 2010) that combine the best features of SSR technologies in combination with RNAi and emerging genome engineering technologies (discussed in **Part II** of this introduction).

F. Next-generation GEMMs and non-germline GEMMs of cancer

All of the technologies that I have described so far can be combined to produce some of the most elegant and sophisticated GEMMs and non-germline GEMMs (nGEMMs) (Heyer et al., 2010) of cancer. By exploiting the recombinase-mediated cassette exchange (RMCE) technology developed by Jürgen Bode's group (O'Gorman et al., 1991, Schlake and Bode 1994, Seibler et al., 1998) and later applied to the ubiquitously-expressed *ColA1* (Beard et al., 2006) and *Rosa26* (Seibler et al., 2007) loci by the groups of Rudolf Jaenisch and Frieder Schwenk, respectively, researchers are now able to rapidly generate ES cells and GEMM-derived ES cells in which expression cassettes (carrying cDNAs, shRNAs, and, more recently, single guide RNAs (see **Part II** of this

introduction)) can be seamlessly introduced at single copy in the genome (for comprehensive reviews on RMCE technologies, see Turan et al., 2013 and Turan et al., 2014). Notably, the ability to knock-in these expression cassettes at single copy overcomes some of the major technical hurdles imposed by traditional transgenic technologies, such as the inability to control the number of copies of the transgene that get integrated (which in turn affects the expression levels of the transgene) and the propensity of transgenes to get silenced. This technology has been exploited by several researchers in the field, such as the groups of Anton Berns and Jos Jonkers (Huijbers et al., 2014, Henneman et al., 2015), as well as the group of Scott Lowe (Premsrirut et al., 2011, Dow et al., 2012, Premsrirut et al., 2013, Livshits and Lowe 2013, Saborowski et al., 2014, Miething et al., 2014, Dow et al., 2015) to develop incredibly sophisticated GEMMs and nGEMMs of cancer that effectively combine tetracycline-regulatable systems, hormone-regulatable systems, Cre/loxP systems, and modern RNAi technologies. The mouse model of colorectal cancer (CRC) developed by Lukas Dow, Scott Lowe, and colleagues earlier this year epitomizes the power of this combined approach (Dow et al., 2015). This elegant mouse model has 4 main parts: (1) an intestinal stem cell-restricted, tamoxifen-inducible Cre^{ER} allele knocked into the *Lgr5* locus (Barker et al., 2007), (2) a LSL-rtTA allele knocked into the *Rosa26* locus via RMCE (Dow et al., 2014), (3) a potent TRE-driven shRNA targeting the *Apc* tumor suppressor gene knocked into the *ColA1* locus via RMCE (Premsrirut et al., 2011, Dow et al., 2012), and (4) endogenous, conditional alleles of Cre-inducible oncogenic $Kras^{G12D}$ and floxed *p53* (Jackson et al., 2001, Marino et al., 2000, Jonkers et al., 2001). Therefore, co-injection of tamoxifen and doxycycline induces the development of

aggressive CRC by activating oncogenic *Kras*^{G12D} and concomitantly inactivating the *p53* and *Apc* tumor suppressor genes via Cre-recombinase- and RNAi-mediated loss of function, respectively, in an intestinal stem cell-restricted fashion. Using this sophisticated mouse model, they demonstrated that established colorectal adenocarcinomas are exquisitely sensitive to *Apc* reactivation (Dow et al., 2015).

G. Synopsis and outlook

The mouse modeling field will keep witnessing the development of increasingly complicated (but highly tractable) models of cancer that will allow for unprecedented control of the genetics of a tumor and its microenvironment. These tools will combine inducibility, reversibility, and exquisite spatiotemporal control of the genetics of both tumor cells and cells of the tumor microenvironment in a way that will allow researchers to tackle some of the most complex problems in cancer biology. As the genetic toolkits keep expanding - for example, one can achieve inducible and reversible control of gene expression using light (Cambridge et al., 2009, Sinha et al., 2010a, Sinha et al., 2010b, Kennedy et al., 2010, Tucker 2012, Wang et al., 2012, Nihongaki 2015a, Nihongaki 2015b, Hemphill et al., 2015, Polstein and Gersbach 2015) or small molecule dimerizers (Jullien et al., 2003, Hirrlinger et al., 2009a, Hirrlinger et al., 2009b, Zetsche et al., 2015) - the rate limiting step will be the imagination and creativity of the researcher. Furthermore, all of the aforementioned technologies could be seamlessly combined with emerging genome engineering approaches, such as those based on the CRISPR-Cas9 system, to gain even further control over the genetic events being modeled *in vivo*. These genome engineering technologies are the focus of the next section.

Chapter 1 – Introduction

Part II – CRISPR-Cas systems

Portions of this chapter contain extended sections of a review published in final form as: **Sánchez-Rivera, F.J.**, and Jacks, T. (2015). Applications of the CRISPR-Cas9 system in cancer biology. Nat Rev Cancer *15*, 387–395.

I - Historical Perspective

A. Genome engineering in ES Cells

As discussed extensively in **Part I** of the introduction, the pioneering work by Mario Capecchi, Oliver Smithies, and others on gene targeting in embryonic stem (ES) cells via homologous recombination (Smithies et al., 1985, Thomas et al., 1986, Mansour et al., 1988) provided the scientific community the means to generate numerous GEMMs harboring precise mutations in tumor suppressor genes and oncogenes, as well as cell lines with defined loss of function or gain of function alterations in genes that are relevant to cancer biology. Moreover, this technology has been successfully employed in combination with SSRs, such as Cre and Flp, to generate conditional alleles of a large number of cancer genes. Although a mainstay of cancer genetics over the past two decades, these gene modification approaches have been limited by the relatively

low efficiency of gene targeting by homologous recombination and the time required for ES cell manipulation and subsequent mouse breeding.

B. DSBs increase the efficiency of genome engineering

One strategy to increase the efficiency of gene targeting, which was catalyzed by groundbreaking research carried out by the groups of James Haber (Rudin et al., 1989, Plessis et al., 1992), Maria Jasin (Rouet et al., 1994), Jean-François Nicolas (Choulika et al., 1995), and Dana Carroll (Bibikova et al., 2002 and Bibikova et al., 2003), among others, is to introduce DNA double-strand breaks (DSBs) at the genomic locus of interest. These DSBs are repaired by cellular DNA repair pathways, particularly by the error-prone non-homologous end joining (NHEJ) pathway, which frequently leads to insertion or deletion mutations (indels). DSBs are also repaired by the homology-directed repair (HDR) pathway, which can mediate precise DNA modifications in the presence of exogenous donor DNA templates (**Figure 1**).

The realization that such endogenous cellular pathways could be harnessed to increase the efficiency of genome engineering led to a scientific revolution that brought to light several experimental platforms that allowed for precise introduction of DSBs. Indeed, it is this single biological phenomenon that unites all of the genome engineering technologies developed so far, and it is also the one that will undoubtedly unite all future experimental platforms. I will briefly describe four of these platforms for historical purposes, with the goal of emphasizing how each technology set up the stage for the technology that followed, and how important lessons learned from each of the preceding

technologies positively impacted the efficiency with which each of the subsequent technologies were firmly established in the genome engineering field.

C. Meganucleases / Homing Endonucleases

Meganucleases (also referred to as homing endonucleases) are site-specific endonucleases that recognize large sequences (typically >12bp long) (Pâques and Duchateau 2007). The prototypical meganucleases are the *Homothallic switching endonuclease* (HO endonuclease) (Kostriken et al., 1983) and the intron-encoded I-SceI endonuclease (Jacquier and Dujon 1985), both of which were initially characterized in *Saccharomyces cerevisiae*. The former is involved in mediating the well-described mating type switching phenomenon in yeast (reviewed in Haber 1998) while the latter is involved in propagating the mitochondrial intron that encodes it among intron-minus copies of the same gene through gene conversion (Jacquier and Dujon 1985). Beyond their biological role in *Saccharomyces cerevisiae*, what united these two endonucleases was their intrinsic capacity to recognize specific DNA sequences in a highly precise manner. For example, the HO endonuclease was shown to recognize and cleave an 18bp DNA sequence within the *MAT* locus (Nickoloff et al., 1986) while the I-SceI endonuclease was shown to recognize an 18bp DNA sequence within the group I intron of the mitochondrial 21S rRNA gene (Colleaux et al., 1988).

The fact that these two endonucleases were able to recognize and cleave a large (18bp) sequence in a very precise manner suggested at least three things: (1) that such endonucleases could be harnessed to increase the rate of homologous recombination in cells via the incorporation of templates harboring their cognate recognition sequences

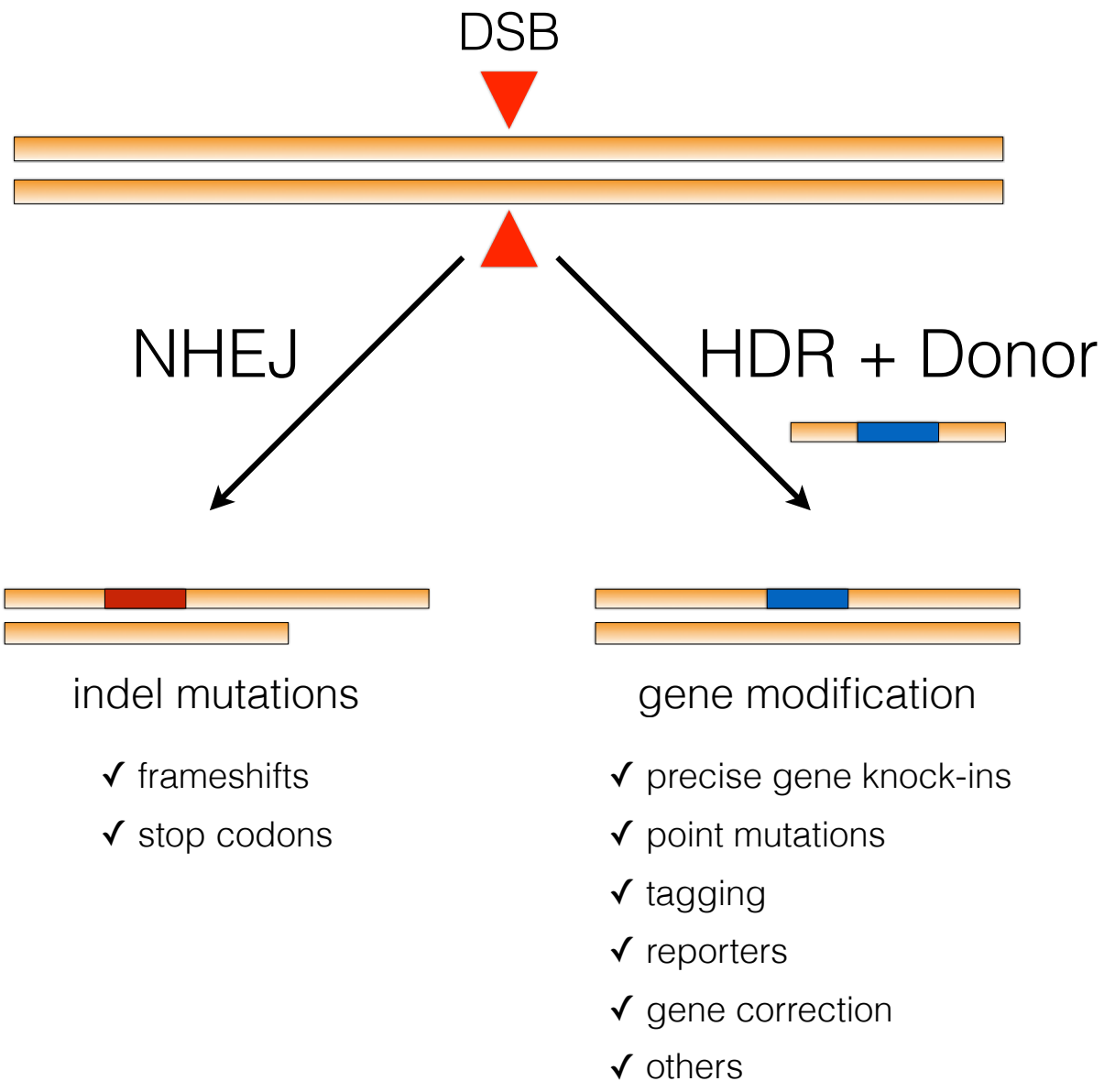


Figure 1: Cellular DNA repair pathways can be harnessed for genome engineering. DNA double-strand breaks (DSBs) (red arrows) can be repaired by two cellular DNA repair pathways: the non-homologous end joining (NHEJ) pathway or the homology-directed repair (HDR) pathway. Repair via the NHEJ pathway, which is error-prone, frequently leads to insertion or deletion mutations (indels) that can lead to disruptive frameshift mutations and the generation of premature stop codons. Alternatively, in the presence of an exogenous donor DNA template, the DSB can be repaired via the HDR pathway, which can be used for engineering precise DNA modifications. Adapted from Sánchez-Rivera & Jacks (2015).

in various configurations, (2) that engineering such sites in the mammalian genome (for example, by introducing two of these sites flanking a gene of interest) would allow for highly efficient gene knockouts, and (3) that altering the specificity of their DNA recognition domains via protein engineering could, in theory, allow for recognition of a multitude of other sequences in the genome. The latter point opened the door to what is known today as the field of programmable nucleases (sometimes referred to as *designer* or *tailored* nucleases. Indeed, multiple groups during the last decade have reported the successful engineering of novel meganucleases derived from the well-described I-SceI (Doyon et al., 2006) and I-CreI (Arnould et al., 2006, Smith et al., 2006, Redondo et al., 2008, and Muñoz et al., 2011) homing endonucleases. Notably, these I-SceI and I-CreI derivatives were able to precisely recognize and cleave a completely different DNA sequence while sparring its natural recognition sequence (reviewed in Pâques and Duchateau 2007). Despite the fact that genome engineering platforms built on meganucleases are undoubtedly powerful, their widespread adoption has been severely hampered by the complexity of designing and evolving these nucleases, as well as an incomplete understanding of the relationship between protein residues in their DNA binding and nuclease domains and their recognition, interaction, and cleavage of their target DNA sequences. Therefore, the genome engineering field shifted its attention to another class of programmable DNA binding proteins – zinc finger proteins.

D. Zinc Finger Proteins and Zinc Finger Nucleases

Zinc finger proteins (ZFPs) were discovered through biochemical approaches by the groups of Cheng-Wen Wu (Hanas et al., 1983) and Aaron Klug (Miller et al., 1985) while studying the transcription factor IIIA from *Xenopus laevis* using the frog's oocytes as a biochemical system. Intriguingly, each zinc finger domain was subsequently found to be composed of 30 amino acid residues sequentially positioned in a very particular order: Cys – Cys – His – His (commonly abbreviated as Cys_2His_2 or C_2H_2) (Ginsberg et al., 1984). A few years later, the crystal structure of a zinc finger-containing protein physically bound to DNA was solved by the group of Carl O. Pabo, which confirmed the C_2H_2 repetitive structure of the protein and further suggested that the "*structure provides a framework for understanding how zinc fingers recognize DNA and suggests that this motif may provide a useful basis for the design of novel DNA-binding proteins*" (Pavletich and Pabo 1991). The latter point combined with the fact that a number of contemporaneous reports demonstrated that the Type IIS restriction enzyme FokI (Kita et al., 1989, Looney et al., 1989) could be fused to individual, sequence-specific DNA binding domains to create site-specific chimeric restriction endonucleases (Kim and Chandrasegaran 1994 and Kim et al., 1996) inspired one of the earliest scientific races in the field of genome engineering, in which numerous investigators employed phage display screening (Pande et al., 2010) as an attempt to identify novel ZFPs that were able to recognize each triplet of the genetic code with high specificity and low promiscuity (Rebar and Pabo 1994, Choo and Klug 1994a, Choo and Klug 1994b, Choo et al., 1994, Greisman and Pabo 1997, and many others). Unfortunately, these efforts had numerous shortcomings mainly because experimentally identifying or

evolving individual ZFP modules to recognize a specific triplet with high specificity and sensitivity was virtually impossible due to the fact that the recognition of DNA by individual ZFPs was not completely modular and instead depended on (and was influenced by) the neighboring ZFP-DNA contact (Isalan et al., 1997). Multiple investigators, including the groups of Carl O. Pabo (Wolfe et al., 1999, Joung et al., 2000, Miller and Pabo 2001), Yen Choo (Moore et al., 2001, Isalan et al., 2001), Keith Joung (Hurt et al., 2003) and Srinivasan Chandrasegaran (Durai et al., 2005 and Durai et al., 2006), among others, tackled these issues in a variety of clever ways by developing novel iterative phage display-based and cell-based selection strategies aimed at circumventing the issues associated with engineering single, individual ZFPs (reviewed in Durai et al., 2005 and Chandrasegaran and Carroll, 2015). Albeit a number of these and other subsequent strategies greatly improved the design specificity and throughput with which new ZFPs and ZFNs could be engineered for genome editing purposes, and despite the fact that ZFNs were used for genome editing in a multitude of organisms – from *Drosophila melanogaster* (Bibikova et al., 2002, Bibikova et al., 2003), to *Caenorhabditis elegans* (Morton et al., 2006, Wood et al., 2011), to *Danio rerio* (Doyon et al., 2008, Meng et al., 2008), to *Sus scrofa* (Hauschild et al., 2011), to *Mus musculus* (Meyer et al., 2010, Carbery et al., 2010), all the way to somatic and pluripotent stem cells derived from *Homo sapiens* (Porteus and Baltimore, 2003, Urnov et al., 2005, Lombardo et al., 2007, Zou et al., 2009, Hockemeyer et al., 2009), their widespread adoption by the scientific community was still limited due to the fact that generating these custom-built nucleases remained highly complex and costly.

Therefore, the genome engineering field shifted its attention to yet another class of programmable DNA binding proteins - transcription activator-like effector proteins.

E. Transcription Activator-Like Effector Proteins

Transcription activator-like effector proteins (TALEs) comprise a large family of type III effector proteins originally identified in plant pathogens of the *Xanthomonas* genus by the group of Robert Stall (Minsavage et al., 1990). In the context of plant pathogenesis, proteobacteria from the *Xanthomonas* genus inject these TALEs into host plant cells, which are then shuttled to the nucleus where they physically bind the host's genomic DNA and mediate transcriptional activation of a plethora of genes important for bacterial colonization (reviewed in Boch and Bonas, 2010). Similar to ZFPs, TALEs are comprised of fully customizable DNA binding domains that can be fused to the FokI endonuclease to create site-specific programmable endonucleases termed TALE nucleases (or TALENs for short) (reviewed in Joung and Sander, 2013). Remarkably, TALE-derived DNA binding domains are composed of variable arrays of 33-35 amino acid repeats, each of which is able to recognize a single base of DNA (Boch et al., 2009). This incredibly specificity is dictated by two amino acids within the variable array, which are typically referred to as the hypervariable residues and are found in positions 12 and 13 of the array (Boch et al., 2009). Due to the fact that the genome engineering field was already primed for such a discovery, several researchers rapidly elucidated the TALE repeat code, both experimentally (Boch et al., 2009) as well as computationally (Moscou and Bogdanove 2009). The TALEN revolution that followed closely echoed the ZFN revolution and, just three years after the TALE repeat code was elucidated, several

groups reported successful TALEN-mediated genome engineering in *Saccharomyces cerevisiae* (Li et al., 2011), *Drosophila melanogaster* (Liu et al., 2012), *Caenorhabditis elegans* (Wood et al., 2011), *Danio rerio* (Sander et al., 2011, Huang et al., 2011 and Bedell et al., 2012), *Rattus norvegicus* (Tesson et al., 2011), *Sus scrofa* (Carlson et al., 2012), and in somatic and pluripotent stem cells derived from *Homo sapiens* (Cermak et al., 2011, Miller et al., 2011, Hockemeyer et al., 2011, Reyon et al., 2012). Despite the fact that TALENs offered greater design flexibility when compared to ZFNs and meganucleases (as exemplified by the publication of multiple relatively user-friendly protocols for TALEN design by the Zhang (Zhang et al., 2011, Sanjana et al., 2012) and Joung (Reyon et al., 2012) groups), generating custom TALENs is still labor intensive and usually requires multiple rounds of design and screening to achieve potent and highly specific reagents. Just like the genome engineering field experienced the ZFN and TALEN revolutions, the CRISPR revolution that ensued took the entire field by storm and has literally transformed every single aspect of biological research (Barrangou 2014).

F. CRISPR-Cas Systems

The recently described clustered regularly interspaced short palindromic repeats (CRISPR)-CRISPR associated (Cas) system and the successful implementation of the *Streptococcus pyogenes*-derived type II CRISPR-Cas9 system in mammalian cells by the Zhang (Cong et al., 2013), Church (Mali et al., 2013a), Doudna (Jinek et al., 2013) and Kim (Cho et al., 2013) groups has rapidly changed the landscape of genome engineering by overcoming many of the technical hurdles of earlier methods. Due to the

fact that this thesis revolves around the development and application of CRISPR-Cas9 technologies for modeling and dissecting tumorigenesis, I will devote the next few sections to a comprehensive (but concise) recount of the key events in the CRISPR timeline that led to the genome engineering revolution that we are experiencing today.

II - Discovery and historical timeline

A remarkable feature that unites all four genome engineering technologies discussed so far is the fact that their roots are firmly established in fundamental biological research. Just like meganucleases were discovered through fundamental yeast molecular genetics, and ZFNs were discovered through basic molecular biology, the origins of TALENs and CRISPR lie in fundamental microbiology research. Due to the fact that genuine scientific curiosity was the main driver behind these transformative technologies, I will devote a few paragraphs to a discussion of the key historical events in the CRISPR-Cas9 field (**Figure 2**).

The first reported event in the CRISPR timeline dates back to 1987 when a group of researchers at Osaka University in Japan that were working on the gene responsible for the isozyme conversion of alkaline phosphatase in *Escherichia coli*, called *iap*, reported a curious stretch of repetitive 29nt long sequences that were interspaced by 32nt long variable sequences at the end of the gene (Ishino et al., 1987) (**Figure 2**). It took 13 years of fundamental microbiology research and extensive prokaryotic genomics for researchers to notice that these repetitive sequences were pervasive in both bacterial and archaeal genomes. This realization led Mojica and colleagues in the year 2000 to

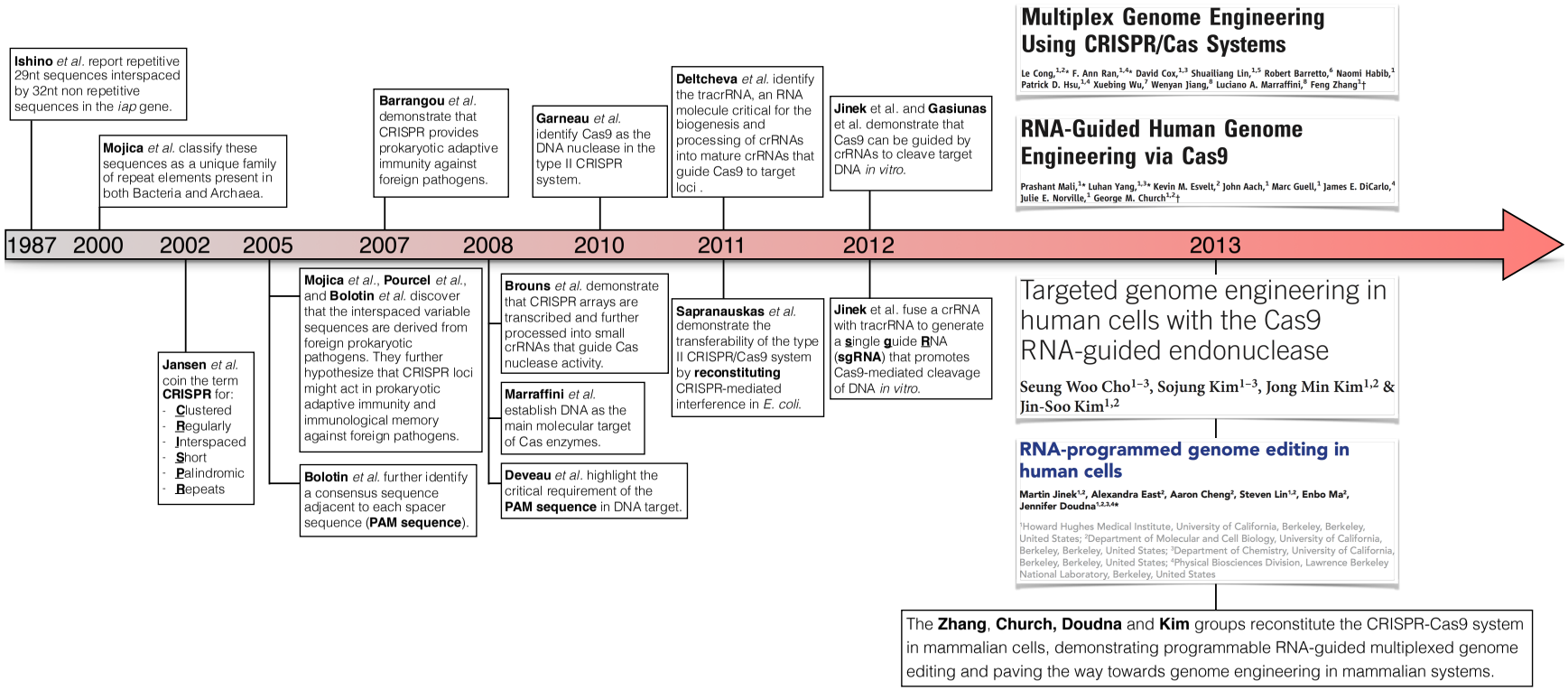


Figure 2: Major milestones in the CRISPR-Cas9 timeline. See text for details.

suggest that such sequences (which they initially termed short regularly spaced repeats (SRSRs)) be classified as a unique family of repeat elements present across the Bacteria and Archaea domains (Mojica et al., 2000) (**Figure 2**). Two years later, Jansen and colleagues coined the term CRISPR for clustered regularly interspaced short palindromic repeats and further noted that CRISPR-associated (Cas) genes were usually located adjacent to CRISPR loci, suggesting a possible functional relationship from this spatial organization of the CRISPR locus (Jansen et al., 2002) (**Figure 2**).

It was in 2005 that, in my opinion, one of the major breakthroughs in the CRISPR timeline took place. In this year, three separate groups discovered that the variable spacer sequences did not correspond to endogenous genomic loci in the prokaryotic genome (**Figure 2**). Rather, they found that these variable spacer sequences were actually derived from foreign prokaryotic pathogens, such as bacteriophages and conjugative plasmids (Mojica et al., 2005, Pourcel et al., 2005 and Bolotin et al., 2005). They further hypothesized that CRISPR loci might act in prokaryotic adaptive immunity and that the variable spacers might confer a form of immunological molecular memory against foreign pathogens (Mojica et al., 2005, Pourcel et al., 2005 and Bolotin et al., 2005). Bolotin and colleagues further noted that there was a consensus sequence adjacent to spacer sequences. Notably, this consensus sequence corresponded to the so-called Protospacer Adjacent Motif (PAM) sequence, whose function was unknown back then (Bolotin et al., 2005).

The year 2007 witnessed yet another major breakthrough in the CRISPR timeline (**Figure 2**) when Rodolphe Barrangou, Philippe Horvath and colleagues published the

seminal study conclusively demonstrating that CRISPR provided prokaryotic adaptive immunity against foreign pathogens (Barrangou et al., 2007). In a series of very elegant experiments, they demonstrated that the removal or addition of particular spacers corresponding to different bacteriophages was sufficient to modify the adaptive immunity of the prokaryote against the foreign pathogens (Barrangou et al., 2007). Subsequently, a flurry of papers in 2008 (**Figure 2**) demonstrated three very important concepts: (1) that CRISPR arrays are transcribed and further processed into small CRISPR RNAs (crRNAs) that guide Cas nuclease activity (Brouns et al., 2008), (2) that DNA is the main molecular target of Cas enzymes (Marraffini and Sontheimer, 2008), and (3) that the PAM sequence present in the target DNA is critical for CRISPR-mediated interference (Deveau et al., 2008). It only took two years for the group of Sylvain Moineau to demonstrate that the Cas9 gene of the type II CRISPR system encoded a DNA nuclease (Garneau et al., 2010) (**Figure 2**).

That CRISPR loci contained (1) variable spacers corresponding to foreign genome invaders that were critical in mediating immunity against such intruders and (2) an enzyme (Cas9) with potent DNA nuclease activity started to give hints of a system whereby the spacers somehow directed Cas9 to cleave foreign genomes. A year later (**Figure 2**), another critical element of CRISPR systems was uncovered by the group of Emmanuelle Charpentier – the *trans*-activating crRNA (tracrRNA) – which is a *trans*-encoded small RNA that is critical for the biogenesis and processing of crRNAs into mature crRNAs that then guide Cas9 to target DNA molecules (Deltcheva et al., 2011). Remarkably, the tracrRNA was found to contain a 24nt long region with perfect complementarity to the repeat regions of precursor crRNA transcripts (pre-crRNAs) that

was critical for promoting the maturation of pre-crRNAs into mature crRNAs through the recruitment of the endogenous RNase III machinery (Deltcheva et al., 2011). During that same year, the group of Virginijus Siksnys reported a major breakthrough when they demonstrated that the type II CRISPR-Cas9 system was transferable between different bacterial species, functionally reconstituting CRISPR-mediated interference in a different host (Sapranauskas et al., 2011) (**Figure 2**).

The fact that such a powerful, DNA-targeted, sequence-specific system could be transferred between completely different hosts to mediate cleavage of a specific sequence of interest prompted a real race towards functionally understanding this novel DNA-interference system. The year 2012 witnessed the breakthroughs that truly brought the CRISPR-Cas9 system to the genome engineering spotlight (**Figure 2**). In a pair of groundbreaking research articles, the group of Virginijus Siksnys and the groups of Jennifer Doudna and Emmanuelle Charpentier conclusively demonstrated that Cas9 could be guided by crRNAs to cleave target DNA *in vitro* (Gasiunas et al., 2012 and Jinek et al., 2012). Furthermore, Doudna and colleagues made the seminal observation that fusing a crRNA with the tracrRNA to generate a chimeric single guide RNA (sgRNA) molecule was sufficient to promote Cas9-mediated cleavage of specific DNA molecules *in vitro* (Jinek et al., 2012). This last observation highlighted the power and simplicity of the CRISPR-Cas9 system for genome editing: an individual sgRNA molecule can guide Cas9 to cleave a genomic site of interest via Watson-Crick base pairing between the RNA molecule and the DNA target sequence (Jinek et al., 2012). Therefore, CRISPR-Cas9-mediated genome editing required only two components. This realization triggered what was inarguably the fastest race in the CRISPR-Cas9 timeline. In 2013,

the groups of Feng Zhang (Cong et al., 2013), George Church (Mali et al., 2013a), Jennifer Doudna (Jinek et al., 2013), and Jin-Soo Kim (Cho et al., 2013) reported that the type II CRISPR-Cas9 system could be reconstituted in mammalian cells, demonstrating efficient RNA-guided multiplexed genome editing and paving the way towards the CRISPR revolution that followed (**Figure 2**).

To date, the CRISPR-Cas9 system has been utilized to engineer the genomes of numerous organisms, including *Mus musculus* (Wang et al., 2013, Yang et al., 2013), *Rattus norvegicus* (Hu et al., 2013, Ma et al., 2014), *Drosophila Melanogaster* (Gratz et al., 2013, Bassett et al., 2013), *Caenorhabditis elegans* (Friedland et al., 2013), *Danio rerio* (Chang et al., 2013, Hwang et al., 2013), *Sus scrofa* (Whitworth et al., 2014), *Macaca fascicularis* (Niu et al., 2014), and multiple somatic and pluripotent stem cells derived from *Homo sapiens* (extensively reviewed in Doudna and Charpentier 2013, Hsu et al., 2013 and Cox et al., 2015, among others). From basic research, to applied biotechnology and medicine, the CRISPR-Cas9 system has truly revolutionized every aspect of biological research. Experimental approaches based on this versatile technology have the potential to transform the field of cancer biology. In the next section, I will discuss current and future approaches for functional studies of cancer genes that are based on CRISPR-Cas9, with emphasis on their applicability for the development of next-generation models of human cancer (**Figure 3**).

III. Applications in Cancer Biology

As discussed in the previous section, the CRISPR-Cas9 system is composed of two biological components: the RNA-guided DNA endonuclease Cas9 and a chimeric

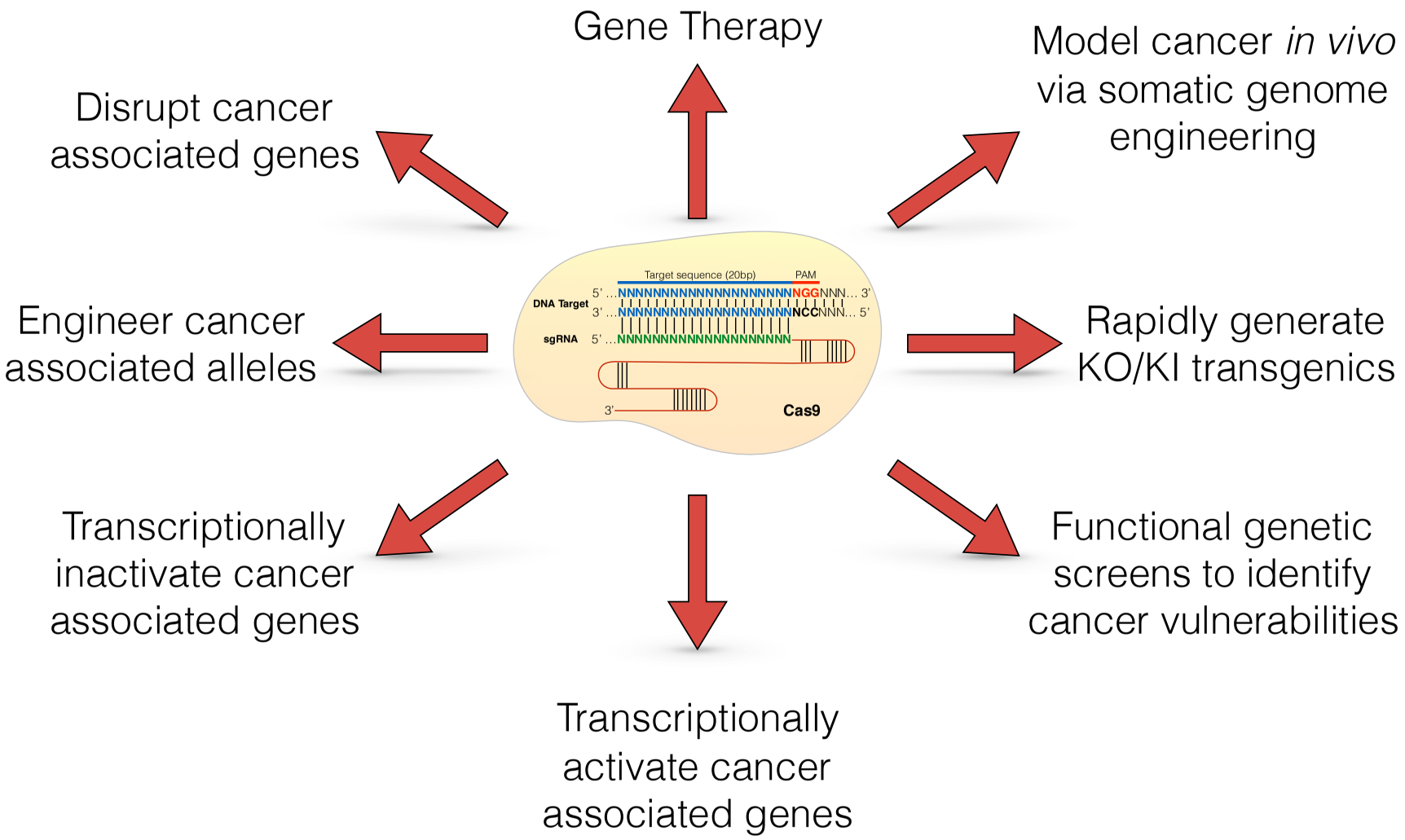


Figure 3: Selected cancer-centric applications of the CRISPR-Cas9 system. See text for details.

sgRNA. The sgRNA molecule contains both a crRNA component and a tracrRNA component, which binds to Cas9 and directs it to a genomic sequence of interest via base pairing to the target sequence (Jinek et al., 2012) (**Figure 4**). The only criterion defining the target sequence is that it be adjacent to a PAM sequence, consisting of either an NGG or NAG trinucleotide (Hsu et al., 2013) for *Streptococcus pyogenes*-derived Cas9 (of note, other Cas9 orthologs (Esvelt et al., 2013, Hou et al., 2013, Ran et al., 2015, Chari et al., 2015, Zhang et al., 2015, Zetsche et al., 2015) and engineered Cas9 derivatives (Kleinstiver et al., 2015a, Kleinstiver et al., 2015b) recognize different PAM sequences). By simply combining the expression of Cas9 with an sgRNA complementary to a target DNA sequence, one can achieve high efficiency cleavage of the target, leading to DSBs, which then get repaired via NHEJ or HDR (**Figure 4**). Below, I will discuss several recent and future applications of the CRISPR-Cas9 system, with particular emphasis on approaches that promise to transform the field of cancer biology by facilitating the engineering of normal and cancer genomes.

A. Rapid modeling of genetic events in cell culture systems

i. LOF/GOF events using the Cas9 nuclease

In the current era of cancer genomics, several large-scale cancer genome sequencing efforts have produced an expanding catalogue of the genetic alterations present in human tumors (Vogelstein et al., 2013). Amongst a background of so-called passenger mutations, which are presumed not to directly affect the tumorigenic process, driver mutations directly or indirectly promote the transformation of normal cells to cancer cells through mutational activation of oncogenes and/or inactivation of tumor suppressor

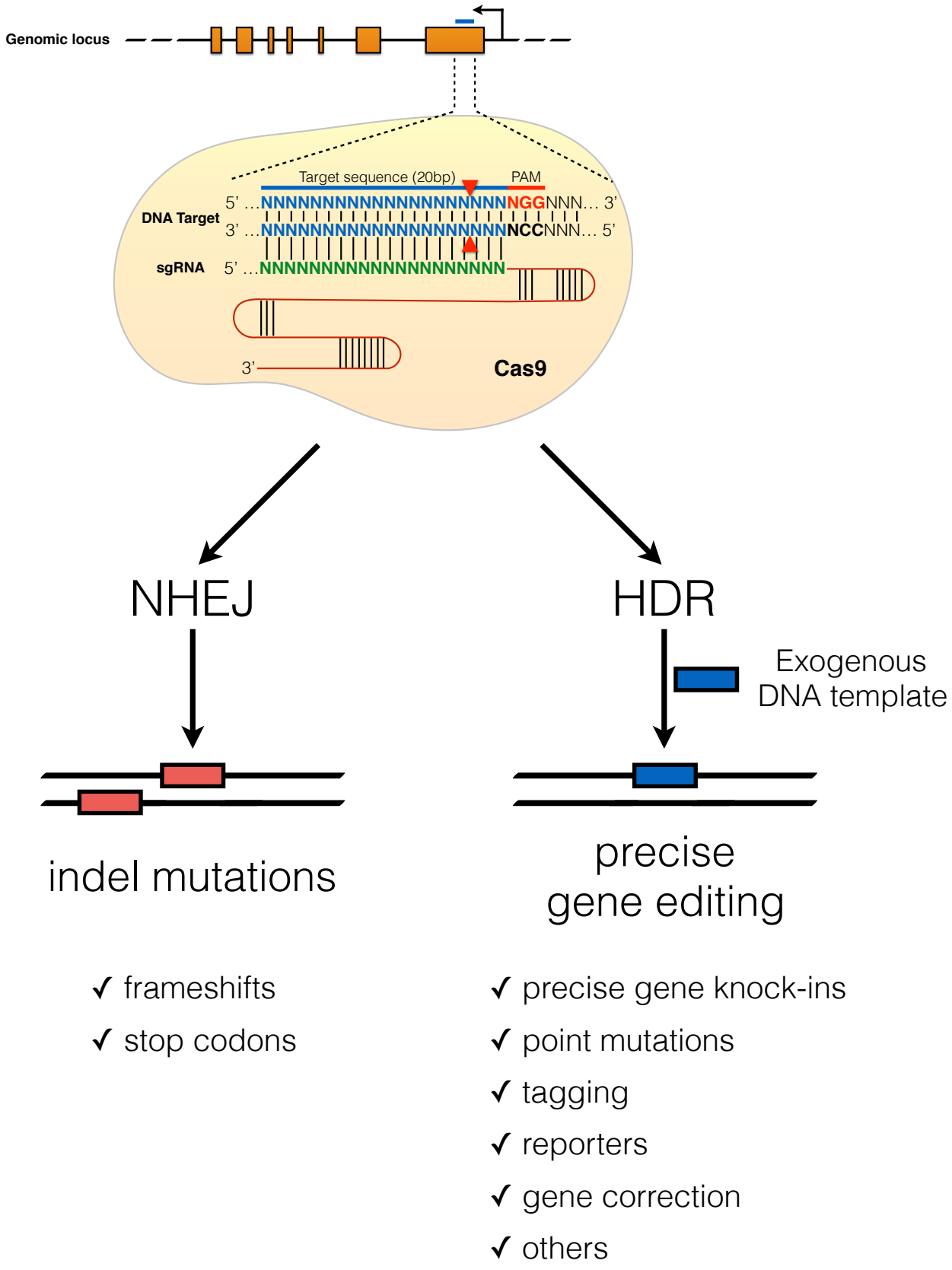

Genomic locus
Target sequence (20bp)
PAM
DNA Target
5' …NNNNNNNNNNNNNNNNNNNNGGNNN… 3'
3' …NNNNNNNNNNNNNNNNNNNNCCNNN… 5'
sgRNA
5' …NNNNNNNNNNNNNNNNNNNN
3'
Cas9
NHEJ
HDR
Exogenous
DNA template
indel mutations
precise
gene editing
✓ frameshifts
✓ stop codons
✓ precise gene knock-ins
✓ point mutations
✓ tagging
✓ reporters
✓ gene correction
✓ others

Figure 4: Genome engineering using the CRISPR-Cas9 system. The *Streptococcus pyogenes*-derived CRISPR-Cas9 RNA-guided DNA endonuclease is localized to a specific DNA sequence via a single guide RNA (sgRNA) sequence, which base-pairs with a specific target sequence that is adjacent to a protospacer adjacent motif (PAM) sequence in the form of NGG or NAG. Cas9-mediated induction of a DSB (red arrows) in the DNA target sequence leads to indel mutations via NHEJ or precise gene modification via HDR. Adapted from Sánchez-Rivera & Jacks (2015).

genes. Oncogenes are typically activated via gain of function (GOF) mutations whereas tumor suppressor genes are usually inactivated via loss of function (LOF) mutations. Moderate- to large-scale functional genetic studies aimed at dissecting the role of putative oncogenes and tumor suppressor genes in cell culture, xenografts, allografts and, in some cases, transgenic mouse models, have traditionally relied on cDNA-based overexpression and RNA interference (RNAi)-mediated knockdown approaches. While these approaches have led to many important discoveries in cancer biology over the last several years, they have a number of important limitations. First, cDNA-based expression systems can lead to supraphysiological levels of gene expression (Denicola et al., 2011), which might cause aberrant and artifactual effects on signaling pathways and cell biological processes. RNAi-based inactivation approaches are limited by the uncertainty of the degree of gene silencing and the stability of the inhibition. This is not problematic for some targets or experimental protocols, but for others complete and permanent inactivation is required to obtain consistent results. RNAi-based approaches can also suffer from substantial off-target effects (Kaelin 2012). The deployment of the CRISPR-Cas9 system for targeted modification of endogenous loci offers a rapid method for overcoming these limitations. In addition to simplifying the study of oncogenes and tumor suppressor genes, the CRISPR-Cas9 system also allows for rapid discrimination between driver and passenger mutations.

Permanent Cas9-mediated modification of single or multiple endogenous loci can be achieved via transient or stable delivery of the CRISPR components. Several groups have reported successful editing of endogenous genes in cells in culture via transient transfection of plasmid DNA encoding Cas9 and sgRNAs (Cong et al., 2013, Mali et al.,

2013a, Jinek et al., 2013, Cho et al., 2013) or Cas9-sgRNA ribonucleoprotein complexes (RNPs) (Kim et al., 2014, Lin et al., 2014). Alternatively, CRISPR components can be stably delivered into cells through the use of retroviruses or lentiviruses (Malina et al., 2013, Shalem et al., 2014). To engineer LOF mutations, one relies on NHEJ, which often results in the generation of indels near the Cas9 cleavage site that frequently lead to frameshift mutations. Engineering GOF mutations requires the inclusion of an HDR template in the form of single-stranded or double-stranded DNA carrying the desired mutation (**Figure 4 and Figure 6)**. Transient expression of the CRISPR components offers the advantage of a hit-and-run strategy, which should allow for unlimited serial editing of endogenous genes without the need of multiple viral integrations or continuous expression of CRISPR components. Cell lines carrying one or more targeted mutations can then be tested using a battery of cell-based and *in vivo* assays to examine the effects of the mutation(s) on cancer-associated phenotypes. This approach can be used on established cancer cell lines and primary cell lines of mouse or human origin, as well as on patient-derived xenografts (PDXs) and organoid cultures, among others (**Figure 5 and Figure 6**).

Indeed, Sato and colleagues recently demonstrated the use of the CRISPR-Cas9 system for systematically engineering both LOF (**Figure 5**) and GOF (**Figure 6**) mutations in untransformed human intestinal organoids in order to model human colorectal cancer (CRC) (Matano et al., 2015). Remarkably, the serial introduction of five independent mutations frequently associated with human CRC (three LOF mutations and two GOF mutations) did not fully recapitulate the tumorigenic and metastatic characteristics of the human disease, suggesting that additional secondary genetic and/

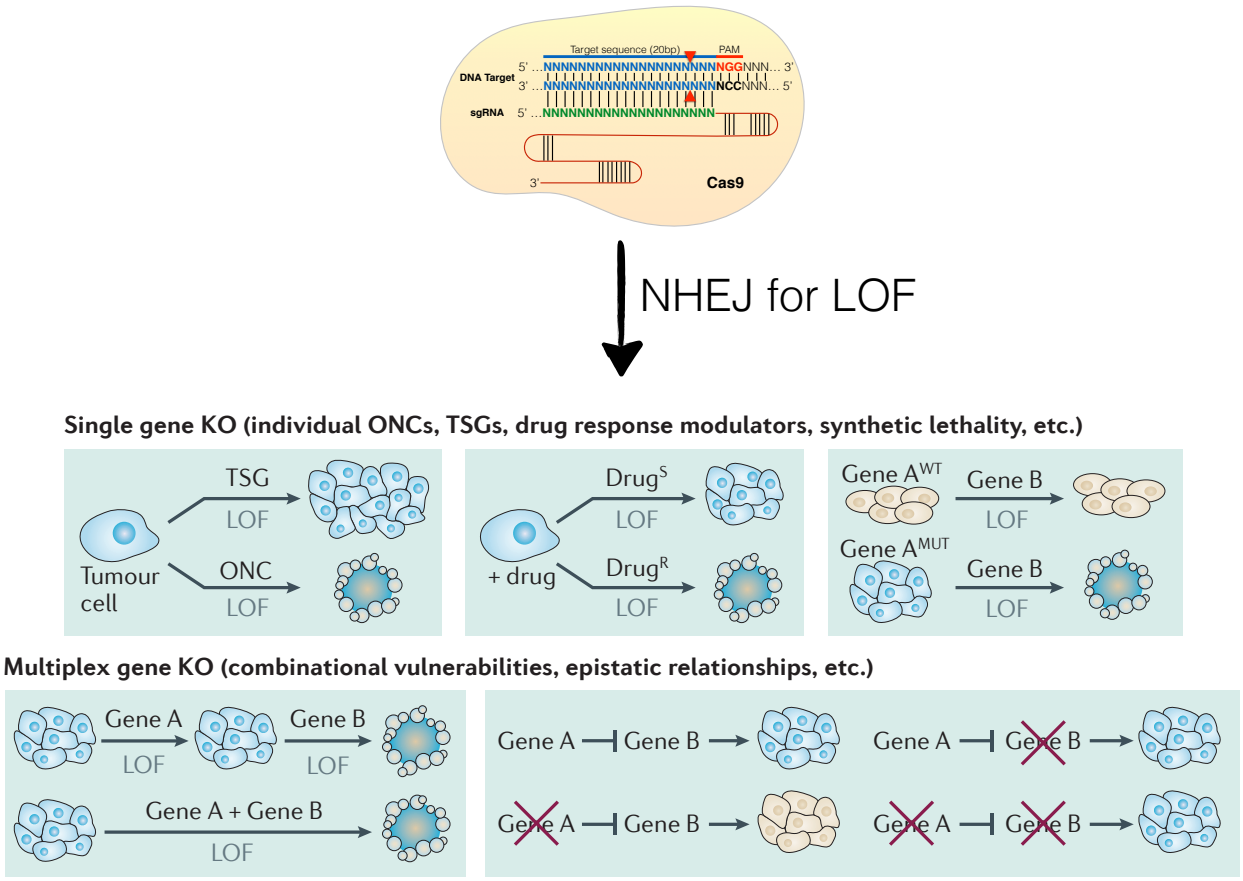
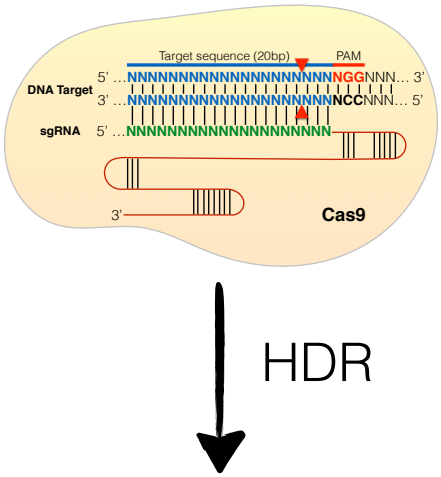
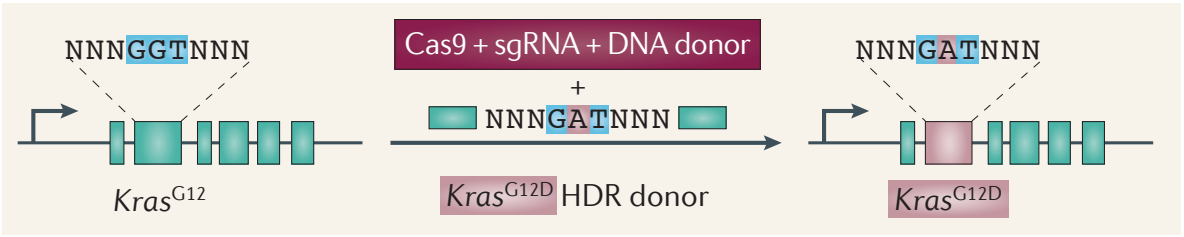


Figure 5: CRISPR-Cas9-mediated disruption of cancer associated genes and related applications. The flexibility and modularity of the CRISPR-Cas9 system allows researchers to dissect the function of cancer associated genes through the efficient disruption of tumor suppressor genes (TSGs), oncogenes (ONCs), and drug response modulators, among other types of genes. Moreover, the multiplexability of the CRISPR-Cas9 system allows researchers to simultaneously disrupt several genes with unprecedented efficiency and precision for tackling questions related to synthetic lethality and epistatic interactions. NHEJ = non-homologous end joining; LOF = loss of function; WT = wild type; MUT = mutant; $Drug^S$ = gene that confers drug sensitivity; $Drug^R$ = gene that confers drug resistance; KO = knockout. Adapted from Sánchez-Rivera & Jacks (2015).



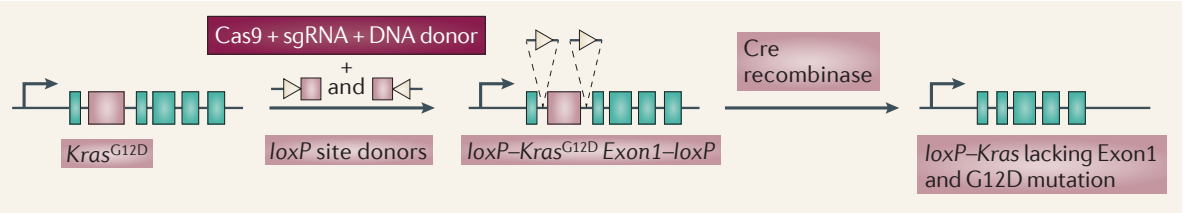
a

Model putative GOF point mutations (TCGA, COSMIC, etc.)



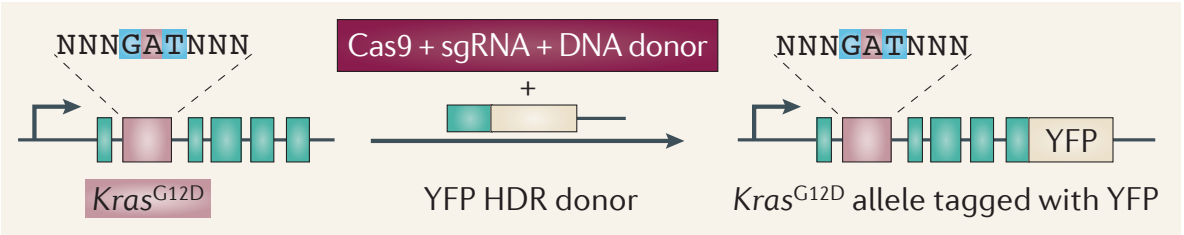
b

Generate endogenous conditional alleles (Cre–*loxP*, Flp–*FRT*, etc.)



c

Tag endogenous alleles (fluorescent reporters, synthetic tags, etc.)



d

Interrogate non-coding DNA elements (enhancers, insulators, promoters, etc.)

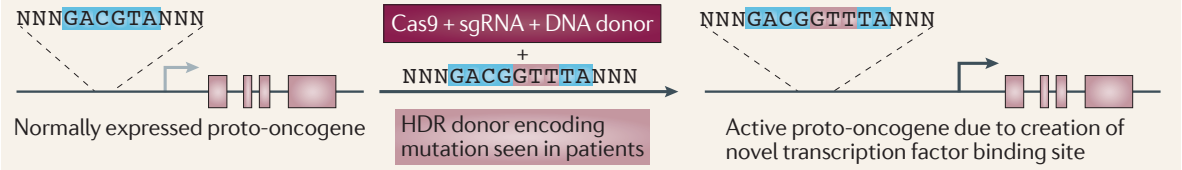


Figure 6: CRISPR-mediated genome engineering through the HDR pathway. (a) The CRISPR-Cas9 system can also be utilized for rapidly and precisely engineering cancer associated gain of function (GOF) mutations in oncogenes and other modulators of cellular transformation or drug response. This technology also allows for the generation of endogenous conditional alleles based on site-specific recombinases (such as Cre and Flp recombinases) **(b)**, tagging of endogenous alleles **(c)**, and interrogation of non-coding DNA elements (d). TCGA = The Cancer Genome Atlas; COSMIC = Catalogue of Somatic Mutations in Cancer. Adapted from Sánchez-Rivera & Jacks (2015).

or epigenetic events are required for full malignancy (Matano et al., 2015). A contemporaneous study by the group of Hans Clevers elegantly confirmed this approach and further demonstrated that co-mutation of the *APC* and *TP53* tumor suppressors was sufficient to trigger extensive chromosomal instability and aneuploidy (Drost et al., 2015).

Beyond the power of the system for modeling individual or sequential mutations, the ability to multiplex the CRISPR-Cas9 system offers the opportunity to investigate combinatorial vulnerabilities in cancer cells, as well as to systematically test epistatic relationships and synthetic lethal interactions (**Figure 5**). This technology also allows for the generation of endogenous conditional alleles based on site-specific recombinases (Yang et al., 2013), tagging of endogenous alleles (Yang et al., 2013), and interrogation of non-coding DNA elements (Mansour et al., 2014) (**Figure 6**).

ii. Chromosomal rearrangements using the Cas9 nuclease

Genomic instability, which is a major hallmark of human cancer, can lead to the development of severe chromosomal rearrangements, such as translocations, inversions and deletions (Mitelman et al., 2007). Chromosomal translocations and inversions frequently lead to the generation of oncogenic fusion alleles, such as the *BCR-ABL1* oncogene, which results from a translocation event and is associated with chronic myeloid leukemia (Rowley et al., 1973) and the *EML4-ALK* oncogene, which results from an inversion event and is associated with a subset of non-small cell lung cancers (Soda et al., 2007). On the other hand, chromosomal deletions can simultaneously inactivate multiple genes, thereby potentially inactivating multiple tumor

suppressor genes with a single hit (Solimini et al., 2012, Davoli et al., 2013). Experimentally modeling these complex genetic rearrangements in cells or animal models has been technically challenging, often requiring serial targeting of ES cells to introduce distant loxP sites in the same chromosome (for modeling inversions and deletions) (Duboule et al., 2000, Smith et al., 2002, Bagchi et al., 2007) or in different chromosomes (for modeling translocations) (Forster et al., 2003) (see **Part I** of this introduction). Furthermore, rapid modeling of variants of these events has required several iterative cycles of ES cell targeting, making it a cumbersome approach for systematically modeling diverse chromosomal rearrangements. Even though ZFNs and TALENs have been utilized for generating targeted chromosomal rearrangements (Brunet et al., 2009 and Ghezraoui et al., 2014), their widespread adaptation has been hampered due to the lack of sufficient design flexibility and feasibility. A series of recent studies have reported the successful adaptation of the CRISPR-Cas9 system to overcome these limitations by exploiting the ability of Cas9 and a pair of sgRNAs to mediate distant DSBs in the same or different chromosomes, which can result in inversions/deletions and translocations, respectively (**Figure 7**). The group of Matthew Meyerson recently demonstrated that co-transfection of a pair of plasmids encoding Cas9 and sgRNAs targeting different regions of human chromosome 2 or chromosome 10 into HEK293T cells resulted in the generation of the oncogenic *EML4-ALK* and *KIF5B-RET* inversions, respectively (Choi & Meyerson 2014). Moreover, pairs of sgRNAs targeting different chromosomes were shown to efficiently generate the oncogenic *CD70-ROS1* translocation (Choi & Meyerson 2014). Using a similar strategy, Torres and colleagues reported the efficient generation of the *EWSR1-FLI1* and

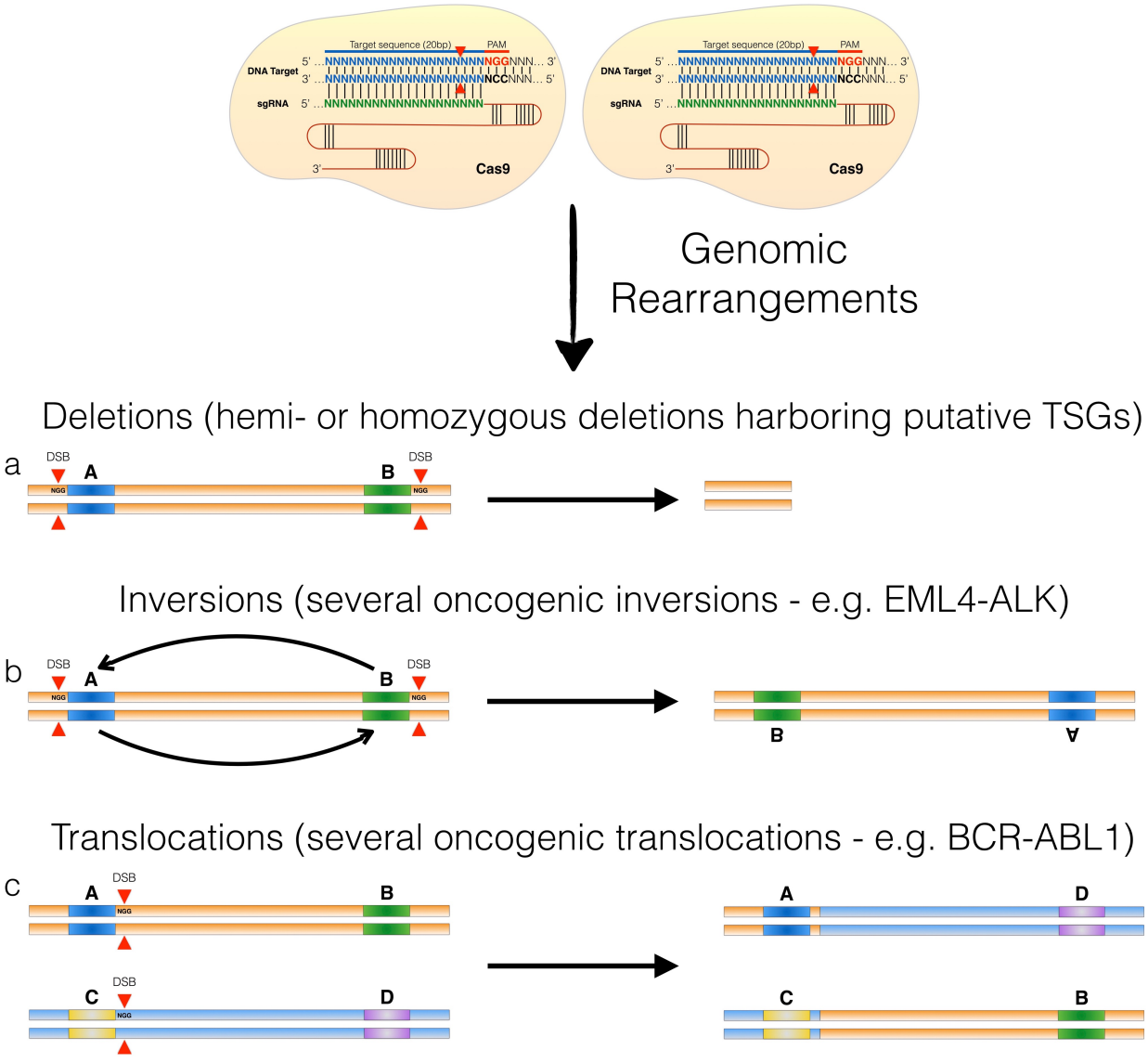


Figure 7: Modeling chromosomal rearrangements using CRISPR-Cas9. The CRISPR-Cas9 system can also be utilized to trigger two distant double-strand breaks (DSBs) (red arrows) in the same or different chromosomes, leading to deletion **(a)**, inversion **(b)** or translocation **(c)** of the target sequences. ALK = anaplastic lymphoma kinase; BCR = breakpoint cluster region. Adapted from Sánchez-Rivera & Jacks (2015).

RUNX1-ETO translocations, which are frequently observed in Ewing's sarcoma and acute myeloid leukemia (AML), respectively (Torres et al., 2014). In addition, the groups of Erika Brunet and Maria Jasin (Ghezraoui et al., 2014) carried out a comprehensive study of chromosomal translocations in human cells in which they reported efficient generation of several translocations, including the *NPM-ALK* translocation characteristic of anaplastic large cell lymphoma utilizing Cas9 and pairs of sgRNAs. These landmark studies have firmly established the CRISPR-Cas9 system as a highly precise and efficient experimental tool to rapidly model chromosomal rearrangements frequently observed across human cancers. In addition to modeling oncogenic inversions and translocations, several groups have reported efficient generation of large chromosomal deletions encompassing coding and non-coding genes using CRISPR-Cas9 (Xiao et al., 2013, Canver et al., 2014, Han et al., 2014, Ho et al., 2014 and Essletzbichler et al., 2014, Lupiáñez et al., 2015), with up to 30 megabases of genetic material being efficiently deleted using this approach (Essletzbichler et al., 2014). The deployment of the CRISPR-Cas9 system for modeling chromosomal rearrangements offers at least two major advantages over previously used approaches. First, it is very rapid and precise, requiring only co-expression of Cas9 and a pair of properly designed sgRNAs. Second, it is a superior approach for studying inversions and translocations due to the fact that manipulation of endogenous loci is more physiologically relevant than cDNA-based overexpression of the fusion products, thus ensuring that the oncogenic fusion protein is still under the control of endogenous regulatory elements. Therefore, this technology has the potential to immediately fill the gap between the identification of numerous chromosomal rearrangements through large-scale cancer genome

resequencing studies and the successful validation of these alterations as *bona fide* cancer promoting events.

iii. LOF/GOF events using dead Cas9

One of the most powerful features of the Cas9 enzyme is the ability to functionally uncouple its exquisitely specific RNA-dependent DNA binding activity from its nuclease activity by mutating its HNH and RuvC-like catalytic domains. This catalytically inactive form of Cas9, often referred to as dead Cas9 (dCas9), retains its RNA-guided DNA binding activity without any detectable DNA endonuclease activity (Jinek et al., 2012) (**Figure 8**). Recent studies have demonstrated that programming of dCas9 with an sgRNA against a gene of interest is sufficient to repress gene expression by inhibiting transcriptional elongation (Qi et al., 2013, Bikard et al., 2013 and Esvelt et al., 2013). Subsequent studies have demonstrated that fusion of dCas9 with transcriptional repressors (Gilbert et al., 2013, Bikard et al., 2013, Gilbert et al., 2014) or transcriptional activators (Gilbert et al., 2013, Bikard et al., 2013, Mali et al., 2013b, Esvelt et al., 2013, Perez-Pinera et al., 2013, Cheng et al., 2013, Maeder et al., 2013, Gilbert et al., 2014, Tanenbaum et al., 2014 and Konermann et al., 2014) can be utilized for potent repression or activation of endogenous genes, respectively. These dCas9-based complementary platforms offer several features that can be exploited for cancer research. First, it takes advantage of the exquisite binding precision of dCas9, which can be efficiently programmed with single or multiple sgRNAs to activate or repress specific cancer-associated loci of interest with low off-target activity (Gilbert et al., 2014). Secondly, the modularity of dCas9 should allow for functional fusions with different

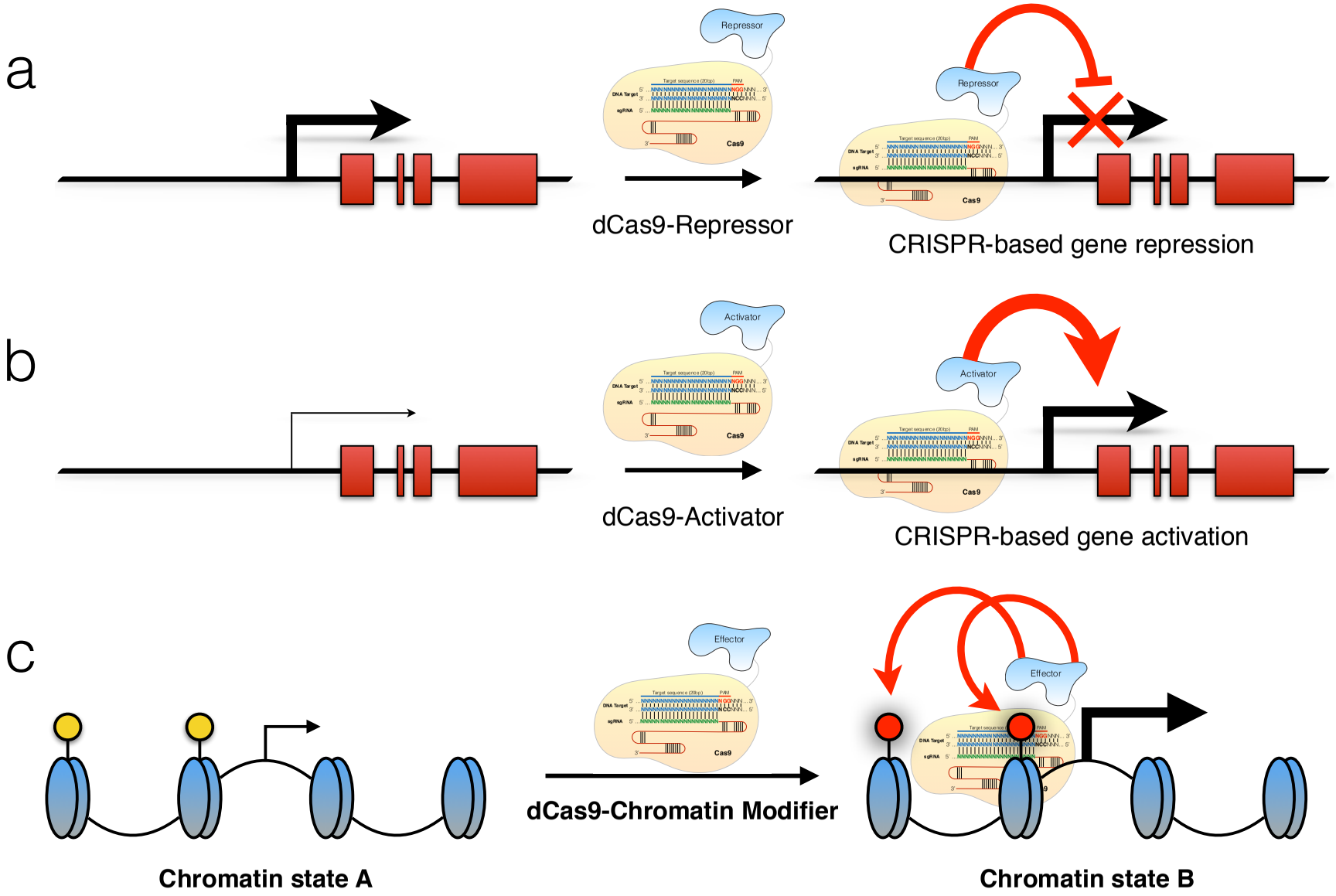

a
Repressor
Cas9
dCas9-Repressor
Repressor
CRISPR-based gene repression
b
Activator
Cas9
dCas9-Activator
Activator
CRISPR-based gene activation
c
Effector
Cas9
dCas9-Chromatin Modifier
Effector
Chromatin state A
Chromatin state B

Figure 8: Applications of dCas9-effector fusions in cancer biology. The ability of Cas9 to bind in a specific RNA-dependent fashion can be uncoupled from its nuclease activity by mutating its HNH and RuvC-like catalytic domains. This catalytically inactive form of Cas9, often referred to as dead Cas9 (dCas9), retains its RNA-guided DNA-binding activity without any detectable DNA endonuclease activity. This dCas9, in turn, can be fused to several different types of effector proteins, such as transcriptional repressors **(a)**, transcriptional activators **(b)** and chromatin modifying enzymes **(c)**. dCas9-repressors and dCas9-activators can be utilized for reversible single/multiplex repression or activation of both coding and non-coding genes, respectively. dCas9-effectors with chromatin modifying activity (c), such as histone acetyltransferases and demethylases, permit the interrogation of various epigenetic modifications in cancer biology. Other effectors, such as the FokI nuclease, have also been fused to dCas9. Adapted from Sánchez-Rivera & Jacks (2015).

functional effectors, such as chromatin modifiers that could write or erase specific epigenetic marks to interrogate the role of various epigenetic modifications in cancer biology (Kouzarides 2007). Indeed, at least two groups have successfully reported the engineering of dCas9-based epigenetic effectors for editing of the epigenome (Hilton et al., 2015, Kearns et al., 2015). The group of Charles Gersbach developed a dCas9-based histone acetyltransferase by fusing dCas9 to the catalytic core of the human acetyltransferase p300, which catalyzes H3K27 acetylation and thereby leads to activation of gene expression (Hilton et al., 2015). Notably, this dCas9-p300 fusion was able to potently activate gene expression even when targeted to distal enhancer regions (Hilton et al., 2015). Similarly, René Maehr's group developed a dCas9-based histone demethylase by fusing dCas9 to LSD1 (Kearns et al., 2015). Remarkably, this dCas9-LSD1 fusion was able to modulate gene expression when targeted to specific enhancer regions and could also be exploited for functionally characterizing novel enhancer elements (Kearns et al., 2015). Thirdly, the fact that dCas9 fusions do not permanently alter the genetic material potentially allows for reversibility of activation or inactivation of gene expression via the implementation of inducible dCas9 alleles. Fourthly, dCas9-mediated activation or repression can also be utilized to regulate the expression of long non-coding RNAs, which frequently act via *cis*-regulatory mechanisms that require local expression from its endogenous locus (see Dimitrova et al., 2014 for an example). More generally, dCas9-mediated activation of endogenous genes takes place within the context of the endogenous locus, potentially maintaining intact chromatin structure and functional regulatory sequences for both coding and non-coding genes. In addition to independent dCas9-effector fusions, the use of scaffold RNAs that encode both

targeting and effector-recruitment functions can be utilized for simultaneous multiplex gene repression and activation within a single cell (Zalatan et al., 2015, Shechner et al., 2015).

Similarly, the recent observation that truncated (also referred to as dead) sgRNAs (<15nt long) allow the catalytically active Cas9 to bind (but not cut) a genomic site of interest, and the implementation of these sgRNAs for CRISPR-mediated modulation of gene expression via sgRNA-mediated recruitment of repressors or activators should allow for even more complex manipulations of the transcriptome (Kiani et al., 2015, Dahlman et al., 2015). Therefore, I anticipate that complementary dCas9-mediated activation and repression approaches will be utilized to systematically interrogate cancer-associated genes that are frequently down regulated (putative tumor suppressor genes) or frequently overexpressed or amplified (putative oncogenes) (**Figure 8**).

B. High-throughput genetic screens using CRISPR-Cas9

Just as the field of cancer genomics has been cataloguing recurrent somatic mutations and copy number alterations frequently observed across all human cancers (Vogelstein et al., 2013), the field of functional cancer genomics has been developing tools and approaches for systematically dissecting the function of cancer genes in a highly parallel manner (Boehm & Hahn 2011). Similar to the way that RNAi technology was adapted to perform numerous high-throughput loss of function screens, both *in vitro* (Berns et al., 2004, Westbrook et al., 2005, Schlabach et al., 2008, Luo et al., 2009, Sun et al., 2011, Solimini et al., 2012, among many others) and *in vivo* (Burgess et al., 2008, Zender et al., 2008, Meacham et al., 2009, Bric et al., 2009, Zuber et al., 2011,

Possemato et al., 2011, Scuoppo et al., 2012, Gargiulo et al., 2013, Miller et al., 2013, Järås et al., 2014, Meacham et al., 2015, among many others), several groups have recently reported the implementation of the CRISPR-Cas9 system for carrying out high-throughput CRISPR screens in mammalian cells using the Cas9 nuclease (Shalem et al., 2014, Wang et al., 2014, Koike-Yusa et al., 2013, Zhou et al., 2014, Chen et al., 2015, Shi et al., 2015, Parnas et al., 2015, Birsoy et al., 2015, Wang et al., 2015) and dCas9–effectors (Gilbert et al., 2014, Konermann et al., 2014) for the systematic identification of genes involved in various biological phenotypes (**Figure 9**). For example, the Sabatini and Lander groups designed and used a library of ~73,000 sgRNAs targeting human genes to screen for genes involved in the DNA-mismatch repair pathway (MMR) in the presence of the nucleotide analogue 6-thioguanine (6-TG) and for genes whose disruption conferred resistance to the topoisomerase 2α (TOP2A) poison etoposide (Wang et al., 2014). Strikingly, both CRISPR screens demonstrated a very high signal-to-noise ratio, with the top scoring sgRNAs from each screen targeting genes involved in the MMR pathway and *TOP2A* itself, respectively. In a parallel study, Zhang and colleagues generated and screened a library of ~65,000 sgRNAs targeting human genes and successfully identified essential genes in both cancer cell lines and pluripotent stem cells (Shalem et al., 2014). Moreover, they used this library for performing a positive selection screen in melanoma cell lines to uncover genes, the deletion of which mediates resistance to the BRAF-V600E inhibitor vemurafenib, successfully identifying several known and novel candidates mediating resistance to this targeted therapy. Additional contemporaneous studies by Koike-Yusa and colleagues (Koike-Yusa et al., 2013) and Zhou and colleagues (Zhou et al., 2014) successfully

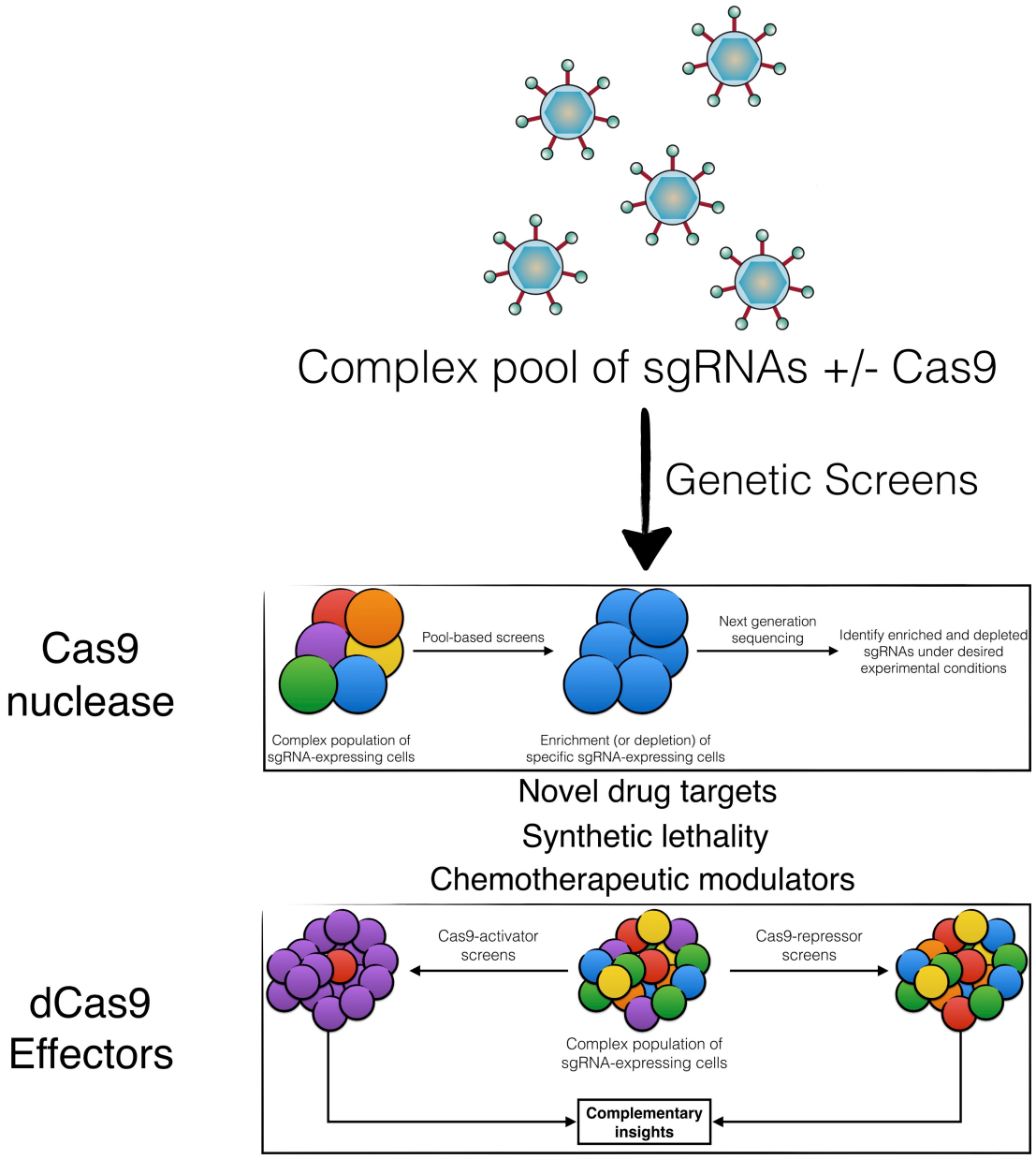


Figure 9: Genetic screens using CRISPR-Cas9. The CRISPR-Cas9 system can also be exploited for carrying out high-throughput pooled-format genetic screens using the Cas9 nuclease or dCas9 effectors for the systematic identification of genes involved in various biological phenotypes. Adapted from Sánchez-Rivera & Jacks (2015).

demonstrated the broad applicability of pooled CRISPR-based screening technologies for identifying host factors mediating toxin susceptibility in mouse embryonic stem cells and human cells, respectively. In addition to CRISPR-based screens utilizing the Cas9 nuclease, the Weissman group adapted dCas9-based activators and repressors to carry out powerful complementary genome-wide gene activation and repression screens, respectively (Gilbert et al., 2014). A subsequent study by Zhang's laboratory also demonstrated the successful adaptation of dCas9-based activators for genome-wide gene activation screens (Konermann et al., 2014). Notably, Zhang's group also demonstrated the feasibility of identifying mediators of vemurafenib resistance. In addition to these pioneering screens, the group of Christopher Vakoc recently demonstrated a novel and very powerful strategy for uncovering cancer drug targets through high-throughput screening of sgRNAs targeting protein domains (Shi et al., 2015). The power of this method lies in the fact that both in-frame and frameshift indels generated by CRISPR-Cas9 almost always lead to complete loss of function of the protein if the sgRNA is targeting a functional domain of the target protein (Shi et al., 2015). On the other hand, in-frame indels (or even frameshift indels present in late exons) that are located in non-functional domains will not necessarily ablate protein function. Therefore, CRISPR screens using domain-focused sgRNA libraries might be more powerful than those using traditional sgRNA libraries. These landmark studies demonstrated the feasibility of carrying out pooled high-throughput screens using CRISPR–Cas9 technologies to uncover genes mediating numerous biological phenotypes, as well as cancer cell vulnerabilities and mechanisms of therapeutic response and resistance. Recent work from the Sabatini and Lander groups

demonstrated the power of CRISPR-based genetic screens even further by employing a 2nd generation human library to uncover the entire set of essential genes in the mammalian genome, as well as cancer cell-specific vulnerabilities (Wang et al., 2015). In addition to *in vitro* screens, the Zhang and Sharp groups recently demonstrated the utility of the CRISPR–Cas9 system for performing genome-wide *in vivo* screens to uncover genes involved in tumor progression and metastasis (Chen et al., 2015).

I envision that both types of CRISPR-based genetic screens utilizing the Cas9 nuclease and dCas9-based activators or repressors, and combinations thereof will rapidly transform the field of functional cancer genomics. For example, both platforms can be systematically deployed to identify synthetic lethal interactions (Guarente, 1993, Kaelin 2005) by carrying out genome-wide screens across several different cancer types utilizing multiple cell lines that harbor a particular genotype of interest (such as those harboring mutant alleles of Kras). In this scenario, CRISPR-mediated suppression or activation of specific genes could lead to selective killing of tumor cells expressing a particular oncogene. In addition to uncovering synthetic lethal interactions with specific oncogenes across multiple cancer types using pre-existing cell lines, CRISPR-based screens can be implemented to identify genotype-selective vulnerabilities by engineering isogenic cell lines harboring specific cancer genotypes frequently observed in specific subtypes of human cancer. In this scenario, I envision the engineering of isogenic cell lines harboring different cancer-associated genotypes via CRISPR-based genome editing and subsequently uncovering genotype-selective vulnerabilities via high-throughput CRISPR-based genetic screens or small molecule screens.

Beyond the aforementioned *in vitro* applications, the CRISPR-Cas9 system holds great promise for modeling and dissecting cancer *in vivo*. Therefore, I will focus below on the utility of this technology for generating animal models for the study of cancer genes *in vivo*.

C. Rapid generation of mouse models

GEMMs (Frese and Tuveson 2007) and non-germline GEMMs (nGEMMs) (Heyer et al., 2010) of cancer have played a critical role in uncovering several fundamental aspects of tumor initiation, maintenance and progression. In addition, they have emerged as faithful models with which to test a variety of anti-cancer agents, as well as for uncovering mechanisms of drug resistance (Engelman et al., 2008, Chen et al., 2012). However, generating GEMMs is a slow and expensive process, requiring complex ES cell manipulation and/or pronuclear injection, as well as extensive mouse husbandry to obtain animals harboring the alleles of interest (Heyer et al., 2010). nGEMMs of cancer can simplify this process by bypassing the need for complex genetic crosses through the serial re-targeting of ES cells (Heyer et al., 2010). Nevertheless, the inability to simultaneously introduce multiple genetic modifications in mice or ES cells remains a considerable barrier.

Rudolf Jaenisch and colleagues have recently demonstrated that the CRISPR-Cas9 system can be utilized to simultaneously disrupt up to eight alleles in mouse ES cells in a single step (Wang et al., 2013). Furthermore, they reported efficient simultaneous disruption of two genes in single-cell mouse embryos and the subsequent one-step generation of double knockout animals (Wang et al., 2013). This group also

demonstrated efficient simultaneous HDR-mediated genome editing of two endogenous genes (Wang et al., 2013). In a subsequent study, they extended their CRISPR-Cas9 methods for rapidly generating mice carrying conditional Cre/loxP-based alleles and reporter alleles, as well as using pairs of sgRNAs to generate mice carrying small deletions (Yang et al., 2013) (**Figure 6**). These studies have demonstrated the ease with which ES cells or mice harboring multiple GOF and LOF mutations can be generated, an advance that has opened the door for the development of novel GEMMs and nGEMMs of cancer with unprecedented speed and precision. Indeed, I predict that there will be an explosion of novel GEMMs and nGEMMs harboring uniquely complex genetic alterations that will allow for detailed analysis of several stages of tumor evolution with unprecedented speed and efficiency (**Figure 10 and Figure 11**). For example, CRISPR-mediated engineering will allow for rapid generation of large repositories of ES cell lines harboring multiple combinations of constitutive or conditional mutations in oncogenes and tumor suppressor genes, as well as large chromosomal rearrangements that will capture some of the genetic heterogeneity that is characteristic of human cancer genomes. These ES cell lines can be utilized to generate GEMMs and nGEMMs of cancer harboring multiple distinct mutant genotypes, which will be highly valuable for testing new therapeutic regimens and for personalized oncology efforts.

It is important to note that the majority of mouse cancer models have been based on a rather limited number of mutant genes or alleles, such as the G12D or G12V mutations in the *Kras* oncogene (Jackson et al., 2001, Guerra et al., 2003). The CRISPR-Cas9 system will allow for systematic generation of models harboring multiple oncogenic

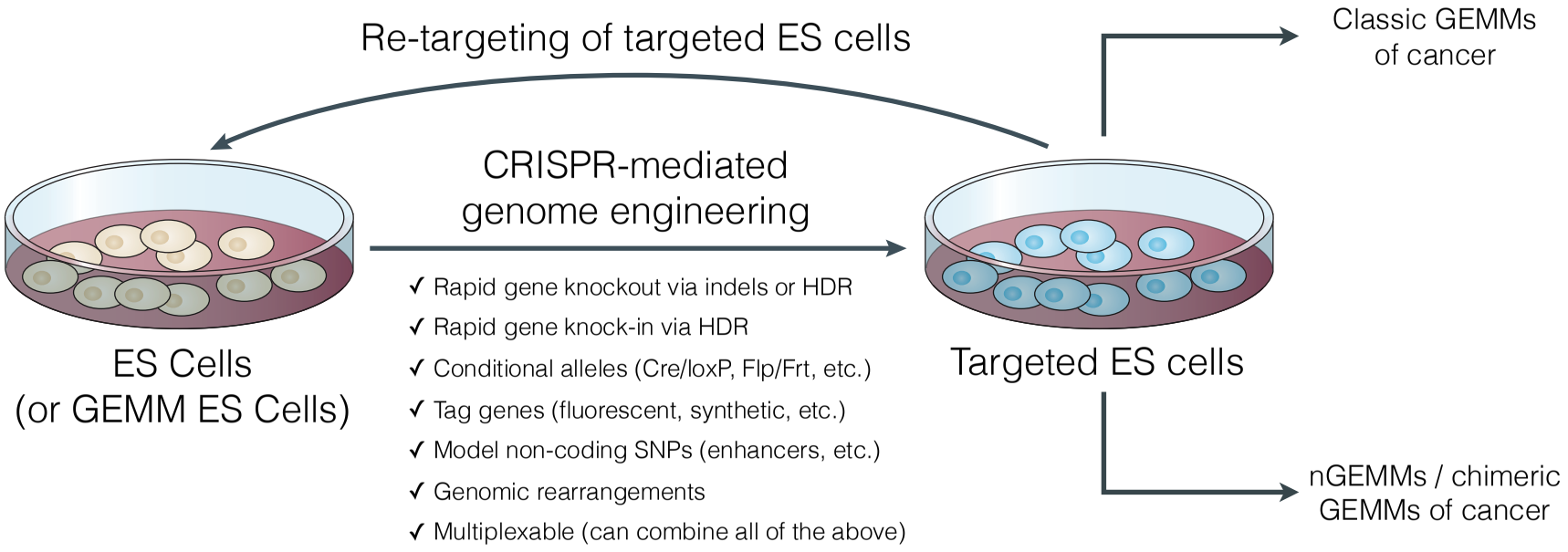


Figure 10: CRISPR-mediated genome engineering of embryonic stem (ES) cells or GEMM-derived ES cells for rapid generation of GEMMs and nGEMMs of cancer. The CRISPR-Cas9 system can be utilized for rapidly generating novel GEMMs or non-germline GEMMs (nGEMMs) of cancer harboring multiple genetic alterations. These alterations include constitutive or conditional knockout and knock-in alleles, endogenous synthetic tags or reporters, non-coding single-nucleotide polymorphisms (SNPs) and genomic rearrangements, as well as a combination of all of these via re-engineering of ES cells or multiplex CRISPR-mediated genome engineering. Adapted from Sánchez-Rivera & Jacks (2015).

alleles, making it possible to investigate allele-specific consequences in tumor progression and therapeutic response. Highly systematic and multiplexable approaches for HDR-mediated editing of specific genomic regions, such as the methods developed by the Shendure laboratory (Findlay et al., 2014), will facilitate rapid analysis of 'hotspot' regions with various combinations of mutations and subsequent generation of GEMMs and nGEMMs.

Beyond new model development, the CRISPR-Cas9 system can also be used to refine existing models of cancer. ES cell lines derived from well-studied GEMMs can be readily reengineered to harbor additional constitutive or conditional mutant alleles of oncogenes and tumor suppressor genes (reviewed in Dow and Lowe 2012) (**Figure 10**). Thus, candidate cooperating mutations can be easily studied and putative synthetic lethal interactions can be validated. Moreover, this approach will allow for pre-clinical studies consisting of cohorts of mice that better represent the genetic heterogeneity of human cancers (**Figure 11**). One can even envision combining comprehensive genomic characterization of tumors from individual patients with the rapid generation of personalized GEMMs, nGEMMs or cell-based xenografts. *In vivo* models carrying the exact complement of driver mutations from a given patient's tumor could then be screened with conventional or experimental anti- cancer agents to identify the most effective therapies.

D. Somatic genome engineering

As outlined above, the efficiency of genome editing by CRISPR-Cas9 makes the process of germline and ES cell line genetic manipulation more rapid and more

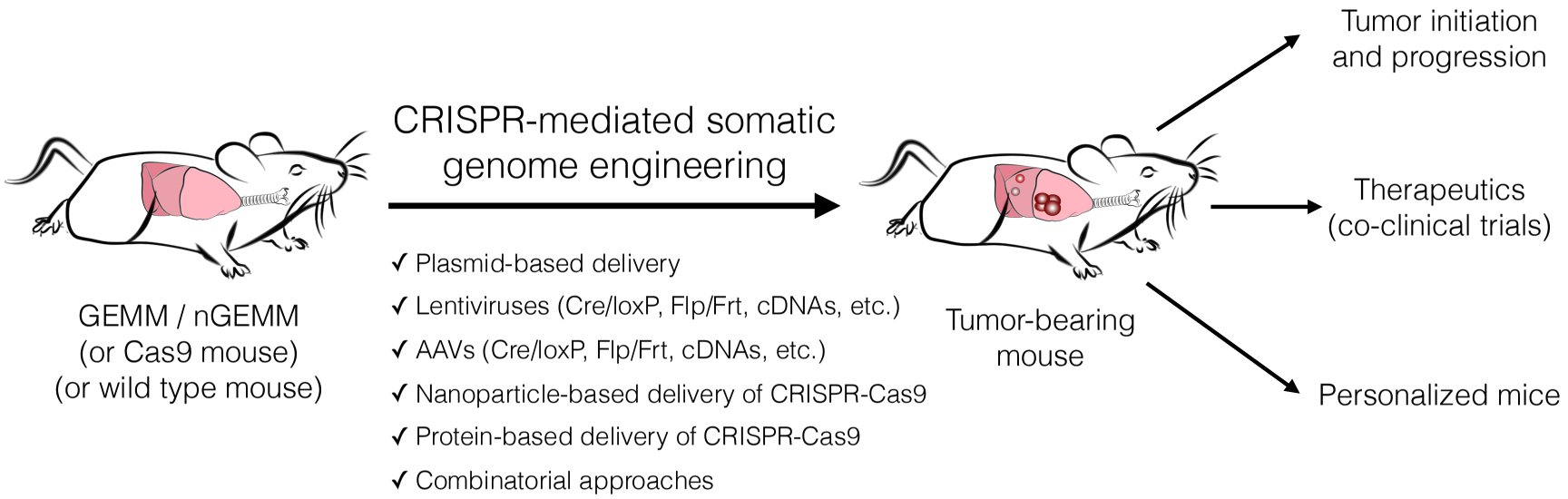


Figure 11: CRISPR-mediated somatic genome engineering *in vivo* for modeling cancer. CRISPR-mediated somatic genome engineering *in vivo* can be used to rapidly generate cohorts of tumor-bearing mice for studying both basic and translational aspects of cancer biology. For example, the CRISPR–Cas9 system can be deployed *in vivo* for generating cohorts of personalized mice based on the exact complement of mutations observed in individual patients. Adapted from Sánchez-Rivera & Jacks (2015).

powerful. The power of the system is even more evident in the ability to perform somatic genome editing *ex vivo* and *in vivo*.

i. *Ex vivo* CRISPR-based somatic genome editing

Three recent studies have demonstrated the power of CRISPR-based *ex vivo* somatic genome editing for rapid modeling of cooperating mutations and the generation of mouse models of hematopoietic malignancies (Malina et al., 2013, Chen et al., 2014, Heckl et al., 2014). The Pelletier group demonstrated efficient *ex vivo* editing of the *Trp53* tumor suppressor gene in *Arf*$^{-/-}$Eμ-*Myc* lymphomas that were subsequently transplanted into syngeneic mice to show that *Arf*$^{-/-}$Eμ-*Myc* cells lacking p53 are substantially enriched upon treatment with doxorubicin (Malina et al., 2013). Using a similar approach, the Lowe laboratory utilized *ex vivo* CRISPR-mediated disruption of the *Mll3* (also known as *Kmt2c*) tumor suppressor gene in *shNf1*;*Trp53*$^{-/-}$ primary mouse hematopoietic stem and progenitor cells (HSPCs) to demonstrate that *Mll3* is a haploinsufficient tumor suppressor in acute myeloid leukemia (AML) (Chen et al., 2014). Strikingly, the vast majority of AML clones sequenced harbored indels in only one of the two *Mll3* alleles, strongly suggesting a role as a haploinsufficient tumor suppressor gene in AML and demonstrating the power of the CRISPR-Cas9 system beyond simple gene editing. The Ebert group employed the CRISPR-Cas9 system to rapidly generate mouse models of AML by lentiviral-mediated *ex vivo* editing of single or multiple genes in primary mouse hematopoietic stem and progenitor cells (Heckl et al., 2014. Subsequent transplantation of engineered HSPCs into lethally irradiated wild type mice led to rapid generation of novel mouse models of AML with simultaneous disruption of

up to five genes. These three studies highlight the potential of the CRISPR-Cas9 system for *ex vivo* somatic genome editing of primary cells, which can be further exploited for the rapid generation of mouse models of a variety of human malignancies (**Figure 10 and Figure 11**).

ii. *In vivo* CRISPR-based somatic genome editing

To explore the use of the CRISPR-Cas9 system for directly mutating genes in living animals, we utilized hydrodynamic gene transfer to simultaneously deliver plasmids encoding Cas9 and sgRNAs targeting the *Pten* and *Trp53* tumor suppressor genes to hepatocytes *in vivo* (Xue et al., 2014). Strikingly, delivery of these CRISPR plasmids to the hepatocytes of adult wild-type mice was sufficient to induce liver tumors with identical histopathology to those observed in $Pten^{fl/fl};Trp53^{fl/fl}$ GEMMs, in which tumors were initiated via delivery of adenoviruses expressing Cre recombinase. These results strongly suggest that CRISPR-mediated somatic genome editing of cancer genes in adult wild-type mice can efficiently substitute for traditional GEMMs, at least for some cancer types. Moreover, we further demonstrated the feasibility of using the CRISPR-Cas9 system to engineer gain of function mutations in the livers of adult wild-type mice via the co-delivery of CRISPR components and a single-stranded DNA template encoding a mutant form of β-catenin, which resulted in the generation of hepatocytes with nuclear β-catenin at a low (0.5%) but detectable frequency (Xue et al., 2014).

Moving beyond the liver, we also developed an all-in-one lentivirus simultaneously encoding CRISPR components and Cre recombinase. This vector was used to mutate three lung cancer tumor suppressor genes in the developing tumors of the well

established *Kras*$^{LSL\text{-}G12D/+}$ and *Kras*$^{LSL\text{-}G12D/+}$*;Trp53*$^{flox/flox}$ GEMMs of lung cancer (Jackson et al., 2001, Jackson et al., 2005, Sánchez-Rivera et al., 2014). Intratracheal delivery of all-in-one lentiviruses expressing sgRNAs targeting a panel of tumor suppressor genes into *Kras*$^{LSL\text{-}G12D/+}$ or *Kras*$^{LSL\text{-}G12D/+}$*;Trp53*$^{flox/flox}$ mice resulted in lung adenocarcinomas with diverse histopathological and molecular features that depended on the tumor suppressor gene targeted. Moreover, a large fraction of the lung tumors harbored indels in predicted sites within the target genes with no detectable off-target editing, strongly supporting Cas9 on-target activity for somatic genome editing *in vivo*. In a parallel study, Andrea Ventura's group demonstrated the feasibility of using the CRISPR-Cas9 system for modeling large oncogenic chromosomal rearrangements (**Figure 7**) in wild-type mice *in vivo* via delivery of an adenovirus encoding Cas9 and two sgRNAs designed to induce an *Eml4-Alk* (echinoderm microtubule associated protein like 4-anaplastic lymphoma kinase) inversion (Maddalo et al., 2014 and Soda et al., 2007). Lung tumors developed with complete penetrance and were exquisitely sensitive to crizotinib, an inhibitor used to treat human lung tumors that harbor this particular oncogenic rearrangement (Shaw et al., 2013). Moreover, the fact that the *Eml4* and *Alk* loci are separated by ~11 megabases strongly supports the feasibility of the CRISPR-Cas9 system for modeling large genomic rearrangements. A subsequent study by Roberto Chiarle's group utilizing lentiviruses also demonstrated the ability of the CRISPR-Cas9 system to induce chromosomal rearrangements *in vivo* (Blasco et al., 2014). These studies demonstrated the potential of rapidly generating mouse models of cancer via somatic genome engineering through delivery of all CRISPR components in the form of plasmids or viruses. In addition to these traditional DNA- or viral-based delivery

methods, recent advances in engineering of non-viral delivery materials have made it possible to deliver Cas9-sgRNA protein-RNA complexes (Zuris et al., 2015) and sgRNA-nanoparticle complexes (Platt et al., 2014) *in vivo* utilizing cationic lipid-mediated delivery or 7C1 nanoparticles, respectively. Future advances in materials science and engineering should make it possible to implement additional types of non-viral delivery platforms for the delivery of Cas9, sgRNAs and HDR donor DNA templates for achieving highly efficient genome modification *in vivo* (non-viral materials extensively reviewed in Yin et al., 2014).

To further streamline the generation of CRISPR-based somatic mouse models of cancer, the Zhang and Sharp laboratories reported the generation of mouse models expressing constitutive or Cre-inducible versions of the Cas9 enzyme (Platt et al., 2014). By intratracheally delivering a novel adeno-associated virus (AAV) encoding six components: a $Kras^{G12D}$ HDR donor DNA template, sgRNAs targeting *Kras*, serine/threonine kinase 11 (*Stk11*; also known as *Lkb1*) and *Trp53*, Cre recombinase and *Renilla* luciferase into mice expressing the Cre-inducible *Cas9* allele, they were able to induce lung tumors in adult mice by simultaneously disrupting both tumor suppressors and engineering the oncogenic $Kras^{G12D}$ mutation. Intriguingly, the absolute co-occurrence of these three events was very low, with the frequency of $Kras^{G12D}$ HDR events consistently being under 2% with the exception of one outlier tumor. Unexpectedly, the vast majority of the resulting tumors harbored indels in both *Lkb1* and *Trp53*, which lead the authors to hypothesize that *Lkb1;Trp53* double-mutant tumors outcompete $Kras^{G12D}$*;Lkb1;Trp53* triple-mutant tumors (Platt et al., 2014).Recently, the Lowe laboratory reported the generation of a highly flexible mouse modeling platform

consisting of transgenic mice co-expressing doxycycline-inducible alleles of *Cas9* or the *Cas9*^{D10A} nickase variant (Ran et al., 2013) and constitutively expressed sgRNA cassettes (Dow et al., 2015). Utilizing this conditional platform, they demonstrated effective gene editing *in vivo* with up to 85% target gene modification. Moreover, they demonstrated efficient simultaneous biallelic modification of up to two genes *in vivo* using a pair of sgRNAs and the Cas9 nuclease. This flexible platform allowed them to accommodate up to six sgRNA cassettes that, when combined with the Cas9^{D10A} nickase, led to simultaneous editing of three genes in mouse ES cells with high efficiency.

The development of mouse models expressing the Cas9 nuclease and Cas9^{D10A} nickase represents a major advancement for CRISPR applications in cancer biology, allowing researchers to focus their efforts on delivering single or multiple sgRNAs with or without synthetic HDR donor DNA templates utilizing viral and/or non-viral carriers, bypassing the need to optimize approaches for co-delivery of this large DNA endonuclease. In addition, expression of Cre-inducible or doxycycline-inducible alleles of Cas9 *in vivo* can be rendered tissue-specific via the incorporation of tissue-specific Cre or reverse tetracycline transactivator alleles, respectively. Moreover, the development of constitutive and conditional mouse models for CRISPR-mediated activation (Tanenbaum et al., 2014) or repression (Gilbert et al., 2013) of gene expression (**Figure 8**) will serve as powerful complementary approaches for functionally studying both coding and non-coding DNA elements without permanent disruption of the endogenous genomic sequence. Beyond the established *Mus musculus* laboratory organism, the flexibility of CRISPR-Cas9 technologies should allow for rapid generation

of novel animal models of cancer utilizing genetically intractable organisms that better recapitulate human tissue architecture and drug metabolism, such as pigs (Whitworth et al., 2014) and non-human primates (Niu et al., 2014).

E. Synopsis and outlook

We envision a new era in cancer biology in which CRISPR-based genome engineering will serve as an important conduit between the bench and the bedside (**Figure 12**). The successful deployment of sophisticated genetic and molecular profiling technologies for comprehensive characterization of a patient's tumor is generating detailed roadmaps to instruct the development of tailored cell-based or whole animal-based experimental systems. These systems will serve as personalized platforms, with which researchers will rapidly and systematically identify genotype-specific vulnerabilities and synthetic lethal interactions via single or multiplex CRISPR-based and small molecule-based approaches. Moreover, such personalized platforms could be studied in parallel to the patients, potentially allowing for the rapid identification of resistance mechanisms and the development of strategies to overcome such shortcomings (Crystal et al., 2014).

Although there are current technical limitations to the use of CRISPR-Cas9 for targeting cancer genes in human patients as a therapeutic strategy, the prospects of this form of gene therapy are nonetheless very exciting. Recent work has demonstrated the potential of this technology to permanently correct genetic mutations *in vivo* in the adult liver of mouse models of a hereditary genetic disease via HDR, successfully alleviating aspects of the disorder (Yin et al., 2014). Therefore, future advancements of this technology for increasing the efficiency of editing and delivery of CRISPR-Cas9

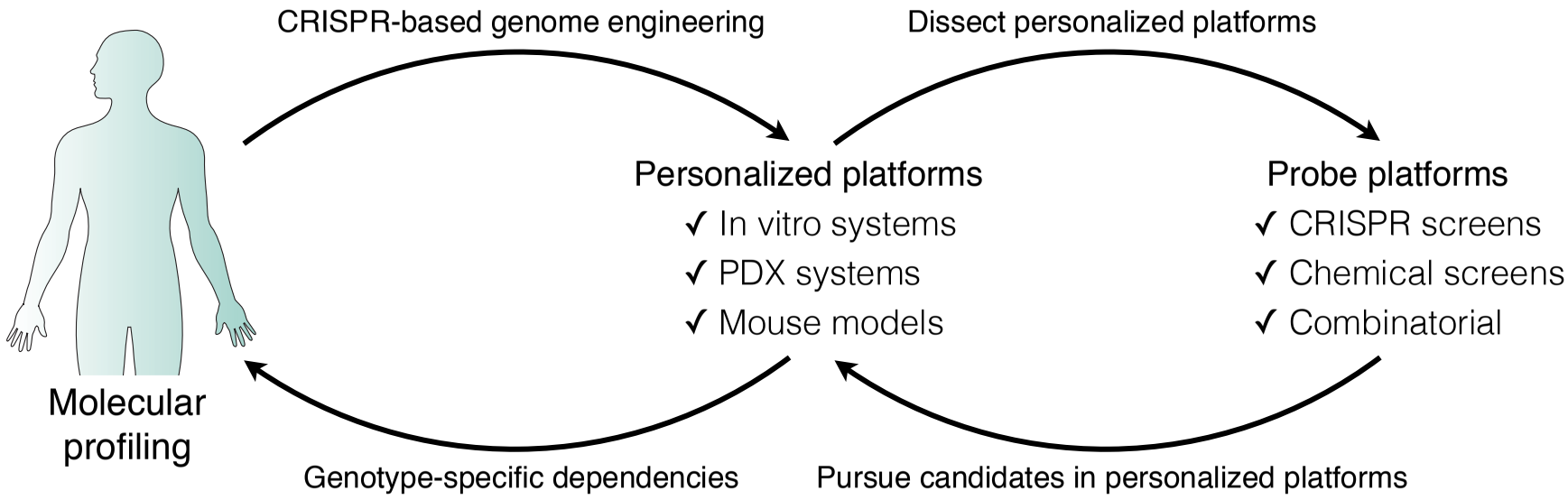


Figure 12: The CRISPR–Cas9 system can serve as an important conduit between the bench and the bedside. CRISPR-mediated somatic genome engineering *in vivo* can be used to rapidly generate cohorts of tumor-bearing mice for studying both basic and translational aspects of cancer biology. For example, the CRISPR–Cas9 system can be deployed *in vivo* for generating cohorts of personalized mice based on the exact complement of mutations observed in individual patients. Adapted from Sánchez-Rivera & Jacks (2015).

components utilizing both viral and non-viral delivery vehicles will allow for therapeutic genetic correction of single or multiple driver mutations. In addition to permanently correcting cancer-associated mutations, the CRISPR-Cas9 system could be employed for precise *ex vivo* engineering of immune cells for immunotherapeutic applications. For example, the CRISPR-Cas9 system could be utilized for the development of novel chimeric antigen receptor (CAR)-modified T cells (Sadelain et al., 2013), in which the CAR is precisely inserted into a safe harbor locus (Sadelain et al., 2012).

Ever since the Doudna and Charpentier groups demonstrated the potential of the CRISPR-Cas9 system as a powerful RNA-programmed genome editing platform (Jinek et al., 2012), the field of genome engineering has rapidly undergone a scientific revolution that promises to transform nearly every aspect of basic biological and biomedical research. The application of this technology to several aspects of cancer biology, ranging from basic research to clinical and translational applications, offers numerous exciting opportunities for better understanding and potentially treating this devastating disease.

Chapter 1 – Introduction

Part III – Cancer genomics

I - Historical Perspective

Over a hundred years ago, Theodor Boveri postulated in his seminal essay "*Concerning the origin of malignant tumors*" that cancer was caused by chromosomal aberrations that arise from a failure of cells to undergo proper cell division (Boveri 1902, Boveri 2007 (translation by Henry Harris)). That cancer was indeed a genetic disease - that is, a disease that arises from defects in the genome of incipient cancer cells - would be demonstrated several decades later through the seminal studies of G. Steven Martin, Michael Bishop, Harold Varmus, Robert Weinberg, Michael Wigler, Mariano Barbacid, and many of their colleagues (Martin 1970, Stehelin et al., 1976, Tabin et al., 1982, Parada et al., 1982, Reddy et al., 1982, Taparowsky et al., 1982, Friend et al., 1986, among other studies). As discussed extensively in **Part I** of this introduction, the existence of two main classes of cellular genes (oncogenes and tumor suppressor genes) whose mutation was causally related to cancer suggested to many researchers in the field that sequencing the entire genomes of both normal and cancer cells would lead to the comprehensive identification and cataloguing of the entire repertoire of oncogenes and tumor suppressor genes in the human genome. Therefore, the groundbreaking studies that led to the discovery of viral oncogenes and cellular proto-

oncogenes and tumor suppressor genes was undoubtedly one of the main catalyzers behind the launching and completion of the Human Genome Project (HGP) (Lander et al., 2001, Venter et al., 2001, IHGSC 2004). The successful completion of the HGP positively impacted biological research at several levels. First and foremost, it produced a highly reliable map that covered ~ 99.7% of the sequence of the euchromatic human genome (IHGSC 2004, Lander 2011). In addition, the HGP substantially accelerated the development and implementation of novel, state of the art DNA sequencing technologies and bioinformatic approaches that had an enormous impact in the post-HGP era. The existence of a reference sequence of the human genome, against which novel mutant alleles of oncogenes and tumor suppressor genes could be compared to, combined with the availability of emerging high-throughput sequencing technologies led to the development and establishment of the field of “cancer genomics”. Due to the fact that a comprehensive discussion of the field of cancer genomics is beyond the scope of the introduction of this thesis, I opted to limit my discussion below to four case studies that illustrate the power and utility of cancer genomics (for excellent reviews, see Stratton et al., 2009, Garraway and Lander 2013, Wheeler and Wang 2013 and Vogelstein et al., 2013)

II - Pioneering cancer genome sequencing studies

A. *BRAF* mutations in melanoma

One of the first studies that firmly established the utility of large scale cancer genome sequencing studies was carried out by the Sanger Institute back in 2002 (Davies et al., 2002). In this pioneering study strictly focused on MAPK pathway genes, the authors

sequenced a collection of 530 specimens spanning several cancer types, such as breast cancer, colorectal cancer, lung cancer, and melanoma, among several others. Notably, they identified somatic missense mutations in *BRAF* in ~ 66% of malignant melanomas, an event that triggered the subsequent development of *BRAF* inhibitors for the treatment of this disease (Davies et al., 2002, Flaherty et al., 2010, Chapman et al., 2011, Holderfield et al., 2014). Nowadays, *BRAF* inhibitors, such as vemurafenib, are routinely used in the clinic to treat patients with *BRAF*-mutant metastatic melanoma (Bollag et al., 2012, Flaherty 2012).

B. *PIK3CA* mutations in colorectal cancer

Another pioneering study was carried out by the group of Victor Velculescu along with Bert Vogelstein, Kenneth Kinzler, and colleagues in 2004 (Samuels et al., 2004). In this study strictly focused on determining whether human cancers harbored mutations in the phosphatidylinositol 3-kinase (PI3K) family of genes, the authors sequenced a total of 234 colorectal cancer specimens and found that the *PIK3CA* gene (which encodes the p110α catalytic subunit of PI3K) was mutated in ~ 32% of the cases. Notably, they extended their p110α sequencing efforts to uncover *PIK3CA* mutations in ~ 27% of glioblastomas, 25% of gastric cancers, 8% of breast cancers, and 4% of lung cancers (Samuels et al., 2004). The identification of these oncogenic *PIK3CA* mutations in multiple tumor types stimulated the scientific community and the private sector to develop PI3K pathway inhibitors, many of which are currently being tested as single agents or in combination with other agents in multiple clinical trials (Thorpe et al., 2015).

C. *EGFR* mutations in lung cancer

In the late 1990s and early 2000s, tyrosine kinase inhibitors targeting the epidermal growth factor receptor (EGFR) were being actively pursued in clinical trials due to the fact that over 60% of non-small cell lung cancer (NSCLC) cases were characterized by *EGFR* overexpression (Ohsaki et al., 2000, Hirsch et al., 2003). Nevertheless, partial responses were being observed in only ~ 10-30% of cases (Fukuoka et al., 2003, Kris et al., 2003). The nature of these "exceptional responders" was subsequently elucidated by the groups of Daniel Haber, Matthew Meyerson, and Harold Varmus when they discovered that NSCLC patients harboring mutations in the kinase domain of *EGFR* were exquisitely sensitive to the tyrosine kinase inhibitors gefitinib or erlotinib (Lynch et al., 2004, Paez et al., 2004, Pao et al., 2004). The fact that these very specific mutations conferred exquisite sensitivity to a particular targeted therapy served as another example (in addition to the classical example of the *BCR-ABL* translocation conferring exquisite sensitivity to Gleevec in chronic myelogenous leukemia - see Sawyers 2003 for an excellent contemporaneous review) of the potential of so-called personalized or precision oncology approaches (Garraway 2013).

D. *IDH1/2* mutations in brain cancer and acute myeloid leukemia

In a remarkable tour de force approach, Victor Velculescu, Bert Vogelstein, Kenneth Kinzler, and colleagues sequenced 20,661 protein coding genes across 22 human glioblastoma multiforme (GBM) tumor samples and identified recurrent mutations tightly clustered around the active site of the metabolic enzyme isocitrate dehydrogenase 1 (*IDH1*) in 12% of the cases analyzed (Parsons et al., 2008). Specifically, all *IDH1*

mutations resulted in an amino acid substitution at position 132 of the enzyme - a residue that is highly conserved across evolution. A follow up study substantially expanded the number of tumor samples to over 400 central nervous system (CNS) tumor samples, including 138 primary GBM tumors and 13 secondary GBM tumors, as well as almost 500 non-CNS tumor samples, including solid tumors of the lung, breast, and pancreas, among others (Yan et al., 2009). Strikingly, more than 70% of GBM tumors, oligodendrogliomas, and high-grade astrocytomas (but 0% of the non-CNS tumor samples examined in this study) harbored mutations in *IDH1* at position 132 or, unexpectedly, at the analogous residue (R172) of its mitochondrial homolog - *IDH2* (Yan et al., 2009). A subsequent study by the group of Timothy Ley uncovered *IDH1* mutations in about 16% of acute myeloid leukemia (AML) patients (Mardis et al., 2009). A series of subsequent studies by Agios Pharmaceuticals and by Craig Thompson's research group demonstrated that these mutations were actually neomorphic - that is, they completely changed the enzymatic activity of IDH1 and IDH2 (Dang et al., 2009, Ward et al., 2010). Specifically, mutant IDH1 and IDH2 were catalyzing the conversion of isocitrate to the "oncometabolite" 2-hydroxyglutarate (2HG) instead of converting it to α-ketoglutarate (α-KG). Remarkably, subsequent studies demonstrated that the 2HG oncometabolite directly impinged on the epigenetic landscape of cancer cells via the simultaneous inhibition of several α-KG-dependent enzymes, including histone demethylases and 5-methylcytosine hydroxylases (Xu et al., 2011, Lu et al., 2012). Recent studies by multiple groups have demonstrated that cancer-associated *IDH1* and *IDH2* mutations are indeed functionally oncogenic (Losman et al., 2013, Chen et al., 2013, Kats et al., 2014, Hirata et al., 2015) and that therapeutic targeting of these

mutant enzymes is a promising approach for treating the malignancies that they promote (Losman et al., 2013, Rohle et al., 2013, Wang et al., 2013). The fact that such an unexpected cancer driver and therapeutic target was identified through unbiased cancer genomics further supported the substantial expansion of these approaches to dozens of other human cancers.

E. Synopsis and outlook

The last 10-15 years have witnessed an explosion of large scale cancer genome sequencing studies. The creation of multi-institutional consortia, such as The Cancer Genome Atlas (TCGA), the International Cancer Genome Consortium (ICGC), and the Cancer Genome Project (CGP) have led to the execution and completion of > 100 whole cancer exome or whole cancer genome sequencing studies (these studies are catalogued in the supplementary material of Vogelstein et al., 2013). These large scale multi-institutional cancer genome sequencing studies have revealed a multitude of recurrent mutations and copy number alterations in several types of human cancer, including colorectal cancer (CRC TCGA 2012), breast cancer (Breast TCGA 2012), pancreatic cancer (Waddell et al., 2015), gastric adenocarcinoma (Gastric TCGA 2014), ovarian carcinoma (Patch et al., 2015), prostate adenocarcinoma (Abeshouse et al., 2015, Robinson et al., 2015), melanoma (Akbani et al., 2015), glioma (Glioma TCGA 2015), squamous cell carcinoma of the lung (Hammerman et al., 2012), small cell lung cancer (George et al., 2015) and lung adenocarcinoma (LUAD TCGA 2014), as well as several types of hematological cancers, including acute myeloid leukemia (AML TCGA 2013), acute lymphoblastic leukemia (Andersson et al., 2015), multiple myeloma (Lohr

et al., 2014) and diffuse large B-cell lymphoma (Pasqualucci et al., 2011), among many others.

The current consensus is that the cancer genome sequencing studies carried out so far have identified ~ 125 driver genes (of which ~ 71 appear to be tumor suppressor genes and ~ 54 appear to be oncogenes) that can be classified into ~ 12 signaling pathways that are invariably centered around core biological processes involved in cellular fate determination, cellular survival, and maintenance of genome integrity (Vogelstein et al., 2013). Despite the fact that this suggests that most strong driver genes have already been identified, recent stringent bioinformatic analyses have proposed that this is likely a major underestimate. In fact, Stephen Elledge and colleagues have recently estimated that there are ~ 320 tumor suppressor genes and ~ 250 oncogenes, suggesting that many more drivers remain to be discovered (Davoli et al., 2013). Importantly, Elledge and colleagues suggested that these additional drivers are likely to be less potent than the main drivers that have already been discovered. Nevertheless, they postulate that the combined inactivation or overactivation of these less potent tumor suppressor genes and oncogenes, respectively, can have major consequences during tumorigenesis. In fact, they demonstrated that a large proportion of these additional drivers are physically linked, such that catastrophic genomic events that trigger chromosomal amplifications or deletions can lead to cumulative triplosensitivity and cumulative haploinsufficiency, respectively. In this scenario, the combined overrepresentation or underrepresentation of less potent oncogenes and tumor suppressor genes can have a major cumulative effect in the context of tumorigenesis (Davoli et al., 2013). This study sheds important light on the evolutionary reasons behind the pervasive amount of somatic copy number

alterations (be it focal or whole chromosomal amplifications or deletions) present in some tumor types (Beroukhim et al., 2010, Gordon et al., 2012). Moreover, it supports a very important conclusion - oncogenes and tumor suppressor genes need not be exclusively activated or inactivated, respectively, via mutation in their coding sequence *per se*; rather, their overrepresentation and underrepresentation via the gain or loss of chromosomal regions, respectively, can have a major effect during tumorigenesis.

The fact that there are likely many more drivers than anticipated raises the following question: when is the catalog of driver mutations present in all human cancers going to be complete? Recent estimates suggest that this will be completed in the not too distant future. In fact, estimates computed by Eric Lander, Gad Getz, and colleagues suggest that this can be achieved by sequencing 600 - 5,000 samples per tumor type (Lawrence et al., 2014). Therefore, developing and applying novel approaches aimed at functionally characterizing this growing list of putative drivers in appropriate experimental models of cancer is of utmost importance in order to fulfill the promise of cancer genomics.

In **Chapter 2** of this thesis, I describe a novel CRISPR-Cas9-based approach for functionally validating novel cancer drivers *in vivo* using well established pre-clinical GEMMs of cancer. Due to the fact that this system was initially established and validated in pre-clinical GEMMs of lung cancer, I will devote the next section of this part of the introduction to a brief description of the past, present, and future of lung cancer genomics.

III - Lung cancer genomics

A. Introduction to lung cancer

Lung cancer is the leading cause of cancer-related deaths in the United States and worldwide (Jemal et al., 2008). This cancer type is typically classified as either non-small cell lung cancer (NSCLC - accounting for ~ 85% of all cases) or small-cell lung cancer (SCLC - accounting for ~ 15% of all cases) (Herbst et al., 2008, Chen et al., 2014). NSCLC can be further divided into three major histological subtypes: adenocarcinoma of the lung, squamous-cell carcinoma of the lung, and large-cell lung cancer (Herbst et al., 2008, Sun et al., 2007). Adenocarcinoma of the lung is the most common histological subtype (Herbst et al., 2008). Notably, this histological subtype is also the most common in patients that are non-smokers, accounting for over 300,000 deaths every year (Sun et al., 2007).

B. Genetics and genomics of lung adenocarcinoma

The aforementioned technological advances that came to light after the completion of the HGP have been extensively applied to the study of lung cancer. In fact, several large scale genomic and molecular studies over the last decade have extensively contributed to the comprehensive molecular characterization of the major lung cancer subtypes, including NSCLC (Weir et al., 2007, Ding et al., 2008, Imielinski et al., 2012, Govindan et al., 2012, Hammerman et al., 2012, LUAD TCGA 2014) and SCLC (Pleasance et al., 2010, Peifer et al., 2012, Rudin et al., 2012).

Past and present efforts focused on adenocarcinomas of the lung have generated a continuously expanding list of genetic drivers, many of which remain largely uncharacterized due to the absence of adequate experimental approaches and model systems (for a current catalog of the most frequent mutations, see LUAD TCGA 2014 and Devarakonda et al., 2015). Below, I will briefly describe the current catalog of lung adenocarcinoma mutations, many of which have been discovered or re-discovered through large scale cancer genome sequencing studies. I will then briefly discuss *KRAS* and *TP53* mutations, which form the basis of the K and KP models used throughout the experiments described in this thesis to model and dissect lung tumorigenesis *in vivo*. Finally, I will present evidence suggesting that a large percentage of human KRAS-driven lung adenocarcinomas harbor mutations in multiple putative tumor suppressor genes, the vast majority of which remain to be validated using well established GEMMs of this disease.

i. Somatic mutations and copy number alterations

The most comprehensive lung adenocarcinoma genomics study carried out so far was published by the TCGA consortium last year (LUAD TCGA 2014). In this study, they performed extensive molecular profiling of 230 lung adenocarcinoma specimens using a combination of whole-transcriptome, whole-exome, DNA methylation and proteomic analyses. They uncovered ~ 18 genes that were significantly mutated (see **Figure 1A** for an up to date catalog of somatic mutations compiled from analysis of 405 patients). Beyond re-discovering known lung adenocarcinoma drivers, such as *KRAS*, *TP53*, *STK11*, *EGFR*, *PIK3CA*, and *BRAF*, among several others, they identified multiple novel

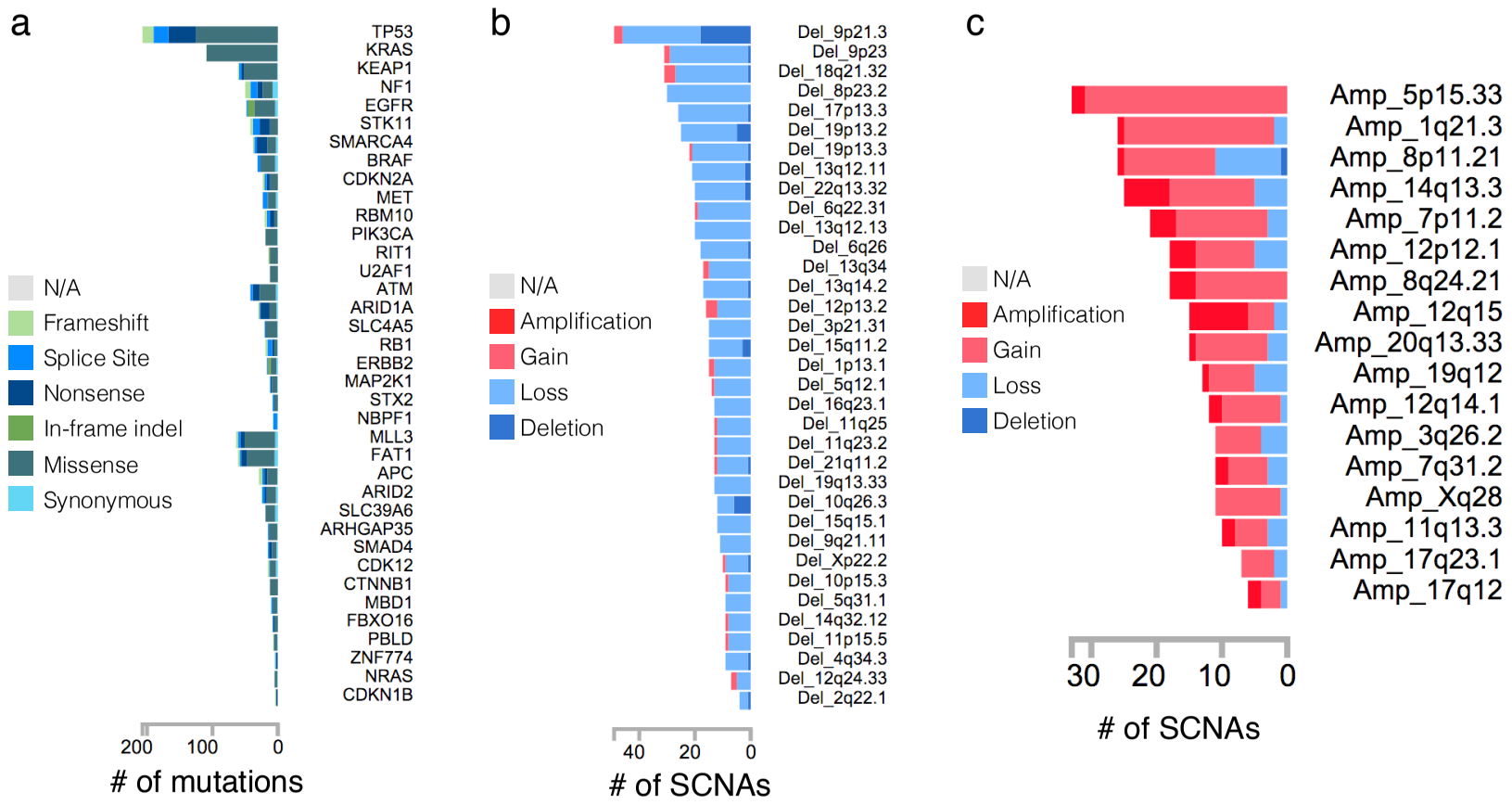


Figure 1: The current catalog of somatic mutations and copy number alterations in human lung adenocarcinoma. (a) Somatic mutations. (b) Focal chromosomal deletions. (c) Focal chromosomal amplifications. Del = Deletion. Amp = Amplification. This figure is based on combined data from 405 lung adenocarcinoma patients and was generated using the TumorPortal website. See Lawrence et al., (2014) and http://www.tumorportal.org/ .

putative drivers, including *RIT1*, *MGA*, *RBM10* and *KEAP1* (more on *KEAP1* on **Part IV** of this introduction). Interestingly, ~ 38% of lung adenocarcinoma tumors - referred to as the *oncogene negative* class of tumors - lacked activating mutations in known oncogenic drivers, suggesting that these tumors might be driven by loss of one or more tumor suppressor genes (such as *NF1* and *KEAP1*). This hypothesis, however, remains to be stringently validated, particularly in the setting of well established GEMMs of lung cancer.

Lung adenocarcinomas typically harbor multiple somatic copy number alterations, including focal chromosomal amplifications and deletions (see **Figure 1B-C** for an up to date catalog of somatic copy number alterations compiled from analysis of 405 patients) (Weir et al., 2007, Beroukhim et al., 2010). The most commonly amplified regions contained multiple well described human lung adenocarcinoma drivers, including *NKX2-1* (14q13.3), *MDM2* (12q15), *KRAS* (12p12.1), *EGFR* (7p11.2), *MET* (7q31.2), and *TERT* (5p15.33), among others. On the other hand, the most commonly deleted regions contained multiple well described tumor suppressor genes, including *APC* (5q11.2, 5q31.1) and *CDKN2A* (9p21.3), with the latter being the most statistically significant event across all samples.

ii. *KRAS* mutations

Lung cancer-associated mutations in the *KRAS* GTPase - which invariably lead to a constitutively GTP-bound state accompanied by incessant activation of RAS effector pathways - were initially identified in the 1980s by the laboratories of Mariano Barbacid, Manuel Perucho, and Johannes Bos (Santos et al., 1984, Nakano et al., 1984,

Rodenhuis et al., 1987). As can be seen in **Figure 1A**, these mutations are observed in ~ 33% of human lung adenocarcinoma patients and are relatively uncommon in other lung cancer subtypes (Hammerman et al., 2012, George et al., , LUAD TCGA 2014, Politi and Herbst 2015). Interestingly, they are almost always associated with codons 12, 13 or 61; however, codon 12 mutations that lead to glycine-to-cysteine (G12C), glycine-to-valine (G12V), or glycine-to-aspartic acid (G12D) changes are the ones most frequently observed in human lung adenocarcinoma patients (Prior et al., 2012, Cox et al., 2014, Devarakonda et al., 2015).

RAS oncogenes (including oncogenic KRAS) activate multiple downstream signaling pathways (such as the PI3K, MAPK, RALGEF, and PLCε effector pathways) that promote several canonical hallmarks of cancer, such as uncontrolled proliferation, suppression of apoptosis, metabolic reprogramming, and evasion of the immune system, among others (Hanahan and Weinberg 2000, 2011, reviewed in Downward 2003, Karnoub and Weinberg 2008, and Pylayeva-Gupta et al., 2011). For example, oncogenic RAS can promote cellular proliferation through the indirect transcriptional upregulation of several growth promoting transcription factors, such as FOS, SRF, JUN, and NF-kB, among others (Stacey et al., 1987, Urich et al., 1997, Westwick et al., 1994, Finco et al., 1997). Similarly, oncogenic RAS can actively suppress cell death via the PI3K and MAPK pathways through downregulation of the pro-apoptotic BCL-2 family member BAK1 and the pro-apoptotic transcriptional repressor PAR4 (Rosen et al., 1998, Nalca et al., 1999). Collectively, these and several other studies have unequivocally demonstrated the oncogenic potential of cancer-associated *RAS* mutations, including the most prevalent lung cancer-associated *KRAS* mutant alleles.

Several studies have demonstrated the oncogenic role of *KRAS* mutant alleles in lung adenocarcinoma using GEMMs of cancer. The first GEMM of Kras-driven lung adenocarcinoma, which employed a latent allele (see **Figure 1C** from **Part I** of this introduction) of oncogenic $Kras^{G12D}$, was developed by the group of Tyler Jacks (Johnson et al., 2001). In this model, the spontaneous somatic activation of oncogenic $Kras^{G12D}$ *in vivo* triggered the development of lung adenocarcinoma with 100% penetrance. Subsequently, Jacks and colleagues generated the well established $Kras^{LSL\text{-}G12D/+}$ (K) and $Kras^{LSL\text{-}G12D/+;\ p53^{flox/flox}}$ (KP) GEMMs of lung adenocarcinoma (see **Figure 4** from **Part I** of this introduction), and used them to demonstrate that activation of the oncogenic $Kras^{G12D}$ allele with or without concomitant deletion of both copies of the *p53* tumor suppressor gene in the murine lung epithelium (via intranasal administration of adenoviruses expressing Cre recombinase) triggered the development of lung adenocarcinoma with 100% penetrance (Jackson et al., 2001, Jackson et al., 2005). These studies unequivocally demonstrated the oncogenic nature of mutant *Kras* *in vivo* and established the K and KP GEMMs as the prototypical models for dissecting the mechanisms of lung adenocarcinoma initiation and progression, as well as for performing pre-clinical and co-clinical studies (Engelman et al., 2008, Singh et al., 2010, Chen et al., 2012).

iii. *TP53* mutations

Lung cancer-associated mutations in the *TP53* tumor suppressor gene, which encodes for the P53 transcription factor, were initially identified in the late 1980s by the groups of John Minna and Bert Vogelstein (Brauch et al., 1987, Takahashi et al., 1989, Nigro et

al., 1989, Chiba et al., 1990). Subsequent studies demonstrated that *TP53* was in fact the most commonly mutated tumor suppressor gene in human cancer, with mutations being observed in colorectal cancer (Baker et al., 1989, Nigro et al., 1989, Shaw et al., 1991), breast cancer (Nigro et al., 1989, Bartek et al., 1990, Varley et al., 1991), bladder cancer (Sidransky et al., 1991), prostate cancer (Isaacs et al., 1991), liver cancer (Hsu et al., 1991, Bressac et al., 1991), soft tissue sarcomas (Toguchida et al., 1992a, Toguchida et al., 1992b), and melanomas (Stretch et al., 1991), among others.

P53 is a pleiotropic transcription factor that sits at the center of a network that responds to a variety of stress signals (Vogelstein et al., 2000). In the absence of stress signals, the P53 protein is constantly undergoing degradation due to a feedback loop in which P53 transcriptionally activates its main negative regulator, MDM2, and MDM2, in turn, promotes the degradation of P53 (Kubbutat et al., 1997, Haupt et al., 1997, Honda et al., 1997). In the presence of stress signals, such as DNA damage (Lowe et al., 1993), oncogenic stress (Serrano et al., 1997), hypoxia (Graeber et al., 1994), and metabolic stress (Okorokov and Milner 1999), among others, this negative feedback loop is inhibited and the P53 protein gets stabilized. Upon stabilization, P53 gets shuttled to the nucleus, where it directly activates a myriad of genes involved in a variety of tumor suppressive programs, such as cell cycle arrest, cellular senescence, and apoptosis, among others (reviewed in Vogelstein et al., 2000 and Brosh and Rotter 2009). These critical tumor suppressive functions, along with the fact that ~ 50% of all human cancers harbor mutations in the *TP53* gene while the remaining cases typically harbor mutations that inactivate other components the P53 pathway, have uniquely positioned P53 as the most important tumor suppressor gene.

As can be seen in **Figure 1A**, *TP53* mutations are observed in ~ 46% of human lung adenocarcinomas (LUAD TCGA 2014) and are often classified in one of two categories: *contact mutations* or *structural mutations* (Brosh and Rotter 2009). Contact mutations, such as R248Q and R273H, occur in the DNA binding domain of P53 and tend to abolish its function as a transcription factor. Structural mutations, such as R249S, G245S, R175H, and R282W tend to disrupt the proper conformation of the P53 protein (Brosh and Rotter 2009). Interestingly, multiple studies throughout the years have demonstrated that some of these mutations can act in a potent gain of function and dominant negative fashion to promote several aspects of tumorigenesis and therapeutic resistance, such as enhanced tumorigenic potential (Dittmer et al., 1993, Olive et al., 2004, Lang et al., 2004, Bossi et al., 2006, Song et al., 2007, Terzian et al., 2008), enhanced metastatic potential (Weissmueller et al., 2014), and resistance to chemotherapy (Chin et al., 1992, Aas et al., 1996, Blandino et al., 1999), among others. In addition, mutant P53 has been shown to antagonize its family members P63 and P73 (which, on their own, have potent tumor suppressive properties) by directly binding to them and interfering with their own transcriptional activity (Di Como et al., 1999, Gaiddon et al., 2001). Moreover, disrupting the interaction between mutant P53 and P63 or P73 has been shown to have a therapeutic effect (Di Agostino et al., 2008, Kravchenko et al., 2008), further supporting the notion that mutant P53 alleles are not just mere loss of function alleles. Recently, mutant P53 was shown to transcriptionally activate multiple chromatin regulators, including histone methyltransferases and acetyltransferases (Zhu et al., 2015), as well as to cooperate with SWI/SNF family members to promote the transcription of *VEGFR2* (Pfister et al., 2015). Collectively,

these studies suggest that therapies designed against mutant P53 or that exploit a synthetic lethal interaction with mutant p53 might be highly effective in treating tumors that harbor these gain of function mutations.

Several studies throughout the years have demonstrated the tumor suppressive role of *P53* in lung adenocarcinoma using GEMMs of cancer. The original studies mentioned above, in which mice bearing the *Kras*^{G12D} latent allele or KP mice were used to model lung adenocarcinoma initiation and progression *in vivo*, also demonstrated that *p53* is a potent tumor suppressor gene in this context (Johnson et al., 2001, Jackson et al., 2005). Indeed, *p53* loss dramatically accelerates the progression of *Kras*^{G12D}-driven lung adenocarcinoma, increasing both the frequency of high grade tumors, as well as the frequency of metastatic dissemination (Jackson et al., 2005, Winslow et al., 2011). Another notable study (described previously in **Part I** of this introduction), utilized *Kras*^{G12D} latent mice expressing a LSL-p53 allele to demonstrate that *p53* mainly acts as a stage-specific tumor suppressor gene in lung adenocarcinoma, whereby p53 suppresses the transition from low grade to high grade tumors (Feldser et al., 2010). These and multiple other studies using GEMMs of cancer (reviewed in Donehower and Lozano 2009) have unequivocally demonstrated the tumor suppressive activity of *p53 in vivo.*

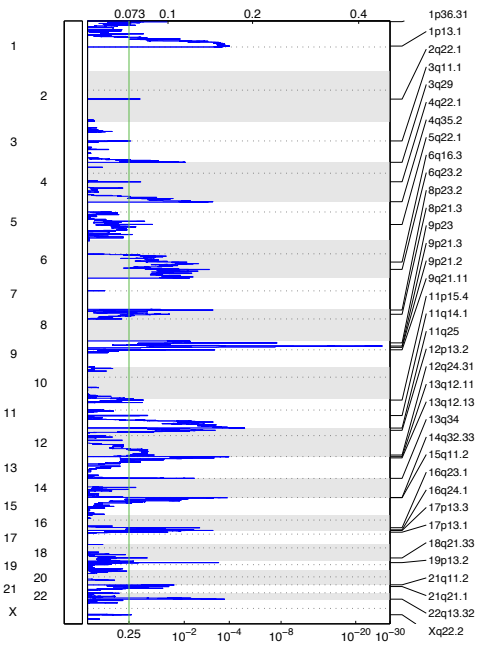
iv. Genes co-mutated with oncogenic *KRAS* and beyond

The current catalog of recurrent somatic mutations and copy number alterations in lung adenocarcinoma can also be used for uncovering novel genetic interactions between well established potent oncogenes, such as *KRAS*, and other oncogenes or tumor

a

Gene	# of patients with gene mutation	# of patients with KRAS mutation and gene mutation	% of patients with KRAS mutation and gene mutation	MutSig q-value	Established LUAD tumor suppressor gene?
TP53	275	52	35.62	0	Yes
STK11 / LKB1	74	36	24.66	0	Yes
RBM10	35	18	12.33	0	No
IL32	33	9	6.16	0	No
KEAP1	89	26	17.81	6.28E-12	No
NBPF1	47	16	10.96	3.46E-03	No
HAX1	19	6	4.11	7.48E-03	No
NUDT11	7	5	3.42	7.48E-03	No
HEBP1	18	5	3.42	1.99E-02	No

b Deletions



c Amplifications

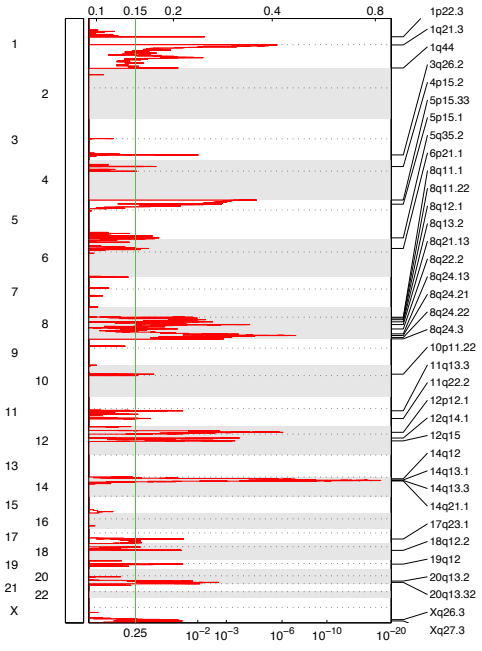


Figure 2: Genes and genomic regions significantly altered in lung adenocarcinoma tumors harboring *KRAS* mutations. (a) Genes significantly co-mutated with oncogenic *KRAS* based on MutSig analysis (see Lawrence et al., (2013) and https://www.broadinstitute.org/cancer/cga/mutsig). Focal chromosomal deletions (b) and amplifications (c) significantly co-occurring in tumors harboring oncogenic *KRAS* based on GISTIC2.0 analysis (see Mermel et al., (2011) and http://www.broadinstitute.org/cgi-bin/cancer/publications/view/216).

suppressor genes. For example, putative oncogenes or tumor suppressor genes that are found to be significantly co-mutated with oncogenic *KRAS* would be predicted to have a functional cooperative interaction during one or more stages of lung adenocarcinoma tumorigenesis. Therefore, the functional interrogation of such candidate cooperating oncogenes and tumor suppressor genes using well defined GEMMs of lung adenocarcinoma would be a powerful approach that can be applied to uncover and validate such genetic relationships. As a first step towards this goal, we have performed a series of preliminary *in silico* analyses that have identified multiple putative lung cancer genes altered by mutation or somatic copy number alteration in *KRAS* mutant lung adenocarcinoma tumors (**Figure 2**). Notably, the vast majority of these genes and genomic regions have not been previously validated as *bona fide* drivers of lung adenocarcinoma, particularly using well defined GEMMs of lung cancer. These initial analyses, combined with the model of cumulative haploinsufficiency and triplosensitivity postulated by Elledge and colleagues (Davoli et al., 2013), suggest that many more cancer drivers remain to be discovered. In **Chapter 2** of this thesis, I describe a novel CRISPR-Cas9-based approach that will allow for the functional identification of single and multiple genetic drivers *in vivo* using well established GEMMs of cancer, including those designed to model lung adenocarcinoma.

Chapter 1 – Introduction

Part IV – The Keap1-Nrf2 pathway

I - Non-oncogene addiction and the stress hallmarks of cancer

In addition to the classical 1st and 2nd generation Hanahan-Weinberg hallmarks of cancer, Stephen Elledge and colleagues have proposed that there are at least 5 additional hallmarks of equal importance, which they refer to as the *stress phenotypes of cancer* or the *stress hallmarks of cancer* (Luo et al., 2009). These additional stress hallmarks comprise DNA damage/replication stress, proteotoxic stress, mitotic stress, metabolic stress, and oxidative stress. Notably, the last two stress hallmarks directly resonate with the work of Otto Warburg and his demonstration of the famous "Warburg Effect", whereby cancer cells consume much larger amounts of glucose than their normal counterparts and do so through aerobic glycolysis, which in turn leads to high levels of oxidative stress (Warburg et al., 1924, Warburg 1956, reviewed in Lunt and Vander Heiden 2011). Consequently, the increased proliferative capacity that is characteristic of cancer cells unavoidably gives rise to a major Achilles heel in the form of high levels of oxidative stress, which are mainly exerted by reactive oxygen species (ROS) that dramatically accumulate in the cells due to a major imbalance in the homeostatic mechanisms that generate and eliminate ROS (reviewed in Trachootham et al., 2009 and Gorrini et al., 2013).

ROS - which are typically defined as oxygen-containing chemical species (including the superoxide and hydroxyl free radicals, as well as the non-radical hydrogen peroxide) - are highly reactive molecules that, when present at high levels, can directly damage cellular components and macromolecules, including DNA, proteins, and lipids (Gorrini et al., 2013). Multiple biological processes, such as hypoxia and metabolism (Chandel et al., 2000, Kaelin 2005, Brunelle et al., 2005, Guzy et al., 2005, Mansfield et al., 2005, Gao et al., 2007), endoplasmic reticulum stress (Santos et al., 2009, Dejeans et al., 2010, Inoue and Suzuki-Karasaki 2013, Eletto et al., 2014), and oncogene activation (Irani et al., 1997, Lee et al., 1999, Tanaka et al., 2002, Vafa et al., 2002), can lead to the production of high levels of ROS, which, if left unbuffered, can lead to severe cellular damage and cell death. On the other hand, moderate levels of ROS can also serve as important signaling molecules in multiple cellular pathways involved in cellular proliferation, survival and differentiation, among others (Gorrini et al., 2013). For example, ROS can play active signaling roles in promoting inflammation (Bulua et al., 2011, Zhou et al., 2011, Nakahira et al., 2011), in activation of Jun N-terminal kinase (JNK) and downstream signaling pathways (Kamata et al., 2005), and in the transcriptional response to hypoxia (Chandel et al., 2000, Gao et al., 2007). Therefore, to prevent the lethal consequences of high levels of oxidative stress without disturbing critical ROS-mediated signaling pathways, cells have evolved intricate and highly regulated cellular pathways to achieve a delicate homeostatic balance of ROS production and clearance. Notably, there is an extensive body of evidence that suggests that cancer cells (which typically produce more ROS than their normal cellular counterparts) become highly dependent on cellular pathways that mediate ROS

detoxification and clearance in order to achieve homeostatic levels that are compatible with rapid cellular proliferation and survival (reviewed in Trachootham et al., 2009). This biological phenomenon - which is an example of what Stephen Elledge and colleagues have termed *non-oncogene addiction* due to the fact that the main players in these pathways are not necessarily oncogenes *per se* - is the focus of this part of the introduction.

What are the genes and pathways that normal cells use to achieve homeostatic levels of ROS in order to avoid the detrimental effects of high oxidative stress? Are these genes and pathways altered in human cancer? If so, do these alterations contribute to one or more of the stress hallmarks of cancer? Consequently, do they create novel cancer dependencies based on the phenomenon of non-oncogene addiction? These are the questions that I will address below focusing on the Keap1-Nrf2 pathway as the prototypical example of a signaling pathway that is prone to being hijacked by cancer cells, and which appears to exemplify the phenomenon of non-oncogene addiction. Due to the fact that this introduction will be exclusively focused on the Keap1-Nrf2 pathway and its emerging role in cancer biology, I would like to direct the reader to a series of excellent reviews that discuss ROS biology, ROS homeostasis, and additional signaling pathways in great depth: Finkel 2011, Gorrini et al., 2013, and Holmstrom and Finkel 2014.

II - The Keap1-Nrf2 pathway and ROS homeostasis

A. ROS detoxification

Remarkably, cells have evolved multiple ROS scavenging and detoxifying pathways to achieve a delicate homeostatic balance between ROS production and ROS clearance. As a preamble to the discussion of the Keap1-Nrf2 pathway, I will briefly describe two of these ROS detoxifying mechanisms, all of which are directly or indirectly controlled by the Keap1-Nrf2 pathway (**Figure 1**).

i. Glutathione-mediated ROS detoxification

Cells employ glutathione (GSH) - the most abundant metabolite and antioxidant in the cell - to eliminate ROS via a reaction cycle controlled by the concerted action of two classes of enzymes: GSH peroxidases (GPXs) and GSH S-transferases (GSTs) (**Figure 1A**) (Meister et al., 1983, Gorrini et al., 2013, Harris 2013, and others). In this scenario, both GPXs and GSTs catalyze the GSH-mediated detoxification of ROS, which leads to the conversion of GSH to its oxidized form, GSSG. Subsequently, GSH is regenerated via the action of GSH reductase (GSR), which utilizes nicotinamide adenine dinucleotide phosphate (NADPH) as a reducing agent to reduce GSSG to GSH, thereby regenerating the pool of GSH available to detoxify additional ROS within the cell (Gorrini et al., 2013, Harris 2013). Importantly, the ratio of GSH:GSSG in a cell is directly indicative of how much oxidative stress the cell is currently experiencing. Therefore, low GSH:GSSG ratios denote conditions of high oxidative stress, whereas high GSH:GSSG ratios denote conditions of a highly reduced environment.

ii. Thioredoxin-mediated ROS detoxification

Cells also employ the thioredoxin (TXN) protein - another major cellular antioxidant - to detoxify ROS via the reduction of peroxiredoxin (PRDX) (Murphy et al., 2012, Gorrini et al., 2013). In this scenario, TXN-mediated reduction of PRDX allows the latter to directly detoxify ROS, which in turn leads to the conversion of PRDX to its oxidized form (Murphy et al., 2012). Subsequently, reduced PRDX is regenerated via the action of TXN, thereby regenerating the pool of PRDX available to detoxify additional ROS within the cell. TXN itself gets oxidized upon reducing PRDX; therefore, the regeneration of reduced TXN is a critical rate-limiting step in this ROS detoxification cycle. Reduced TXN is regenerated via the action of TXN reductase (TXNRD) in an NADPH-dependent manner (Figure 1B) (Gorrini et al., 2013, Harris 2013).

iii. Nrf2-mediated ROS detoxification

Cells also employ a pleiotropic transcription factor - the Nuclear Factor Erythroid 2-related factor 2 (Nrf2) transcription factor - which regulates a large repertoire of genes that coordinate the cellular response to oxidative stress and oxidative damage (Mitsuishi et al., 2012a). Nrf2 is widely considered the master regulator of the antioxidant response, and deservedly so: in conditions of high oxidative stress, such as those triggered by high levels of ROS, Nrf2 activates a battery of genes that regulate multiple cytoprotective pathways, including GSH synthesis and conjugation, ROS detoxification and elimination, and drug excretion, among others (reviewed in Mitsuishi et al., 2012a and Suzuki and Yamamoto 2015) (**Figure 2**). Indeed, Nrf2 controls both GSH-mediated

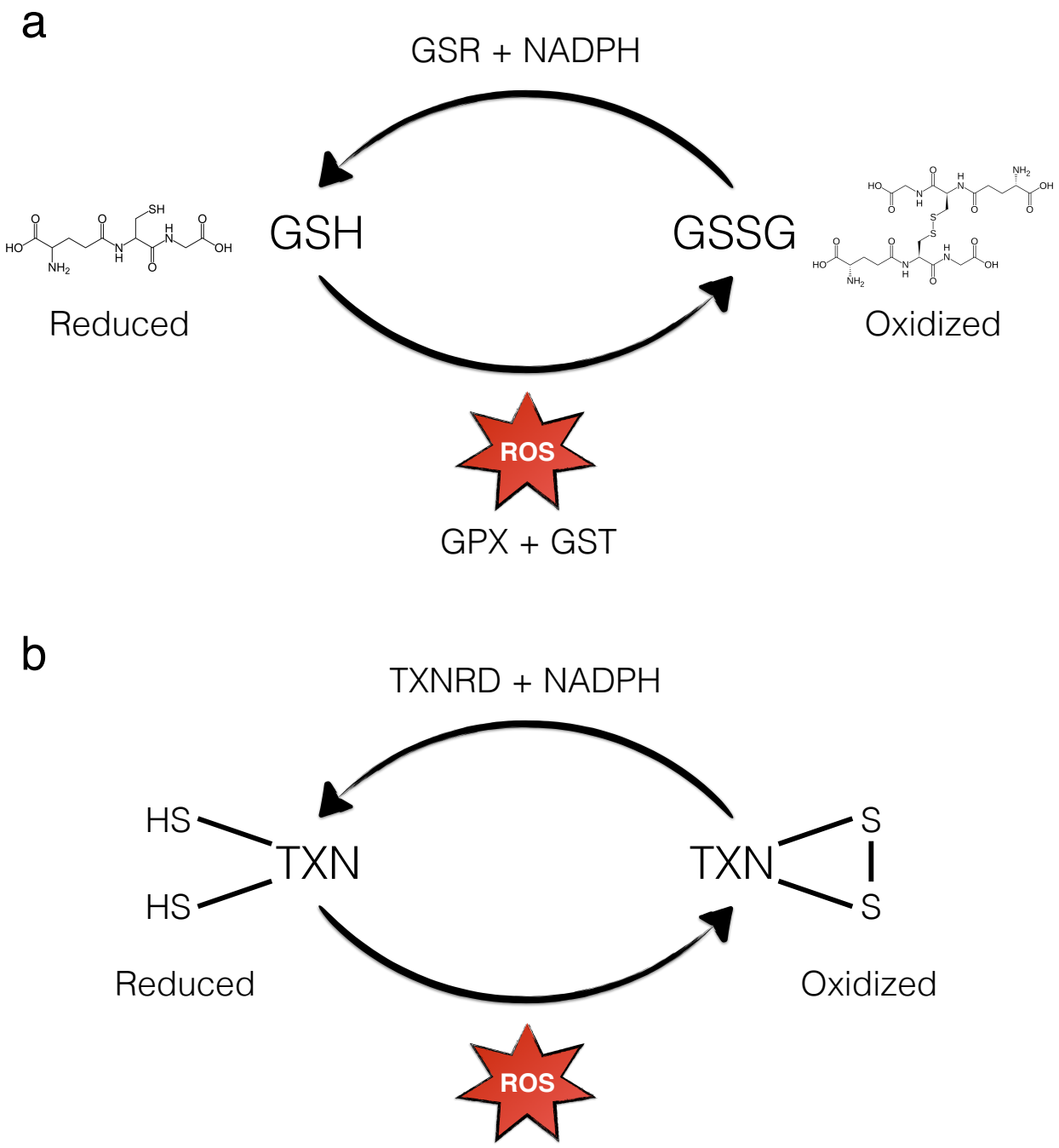


Figure 1: GSH- and TXN-mediated ROS detoxification pathways. (a) GSH detoxifies ROS via the concerted action of GPXs and GSTs. GSH is regenerated by GSR in an NADPH-dependent manner. (b) TXN detoxifies ROS via PRDX (not shown). Reduced TXN is regenerated by TXNRD in an NADPH-dependent manner. See text for additional details.

and TXN-mediated detoxification pathways at multiple levels through its pleiotropic transcriptional activity, which will be described in detail in the next few sections.

B. Nrf2: master regulator of the antioxidant response

i. Discovery and historical perspective

A series of studies by Thomas Rushmore and Cecil Pickett in the early 1990s led to the identification and functional characterization of a novel type of DNA regulatory element that is now known to be present in the promoter of a large number of genes whose expression is induced in conditions of high oxidative stress (Rushmore and Pickett 1990, Rushmore et al., 1991, Nguyen et al., 2009). These regulatory elements, termed the antioxidant response elements (AREs), are characterized by containing the following consensus sequence: 5' - RGTGACNNNGC - 3', where R is a purine and N is any other nucleotide (Rushmore et al., 1991). Notably, the first ARE to be identified by Pickett and colleagues was the one present in the gene encoding for the GSH S-transferase (GST) (Rushmore and Pickett 1990), emphasizing the importance of this 11bp DNA sequence in potentially regulating the inducible expression of a battery of genes involved in pathways mediating cellular protection against high levels of oxidative stress. The transcription factor(s) that recognized this particular regulatory element were unknown at the time.

Nrf2 was initially identified and cloned by the group of Yuet Wai Kan in 1994 as a 66-kDa human protein that contained a basic region-leucine zipper (bZip) DNA binding domain with substantial homology to that of Nuclear Factor Erythroid 2 (NF-E2) (Moi et al., 1994). Indeed, Nrf2 was able to recognize and bind the NF-E2 binding site present

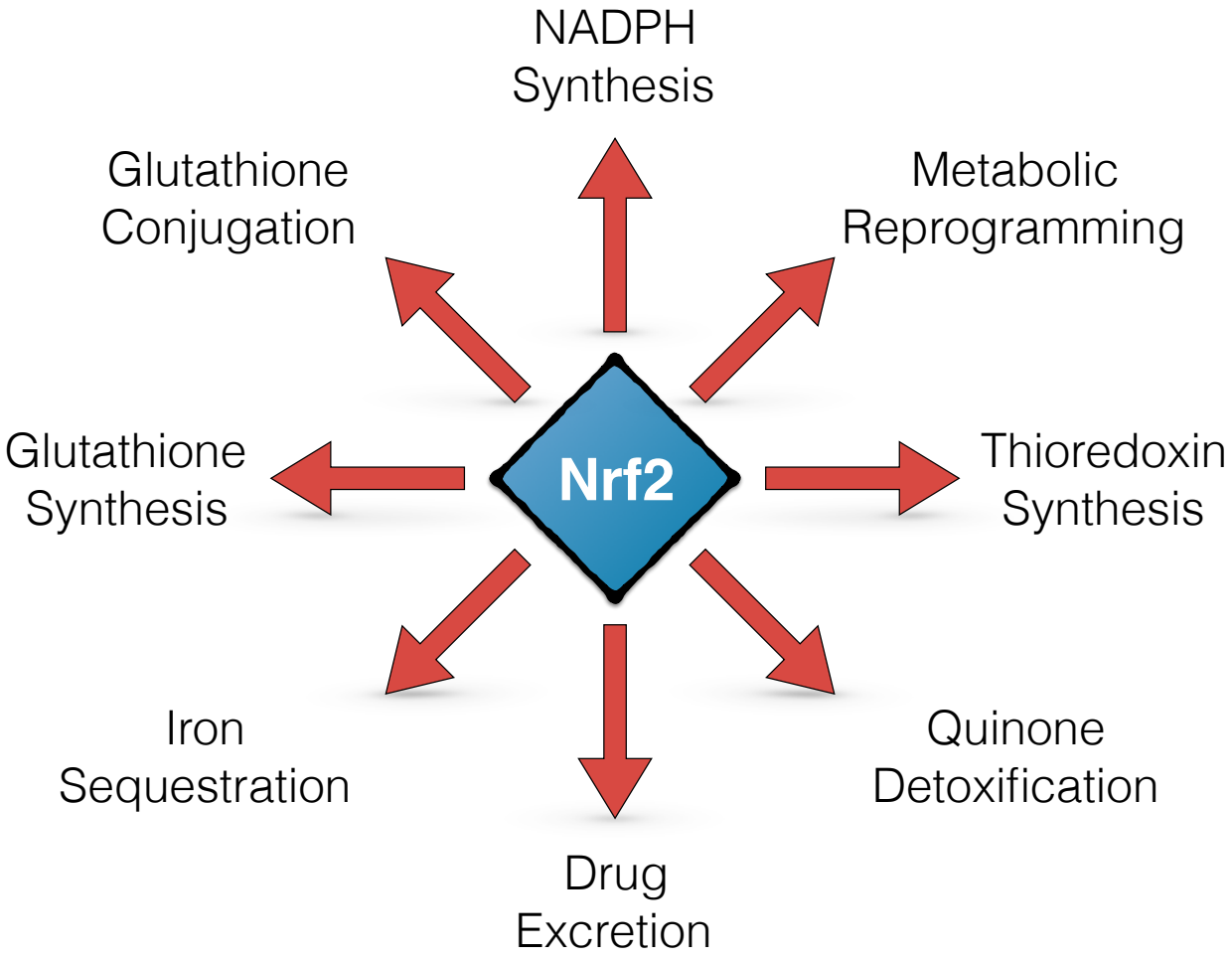


Figure 2: The pleiotropic Nrf2 transcription factor controls multiple cytoprotective pathways. See text for details.

in human β-globin genes, an observation that led Yuet Wai Kan and colleagues to hypothesize that Nrf2 could play a role in the transcriptional activation of β-globin genes (Moi et al., 1994). In a contemporaneous study, Masayuki Yamamoto and colleagues identified and cloned the chicken homolog of Nrf2 by screening chicken cDNA libraries with a cDNA probe corresponding to murine p45 NF-E2 (which was initially identified and cloned by Stuart Orkin and colleagues in 1993) (Andrews et al., 1993a, Peters et al., 1993, Itoh et al., 1995). They initially named this protein Erythroid-Derived Protein with CNC Homology (ECH) and further demonstrated that it could heterodimerize with small Maf family proteins (Itoh et al., 1995), which were known to recognize and bind a sequence that was remarkably similar to the ARE elements despite lacking canonical transcriptional activation domains (Andrews et al., 1993b, Igarashi et al., 1994, Kataoka et al., 1994, Engel 1994). Nevertheless, a direct functional connection between Nrf2 and transcriptional activation of ARE-containing genes was still lacking.

The first functional connection between Nrf2 and transcriptional activation of ARE-containing genes was reported by the group of Anil Jaiswal in 1996, in which they demonstrated that the inducible expression of the gene encoding NADPH:quinone oxidoreductase-1 (NQO1) required the presence of an ARE element, and that Nrf2 was able to directly bind to this element and mediate potent transcriptional induction of *NQO1* (Venugopal and Jaiswal 1996). Subsequent studies by the groups of Yuet Wai Kan and Masayuki Yamamoto using gene targeting via homologous recombination in murine ES cells demonstrated that *Nrf2* knockout mice were viable and fertile (Chan et al., 1996, Itoh et al., 1997). Nevertheless, Yamamoto and colleagues unequivocally confirmed that the Nrf2 transcription factor was critical for the inducible expression of

multiple detoxifying enzymes, including GSTs and *NQO1* (Itoh et al., 1997). These and other contemporaneous studies positioned the Nrf2 transcription factor at the center of a cellular network that responded to conditions of high oxidative stress via the rapid transcriptional induction of multiple ARE-containing genes involved in ROS detoxification and elimination, among other important biological processes (for a recent review, see Suzuki and Yamamoto 2015).

ii. Function

As discussed above, the Nrf2 transcription factor is a potent transcriptional activator that plays a critical role in the inducible expression of a large repertoire of genes involved in the cellular response to multiple extrinsic and intrinsic insults, such as oxidative stress (Itoh et al., 1997). The Nrf2 protein contains seven Nrf2-ECH homology (Neh) domains, commonly referred to as the Neh1-Neh7 domains (Itoh et al., 1999, Wang et al., 2013) (**Figure 3A**). The Neh1 domain contains the CNC homology region, as well as the bZIP DNA-binding domain; therefore, this is the domain that allows Nrf2 to bind DNA and heterodimerize with other transcription factors, including the small Maf family of proteins. The Neh2 domain, which is the major regulatory domain, contains a string of seven lysine residues that are critical for Nrf2 regulation through polyubiquitination, as well as the DLG and ETGE motifs, which are involved in regulating Nrf2 stability (Zhang et al., 2004, McMahon et al., 2006, Tong et al., 2006). The Neh3 domain has been shown to interact with the CHD6 transcriptional coactivator protein; this interaction appears to be critical for proper induction of *NQO1* and other Nrf2 target genes (Nioi et al., 2005). The Neh4 and Neh5 domains have been shown to interact with CBP; this

interaction appears to be important for proper induction of multiple Nrf2 target genes (Katoh et al., 2001, Zhu and Fahl 2001). On the other hand, the recently described Neh7 domain appears to be important for additional negative regulation of Nrf2 via its interaction with the Nrf2 repressor, RXRα (Wang et al., 2013).

iii. Regulation by Keap1

Nrf2 activity is tightly regulated at the post-transcriptional level by its main negative regulator, Kelch-like ECH-associated protein 1 (Keap1) (**Figure 4**). This protein, which was originally identified in 1998 by Yamamoto and colleagues through a yeast two-hybrid screen using the Neh2 domain as bait (Itoh et al., 1999), functions as a substrate adaptor for a Cul3-containing E3 ubiquitin ligase. This Keap1-Cul3-E3 ubiquitin ligase complex promotes the polyubiquitination of the seven lysine residues in the Neh2 domain of Nrf2, thereby promoting its subsequent degradation via the 26S proteasome (Zhang et al., 2004, McMahon et al., 2006, Tong et al., 2006). Therefore, in the absence of oxidative stress, Nrf2 is constitutively degraded in the cytoplasm by the action of the Keap1-Cul3-E3 ubiquitin ligase complex. However, in the presence of oxidative stress, Keap1 is oxidized at multiple critical cysteine residues, which in turn abolishes the Keap1-Nrf2 interaction. Consequently, free Nrf2 gets shuttled to the nucleus where it activates a large repertoire of genes that orchestrate the cellular antioxidant response (**Figure 4**).

The negative regulation imposed by Keap1 is significantly complex. The Keap1 protein (**Figure 3B**) contains at least three functional domains, including the bric-a-brac/ tramtrack/broad complex (BTB) and Kelch/DGR protein interaction domains, as well as

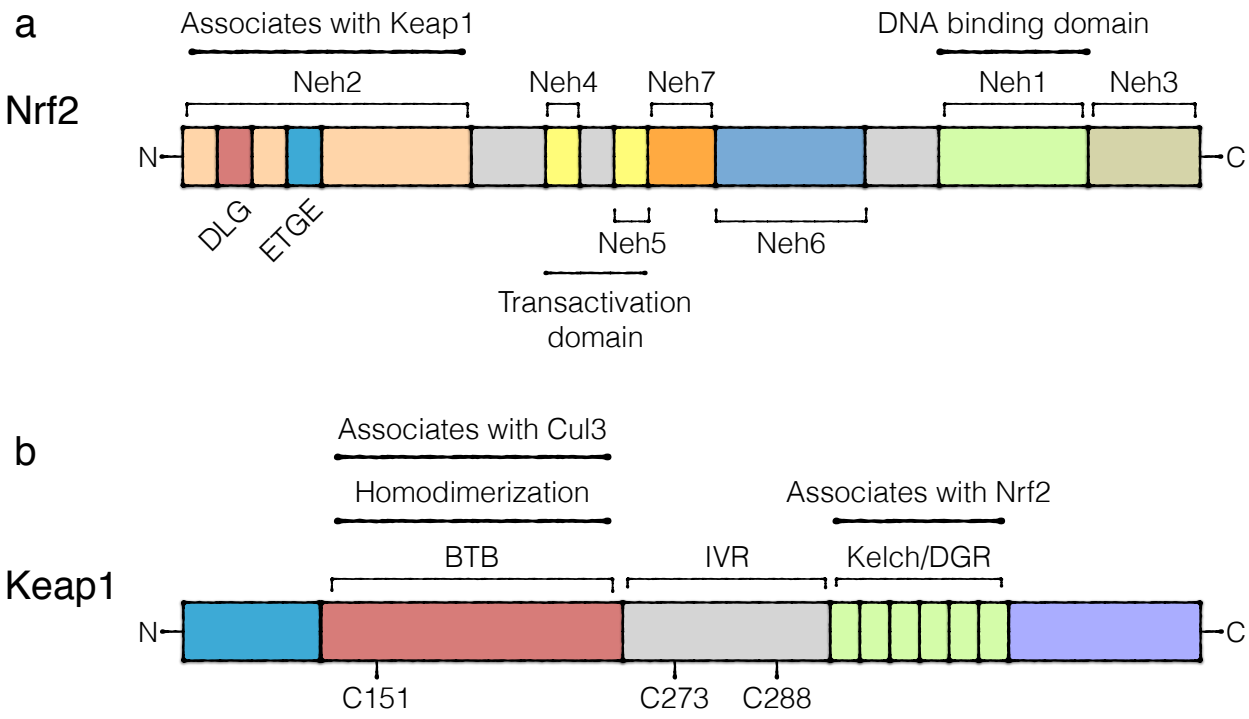


Figure 3: Functional domain structures of Nrf2 and Keap1. (a) Nrf2 contains seven functional domains, termed Neh1-7, which are involved in multiple aspects of Nrf2 function and regulation, including association with Keap1, DNA binding, and transactivation of target genes. (b) Keap1 contains three functional domains, termed the BTB, IVR, and DGR domains, which are involved in multiple aspects of Keap1 function and regulation, including association with the Cul3-E3 ubiquitin ligase complex, homodimerization, and association with Nrf2. Notably, Keap1 contains multiple cysteine residues, most of which get oxidized in conditions of high oxidative stress. These cysteine residues play active roles in the regulation of Keap1 function and its negative regulation of Nrf2. Adapted from Jaramillo and Zhang 2013.

an intervening region (IVR) domain (Itoh et al., 1999, Zipper and Mulcahy 2002, McMahon et al., 2004, Lo et al., 2006, Itoh et al., 2010, Jaramillo and Zhang 2013). The BTB domain plays two major functional roles. On one hand, the BTB domain is required for Keap1 homodimerization, which is critical for negatively regulating Nrf2 (Zipper and Mulcahy 2002). On the other hand, Keap1 binds Cul3 through the BTB domain and this interaction is critical for mediating Nrf2 polyubiquitination (Cullinan et al., 2004, Kobayashi et al., 2004, Zhang et al., 2004). The Kelch/DGR domain interacts with the Neh2 domain of Nrf2; this interaction is critical for maintaining a stable Keap1-Nrf2 interaction (Itoh et al., 1999, McMahon et al., 2004). Lastly, the IVR domain connects the other two domains via a linker region that contains multiple cysteine residues that might be involved in regulating Keap1 activity (Kobayashi et al., 2004). Notably, the murine and human Keap1 proteins contain 27 and 25 cysteine residues, respectively, many of which undergo oxidation in the presence of high levels of oxidative stress (Itoh et al., 1999).

A very interesting feature of the negative regulation imposed on Nrf2 by Keap1 is the fact that it is *newly* synthesized free Nrf2 that translocates to the nucleus in conditions of high oxidative stress (Jaramillo and Zhang 2013). Specifically, oxidized Keap1 is unable to bind Nrf2 with the same high affinity that non-oxidized Keap1 does. Nevertheless, this relatively weak Keap1-Nrf2 interaction persists and prevents Nrf2 polyubiquitination. Due to the fact that Nrf2 is not being degraded, all of the available Keap1 molecules in the cell become saturated with nondegradable Nrf2. Consequently, newly synthesized Nrf2 molecules are able to escape Keap1-mediated sequestration and readily shuttle to the nucleus to activate the transcription of genes that mediate the cellular antioxidant

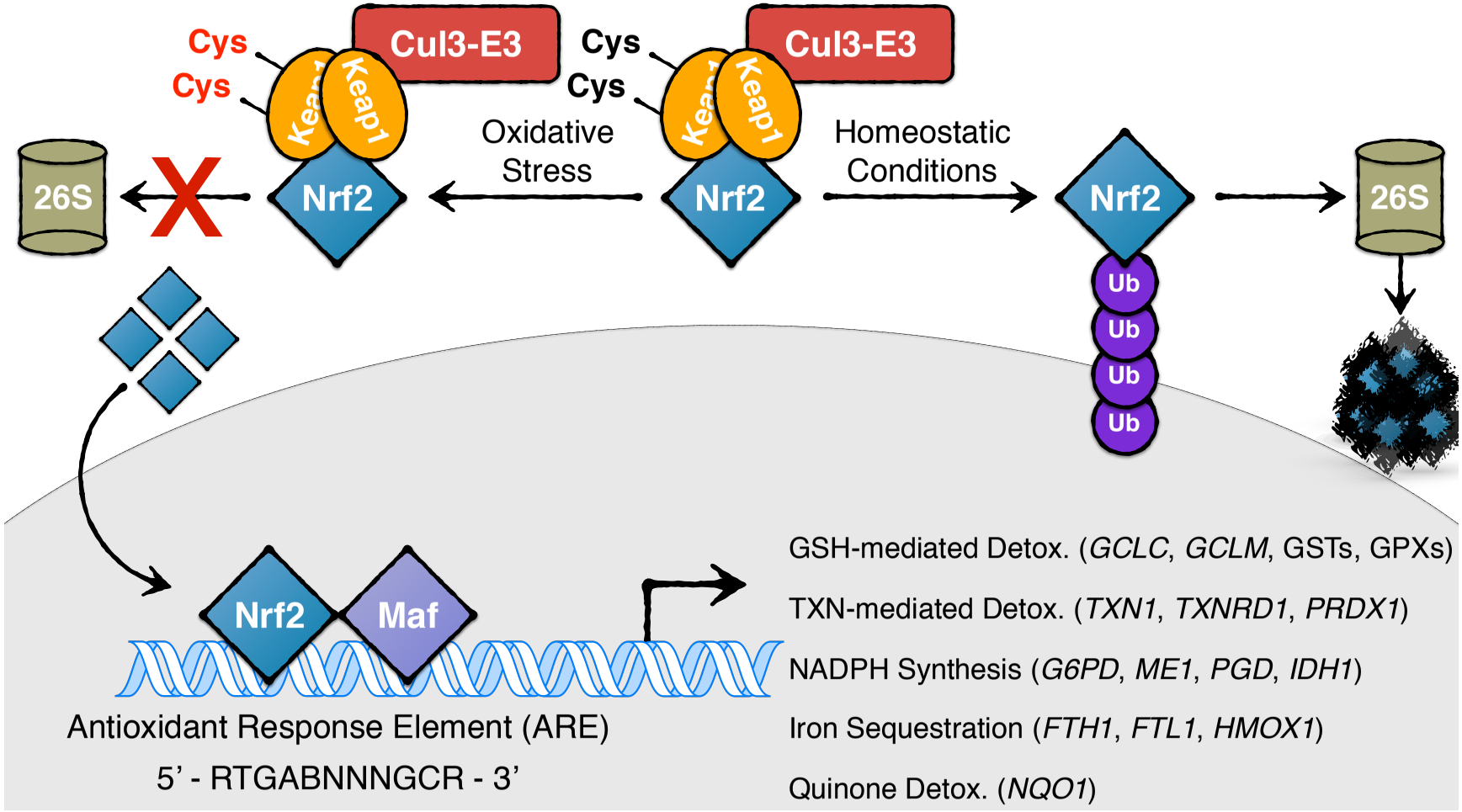


Figure 4: The Keap1-Nrf2 signaling pathway. Under homeostatic conditions, Nrf2 is constantly being degraded via the action of the Keap1-Cul3-E3 ubiquitin ligase complex. In conditions of oxidative stress, multiple cysteine residues in the Keap1 protein get oxidized, which causes a conformational change in the protein that decreases the strength of its interaction with Nrf2. This decreased interaction prevents the polyubiquitination and subsequent degradation of Nrf2. Consequently, free Nrf2 shuttles to the nucleus, where it heterodimerizes with small Maf family proteins and transcriptionally activates multiple cytoprotective genes. See text for details.

response (**Figure 4**).

iv. The Nrf2-driven antioxidant response

In conditions of high oxidative stress, Nrf2 induces the expression of a large repertoire of genes, including those encoding for enzymes involved in GSH synthesis and GSH-mediated ROS detoxification, TXN synthesis and TXN-mediated ROS detoxification, antioxidant enzymes, drug-metabolizing enzymes, drug efflux pumps, and multiple metabolic enzymes, including those involved in the pentose phosphate pathway (**Figure 4**) (reviewed in Suzuki and Yamamoto 2015). Below, I will discuss a select few of these Nrf2-regulated signaling arms of the cellular antioxidant response.

a. GSH synthesis and utilization

As discussed before, cells employ GSH to eliminate ROS via the GSH-GSSG cycle shown in **Figure 1A**. But where does GSH come from? GSH is synthesized *de novo* through two steps (Griffith and Mulcahy 2006). The first step - which is the rate-limiting step in the synthesis of GSH - is catalyzed by the heterodimeric glutathione-cysteine ligase enzymatic complex (GCL), which is composed of a catalytic subunit (GCLC) and a modifier subunit (GCLM). In this step, GCL catalyzes the reaction of glutamate with cysteine, generating glutamylcysteine. In the second step, glutathione synthetase (GSS) catalyzes the addition of glycine to glutamylcysteine, generating GSH. Remarkably, Nrf2 transcriptionally activates both *GCLC* and *GCLM*, thereby exerting major control over the synthesis of GSH (Wild et al., 1999). Moreover, Nrf2 regulates cysteine abundance via the transcriptional activation of the solute carrier family 7 member 11 (*SLC7A11*) gene, which encodes for the cystine/glutamate transporter XCT (Sasaki et al., 2002). In

addition to regulating the synthesis of GSH at multiple levels, Nrf2 regulates the regeneration of GSH through the transcriptional activation of *GSR* (Harvey et al., 2009). Beyond regulating the synthesis and regeneration of GSH at multiple levels, Nrf2 also regulates the utilization of GSH via the transcriptional activation of multiple GPXs and GSTs. For example, Nrf2 has been shown to directly activate the transcription of *GPX2*, *GSTA1*, *GSTA2*, *GSTA3*, *GSTA5, GSTM1*, *GSTM2*, *GSTM3*, and *GSTP1* (McMahon et al., 2001, Chanas et al., 2002, Thimmulappa et al., 2002, Banning et al., 2005, Singh et al., 2006, Lii et al., 2010). Therefore, the Nrf2 transcription factor regulates GSH synthesis and utilization at multiple levels via the transcriptional activation of several genes involved in each of these biological processes.

b. TXN synthesis and utilization

As discussed before, cells employ reduced TXN to eliminate ROS via the reduction of PRDX, which in turn results in the oxidation of TXN. Reduced TXN is regenerated via the action of TXNRD (**Figure 1B**). Remarkably, Nrf2 transcriptionally activates *TXN1*, *TXNRD1*, and *PRDX1* (Kim et al., 2001, Tanito et al., 2005, Sakurai et al., 2005, Kim et al., 2007), thereby controlling the TXN antioxidant pathway at multiple levels.

c. NADPH synthesis and utilization

As discussed before, both GSH-mediated and TXN-mediated ROS detoxification pathways require NADPH in order to regenerate GSH and reduced TXN, respectively. Therefore, NADPH synthesis is of utmost importance for what are arguably the two most important antioxidant pathways in the cell. Strikingly, Nrf2 acts as a major regulator of NADPH biosynthesis by transcriptionally activating the genes encoding four key

metabolic enzymes: glucose-6-phosphate dehydrogenase (G6PD), phosphogluconate dehydrogenase (PGD), malic enzyme (ME1), and isocitrate dehydrogenase 1 (IDH1) (Mitsuishi et al., 2012b). The most remarkable fact is that NADPH biosynthesis can only proceed in one of three ways: (1) via the pentose phosphate pathway (PPP), in which G6PD and PGD metabolize glucose-6-phosphate and 6-phosphogluconate in the 1st and 3rd steps of the metabolic pathway, respectively; (2) via the ME1-mediated conversion of malate to pyruvate; and (3) via the IDH1/2-mediated conversion of isocitrate to α-ketoglutarate. All of these metabolic reactions readily produce NADPH, which can be utilized for regenerating GSH and reduced TXN (Wu et al., 2011, Mitsuishi et al., 2012b, Dinkova-Kostova and Abramov 2015).

d. Iron sequestration

Heme molecules - which contain highly reactive ferrous Fe(II) molecules within the core of their structures - can serve as a major source of free radicals in the cell (reviewed in Gozzelino et al., 2010). Specifically, the Fe(II) molecules can catalyze the production of free radicals through the Fenton reaction, in which hydrogen peroxide is converted to the highly reactive hydroxyl free radical (Gutteridge 1986). These hydroxyl free radicals can trigger massive cellular toxicity by directly damaging multiple organelles and macromolecules, such as DNA, proteins, and lipids (Gozzelino et al., 2010, Gorrini et al., 2013). Under homeostatic conditions, Fe(II)-containing heme molecules are inserted in the heme pockets of hemoproteins; however, in the presence of oxidative stress, hemoproteins release these Fe(II)-containing heme molecules, which are subsequently broken down by an enzyme called heme oxygenase (HMOX1, also called HO-1),

thereby releasing potentially toxic Fe(II) molecules. Cells have evolved a built-in safety mechanism to deal with these Fe(II) molecules: the ferritin complex. This complex is composed of a ferritin heavy chain (FTH) and a ferritin light chain (FTL) and it catalyzes the conversion of Fe(II) into Fe(III) accompanied by the incorporation of Fe(III) within its own protein structure (Gozzelino et al., 2010).

As can be surmised from the paragraph above, HMOX1 and FTH/FTL appear to be antagonistic to each other. Under conditions of oxidative stress, HMOX1 is constantly promoting the release of highly reactive Fe(II), which in turn gets modified to its non-reactive Fe(III) form by the ferritin complex. Perhaps unexpectedly, Nrf2 directly regulates the levels of HMOX1, FTH, and FTL by transcriptionally activating the genes that encode them (Alam et al., 1999, Tsuji et al., 2000, Thimmulappa et al., 2002, Pietsch et al., 2003). Therefore, Nrf2 appears to orchestrate what looks to be a futile cycle during conditions of oxidative stress; the biological and evolutionary reasons behind the design of this cycle are not completely understood at this point (Gorrini et al., 2013). Notably, this futile cycle has the potential to create a unique Achilles heel in cells that are highly dependent on the Nrf2 pathway (this idea will be discussed later).

C. The Keap1-Nrf2 pathway in cancer biology

Several small- and large-scale sequencing studies carried out over the last decade have unequivocally demonstrated that the Keap1-Nrf2 pathway is frequently altered in a large proportion of human cancers (**Figure 5)**. Interestingly, alterations in this pathway can occur through both mutational and non-mutational mechanisms, underscoring the potential importance of this antioxidant program in multiple aspects of tumorigenesis.

Below, I will discuss the current state of the Keap1-Nrf2 mutational landscape in human cancers and the functional implications of these alterations in the genesis and progression of the disease.

i. *KEAP1* mutations in human cancer

Cancer-associated mutations in *KEAP1* were initially identified in 2006 by Masayuki Yamamoto and colleagues (Padmanabhan et al., 2006). In this study, they solved the structure of the DGR domain of murine Keap1 (see **Figure 3B**) and found that a Nrf2 peptide fragment that contained the ETGE domain (see **Figure 3A**) could strongly associate with Keap1 via its DGR domain. In addition to solving the structure of Keap1 and demonstrating the mechanistic basis for its interaction with Nrf2 at the atomic level, Yamamoto and colleagues identified *KEAP1* somatic missense mutations in the H1184 and H1648 human lung cancer cell lines. These mutations, which led to a substantial decrease in the strength of the Keap1-Nrf2 interaction, were present in the DGR domain of KEAP1. Therefore, cancer-associated somatic mutations that impair the interaction between these proteins have the potential to compromise the negative regulation imposed by Keap1 on Nrf2. The net result is hyperactivation of the Nrf2-driven antioxidant response pathway, which potentially endows cancer cells with the ability to counteract one of the major stress hallmarks of cancer - oxidative stress (Padmanabhan et al., 2006, Luo et al., 2009). A contemporaneous study by Shyam Biswal and colleagues identified multiple *KEAP1* somatic missense and indel mutations in multiple human lung cancer cell lines (~ 50%) and primary NSCLC patient specimens (~ 19%), all of which were located in the DGR or IVR domains of KEAP1 (Singh et al., 2006).

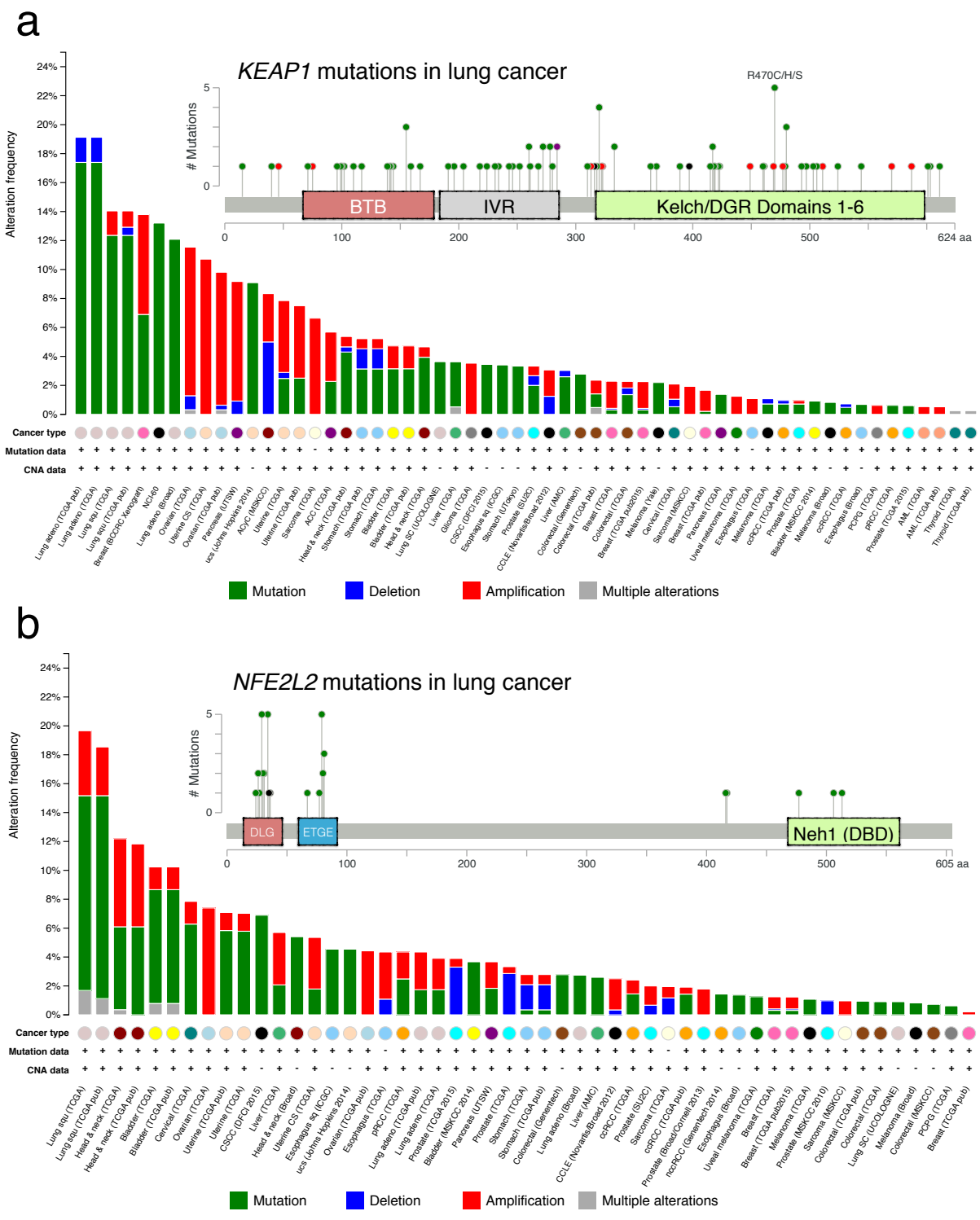


Figure 5: *KEAP1* and *NFE2L2* mutations across human cancer. Somatic cancer-associated mutations in (a) *KEAP1* and (b) *NFE2L2* predominantly affect functional protein domains. Data obtained from cBioPortal (http://www.cbioportal.org/) (Cerami et al., 2012 and Gao et al., 2013).

Moreover, Biswal and colleagues demonstrated that a large percentage of samples (~ 61% of lung cancer cell lines and ~ 41% of patient samples) showed clear evidence of loss of heterozygosity of the remaining wild type allele of *KEAP1*, further supporting the hypothesis that Keap1-mediated negative regulation of Nrf2 is a potential tumor suppressor mechanism in human cancer (Singh et al., 2006). Again, the net result of these mutations is hyperactivation of the Nrf2 antioxidant program. Subsequent studies by Tsutomu Ohta, Masayuki Yamamoto, and colleagues substantially extended these sequencing efforts and confirmed that somatic mutation of *KEAP1* is a relatively common event in human lung cancer (Ohta et al., 2008). Importantly, they demonstrated that *KEAP1* mutations confer resistance to cisplatin via the NRF2-dependent activation of multiple genes encoding drug efflux pumps and other detoxifying enzymes, thereby establishing a functional link between hyperactivation of the NRF2 antioxidant response pathway and chemoresistance (Ohta et al., 2008). In addition to these cancer-associated somatic mutations that lead to loss of function of *KEAP1*, multiple studies have reported that *KEAP1* can also be inactivated via promoter hypermethylation (Wang et al., 2008). More recently, multiple large-scale cancer genome sequencing studies have identified somatic mutations in *KEAP1* in ~ 17% of lung adenocarcinomas and ~ 12% of squamous cell lung cancers (LUAD TCGA 2014, Hammerman et al., 2012). Notably, these mutations are almost invariably found in functionally important domains of the KEAP1 protein, such as the BTB, IVR, and DGR domains (**Figure 5A**). This data strongly suggests that *KEAP1* is a potent tumor suppressor in lung cancer.

ii. *NFE2L2* mutations in human cancer

Cancer-associated mutations in *NFE2L2* were initially identified in 2008 by Masayuki Yamamoto and colleagues (Shibata et al., 2008a). By analyzing a large collection of human cancer cell lines and primary lung cancer specimens, they identified multiple somatic missense mutations in the *NFE2L2* coding region. Approximately 11% of primary lung cancer specimens harbored missense mutations in *NFE2L2*. These mutations were found to be mutually exclusive with *KEAP1* mutations, supporting the hypothesis that somatic mutation of *KEAP1* triggers hyperactivation of NRF2, and that this is the main biological property that gets selected for in the context of tumorigenesis. Interestingly, they identified one NSCLC cell line that harbored a single mutant allele of *NFE2L2*, and one squamous cell lung cancer cell line that harbored a focal amplification of the *NFE2L2* locus, suggesting that tumors can acquire somatic activation of the NRF2 pathway through multiple genetic mechanisms. Notably, somatic *NFE2L2* mutations were found to frequently alter the DLG or ETGE motifs, once again suggesting that the abolishment of KEAP1-mediated negative regulation and subsequent hyperactivation of the NRF2 antioxidant pathway is strongly selected for during tumorigenesis. Furthermore, they experimentally demonstrated that these mutant *NFE2L2* alleles do impair KEAP1-mediate negative regulation and that siRNA-mediated downregulation of NRF2 sensitizes cancer cells to oxidative stress and cisplatin (Shibata et al., 2008a). More recently, multiple large-scale cancer genome sequencing studies have identified somatic mutations in *NFE2L2* in ~ 3% of lung adenocarcinomas and ~ 19% of squamous cell lung cancers (LUAD TCGA 2014, Hammerman et al., 2012). Notably, these mutations are almost invariably found in functionally important

domains of the NRF2 protein, such as the DLG and ETGE domains (**Figure 5B**). This data strongly suggests that *NFE2L2* mutation and subsequent hyperactivation of the NRF2 antioxidant program can have major oncogenic properties in the context of tumorigenesis.

The identification and functional characterization of cancer-associated *KEAP1* and *NFE2L2* mutations strongly suggests that the Keap1-Nrf2 pathway can play a very important role in the context of tumorigenesis. Indeed, it is estimated that ~ 20-30% of all lung cancers harbor mutations in components of the Keap1-Nrf2 pathway (LUAD TCGA 2014, Hammerman et al., 2012). Beyond lung malignancies, several other cancer types, including head and neck cancers and cancers of the bladder, liver, stomach, prostate, and breast, among others, harbor mutations in either *KEAP1* or *NFE2L2* , the vast majority of which are predicted to lead to hyperactivation of the NRF2 antioxidant pathway (**Figure 5**) (Nioi and Nguyen 2007, Shibata et al., 2008a, Shibata et al., 2008b, Kim et al., 2010, Zhang et al., 2010, Eichenmüller et al., 2014, Tanase et al., 2015, and others). Importantly, hyperactivation of the NRF2 antioxidant pathway can happen in the absence of somatic mutations in *KEAP1* or *NFE2L2*. For example, multiple studies have shown *NFE2L2* overexpression in many different types of cancers, such as endometrial cancer and lung cancer (Jiang et al., 2010, Solis et al., 2010). Moreover, pathway hyperactivation can also be achieved via somatic mutation of additional pathway components, such as *CUL3*. Indeed, loss of function mutations in *CUL3* have been observed in ~ 3% of human lung adenocarcinomas and ~ 7% of squamous cell lung cancers (LUAD TCGA 2014, Hammerman et al., 2012). These and multiple other studies have sparked a great amount of interest in understanding the role of this

pathway during cancer initiation, progression, and therapeutic response. Below, I will describe a select few of these studies. For comprehensive cancer-centric reviews, see Hayes and McMahon 2009, Taguchi et al., 2011, Sporn and Liby 2012, Jaramillo and Zhang 2013, and Suzuki and Yamamoto 2015.

iii. Functional studies of the Keap1-Nrf2 pathway in cancer

In addition to the pioneering studies described above, multiple studies over the last few years have demonstrated the functional importance of the Keap1-Nrf2 pathway in several aspects of cancer biology. One notable study by David Tuveson and colleagues demonstrated that potent oncogenes can induce the transcription of *NFE2L2* (DeNicola et al., 2011). By employing multiple GEMM-based models harboring either (1) Cre-inducible alleles of oncogenic *Kras*^{G12D} or *Braf*^{V619E}, or (2) a hormone-inducible allele of *Myc* (*Myc*^{ERT2}), they demonstrated that activation of these endogenous oncogenes led to a substantial transcriptional induction of the endogenous *Nfe2l2* gene. Therefore, oncogene activation led to a significant, Nrf2-dependent decrease in the intracellular levels of ROS, which was accompanied by a high GSH/GSSG ratio indicative of a highly reduced intracellular environment. As expected, endogenous expression of these oncogenes led to Nrf2-mediated induction of the antioxidant program, as gauged by the transcriptional induction of multiple Nrf2 target genes, including *Hmox1*, *Nqo1*, *Gclc*, and *Gclm*. Mechanistically, they demonstrated that oncogenic Kras^{G12D} activates the Raf-Mek-Erk arm of the MAPK signaling pathway, which in turn leads to the activation of the Jun and Myc transcription factors. Jun and Myc directly activate the transcription of Nrf2 by binding to a series of regulatory elements located in the Nrf2 promoter.

Therefore, the elegant studies of Tuveson and colleagues led to the elucidation of a novel oncogene-induced ROS detoxification pathway, which can be summarized as: Kras → Braf → Mek → Erk → Jun/Myc → Nrf2 → ROS detoxification. Furthermore, using well established pre-clinical GEMMs of cancer that were deficient for endogenous *Nfe2l2* (germline knockout), they went on to demonstrate that both pancreatic and lung adenocarcinomas are highly dependent on this novel oncogene-induced Nrf2-mediated ROS detoxification pathway. Hyperactivation of the NRF2 antioxidant program was also observed in multiple human pancreatic cancer specimens, *in the absence of any somatic mutations in NFE2L2 or KEAP1*. That both of these aggressive human malignancies become highly dependent on NRF2-mediated ROS detoxification to survive in conditions of high oxidative stress strongly resonates with the concept of non-oncogene addiction proposed by Stephen Elledge and colleagues (Luo et al., 2009). One very important methodological side-note is that these phenotypes were only seen when using GEMM-based models that are driven by activation of *endogenous* oncogenes. In fact, when they utilized overexpression constructs encoding each of these potent oncogenes, they obtained completely opposite results, presumably due to the supraphysiological levels of expression obtained when using exogenous retroviral constructs (DeNicola et al., 2011).

Another notable study that demonstrated the functional importance of the Nrf2 antioxidant program in promoting several aspects of the cancer phenotype was published by the Yamamoto group (Mitsuishi et al., 2012b). By employing a combination of gene expression profiling, chromatin immunoprecipitation followed by high-throughput sequencing (ChIP-Seq), and extensive metabolomic profiling using mass spectrometry,

Yamamoto and colleagues demonstrated that NRF2 plays a critical role in promoting multiple anabolic pathways by transcriptionally activating a repertoire of genes involved in the PPP pathway (*G6PD*, *PGD*, *TKT*, and *TALDO1*), de novo nucleotide biosynthesis pathway (*PPAT* and *MTHFD2*), and NADPH production (*ME1* and *IDH1*). The net effect of this NRF2-dependent metabolic reprogramming is an increase in the biosynthetic and proliferative capacities of cancer cells accompanied by redox homeostasis facilitated by the NADPH-dependent regeneration of GSH (and TXN) (Mitsuishi et al., 2012b).

Recently, Tak Mak and colleagues demonstrated that the GSH and TXN arms of the Nrf2 antioxidant program strongly cooperate to drive tumor initiation and progression in multiple GEMMs of cancer (Harris et al., 2015). By employing the well-established MMTV-PyMT and KP mouse models of breast cancer and sarcoma, respectively, Tak Mak and colleagues demonstrated that genetic deletion of *Gclm* (which led to a ~ 75% reduction in GSH levels) markedly impaired tumor initiation, presumably because the cells were experiencing lethal conditions of high oxidative stress. Similarly, pharmacological inhibition of GSH synthesis also impaired tumor initiation, underscoring the cancer promoting role of this potent cellular antioxidant. In addition to demonstrating the pro-tumorigenic role of GSH-mediated ROS detoxification *in vivo*, Tak Mak and colleagues also showed that the cellular antioxidant TXN is able to compensate for loss of GSH in cancer cells. Indeed, combined pharmacological inhibition of GSH and TXN synthesis exhibited strong synergistic effects in blocking tumor growth, underscoring the importance of these two major antioxidant pathways in promoting multiple aspects of the cancer phenotype and lending support for novel therapeutic strategies aimed at simultaneously targeting these two potent cellular antioxidants. Notably, the authors

found evidence of *GCLM* overexpression in multiple breast cancer patient cohorts, underscoring the importance of this antioxidant pathway in promoting tumorigenesis.

D. Synopsis and outlook

Collectively, the studies described above (as well as other contemporaneous studies, such as Sayin et al., 2014, DeNicola et al., 2015, and Piskounova et al., 2015) have convincingly demonstrated that the Nrf2 antioxidant program and cellular antioxidants in general can play a major role in promoting several aspects of the cancer phenotype. These studies have sparked a great amount of interest in understanding the contributions of the Keap1-Nrf2 pathway during carcinogenesis, tumor progression, and therapeutic response. A multitude of outstanding questions in the field remain unanswered, such as the ones listed below:

1) At which stage(s) of tumor progression is Nrf2 hyperactivation critical?

2) Similarly, what is the stage(s) of tumor progression that Keap1 suppresses?

3) Are any of the above situations context-specific? In other words, are there situations where Nrf2 suppresses or Keap1 promotes certain aspects of tumorigenesis?

4) Does hyperactivation of the Nrf2 antioxidant pathway represent an Achilles heel that can be exploited for therapeutic purposes?

In **Chapter III** of this thesis, I present evidence that suggests that hyperactivation of the Nrf2 antioxidant pathway in cancer cells does indeed create an Achilles heel that might be exploited therapeutically for selectively targeting tumors that are highly dependent on this antioxidant pathway.

REFERENCES

Aas, T., Børresen, A.-L., Geisler, S., Smith-Sørensen, B., Johnsen, H., Varhaug, J.E., Akslen, L.A., and Lønning, P.E. (1996). Specific P53 mutations are associated with de novo resistance to doxorubicin in breast cancer patients. Nat Med 2, 811–814.

Abeshouse, A., Ahn, J., Akbani, R., Ally, A., Amin, S., Andry, C.D., Annala, M., Aprikian, A., Armenia, J., Arora, A., et al. (2015). The Molecular Taxonomy of Primary Prostate Cancer. Cell 163, 1011–1025.

Adams, D.J., Biggs, P.J., Cox, T., Davies, R., van der Weyden, L., Jonkers, J., Smith, J., Plumb, B., Taylor, R., Nishijima, I., et al. (2004). Mutagenic insertion and chromosome engineering resource (MICER). Nat Genet 36, 867–871.

Adams, J.M., Harris, A.W., Pinkert, C.A., Corcoran, L.M., Alexander, W.S., Cory, S., Palmiter, R.D., and Brinster, R.L. (1985). The c-myc oncogene driven by immunoglobulin enhancers induces lymphoid malignancy in transgenic mice. Nature 318, 533–538.

Akbani, R., Akdemir, K.C., Aksoy, B.A., Albert, M., Ally, A., Amin, S.B., Arachchi, H., Arora, A., Auman, J.T., Ayala, B., et al. (2015). Genomic Classification of Cutaneous Melanoma. Cell 161, 1681–1696.

Alam, J., Stewart, D., Touchard, C., Boinapally, S., Choi, A.M., and Cook, J.L. (1999). Nrf2, a Cap“n”Collar transcription factor, regulates induction of the heme oxygenase-1 gene. J Biol Chem 274, 26071–26078.

Andersson, A.K., Ma, J., Wang, J., Chen, X., Gedman, A.L., Dang, J., Nakitandwe, J., Holmfeldt, L., Parker, M., Easton, J., et al. (2015). The landscape of somatic mutations in infant MLL-rearranged acute lymphoblastic leukemias. Nat Genet 47, 330–337.

Andres, A.C., Schönenberger, C.A., Groner, B., Hennighausen, L., LeMeur, M., and Gerlinger, P. (1987). Ha-ras oncogene expression directed by a milk protein gene promoter: tissue specificity, hormonal regulation, and tumor induction in transgenic mice. Proc Natl Acad Sci USA 84, 1299–1303.

Andrews, B.J., Proteau, G.A., Beatty, L.G., and Sadowski, P.D. (1985). The FLP recombinase of the 2 micron circle DNA of yeast: interaction with its target sequences. Cell 40, 795–803.

Andrews, N.C., Erdjument-Bromage, H., Davidson, M.B., Tempst, P., and Orkin, S.H. (1993a). Erythroid transcription factor NF-E2 is a haematopoietic-specific basic-leucine zipper protein. Nature 362, 722–728.

Andrews, N.C., Kotkow, K.J., Ney, P.A., Erdjument-Bromage, H., Tempst, P., and Orkin, S.H. (1993b). The ubiquitous subunit of erythroid transcription factor NF-E2 is a small basic-leucine zipper protein related to the v-maf oncogene. Proc Natl Acad Sci USA 90, 11488–11492.

Arnould, S., Chames, P., Perez, C., Lacroix, E., Duclert, A., Epinat, J.-C., Stricher, F., Petit, A.-S., Patin, A., Guillier, S., et al. (2006). Engineering of large numbers of highly specific homing endonucleases that induce recombination on novel DNA targets. J Mol Biol 355, 443–458.

Bagchi, A., Papazoglu, C., Wu, Y., Capurso, D., Brodt, M., Francis, D., Bredel, M., Vogel, H., and Mills, A.A. (2007). CHD5 is a tumor suppressor at human 1p36. Cell 128, 459–475.

Baker, S.J., Fearon, E.R., Nigro, J.M., Hamilton, S.R., Preisinger, A.C., Jessup, J.M., vanTuinen, P., Ledbetter, D.H., Barker, D.F., Nakamura, Y., et al. (1989). Chromosome 17 deletions and p53 gene mutations in colorectal carcinomas. Science 244, 217–221.

Banning, A., Deubel, S., Kluth, D., Zhou, Z., and Brigelius-Flohe, R. (2005). The GI-GPx gene is a target for Nrf2. Mol Cell Biol 25, 4914–4923.

Barker, N., van Es, J.H., Kuipers, J., Kujala, P., den Born, van, M., Cozijnsen, M., Haegebarth, A., Korving, J., Begthel, H., Peters, P.J., et al. (2007). Identification of stem cells in small intestine and colon by marker gene Lgr5. Nature 449, 1003–1007.

Barlow, C., Hirotsune, S., Paylor, R., Liyanage, M., Eckhaus, M., Collins, F., Shiloh, Y., Crawley, J.N., Ried, T., Tagle, D., et al. (1996). Atm-deficient mice: a paradigm of ataxia telangiectasia. Cell 86, 159–171.

Barlow, J.L., Drynan, L.F., Hewett, D.R., Holmes, L.R., Lorenzo-Abalde, S., Lane, A.L., Jolin, H.E., Pannell, R., Middleton, A.J., Wong, S.H., et al. (2009). A p53-dependent mechanism underlies macrocytic anemia in a mouse model of human 5q- syndrome. Nat Med 16, 59–66.

Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., Romero, D.A., and Horvath, P. (2007). CRISPR Provides Acquired Resistance Against Viruses in Prokaryotes. Science 315, 1709–1712.

Barrangou, R. (2014). Cas9 Targeting and the CRISPR Revolution. Science 344, 707–708.

Bartek, J., Iggo, R., Gannon, J., and Lane, D.P. (1990). Genetic and immunochemical analysis of mutant p53 in human breast cancer cell lines. Oncogene 5, 893–899.

Bartel, D.P. (2004). MicroRNAs: genomics, biogenesis, mechanism, and function. Cell 116, 281–297.

Bartel, D.P. (2009). MicroRNAs: target recognition and regulatory functions. Cell 136, 215–233.

Bassett, A.R., Tibbit, C., Ponting, C.P., and Liu, J.-L. (2013). Highly efficient targeted mutagenesis of Drosophila with the CRISPR/Cas9 system. Cell Rep 4, 220–228.

Bayascas, J.R., Sakamoto, K., Armit, L., Arthur, J.S.C., and Alessi, D.R. (2006). Evaluation of approaches to generation of tissue-specific knock-in mice. J Biol Chem 281, 28772–28781.

Beard, C., Hochedlinger, K., Plath, K., Wutz, A., and Jaenisch, R. (2006). Efficient method to generate single-copy transgenic mice by site-specific integration in embryonic stem cells. Genesis 44, 23–28.

Bedell, V.M., Wang, Y., Campbell, J.M., Poshusta, T.L., Starker, C.G., Krug, R.G., II, Tan, W., Penheiter, S.G., Ma, A.C., Leung, A.Y.H., et al. (2012). In vivo genome editing using a high-efficiency TALEN system. Nature 491, 114–118.

Bernardi, R., Grisendi, S., and Pandolfi, P.P. (2002). Modelling haematopoietic malignancies in the mouse and therapeutical implications. Oncogene 21, 3445–3458.

Berns, K., Hijmans, E.M., Mullenders, J., Brummelkamp, T.R., Velds, A., Heimerikx, M., Kerkhoven, R.M., Madiredjo, M., Nijkamp, W., Weigelt, B., et al. (2004). A large-scale RNAi screen in human cells identifies new components of the p53 pathway. Nature 428, 431–437.

Bernstein, E., Caudy, A.A., Hammond, S.M., and Hannon, G.J. (2001). Role for a bidentate ribonuclease in the initiation step of RNA interference. Nature 409, 363–366.

Beronja, S., and Fuchs, E. (2013). RNAi-mediated gene function analysis in skin. Methods Mol Biol 961, 351–361.

Beronja, S., Janki, P., Heller, E., Lien, W.-H., Keyes, B.E., Oshimori, N., and Fuchs, E. (2013). RNAi screens in mice identify physiological regulators of oncogenic growth. Nature 501, 185–190.

Beronja, S., Livshits, G., Williams, S., and Fuchs, E. (2010). Rapid functional dissection of genetic networks via tissue-specific transduction and RNAi in mouse embryos. Nat Med 16, 821–827.

Beroukhim, R., Mermel, C.H., Porter, D., Wei, G., Raychaudhuri, S., Donovan, J., Barretina, J., Boehm, J.S., Dobson, J., Urashima, M., et al. (2010). The landscape of somatic copy-number alteration across human cancers. Nature 463, 899–905.

Bhattacharyya, S.N., Habermacher, R., Martine, U., Closs, E.I., and Filipowicz, W. (2006). Relief of microRNA-mediated translational repression in human cells subjected to stress. Cell 125, 1111–1124.

Bi, L., Okabe, I., Bernard, D.J., Wynshaw-Boris, A., and Nussbaum, R.L. (1999). Proliferative Defect and Embryonic Lethality in Mice Homozygous for a Deletion in the p110α Subunit of Phosphoinositide 3-Kinase. Journal of Biological Chemistry 274, 10963–10968.

Bibikova, M., Beumer, K., Trautman, J.K., and Carroll, D. (2003). Enhancing Gene Targeting with Designed Zinc Finger Nucleases. Science 300, 764–764.

Bibikova, M., Golic, M., Golic, K.G., and Carroll, D. (2002). Targeted chromosomal cleavage and mutagenesis in Drosophila using zinc-finger nucleases. 161, 1169–1175.

Bikard, D., Jiang, W., Samai, P., Hochschild, A., Zhang, F., and Marraffini, L.A. (2013). Programmable repression and activation of bacterial gene expression using an engineered CRISPR-Cas system. Nucleic Acids Res 41, 7429–7437.

Birsoy, K., Wang, T., Chen, W.W., Freinkman, E., Abu-Remaileh, M., and Sabatini, D.M. (2015). An Essential Role of the Mitochondrial Electron Transport Chain in Cell Proliferation Is to Enable Aspartate Synthesis. Cell 162, 540–551.

Blandino, G., Levine, A.J., and Oren, M. (1999). Mutant p53 gain of function: differential effects of different p53 mutants on resistance of cultured cells to chemotherapy. 18, 477–485.

Blasco, R.B., Karaca, E., Ambrogio, C., Cheong, T.-C., Karayol, E., Minero, V.G., Voena, C., and Chiarle, R. (2014). Simple and Rapid In Vivo Generation of Chromosomal Rearrangements using CRISPR/Cas9 Technology. Cell Rep 9, 1219–1227.

Boch, J., and Bonas, U. (2010). XanthomonasAvrBs3 Family-Type III Effectors: Discovery and Function. Annu. Rev. Phytopathol. 48, 419–436.

Boch, J., Scholze, H., Schornack, S., Landgraf, A., Hahn, S., Kay, S., Lahaye, T., Nickstadt, A., and Bonas, U. (2009). Breaking the Code of DNA Binding Specificity of TAL-Type III Effectors. Science 326, 1509–1512.

Boehm, J.S., and Hahn, W.C. (2011). Towards systematic functional characterization of cancer genomes. Nat Rev Genet 12, 487–498.

Boehmelt, G., Walker, A., Kabrun, N., Mellitzer, G., Beug, H., Zenke, M., and Enrietto, P.J. (1992). Hormone-regulated v-rel estrogen receptor fusion protein: reversible induction of cell transformation and cellular gene expression. Embo J 11, 4641–4652.

Bohnsack, M.T., Czaplinski, K., and Gorlich, D. (2004). Exportin 5 is a RanGTP-dependent dsRNA-binding protein that mediates nuclear export of pre-miRNAs. Rna 10, 185–191.

Bollag, G., Tsai, J., Zhang, J., Zhang, C., Ibrahim, P., Nolop, K., and Hirth, P. (2012). Vemurafenib: the first drug approved for BRAF-mutant cancer. Nature Reviews Drug Discovery 11, 873–886.

Bolotin, A., Quinquis, B., Sorokin, A., and Ehrlich, S.D. (2005). Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. Microbiology (Reading, Engl.) 151, 2551–2561.

Bossi, G., Lapi, E., Strano, S., Rinaldo, C., Blandino, G., and Sacchi, A. (2006). Mutant p53 gain of function: reduction of tumor malignancy of human cancer cell lines through abrogation of mutant p53 expression. Oncogene 25, 304–309.

Boveri, T. (1902). Über mehrpolige Mitosen als Mittel zur Analyse des Zellkerns.

Boveri, T. (2007). Concerning the Origin of Malignant Tumours by Theodor Boveri. Translated and annotated by Henry Harris. Journal of Cell Science 121, 1–84.

Brannan, C.I., Perkins, A.S., Vogel, K.S., Ratner, N., Nordlund, M.L., Reid, S.W., Buchberg, A.M., Jenkins, N.A., Parada, L.F., and Copeland, N.G. (1994). Targeted disruption of the neurofibromatosis type-1 gene leads to developmental abnormalities in heart and various neural crest-derived tissues. Genes Dev 8, 1019–1029.

Brauch, H., Johnson, B., Hovis, J., Yano, T., Gazdar, A., Pettengill, O.S., Graziano, S., Sorenson, G.D., Poiesz, B.J., and Minna, J. (1987). Molecular analysis of the short arm of chromosome 3 in small-cell and non-small-cell carcinoma of the lung. Nejm 317, 1109–1113.

Bressac, B., Kew, M., Wands, J., and Ozturk, M. (1991). Selective G to T mutations of p53 gene in hepatocellular carcinoma from southern Africa. Nature 350, 429–431.

Bric, A., Miething, C., Bialucha, C.U., Scuoppo, C., Zender, L., Krasnitz, A., Xuan, Z., Zuber, J., Wigler, M., Hicks, J., et al. (2009). Functional Identification of Tumor-Suppressor Genes through an In Vivo RNA Interference Screen in a Mouse Lymphoma Model. Cancer Cell 16, 324–335.

Brinster, R.L., Chen, H.-Y., Messing, A., Van Dyke, T., Levine, A.J., and Palmiter, R.D. (1984). Transgenic mice harboring SV40 T-antigen genes develop characteristic brain tumors. Cell 37, 367–379.

Brosh, R., and Rotter, V. (2009). When mutants gain new powers: news from the mutant p53 field. Nat Rev Cancer 9, 701–713.

Brouns, S.J.J., Jore, M.M., Lundgren, M., Westra, E.R., Slijkhuis, R.J.H., Snijders, A.P.L., Dickman, M.J., Makarova, K.S., Koonin, E.V., and van der Oost, J. (2008). Small CRISPR RNAs Guide Antiviral Defense in Prokaryotes. Science 321, 960–964.

Brunelle, J.K., Bell, E.L., Quesada, N.M., Vercauteren, K., Tiranti, V., Zeviani, M., Scarpulla, R.C., and Chandel, N.S. (2005). Oxygen sensing requires mitochondrial ROS but not oxidative phosphorylation. Cell Metabolism 1, 409–414.

Brunet, E., Simsek, D., Tomishima, M., DeKelver, R., Choi, V.M., Gregory, P., Urnov, F., Weinstock, D.M., and Jasin, M. (2009). Chromosomal translocations induced at specified loci in human stem cells. Proceedings of the National Academy of Sciences 106, 10620–10625.

Buchholz, F., Refaeli, Y., Trumpp, A., and Bishop, J.M. (2000). Inducible chromosomal translocation of AML1 and ETO genes through Cre/loxP-mediated recombination in the mouse. EMBO Rep 1, 133–139.

Bulua, A.C., Simon, A., Maddipati, R., Pelletier, M., Park, H., Kim, K.-Y., Sack, M.N., Kastner, D.L., and Siegel, R.M. (2011). Mitochondrial reactive oxygen species promote production of proinflammatory cytokines and are elevated in TNFR1-associated periodic syndrome (TRAPS). J Exp Med 208, 519–533.

Burgess, D.J., Doles, J., Zender, L., Xue, W., Ma, B., McCombie, W.R., Hannon, G.J., Lowe, S.W., and Hemann, M.T. (2008). Topoisomerase levels determine chemotherapy response in vitro and in vivo. Proc Natl Acad Sci USA 105, 9053–9058.

Burk, O., and Klempnauer, K.H. (1991). Estrogen-dependent alterations in differentiation state of myeloid cells caused by a v-myb/estrogen receptor fusion protein. Embo J 10, 3713–3719.

Cambridge, S.B., Geissler, D., Calegari, F., Anastassiadis, K., Hasan, M.T., Stewart, A.F., Huttner, W.B., Hagen, V., and Bonhoeffer, T. (2009). Doxycycline-dependent photoactivated gene expression in eukaryotic systems. Nat Methods 6, 527–531.

Cancer Genome Atlas Network (2012a). Comprehensive molecular characterization of human colon and rectal cancer. Nature 487, 330–337.

Cancer Genome Atlas Network (2012b). Comprehensive molecular portraits of human breast tumours. Nature 490, 61–70.

Cancer Genome Atlas Research Network (2013). Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. N Engl J Med 368, 2059–2074.

Cancer Genome Atlas Research Network (2014). Comprehensive molecular characterization of gastric adenocarcinoma. Nature 513, 202–209.

Cancer Genome Atlas Research Network (2014). Comprehensive molecular profiling of lung adenocarcinoma. Nature 511, 543–550.

Canver, M.C., Bauer, D.E., Dass, A., Yien, Y.Y., Chung, J., Masuda, T., Maeda, T., Paw, B.H., and Orkin, S.H. (2014). Characterization of Genomic Deletion Efficiency Mediated by Clustered Regularly Interspaced Palindromic Repeats (CRISPR)/Cas9 Nuclease System in Mammalian Cells. Journal of Biological Chemistry 289, 21312–21324.

Carbery, I.D., Ji, D., Harrington, A., Brown, V., Weinstein, E.J., Liaw, L., and Cui, X. (2010). Targeted genome modification in mice using zinc-finger nucleases. Genetics 186, 451–459.

Carlson, D.F., Tan, W., Lillico, S.G., Stverakova, D., Proudfoot, C., Christian, M., Voytas, D.F., Long, C.R., Whitelaw, C.B.A., and Fahrenkrug, S.C. (2012). Efficient TALEN-mediated gene knockout in livestock. Proceedings of the National Academy of Sciences 109, 17382–17387.

Carmell, M.A., Zhang, L., Conklin, D.S., Hannon, G.J., and Rosenquist, T.A. (2003). Germline transmission of RNAi in mice. Nat Struct Biol 10, 91–92.

Cermak, T., Doyle, E.L., Christian, M., Wang, L., Zhang, Y., Schmidt, C., Baller, J.A., Somia, N.V., Bogdanove, A.J., and Voytas, D.F. (2011). Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. Nucleic Acids Res 39, e82–e82.

Chan, K., Lu, R., Chang, J.C., and Kan, Y.W. (1996). NRF2, a member of the NFE2 family of transcription factors, is not essential for murine erythropoiesis, growth, and development. Proc Natl Acad Sci USA 93, 13943–13948.

Chanas, S.A., Jiang, Q., McMahon, M., McWalter, G.K., McLellan, L.I., Elcombe, C.R., Henderson, C.J., Wolf, C.R., Moffat, G.J., Itoh, K., et al. (2002). Loss of the Nrf2 transcription factor causes a marked reduction in constitutive and inducible expression of the glutathione S-transferase Gsta1, Gsta2, Gstm1, Gstm2, Gstm3 and Gstm4 genes in the livers of male and female mice. Biochem. J. 365, 405–416.

Chandel, N.S., McClintock, D.S., Feliciano, C.E., Wood, T.M., Melendez, J.A., Rodriguez, A.M., and Schumacker, P.T. (2000). Reactive oxygen species generated at mitochondrial complex III stabilize hypoxia-inducible factor-1alpha during hypoxia: a mechanism of O2 sensing. J Biol Chem 275, 25130–25138.

Chandrasegaran, S., and Carroll, D. (2015). Origins of Programmable Nucleases for Genome Engineering. J Mol Biol.

Chang, N., Sun, C., Gao, L., Zhu, D., Xu, X., Zhu, X., Xiong, J.-W., and Xi, J.J. (2013). Genome editing with RNA-guided Cas9 nuclease in zebrafish embryos. Cell Research 23, 465–472.

Chapman, P.B., Hauschild, A., Robert, C., Haanen, J.B., Ascierto, P., Larkin, J., Dummer, R., Garbe, C., Testori, A., Maio, M., et al. (2011). Improved Survival with Vemurafenib in Melanoma with BRAF V600E Mutation. N Engl J Med 364, 2507–2516.

Chari, R., Mali, P., Moosburner, M., and Church, G.M. (2015). Unraveling CRISPR-Cas9 genome engineering parameters via a library-on-library approach. Nat Meth 12, 823–826.

Chen, C., Liu, Y., Lu, C., Cross, J.R., Morris, J.P., Shroff, A.S., Ward, P.S., Bradner, J.E., Thompson, C., and Lowe, S.W. (2013). Cancer-associated IDH2 mutants drive an acute myeloid leukemia that is susceptible to Brd4 inhibition. Genes Dev 27, 1974–1985.

Chen, C., Liu, Y., Rappaport, A.R., Kitzing, T., Schultz, N., Zhao, Z., Shroff, A.S., Dickins, R.A., Vakoc, C.R., Bradner, J.E., et al. (2014a). MLL3 is a haploinsufficient 7q tumor suppressor in acute myeloid leukemia. Cancer Cell 25, 652–665.

Chen, S., Sanjana, N.E., Zheng, K., Shalem, O., Lee, K., Shi, X., Scott, D.A., Song, J., Pan, J.Q., Weissleder, R., et al. (2015). Genome-wide CRISPR Screen in a Mouse Model of Tumor Growth and Metastasis. Cell.

Chen, Z., Cheng, K., Walton, Z., Wang, Y., Ebi, H., Shimamura, T., Liu, Y., Tupper, T., Ouyang, J., Li, J., et al. (2012). A murine lung cancer co-clinical trial identifies genetic modifiers of therapeutic response. Nature 483, 613–617.

Chen, Z., Fillmore, C.M., Hammerman, P.S., Kim, C.F., and Wong, K.-K. (2014b). Non-small-cell lung cancers: a heterogeneous set of diseases. Nat Rev Cancer 14, 535–546.

Cheng, A.W., Wang, H., Yang, H., Shi, L., Katz, Y., Theunissen, T.W., Rangarajan, S., Shivalila, C.S., Dadon, D.B., and Jaenisch, R. (2013). Multiplexed activation of endogenous genes by CRISPR-on, an RNA-guided transcriptional activator system. Cell Research 23, 1163–1171.

Chiba, I., Takahashi, T., Nau, M.M., D'Amico, D., Curiel, D.T., Mitsudomi, T., Buchhagen, D.L., Carbone, D., Piantadosi, S., Koga, H., et al. (1990). Mutations in the p53 gene are frequent in primary, resected non-small cell lung cancer. Oncogene 5, 1603–1610.

Chin, K.V., Ueda, K., Pastan, I., and Gottesman, M.M. (1992). Modulation of activity of the promoter of the human MDR1 gene by Ras and p53. Science 255, 459–462.

Chin, L., Tam, A., Pomerantz, J., Wong, M., Holash, J., Bardeesy, N., Shen, Q., O'Hagan, R., Pantginis, J., Zhou, H., et al. (1999). Essential role for oncogenic Ras in tumour maintenance. Nature 400, 468–472.

Chiou, S.-H., Winters, I.P., Wang, J., Naranjo, S., Dudgeon, C., Tamburini, F.B., Brady, J.J., Yang, D., Grüner, B.M., Chuang, C.-H., et al. (2015). Pancreatic cancer modeling using retrograde viral vector delivery and in vivo CRISPR/Cas9-mediated somatic genome editing. Genes Dev.

Cho, S.W., Kim, S., Kim, J.M., and Kim, J.-S. (2013). Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease. Nat Biotechnol 31, 230–232.

Choi, P.S., and Meyerson, M. (2014). Targeted genomic rearrangements using CRISPR/Cas technology. Nature Communications 5, 3728.

Choo, Y., and Klug, A. (1994a). Selection of DNA binding sites for zinc fingers using rationally randomized DNA reveals coded interactions. Proc Natl Acad Sci USA 91, 11168–11172.

Choo, Y., and Klug, A. (1994b). Toward a code for the interactions of zinc fingers with DNA: Selection of randomized fingers displayed on phage. Proc Natl Acad Sci USA 91, 11163–11167.

Choulika, A., Perrin, A., Dujon, B., and Nicolas, J.F. (1995). Induction of homologous recombination in mammalian chromosomes by using the I-SceI system of Saccharomyces cerevisiae. Mol Cell Biol 15, 1968–1973.

Clarke, A.R., Maandag, E.R., van Roon, M., der Lugt, van, N.M., der Valk, van, M., Hooper, M.L., Berns, A., and Riele, te, H. (1992). Requirement for a functional Rb-1 gene in murine development. Nature 359, 328–330.

Cohen-Tannoudji, M., and Babinet, C. (1998). Beyond “knock-out” mice: new perspectives for the programmed modification of the mammalian genome. Mol. Hum. Reprod. 4, 929–938.

Colleaux, L., D'Auriol, L., Galibert, F., and Dujon, B. (1988). Recognition and cleavage site of the intron-encoded omega transposase. Proc Natl Acad Sci USA 85, 6022–6026.

Collins, E.C., Pannell, R., Simpson, E.M., Forster, A., and Rabbitts, T.H. (2000). Inter-chromosomal recombination of Mlland Af9genes mediated by cre- loxPin mouse development. EMBO Rep 1, 127–132.

Cong, L., Ran, F.A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P.D., Wu, X., Jiang, W., Marraffini, L.A., et al. (2013). Multiplex Genome Engineering Using CRISPR/Cas Systems. Science 339, 819–823.

Cory, S., and Adams, J.M. (1988). Transgenic Mice and Oncogenesis. Annu Rev Immunol 6, 25–48.

Cox, A.D., Fesik, S.W., Kimmelman, A.C., Luo, J., and Der, C.J. (2014). Drugging the undruggable RAS: Mission possible? Nat Rev Drug Discov 13, 828–851.

Cox, D.B.T., Platt, R.J., and Zhang, F. (2015). Therapeutic genome editing: prospects and challenges. Nature Medicine 21, 121–131.

Crotty, S., and Pipkin, M.E. (2015). In vivo RNAi screens: concepts and applications. Trends in Immunology 36, 315–322.

Crystal, A.S., Shaw, A.T., Sequist, L.V., Friboulet, L., Niederst, M.J., Lockerman, E.L., Frias, R.L., Gainor, J.F., Amzallag, A., Greninger, P., et al. (2014). Patient-derived models of acquired resistance can identify effective drug combinations for cancer. Science.

Cullinan, S.B., Gordan, J.D., Jin, J., Harper, J.W., and Diehl, J.A. (2004). The Keap1-BTB Protein Is an Adaptor That Bridges Nrf2 to a Cul3-Based E3 Ligase: Oxidative Stress Sensing by a Cul3-Keap1 Ligase. Mol Cell Biol 24, 8477–8486.

Dahlman, J.E., Abudayyeh, O.O., Joung, J., Gootenberg, J.S., Zhang, F., and Konermann, S. (2015). Orthogonal gene knockout and activation with a catalytically active Cas9 nuclease. Nat Biotechnol 33, 1159–1161.

Dang, L., White, D.W., Gross, S., Bennett, B.D., Bittinger, M.A., Driggers, E.M., Fantin, V.R., Jang, H.G., Jin, S., Keenan, M.C., et al. (2009). Cancer-associated IDH1 mutations produce 2-hydroxyglutarate. Nature 462, 739–744.

Dankort, D., Filenova, E., Collado, M., Serrano, M., Jones, K., and Mcmahon, M. (2007). A new mouse model to explore the initiation, progression, and therapy of BRAFV600E-induced lung tumors. Genes Dev 21, 379–384.

Davies, H., Bignell, G.R., Cox, C., Stephens, P., Edkins, S., Clegg, S., Teague, J., Woffendin, H., Garnett, M.J., Bottomley, W., et al. (2002). Mutations of the BRAF gene in human cancer. Nature 417, 949–954.

Davoli, T., Xu, A.W., Mengwasser, K.E., Sack, L.M., Yoon, J.C., Park, P.J., and Elledge, S.J. (2013). Cumulative Haploinsufficiency and Triplosensitivity Drive Aneuploidy Patterns and Shape the Cancer Genome. Cell 155, 948–962.

Dejeans, N., Tajeddine, N., Beck, R., Verrax, J., Taper, H., Gailly, P., and Calderon, P.B. (2010). Endoplasmic reticulum calcium release potentiates the ER stress and cell death caused by an oxidative stress in MCF-7 cells. Biochem Pharmacol 79, 1221–1230.

Deltcheva, E., Chylinski, K., Sharma, C.M., Gonzales, K., Chao, Y., Pirzada, Z.A., Eckert, M.R., Vogel, J., and Charpentier, E. (2011). CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. Nature 471, 602–607.

Denicola, G.M., Chen, P.-H., Mullarky, E., Sudderth, J.A., Hu, Z., Wu, D., Tang, H., Xie, Y., Asara, J.M., Huffman, K.E., et al. (2015). NRF2 regulates serine biosynthesis in non-small cell lung cancer. Nat Genet.

Denicola, G.M., Karreth, F.A., Humpton, T.J., Gopinathan, A., Wei, C., Frese, K., Mangal, D., Yu, K.H., Yeo, C.J., Calhoun, E.S., et al. (2011). Oncogene-induced Nrf2 transcription promotes ROS detoxification and tumorigenesis. Nature 475, 106–109.

Denli, A.M., Tops, B.B.J., Plasterk, R.H.A., Ketting, R.F., and Hannon, G.J. (2004). Processing of primary microRNAs by the Microprocessor complex. Nature 432, 231–235.

Devarakonda, S., Morgensztern, D., and Govindan, R. (2015). Genomic alterations in lung adenocarcinoma. Lancet Oncol. 16, e342–e351.

Deveau, H., Barrangou, R., Garneau, J.E., Labonté, J., Fremaux, C., Boyaval, P., Romero, D.A., Horvath, P., and Moineau, S. (2008). Phage Response to CRISPR-Encoded Resistance in Streptococcus thermophilus. J Bacteriol 190, 1390–1400.

Dhomen, N., Reis-Filho, J.S., da Rocha Dias, S., Hayward, R., Savage, K., Delmas, V., Larue, L., Pritchard, C., and Marais, R. (2009). Oncogenic Braf Induces Melanocyte Senescence and Melanoma in Mice. Cancer Cell 15, 294–303.

Di Agostino, S., Cortese, G., Monti, O., Dell'Orso, S., Sacchi, A., Eisenstein, M., Citro, G., Strano, S., and Blandino, G. (2008). The disruption of the protein complex mutantp53/p73 increases selectively the response of tumor cells to anticancer drugs. Cell Cycle 7, 3440–3447.

Di Como, C.J., Gaiddon, C., and Prives, C. (1999). p73 function is inhibited by tumor-derived p53 mutants in mammalian cells. Mol Cell Biol 19, 1438–1449.

Di Cristofano, A., Pesce, B., Cordon-Cardo, C., and Pandolfi, P.P. (1998). Pten is essential for embryonic development and tumour suppression. Nat Genet 19, 348–355.

Dickins, R.A., Hemann, M.T., Zilfou, J.T., Simpson, D.R., Ibarra, I., Hannon, G.J., and Lowe, S.W. (2005). Probing tumor phenotypes using stable and regulated synthetic microRNA precursors. Nat Genet 37, 1289–1295.

Dickins, R.A., McJunkin, K., Hernando, E., Premsrirut, P.K., Krizhanovsky, V., Burgess, D.J., Kim, S.Y., Cordon-Cardo, C., Zender, L., Hannon, G.J., et al. (2007). Tissue-specific and reversible RNA interference in transgenic mice. Nat Genet 39, 914–921.

Dimitrova, N., Zamudio, J.R., Jong, R.M., Soukup, D., Resnick, R., Sarma, K., Ward, A.J., Raj, A., Lee, J.T., Sharp, P.A., et al. (2014). LincRNA-p21 activates p21 in cis to promote Polycomb target gene expression and to enforce the G1/S checkpoint. Mol Cell 54, 777–790.

Ding, L., Getz, G., Wheeler, D.A., Mardis, E.R., McLellan, M.D., Cibulskis, K., Sougnez, C., Greulich, H., Muzny, D.M., Morgan, M.B., et al. (2008). Somatic mutations affect key pathways in lung adenocarcinoma. Nature 455, 1069–1075.

Dinkova-Kostova, A.T., and Abramov, A.Y. (2015). The emerging role of Nrf2 in mitochondrial function. Free Radic Biol Med 88, Part B IS -, 179–188.

Dittmer, D., Pati, S., Zambetti, G., Chu, S., Teresky, A.K., Moore, M., Finlay, C., and Levine, A.J. (1993). Gain of function mutations in p53. Nat Genet 4, 42–46.

Doetschman, T., Gregg, R.G., Maeda, N., Hooper, M.L., Melton, D.W., Thompson, S., and Smithies, O. (1987). Targetted correction of a mutant HPRT gene in mouse embryonic stem cells. Nature 330, 576–578.

Donehower, L.A., Harvey, M., Slagle, B.L., McArthur, M.J., Montgomery, C.A., Butel, J.S., and Bradley, A. (1992). Mice deficient for p53 are developmentally normal but susceptible to spontaneous tumours. Nature 356, 215–221.

Donehower, L.A., and Lozano, G. (2009). 20 years studying p53 functions in genetically engineered mice. Nat Rev Cancer 1–11.

Doudna, J.A., and Charpentier, E. (2014). Genome editing. The new frontier of genome engineering with CRISPR-Cas9. Science 346, 1258096.

Dow, L.E., and Lowe, S.W. (2012). Life in the Fast Lane: Mammalian Disease Models in the Genomics Era. Cell 148, 1099–1109.

Dow, L.E., Fisher, J., O'Rourke, K.P., Muley, A., Kastenhuber, E.R., Livshits, G., Tschaharganeh, D.F., Socci, N.D., and Lowe, S.W. (2015a). Inducible in vivo genome editing with CRISPR-Cas9. Nat Biotechnol.

Dow, L.E., Nasr, Z., Saborowski, M., Ebbesen, S.H., Manchado, E., Tasdemir, N., Lee, T., Pelletier, J., and Lowe, S.W. (2014). Conditional reverse tet-transactivator mouse strains for the efficient induction of TRE-regulated transgenes in mice. PLoS ONE 9, e95236.

Dow, L.E., O'Rourke, K.P., Simon, J., Tschaharganeh, D.F., van Es, J.H., Clevers, H., and Lowe, S.W. (2015b). Apc Restoration Promotes Cellular Differentiation and Reestablishes Crypt Homeostasis in Colorectal Cancer. Cell 161, 1539–1552.

Dow, L.E., Premsrirut, P.K., Zuber, J., Fellmann, C., McJunkin, K., Miething, C., Park, Y., Dickins, R.A., Hannon, G.J., and Lowe, S.W. (2012). A pipeline for the generation of shRNA transgenic mice. Nat Protoc 7, 374–393.

Downward, J. (2003). Targeting RAS signalling pathways in cancer therapy. Nat Rev Cancer 3, 11–22.

Doyon, J.B., Pattanayak, V., Meyer, C.B., and Liu, D.R. (2006). Directed Evolution and Substrate Specificity Profile of Homing Endonuclease I-SceI. J Am Chem Soc 128, 2477–2484.

Doyon, Y., McCammon, J.M., Miller, J.C., Faraji, F., Ngo, C., Katibah, G.E., Amora, R., Hocking, T.D., Zhang, L., Rebar, E.J., et al. (2008). Heritable targeted gene disruption in zebrafish using designed zinc-finger nucleases. Nat Biotechnol 26, 702–708.

Drost, J., van Jaarsveld, R.H., Ponsioen, B., Zimberlin, C., van Boxtel, R., Buijs, A., Sachs, N., Overmeer, R.M., Offerhaus, G.J., Begthel, H., et al. (2015). Sequential cancer mutations in cultured human intestinal stem cells. Nature 521, 43–47.

Duboule, D., Kmita, M., and Kondo, T. (2000). Targeted inversion of a polar silencer within the HoxD complex re-allocates domains of enhancer sharing - Nature Genetics. Nat Genet 26, 451–454.

Dupage, M., Dooley, A.L., and Jacks, T. (2009). Conditional mouse lung cancer models using adenoviral or lentiviral delivery of Cre recombinase. Nat Protoc 4, 1064–1072.

Durai, S., Bosley, A., Abulencia, A.B., Chandrasegaran, S., and Ostermeier, M. (2006). A bacterial one-hybrid selection system for interrogating zinc finger-DNA interactions. Combinatorial Chemistry and High Throughput Screening 9, 301–311.

Durai, S., Mani, M., Kandavelou, K., Wu, J., Porteus, M.H., and Chandrasegaran, S. (2005). Zinc finger nucleases: Custom-designed molecular scissors for genome engineering of plant and mammalian cells. Nucleic Acids Res 33, 5978–5990.

Eichenmuller, M., Trippel, F., Kreuder, M., Beck, A., Schwarzmayr, T., Haberle, B., Cairo, S., Leuschner, I., Schweinitz, von, D., Strom, T.M., et al. (2014). The genomic landscape of hepatoblastoma and their progenies with HCC-like features. J Hepatol 61, 1312–1320.

Eilers, M., Picard, D., Yamamoto, K.R., and Bishop, J.M. (1989). Chimaeras of Myc oncoprotein and steroid receptors cause hormone-dependent transformation of cells. Nature 340, 66–68.

Eischen, C.M., Weber, J.D., Roussel, M.F., Sherr, C.J., and Cleveland, J.L. (1999). Disruption of the ARF–Mdm2–p53 tumor suppressor pathway in Myc-induced lymphomagenesis. Genes Dev 13, 2658–2669.

Eletto, D., Chevet, E., Argon, Y., and Appenzeller-Herzog, C. (2014). Redox controls UPR to control redox. Journal of Cell Science 127, 3649–3658.

Elson, A., Wang, Y., Daugherty, C.J., Morton, C.C., Zhou, F., Campos-Torres, J., and Leder, P. (1996). Pleiotropic defects in ataxia-telangiectasia protein-deficient mice. Proc Natl Acad Sci USA 93, 13084–13089.

Engel, J.D. (1994). Meticulous AP-1 factors. Nature 367, 516–517.

Engelman, J.A., Chen, L., Tan, X., Crosby, K., Guimaraes, A.R., Upadhyay, R., Maira, M., McNamara, K., Perera, S.A., Song, Y., et al. (2008). Effective use of PI3K and MEK inhibitors to treat mutant Kras G12D and PIK3CA H1047R murine lung cancers. Nat Med 14, 1351–1356.

Essletzbichler, P., Konopka, T., Santoro, F., Chen, D., Gapp, B.V., Kralovics, R., Brummelkamp, T.R., Nijman, S.M.B., and Bürckstümmer, T. (2014). Megabase-scale deletion using CRISPR/Cas9 to generate a fully haploid human cell line. Genome Research 24, 2059–2065.

Esvelt, K.M., Mali, P., Braff, J.L., Moosburner, M., Yaung, S.J., and Church, G.M. (2013). Orthogonal Cas9 proteins for RNA-guided gene regulation and editing. Nat Methods 10, 1116–1121.

Evans, M.J., and Kaufman, M.H. (1981). Establishment in culture of pluripotential cells from mouse embryos. Nature 292, 154–156.

Ewald, D., Li, M., Efrat, S., Auer, G., Wall, R.J., Furth, P.A., and Hennighausen, L. (1996). Time-sensitive reversal of hyperplasia in transgenic mice expressing SV40 T antigen. Science 273, 1384–1386.

Feldser, D.M., Kostova, K.K., Winslow, M.M., Taylor, S.E., Cashman, C., Whittaker, C.A., Sanchez-Rivera, F.J., Resnick, R., Bronson, R., Hemann, M.T., et al. (2010). Stage-specific sensitivity to p53 restoration during lung cancer progression. Nature 468, 572–575.

Fellmann, C., and Lowe, S.W. (2014). Stable RNA interference rules for silencing. Nat Cell Biol 16, 10–18.

Felsher, D.W., and Bishop, J.M. (1999). Reversible tumorigenesis by MYC in hematopoietic lineages. Mol Cell 4, 199–207.

Filipowicz, W., Bhattacharyya, S.N., and Sonenberg, N. (2008). Mechanisms of post-transcriptional regulation by microRNAs: are the answers in sight? Nat Rev Genet 9, 102–114.

Finco, T.S., Westwick, J.K., Norris, J.L., Beg, A.A., Der, C.J., and Baldwin, A.S.J. (1997). Oncogenic Ha-Ras-induced signaling activates NF-kappaB transcriptional activity, which is required for cellular transformation. J Biol Chem 272, 24113–24116.

Findlay, G.M., Boyle, E.A., Hause, R.J., Klein, J.C., and Shendure, J. (2014). Saturation editing of genomic regions by multiplex homology-directed repair. Nature 513, 120–123.

Finkel, T. (2011). Signal transduction by reactive oxygen species. The Journal of Cell Biology 194, 7–15.

Fire, A., Xu, S., Montgomery, M.K., Kostas, S.A., Driver, S.E., and Mello, C.C. (1998). Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. Nature 391, 806–811.

Flaherty, K.T. (2012). Targeting Metastatic Melanoma. Annu. Rev. Med. 63, 171–183.

Flaherty, K.T., Puzanov, I., Kim, K.B., Ribas, A., Mcarthur, G.A., Sosman, J.A., O'Dwyer, P.J., Lee, R.J., Grippo, J.F., Nolop, K., et al. (2010). Inhibition of Mutated, Activated BRAF in Metastatic Melanoma. N Engl J Med 363, 809–819.

Fletcher, B.S. (2010). Delivery of Small Interfering RNA (siRNA) Using the Sleeping Beauty Transposon. Cold Spring Harb Protoc 2010, pdb.prot5521.

Fodde, R., Edelmann, W., Yang, K., van Leeuwen, C., Carlson, C., Renault, B., Breukel, C., Alt, E., Lipkin, M., and Khan, P.M. (1994). A targeted chain-termination mutation in the mouse Apc gene results in multiple intestinal tumors. Proceedings of the National Academy of Sciences 91, 8969–8973.

Forster, A., Pannell, R., Drynan, L.F., Codrington, R., Daser, A., Metzler, M., Lobato, M.N., and Rabbitts, T.H. (2005). The invertor knock-in conditional chromosomal translocation mimic. Nat Meth 2, 27–30.

Forster, A., Pannell, R., Drynan, L.F., McCormack, M., Collins, E.C., Daser, A., and Rabbitts, T.H. (2003). Engineering de novo reciprocal chromosomal translocations associated with Mll to replicate primary events of human cancer. Cancer Cell 3, 449–458.

Frese, K.K., and Tuveson, D.A. (2007). Maximizing mouse cancer models. Nat Rev Cancer 7, 645–658.

Friedland, A.E., Tzur, Y.B., Esvelt, K.M., Colaiacovo, M.P., Church, G.M., and Calarco, J.A. (2013). Heritable genome editing in C. elegans via a CRISPR-Cas9 system. Nat Meth 10, 741–743.

Friedlander, S.Y.G., Chu, G.C., Snyder, E.L., Girnius, N., Dibelius, G., Crowley, D., Vasile, E., DePinho, R.A., and Jacks, T. (2009). Context-Dependent Transformation of Adult Pancreatic Cells by Oncogenic K-Ras. Cancer Cell 16, 379–389.

Friend, S.H., Bernards, R., Rogelj, S., Weinberg, R.A., Rapaport, J.M., Albert, D.M., and Dryja, T.P. (1986). A human DNA segment with properties of the gene that predisposes to retinoblastoma and osteosarcoma. Nature 323, 643–646.

Fukuoka, M., Yano, S., Giaccone, G., Tamura, T., Nakagawa, K., Douillard, J.-Y., Nishiwaki, Y., Vansteenkiste, J., Kudoh, S., Rischin, D., et al. (2003). Multi-institutional randomized phase II trial of gefitinib for previously treated patients with advanced non-small-cell lung cancer (The IDEAL 1 Trial) [corrected]. J Clin Oncol 21, 2237–2246.

Gaiddon, C., Lokshin, M., Ahn, J., Zhang, T., and Prives, C. (2001). A subset of tumor-derived mutant forms of p53 down-regulate p63 and p73 through a direct interaction with the p53 core domain. Mol Cell Biol 21, 1874–1887.

Gao, P., Zhang, H., Dinavahi, R., Li, F., Xiang, Y., Raman, V., Bhujwalla, Z.M., Felsher, D.W., Cheng, L., Pevsner, J., et al. (2007). HIF-dependent antitumorigenic effect of antioxidants in vivo. Cancer Cell 12, 230–238.

Gargiulo, G., Cesaroni, M., Serresi, M., de Vries, N., Hulsman, D., Bruggeman, S.W., Lancini, C., and van Lohuizen, M. (2013). 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.

Garneau, J.E., Dupuis, M.-È., Villion, M., Romero, D.A., Barrangou, R., Boyaval, P., Fremaux, C., Horvath, P., Magadán, A.H., and Moineau, S. (2010). The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. Nature 468, 67–71.

Garraway, L.A., and Lander, E.S. (2013). Lessons from the cancer genome. Cell 153, 17–37.

Garraway, L.A., Verweij, J., and Ballman, K.V. (2013). Precision oncology: an overview. J Clin Oncol 31, 1803–1805.

Garrick, D., Sutherland, H., Robertson, G., and Whitelaw, E. (1996). Variegated expression of a globin transgene correlates with chromatin accessibility but not methylation status. Nucleic Acids Res 24, 4902–4909.

Gasiunas, G., Barrangou, R., Horvath, P., and Siksnys, V. (2012). Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. Proc Natl Acad Sci USA 109, E2579–E2586.

Gatz, C., Frohberg, C., and Wendenburg, R. (1992). Stringent repression and homogeneous de-repression by tetracycline of a modified CaMV 35S promoter in intact transgenic tobacco plants. Plant J 2, 397–404.

George, J., Lim, J.S., Jang, S.J., Cun, Y., Ozretić, L., Kong, G., Leenders, F., Lu, X., Fernández-Cuesta, L., Bosco, G., et al. (2015). Comprehensive genomic profiles of small cell lung cancer. Nature 524, 47–53.

Ghezraoui, H., Piganeau, M., Renouf, B., Renaud, J.-B., Sallmyr, A., Ruis, B., Oh, S., Tomkinson, A.E., Hendrickson, E.A., Giovannangeli, C., et al. (2014). Chromosomal translocations in human cells are generated by canonical nonhomologous end-joining. Mol Cell 55, 829–842.

Gilbert, L.A., Horlbeck, M.A., Adamson, B., Villalta, J.E., Chen, Y., Whitehead, E.H., Guimaraes, C., Panning, B., Ploegh, H.L., Bassik, M.C., et al. (2014). Genome-Scale CRISPR-Mediated Controlof Gene Repression and Activation. Cell 1–28.

Gilbert, L.A., Larson, M.H., Morsut, L., Liu, Z., Brar, G.A., Torres, S.E., Stern-Ginossar, N., Brandman, O., Whitehead, E.H., Doudna, J.A., et al. (2013). CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. Cell 154, 442–451.

Ginsberg, A.M., King, B.O., and Roeder, R.G. (1984). Xenopus 5S gene transcription factor, TFIIIA: characterization of a cDNA clone and measurement of RNA levels throughout development. Cell 39, 479–489.

Gnarra, J.R., Ward, J.M., Porter, F.D., Wagner, J.R., Devor, D.E., Grinberg, A., Emmert-Buck, M.R., Westphal, H., Klausner, R.D., and Linehan, W.M. (1997). Defective placental vasculogenesis causes embryonic lethality in VHL-deficient mice. Proc Natl Acad Sci USA 94, 9102–9107.

Gordon, D.J., Resio, B., and Pellman, D. (2012). Causes and consequences of aneuploidy in cancer. Nat Rev Genet.

Gorrini, C., Harris, I.S., and Mak, T.W. (2013). Modulation of oxidative stress as an anticancer strategy. Nature Reviews Drug Discovery 12, 931–947.

Gossen, M., and Bujard, H. (1992). Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. Proc Natl Acad Sci USA 89, 5547–5551.

Gossen, M., Freundlieb, S., Bender, G., Muller, G., Hillen, W., and Bujard, H. (1995). Transcriptional activation by tetracyclines in mammalian cells. Science 268, 1766–1769.

Govindan, R., Govindan, R., Ding, L., Ding, L., Griffith, M., Griffith, M., Subramanian, J., Subramanian, J., Dees, N.D., Dees, N.D., et al. (2012). Genomic landscape of non-small cell lung cancer in smokers and never-smokers. Cell 150, 1121–1134.

Gowen, L.C., Johnson, B.L., Latour, A.M., Sulik, K.K., and Koller, B.H. (1996). Brca1 deficiency results in early embryonic lethality characterized by neuroepithelial abnormalities. Nat Genet 12, 191–194.

Gozzelino, R., Jeney, V., and Soares, M.P. (2010). Mechanisms of Cell Protection by Heme Oxygenase-1. Annu. Rev. Pharmacol. Toxicol. 50, 323–354.

Graeber, T.G., Peterson, J.F., Tsai, M., Monica, K., Fornace, A.J.J., and Giaccia, A.J. (1994). Hypoxia induces accumulation of p53 protein, but activation of a G1-phase checkpoint by low-oxygen conditions is independent of p53 status. Mol Cell Biol 14, 6264–6277.

Gratz, S.J., Cummings, A.M., Nguyen, J.N., Hamm, D.C., Donohue, L.K., Harrison, M.M., Wildonger, J., and O'Connor-Giles, K.M. (2013). Genome Engineering of Drosophila with the CRISPR RNA-Guided Cas9 Nuclease. Genetics 194, 1029–1035.

Gregory, R.I., Yan, K.-P., Amuthan, G., Chendrimada, T., Doratotaj, B., Cooch, N., and Shiekhattar, R. (2004). The Microprocessor complex mediates the genesis of microRNAs. Nature 432, 235–240.

Greisman, H.A., and Pabo, C.O. (1997). A General Strategy for Selecting High-Affinity Zinc Finger Proteins for Diverse DNA Target Sites. Science 275, 657–661.

Griffith, O.W., and Mulcahy, R.T. (2006). The Enzymes of Glutathione Synthesis: γ-Glutamylcysteine Synthetase. In Advances in Enzymology and Related Areas of Molecular Biology, (John Wiley & Sons, Inc.), pp. 209–267.

Gu, H., Marth, J.D., Orban, P.C., Mossmann, H., and Rajewsky, K. (1994). Deletion of a DNA polymerase beta gene segment in T cells using cell type-specific gene targeting. Science 265, 103–106.

Guarente, L. (1993). Synthetic enhancement in gene interaction: a genetic tool come of age. Trends Genet. 9, 362–366.

Guerra, C., Mijimolle, N., Dhawahir, A., Dubus, P., Barradas, M., Serrano, M., Campuzano, V., and Barbacid, M. (2003). Tumor induction by an endogenous K-ras oncogene is highly dependent on cellular context. Cancer Cell 4, 111–120.

Guo, H., Ingolia, N.T., Weissman, J.S., and Bartel, D.P. (2010). Mammalian microRNAs predominantly act to decrease target mRNA levels. Nature 466, 835–840.

Gutteridge, J.M. (1986). Iron promoters of the Fenton reaction and lipid peroxidation can be released from haemoglobin by peroxides. FEBS Lett 201, 291–295.

Guzy, R.D., Hoyos, B., Robin, E., Chen, H., Liu, L., Mansfield, K.D., Simon, M.C., Hammerling, U., and Schumacker, P.T. (2005). Mitochondrial complex III is required for hypoxia-induced ROS production and cellular oxygen sensing. Cell Metabolism 1, 401–408.

Haber, J.E. (1998). MATING-TYPE GENE SWITCHING IN SACCHAROMYCES CEREVISIAE. Annu Rev Genet 32, 561–599.

Hakem, R., la Pompa, de, J.L., Sirard, C., Mo, R., Woo, M., Hakem, A., Wakeham, A., Potter, J., Reitmair, A., Billia, F., et al. (1996). The tumor suppressor gene Brca1 is required for embryonic cellular proliferation in the mouse. Cell 85, 1009–1023.

Hammerman, P.S., Lawrence, M.S., Voet, D., Jing, R., Cibulskis, K., Sivachenko, A., Stojanov, P., McKenna, A., Lander, E.S., Gabriel, S., et al. (2012). Comprehensive genomic characterization of squamous cell lung cancers. Nature 489, 519–525.

Hammond, S.M., Caudy, A.A., and Hannon, G.J. (2001). Post-transcriptional gene silencing by double-stranded RNA. Nature Publishing Group 2, 110–119.

Han, J., Zhang, J., Chen, L., Shen, B., Zhou, J., Hu, B., Du, Y., Tate, P.H., Huang, X., and Zhang, W. (2014). Efficient in vivo deletion of a large imprinted lncRNA by CRISPR/Cas9. RNA Biol 11, 829–835.

Hanahan, D. (1988). Dissecting Multistep Tumorigenesis in Transgenic Mice. Annu Rev Genet 22, 479–519.

Hanahan, D. (1989). Transgenic mice as probes into complex systems. Science 246, 1265–1275.

Hanahan, D., and Weinberg, R.A. (2000). The hallmarks of cancer. Cell 100, 57–70.

Hanahan, D. (1985). Heritable formation of pancreatic β-cell tumours in transgenic mice expressing recombinant insulin/simian virus 40 oncogenes. Nature 315, 115–122.

Hanas, J.S., Hazuda, D.J., Bogenhagen, D.F., Wu, F.Y., and Wu, C.W. (1983). Xenopus transcription factor A requires zinc for binding to the 5 S RNA gene. Journal of Biological Chemistry 258, 14120–14125.

Hannon, G.J. (2002). RNA interference. Nature 418, 244–251.

Hannon, G.J., and Rossi, J.J. (2004). Unlocking the potential of the human genome with RNA interference. Nat Cell Biol 431, 371–378.

Harris, I.S. (2013). Investigation into the Role of Antioxidants in Tumorigenesis.

Harris, I.S., Treloar, A.E., Inoue, S., Sasaki, M., Gorrini, C., Lee, K.C., Yung, K.Y., Brenner, D., Knobbe-Thomsen, C.B., Cox, M.A., et al. (2015). Glutathione and Thioredoxin Antioxidant Pathways Synergize to Drive Cancer Initiation and Progression. Cancer Cell.

Harvey, C.J., Thimmulappa, R.K., Singh, A., Blake, D.J., Ling, G., Wakabayashi, N., Fujii, J., Myers, A., and Biswal, S. (2009). Nrf2-regulated glutathione recycling independent of biosynthesis is critical for cell survival during oxidative stress. Free Radic Biol Med 46, 443–453.

Hasuwa, H., Kaseda, K., Einarsdottir, T., and Okabe, M. (2002). Small interfering RNA and gene silencing in transgenic mice and rats. FEBS Lett 532, 227–230.

Haupt, Y., Maya, R., Kazaz, A., and Oren, M. (1997). Mdm2 promotes the rapid degradation of p53. Nature 387, 296–299.

Hauschild, J., Petersen, B., Santiago, Y., Queisser, A.-L., Carnwath, J.W., Lucas-Hahn, A., Zhang, L., Meng, X., Gregory, P.D., Schwinzer, R., et al. (2011). Efficient generation of a biallelic knockout in pigs using zinc-finger nucleases. Proceedings of the National Academy of Sciences 108, 12013–12017.

Hayes, J.D., and McMahon, M. (2009). NRF2 and KEAP1 mutations: permanent activation of an adaptive response in cancer. Trends Biochem Sci 34, 176–188.

Heckl, D., Kowalczyk, M.S., Yudovich, D., Belizaire, R., Puram, R.V., McConkey, M.E., Thielke, A., Aster, J.C., Regev, A., and Ebert, B.L. (2014). Generation of mouse models of myeloid malignancy with combinatorial genetic lesions using CRISPR-Cas9 genome editing. Nat Biotechnol 32, 941–946.

Heggestad, A.D., Notterpek, L., and Fletcher, B.S. (2004). Transposon-based RNAi delivery system for generating knockdown cell lines. Biochem Biophys Res Commun 316, 643–650.

Hemann, M.T., Fridman, J.S., Zilfou, J.T., Hernando, E., Paddison, P.J., Cordon-Cardo, C., Hannon, G.J., and Lowe, S.W. (2003). An epi-allelic series of p53 hypomorphs created by stable RNAi produces distinct tumor phenotypes in vivo. Nat Genet 33, 396–400.

Hemphill, J., Borchardt, E.K., Brown, K., Asokan, A., and Deiters, A. (2015). Optical Control of CRISPR/Cas9 Gene Editing. J Am Chem Soc 137, 5642–5645.

Henikoff, S. (1998). Conspiracy of silence among repeated transgenes. Bioessays 20, 532–535.

Henneman, L., van Miltenburg, M.H., Michalak, E.M., Braumuller, T.M., Jaspers, J.E., Drenth, A.P., de Korte-Grimmerink, R., Gogola, E., Szuhai, K., Schlicker, A., et al. (2015). Selective resistance to the PARP inhibitor olaparib in a mouse model for BRCA1-deficient metaplastic breast cancer. Proceedings of the National Academy of Sciences 112, 8409–8414.

Herbst, R.S., Heymach, J.V., and Lippman, S.M. (2008). Lung Cancer. Nejm 359, 1367–1380.

Heyer, J., Kwong, L.N., Lowe, S.W., and Chin, L. (2010). Non-germline genetically engineered mouse models for translational cancer research. Nat Rev Cancer 10, 470–480.

Hilton, I.B., D'Ippolito, A.M., Vockley, C.M., Thakore, P.I., Crawford, G.E., Reddy, T.E., and Gersbach, C.A. (2015). Epigenome editing by a CRISPR-Cas9-based acetyltransferase activates genes from promoters and enhancers. Nat Biotechnol 33, 510–517.

Hinnen, A., Hicks, J.B., and Fink, G.R. (1978). Transformation of yeast. Proceedings of the National Academy of Sciences 75, 1929–1933.

Hirata, M., Sasaki, M., Cairns, R.A., Inoue, S., Puviindran, V., Li, W.Y., Snow, B.E., Jones, L.D., Wei, Q., Sato, S., et al. (2015). Mutant IDH is sufficient to initiate enchondromatosis in mice. Proc Natl Acad Sci USA 112, 2829–2834.

Hirrlinger, J., Requardt, R.P., Winkler, U., Wilhelm, F., Schulze, C., and Hirrlinger, P.G. (2009a). Split-CreERT2: temporal control of DNA recombination mediated by split-Cre protein fragment complementation. PLoS ONE 4, e8354.

Hirrlinger, J., Scheller, A., Hirrlinger, P.G., Kellert, B., Tang, W., Wehr, M.C., Goebbels, S., Reichenbach, A., Sprengel, R., Rossner, M.J., et al. (2009b). Split-cre complementation indicates coincident activity of different genes in vivo. PLoS ONE 4, e4286.

Hirsch, F.R., Varella-Garcia, M., Bunn, P.A.J., Di Maria, M.V., Veve, R., Bremmes, R.M., Baron, A.E., Zeng, C., and Franklin, W.A. (2003). Epidermal growth factor receptor in non-small-cell lung carcinomas: correlation between gene copy number and protein expression and impact on prognosis. J Clin Oncol 21, 3798–3807.

Ho, T.-T., Zhou, N., Huang, J., Koirala, P., Xu, M., Fung, R., Wu, F., and Mo, Y.-Y. (2014). Targeting non-coding RNAs with the CRISPR/Cas9 system in human cell lines. Nucleic Acids Res.

Hockemeyer, D., Soldner, F., Beard, C., Gao, Q., Mitalipova, M., DeKelver, R.C., Katibah, G.E., Amora, R., Boydston, E.A., Zeitler, B., et al. (2009). Efficient targeting of expressed and silent genes in human ESCs and iPSCs using zinc-finger nucleases. Nat Biotechnol 27, 851–857.

Hockemeyer, D., Wang, H., Kiani, S., Lai, C.S., Gao, Q., Cassady, J.P., Cost, G.J., Zhang, L., Santiago, Y., Miller, J.C., et al. (2011). Genetic engineering of human pluripotent cells using TALE nucleases. Nat Biotechnol 29, 731–734.

Holderfield, M., Deuker, M.M., McCormick, F., and Mcmahon, M. (2014). Targeting RAF kinases for cancer therapy: BRAF-mutated melanoma and beyond. Nat Rev Cancer 14, 455–467.

Holmstrom, K.M., and Finkel, T. (2014). Cellular mechanisms and physiological consequences of redox-dependent signalling. Nature Publishing Group 15, 411–421.

Honda, R., Tanaka, H., and Yasuda, H. (1997). Oncoprotein MDM2 is a ubiquitin ligase E3 for tumor suppressor p53. FEBS Lett 420, 25–27.

Hou, Z., Zhang, Y., Propson, N.E., Howden, S.E., Chu, L.-F., Sontheimer, E.J., and Thomson, J.A. (2013). Efficient genome engineering in human pluripotent stem cells using Cas9 from Neisseria meningitidis. Proceedings of the National Academy of Sciences 110, 15644–15649.

Houbracken, I., Baeyens, L., Ravassard, P., Heimberg, H., and Bouwens, L. (2012). Gene delivery to pancreatic exocrine cells in vivo and in vitro. BMC Biotechnol 12, 74.

Hsu, I.C., Metcalf, R.A., Sun, T., Welsh, J.A., Wang, N.J., and Harris, C.C. (1991). Mutational hotspot in the p53 gene in human hepatocellular carcinomas. Nature 350, 427–428.

Hsu, P.D., Lander, E.S., and Zhang, F. (2014). Development and Applications of CRISPR-Cas9 for Genome Engineering. Cell 157, 1262–1278.

Hsu, P.D., Scott, D.A., Weinstein, J.A., Ran, F.A., Konermann, S., Agarwala, V., Li, Y., Fine, E.J., Wu, X., Shalem, O., et al. (2013). DNA targeting specificity of RNA-guided Cas9 nucleases. Nat Biotechnol 31, 827–832.

Hu, X., Chang, N., Wang, X., Zhou, F., Zhou, X., Zhu, X., and Xiong, J.-W. (2013). Heritable gene-targeting with gRNA/Cas9 in rats. Cell Res 23, 1322–1325.

Huang, P., Xiao, A., Zhou, M., Zhu, Z., Lin, S., and Zhang, B. (2011). Heritable gene targeting in zebrafish using customized TALENs. Nat Biotechnol 29, 699–700.

Huijbers, I.J., Bin Ali, R., Pritchard, C., Cozijnsen, M., Kwon, M.-C., Proost, N., Song, J.-Y., de Vries, H., Badhai, J., Sutherland, K., et al. (2014). Rapid target gene validation in complex cancer mouse models using re-derived embryonic stem cells. EMBO Mol Med 6, 212–225.

Humphreys, D.T., Westman, B.J., Martin, D.I.K., and Preiss, T. (2005). MicroRNAs control translation initiation by inhibiting eukaryotic initiation factor 4E/cap and poly(A) tail function. Proceedings of the National Academy of Sciences 102, 16961–16966.

Huntzinger, E., and Izaurralde, E. (2011). Gene silencing by microRNAs: contributions of translational repression and mRNA decay. Nature Publishing Group 12, 99–110.

Hurt, J.A., Thibodeau, S.A., Hirsh, A.S., Pabo, C.O., and Joung, J.K. (2003). Highly specific zinc finger proteins obtained by directed domain shuffling and cell-based selection. Proc Natl Acad Sci USA 100, 12271–12276.

Hutvagner, G., McLachlan, J., Pasquinelli, A.E., Balint, E., Tuschl, T., and Zamore, P.D. (2001). A cellular function for the RNA-interference enzyme Dicer in the maturation of the let-7 small temporal RNA. Science 293, 834–838.

Hutvagner, G., and Zamore, P.D. (2002). A microRNA in a multiple-turnover RNAi enzyme complex. Science 297, 2056–2060.

Hwang, W.Y., Fu, Y., Reyon, D., Maeder, M.L., Tsai, S.Q., Sander, J.D., Peterson, R.T., Yeh, J.-R.J., and Joung, J.K. (2013). Efficient genome editing in zebrafish using a CRISPR-Cas system. Nat Biotechnol 31, 227–229.

Igarashi, K., Kataoka, K., Itoh, K., Hayashi, N., Nishizawa, M., and Yamamoto, M. (1994). Regulation of transcription by dimerization of erythroid factor NF-E2 p45 with small Maf proteins. Nature 367, 568–572.

Ikeda, H., and Tomizawa, J. (1968). Prophage P1, and extrachromosomal replication unit. Cold Spring Harb Symp Quant Biol 33, 791–798.

Imielinski, M., Berger, A.H., Hammerman, P.S., Hernandez, B., Pugh, T.J., Hodis, E., Cho, J., Suh, J., Capelletti, M., Sivachenko, A., et al. (2012). Mapping the hallmarks of lung adenocarcinoma with massively parallel sequencing. Cell 150, 1107–1120.

Inoue, T., and Suzuki-Karasaki, Y. (2013). Mitochondrial superoxide mediates mitochondrial and endoplasmic reticulum dysfunctions in TRAIL-induced apoptosis in Jurkat cells. Free Radic Biol Med 61, 273–284.

International Human Genome Sequencing Consortium (2004). Finishing the euchromatic sequence of the human genome. Nature 431, 931–945.

Irani, K., Xia, Y., Zweier, J.L., Sollott, S.J., Der, C.J., Fearon, E.R., Sundaresan, M., Finkel, T., and Goldschmidt-Clermont, P.J. (1997). Mitogenic signaling mediated by oxidants in Ras-transformed fibroblasts. Science 275, 1649–1652.

Isaacs, W.B., Carter, B.S., and Ewing, C.M. (1991). Wild-type p53 suppresses growth of human prostate cancer cells containing mutant p53 alleles. Cancer Res 51, 4716–4720.

Isalan, M., Choo, Y., and Klug, A. (1997). Synergy between adjacent zinc fingers in sequence-specific DNA recognition. Proc Natl Acad Sci USA 94, 5617–5621.

Isalan, M., Klug, A., and Choo, Y. (2001). A rapid, generally applicable method to engineer zinc fingers illustrated by targeting the HIV-1 promoter. Nat Biotechnol 19, 656–660.

Ishino, Y., Shinagawa, H., Makino, K., Amemura, M., and Nakata, A. (1987). Nucleotide sequence of the iap gene, responsible for alkaline phosphatase isozyme conversion in Escherichia coli, and identification of the gene product. J Bacteriol 169, 5429–5433.

Itoh, K., Chiba, T., Takahashi, S., Ishii, T., Igarashi, K., Katoh, Y., Oyake, T., Hayashi, N., Satoh, K., Hatayama, I., et al. (1997). An Nrf2/small Maf heterodimer mediates the induction of phase II detoxifying enzyme genes through antioxidant response elements. Biochem Biophys Res Commun 236, 313–322.

Itoh, K., Igarashi, K., Hayashi, N., Nishizawa, M., and Yamamoto, M. (1995). Cloning and characterization of a novel erythroid cell-derived CNC family transcription factor heterodimerizing with the small Maf family proteins. Mol Cell Biol 15, 4184–4193.

Itoh, K., Wakabayashi, N., Katoh, Y., Ishii, T., Igarashi, K., Engel, J.D., and Yamamoto, M. (1999). Keap1 represses nuclear activation of antioxidant responsive elements by Nrf2 through binding to the amino-terminal Neh2 domain. Genes Dev 13, 76–86.

Itoh, K., Mimura, J., and Yamamoto, M. (2010). Discovery of the Negative Regulator of Nrf2, Keap1: A Historical Overview. Antioxid Redox Signal 13, 1665–1678.

Jacks, T. (1996). Tumor suppressor gene mutations in mice. Annu Rev Genet 30, 603–636.

Jacks, T., Fazeli, A., Schmitt, E.M., Bronson, R.T., Goodell, M.A., and Weinberg, R.A. (1992). Effects of an Rb mutation in the mouse. Nature 359, 295–300.

Jacks, T., Remington, L., Williams, B.O., Schmitt, E.M., Halachmi, S., Bronson, R.T., and Weinberg, R.A. (1994). Tumor spectrum analysis in p53-mutant mice. Curr Biol 4, 1–7.

Jackson, E.L., Willis, N., Mercer, K., Bronson, R.T., Crowley, D., Montoya, R., Jacks, T., and Tuveson, D.A. (2001). Analysis of lung tumor initiation and progression using conditional expression of oncogenic K-ras. Genes Dev 15, 3243–3248.

Jackson, E., Olive, K., Tuveson, D., Bronson, R., Crowley, D., Brown, M., and Jacks, T. (2005). The Differential Effects of Mutant p53 Alleles on Advanced Murine Lung Cancer. Cancer Res 65, 10280–10288.

Jackson, P., Baltimore, D., and Picard, D. (1993). Hormone-conditional transformation by fusion proteins of c-Abl and its transforming variants. Embo J 12, 2809–2819.

Jacobs, J.J.L., Scheijen, B., Voncken, J.-W., Kieboom, K., Berns, A., and van Lohuizen, M. (1999). Bmi-1 collaborates with c-Myc in tumorigenesis by inhibiting c-Myc-induced apoptosis via INK4a/ARF. Genes Dev 13, 2678–2690.

Jacquier, A., and Dujon, B. (1985). An intron-encoded protein is active in a gene conversion process that spreads an intron into a mitochondrial gene. Cell 41, 383–394.

Jaenisch, R. (1976). Germ line integration and Mendelian transmission of the exogenous Moloney leukemia virus. Proceedings of the National Academy of Sciences 73, 1260–1264.

Jaenisch, R., and Mintz, B. (1974). Simian Virus 40 DNA Sequences in DNA of Healthy Adult Mice Derived from Preimplantation Blastocysts Injected with Viral DNA. Proceedings of the National Academy of Sciences 71, 1250–1254.

Jain, M., Arvanitis, C., Chu, K., Dewey, W., Leonhardt, E., Trinh, M., Sundberg, C.D., Bishop, J.M., and Felsher, D.W. (2002). Sustained loss of a neoplastic phenotype by brief inactivation of MYC. Science 297, 102–104.

James, R.M., Arends, M.J., Plowman, S.J., Brooks, D.G., Miles, C.G., West, J.D., and Patek, C.E. (2003). K-ras Proto-Oncogene Exhibits Tumor Suppressor Activity As Its Absence Promotes Tumorigenesis in Murine Teratomas11The Cancer Research Campaign (now Cancer Research UK). Molecular Cancer Research 1, 820–825.

Jansen, R., Embden, J.D.A.V., Gaastra, W., and Schouls, L.M. (2002). Identification of genes that are associated with DNA repeats in prokaryotes. Mol. Microbiol. 43, 1565–1575.

Jaramillo, M.C., and Zhang, D.D. (2013). The emerging role of the Nrf2–Keap1 signaling pathway in cancer. Genes Dev 27, 2179–2191.

Järås, M., Miller, P.G., Chu, L.P., Puram, R.V., Fink, E.C., Schneider, R.K., Al-Shahrour, F., Peña, P., Breyfogle, L.J., Hartwell, K.A., et al. (2014). Csnk1a1 inhibition has p53-dependent therapeutic efficacy in acute myeloid leukemia. J Exp Med 211, 605–612.

Jemal, A., Siegel, R., Ward, E., Hao, Y., Xu, J., Murray, T., and Thun, M.J. (2008). Cancer statistics, 2008. CA Cancer J Clin 58, 71–96.

Jiang, T., Chen, N., Zhao, F., Wang, X.-J., Kong, B., Zheng, W., and Zhang, D.D. (2010). High Levels of Nrf2 Determine Chemoresistance in Type II Endometrial Cancer. Cancer Research 70, 5486–5496.

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A., and Charpentier, E. (2012). A Programmable Dual-RNA-Guided DNA Endonuclease in Adaptive Bacterial Immunity. Science 337, 816–821.

Jinek, M., East, A., Cheng, A., Lin, S., Ma, E., and Doudna, J. (2013). RNA-programmed genome editing in human cells. eLife Sciences 2, e00471.

Johnson, L., Greenbaum, D., Cichowski, K., Mercer, K., Murphy, E., Schmitt, E., Bronson, R.T., Umanoff, H., Edelmann, W., Kucherlapati, R., et al. (1997). K-ras is an essential gene in the mouse with partial functional overlap with N-ras. Genes Dev 11, 2468–2481.

Johnson, L., Mercer, K., Greenbaum, D., Bronson, R.T., Crowley, D., Tuveson, D.A., and Jacks, T. (2001). Somatic activation of the K-ras oncogene causes early onset lung cancer in mice. Nature 410, 1111–1116.

Jones, S.N., Roe, A.E., Donehower, L.A., and Bradley, A. (1995). Rescue of embryonic lethality in Mdm2-deficient mice by absence of p53. Nature 378, 206–208.

Jonkers, J., and Berns, A. (2002). Conditional mouse models of sporadic cancer. Nat Rev Cancer 2, 251–265.

Jonkers, J., Meuwissen, R., van der Gulden, H., Peterse, H., van der Valk, M., and Berns, A. (2001). Synergistic tumor suppressor activity of BRCA2 and p53 in a conditional mouse model for breast cancer. Nat Genet 29, 418–425.

Joung, J.K., and Sander, J.D. (2013). TALENs: a widely applicable technology for targeted genome editing. Nat Rev Mol Cell Biol 14, 49–55.

Joung, J.K., Ramm, E.I., and Pabo, C.O. (2000). A bacterial two-hybrid selection system for studying protein-DNA and protein-protein interactions. Proc Natl Acad Sci USA 97, 7382–7387.

Jullien, N., Sampieri, F., Enjalbert, A., and Herman, J.-P. (2003). Regulation of Cre recombinase by ligand-induced complementation of inactive fragments. Nucleic Acids Res 31, 131e–131.

Junttila, M.R., Karnezis, A.N., Garcia, D., Madriles, F., Kortlever, R.M., Rostker, F., Swigart, L.B., Pham, D.M., Seo, Y., Evan, G.I., et al. (2010). Selective activation of p53-mediated tumour suppression in high-grade tumours. Nature 468, 567–571.

Kaelin, W.G. (2005a). The Concept of Synthetic Lethality in the Context of Anticancer Therapy. Nat Rev Cancer 5, 689–698.

Kaelin, W.G. (2012). Molecular biology. Use and abuse of RNAi to study mammalian gene function. Science 337, 421–422.

Kaelin, W.G.J. (2005b). ROS: really involved in oxygen sensing. Cell Metabolism 1, 357–358.

Kamata, H., Honda, S.-I., Maeda, S., Chang, L., Hirata, H., and Karin, M. (2005). Reactive oxygen species promote TNFalpha-induced death and sustained JNK activation by inhibiting MAP kinase phosphatases. Cell 120, 649–661.

Kamijo, T., Zindy, F., Roussel, M.F., Quelle, D.E., Downing, J.R., Ashmun, R.A., Grosveld, G., and Sherr, C.J. (1997). Tumor suppression at the mouse INK4a locus mediated by the alternative reading frame product p19ARF. Cell 91, 649–659.

Karnoub, A., and Weinberg, R. (2008). Ras oncogenes: split personalities. Nat Rev Mol Cell Biol 9, 517–531.

Kataoka, K., Noda, M., and Nishizawa, M. (1994). Maf nuclear oncoprotein recognizes sequences related to an AP-1 site and forms heterodimers with both Fos and Jun. Mol Cell Biol 14, 700–712.

Katoh, Y., Itoh, K., Yoshida, E., Miyagishi, M., Fukamizu, A., and Yamamoto, M. (2001). Two domains of Nrf2 cooperatively bind CBP, a CREB binding protein, and synergistically activate transcription. Genes Cells 6, 857–868.

Kats, L.M., Reschke, M., Taulli, R., Pozdnyakova, O., Burgess, K., Bhargava, P., Straley, K., Karnik, R., Meissner, A., Small, D., et al. (2014). Proto-oncogenic role of mutant IDH2 in leukemia initiation and maintenance. Cell Stem Cell 14, 329–341.

Kearns, N.A., Pham, H., Tabak, B., Genga, R.M., Silverstein, N.J., Garber, M., and Maehr, R. (2015). Functional annotation of native enhancers with a Cas9–histone demethylase fusion. Nat Methods.

Kennedy, M.J., Hughes, R.M., Peteya, L.A., Schwartz, J.W., Ehlers, M.D., and Tucker, C.L. (2010). Rapid blue-light–mediated induction of protein interactions in living cells. Nat Methods 7, 973–975.

Kiani, S., Chavez, A., Tuttle, M., Hall, R.N., Chari, R., Ter-Ovanesyan, D., Qian, J., Pruitt, B.W., Beal, J., Vora, S., et al. (2015). Cas9 gRNA engineering for genome editing, activation and repression. Nat Methods 12, 1051–1054.

Kim, S., Kim, D., Cho, S.W., Kim, J., and Kim, J.-S. (2014). Highly efficient RNA-guided genome editing in human cells via delivery of purified Cas9 ribonucleoproteins. Genome Research 24, 1012–1019.

Kim, Y.C., Masutani, H., Yamaguchi, Y., Itoh, K., Yamamoto, M., and Yodoi, J. (2001). Hemin-induced activation of the thioredoxin gene by Nrf2. A differential regulation of the antioxidant responsive element by a switch of its binding factors. J Biol Chem 276, 18399–18406.

Kim, Y.G., and Chandrasegaran, S. (1994). Chimeric restriction endonuclease. Proc Natl Acad Sci USA 91, 883–887.

Kim, Y.G., Cha, J., and Chandrasegaran, S. (1996). Hybrid restriction enzymes: Zinc finger fusions to Fok I cleavage domain. Proc Natl Acad Sci USA 93, 1156–1160.

Kim, Y.R., Oh, J.E., Kim, M.S., Kang, M.R., Park, S.W., Han, J.Y., Eom, H.S., Yoo, N.J., and Lee, S.H. (2010). Oncogenic NRF2 mutations in squamous cell carcinomas of oesophagus and skin. The Journal of Pathology 220, 446–451.

Kim, Y.-J., Ahn, J.-Y., Liang, P., Ip, C., Zhang, Y., and Park, Y.-M. (2007). Human prx1 gene is a target of Nrf2 and is up-regulated by hypoxia/reoxygenation: implication to tumor biology. Cancer Res 67, 546–554.

Kiriakidou, M., Tan, G.S., Lamprinaki, S., De Planell-Saguer, M., Nelson, P.T., and Mourelatos, Z. (2007). An mRNA m7G cap binding-like motif within human Ago2 represses translation. Cell 129, 1141–1151.

Kistner, A., Gossen, M., Zimmermann, F., Jerecic, J., Ullmer, C., Lübbert, H., and Bujard, H. (1996). Doxycycline-mediated quantitative and tissue-specific control of gene expression in transgenic mice. Proc Natl Acad Sci USA 93, 10933–10938.

Kita, K., Kotani, H., Sugisaki, H., and Takanami, M. (1989). The FokI restriction-modification system. I. Organization and nucleotide sequences of the restriction and modification genes. Journal of Biological Chemistry 264, 5751–5756.

Kleinstiver, B.P., Prew, M.S., Tsai, S.Q., Nguyen, N.T., Topkar, V.V., Zheng, Z., and Joung, J.K. (2015a). Broadening the targeting range of Staphylococcus aureus CRISPR-Cas9 by modifying PAM recognition. Nat Biotechnol.

Kleinstiver, B.P., Prew, M.S., Tsai, S.Q., Topkar, V.V., Nguyen, N.T., Zheng, Z., Gonzales, A.P.W., Li, Z., Peterson, R.T., Yeh, J.-R.J., et al. (2015b). Engineered CRISPR-Cas9 nucleases with altered PAM specificities. Nature –.

Knudson, A.G. (1971). Mutation and Cancer: Statistical Study of Retinoblastoma. Proceedings of the National Academy of Sciences 68, 820–823.

Kobayashi, A., Kang, M.-I., Okawa, H., Ohtsuji, M., Zenke, Y., Chiba, T., Igarashi, K., and Yamamoto, M. (2004). Oxidative Stress Sensor Keap1 Functions as an Adaptor for Cul3-Based E3 Ligase To Regulate Proteasomal Degradation of Nrf2. Mol Cell Biol 24, 7130–7139.

Koike-Yusa, H., Li, Y., Tan, E.-P., Velasco-Herrera, M.D.C., and Yusa, K. (2013). Genome-wide recessive genetic screening in mammalian cells with a lentiviral CRISPR-guide RNA library. Nat Biotechnol 32, 267–273.

Koller, B.H., and Smithies, O. (1989). Inactivating the beta 2-microglobulin locus in mouse embryonic stem cells by homologous recombination. Proceedings of the National Academy of Sciences 86, 8932–8935.

Konermann, S., Brigham, M.D., Trevino, A.E., Joung, J., Abudayyeh, O.O., Barcena, C., Hsu, P.D., Habib, N., Gootenberg, J.S., Nishimasu, H., et al. (2014). Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. Nature 517, 583–588.

Kostriken, R., Strathern, J.N., Klar, A.J., Hicks, J.B., and Heffron, F. (1983). A site-specific endonuclease essential for mating-type switching in Saccharomyces cerevisiae. Cell 35, 167–174.

Kouzarides, T. (2007). Chromatin modifications and their function. Cell 128, 693–705.

Kravchenko, J.E., Ilyinskaya, G.V., Komarov, P.G., Agapova, L.S., Kochetkov, D.V., Strom, E., Frolova, E.I., Kovriga, I., Gudkov, A.V., Feinstein, E., et al. (2008). Small-molecule RETRA suppresses mutant p53-bearing cancer cells through a p73-dependent salvage pathway. Proc Natl Acad Sci USA 105, 6302–6307.

Kreidberg, J.A., Sariola, H., Loring, J.M., Maeda, M., Pelletier, J., Housman, D., and Jaenisch, R. (1993). WT-1 is required for early kidney development. Cell 74, 679–691.

Kris, M.G., Natale, R.B., Herbst, R.S., Lynch, T.J.J., Prager, D., Belani, C.P., Schiller, J.H., Kelly, K., Spiridonidis, H., Sandler, A., et al. (2003). Efficacy of gefitinib, an inhibitor of the epidermal growth factor receptor tyrosine kinase, in symptomatic patients with non-small cell lung cancer: a randomized trial. Jama 290, 2149–2158.

Kubbutat, M.H., Jones, S.N., and Vousden, K.H. (1997). Regulation of p53 stability by Mdm2. Nature 387, 299–303.

Kunath, T., Gish, G., Lickert, H., Jones, N., Pawson, T., and Rossant, J. (2003). Transgenic RNA interference in ES cell-derived embryos recapitulates a genetic null phenotype. Nat Biotechnol 21, 559–561.

Lakso, M., Sauer, B., Mosinger, B.J., Lee, E.J., Manning, R.W., Yu, S.H., Mulder, K.L., and Westphal, H. (1992). Targeted oncogene activation by site-specific recombination in transgenic mice. Proc Natl Acad Sci USA 89, 6232–6236.

Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., et al. (2001). Initial sequencing and analysis of the human genome. Nature 409, 860–921.

Lang, G.A., Iwakuma, T., Suh, Y.-A., Liu, G., Rao, V.A., Parant, J.M., Valentin-Vega, Y.A., Terzian, T., Caldwell, L.C., Strong, L.C., et al. (2004). Gain of function of a p53 hot spot mutation in a mouse model of Li-Fraumeni syndrome. Cell 119, 861–872.

Lawrence, M.S., Stojanov, P., Mermel, C.H., Robinson, J.T., Garraway, L.A., Golub, T.R., Meyerson, M., Gabriel, S.B., Lander, E.S., and Getz, G. (2014). Discovery and saturation analysis of cancer genes across 21 tumour types. Nature 505, 495–501.

Lee, A.C., Fenster, B.E., Ito, H., Takeda, K., Bae, N.S., Hirai, T., Yu, Z.X., Ferrans, V.J., Howard, B.H., and Finkel, T. (1999). Ras proteins induce senescence by altering the intracellular levels of reactive oxygen species. J Biol Chem 274, 7936–7940.

Lee, E.Y., Chang, C.Y., Hu, N., Wang, Y.C., Lai, C.C., Herrup, K., Lee, W.H., and Bradley, A. (1992). Mice deficient for Rb are nonviable and show defects in neurogenesis and haematopoiesis. Nature 359, 288–294.

Lee, Y., Ahn, C., Han, J., Choi, H., Kim, J., Yim, J., Lee, J., Provost, P., Rådmark, O., Kim, S., et al. (2003). The nuclear RNase III Drosha initiates microRNA processing. Nat Cell Biol 425, 415–419.

Li, T., Huang, S., Zhao, X., Wright, D.A., Carpenter, S., Spalding, M.H., Weeks, D.P., and Yang, B. (2011). Modularly assembled designer TAL effector nucleases for targeted gene knockout and gene replacement in eukaryotes. Nucleic Acids Res 39, 6315–6325.

Lii, C.-K., Liu, K.-L., Cheng, Y.-P., Lin, A.-H., Chen, H.-W., and Tsai, C.-W. (2010). Sulforaphane and alpha-lipoic acid upregulate the expression of the pi class of glutathione S-transferase through c-jun and Nrf2 activation. J Nutr 140, 885–892.

Lin, S., Staahl, B.T., Alla, R.K., and Doudna, J.A. (2014). Enhanced homology-directed human genome engineering by controlled timing of CRISPR/Cas9 delivery. eLife Sciences 4.

Littlewood, T.D., Hancock, D.C., Danielian, P.S., Parker, M.G., and Evan, G.I. (1995). A modified oestrogen receptor ligand-binding domain as an improved switch for the regulation of heterologous proteins. Nucleic Acids Res 23, 1686–1690.

Liu, J., Carmell, M.A., Rivas, F.V., Marsden, C.G., Thomson, J.M., Song, J.-J., Hammond, S.M., Joshua-Tor, L., and Hannon, G.J. (2004). Argonaute2 is the catalytic engine of mammalian RNAi. Science 305, 1437–1441.

Liu, J., Li, C., Yu, Z., Huang, P., Wu, H., Wei, C., Zhu, N., Shen, Y., Chen, Y., Zhang, B., et al. (2012). Efficient and specific modifications of the Drosophila genome by means of an easy TALEN strategy. J Genet Genomics 39, 209–215.

Livshits, G., and Lowe, S.W. (2013). Accelerating cancer modeling with RNAi and nongermline genetically engineered mouse models. Cold Spring Harb Protoc 2013.

Lo, S.-C., Li, X., Henzl, M.T., Beamer, L.J., and Hannink, M. (2006). Structure of the Keap1:Nrf2 interface provides mechanistic insight into Nrf2 signaling. Embo J 25, 3605–3617.

Lobe, C.G., and Nagy, A. (1998). Conditional genome alteration in mice. Bioessays 20, 200–208.

Lohr, J.G., Stojanov, P., Carter, S.L., Cruz-Gordillo, P., Lawrence, M.S., Auclair, D., Sougnez, C., Knoechel, B., Gould, J., Saksena, G., et al. (2014). Widespread genetic heterogeneity in multiple myeloma: implications for targeted therapy. Cancer Cell 25, 91–101.

Lombardo, A., Genovese, P., Beausejour, C.M., Colleoni, S., Lee, Y.-L., Kim, K.A., Ando, D., Urnov, F.D., Galli, C., Gregory, P.D., et al. (2007). Gene editing in human stem cells using zinc finger nucleases and integrase-defective lentiviral vector delivery. Nat Biotechnol 25, 1298–1306.

Looney, M.C., Moran, L.S., Jack, W.E., Feehery, G.R., Benner, J.S., Slatko, B.E., and Wilson, G.G. (1989). Nucleotide sequence of the FokI restriction-modification system: separate strand-specificity domains in the methyltransferase. Gene 80, 193–208.

Losman, J.-A., Looper, R.E., Koivunen, P., Lee, S., Schneider, R.K., McMahon, C., Cowley, G.S., Root, D.E., Ebert, B.L., and Kaelin, W.G.J. (2013). (R)-2-hydroxyglutarate is sufficient to promote leukemogenesis and its effects are reversible. Science 339, 1621–1625.

Lowe, S.W., Schmitt, E.M., Smith, S.W., Osborne, B.A., and Jacks, T. (1993). p53 is required for radiation-induced apoptosis in mouse thymocytes. Nature 362, 847–849.

Lu, C., Ward, P.S., Kapoor, G.S., Rohle, D., Turcan, S., Abdel-Wahab, O., Edwards, C.R., Khanin, R., Figueroa, M.E., Melnick, A., et al. (2012). IDH mutation impairs histone demethylation and results in a block to cell differentiation. Nature 483, 474–478.

Lund, E., Güttinger, S., Calado, A., Dahlberg, J.E., and Kutay, U. (2004). Nuclear Export of MicroRNA Precursors. Science 303, 95–98.

Lunt, S.Y., and Vander Heiden, M.G. (2011). Aerobic glycolysis: meeting the metabolic requirements of cell proliferation. Annu Rev Cell Dev Biol 27, 441–464.

Luo, F., Poulogiannis, G., Ye, H., Hamoudi, R., Dong, G., Zhang, W., Ibrahim, A.E.K., and Arends, M.J. (2014). Wild-type K-ras has a tumour suppressor effect on carcinogen-induced murine colorectal adenoma formation. Int J Exp Pathol 95, 8–15.

Luo, J., Emanuele, M., Li, D., Creighton, C., Schlabach, M., Westbrook, T., Wong, K.-K., and Elledge, S. (2009a). A Genome-wide RNAi Screen Identifies Multiple Synthetic Lethal Interactions with the Ras Oncogene. Cell 137, 835–848.

Luo, J., Solimini, N., and Elledge, S. (2009b). Principles of Cancer Therapy: Oncogene and Non-oncogene Addiction. Cell 136, 823–837.

Lupiáñez, D.G., Kraft, K., Heinrich, V., Krawitz, P., Brancati, F., Klopocki, E., Horn, D., Kayserili, H., Opitz, J.M., Laxova, R., et al. (2015). Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer interactions. Cell 161, 1012–1025.

Lynch, T.J., Bell, D.W., Sordella, R., Gurubhagavatula, S., Okimoto, R.A., Brannigan, B.W., Harris, P.L., Haserlat, S.M., Supko, J.G., Haluska, F.G., et al. (2004). Activating Mutations in the Epidermal Growth Factor Receptor Underlying Responsiveness of Non–Small-Cell Lung Cancer to Gefitinib. N Engl J Med 350, 2129–2139.

Ma, Y., Shen, B., Zhang, X., Lu, Y., Chen, W., Ma, J., Huang, X., and Zhang, L. (2014). Heritable Multiplex Genetic Engineering in Rats Using CRISPR/Cas9. PLoS ONE 9, e89413.

Maddalo, D., Manchado, E., Concepcion, C.P., Bonetti, C., Vidigal, J.A., Han, Y.-C., Ogrodowski, P., Crippa, A., Rekhtman, N., De Stanchina, E., et al. (2014). In vivo engineering of oncogenic chromosomal rearrangements with the CRISPR/Cas9 system. Nature 516, 423–427.

Maeder, M.L., Linder, S.J., Cascio, V.M., Fu, Y., Ho, Q.H., and Joung, J.K. (2013). CRISPR RNA-guided activation of endogenous human genes. Nat Methods 10, 977–979.

Mali, P., Yang, L., Esvelt, K.M., Aach, J., Guell, M., DiCarlo, J.E., Norville, J.E., and Church, G.M. (2013a). RNA-Guided Human Genome Engineering via Cas9. Science 339, 823–826.

Mali, P., Aach, J., Stranges, P.B., Esvelt, K.M., Moosburner, M., Kosuri, S., Yang, L., and Church, G.M. (2013b). CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. Nat Biotechnol 31, 833–838.

Malina, A., Mills, J.R., Cencic, R., Yan, Y., Fraser, J., Schippers, L.M., Paquet, M., Dostie, J., and Pelletier, J. (2013). Repurposing CRISPR/Cas9 for in situ functional assays. Genes Dev 27, 2602–2614.

Mansfield, K.D., Guzy, R.D., Pan, Y., Young, R.M., Cash, T.P., Schumacker, P.T., and Simon, M.C. (2005). Mitochondrial dysfunction resulting from loss of cytochrome c impairs cellular oxygen sensing and hypoxic HIF-alpha activation. Cell Metabolism 1, 393–399.

Mansour, M.R., Abraham, B.J., Anders, L., Berezovskaya, A., Gutierrez, A., Durbin, A.D., Etchin, J., Lawton, L., Sallan, S.E., Silverman, L.B., et al. (2014). An oncogenic super-enhancer formed through somatic mutation of a noncoding intergenic element. Science 346, 1373–1377.

Mansour, S.L., Thomas, K.R., and Capecchi, M.R. (1988). Disruption of the proto-oncogene int-2 in mouse embryo-derived stem cells: a general strategy for targeting mutations to non-selectable genes. Nature 336, 348–352.

Mardis, E.R., Ding, L., Dooling, D.J., Larson, D.E., McLellan, M.D., Chen, K., Koboldt, D.C., Fulton, R.S., Delehaunty, K.D., McGrath, S.D., et al. (2009). Recurring Mutations Found by Sequencing an Acute Myeloid Leukemia Genome. N Engl J Med 361, 1058–1066.

Marino, S., Vooijs, M., van Der Gulden, H., Jonkers, J., and Berns, A. (2000). Induction of medulloblastomas in p53-null mutant mice by somatic inactivation of Rb in the external granular layer cells of the cerebellum. Genes Dev 14, 994–1004.

Marraffini, L.A., and Sontheimer, E.J. (2008). CRISPR Interference Limits Horizontal Gene Transfer in Staphylococci by Targeting DNA. Science 322, 1843–1845.

Martin, G.S. (1970). Rous Sarcoma Virus: a Function required for the Maintenance of the Transformed State. Nature 227, 1021–1023.

Martins, C.P., Brown-Swigart, L., and Evan, G.I. (2006). Modeling the therapeutic efficacy of p53 restoration in tumors. Cell 127, 1323–1334.

Marumoto, T., Tashiro, A., Friedmann-Morvinski, D., Scadeng, M., Soda, Y., Gage, F.H., and Verma, I.M. (2009). Development of a novel mouse glioma model using lentiviral vectors. Nat Med 15, 110–116.

Matano, M., Date, S., Shimokawa, M., Takano, A., Fujii, M., Ohta, Y., Watanabe, T., Kanai, T., and Sato, T. (2015). Modeling colorectal cancer using CRISPR-Cas9–mediated engineering of human intestinal organoids. Nature Medicine –.

McClatchey, A.I., Saotome, I., Ramesh, V., Gusella, J.F., and Jacks, T. (1997). The Nf2 tumor suppressor gene product is essential for extraembryonic development immediately prior to gastrulation. Genes Dev 11, 1253–1265.

McMahon, M., Itoh, K., Yamamoto, M., Chanas, S.A., Henderson, C.J., McLellan, L.I., Wolf, C.R., Cavin, C., and Hayes, J.D. (2001). The cap “n” collar basic leucine zipper transcription factor Nrf2 (NF-E2 p45-related factor 2) controls both constitutive and inducible expression of intestinal detoxification and glutathione biosynthetic enzymes. Cancer Research 61, 3299–3307.

McMahon, M., Thomas, N., Itoh, K., Yamamoto, M., and Hayes, J.D. (2004). Redox-regulated Turnover of Nrf2 Is Determined by at Least Two Separate Protein Domains, the Redox-sensitive Neh2 Degron and the Redox-insensitive Neh6 Degron. Journal of Biological Chemistry 279, 31556–31567.

McMahon, M., Thomas, N., Itoh, K., Yamamoto, M., and Hayes, J.D. (2006). Dimerization of Substrate Adaptors Can Facilitate Cullin-mediated Ubiquitylation of Proteins by a “Tethering” Mechanism: A TWO-

SITE INTERACTION MODEL FOR THE Nrf2-Keap1 COMPLEX. Journal of Biological Chemistry 281, 24756–24768.

McManus, M.T., and Sharp, P.A. (2002). Gene silencing in mammals by small interfering RNAs. Nat Rev Genet 3, 737–747.

Meacham, C.E., Ho, E.E., Dubrovsky, E., Gertler, F.B., and Hemann, M.T. (2009). In vivo RNAi screening identifies regulators of actin dynamics as key determinants of lymphoma progression. Nat Genet 41, 1133–1137.

Meacham, C.E., Lawton, L.N., Soto-Feliciano, Y.M., Pritchard, J.R., Joughin, B.A., Ehrenberger, T., Fenouille, N., Zuber, J., Williams, R.T., Young, R.A., et al. (2015). A genome-scale in vivo loss-of-function screen identifies Phf6as a lineage-specific regulator of leukemia cell growth. Genes Dev 29, 483–488.

Meister, A. (1983). Selective modification of glutathione metabolism. Science 220, 472–477.

Meister, G., and Tuschl, T. (2004). Mechanisms of gene silencing by double-stranded RNA. Nat Cell Biol 431, 343–349.

Mello, C.C., and Conte, D. (2004). Revealing the world of RNA interference. Nat Cell Biol 431, 338–342.

Meng, X., Noyes, M.B., Zhu, L.J., Lawson, N.D., and Wolfe, S.A. (2008). Targeted gene inactivation in zebrafish using engineered zinc-finger nucleases. Nat Biotechnol 26, 695–701.

Mercer, K., Giblett, S., Green, S., Lloyd, D., Dias, S.D., Plumb, M., Marais, R., and Pritchard, C. (2005). Expression of endogenous oncogenic V600EB-raf induces proliferation and developmental defects in mice and transformation of primary fibroblasts. Cancer Research 65, 11493–11500.

Messing, A., Chen, H.Y., Palmiter, R.D., and Brinster, R.L. (1985). Peripheral neuropathies, hepatocellular carcinomas and islet cell adenomas in transgenic mice. Nature 316, 461–463.

Meuwissen, R., Linn, S.C., Linnoila, R.I., Zevenhoven, J., Mooi, W.J., and Berns, A. (2003). Induction of small cell lung cancer by somatic inactivation of both Trp53 and Rb1 in a conditional mouse model. Cancer Cell 4, 181–189.

Meyer, M., de Angelis, M.H., Wurst, W., and Kühn, R. (2010). Gene targeting by homologous recombination in mouse zygotes mediated by zinc-finger nucleases. Proc Natl Acad Sci USA 107, 15022–15026.

Meylan, E., Dooley, A.L., Feldser, D.M., Shen, L., Turk, E., Ouyang, C., and Jacks, T. (2009). Requirement for NF-kappaB signalling in a mouse model of lung adenocarcinoma. Nature 462, 104–107.

Miething, C., Scuoppo, C., Bosbach, B., Appelmann, I., Nakitandwe, J., Ma, J., Wu, G., Lintault, L., Auer, M., Premsrirut, P.K., et al. (2014). PTEN action in leukaemia dictated by the tissue microenvironment. Nature 510, 402–406.

Miller, J.C., and Pabo, C.O. (2001). Rearrangement of Side-chains in a Zif268 Mutant Highlights the Complexities of Zinc Finger-DNA Recognition. J Mol Biol 313, 309–315.

Miller, J., McLachlan, A.D., and Klug, A. (1985). Repetitive zinc-binding domains in the protein transcription factor IIIA from Xenopus oocytes. Embo J 4, 1609–1614.

Miller, J.C., Tan, S., Qiao, G., Barlow, K.A., Wang, J., Xia, D.F., Meng, X., Paschon, D.E., Leung, E., Hinkley, S.J., et al. (2011). A TALE nuclease architecture for efficient genome editing. Nat Biotechnol 29, 143–148.

Miller, P.G., Al-Shahrour, F., Hartwell, K.A., Chu, L.P., Järås, M., Puram, R.V., Puissant, A., Callahan, K.P., Ashton, J., McConkey, M.E., et al. (2013). In Vivo RNAi screening identifies a leukemia-specific dependence on integrin beta 3 signaling. Cancer Cell 24, 45–58.

Minsavage, G.V., Dahlbeck, D., Whalen, M.C., Kearney, B., Bonas, U., Staskawicz, B.J., and Stall, R.E. (1990). Gene-for-gene relationships specifying disease resistance in Xanthomonas campestris pv. vesicatoria—pepper interactions. Molecular Plant-Microbe Interactions 3, 41–47.

Mitelman, F., Johansson, B., and Mertens, F. (2007). The impact of translocations and gene fusions on cancer causation. Nat Rev Cancer 7, 233–245.

Mitsuishi, Y., Motohashi, H., and Yamamoto, M. (2012). The Keap1-Nrf2 system in cancers: stress response and anabolic metabolism. Front. Oncol. 2, 200.

Mitsuishi, Y., Taguchi, K., Kawatani, Y., Shibata, T., Nukiwa, T., Aburatani, H., Yamamoto, M., and Motohashi, H. (2012b). Nrf2 redirects glucose and glutamine into anabolic pathways in metabolic reprogramming. Cancer Cell 22, 66–79.

Mohr, S.E., Smith, J.A., Shamu, C.E., Neumüller, R.A., and Perrimon, N. (2014). RNAi screening comes of age: improved techniques and complementary approaches. Nat Rev Mol Cell Biol 15, 591–600.

Moi, P., Chan, K., Asunis, I., Cao, A., and Kan, Y.W. (1994). Isolation of NF-E2-related factor 2 (Nrf2), a NF-E2-like basic leucine zipper transcriptional activator that binds to the tandem NF-E2/AP1 repeat of the beta-globin locus control region. Proc Natl Acad Sci USA 91, 9926–9930.

Mojica, F.J., Díez-Villaseñor, C., Soria, E., and Juez, G. (2000). Biological significance of a family of regularly spaced repeats in the genomes of Archaea, Bacteria and mitochondria. Mol. Microbiol. 36, 244–246.

Mojica, F.J.M., Díez-Villaseñor, C., García-Martínez, J., and Soria, E. (2005). Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. J. Mol. Evol. 60, 174–182.

Montes de Oca Luna, R., Wagner, D.S., and Lozano, G. (1995). Rescue of early embryonic lethality in mdm2-deficient mice by deletion of p53. Nature 378, 203–206.

Moore, M., Klug, A., and Choo, Y. (2001). Improved DNA binding specificity from polyzinc finger peptides by using strings of two-finger units. Proc Natl Acad Sci USA 98, 1437–1441.

Morton, J., Davis, M.W., Jorgensen, E.M., and Carroll, D. (2006). Induction and repair of zinc-finger nuclease-targeted double-strand breaks in Caenorhabditis elegans somatic cells. Proceedings of the National Academy of Sciences 103, 16370–16375.

Moscou, M.J., and Bogdanove, A.J. (2009). A Simple Cipher Governs DNA Recognition by TAL Effectors. Science 326, 1501–1501.

Moser, A.R., Shoemaker, A.R., Connelly, C.S., Clipson, L., Gould, K.A., Luongo, C., Dove, W.F., Siggers, P.H., and Gardner, R.L. (1995). Homozygosity for the Min allele of Apc results in disruption of mouse development prior to gastrulation. Dev. Dyn. 203, 422–433.

Munoz, I.G., Prieto, J., Subramanian, S., Coloma, J., Redondo, P., Villate, M., Merino, N., Marenchino, M., D'Abramo, M., Gervasio, F.L., et al. (2011). Molecular basis of engineered meganuclease targeting of the endogenous human RAG1 locus. Nucleic Acids Res 39, 729–743.

Murphy, M.P. (2012). Mitochondrial Thiols in Antioxidant Protection and Redox Signaling: Distinct Roles for Glutathionylation and Other Thiol Modifications. Antioxid Redox Signal 16, 476–495.

Nakahira, K., Haspel, J.A., Rathinam, V.A.K., Lee, S.-J., Dolinay, T., Lam, H.C., Englert, J.A., Rabinovitch, M., Cernadas, M., Kim, H.P., et al. (2011). Autophagy proteins regulate innate immune responses by inhibiting the release of mitochondrial DNA mediated by the NALP3 inflammasome. Nature Publishing Group 12, 222–230.

Nakano, H., Yamamoto, F., Neville, C., Evans, D., Mizuno, T., and Perucho, M. (1984). Isolation of transforming sequences of two human lung carcinomas: structural and functional analysis of the activated c-K-ras oncogenes. Proc Natl Acad Sci USA 81, 71–75.

Nalca, A., Qiu, S.G., El-Guendy, N., Krishnan, S., and Rangnekar, V.M. (1999). Oncogenic Ras sensitizes cells to apoptosis by Par-4. J Biol Chem 274, 29976–29983.

Nguyen, T., Nioi, P., and Pickett, C.B. (2009). The Nrf2-Antioxidant Response Element Signaling Pathway and Its Activation by Oxidative Stress. Journal of Biological Chemistry 284, 13291–13295.

Nickoloff, J.A., Chen, E.Y., and Heffron, F. (1986). A 24-base-pair DNA sequence from the MAT locus stimulates intergenic recombination in yeast. Proc Natl Acad Sci USA 83, 7831–7835.

Nigro, J.M., Baker, S.J., Preisinger, A.C., Jessup, J.M., Hosteller, R., Cleary, K., Signer, S.H., Davidson, N., Baylin, S., Devilee, P., et al. (1989). Mutations in the p53 gene occur in diverse human tumour types. Nature 342, 705–708.

Nihongaki, Y., Kawano, F., Nakajima, T., and Sato, M. (2015a). Photoactivatable CRISPR-Cas9 for optogenetic genome editing. Nat Biotechnol 33, 755–760.

Nihongaki, Y., Yamamoto, S., Kawano, F., Suzuki, H., and Sato, M. (2015b). CRISPR-Cas9-based photoactivatable transcription system. Chemistry & Biology 22, 169–174.

Nioi, P., and Nguyen, T. (2007). A mutation of Keap1 found in breast cancer impairs its ability to repress Nrf2 activity. Biochem Biophys Res Commun 362, 816–821.

Nioi, P., Nguyen, T., Sherratt, P.J., and Pickett, C.B. (2005). The carboxy-terminal Neh3 domain of Nrf2 is required for transcriptional activation. Mol Cell Biol 25, 10895–10906.

Niu, Y., Shen, B., Cui, Y., Chen, Y., Wang, J., Wang, L., Kang, Y., Zhao, X., Si, W., Li, W., et al. (2014). Generation of gene-modified cynomolgus monkey via Cas9/RNA-mediated gene targeting in one-cell embryos. Cell 156, 836–843.

O'Gorman, S., Fox, D.T., and Wahl, G.M. (1991). Recombinase-mediated gene activation and site-specific integration in mammalian cells. Science 251, 1351–1355.

Ohsaki, Y., Tanno, S., Fujita, Y., Toyoshima, E., Fujiuchi, S., Nishigaki, Y., Ishida, S., Nagase, A., Miyokawa, N., Hirata, S., et al. (2000). Epidermal growth factor receptor expression correlates with poor prognosis in non-small cell lung cancer patients with p53 overexpression. Oncol Rep 7, 603–607.

Ohta, T., Iijima, K., Miyamoto, M., Nakahara, I., Tanaka, H., Ohtsuji, M., Suzuki, T., Kobayashi, A., Yokota, J., Sakiyama, T., et al. (2008). Loss of Keap1 function activates Nrf2 and provides advantages for lung cancer cell growth. Cancer Research 68, 1303–1309.

Okorokov, A.L., and Milner, J. (1999). An ATP/ADP-dependent molecular switch regulates the stability of p53-DNA complexes. Mol Cell Biol 19, 7501–7510.

Olive, K., Tuveson, D., Ruhe, Z., Yin, B., Willis, N., Bronson, R., Crowley, D., and Jacks, T. (2004). Mutant p53 Gain of Function in Two Mouse Models of Li-Fraumeni Syndrome. Cell 119, 847–860.

Oliver, T.G., Meylan, E., Chang, G.P., Xue, W., Burke, J.R., Humpton, T.J., Hubbard, D., Bhutkar, A., and Jacks, T. (2011). Caspase-2-Mediated Cleavage of Mdm2 Creates a p53-Induced Positive Feedback Loop. Mol Cell 43, 57–71.

Orban, P.C., Chui, D., and Marth, J.D. (1992). Tissue- and site-specific DNA recombination in transgenic mice. Proc Natl Acad Sci USA 89, 6861–6865.

Ornitz, D.M., Hammer, R.E., Messing, A., Palmiter, R.D., and Brinster, R.L. (1987). Pancreatic neoplasia induced by SV40 T-antigen expression in acinar cells of transgenic mice. Science 238, 188–193.

Orr-Weaver, T.L., Szostak, J.W., and Rothstein, R.J. (1981). Yeast transformation: a model system for the study of recombination. Proceedings of the National Academy of Sciences 78, 6354–6358.

Orr-Weaver, T.L., Szostak, J.W., and Rothstein, R.J. (1983). Genetic applications of yeast transformation with linear and gapped plasmids. Meth Enzymol 101, 228–245.

Oshima, M., Oshima, H., Kitagawa, K., Kobayashi, M., Itakura, C., and Taketo, M. (1995). Loss of Apc heterozygosity and abnormal tissue building in nascent intestinal polyps in mice carrying a truncated Apc gene. Proc Natl Acad Sci USA 92, 4482–4486.

Paddison, P.J., and Hannon, G.J. (2002). RNA interference: the new somatic cell genetics? Cancer Cell 2, 17–23.

Padmanabhan, B., Tong, K.I., Ohta, T., Nakamura, Y., Scharlock, M., Ohtsuji, M., Kang, M.-I., Kobayashi, A., Yokoyama, S., and Yamamoto, M. (2006). Structural basis for defects of Keap1 activity provoked by its point mutations in lung cancer. Mol Cell 21, 689–700.

Paez, J.G., Jänne, P.A., Lee, J.C., Tracy, S., Greulich, H., Gabriel, S., Herman, P., Kaye, F.J., Lindeman, N., Boggon, T.J., et al. (2004). EGFR mutations in lung cancer: correlation with clinical response to gefitinib therapy. Science 304, 1497–1500.

Pande, J., Szewczyk, M.M., and Grover, A.K. (2010). Phage display: Concept, innovations, applications and future. From Plants to Pharmacy Shelf: Recent Trends in Leads Finding and Bioproduction 28, 849–858.

Pao, W., Miller, V., Zakowski, M., Doherty, J., Politi, K., Sarkaria, I., Singh, B., Heelan, R., Rusch, V., Fulton, L., et al. (2004). EGF receptor gene mutations are common in lung cancers from “never smokers” and are associated with sensitivity of tumors to gefitinib and erlotinib. Proc Natl Acad Sci USA 101, 13306–13311.

Parada, L.F., Tabin, C.J., Shih, C., and Weinberg, R.A. (1982). Human EJ bladder carcinoma oncogene is homologue of Harvey sarcoma virus ras gene. Nature 297, 474–478.

Parnas, O., Jovanovic, M., Eisenhaure, T.M., Herbst, R.H., Dixit, A., Ye, C.J., Przybylski, D., Platt, R.J., Tirosh, I., Sanjana, N.E., et al. (2015). A Genome-wide CRISPR Screen in Primary Immune Cells to Dissect Regulatory Networks. Cell 162, 675–686.

Parsons, D.W., Jones, S., Zhang, X., Lin, J.C.-H., Leary, R.J., Angenendt, P., Mankoo, P., Carter, H., Siu, I.-M., Gallia, G.L., et al. (2008). An Integrated Genomic Analysis of Human Glioblastoma Multiforme. Science 321, 1807–1812.

Pasqualucci, L., Trifonov, V., Fabbri, G., Ma, J., Rossi, D., Chiarenza, A., Wells, V.A., Grunn, A., Messina, M., Elliot, O., et al. Analysis of the coding genome of diffuse large B-cell lymphoma. Nat Genet 43, 830–837.

Patch, A.-M., Christie, E.L., Etemadmoghadam, D., Garsed, D.W., George, J., Fereday, S., Nones, K., Cowin, P., Alsop, K., Bailey, P.J., et al. (2015). Whole–genome characterization of chemoresistant ovarian cancer. Nature 521, 489–494.

Pavletich, N.P., and Pabo, C.O. (1991). Zinc finger-DNA recognition: crystal structure of a Zif268-DNA complex at 2.1 A. Science 252, 809–817.

Pâques, F., and Duchateau, P. (2007). Meganucleases and DNA double-strand break-induced recombination: perspectives for gene therapy. Cgt 7, 49–66.

Peifer, M., Fernández-Cuesta, L., Sos, M.L., George, J., Seidel, D., Kasper, L.H., Plenker, D., Leenders, F., Sun, R., Zander, T., et al. (2012). Integrative genome analyses identify key somatic driver mutations of small-cell lung cancer. Nat Genet 44, 1104–1110.

Pelengaris, S., Littlewood, T., Khan, M., Elia, G., and Evan, G. (1999). Reversible activation of c-Myc in skin: induction of a complex neoplastic phenotype by a single oncogenic lesion. Mol Cell 3, 565–577.

Perez-Pinera, P., Kocak, D.D., Vockley, C.M., Adler, A.F., Kabadi, A.M., Polstein, L.R., Thakore, P.I., Glass, K.A., Ousterout, D.G., Leong, K.W., et al. (2013). RNA-guided gene activation by CRISPR-Cas9-based transcription factors. Nat Methods 10, 973–976.

Peters, L.L., Andrews, N.C., Eicher, E.M., Davidson, M.B., Orkin, S.H., and Lux, S.E. (1993). Mouse microcytic anaemia caused by a defect in the gene encoding the globin enhancer-binding protein NF-E2. Nature 362, 768–770.

Pfister, N.T., Fomin, V., Regunath, K., Zhou, J.Y., Zhou, W., Silwal-Pandit, L., Freed-Pastor, W.A., Laptenko, O., Neo, S.P., Bargonetti, J., et al. (2015). Mutant p53 cooperates with the SWI/SNF chromatin remodeling complex to regulate VEGFR2 in breast cancer cells. Genes Dev.

Picard, D. (1993). Steroid-binding domains for regulating the functions of heterologous proteins in cis. Trends Cell Biol 3, 278–280.

Picard, D., Salser, S.J., and Yamamoto, K.R. (1988). A movable and regulable inactivation function within the steroid binding domain of the glucocorticoid receptor. Cell 54, 1073–1080.

Pietsch, E.C., Chan, J.Y., Torti, F.M., and Torti, S.V. (2003). Nrf2 mediates the induction of ferritin H in response to xenobiotics and cancer chemopreventive dithiolethiones. Journal of Biological Chemistry 278, 2361–2369.

Pillai, R.S., Bhattacharyya, S.N., Artus, C.G., Zoller, T., Cougot, N., Basyuk, E., Bertrand, E., and Filipowicz, W. (2005). Inhibition of Translational Initiation by Let-7 MicroRNA in Human Cells. Science 309, 1573–1576.

Piskounova, E., Agathocleous, M., Murphy, M.M., Hu, Z., Huddlestun, S.E., Zhao, Z., Leitch, A.M., Johnson, T.M., DeBerardinis, R.J., and Morrison, S.J. (2015). Oxidative stress inhibits distant metastasis by human melanoma cells. Nature advance online publication SP - EP .

Platt, R.J., Chen, S., Zhou, Y., Yim, M.J., Swiech, L., Kempton, H.R., Dahlman, J.E., Parnas, O., Eisenhaure, T.M., Jovanovic, M., et al. (2014). CRISPR-Cas9 knockin mice for genome editing and cancer modeling. Cell 159, 440–455.

Pleasance, E.D., Stephens, P.J., O meara, S., Mcbride, D.J., Meynert, A., Jones, D., Lin, M.-L., Beare, D., Lau, K.W., Greenman, C., et al. (2010). A small-cell lung cancer genome with complex signatures of tobacco exposure. Nature 463, 184–190.

Plessis, A., Perrin, A., Haber, J.E., and Dujon, B. (1992). Site-specific recombination determined by I-SceI, a mitochondrial group I intron-encoded endonuclease expressed in the yeast nucleus. 130, 451–460.

Politi, K., and Herbst, R.S. (2015). Lung Cancer in the Era of Precision Medicine. Clinical Cancer Research 21, 2213–2220.

Polstein, L.R., and Gersbach, C.A. (2015). A light-inducible CRISPR-Cas9 system for control of endogenous gene activation. Nat Chem Biol 11, 198–200.

Porteus, M.H., and Baltimore, D. (2003). Chimeric Nucleases Stimulate Gene Targeting in Human Cells. Science 300, 763–763.

Possemato, R., Marks, K.M., Shaul, Y.D., Pacold, M.E., Kim, D., Birsoy, K., Sethumadhavan, S., Woo, H.-K., Jang, H.G., Jha, A.K., et al. (2011). Functional genomics reveal that the serine synthesis pathway is essential in breast cancer. Nature 476, 346–350.

Pourcel, C., Salvignol, G., and Vergnaud, G. (2005). CRISPR elements in Yersinia pestis acquire new repeats by preferential uptake of bacteriophage DNA, and provide additional tools for evolutionary studies. Microbiology (Reading, Engl.) 151, 653–663.

Premsrirut, P.K., Dow, L.E., Kim, S.Y., Camiolo, M., Malone, C.D., Miething, C., Scuoppo, C., Zuber, J., Dickins, R.A., Kogan, S.C., et al. (2011). A Rapid and Scalable System for Studying Gene Function in Mice Using Conditional RNA Interference. Cell 145, 145–158.

Premsrirut, P.K., Dow, L.E., Park, Y., Hannon, G.J., and Lowe, S.W. (2013). Creating transgenic shRNA mice by recombinase-mediated cassette exchange. Cold Spring Harb Protoc 2013, 835–842.

Prior, I.A., Lewis, P.D., and Mattos, C. (2012). A comprehensive survey of Ras mutations in cancer. Cancer Res 72, 2457–2467.

Pylayeva-Gupta, Y., Grabocka, E., and Bar-Sagi, D. (2011). RAS oncogenes: weaving a tumorigenic web. Nat Rev Cancer 11, 761–774.

Qi, L.S., Larson, M.H., Gilbert, L.A., Doudna, J.A., Weissman, J.S., Arkin, A.P., and Lim, W.A. (2013). Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. Cell 152, 1173–1183.

Quaife, C.J., Pinkert, C.A., Ornitz, D.M., Palmiter, R.D., and Brinster, R.L. (1987). Pancreatic neoplasia induced by ras expression in acinar cells of transgenic mice. Cell 48, 1023–1034.

Ran, F.A., Cong, L., Yan, W.X., Scott, D.A., Gootenberg, J.S., Kriz, A.J., Zetsche, B., Shalem, O., Wu, X., Makarova, K.S., et al. (2015). In vivo genome editing using Staphylococcus aureus Cas9. Nature –.

Ran, F.A., Hsu, P.D., Lin, C.-Y., Gootenberg, J.S., Konermann, S., Trevino, A.E., Scott, D.A., Inoue, A., Matoba, S., Zhang, Y., et al. (2013). Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. Cell 154, 1380–1389.

Rand, T.A., Petersen, S., Du, F., and Wang, X. (2005). Argonaute2 cleaves the anti-guide strand of siRNA during RISC activation. Cell 123, 621–629.

Raoul, C., Abbas-Terki, T., Bensadoun, J.-C., Guillot, S., Haase, G., Szulc, J., Henderson, C.E., and Aebischer, P. (2005). Lentiviral-mediated silencing of SOD1 through RNA interference retards disease onset and progression in a mouse model of ALS. Nat Med 11, 423–428.

Rebar, E.J., and Pabo, C.O. (1994). Zinc finger phage: Affinity selection of fingers with new DNA-binding specificities. Science 263, 671–673.

Reddy, E.P., Reynolds, R.K., Santos, E., and Barbacid, M. (1982). A point mutation is responsible for the acquisition of transforming properties by the T24 human bladder carcinoma oncogene. Nature 300, 149–152.

Redondo, P., Prieto, J., Munoz, I.G., Alibes, A., Stricher, F., Serrano, L., Cabaniols, J.-P., Daboussi, F., Arnould, S., Perez, C., et al. (2008). Molecular basis of xeroderma pigmentosum group C DNA recognition by engineered meganucleases. Nature 456, 107–111.

Reyon, D., Tsai, S.Q., Khayter, C., Foden, J.A., Sander, J.D., and Joung, J.K. (2012). FLASH assembly of TALENs for high-throughput genome editing. Nat Biotechnol 30, 460–465.

Robinson, D., Van Allen, E.M., Wu, Y.-M., Schultz, N., Lonigro, R.J., Mosquera, J.-M., Montgomery, B., Taplin, M.-E., Pritchard, C.C., Attard, G., et al. (2015). Integrative clinical genomics of advanced prostate cancer. Cell 161, 1215–1228.

Robles-Oteiza, C., Taylor, S., Yates, T., Cicchini, M., Lauderback, B., Cashman, C.R., Burds, A.A., Winslow, M.M., Jacks, T., and Feldser, D.M. (2015). Recombinase-based conditional and reversible gene regulation via XTR alleles. Nature Communications 6, 8783.

Rodenhuis, S., van de Wetering, M.L., Mooi, W.J., Evers, S.G., van Zandwijk, N., and Bos, J.L. (1987). Mutational activation of the K-ras oncogene. A possible pathogenetic factor in adenocarcinoma of the lung. Nejm 317, 929–935.

Rohle, D., Popovici-Muller, J., Palaskas, N., Turcan, S., Grommes, C., Campos, C., Tsoi, J., Clark, O., Oldrini, B., Komisopoulou, E., et al. (2013). An inhibitor of mutant IDH1 delays growth and promotes differentiation of glioma cells. Science 340, 626–630.

Rosen, K., Rak, J., Jin, J., Kerbel, R.S., Newman, M.J., and Filmus, J. (1998). Downregulation of the pro-apoptotic protein Bak is required for the ras-induced transformation of intestinal epithelial cells. Curr Biol 8, 1331–1334.

Rossant, J., and Nagy, A. (1995). Genome engineering: the new mouse genetics. Nat Med 1, 592–594.

Rouet, P., Smih, F., and Jasin, M. (1994). Expression of a site-specific endonuclease stimulates homologous recombination in mammalian cells. Proc Natl Acad Sci USA 91, 6064–6068.

ROWLEY, J.D. (1973). A New Consistent Chromosomal Abnormality in Chronic Myelogenous Leukaemia identified by Quinacrine Fluorescence and Giemsa Staining. Nature 243, 290–293.

Rudalska, R., Dauch, D., Longerich, T., McJunkin, K., Wuestefeld, T., Kang, T.-W., Hohmeyer, A., Pesic, M., Leibold, J., Thun, von, A., et al. (2014). In vivo RNAi screening identifies a mechanism of sorafenib resistance in liver cancer. Nat Med 20, 1138–1146.

Rudin, C.M., Durinck, S., Stawiski, E.W., Poirier, J.T., Modrusan, Z., Shames, D.S., Bergbower, E.A., Guan, Y., Shin, J., Guillory, J., et al. (2012). Comprehensive genomic analysis identifies SOX2 as a frequently amplified gene in small-cell lung cancer. Nat Genet 44, 1111–1116.

Rudin, N., Sugarman, E., and Haber, J.E. (1989). Genetic and physical analysis of double-strand break repair and recombination in Saccharomyces cerevisiae. 122, 519–534.

Rushmore, T.H., and Pickett, C.B. (1990). Transcriptional regulation of the rat glutathione S-transferase Ya subunit gene. Characterization of a xenobiotic-responsive element controlling inducible expression by phenolic antioxidants. J Biol Chem 265, 14648–14653.

Rushmore, T.H., Morton, M.R., and Pickett, C.B. (1991). The antioxidant responsive element. Activation by oxidative stress and identification of the DNA consensus sequence required for functional activity. Journal of Biological Chemistry 266, 11632–11639.

Rytlewski, J.A., and Beronja, S. (2015). RNAi in the mouse: rapid and affordable gene function studies in a vertebrate system. WIREs Dev Biol 4, 45–57.

Saborowski, M., Saborowski, A., Morris, J.P.4., Bosbach, B., Dow, L.E., Pelletier, J., Klimstra, D.S., and Lowe, S.W. (2014). A modular and flexible ESC-based mouse model of pancreatic cancer. Genes Dev 28, 85–97.

Sadelain, M., Papapetrou, E.P., and Bushman, F.D. (2012). Safe harbours for the integration of new DNA in the human genome. Nat Rev Cancer 12, 51–58.

Sakurai, A., Nishimoto, M., Himeno, S., Imura, N., Tsujimoto, M., Kunimoto, M., and Hara, S. (2005). Transcriptional regulation of thioredoxin reductase 1 expression by cadmium in vascular endothelial cells: role of NF-E2-related factor-2. J Cell Physiol 203, 529–537.

Samuels, M.L., Weber, M.J., Bishop, J.M., and McMahon, M. (1993). Conditional transformation of cells and rapid activation of the mitogen-activated protein kinase cascade by an estradiol-dependent human raf-1 protein kinase. Mol Cell Biol 13, 6241–6252.

Samuels, Y., Wang, Z., Bardelli, A., Silliman, N., Ptak, J., Szabo, S., Yan, H., Gazdar, A., Powell, S.M., Riggins, G.J., et al. (2004). High Frequency of Mutations of the PIK3CA Gene in Human Cancers. Science 304, 554–554.

Sanchez-Rivera, F.J., Papagiannakopoulos, T., Romero, R., Tammela, T., Bauer, M.R., Bhutkar, A., Joshi, N.S., Subbaraj, L., Bronson, R.T., Xue, W., et al. (2014). Rapid modelling of cooperating genetic events in cancer through somatic genome editing. Nature 516, 428–431.

Sander, J.D., Cade, L., Khayter, C., Reyon, D., Peterson, R.T., Joung, J.K., and Yeh, J.-R.J. (2011). Targeted gene disruption in somatic zebrafish cells using engineered TALENs. Nat Biotechnol 29, 697–698.

Sanjana, N.E., Cong, L., Zhou, Y., Cunniff, M.M., Feng, G., and Zhang, F. (2012). A transcription activator-like effector toolbox for genome engineering. Nat Protoc 7, 171–192.

Santos, C.X.C., Tanaka, L.Y., Wosniak, J., and Laurindo, F.R.M. (2009). Mechanisms and implications of reactive oxygen species generation during the unfolded protein response: roles of endoplasmic reticulum oxidoreductases, mitochondrial electron transport, and NADPH oxidase. Antioxid Redox Signal 11, 2409–2427.

Santos, E., Martin-Zanca, D., Reddy, E.P., Pierotti, M.A., Porta, Della, G., and Barbacid, M. (1984). Malignant activation of a K-ras oncogene in lung carcinoma but not in normal tissue of the same patient. Science 223, 661–664.

Sapranauskas, R., Gasiunas, G., Fremaux, C., Barrangou, R., Horvath, P., and Siksnys, V. (2011). The Streptococcus thermophilus CRISPR/Cas system provides immunity in Escherichia coli. Nucleic Acids Res 39, 9275–9282.

Sasaki, H., Sato, H., Kuriyama-Matsumura, K., Sato, K., Maebara, K., Wang, H., Tamba, M., Itoh, K., Yamamoto, M., and Bannai, S. (2002). Electrophile response element-mediated induction of the cystine/glutamate exchange transporter gene expression. J Biol Chem 277, 44765–44771.

Sauer, B., and Henderson, N. (1988). Site-specific DNA recombination in mammalian cells by the Cre recombinase of bacteriophage P1. Proc Natl Acad Sci USA 85, 5166–5170.

Sawyers, C.L. (2003). Opportunities and challenges in the development of kinase inhibitor therapy for cancer. Genes Dev 17, 2998–3010.

Sayin, V.I., Ibrahim, M.X., Larsson, E., Nilsson, J.A., Lindahl, P., and Bergo, M.O. (2014). Antioxidants Accelerate Lung Cancer Progression in Mice. Science Translational Medicine 6, 221ra15–221ra15.

Schlabach, M., Luo, J., Solimini, N., Hu, G., Xu, Q., Li, M., Zhao, Z., Smogorzewska, A., Sowa, M., Ang, X., et al. (2008). Cancer Proliferation Gene Discovery Through Functional Genomics. Science 319, 620–624.

Schlake, T., and Bode, J. (1994). Use of Mutated FLP Recognition Target (FRT) Sites for the Exchange of Expression Cassettes at Defined Chromosomal Loci. Biochemistry 33, 12746–12751.

Schmitt, C.A., McCurrach, M.E., de Stanchina, E., Wallace-Brodeur, R.R., and Lowe, S.W. (1999). INK4a/ARF mutations accelerate lymphomagenesis and promote chemoresistance by disabling p53. Genes Dev 13, 2670–2677.

Schramek, D., Sendoel, A., Segal, J.P., Beronja, S., Heller, E., Oristian, D., Reva, B., and Fuchs, E. (2014). Direct in vivo RNAi screen unveils myosin IIa as a tumor suppressor of squamous cell carcinomas. Science 343, 309–313.

Scuoppo, C., Miething, C., Lindqvist, L., Reyes, J., Ruse, C., Appelmann, I., Yoon, S., Krasnitz, A., Teruya-Feldstein, J., Pappin, D., et al. (2012). A tumour suppressor network relying on the polyamine-hypusine axis. Nature 487, 244–248.

Seibler, J., Kleinridders, A., Küter-Luks, B., Niehaves, S., Brüning, J.C., and Schwenk, F. (2007). Reversible gene knockdown in mice using a tight, inducible shRNA expression system. Nucleic Acids Res 35, e54.

Seibler, J., Schübeler, D., Fiering, S., Groudine, M., and Bode, J. (1998). DNA Cassette Exchange in ES Cells Mediated by FLP Recombinase: An Efficient Strategy for Repeated Modification of Tagged Loci by Marker-Free Constructs †. Biochemistry 37, 6229–6234.

Serrano, M., Lee, H., Chin, L., Cordon-Cardo, C., Beach, D., and DePinho, R.A. (1996). Role of the INK4a locus in tumor suppression and cell mortality. Cell 85, 27–37.

Serrano, M., Lin, A.W., McCurrach, M.E., Beach, D., and Lowe, S.W. (1997). Oncogenic ras provokes premature cell senescence associated with accumulation of p53 and p16INK4a. Cell 88, 593–602.

Shalem, O., Sanjana, N.E., Hartenian, E., Shi, X., Scott, D.A., Mikkelsen, T.S., Heckl, D., Ebert, B.L., Root, D.E., Doench, J.G., et al. (2014). Genome-scale CRISPR-Cas9 knockout screening in human cells. Science 343, 84–87.

Sharma, S.V., and Settleman, J. (2007). Oncogene addiction: setting the stage for molecularly targeted cancer therapy. Genes Dev 21, 3214–3231.

Sharpless, N.E., Bardeesy, N., Lee, K.-H., Carrasco, D., Castrillon, D.H., Aguirre, A.J., Wu, E.A., Horner, J.W., and DePinho, R.A. (2001). Loss of p16Ink4a with retention of p19Arf predisposes mice to tumorigenesis. Nature 413, 86–91.

Shaw, A.T., Kim, D.-W., Nakagawa, K., Seto, T., Crinó, L., Ahn, M.-J., De Pas, T., Besse, B., Solomon, B.J., Blackhall, F., et al. (2013). Crizotinib versus Chemotherapy in Advanced ALK-Positive Lung Cancer. N Engl J Med 368, 2385–2394.

Shaw, P., Tardy, S., Benito, E., Obrador, A., and Costa, J. (1991). Occurrence of Ki-ras and p53 mutations in primary colorectal tumors. Oncogene 6, 2121–2128.

Shechner, D.M., Hacisuleyman, E., Younger, S.T., and Rinn, J.L. (2015). Multiplexable, locus-specific targeting of long RNAs with CRISPR-Display. Nat Methods 12, 664–670.

Shi, J., Wang, E., Milazzo, J.P., Wang, Z., Kinney, J.B., and Vakoc, C.R. (2015). Discovery of cancer drug targets by CRISPR-Cas9 screening of protein domains. Nat Biotechnol 33, 661–667.

Shibata, T., Kokubu, A., Gotoh, M., Ojima, H., Ohta, T., Yamamoto, M., and Hirohashi, S. (2008a). Genetic alteration of Keap1 confers constitutive Nrf2 activation and resistance to chemotherapy in gallbladder cancer. Gastroenterology 135, 1358–1368–1368.e1–4.

Shibata, T., Ohta, T., Tong, K.I., Kokubu, A., Odogawa, R., Tsuta, K., Asamura, H., Yamamoto, M., and Hirohashi, S. (2008b). Cancer related mutations in NRF2 impair its recognition by Keap1-Cul3 E3 ligase and promote malignancy. Proceedings of the National Academy of Sciences 105, 13568–13573.

Sidransky, D., Eschenbach, Von, A., Tsai, Y.C., Jones, P., Summerhayes, I., Marshall, F., Paul, M., Green, P., Hamilton, S.R., Frost, P., et al. (1991). Identification of p53 gene mutations in bladder cancers and urine samples. Science 252, 706–709.

Singh, A., Misra, V., Thimmulappa, R.K., Lee, H., Ames, S., Hoque, M.O., Herman, J.G., Baylin, S.B., Sidransky, D., Gabrielson, E., et al. (2006a). Dysfunctional KEAP1–NRF2 Interaction in Non-Small-Cell Lung Cancer. PLoS Med 3, e420.

Singh, A., Rangasamy, T., Thimmulappa, R.K., Lee, H., Osburn, W.O., Brigelius-Flohe, R., Kensler, T.W., Yamamoto, M., and Biswal, S. (2006b). Glutathione peroxidase 2, the major cigarette smoke-inducible isoform of GPX in lungs, is regulated by Nrf2. Am J Respir Cell Mol Biol 35, 639–650.

Singh, M., Lima, A., Molina, R., Hamilton, P., Clermont, A.C., Devasthali, V., Thompson, J.D., Cheng, J.H., Bou Reslan, H., Ho, C.C.K., et al. (2010). Assessing therapeutic responses in Kras mutant cancers using genetically engineered mouse models. Nat Biotechnol 28, 585–593.

Sinha, D.K., Neveu, P., Gagey, N., Aujard, I., Benbrahim-Bouzidi, C., Le Saux, T., Rampon, C., Gauron, C., Goetz, B., Dubruille, S., et al. (2010a). Photocontrol of Protein Activity in Cultured Cells and Zebrafish with One- and Two-Photon Illumination. Chem. Eur. J. of Chem. Bio. 11, 653–663.

Sinha, D.K., Neveu, P., Gagey, N., Aujard, I., Le Saux, T., Rampon, C., Gauron, C., Kawakami, K., Leucht, C., Bally-Cuif, L., et al. (2010b). Photoactivation of the CreER T2 recombinase for conditional site-specific recombination with high spatiotemporal resolution. Zebrafish 7, 199–204.

Sinn, E., Muller, W., Pattengale, P., Tepler, I., Wallace, R., and Leder, P. (1987). Coexpression of MMTV/v-Ha-ras and MMTV/c-myc genes in transgenic mice: synergistic action of oncogenes in vivo. Cell 49, 465–475.

Skarnes, W.C., Rosen, B., West, A.P., Koutsourakis, M., Bushell, W., Iyer, V., Mujica, A.O., Thomas, M., Harrow, J., Cox, T., et al. (2011). A conditional knockout resource for the genome-wide study of mouse gene function. Nature 474, 337–342.

Smith, A.J.H., De Sousa, M.A., Kwabi-Addo, B., Heppell-Parton, A., Impey, H., and Rabbitts, P. (1995). A site–directed chromosomal translocation induced in embryonic stem cells by Cre-loxP recombination. Nat Genet 9, 376–385.

Smith, A.J.H., Xian, J., Richardson, M., Johnstone, K.A., and Rabbitts, P.H. (2002). Cre-loxP chromosome engineering of a targeted deletion in the mouse corresponding to the 3p21.3 region of homozygous loss in human tumours. Oncogene 21, 4521–4529.

Smith, J., Grizot, S., Arnould, S., Duclert, A., Epinat, J.-C., Chames, P., Prieto, J., Redondo, P., Blanco, F.J., Bravo, J., et al. (2006). A combinatorial approach to create artificial homing endonucleases cleaving chosen sequences. Nucleic Acids Res 34, e149–e149.

Smithies, O., Gregg, R.G., Boggs, S.S., Koralewski, M.A., and Kucherlapati, R.S. (1985). Insertion of DNA sequences into the human chromosomal β-globin locus by homologous recombination. Nature 317, 230–234.

Soda, M., Choi, Y.L., Enomoto, M., Takada, S., Yamashita, Y., Ishikawa, S., Fujiwara, S.-I., Watanabe, H., Kurashina, K., Hatanaka, H., et al. (2007). Identification of the transforming EML4–ALK fusion gene in non-small-cell lung cancer. Nature 448, 561–566.

Solimini, N.L., Xu, Q., Mermel, C.H., Liang, A.C., Schlabach, M.R., Luo, J., Burrows, A.E., Anselmo, A.N., Bredemeyer, A.L., Li, M.Z., et al. (2012). Recurrent Hemizygous Deletions in Cancers May Optimize Proliferative Potential. Science 337, 104–109.

Solis, L.M., Behrens, C., Dong, W., Suraokar, M., Ozburn, N.C., Moran, C.A., Corvalan, A.H., Biswal, S., Swisher, S.G., Bekele, B.N., et al. (2010). Nrf2 and Keap1 Abnormalities in Non–Small Cell Lung Carcinoma and Association with Clinicopathologic Features. Clinical Cancer Research 16, 3743–3753.

Song, H., Hollstein, M., and Xu, Y. (2007). p53 gain-of-function cancer mutants induce genetic instability by inactivating ATM. Nat Cell Biol 9, 573–580.

Sporn, M.B., and Liby, K.T. (2012). NRF2 and cancer: the good, the bad and the importance of context. Nat Rev Cancer 12, 564–571.

Stacey, D.W., Watson, T., Kung, H.F., and Curran, T. (1987). Microinjection of transforming ras protein induces c-fos expression. Mol Cell Biol 7, 523–527.

Staffas, A., Karlsson, C., Persson, M., Palmqvist, L., and Bergo, M.O. (2015). Wild-type KRAS inhibits oncogenic KRAS-induced T-ALL in mice. Leukemia 29, 1032–1040.

Stambolic, V., Suzuki, A., la Pompa, de, J.L., Brothers, G.M., Mirtsos, C., Sasaki, T., Ruland, J., Penninger, J.M., Siderovski, D.P., and Mak, T.W. (1998). Negative regulation of PKB/Akt-dependent cell survival by the tumor suppressor PTEN. Cell 95, 29–39.

Stegmeier, F., Hu, G., Rickles, R.J., Hannon, G.J., and Elledge, S.J. (2005). A lentiviral microRNA-based system for single-copy polymerase II-regulated RNA interference in mammalian cells. Proc Natl Acad Sci USA 102, 13212–13217.

Stehelin, D., Varmus, H.E., Bishop, J.M., and Vogt, P.K. (1976). DNA related to the transforming gene(s) of avian sarcoma viruses is present in normal avian DNA. Nature 260, 170–173.

Sternberg, N., and Hamilton, D. (1981). Bacteriophage P1 site-specific recombination. I. Recombination between loxP sites. J Mol Biol 150, 467–486.

Sternberg, N., Hamilton, D., and Hoess, R. (1981a). Bacteriophage P1 site-specific recombination. II. Recombination between loxP and the bacterial chromosome. J Mol Biol 150, 487–507.

Sternberg, N., Hamilton, D., Austin, S., Yarmolinsky, M., and Hoess, R. (1981b). Site-specific Recombination and Its Role in the Life Cycle of Bacteriophage P1. Cold Spring Harb Symp Quant Biol 45, 297–309.

Stewart, T.A., Pattengale, P.K., and Leder, P. (1984). Spontaneous mammary adenocarcinomas in transgenic mice that carry and express MTV/myc fusion genes. Cell 38, 627–637.

Strasser, A., Harris, A.W., Bath, M.L., and Cory, S. (1990). Novel primitive lymphoid tumours induced in transgenic mice by cooperation between myc and bcl-2. Nature 348, 331–333.

Stratton, M.R., Campbell, P.J., and Futreal, P.A. (2009). The cancer genome. Nature 458, 719–724.

Stretch, J.R., Gatter, K.C., Ralfkiaer, E., Lane, D.P., and Harris, A.L. (1991). Expression of mutant p53 in melanoma. Cancer Res 51, 5976–5979.

Sun, S., Schiller, J.H., and Gazdar, A.F. (2007). Lung cancer in never smokers--a different disease. Nat Rev Cancer 7, 778–790.

Sun, T., Aceto, N., Meerbrey, K.L., Kessler, J.D., Zhou, C., Migliaccio, I., Nguyen, D.X., Pavlova, N.N., Botero, M., Huang, J., et al. (2011). Activation of Multiple Proto-oncogenic Tyrosine Kinases in Breast Cancer via Loss of the PTPN12 Phosphatase. Cell 144, 703–718.

Sutherland, K.D., Song, J.-Y., Kwon, M.-C., Proost, N., Zevenhoven, J., and Berns, A. (2014). Multiple cells-of-origin of mutant K-Ras-induced mouse lung adenocarcinoma. Proc Natl Acad Sci USA 111, 4952–4957.

Suzuki, T., and Yamamoto, M. (2015). Molecular basis of the Keap1–Nrf2 system. Free Radic Biol Med 88, Part B IS -, 93–100.

Szostak, J.W., Orr-Weaver, T.L., Rothstein, R.J., and Stahl, F.W. (1983). The double-strand-break repair model for recombination. Cell 33, 25–35.

Tabin, C.J., Bradley, S.M., Bargmann, C.I., Weinberg, R.A., Papageorge, A.G., Scolnick, E.M., Dhar, R., Lowy, D.R., and Chang, E.H. (1982). Mechanism of activation of a human oncogene. Nature 300, 143–149.

Taguchi, K., Motohashi, H., and Yamamoto, M. (2011). Molecular mechanisms of the Keap1–Nrf2 pathway in stress response and cancer evolution. Genes Cells 16, 123–140.

Takahashi, T., Nau, M.M., Chiba, I., Birrer, M.J., Rosenberg, R.K., Vinocour, M., Levitt, M., Pass, H., Gazdar, A.F., and Minna, J.D. (1989). p53: a frequent target for genetic abnormalities in lung cancer. Science 246, 491–494.

Tanaka, H., Matsumura, I., Ezoe, S., Satoh, Y., Sakamaki, T., Albanese, C., Machii, T., Pestell, R.G., and Kanakura, Y. (2002). E2F1 and c-Myc potentiate apoptosis through inhibition of NF-kappaB activity that facilitates MnSOD-mediated ROS elimination. Mol Cell 9, 1017–1029.

Tanase, A.-M., Marchio, A., Dumitrascu, T., Dima, S., Herlea, V., Oprisan, G., Dejean, A., Popescu, I., and Pineau, P. (2015). Mutation spectrum of hepatocellular carcinoma from eastern-European patients betrays the impact of a complex exposome. J Expo Sci Environ Epidemiol 25, 256–263.

Tanenbaum, M.E., Gilbert, L.A., Qi, L.S., Weissman, J.S., and Vale, R.D. (2014). A protein-tagging system for signal amplification in gene expression and fluorescence imaging. Cell 159, 635–646.

Taniguchi, H., Yamato, E., Tashiro, F., Ikegami, H., Ogihara, T., and Miyazaki, J. (2003). beta-cell neogenesis induced by adenovirus-mediated gene delivery of transcription factor pdx-1 into mouse pancreas. Gene Therapy 10, 15–23.

Tanito, M., Masutani, H., Kim, Y.-C., Nishikawa, M., Ohira, A., and Yodoi, J. (2005). Sulforaphane induces thioredoxin through the antioxidant-responsive element and attenuates retinal light damage in mice. Invest Ophthalmol Vis Sci 46, 979–987.

Taparowsky, E., Suard, Y., Fasano, O., Shimizu, K., Goldfarb, M., and Wigler, M. (1982). Activation of the T24 bladder carcinoma transforming gene is linked to a single amino acid change. Nature 300, 762–765.

Terzian, T., Suh, Y.-A., Iwakuma, T., Post, S.M., Neumann, M., Lang, G.A., Van Pelt, C.S., and Lozano, G. (2008). The inherent instability of mutant p53 is alleviated by Mdm2 or p16INK4a loss. Genes Dev 22, 1337–1344.

Tesson, L., Usal, C., Ménoret, S., Leung, E., Niles, B.J., Remy, S., Santiago, Y., Vincent, A.I., Meng, X., Zhang, L., et al. (2011). Knockout rats generated by embryo microinjection of TALENs. Nat Biotechnol 29, 695–696.

The Cancer Genome Atlas Research Network (2015). Comprehensive, Integrative Genomic Analysis of Diffuse Lower-Grade Gliomas. N Engl J Med 372, 2481–2498.

Thimmulappa, R.K., Mai, K.H., Srisuma, S., Kensler, T.W., Yamamoto, M., and Biswal, S. (2002). Identification of Nrf2-regulated Genes Induced by the Chemopreventive Agent Sulforaphane by Oligonucleotide Microarray. Cancer Research 62, 5196–5203.

Thomas, K.R., and Capecchi, M.R. (1987). Site-directed mutagenesis by gene targeting in mouse embryo-derived stem cells. Cell 51, 503–512.

Thomas, K.R., Folger, K.R., and Capecchi, M.R. (1986). High frequency targeting of genes to specific sites in the mammalian genome. Cell 44, 419–428.

Thorpe, L.M., Yuzugullu, H., and Zhao, J.J. (2015). PI3K in cancer: divergent roles of isoforms, modes of activation and therapeutic targeting. Nat Rev Cancer 15, 7–24.

Threadgill, D.W., Dlugosz, A.A., Hansen, L.A., Tennenbaum, T., Lichti, U., Yee, D., LaMantia, C., Mourton, T., Herrup, K., and Harris, R.C. (1995). Targeted disruption of mouse EGF receptor: effect of genetic background on mutant phenotype. Science 269, 230–234.

Tiscornia, G., Singer, O., Ikawa, M., and Verma, I.M. (2003). A general method for gene knockdown in mice by using lentiviral vectors expressing small interfering RNA. Proc Natl Acad Sci USA 100, 1844–1848.

Toguchida, J., Yamaguchi, T., Dayton, S.H., Beauchamp, R.L., Herrera, G.E., Ishizaki, K., Yamamuro, T., Meyers, P.A., Little, J.B., Sasaki, M.S., et al. (1992a). Prevalence and spectrum of germline mutations of the p53 gene among patients with sarcoma. New England Journal of Medicine 326, 1301–1308.

Toguchida, J., Yamaguchi, T., Ritchie, B., Beauchamp, R.L., Dayton, S.H., Herrera, G.E., Yamamuro, T., Kotoura, Y., Sasaki, M.S., Little, J.B., et al. (1992b). Mutation spectrum of the p53 gene in bone and soft tissue sarcomas. Cancer Research 52, 6194–6199.

Tong, K.I., Katoh, Y., Kusunoki, H., Itoh, K., Tanaka, T., and Yamamoto, M. (2006). Keap1 Recruits Neh2 through Binding to ETGE and DLG Motifs: Characterization of the Two-Site Molecular Recognition Model. Mol Cell Biol 26, 2887–2900.

Torres, R., Martin, M.C., Garcia, A., Cigudosa, J.C., Ramirez, J.C., and Rodriguez-Perales, S. (2014). Engineering human tumour-associated chromosomal translocations with the RNA-guided CRISPR–Cas9 system. Nature Communications 5.

Trachootham, D., Alexandre, J., and Huang, P. (2009). Targeting cancer cells by ROS-mediated mechanisms: a radical therapeutic approach? Nature Reviews Drug Discovery 8, 579–591.

Tschaharganeh, D.F., Xue, W., Calvisi, D.F., Evert, M., Michurina, T.V., Dow, L.E., Banito, A., Katz, S.F., Kastenhuber, E.R., Weissmueller, S., et al. (2014). p53-dependent Nestin regulation links tumor suppression to cellular plasticity in liver cancer. Cell 158, 579–592.

Tsuji, Y., Ayaki, H., Whitman, S.P., Morrow, C.S., Torti, S.V., and Torti, F.M. (2000). Coordinate transcriptional and translational regulation of ferritin in response to oxidative stress. Mol Cell Biol 20, 5818–5827.

Tucker, C.L. (2012). Manipulating cellular processes using optical control of protein-protein interactions. Prog. Brain Res. 196, 95–117.

Turan, S., Qiao, J., Madden, S., Benham, C., Kotz, M., Schambach, A., and Bode, J. (2014). Expanding Flp-RMCE options: the potential of Recombinase Mediated Twin-Site Targeting (RMTT). Gene 546, 135–144.

Turan, S., Zehe, C., Kuehle, J., Qiao, J., and Bode, J. (2013). Recombinase-mediated cassette exchange (RMCE) - a rapidly-expanding toolbox for targeted genomic modifications. Gene 515, 1–27.

Urich, M., Senften, M., Shaw, P.E., and Ballmer-Hofer, K. (1997). A role for the small GTPase Rac in polyomavirus middle-T antigen-mediated activation of the serum response element and in cell transformation. Oncogene 14, 1235–1241.

Urnov, F.D., Miller, J.C., Lee, Y.-L., Beausejour, C.M., Rock, J.M., Augustus, S., Jamieson, A.C., Porteus, M.H., Gregory, P.D., and Holmes, M.C. (2005). Highly efficient endogenous human gene correction using designed zinc-finger nucleases. Nat Cell Biol 435, 646–651.

Vafa, O., Wade, M., Kern, S., Beeche, M., Pandita, T.K., Hampton, G.M., and Wahl, G.M. (2002). c-Myc can induce DNA damage, increase reactive oxygen species, and mitigate p53 function: a mechanism for oncogene-induced genetic instability. Mol Cell 9, 1031–1044.

Van den Haute, C., Eggermont, K., Nuttin, B., Debyser, Z., and Baekelandt, V. (2003). Lentiviral vector-mediated delivery of short hairpin RNA results in persistent knockdown of gene expression in mouse brain. Hum Gene Ther 14, 1799–1807.

Van Deursen, J., Fornerod, M., Van Rees, B., and Grosveld, G. (1995). Cre-mediated site-specific translocation between nonhomologous mouse chromosomes. Proceedings of the National Academy of Sciences 92, 7376–7380.

Van Dyke, T., and Jacks, T. (2002). Cancer modeling in the modern era: progress and challenges. Cell 108, 135–144.

Varley, J.M., Brammar, W.J., Lane, D.P., Swallow, J.E., Dolan, C., and Walker, R.A. (1991). Loss of chromosome 17p13 sequences and mutation of p53 in human breast carcinomas. Oncogene 6, 413–421.

Vaux, D.L., Cory, S., and Adams, J.M. (1988). Bcl-2 gene promotes haemopoietic cell survival and cooperates with c-myc to immortalize pre-B cells. Nature 335, 440–442.

Venter, J.C., Adams, M.D., Myers, E.W., Li, P.W., Mural, R.J., Sutton, G.G., Smith, H.O., Yandell, M., Evans, C.A., Holt, R.A., et al. (2001). The sequence of the human genome. Science 291, 1304–1351.

Ventura, A., Kirsch, D., McLaughlin, M., Tuveson, D., Grimm, J., Lintault, L., Newman, J., Reczek, E., Weissleder, R., and Jacks, T. (2007). Restoration of p53 function leads to tumour regression in vivo. Nature 445, 661–665.

Ventura, A., Meissner, A., Dillon, C.P., McManus, M., Sharp, P.A., Van Parijs, L., Jaenisch, R., and Jacks, T. (2004). Cre-lox-regulated conditional RNA interference from transgenes. Proc Natl Acad Sci USA 101, 10380–10385.

Venugopal, R., and Jaiswal, A.K. (1996). Nrf1 and Nrf2 positively and c-Fos and Fra1 negatively regulate the human antioxidant response element-mediated expression of NAD(P)H:quinone oxidoreductase1 gene. Proc Natl Acad Sci USA 93, 14960–14965.

Vogelstein, B., Lane, D., and Levine, A.J. (2000). Surfing the p53 network. Nature 408, 307–310.

Vogelstein, B., Papadopoulos, N., Velculescu, V.E., Zhou, S., Diaz, L.A., and Kinzler, K.W. (2013). Cancer genome landscapes. Science 339, 1546–1558.

Waddell, N., Pajic, M., Patch, A.-M., Chang, D.K., Kassahn, K.S., Bailey, P., Johns, A.L., Miller, D., Nones, K., Quek, K., et al. (2015). Whole genomes redefine the mutational landscape of pancreatic cancer. Nature 518, 495–501.

Wang, F., Travins, J., DeLaBarre, B., Penard-Lacronique, V., Schalm, S., Hansen, E., Straley, K., Kernytsky, A., Liu, W., Gliser, C., et al. (2013a). Targeted inhibition of mutant IDH2 in leukemia cells induces cellular differentiation. Science 340, 622–626.

Wang, H., Yang, H., Shivalila, C.S., Dawlaty, M.M., Cheng, A.W., Zhang, F., and Jaenisch, R. (2013b). One-Step Generation of Mice Carrying Mutations in Multiple Genes by CRISPR/Cas-Mediated Genome Engineering. Cell 153, 910–918.

Wang, H., Liu, K., Geng, M., Gao, P., Wu, X., Hai, Y., Li, Y., Li, Y., Luo, L., Hayes, J.D., et al. (2013c). RXRα Inhibits the NRF2-ARE Signaling Pathway through a Direct Interaction with the Neh7 Domain of NRF2. Cancer Research 73, 3097–3108.

Wang, R., An, J., Ji, F., Jiao, H., Sun, H., and Zhou, D. (2008). Hypermethylation of the Keap1 gene in human lung cancer cell lines and lung cancer tissues. Biochem Biophys Res Commun 373, 151–154.

Wang, T., Birsoy, K., Hughes, N.W., Krupczak, K.M., Post, Y., Wei, J.J., Lander, E.S., and Sabatini, D.M. (2015). Identification and characterization of essential genes in the human genome. Science 350, 1096–1101.

Wang, T., Wei, J.J., Sabatini, D.M., and Lander, E.S. (2014). Genetic screens in human cells using the CRISPR-Cas9 system. Science 343, 80–84.

Wang, X., Chen, X., and Yang, Y. (2012). Spatiotemporal control of gene expression by a light-switchable transgene system. Nat Methods 9, 266–269.

Warburg, O. (1956). On the Origin of Cancer Cells. Science 123, 309–314.

Warburg, O., Wind, F., and Negelein, E. (1924). Über den Stoffwechsel von Tumoren im Körper. Klin Wochenschr 5, 829–832.

Ward, P.S., Patel, J., Wise, D.R., Abdel-Wahab, O., Bennett, B.D., Coller, H.A., Cross, J.R., Fantin, V.R., Hedvat, C.V., Perl, A.E., et al. (2010). The common feature of leukemia-associated IDH1 and IDH2 mutations is a neomorphic enzyme activity converting alpha-ketoglutarate to 2-hydroxyglutarate. Cancer Cell 17, 225–234.

Weinstein, I.B. (2002). Cancer. Addiction to oncogenes--the Achilles heal of cancer. Science 297, 63–64.

Weir, B.A., Woo, M.S., Getz, G., Perner, S., Ding, L., Beroukhim, R., Lin, W.M., Province, M.A., Kraja, A., Johnson, L.A., et al. (2007). Characterizing the cancer genome in lung adenocarcinoma. Nature 450, 893–898.

Weissmueller, S., Manchado, E., Saborowski, M., Morris, J.P.4., Wagenblast, E., Davis, C.A., Moon, S.-H., Pfister, N.T., Tschaharganeh, D.F., Kitzing, T., et al. (2014). Mutant p53 drives pancreatic cancer metastasis through cell-autonomous PDGF receptor beta signaling. Cell 157, 382–394.

Westbrook, T.F., Martin, E.S., Schlabach, M.R., Leng, Y., Liang, A.C., Feng, B., Zhao, J.J., Roberts, T.M., Mandel, G., Hannon, G.J., et al. (2005). A genetic screen for candidate tumor suppressors identifies REST. Cell 121, 837–848.

Westwick, J.K., Cox, A.D., Der, C.J., Cobb, M.H., Hibi, M., Karin, M., and Brenner, D.A. (1994). Oncogenic Ras activates c-Jun via a separate pathway from the activation of extracellular signal-regulated kinases. Proc Natl Acad Sci USA 91, 6030–6034.

Wheeler, D.A., and Wang, L. (2013). From human genome to cancer genome: The first decade. Genome Research 23, 1054–1062.

Whitworth, K.M., Lee, K., Benne, J.A., Beaton, B.P., Spate, L.D., Murphy, S.L., Samuel, M.S., Mao, J., O'Gorman, C., Walters, E.M., et al. (2014). Use of the CRISPR/Cas9 System to Produce Genetically Engineered Pigs from In Vitro-Derived Oocytes and Embryos. Biology of Reproduction 91, 78–78.

Wild, A.C., Moinova, H.R., and Mulcahy, R.T. (1999). Regulation of γ-Glutamylcysteine Synthetase Subunit Gene Expression by the Transcription Factor Nrf2. Journal of Biological Chemistry 274, 33627–33636.

Winslow, M.M., Dayton, T.L., Verhaak, R.G.W., Kim-Kiselak, C., Snyder, E.L., Feldser, D.M., Hubbard, D.D., DuPage, M.J., Whittaker, C.A., Hoersch, S., et al. (2011). Suppression of lung adenocarcinoma progression by Nkx2-1. Nature 473, 101–104.

Wolfe, S.A., Greisman, H.A., Ramm, E.I., and Pabo, C.O. (1999). Analysis of zinc fingers optimized via phage display: evaluating the utility of a recognition code. J Mol Biol 285, 1917–1934.

Wong, J.C., Weinfurtner, K.M., Alzamora, M.D.P., Kogan, S.C., Burgess, M.R., Zhang, Y., Nakitandwe, J., Ma, J., Cheng, J., Chen, S.-C., et al. (2015). Functional evidence implicating chromosome 7q22 haploinsufficiency in myelodysplastic syndrome pathogenesis. eLife Sciences 4, e07839.

Wong, S., and Witte, O.N. (2001). Modeling Philadelphia chromosome positive leukemias. Oncogene 20, 5644–5659.

Wood, A.J., Lo, T.-W., Zeitler, B., Pickle, C.S., Ralston, E.J., Lee, A.H., Amora, R., Miller, J.C., Leung, E., Meng, X., et al. (2011). Targeted genome editing across species using ZFNs and TALENs. Science 333, 307.

Wu, K.C., Cui, J.Y., and Klaassen, C.D. (2011). Beneficial Role of Nrf2 in Regulating NADPH Generation and Consumption. Toxicological Sciences 123, 590–600.

Wu, X., and Pandolfi, P.P. (2001). Mouse models for multistep tumorigenesis. Trends Cell Biol 11, S2–S9.

Wuestefeld, T., Pesic, M., Rudalska, R., Dauch, D., Longerich, T., Kang, T.-W., Yevsa, T., Heinzmann, F., Hoenicke, L., Hohmeyer, A., et al. (2013). A Direct In Vivo RNAi Screen Identifies MKK4 as a Key Regulator of Liver Regeneration. Cell 153, 389–401.

Xia, Y., Yeddula, N., Leblanc, M., Ke, E., Zhang, Y., Oldfield, E., Shaw, R.J., and Verma, I.M. (2012). Reduced cell proliferation by IKK2 depletion in a mouse lung-cancer model. Nat Cell Biol 14, 257–265.

Xiao, A., Wang, Z., Hu, Y., Wu, Y., Luo, Z., Yang, Z., Zu, Y., Li, W., Huang, P., Tong, X., et al. (2013). Chromosomal deletions and inversions mediated by TALENs and CRISPR/Cas in zebrafish. Nucleic Acids Res 41, e141.

Xu, W., Yang, H., Liu, Y., Yang, Y., Wang, P., Kim, S.-H., Ito, S., Yang, C., Wang, P., Xiao, M.-T., et al. (2011). Oncometabolite 2-hydroxyglutarate is a competitive inhibitor of α-ketoglutarate-dependent dioxygenases. Cancer Cell 19, 17–30.

Xu, Y., Ashley, T., Brainerd, E.E., Bronson, R.T., Meyn, M.S., and Baltimore, D. (1996). Targeted disruption of ATM leads to growth retardation, chromosomal fragmentation during meiosis, immune defects, and thymic lymphoma. Genes Dev 10, 2411–2422.

Xue, W., Meylan, E., Oliver, T.G., Feldser, D.M., Winslow, M.M., Bronson, R., and Jacks, T. (2011). Response and Resistance to NF- B Inhibitors in Mouse Models of Lung Adenocarcinoma. Cancer Discovery.

Xue, W., Chen, S., Yin, H., Tammela, T., Papagiannakopoulos, T., Joshi, N.S., Cai, W., Yang, G., Bronson, R., Crowley, D.G., et al. (2014). CRISPR-mediated direct mutation of cancer genes in the mouse liver. Nature 514, 380–384.

Xue, W., Zender, L., Miething, C., Dickins, R.A., Hernando, E., Krizhanovsky, V., Cordon-Cardo, C., and Lowe, S.W. (2007). Senescence and tumour clearance is triggered by p53 restoration in murine liver carcinomas. Nature 445, 656–660.

Yan, H., Parsons, D.W., Jin, G., McLendon, R., Rasheed, B.A., Yuan, W., Kos, I., Batinic-Haberle, I., Jones, S., Riggins, G.J., et al. (2009). IDH1and IDH2Mutations in Gliomas. N Engl J Med 360, 765–773.

Yang, H., Wang, H., Shivalila, C.S., Cheng, A.W., Shi, L., and Jaenisch, R. (2013). One-Step Generation of Mice Carrying Reporter and Conditional Alleles by CRISPR/Cas-Mediated Genome Engineering. Cell 154, 1370–1379.

Yeddula, N., Xia, Y., Ke, E., Beumer, J., and Verma, I.M. (2015). Screening for tumor suppressors: Loss of ephrin receptor A2 cooperates with oncogenic KRas in promoting lung adenocarcinoma. Proceedings of the National Academy of Sciences.

Yeh, E.S., Vernon-Grey, A., Martin, H., and Chodosh, L.A. (2014). Tetracycline-regulated mouse models of cancer. Cold Spring Harb Protoc 2014, pdb.top069823.

Yekta, S., Shih, I.-H., and Bartel, D.P. (2004). MicroRNA-Directed Cleavage of HOXB8 mRNA. Science 304, 594–596.

Yi, R., Qin, Y., Macara, I.G., and Cullen, B.R. (2003). Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs. Genes Dev 17, 3011–3016.

Yin, H., Kanasty, R.L., Eltoukhy, A.A., Vegas, A.J., Dorkin, J.R., and Anderson, D.G. (2014a). Non-viral vectors for gene-based therapy. Nat Rev Genet 15, 541–555.

Yin, H., Xue, W., Chen, S., Bogorad, R.L., Benedetti, E., Grompe, M., Koteliansky, V., Sharp, P.A., Jacks, T., and Anderson, D.G. (2014b). Genome editing with Cas9 in adult mice corrects a disease mutation and phenotype. Nat Biotechnol 32, 551–553.

Zalatan, J.G., Lee, M.E., Almeida, R., Gilbert, L.A., Whitehead, E.H., La Russa, M., Tsai, J.C., Weissman, J.S., Dueber, J.E., Qi, L.S., et al. (2015). Engineering complex synthetic transcriptional programs with CRISPR RNA scaffolds. Cell 160, 339–350.

Zambrowicz, B.P., Imamoto, A., Fiering, S., Herzenberg, L.A., Kerr, W.G., and Soriano, P. (1997). Disruption of overlapping transcripts in the ROSA beta geo 26 gene trap strain leads to widespread expression of beta-galactosidase in mouse embryos and hematopoietic cells. Proc Natl Acad Sci USA 94, 3789–3794.

Zender, L., Xue, W., Zuber, J., Semighini, C., Alex, Krasnitz, E., Ma, B., Zender, P., Kubicka, S., Luk, J., et al. (2008). An Oncogenomics-Based In Vivo RNAi Screen Identifies Tumor Suppressors in Liver Cancer. Cell 135, 852–864.

Zeng, Y., Cai, X., and Cullen, B.R. (2005). Use of RNA polymerase II to transcribe artificial microRNAs. Meth Enzymol 392, 371–380.

Zeng, Y., Wagner, E.J., and Cullen, B.R. (2002). Both natural and designed micro RNAs can inhibit the expression of cognate mRNAs when expressed in human cells. Mol Cell 9, 1327–1333.

Zetsche, B., Gootenberg, J.S., Abudayyeh, O.O., Slaymaker, I.M., Makarova, K.S., Essletzbichler, P., Volz, S.E., Joung, J., van der Oost, J., Regev, A., et al. (2015). Cpf1 Is a Single RNA-Guided Endonuclease of a Class 2 CRISPR-Cas System. Cell 163, 759–771.

Zetsche, B., Volz, S.E., and Zhang, F. (2015b). A split-Cas9 architecture for inducible genome editing and transcription modulation. Nat Biotechnol 33, 139–142.

Zhang, D.D., Lo, S.-C., Cross, J.V., Templeton, D.J., and Hannink, M. (2004). Keap1 Is a Redox-Regulated Substrate Adaptor Protein for a Cul3-Dependent Ubiquitin Ligase Complex. Mol Cell Biol 24, 10941–10953.

Zhang, F., Cong, L., Lodato, S., Kosuri, S., Church, G.M., and Arlotta, P. (2011). Efficient construction of sequence-specific TAL effectors for modulating mammalian transcription. Nat Biotechnol 29, 149–153.

Zhang, P., Singh, A., Yegnasubramanian, S., Esopi, D., Kombairaju, P., Bodas, M., Wu, H., Bova, S.G., and Biswal, S. (2010). Loss of Kelch-like ECH-associated protein 1 function in prostate cancer cells causes chemoresistance and radioresistance and promotes tumor growth. Mol. Cancer Ther. 9, 336–346.

Zhang, Y., Rajan, R., Seifert, H.S., Mondragón, A., and Sontheimer, E.J. (2015). DNase H Activity of Neisseria meningitidis Cas9. Mol Cell 60, 242–255.

Zhang, Z., Wang, Y., Vikis, H.G., Johnson, L., Liu, G., Li, J., Anderson, M.W., Sills, R.C., Hong, H.L., Devereux, T.R., et al. (2001). Wildtype Kras2 can inhibit lung carcinogenesis in mice. Nat Genet 29, 25–33.

Zhou, R., Yazdi, A.S., Menu, P., and Tschopp, J. (2011). A role for mitochondria in NLRP3 inflammasome activation. Nature 469, 221–225.

Zhou, Y., Zhu, S., Cai, C., Yuan, P., Li, C., Huang, Y., and Wei, W. (2014). High-throughput screening of a CRISPR/Cas9 library for functional genomics in human cells. Nature 509, 487–491.

Zhu, J., Sammons, M.A., Donahue, G., Dou, Z., Vedadi, M., Getlik, M., Barsyte-Lovejoy, D., Al-awar, R., Katona, B.W., Shilatifard, A., et al. (2015). Gain-of-function p53 mutants co-opt chromatin pathways to drive cancer growth. Nature 525, 206–211.

Zhu, M., and Fahl, W.E. (2001). Functional characterization of transcription regulators that interact with the electrophile response element. Biochem Biophys Res Commun 289, 212–219.

Zipper, L.M., and Mulcahy, R.T. (2002). The Keap1 BTB/POZ Dimerization Function Is Required to Sequester Nrf2 in Cytoplasm. Journal of Biological Chemistry 277, 36544–36552.

Zou, J., Maeder, M.L., Mali, P., Pruett-Miller, S.M., Thibodeau-Beganny, S., Chou, B.-K., Chen, G., Ye, Z., Park, I.H., Daley, G.Q., et al. (2009). Gene targeting of a disease-related gene in human induced pluripotent stem and embryonic stem cells. Cell Stem Cell 5, 97–110.

Zuber, J., Rappaport, A.R., Luo, W., Wang, E., Chen, C., Vaseva, A.V., Shi, J., Weissmueller, S., Fellman, C., Taylor, M.J., et al. (2011a). An integrated approach to dissecting oncogene addiction implicates a Myb-coordinated self-renewal program as essential for leukemia maintenance. Genes Dev 25, 1628–1640.

Zuber, J., Shi, J., Wang, E., Rappaport, A.R., Herrmann, H., Sison, E.A., Magoon, D., Qi, J., Blatt, K., Wunderlich, M., et al. (2011b). RNAi screen identifies Brd4 as a therapeutic target in acute myeloid leukaemia. Nature 478, 524–528.

Zuris, J.A., Thompson, D.B., Shu, Y., Guilinger, J.P., Bessen, J.L., Hu, J.H., Maeder, M.L., Joung, J.K., Chen, Z.-Y., and Liu, D.R. (2015). Cationic lipid-mediated delivery of proteins enables efficient protein-based genome editing in vitro and in vivo. Nat Biotechnol 33, 73–80.

Chapter 2

Rapid modeling of cooperating genetic events in cancer through somatic genome editing

Francisco J. Sánchez-Rivera[1,2*], Thales Papagiannakopoulos[1*], Rodrigo Romero[1,2], Tuomas Tammela[1], Matthew R. Bauer[1], Arjun Bhutkar[1], Nikhil S. Joshi[1], Lakshmipriya Subbaraj[1], Roderick T. Bronson[4], Wen Xue[1] and Tyler Jacks[1,2,3].

[1] David H. Koch Institute for Integrative Cancer Research, Massachusetts Institute of Technology, Cambridge, MA 02142
[2] Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02142
[3] Howard Hughes Medical Institute, Massachusetts Institute of Technology, Cambridge, MA 02139
[4] Tufts University and Harvard Medical School, Boston, MA 02115
[*] These authors contributed equally to this work.

This chapter is an extended version of a manuscript published in final form in:
Nature ***516,*** *428–431 (2014).*

Author Contributions

F.J.S.R., T.P. and T.J. designed the study; F.J.S.R, T.P., R.R., M.R.B. and L.S. performed experiments; T.T. generated *Kras*$^{LSL\text{-}G12D/+}$*; Apc*$^{flox/flox}$ data; A.B. conducted bioinformatic analyses; N.S.J. generated Green-Go cells; R.T.B. provided pathology assistance; W.X. gave conceptual advice; F.J.S.R, T.P. and T.J. wrote the manuscript with comments from all authors.

ABSTRACT

Cancer is thought to arise through the sequential acquisition of both genetic and epigenetic alterations in oncogenes and tumor suppressor genes. Large-scale efforts to re-sequence protein-coding genes from human cancer cell lines and biopsy samples have begun to catalog the spectrum of mutations existing in human tumors. One major limitation of these studies has been the inability to rapidly and systematically determine which of these mutations are causally related to tumorigenesis, particularly in the context of *in vivo* models of the disease. Although existing genetically engineered mouse models have led to critical insights into the initiation and progression of human cancer, their use for rapid functional characterization of new cancer genes is currently lacking, partly due to the cost and time required to generate appropriate murine models. Here, we describe a novel CRISPR-Cas9-based approach for rapid functional investigation of candidate genes in well-established autochthonous mouse models of cancer. Using a $Kras^{G12D}$-driven lung cancer model, we performed functional characterization of a panel of tumor suppressor genes with known loss of function alterations in human lung cancer. Cre-dependent somatic activation of oncogenic $Kras^{G12D}$ combined with CRISPR-Cas9-mediated genome editing of tumor suppressor genes resulted in lung adenocarcinomas with distinct histopathological and molecular features. This rapid somatic genome engineering approach enables functional characterization of putative cancer genes in the lung and other tissues using autochthonous mouse models. We anticipate that this approach can be used to systematically dissect the complex catalog of mutations identified in cancer genome sequencing studies.

INTRODUCTION

Large-scale multi-institutional cancer genome sequencing studies have revealed a multitude of recurrent mutations and copy number alterations in several types of human cancer, including colorectal cancer (CRC TCGA 2012), breast cancer (Breast TCGA 2012), pancreatic cancer (Waddell et al., 2015), gastric adenocarcinoma (Gastric TCGA 2014), ovarian carcinoma (Patch et al., 2015), prostate adenocarcinoma (Abeshouse et al., 2015, Robinson et al., 2015), melanoma (Akbani et al., 2015), glioma (Glioma TCGA 2015), squamous cell carcinoma of the lung (Hammerman et al., 2012), small cell lung cancer (George et al., 2015) and lung adenocarcinoma (LUAD TCGA 2014), as well as several types of hematological cancers, including acute myeloid leukemia (AML TCGA 2013), acute lymphoblastic leukemia (Andersson et al., 2015), multiple myeloma (Lohr et al., 2014) and diffuse large B-cell lymphoma (Pasqualucci et al., 2011), among many others. However, the determination of which mutations are causally related to tumorigenesis (particularly in the context of *in vivo* models of the disease) remains a major challenge. This challenge is even greater for lung cancer due to the fact that lung tumors (as is also the case for melanomas (Akbani et al., 2015)) have substantially high mutation frequencies, which represents a major obstacle when trying to distinguish *bona fide* driver events from mere passenger events.

Genetically engineered mouse models (GEMMs) of lung cancer have greatly assisted in the functional characterization of putative driver events identified in human lung tumors (reviewed in Frese and Tuveson 2007, Kwon and Berns 2013, and Farago et al., 2012). However, such approaches typically require modification of the germline, which is both expensive and time consuming. Therefore, these methods cannot be performed in a

highly parallel manner. The Jacks laboratory has established GEMMs of human lung adenoma and adenocarcinoma that faithfully mimic human lung tumors in their progression (Feldser et al., 2010, Jackson et al., 2005, Jackson et al., 2001, Sweet-Cordero et al., 2005, Sweet-Cordero et al., 2006, Winslow et al., 2011), showing similarities both at the molecular and histopathological levels. In these models, lung tumors are induced in *Kras*$^{LSL\text{-}G12D/+}$ (K) (Jackson et al., 2001) or *Kras*$^{LSL\text{-}G12D/+}$*; p53*$^{flox/flox}$ (KP) (Jackson et al., 2005) mice after intratracheal administration of viral vectors expressing Cre-recombinase (DuPage et al., 2009), which activates a oncogenic *Kras*^{G12D} allele with or without concomitant deletion of the tumor suppressor *p53* in the murine lung epithelium.

The combination of emerging CRISPR-Cas9 approaches with faithful GEMMs of human cancer has the potential to fulfill the immediate need for the rapid and efficient systematic interrogation of the complex catalog of mutations obtained from large-scale cancer genome sequencing studies (reviewed in Sánchez-Rivera and Jacks 2015). Towards that goal, we developed and applied a novel CRISPR-Cas9-based lentiviral platform for rapidly interrogating putative cancer genes using pre-clinical GEMMs of lung adenocarcinoma. Our data demonstrate the feasibility of this approach and suggests that it can be readily adapted to many existing Cre/loxP-based GEMMs of several cancer types to facilitate the rapid functional assessment of new hypotheses generated by cancer genome studies.

RESULTS

Design, construction and characterization of the pSECC lentiviral platform

Recent work from the Jacks laboratory has demonstrated the feasibility of using the CRISPR-Cas9 system to directly mutate cancer genes in the liver following hydrodynamic delivery of plasmids carrying the CRISPR components (Xue et al., 2014), which relies on the efficient transfection of hepatocytes. To rapidly interrogate cancer genes in the lung and other tissues, we developed pSECC **(Figure 1)**, a lentiviral-based system that delivers both the CRISPR system and Cre recombinase. In this setting, CRISPR-induced mutation of genes can be examined in the context of several of the well-studied conditional Cre/loxP mouse models of lung cancer and other cancer types.

To validate pSECC, we developed the Green-Go (GG) reporter cell line, which expresses GFP following exposure to Cre (**Figure 2a**). To assess the efficiency of Cas9 in tumors *in vivo,* we targeted a Cre-activatable tdTomato knock-in reporter allele (Madisen et al., 2010) with pSECC lentiviruses expressing an sgRNA against tdTomato (sgTom) or an empty vector control (**Figure 2d-e**). At 10 weeks post-infection, we assessed knockdown of tdTomato expression by immunohistochemistry (IHC). We observed that 28% of tumors lacked tdTomato expression, suggesting that the system was functional *in vivo* by editing an endogenous allele in the context of a lung tumor (**Figure 3a-e**). Importantly, animals infected with empty pSECC rarely contained non-tumor Tomato-expressing cells (data not shown), indicating that there is minimal infection of non-epithelial cells when using a low lentiviral titer.

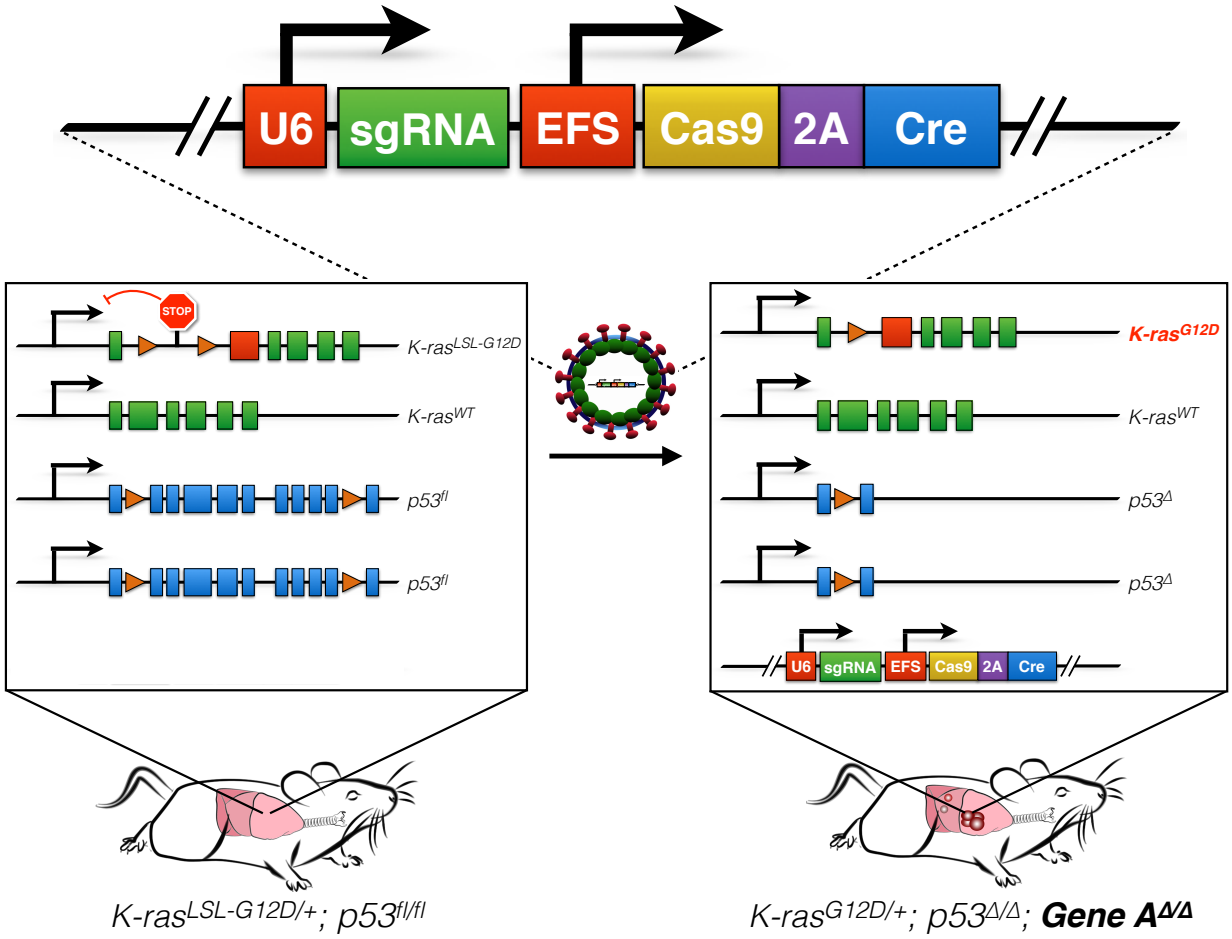


Figure 1: CRISPR-Cas9-mediated somatic genome editing in autochthonous mouse models of lung adenocarcinoma. The pSECC (U6-sgRNA-EFS-Cas9-2A-Cre) lentiviral platform allows for simultaneous delivery of both the CRISPR system and Cre recombinase. Lentiviral constructs expressing sgRNAs of interest are intratracheally delivered into the lungs of Cre/loxP-based mouse models (such as the *Kras*$^{LSL\text{-}G12D/+}$*; p53*$^{flox/flox}$ mouse model shown here (Jackson et al., 2005)). In this context, tumor induction is dependent on Cre, which is co-expressed with Cas9, thus allowing for tumor-specific deletion of any gene of interest with concomitant activation of oncogenic Kras^{G12D} and deletion of both copies of the p53 tumor suppressor.

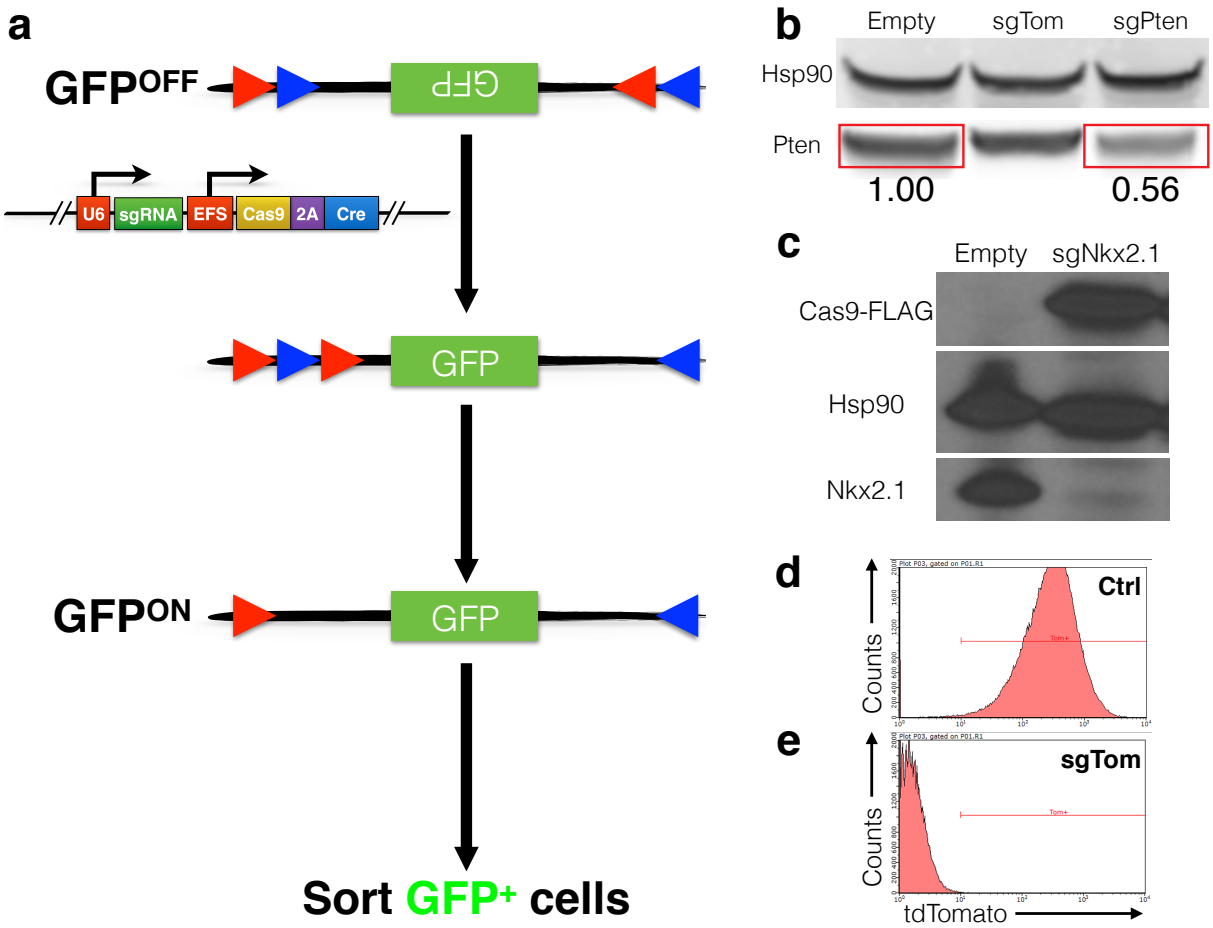


Figure 2: *In vitro* validation of pSECC. (a) The Green-Go Cre-reporter cell line used to validate pSECC lentiviruses *in vitro*. Upon infection with a Cre-containing lentivirus, such as pSECC, cells become GFP+, allowing for purification of pSECC-containing cells by FACS. Red and blue triangles denote pairs of loxP sites, with red loxP sites being able to recombine only with other red loxP sites and blue loxP sites being able to recombine only with other blue loxP sites. (b) Validation of sgPten-pSECC. Numbers below the bands denote quantitation of protein level relative to empty vector control. (c) Validation of sgNkx2.1-pSECC in a cell line that expresses Nkx2.1. (d-e) Validation of sgTom-pSECC by Fluorescence Activated Cell Sorting (FACS). Briefly, a cell line obtained from a *Kras*$^{LSL\text{-}G12D/+}$*; p53*$^{flox/flox}$;*Rosa26*$^{LSL\text{-}tdTomato/LSL\text{-}tdTomato}$ (KPT) mouse was infected with either Empty-pSECC (d) or sgTom-pSECC (e) and cultured for 10 days post-infection, after which the cells were harvested and analyzed by FACS.

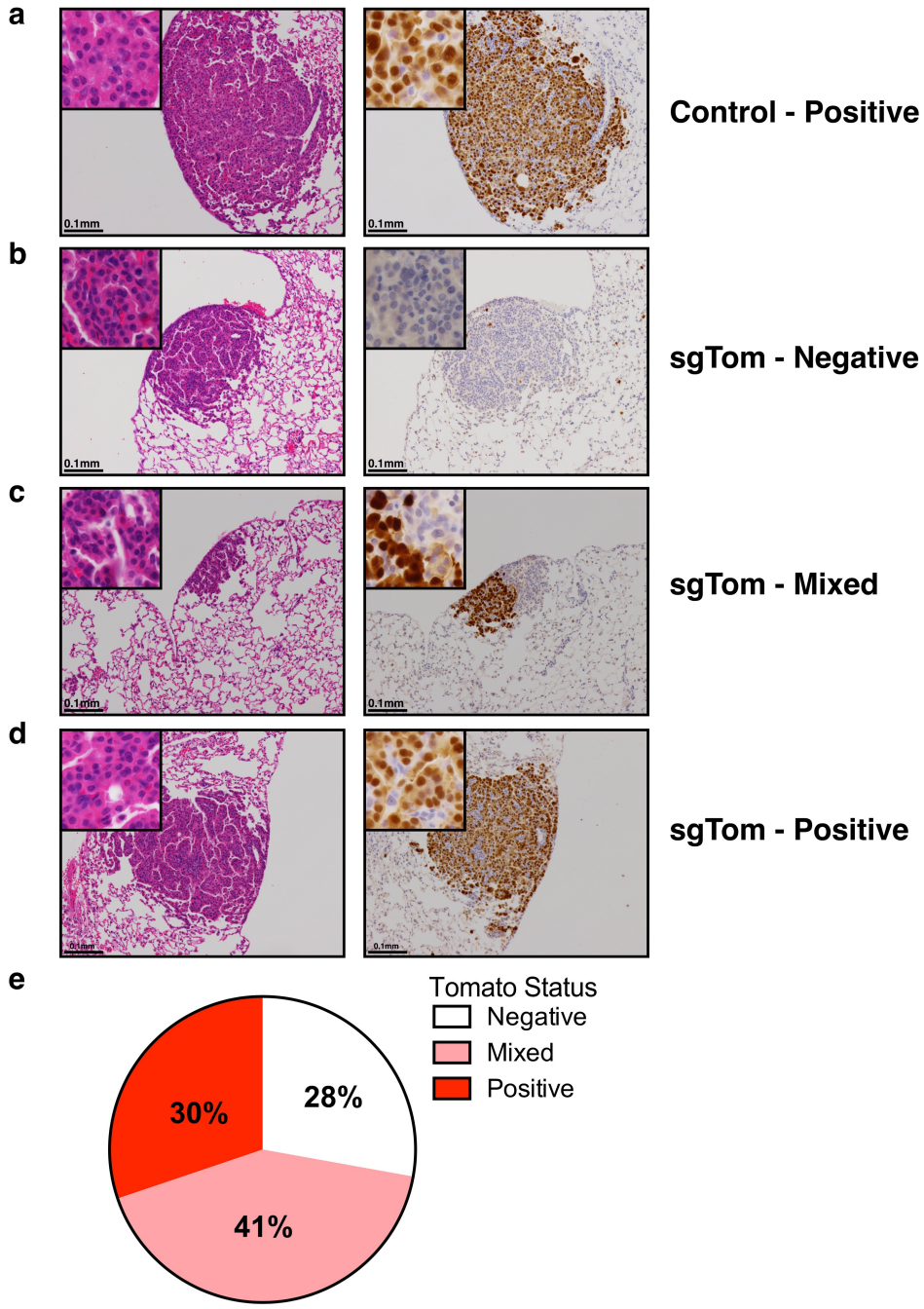


Figure 3: *In vivo* validation of pSECC. (a) Representative H&E and tdTomato IHC staining of serial sections from lung tumors of KPT mice infected with Empty-pSECC. (b-d) Representative H&E and IHC staining of serial sections from (b) negative, (c) mixed and (d) positive lung tumors of KPT mice infected with sgTom-pSECC (n=6). (e) Distribution of lung tumors from all mice infected with sgTom-pSECC (n=6) that were scored as negative, mixed or positive based on tdTomato IHC.

pSECC-mediated gene disruption *in vivo* is highly efficient and mimics Cre/loxP approaches

We then proceeded to functionally characterize tumor suppressor genes using this approach. Loss of NK2 homeobox 1 (*Nkx2.1*), a master regulator of lung development (Rock & Hogan 2011), or Phosphatase and tensin homolog (*Pten*), a negative regulator of oncogenic PI3K/Akt signaling (Song et al., 2012), accelerates lung tumorigenesis in K and KP lung tumor models (Winslow et al., 2011, Curry et al., 2013, Snyder et al., 2013). We infected K and KP animals with pSECC vectors expressing validated sgPten, sgNkx2.1 and controls (sgTom and empty vector) to induce lung tumors. Ten weeks post-infection, we sacrificed animals to assess the effects of CRISPR-Cas9-mediated gene editing in tumors by histopathology, surveyor assays and deep sequencing of the targeted alleles (**Figure 4a**). All animals expressing sgRNAs targeting *Pten* or *Nkx2.1* contained tumors with marked histopathological differences compared to controls **(Figure 4b,d and Figure 5a-d)**.

Animals infected with sgNkx2.1-pSECC developed mucinous adenocarcinomas (MA) typified by the presence of elongated cells, mucin production and glandular rearrangements, in agreement with previous Cre/loxP-based (*Nkx2.1*$^{flox/flox}$) data (Snyder et al., 2013) (**Figure 4b**). The majority of tumors (61%, 54/88 tumors) from sgNkx2.1-pSECC animals lacked Nkx2.1 expression (compared to 0/33 tumors from controls) (**Figure 4b-c**). Importantly, 85% (46/54 tumors) of these Nkx2.1-negative tumors stained positively for mucin (**Figure 4c**), a biomarker of mucinous adenocarcinomas (Snyder et al., 2013). Thus, although a subset of tumors appeared to partially or fully escape CRISPR-mediated deletion of Nkx2.1, we were able to observe

clear phenotypes by examining the full spectrum of tumors generated by sgNkx2.1-pSECC.

Animals infected with sgPten-pSECC demonstrated complete loss of Pten protein in 74% of tumors (40/54 tumors), which was accompanied by a concomitant increase in pAkt (S473), a downstream biomarker of increased PI3-kinase pathway activity (**Figure 4d-e**). These results mimic previously published data using a *Pten*$^{flox/flox}$ allele in K mice (Curry et al., 2013). Collectively, these data indicate that CRISPR-Cas9-based gene editing leads to loss of function mutations in this model and closely parallels what is seen with the use of traditional conditional alleles.

Apc is a potent lung adenocarcinoma tumor suppressor

We next utilized this system to study adenomatous polyposis coli (*Apc*), a tumor suppressor whose functional role in lung adenocarcinoma has not been characterized. Of note, *Apc* is mutated in ~4% of human lung adenocarcinomas and is also found in a region that frequently undergoes copy number loss in human lung cancer (LUAD TCGA 2014, Cerami et al., 2012, Gao et al., 2013). We infected animals with pSECC lentiviruses expressing a validated sgRNA (Schwank et al., 2013) targeting *Apc*. At 10 weeks post-infection, we observed a dramatic difference in the histopathology of sgApc tumors compared to controls (**Figure 4f and Figure 5e**). Importantly, tumors from *Kras*$^{LSL\text{-}G12D/+}$*; Apc*$^{flox/flox}$ mice, which express a conditional allele of *Apc* (Cheung et al., 2010), exhibited identical histopathology (**Figure 4f**). Tumors with Cas9-mediated deletion of *Apc* were highly de-differentiated, invasive and had a significant stromal component (**Figure 4f**). The majority of these tumors (78%, 91/117) stained strongly for nuclear β-catenin, a marker of *Apc* mutation in colon cancer and other settings (Moon et

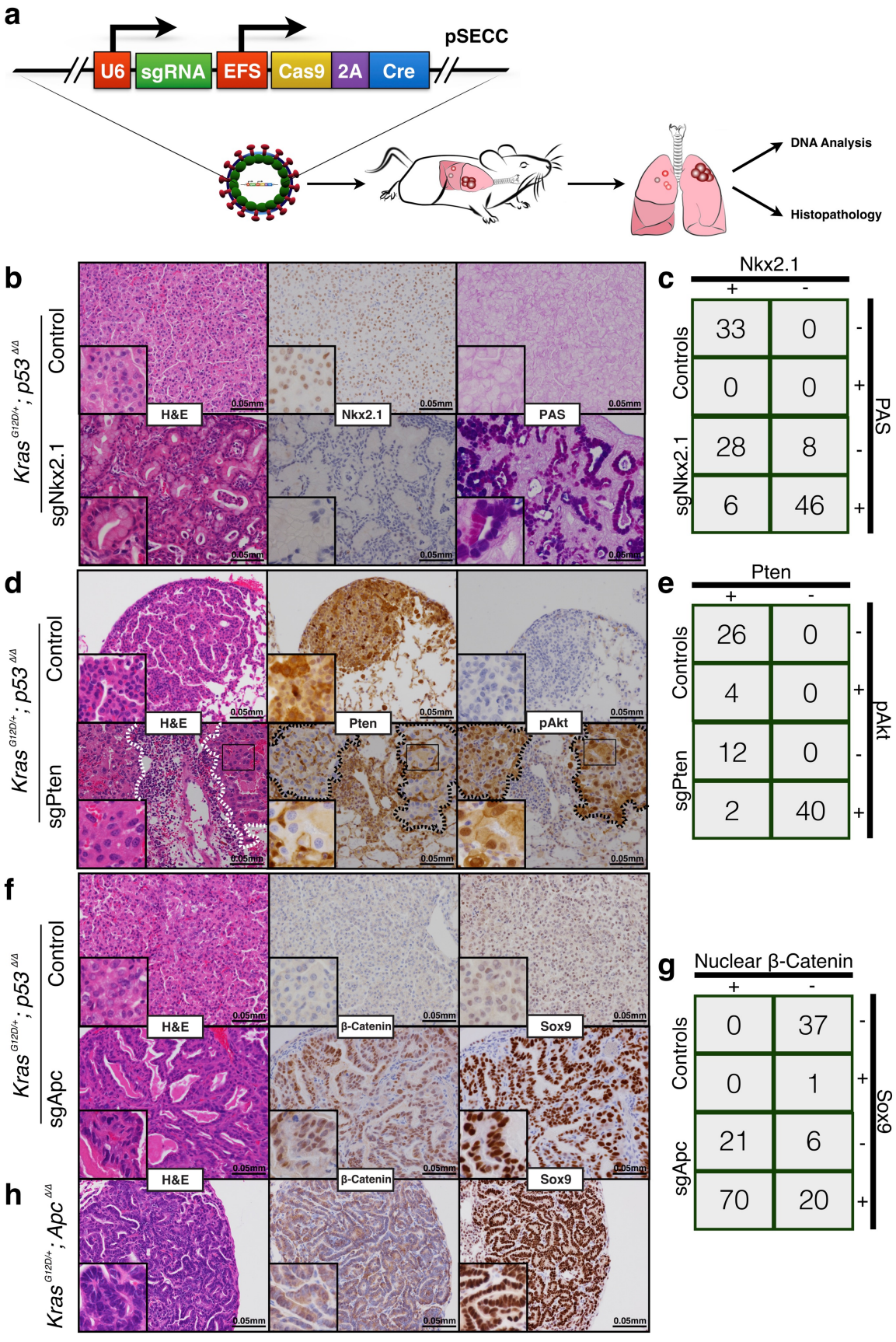

a
pSECC
U6
sgRNA
EFS
Cas9
2A
Cre
DNA Analysis
Histopathology
b
Kras G12D/+; p53 Δ/Δ
Control
sgNkx2.1
H&E
Nkx2.1
PAS
0.05mm
c
Nkx2.1
+
-
Controls
sgNkx2.1
33
0
0
0
28
8
6
46
PAS
d
Kras G12D/+; p53 Δ/Δ
Control
sgPten
H&E
Pten
pAkt
e
Pten
Controls
sgPten
26
0
4
0
12
0
2
40
pAkt
f
Kras G12D/+; p53 Δ/Δ
Control
sgApc
H&E
β-Catenin
Sox9
g
Nuclear β-Catenin
Controls
sgApc
0
37
0
1
21
6
70
20
Sox9
h
Kras G12D/+; Apc Δ/Δ
H&E
β-Catenin
Sox9

Figure 4: CRISPR-Cas9-mediated disruption of a panel of tumor suppressor genes in an autochthonous mouse model of lung adenocarcinoma. (a) pSECC lentiviruses are intratracheally delivered into the lungs of mice to delete genes of interest. DNA extracted from tumor-bearing lungs is analyzed by high-throughput sequencing and surveyor assays to identify gene-editing events. The remaining tissue is analyzed by histopathology. (b) Representative H&E and IHC stainings of serial sections from lung tumors of mice 10 weeks after infection with sgTom-pSECC (left panel) or sgNkx2.1-pSECC (right panel). Alcian Blue/PAS (Periodic Acid-Schiff) stain for mucin. Note the accumulation of mucin only in tumors from sgNkx2.1-pSECC mice. (c) Contingency tables demonstrating anti-correlation between Nkx2.1 expression and mucin production (PAS stain) (two-sided Fisher's exact test, $p<0.0001$). (d) Representative H&E and IHC stainings of serial sections from lung tumors of mice 10 weeks after infection with sgTom-pSECC (left panel) or sgPten-pSECC (right panel). Note: dashed lines demarcate tumor boundaries on each consecutive histological section. (e) Contingency tables demonstrating anti-correlation between Pten expression and Akt phosphorylation (two-sided Fisher's exact test, $p<0.0001$). (f) Representative H&E and IHC stainings of serial sections from lung tumors of mice 10 weeks after infection with sgTom-pSECC (left panel) or sgApc-pSECC (middle panel). The far right panel corresponds to serial sections from lung tumors of $Kras^{LSL\text{-}G12D/+}$; $Apc^{flox/flox}$ mice 18 weeks after infection with Adeno-Cre. (g) Contingency tables demonstrating positive correlation between β-Catenin expression and Sox9 expression (two-sided Fisher's exact test, $p<0.0001$). These data are representative of at least 3 independent K or KP mice infected with each pSECC sgRNA. See **Methods** section for more details

al., 2004) (**Figure 4f-g**). Furthermore, 77% (70/91) of tumors with nuclear β-catenin stained positive for the transcription factor Sox9, which might reflect a distal embryonic differentiation state (Kormish et al., 2010, Pacheco-Pinedo et al., 2011). Of note, we observed a statistically significantly higher number of Sox9-positive tumors in KP-sgApc (29/33, or 88%) than in K-sgApc mice (41/58 tumors, or 71%), suggesting a possible role for p53 in regulating this change in differentiation (**Figure 6b-c**).

To further characterize the differentiation state of sgApc tumors, we stained serial sections for lung differentiation markers, including Sox2, Clara Cell Secretory Protein (CCSP), Surfactant Protein C (SP-C), p63, Nkx2.1 and Sox9 (**Figure 6a**) (Hogan et al., 2014). Tumors from KP-sgTom mice stained positively for CCSP, SP-C and Nkx2.1 and negatively for Sox2, p63 and Sox9. In contrast, tumors from sgApc mice frequently stained positively for SP-C, Nkx2.1 and Sox9 and negatively for CCSP, Sox2 and p63. A large number of tumors from sgApc mice had areas with low levels or complete absence of Nkx2.1, which correlated with the levels of the Nkx2.1 transcriptional target SP-C (Hogan et al., 2014) (16/52 tumors, or 31%) (**Figure 6d**). These data indicate that these tumors are poorly differentiated and that hyperactivation of the canonical Wnt signaling pathway through loss of *Apc* in Kras-driven lung adenocarcinomas results in tumors with varying degrees of differentiation. These results also mimic what we observed in tumors from *Apc* conditional mice (**Figure 4f** and **Figure 6e**) and recapitulate recent observations in a $Braf^{V600E}$-driven mouse model of lung adenocarcinoma upon Wnt pathway hyperactivation (Juan et al., 2014).

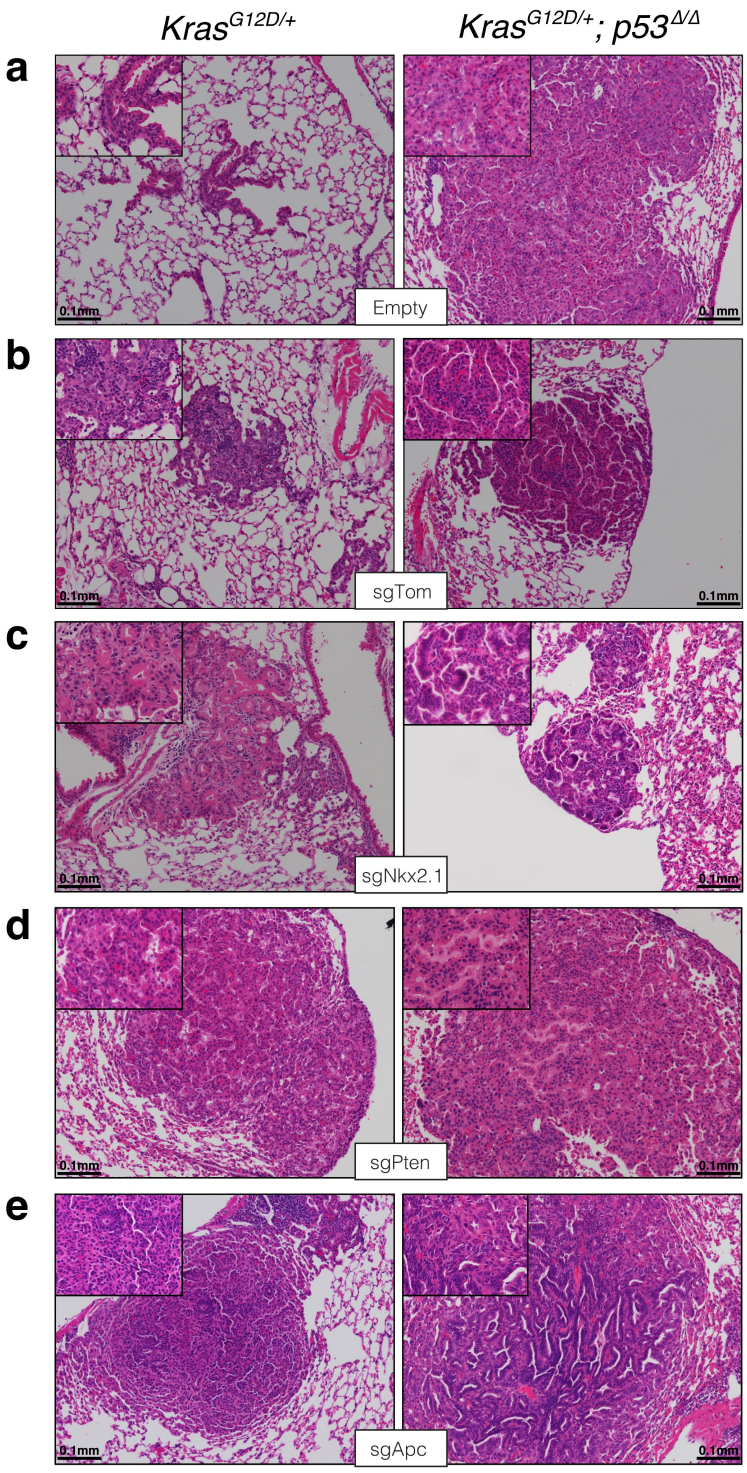


Figure 5: Histological analysis of lung tumors obtained from mice infected with pSECC lentiviruses. Representative H&E images of lung tumors from K and KP mice 10 weeks after infection with (a) Empty-pSECC, (b) sgTom-pSECC, (c) sgNkx2.1-pSECC, (d) sgPten-pSECC and (e) sgApc-pSECC.

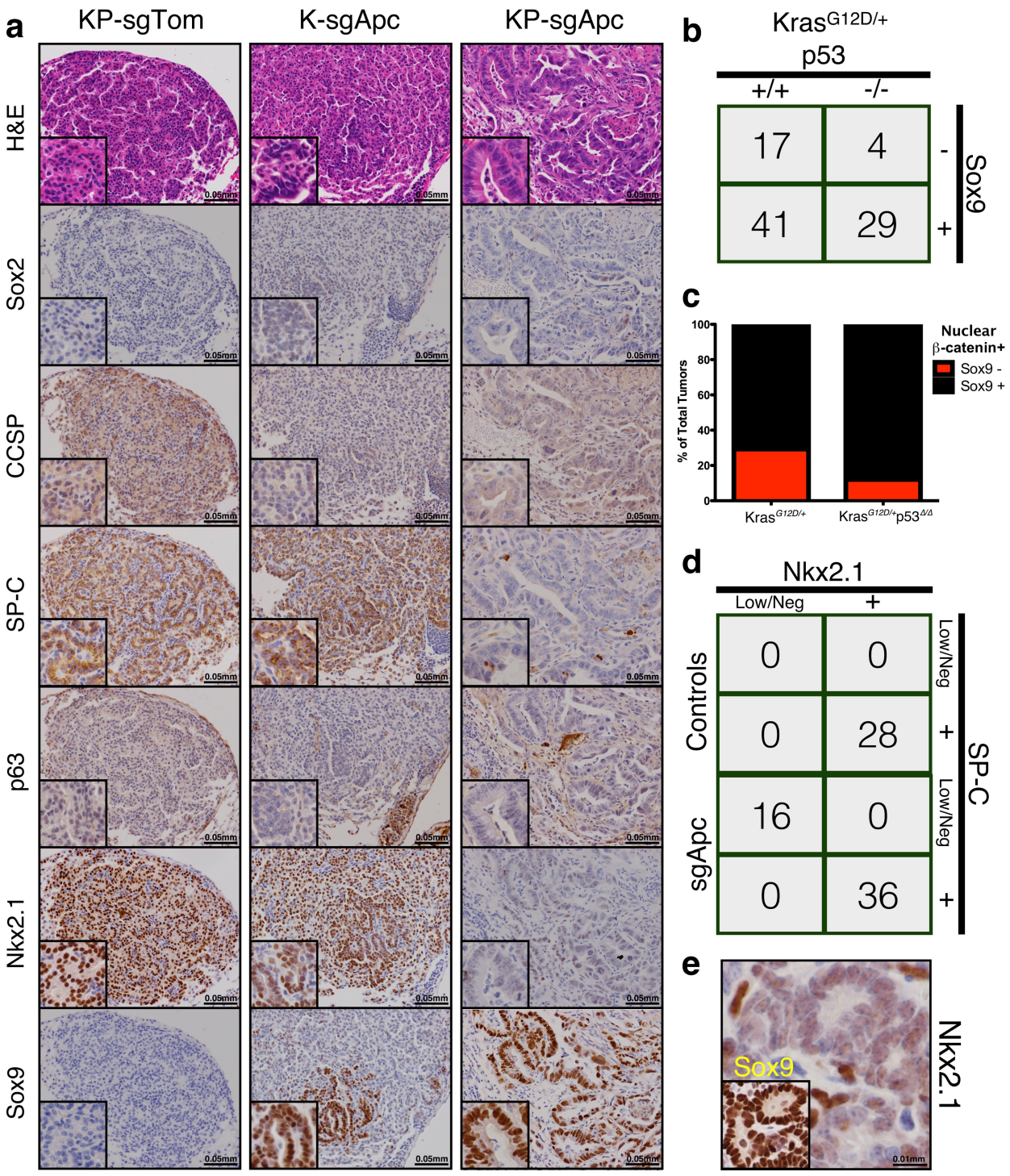

a
KP-sgTom
K-sgApc
KP-sgApc
H&E
Sox2
CCSP
SP-C
p63
Nkx2.1
Sox9
0.05mm
b
Kras G12D/+
p53
+/+
-/-
17
4
41
29
Sox9
-
+
c
% of Total Tumors
100
80
60
40
20
0
Kras G12D/+
Kras G12D/+ p53 Δ/Δ
Nuclear β-catenin+
Sox9 -
Sox9 +
d
Nkx2.1
Low/Neg
+
Controls
sgApc
0
0
0
28
16
0
0
36
Low/Neg
+
Low/Neg
+
SP-C
e
Sox9
Nkx2.1
0.01mm

Figure 6: IHC-based analysis of K- and KP-sgApc tumors. (a) Representative H&E and IHC staining of serial sections from KP-sgTom (control), K-sgApc and KP-sgApc lung tumors. CCSP = Clara Cell Secretory Protein, SP-C = Surfactant Protein C. (b) Contingency table demonstrating a statistically significantly higher number of β-catenin/Sox9 double-positive tumors in KP-sgApc mice (29/33 tumors, 88%) vs K-sgApc mice (41/58 tumors, 71%) (one-sided Chi-square test, $p<0.05$). (c) Percentage of all tumors that stained positive for nuclear β-catenin that stained positive or negative for Sox9 in K- and KP-sgApc mice. (d) Contingency table demonstrating a statistically significantly higher number of tumors with Nkx2.1 Low/Negative areas (which are also SP-C Low/Negative) in sgApc-pSECC animals compared to sgTom-pSECC control animals (two-sided Fisher's exact test, $p<0.0001$). (e) Representative IHC staining of serial sections from an Nkx2.1 Low/Neg lung tumor obtained from a *Kras*$^{LSL\text{-}G12D/+}$; *Apc*$^{flox/flox}$ mouse 18 weeks after infection with Adeno-Cre. Inset shows Sox9 staining. Low/Neg = tumor that had areas with clear downregulation or complete loss of Nkx2.1 or SP-C as assessed by IHC staining.

Disruption of *Nkx2.1*, *Pten* and *Apc* accelerates lung adenocarcinoma initiation and progression

Our initial analysis demonstrated significant histological and pathway-specific differences upon deletion of these tumor suppressors in lung tumors. To assess the overall impact of these alterations on tumorigenesis, we measured tumor burden and grade in both K and KP animals. Deletion of *Pten* and *Apc* significantly increased overall tumor burden, which correlated with higher tumor grades (Grade 3 and 4) (**Figure 7a-c**). *Nkx2.1* deletion had a significant effect on overall tumor burden only in KP animals; however, we observed a striking transition to highly dedifferentiated MA tumors in both K and KP mice (**Figure 7a-c** and **Figure 8**). Conversely, *Apc* deletion had a significant effect on tumor burden only in K mice (**Figure 7a**). Deletion of all three genes led to increased BrdU incorporation, suggesting that the increased tumor burden is partly due to increased proliferation (**Figure 7f**). These data demonstrate the tumor suppressive role of *Nkx2.1*, *Pten*, and *Apc* in the context of oncogenic *Kras*. Furthermore, the unique histopathology observed for each targeted tumor suppressor gene in this Kras-driven model illustrates the potential of this approach to rapidly model the effects of cooperative genetic events in lung tumorigenesis and progression.

CRISPR-based somatic genome editing results in inter- and intra-tumoral heterogeneity

Using this *in vivo* somatic genome editing approach, we observed inter- and intra-tumoral heterogeneity in terms of CRISPR-based loss of function of *Pten* in sgPten animals (**Figure 7d-e** and **Figure 9**). Clones that acquired loss of *Pten* had increased

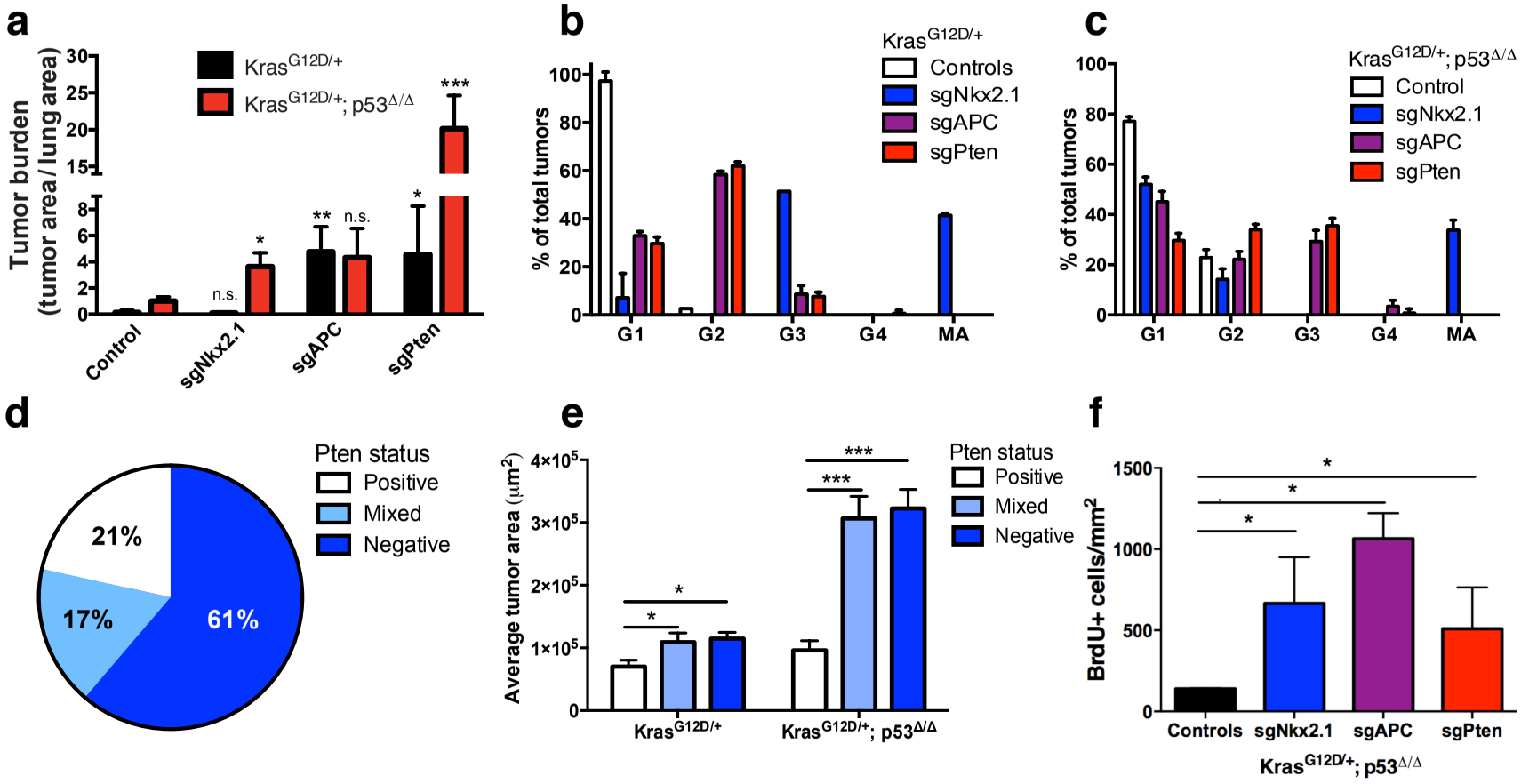

a
Tumor burden (tumor area / lung area)
Kras^G12D/+
Kras^G12D/+; p53^Δ/Δ
n.s.
n.s.
Control
sgNkx2.1
sgAPC
sgPten
b
% of total tumors
Kras^G12D/+
Controls
sgNkx2.1
sgAPC
sgPten
G1
G2
G3
G4
MA
c
% of total tumors
Kras^G12D/+; p53^Δ/Δ
Control
sgNkx2.1
sgAPC
sgPten
G1
G2
G3
G4
MA
d
Pten status
Positive
Mixed
Negative
21%
17%
61%
e
Average tumor area (μm²)
Pten status
Positive
Mixed
Negative
Kras^G12D/+
Kras^G12D/+; p53^Δ/Δ
f
BrdU+ cells/mm²
Controls
sgNkx2.1
sgAPC
sgPten
Kras^G12D/+; p53^Δ/Δ

Figure 7: Histopathological characterization of tumors from pSECC infected animals. (a) Combined quantitation of tumor burden (total tumor area/total lung area) in both K and KP animals 10 weeks after infection with pSECC lentiviruses expressing: control (empty or sgTom, n=4 K and 7 KP), sgNkx2.1 (n=2 K and 6 KP), sgApc (n=3 K and 6 KP) and sgPten (n=4 K and 3 KP). The asterisks indicate statistical significance obtained from comparing K-sgTarget samples to K-Control samples or KP-sgTarget samples to KP-Control samples using Student's t-test (two-sided). (b-c) Distribution of tumor grades in the K (b) or KP (c) animals shown in (a). G1: grade 1, G2: grade 2, G3: grade 3, G4: grade 4, MA: mucinous adenocarcinoma. (d) Distribution of Pten IHC staining status in all sgPten-pSECC infected animals (n=9) represented as percent of negative, mixed and positive tumors. (e) Quantitation of average tumor area (μm^2) of tumors staining negative, mixed or positive in all sgPten-pSECC infected animals (n=9). Positive tumor = ~100% of the tumor cells stained positive for Pten. Mixed tumor = at least ~30% of tumor cells stained positive for Pten. Negative tumor = <25% of the tumor cells stained positive for Pten. (f) Quantitation of BrdU incorporation (BrdU+ cells/mm^2) to assess proliferation of tumor cells from lung tumors in KP animals 10 weeks after infection with pSECC lentiviruses expressing: control (empty or sgTom, n=2), sgNkx2.1 (n=3), sgApc (n=3) and sgPten (n=2). Note: n.s. = not significant, * = $p<0.05$, ** = $p<0.01$, *** = $p< 0.001$ obtained from two-sided Student's t-test. All error bars denote s.e.m.

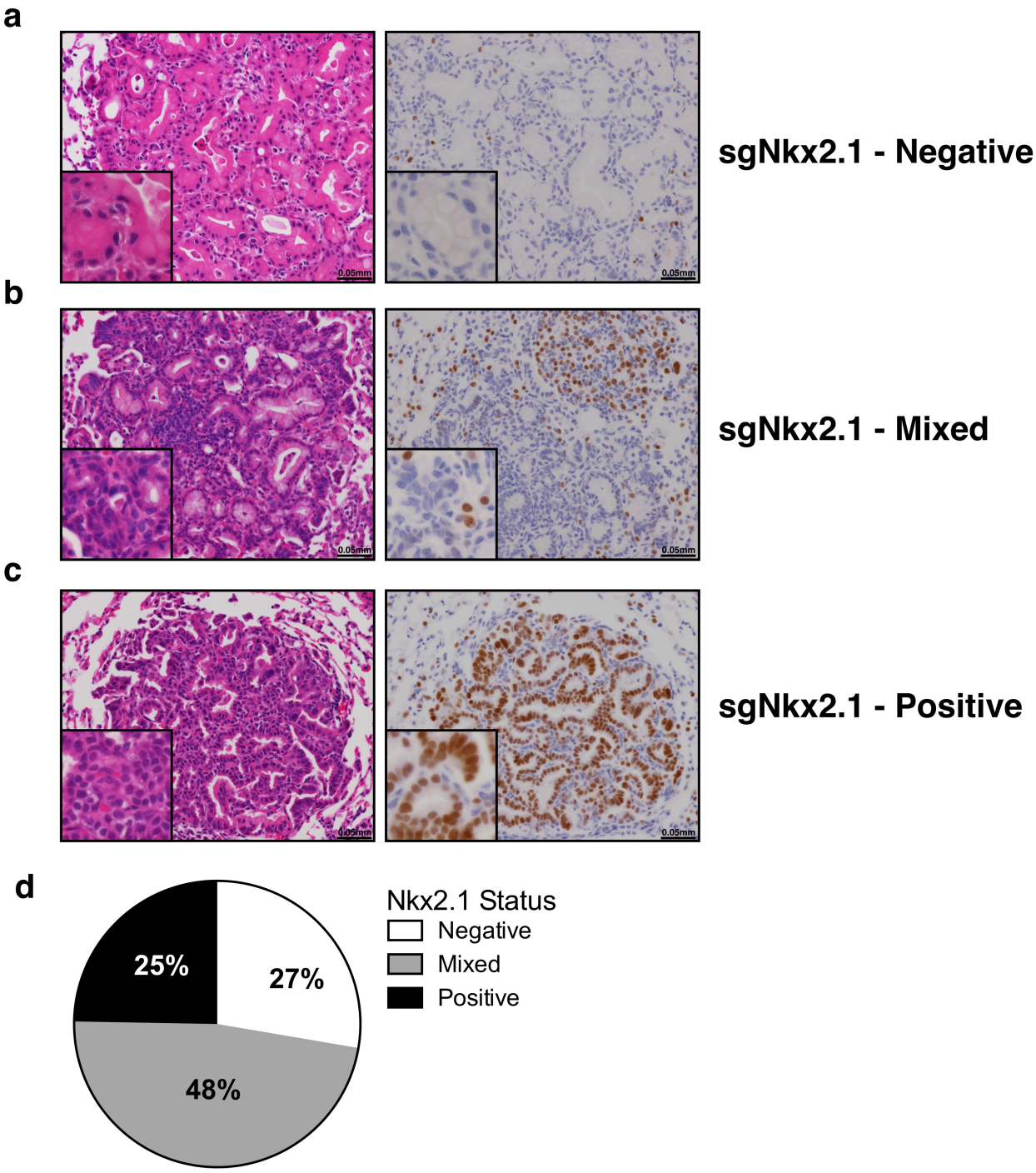


Figure 8: IHC-based analysis of mice infected with sgNkx2.1-pSECC. (a) Negative, (b) mixed and (c) positive lung tumors of mice infected with sgNkx2.1-pSECC. (d) Distribution of Nkx2.1 IHC staining status in all sgNkx2.1-pSECC infected animals (n=8) represented as percent of negative, mixed and positive tumors (staining classification was done identically as described in Figure 7).

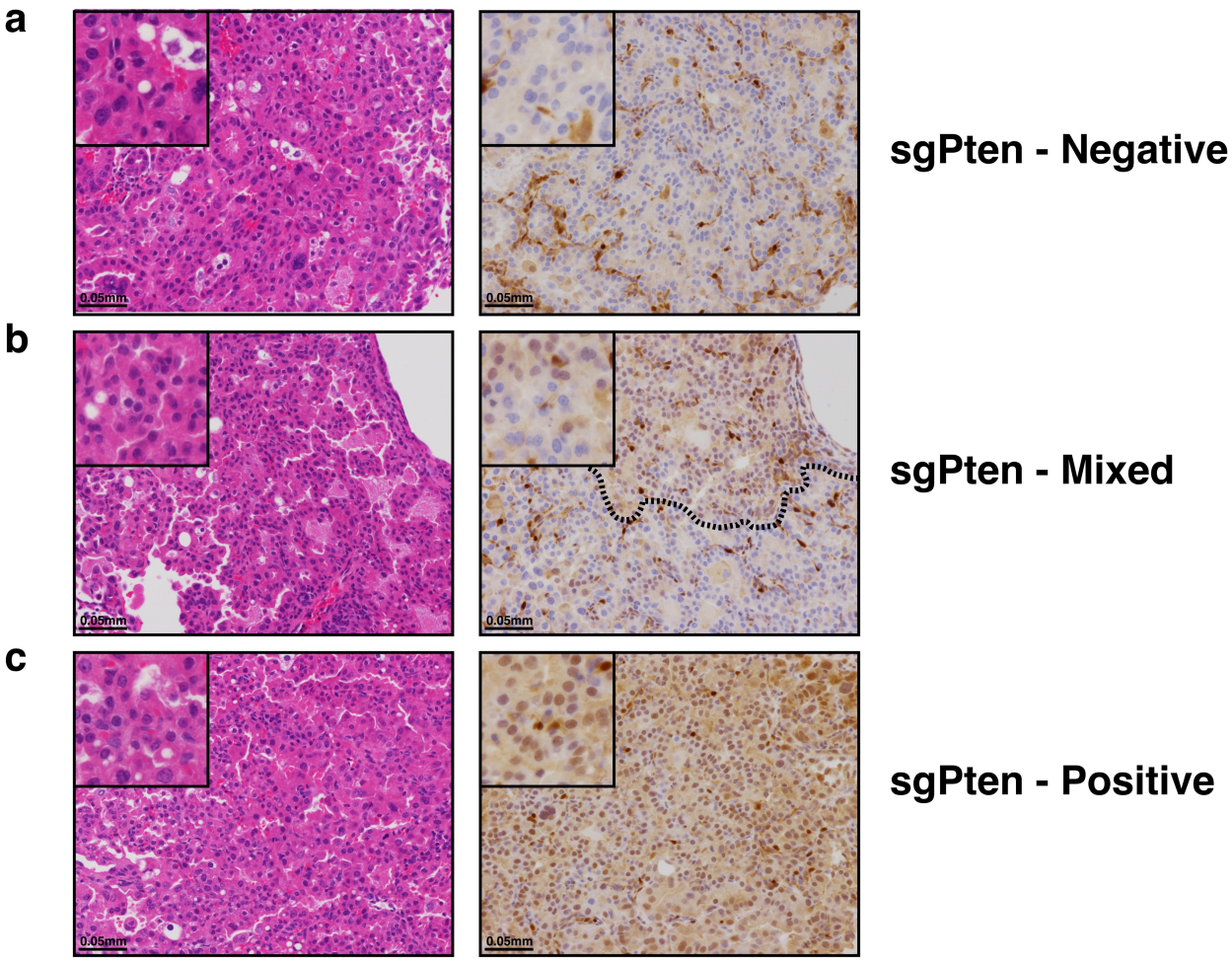


Figure 9: IHC-based analysis of mice infected with sgPten-pSECC. (a) Negative, (b) mixed and (c) positive lung tumors of mice infected with sgPten-pSECC (n=9).

PI3K/Akt signaling and may, therefore, have had a selective advantage over tumors that retained wild-type *Pten* within the same animal. Notably, we observed that tumors with complete or sub-clonal loss of Pten were significantly larger than tumors that retained Pten (**Figure 7d-e**).

The pSECC platform efficiently generates insertions and deletions *in vivo*

The histopathological and IHC analyses indicate that the pSECC system is highly efficient *in vivo*, leading to robust target-specific phenotypic differences in lung tumors. To confirm Cas9-mediated editing of the alleles and precisely characterize the events at single-nucleotide resolution, we performed deep sequencing of target loci from whole lung and tumor DNA. Within a 23bp window (± 10bp flanking the Protospacer Adjacent Motif [PAM] sequence at each locus), the rate of mutations observed in sgTarget samples was significantly greater than in the control samples (**Figure 10a-c**). Using the control samples as a background model to analyze the mutational rate revealed that sgTarget samples were enriched for mutations within 7bp upstream of the PAM sequences in predicted cutting sites, strongly suggesting that they are not secondary consequences of tumor progression (**Figure 10d-f**). The maximum per-base mutation frequency observed in sgTarget samples was 71.7% in *Nkx2.1*, 66.06% in *Pten* and 39.91% in *Apc* (in contrast to control samples: 0.11%, 0.73% and 0.14%, respectively). On average, 27.48% ± 10.3 (*Nkx2.1*), 44.64% ± 5.3 (*Pten*) and 13.54% ± 5.3 (*Apc*) read fragments covering this 7bp locus harbored indels in sgTarget samples. Across all sgTarget samples, > 94% of observed indels constituted nonsynonymous frame-altering events (**Figure 10e**).

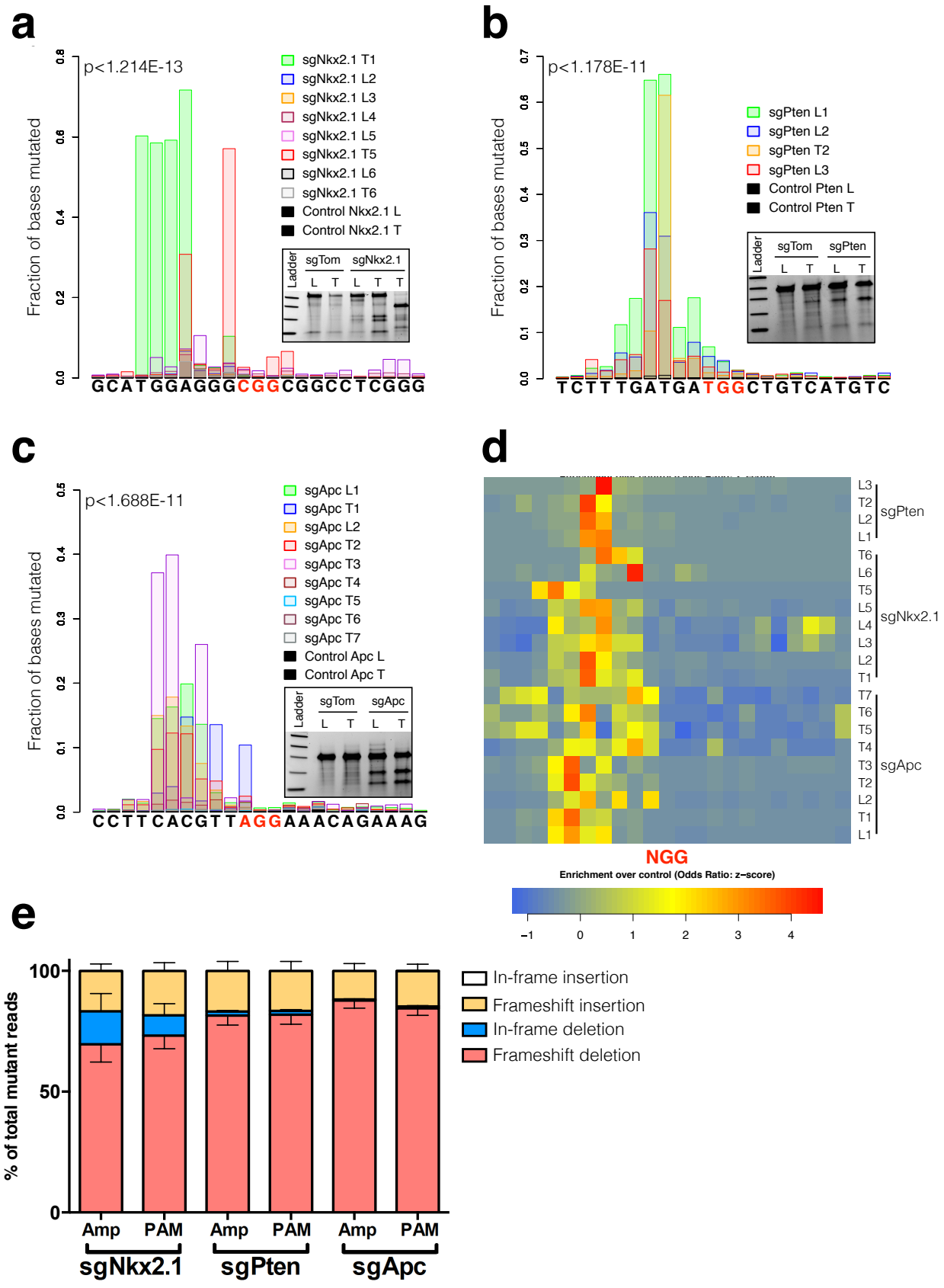

a
p<1.214E-13
Fraction of bases mutated
sgNkx2.1 T1
sgNkx2.1 L2
sgNkx2.1 L3
sgNkx2.1 L4
sgNkx2.1 L5
sgNkx2.1 T5
sgNkx2.1 L6
sgNkx2.1 T6
Control Nkx2.1 L
Control Nkx2.1 T
Ladder
sgTom
sgNkx2.1
GCATGGAGGGCGGCGGCCTCGGG
b
p<1.178E-11
Fraction of bases mutated
sgPten L1
sgPten L2
sgPten T2
sgPten L3
Control Pten L
Control Pten T
Ladder
sgTom
sgPten
TCTTTGATGATGGCTGTCATGTC
c
p<1.688E-11
Fraction of bases mutated
sgApc L1
sgApc T1
sgApc L2
sgApc T2
sgApc T3
sgApc T4
sgApc T5
sgApc T6
sgApc T7
Control Apc L
Control Apc T
Ladder
sgTom
sgApc
CCTTCACGTTAGGAAACAGAAAG
d
sgPten
sgNkx2.1
sgApc
NGG
Enrichment over control (Odds Ratio: z−score)
e
% of total mutant reads
In-frame insertion
Frameshift insertion
In-frame deletion
Frameshift deletion
Amp
PAM
sgNkx2.1
sgPten
sgApc

Figure 10: The pSECC system efficiently generates indels in autochthonous tumors. (a-c) Fraction of bases mutated per position in 10bp flanks on either side of the Protospacer Adjacent Motif (PAM) sequence (highlighted in red). Samples were obtained from entire lobes (L) or microdissected tumors (T) from mice 10 weeks after infection with pSECC lentiviruses targeting (a) *Nkx2.1*, (b) *Pten* or (c) *Apc*. P-values denote enrichment of mutation rate in sgTarget-pSECC samples compared to sgTom-pSECC control samples (Wilcoxon rank sum test). Insets depict surveyor assays for each of the targets from either entire lobes (L) or microdissected tumors (T) from mice. Samples obtained from mice infected with sgTom-pSECC were used as controls. (d) Positional enrichment of mutations in sgTarget-pSECC samples compared to sgTom-pSECC control samples based on all mutations considered at a given position (SNPs, indels). Each row represents a different sgRNA Lung (L) or Tumor (T) sample. Each cell represents the row-normalized (z-score) odds ratio estimate of mutational enrichment over an associated control sample (Fisher's exact test) upstream (+) or downstream (-) of the PAM sequence. (e) Distribution of indels (in-frame insertions, frameshift insertions, in-frame deletions and frameshift deletions) observed in samples from mice infected with sgNkx2.1-pSECC, sgPten-pSECC and sgApc-pSECC. Amp: mutations across whole PCR amplicon, PAM: Mutations 7 base pairs upstream of PAM sequence.

The pSECC platform shows negligible off-target activity

Several studies have reported that Cas9 can bind to sites in the genome other than the intended target site (Wu et al., 2014, Kuscu et al., 2014, Fu et al., 2013, Hsu et al., 2013), which could result in unintended editing at an off-target (OT) site. To assess OT editing, we analyzed the top three predicted (Hsu et al., 2013) loci (**Supplementary Tables 3-5**) for each sgRNA by deep sequencing. We observed negligible OT editing (**Figure 11**). On average, 0.048% ± 0.031 (for sgNkx2.1), 0.26% ± 0.096 (for sgPten) and 0.051% ± 0.027 (for sgApc) of read fragments harbored indels in the OT sites. This data suggests that the observations reported for each of the sgRNAs arise from deletion of the intended target and not from editing of another gene.

DISCUSSION

As discussed extensively in **Part III** of the introduction to this thesis, the goal of cancer genomics is to identify genetic events that underlie cancer initiation and progression. The functional interrogation of putative cancer genes in appropriate experimental models will elucidate which mutations identify *bona fide* cancer genes. This study presents a novel approach to rapidly evaluate human cancer genome candidates and assess cooperativity between genetic events in the context of well-established mouse models of lung cancer. Moreover, our ability to model different lung adenocarcinoma subtypes allows for the detailed study of subtype-specific molecular mechanisms controlling disease initiation and progression. We anticipate that this approach can be readily adapted to many existing Cre/loxP-based GEMMs of several cancer types to facilitate the rapid functional assessment of new hypotheses generated by cancer genome studies. These and other implications are discussed extensively in **Chapter 4**.

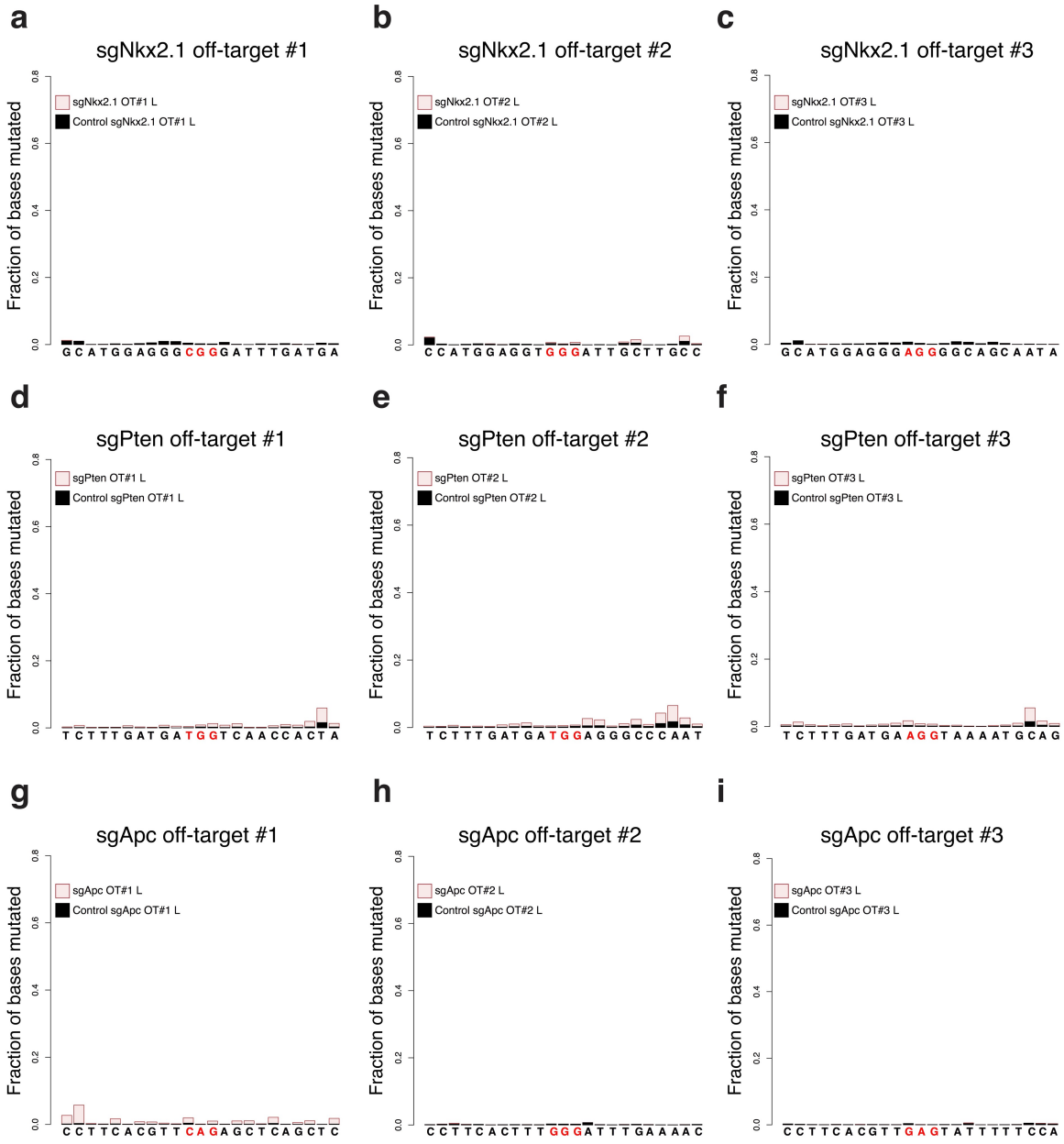


Figure 11: Off-target analysis. Analysis of off-target editing for (a-c) sgNkx2.1, (d-f) sgPten and (g-i) sgApc. Briefly, potential off-target cutting at the top three predicted off-target sites (obtained from http://crispr.mit.edu/; see Supplementary Tables 3-5) for each sgRNA was assayed by Illumina MiSeq. Each plot corresponds to the fraction of bases mutated per position in 10bp flanks on either side of the PAM sequence (highlighted in red). Samples were obtained from entire lobes (L) from mice 10 weeks after infection with pSECC lentiviruses expressing sgNkx2.1, sgPten, sgApc or sgTom (control).

MATERIALS AND METHODS

Lentiviral vectors and sgRNA cloning

The U6-sgRNA-EFS-Cas9-2A-Cre (pSECC) lentiviral vector was constructed by assembling four parts with overlapping DNA ends using Gibson assembly. Briefly, a 2.2kb part (corresponding to the U6-Filler fragment from LentiCRISPR (Shalem et al., 2014)), a 0.3kb part (corresponding to the EFS promoter from LentiCRISPR (Shalem et al., 2014)), a 5.3kb part (corresponding to a Cas9-2A-Cre fragment, which was generated by assembly PCR) and a 5.7kb lentiviral backbone were assembled using Gibson assembly following manufacturer guidelines. For sgRNA cloning, the pSECC vector was digested with BsmBI and ligated with BsmBI-compatible annealed oligos (**Supplementary Table 1**). sgRNAs were designed using CRISPR Design (Hsu et al., 2013) (which was also used to predict potential off-target sites; see **Figure 11** and **Supplementary Tables 3-5**) or E-CRISP (Heigwer et al., 2014), except for sgApc, which was previously reported (Schwank et al., 2013). An extra G (required for U6 transcriptional initiation) was added to the 5' end of sgRNAs that lacked it.

Lentiviral production

Lentiviruses were produced by co-transfection of 293T cells with lentiviral backbone constructs and packaging vectors (delta8.2 and VSV-G) using TransIT-LT1 (Mirus Bio). Supernatant was collected 48 and 72 hours post-transfection, concentrated by ultracentrifugation at 25,000 RPM for 90 minutes and resuspended in an appropriate volume of OptiMEM (Gibco).

Cell culture and generation of Green-Go cells

Cells were maintained in DMEM supplemented with 10% Fetal Bovine Serum and gentamicin. Green-Go cells were generated by transducing 3TZ cells (Psarras et al., 2004) with a bicistronic retrovirus containing an LTR promoter-driven inverted GFP (flanked by two sets of incompatible loxP sites) and a PGK-driven puromycin resistance cassette. Transduced cells were selected with puromycin and a single cell clone that expressed high levels of GFP 2-3 days after infection with a lentivirus expressing Cre recombinase was chosen.

Immunobloting

Cells were lysed with ice-cold RIPA buffer (Pierce, #89900) supplemented with 1× Complete Mini inhibitor mixture (Roche, #11 836 153 001) and mixed on a rotator at 4° C for 30 minutes. Protein concentration of the cell lysates was quantified using the Bio-Rad DC Protein Assay (Catalog #500-0114). 50-80 μg of total protein was separated on 4-12% Bis-Tris gradient gels (Life Technologies) by SDS-PAGE and then transferred to nitrocellulose membranes. The following antibodies were used for immunoblotting: anti-FLAG (Sigma, F1804, 1:1,000), anti-Hsp90 (BD, #610418, 1:10,000), anti-Pten (Cell Signaling, 9188, 1:1,000), anti-TTF1 / Nkx-2.1 (Epitomics, EP1584Y, 1:1,000).

Mice

All animal studies described in this study were approved by the MIT Institutional Animal Care and Use Committee. All animals were maintained on a mixed C57BL/6J x 129SvJ genetic background. *Kras*$^{LSL\text{-}G12D/+}$ and *p53*$^{flox/flox}$ mice have already been described (Jackson et al., 2005; Jackson et al., 2001). Mice were infected intratracheally with lentiviruses as described (Dupage et al., 2009). We infected a total of 7 mice with

Empty-pSECC and 6 mice with sgTom-pSECC (for a total of 13 control mice), as well as 8 mice with sgNkx2.1-pSECC, 9 mice with Pten-pSECC and 9 mice with Apc-pSECC. No randomization or blinding was used. Total lung area occupied by tumor was measured on hematoxylin and eosin (H&E) stained slides using NIS-elements software.

Immunohistochemistry

Mice were euthanized by carbon dioxide asphyxiation. Lungs were perfused through the trachea with 4% paraformaldehyde (PFA), fixed overnight, transferred to 70% ethanol and subsequently embedded in paraffin. Sections were cut at a thickness of four micrometers and stained with H&E for pathological examination. Immunohistochemistry (IHC) was performed on a Thermo Autostainer 360 machine. Slides were antigen retrieved using Thermo citrate buffer, pH 6.0 in the Pretreatment Module. Sections were treated with Biocare rodent block, primary antibody, and anti-Mouse (Biocare) or anti-Rabbit (Vector Labs) HRP-polymer. The slides were developed with Thermo Ultra DAB and counterstained with hematoxylin in a Thermo Gemini stainer and coverslipped using the Thermo Consul cover slipper. The following antibodies were used for IHC: anti-TTF1 / Nkx-2.1 (Epitomics, EP1584Y, 1:1,200), anti-Pten (Cell Signaling, 9559, 1:100), anti-pAkt S473 (Cell Signaling, 4060, 1:100), anti-BrdU (Abcam, 6326, 1:100), anti-β-catenin (BD, 610154, 1:100), anti-Sox9 (Millipore, AB5535, 1:500), anti-RFP (Rockland, 600-401-379, 1:400), anti-Sox2 (Cell Signaling, 3728, 1:250), anti-CCSP (Millipore, 07-623, 1:2,000), anti-SP-C (Chemicon, AB3786, 1:1,000) and anti-p63 (Neomarkers, MS-1081, 1:200). To detect mucin, sections were stained with 1% Alcian Blue pH 2.5 and Periodic Acid-Schiff reagent. All pictures were obtained using a Nikon 80i microscope with a DS-U3 camera and NIS-elements software.

Genomic DNA isolation and Surveyor assay

Genomic DNA from entire snap-frozen left lung lobes or microdissected tumors was isolated using the High Pure PCR Template Preparation Kit (Roche) following manufacturer guidelines. PCR products for surveyor assay were amplified using Herculase II Fusion DNA polymerase (Agilent) (see Supplementary Table 2 for primers used for surveyor assay), gel purified and subsequently assayed with the Surveyor Mutation Detection Kit (Transgenomic). DNA was separated on 4-20% Novex TBE Gels (Life Technologies) and stained with ethidium bromide.

Deep sequencing and bioinformatic analysis of Cas9 target loci

For each target gene or potential off-target site, a genomic region containing the target sequence was amplified using Herculase II Fusion DNA polymerase and gel purified (primer sequences are shown in **Supplementary Table 2**). Sequencing libraries were prepared from 50ng of PCR product using the Nextera DNA Sample Preparation Kit (Illumina) and sequenced on Illumina MiSeq machines. In order to retain high quality sequence for mutation analysis, Illumina MiSeq reads (150bp paired-end) were trimmed to 100mer paired end reads to drop lower quality 3' ends of reads. Traces of Nextera adapters were clipped from PE1 and PE2 100mer reads using the FASTX toolkit (Hannon Lab, CSHL). Reads greater than 15 nucleotides in length were retained. Additionally, reads with 50% or more bases below a base quality threshold of Q30 were dropped from subsequent analysis. Reference sequences with 10bp genomic flanks were indexed using the Burrows-Wheeler Aligner (BWA) IS linear time algorithm (Li & Durbin 2010) and reads were aligned using the BWA aligner. Reads with mapping quality greater than zero were retained. Overlapping alignments of paired end reads

due to short inserts were resolved in order to avoid double counting of coverage and/or mutations observed in a single fragment. In order to minimize alignment ambiguity in the presence of mutations (including indels), the GATK Toolkit (McKenna et al., 2010) was used to realign pooled cohorts mapping to a given locus. Mutations (base substitutions, insertions and deletions) were assessed using a combination of Samtools (Li et al., 2009) and Annovar (Wang et al., 2010) (indel quantification and annotation), NGSUtils/ BAMutils software suite (Breese & Liu 2013) (total mutations per position), and custom scripts. Mutation frequencies were adjusted for sample purity (see next section) and per base substitution, insertion, and deletion frequencies were determined. Significance of overall mutation rates across 10bp flanking the target locus was assessed using the Wilcoxon rank sum test comparing control and sgTarget sample events. Positional enrichment for mutation frequency compared to control samples was assessed using the conditional maximum likelihood odds ratio estimate (Fisher's exact test) and was mean centered and scaled (z-scores) across a 10bp flank on either side of the PAM sequence in each sample. A number of other utilities/tools were used to enable various parts of the analysis, including: BEDTools (Quinlan & Hall 2010), the Integrated Genome Viewer (IGV) (Thorvaldsdottir et al., 2013), and Picard (http:// picard.sourceforge.net/). Statistical analyses and sequence enrichment plots were implemented in R (http://www.R-project.org). Illumina MiSeq sequence datasets have been deposited into the NCBI repository under **BioProjectID PRJNA256245**.

Tumor purity correction

Lung lobe and microdissected tumor genomic DNA was used to perform real-time PCR based analysis to detect the relative levels of the un-recombined *Kras*$^{LSL\text{-}G12D}$ allele

(from non-tumor tissue) using forward primer 5'- CTCTTGCCTACGCCACCAGCTC -3' and reverse primer 5'- AGCTAGCCACCATGGCTTGAGTAAGTCTGCA -3'. To correct for DNA loading of each sample, we amplified the chr5:10054507-10054621 region using forward primer 5'- GAAGAAATTAGAGGGCATGCTTC -3' and reverse primer 5'- CTTCTCCCAGTGACCTTATGTA -3'. Real-time PCR reactions were performed using KAPA Fast SYBR master mix in a Roche LightCycler Real-Time PCR instrument. To calculate percent purity we performed the following calculations for each sample: $\Delta Cp^{tumorX} = Cp^{Chr5} - Cp^{K\text{-}rasLSL\text{-}G12D/+}$ to normalize for sample loading followed by $1/\Delta\Delta Cp = 1/(\Delta Cp^{tumorX} - \Delta Cp^{LungControl})$ for each sample to compare relative purity to lung tissue from *Kras*$^{LSL\text{-}G12D/+}$ animals that were not infected with Cre. To validate the assay, we generated mouse embryonic fibroblasts from *Kras*$^{LSL\text{-}G12D/+}$ mice treated with Cre recombinase (or control FlpO recombinase).

Statistics

P-values were determined by Student's t-test for all measurements of tumor burden and IHC quantifications except for contingency tables, in which Fisher's exact test or Chi-square test were used. All error bars denote s.e.m.

ACKNOWLEDGEMENTS

We thank D. McFadden and Y. Soto-Feliciano for critical reading of the manuscript, H. Yin, S. Levine, and T. Mason for MiSeq sequencing, R. Stott, J. Bartlebaugh, and C. Shivalila for technical assistance and K. Cormier and C. Condon from the Hope Babette Tang (1983) Histology Facility for technical support. This work was supported by the Howard Hughes Medical Institute, the Ludwig Center for Molecular Oncology at MIT and in part by Cancer Center Support (core) grant P30-CA14051 from the National Cancer

Institute. T.J. is a Howard Hughes Medical Institute Investigator, the David H. Koch Professor of Biology, and a Daniel K. Ludwig Scholar.

SUPPLEMENTARY TABLES

Supplementary Table 1. sgRNA sequences (PAM sequence is in bold)

Target gene	sgRNA ID	Target sequence (5' to 3')	Strand	Reference
tdTomato	sgTom	GGCCACGAGTTCGAGATCGA**GGG**	+	This thesis
Nkx2.1	sgNkx2.1	AAGAAAGTGGGCATGGAGGG**CGG**	-	This thesis
Pten	sgPten	GCTAACGATCTCTTTGATGA**TGG**	-	Xue et al., 2014
Apc	sgApc	GTCTGCCATCCCTTCACGTT**AGG**	+	Schwank et al., 2013

Supplementary Table 2. Primer Sequences

Primer ID	Sequence (5' to 3')	Use
Nkx2.1-F	CATCCAACAAGATCGGCGTT	Amplify Nkx2.1 locus for Surveyor + MiSeq
Nkx2.1-R	CTCATGTTGCCCAGGTTGC	
Pten-F	GAGCCATTTCCATCCTGCAG	Amplify Pten locus for Surveyor + MiSeq
Pten-R	CTAGCCGAACACTCCCTAGG	
Apc-F	AAGACCAGGAAGCCTTGTGG	Amplify Apc locus for Surveyor + MiSeq
Apc-R	GCTTGTGTCTCTGCTTACTCC	
Nkx2.1-OT.1-F	AAGACAGGCTTCTGGCAACA	Amplify sgNkx2.1 top three off-targets for MiSeq analysis
Nkx2.1-OT.1-R	TTGGATCCTCTGCCCAGTCA	
Nkx2.1-OT.2-F	TGTTTTCCTGTCTGTGTGCAT	
Nkx2.1-OT.2-R	GTGAGCTTCGGATGAGAGCA	
Nkx2.1-OT.3-F	GTCCCCAGTGACATGCTGAA	
Nkx2.1-OT.3-R	AGGACAGCCAGGGCTACATA	
Pten-OT.1-F	CCTGGGAGTCACCATACCAT	Amplify sgPten top three off-targets for MiSeq analysis
Pten-OT.1-R	CCAGTGAGGCAAAAGGCAGA	
Pten-OT.2-F	AGATGAAACCCTGGAGCAGC	
Pten-OT.2-R	TGAAGGGATGGGGACTCCTT	
Pten-OT.3-F	GAAGCCCGGTCTTTGGTGTA	
Pten-OT.3-R	GGTGGCAGAACGGGTAACAT	

Apc-OT.1-F	AGCTAAGCTGCTTCTGCTCC	Amplify sgApc top three off-targets for MiSeq analysis
Apc-OT.1-R	TCCCACGAGGAGTCAGAGTT	
Apc-OT.2-F	GCTTTACATGAGGGTGGCCT	
Apc-OT.2-R	TGCTCATTGGAAGTGGCTCC	
Apc-OT.3-F	GATGCTACTTGCCGATGGGA	
Apc-OT.3-R	ACTCTGCAGTCCCTACGCTA	

Supplementary Table 3. Top 25 predicted off-targets for sgNkx2.1

Off Target I.D.	Sequence (20nt + PAM)	Score	Mismatches	UCSC Gene	Locus
1	CTTAAAGTGGGCATGGAGGGCGG	2.3	3MMs [1:2:3]		chrX:-131618288
2	AAGAAAGTGGCCATGGAGGTGGG	1.9	2MMs [11:20]		chr4:+89481909
3	ATGAAAATGTGCATGGAGGGAGG	1.7	3MMs [2:7:10]		chr9:-63652289
4	CATAAAGTTGGCATGGAGGGAGG	1.6	3MMs [1:3:9]		chr4:+74763097
5	TGGAAAGTGGCCATGGAGGGGGG	1.6	3MMs [1:2:11]		chr9:-96847779
6	CAGAAAGGGGACATGGAGGGTGG	1.6	3MMs [1:8:11]		chr18:-82166127
7	AGGAAGGAGGGCATGGAGGGTGG	1.5	3MMs [2:6:8]		chr3:+86355463
8	AACAAAGATGGCATGGAGGGAAG	1.5	3MMs [3:8:9]		chr6:+86507745
9	ATGAGAGTGGGCATGGAGGTGGG	1.5	3MMs [2:5:20]		chr12:-58293995
10	AGTAAAGTGGGCATGGAGGTAGG	1.5	3MMs [2:3:20]		chr11:+84599012
11	AAACAAGTGGGCATGGAGGTCAG	1.4	3MMs [3:4:20]		chr5:+39006974
12	GAGAGAGTGGGAATGGAGGGGGG	1.4	3MMs [1:5:12]		chr2:+35910920
13	GTGAGAGGGGGCATGGAGGGAGG	1.4	4MMs [1:2:5:8]		chr5:-121237232
14	GTGAGAGGGGGCATGGAGGGAGG	1.4	4MMs [1:2:5:8]		chr5:-121251360
15	AGGTGAGGGGGCATGGAGGGGGG	1.4	4MMs [2:4:5:8]		chr1:-40351347
16	AAGGAAGTGAACATGGAGGGTGG	1.3	3MMs [4:10:11]		chr11:-121735416
17	AAACCAGAGGGCATGGAGGGTGG	1.3	4MMs [3:4:5:8]		chr16:+24461138
18	AAGAAAGTGGTCATGGAGTGAGG	1.3	2MMs [11:19]		chr4:+126181104
19	CAGAGAGCGAGCATGGAGGGCGG	1.3	4MMs [1:5:8:10]	NM_001004362	chr10:-78560360
20	GAGTGAGTGAGCATGGAGGGGGG	1.3	4MMs [1:4:5:10]		chr8:+91567771
21	AGGATAGAGAGCATGGAGGGCAG	1.3	4MMs [2:5:8:10]		chr5:+132505599
22	AGAAAAGGGAGCATGGAGGGCAG	1.3	4MMs [2:3:8:10]		chr14:-122320136
23	AAGAAAGGGCACATGGAGGGTGG	1.2	3MMs [8:10:11]		chr16:+5095317
24	AAGAAAGTGAGCATGAAGGGGGG	1.1	2MMs [10:16]		chr11:-63441016
25	AAGTAAGTGTGCCTGGAGGGGGG	1	3MMs [4:10:13]	NM_023739	chr4:+40924376

Supplementary Table 4. Top 25 predicted off-targets for sgPten

Off Target I.D.	Sequence (20nt + PAM)	Score	Mismatches	UCSC Gene	Locus
1	CCTATCGATTTCTTTGATGATGG	2.5	3MMs [1:5:10]		chr1:-98296786
2	AATACCGGTCTCTTTGATGATGG	1.4	4MMs [1:2:5:8]		chr10:+11506619
3	TGTCACGATGTCTTTGATGAAGG	1.3	4MMs [1:2:4:10]		chr6:+110090640
4	GCTTACGATGTATTTGATGATGG	1.2	3MMs [4:10:12]		chr1:-148546226
5	GGTGACTTTCTCTTTGATGACAG	0.9	4MMs [2:4:7:8]		chr2:-42473425
6	AGTAGCTATCTCTTTGATGAGAG	0.9	4MMs [1:2:5:7]		chr12:-8417198
7	TGTAACAATGTCTTTGATGAAAG	0.9	4MMs [1:2:7:10]	NM_146253	chr2:+37283154
8	ACTAACAATCTCTTTGGTGAGAG	0.9	3MMs [1:7:17]		chrX:-95318736
9	GCTGACACTGTCTTTGATGATAG	0.9	4MMs [4:7:8:10]	NM_007410	chr3:+138114183
10	ACTCATGTTCTCTTTGATGAAGG	0.8	4MMs [1:4:6:8]		chr17:-72348052
11	AATACTGATCTCTTTGATGAAAG	0.8	4MMs [1:2:5:6]		chr4:+144606221
12	AGTCATGATCTCTTTGATGATAG	0.8	4MMs [1:2:4:6]		chr13:-86124974
13	TCTATGGATTTCTTTGATGAAAG	0.8	4MMs [1:5:6:10]		chr14:+120287042
14	GCTAGAGCTATCTTTGATGACAG	0.7	4MMs [5:6:8:10]		chr12:+72133526
15	GGAAACGATGGCTTTGATGACAG	0.7	4MMs [2:3:10:11]	NM_001079824	chr10:-62480284
16	ACTCACGTTCTCTTTGGTGACAG	0.6	4MMs [1:4:8:17]		chr11:+23874780
17	GGTTACGTTCTCTTTGCTGAGAG	0.6	4MMs [2:4:8:17]		chr11:+87251302
18	CATAACGATTTCTTTGTTGAAAG	0.6	4MMs [1:2:10:17]		chr7:+68979715
19	AATAACGGTCTCTTTGATCAGGG	0.5	4MMs [1:2:8:19]		chr18:+58177529
20	GCTTAGTATATCTTTGATGAGGG	0.5	4MMs [4:6:7:10]		chr19:+29505877
21	GCTAAAGACCACTTTGATGAAGG	0.5	3MMs [6:9:11]	NM_001170912	chr7:+116618495
22	GCTAACTATCCCTTTGATGCTGG	0.5	3MMs [7:11:20]		chr3:-56236438
23	GGTAATGAGATCTTTGATGATGG	0.5	4MMs [2:6:9:10]		chr10:-70527111
24	GATAGCCATCTCTTTGATGTGAG	0.5	4MMs [2:5:7:20]		chr8:+45776715
25	GCTACCAGTCTCTTTGATGGGAG	0.5	4MMs [5:7:8:20]		chr5:+51728529

Supplementary Table 5. Top 25 predicted off-targets for sgApc

Off Target I.D.	Sequence (20nt + PAM)	Score	Mismatches	UCSC Gene	Locus
1	GTCTTACCTCCCTTCACGTTCAG	1.4	3MMs [5:6:8]		chr10:-60270239
2	GTCTGCCATGCCTTCACTTTGGG	1.4	2MMs [10:18]		chr8:+44149079
3	GTCAGCATTTCCTTCACGTTGAG	0.9	4MMs [4:7:8:10]	NM_027493	chr14:+30797330
4	TTCTCTGATCCCTTCACGTTCAG	0.6	4MMs [1:5:6:7]		chr4:-64804834
5	GAGTGCCATACCTTCAGGTTGGG	0.6	4MMs [2:3:10:17]		chr5:+39162358
6	CTCATCCATCCCTTCACGCTAGG	0.5	4MMs [1:4:5:19]		chr1:+92135983
7	GGCTAGCATCCTTTCACGTTTGG	0.4	4MMs [2:5:6:12]	NM_028325	chrX:+33737427

8	GTTTGACATTACTTCACGTTAAG	0.4	4MMs [3:6:10:11]		chr8:+39623910
9	GCCTGGCATTCCTTCATGTTAGG	0.3	4MMs [2:6:10:17]	NM_001033293	chr2:-25217737
10	GTTTGGCATTCCTTCATGTTGGG	0.3	4MMs [3:6:10:17]	NM_133167	chr15:-84101558
11	TTCTGCCATGTCTTCAAGTTGGG	0.3	4MMs [1:10:11:17]		chr11:-22688151
12	ATATGCCATCGCTTCACGATCAG	0.3	4MMs [1:3:11:19]		chr10:-117471455
13	GTCTGCCAGCCCTTCAGGTAGAG	0.3	3MMs [9:17:20]		chr11:-22189861
14	ATCTGCCTTCCCTTCAAGTACAG	0.3	4MMs [1:8:17:20]		chr12:-10429050
15	GTCTGTACTCCCTTCACGTGGGG	0.3	4MMs [6:7:8:20]		chr11:+10271639
16	GTATGCCATACTTTCATGTTGAG	0.3	4MMs [3:10:12:17]		chr8:-111960439
17	GCCTGTTATCCCTTCAAGTTTAG	0.3	4MMs [2:6:7:17]		chr10:+122298126
18	ATCTGTGATCCCATCACGTTTGG	0.2	4MMs [1:6:7:13]		chr5:+130464081
19	GACTGCAGTCCCTTCACCTTTGG	0.2	4MMs [2:7:8:18]		chr12:-70637219
20	GGCTGCTTTCCCTTCACTTTCAG	0.2	4MMs [2:7:8:18]		chr9:-45088057
21	TTCTGACATCTCTTCAGGTTTAG	0.2	4MMs [1:6:11:17]		chr8:-92427843
22	GTCTCCCATGCCATCAAGTTAGG	0.2	4MMs [5:10:13:17]		chr6:-16030290
23	CTCTTCCATCCCTTCATGGTGGG	0.2	4MMs [1:5:17:19]		chr15:+75367728
24	CTCTTCCATCCCTTCATGGTGGG	0.2	4MMs [1:5:17:19]		chr15:+75295342
25	GGCTTCCATCCCTTCAGGGTGAG	0.2	4MMs [2:5:17:19]		chr12:-5215121

Supplementary sequences

Reference sequences used for MiSeq Analysis of Nkx2.1, Pten, Apc and the top three predicted off-targets for each sgRNA. gRNA sequence and PAM sequence (or potential off-target 20bp sequence and potential PAM sequence) are highlighted in yellow and red, respectively.

> Nkx2.1 Reference Sequence (chr12:57635764-57636182) (419bp)

CATCCAACAAGATCGGCGTTAAGGTAACACCAGAATATTTGGCAAAGGGAGAAAAAA
AAAGTAGCGAGGCTTCGCCTTCCCCCTCTCCCTTTTTTTTTTCCTCCTCTTCCTTCC
TCCTCCAGCCGACGCCGAATCATGTCGATGAGTCCAAAGCACACGACTCCGTTCTC
AGTGTCTGACATCTTGAGTCCCCTGGAGGAAAGCTACAAGAAAGTGGGCATGGAG
GGCGGCGGCCTCGGGGCTCCGCTCGCAGCGTACAGACAGGGCCAGGCGGCCCC
ACCGGCCGCGGCCATGCAGCAGCACGCCGTGGGGCACCACGGCGCCGTCACCG
CCGCCTACCACATGACGGCGGCGGGGGTGCCCCAGCTCTCGCACTCCGCCGTGG
GGGGCTACTGCAACGGCAACCTGGGCAACATGAG

> Pten Reference Sequence (chr19:32832846+32833190) (345bp)

GAGCCATTTCCATCCTGCAGAAGAAGCCTCGCCACCAGCAGCTTCTGCCATCTCTC
TCCTCCTTTTTCTTCAGCCACAGGCTCCCAGACATGACAGCCATCATCAAAGAGATC
GTTAGCAGAAACAAAAGGAGATATCAAGAGGATGGATTCGACTTAGACTTGACCTGT
ATCCATTTCTGCGGCTGTTCCTCTTTGCTTTTCTGTCACTCTGATAACGTGGGAGTA

GACGGATGCGAAAAATTTCTGTAGTTGGGGTGACTATAACGTTTAATTCTGGGCGCA
TTTCTAGATCGTGCATATTGTGTCTCTTCCAGTGTATTCAACCTAGGGAGTGTTCGG
CTAG

> Apc Reference Sequence (chr18:34471778+34472077) (300bp)

AAGACCAGGAAGCCTTGTGGGACATGGGGGCAGTGAGCATGCTCAAGAACCTCAT
TCATTCCAAGCACAAAATGATTGCCATGGGAAGTGCAGCAGCTTTAAGGAATCTCAT
GGCAAACAGACCTGCAAAGTATAAGGATGCCAATATCATGTCTCCCGGCTCAAGTCT
GCCATCCCTTCACGTTAGGAAACAGAAAGCTCTAGAAGCTGAGCTAGATGCTCAGC
ATTTATCAGAAACCTTCGACAACATTGACAACCTAAGTCCCAAGGCCTCTCACCGGA
GTAAGCAGAGACACAAGC

> Nkx2.1-OT.1 (chrX:131618130+131618500) (371bp)

AAGACAGGCTTCTGGCAACAGCCTCGTGGACTGATCTGGAGAACTCAGATATTTCA
AGAAGCTACTAGTGGTGCCTGGCACAGAACAATTCCTCCGTGTGTATGCTTGCAAA
GATGACAAGCATTGTCCTCCCAAACATCAAACAACATACAGTCATCAAATCCCGCCC
TCCATGCCCACTTTAAGGACAGAGGTGAAGTAGCCTGCCCAGGACACCATAGCAG
GTCAAAACCAGAGCCTCATTCATGGCTTTGTTCTCCACAAACAGTGGCTTTCTCCTA
TAGAGCATGGACTTAGTTTGGTTAGAGTGACGTCACAGTATTCGCTGAAGCTCAGG
GACCAGCCAGGATCTGACTGGGCAGAGGATCCAA

> Nkx2.1-OT.2 (chr4:89481779+89482228) (450bp)

TGTTTTCCTGTCTGTGTGCATAGTAAAACGACATATTCAATAACAAAAGCCAATTGTAT
CCCAAGTATGCTTATCCAATTATGCATACATCTATACAATTTGCATTATATTGATGATCA
TAATGTGCTATTTAAGAAAGTGGCCATGGAGGTGGGATTGCTTGCCCCCATCAGCCA
TCTTTGATCCAAGTATGCTTCTTTTAGTGTCCTCCAAGATCACTTGTATTAGAATAGGT
TGGCTCTGATAACTGCTAAATAGGAGGGAAGAAAACAAGAAAACAAAGTCAAATCAA
CACCACTGTGTCCTTGAGACTAGCTAGCTGGTAAACCATTTCTGATAGATATTATTGT
TTATCCTGTGATATAAATGTCTTTGTAGTGAGGGATGGGAGTTAACAGTTCAACCCCA
GAAAGCTCATCCATGGTGGAGCTGCTCTCATCCGAAGCTCAC

> Nkx2.1-OT.3 (chr9:63652050+63652590) (541bp)

GTCCCCAGTGACATGCTGAAGGCACAGCAAGAGAATGGTCTGTCATCCAGTCAGC
AGAGCAGGTACCAACCACATGGACTAGGTCCAGTCAAAGCACCTTGCCTCTCAACT
AATATGTTTATGTCAACCCAGCAAGAGGAGACCATGGCTCAGGCAAGGAAGGAAAG
AACTAAGAAAATGTATACATGCTCACAAGACATGAGGAAGACAGACAGACAGACAGA
TAGATGTTATTATTGCTGCCCCTCCCTCCATGCACATTTTCATCAATATCAGTTATGAA
CACAGCACATGGGCACAGGGCTCTAACCCATGTCTTCCAGTCCATGTGGTGTTCCT
TTTATGTCACATGCATTCTATGAGGCTGAGTTGTGAACACCAGGAGTCACAGTCTGT
CTGTCTGTCTGTTACCTGTCTGTCTGCCCACTCCCCTTTCTCTGAGACAAGGTCTG
TGTGTGTCTCTCTGTGTCTCTGTCTCTTGATCTCTGTCTTTCTCTCATTCTCTAAGAC
AAGGGCCTCACTATGTAGCCCTGGCTGTCCT

> Pten-OT.1 (chr1:98296490+98297140) (651bp)

CCTGGGAGTCACCATACCATGTGCATAATAGAATCGAAATGAATCATTTTCAAGTGAC
ACAAAAGACAATCGTATGTGGATATTGATGATTAAAACATAATAATAATAGCTAACTAGT
TTTCCTAAGCAAATAGAATAATGAGATTTCAAGAGAAAGACAAGGCATGGTAAAAAAC
ATTTTCCTAAATTTATATGTTCTTAATTCCAAACAAAACTCTAGAAAATCTTTAAGGGAA
ACTAATATTTAATTTAGTTAGTTAGTTTGTTTGTTTGTAAATGAAAATATTGAGATAGTG
GTTGACCATCATCAAAGAAATCGATAGGCTCATCTTTTATTTTCTCCGTTAATTTAATT
ATTATTTTGACATCCCAATCTTTGCCCTCCCCCCAGTCTCCCTCCTTGCCAGAGTCC
CTCCTTTTAGCCTTCCCCCATATCTCCCCATAAGATAACTTAGCCACTGTCTTTTGTC
CTGTCTCTTTTTTTCTTTCTTTTATATCTCTTTCTTTCAATCTCTCATTTTCCTTTCTTC
CTTCCCTCCTCCCTTCTTTCTTTTCTGCTTCCTTTTATTTTGTGGATTTAGCGTGTCT
GAAACTCTCTATTAAGACAAGATTGGTCTCAAGTCTACAGAGATCTGCCTTTTGCCT
CACTGG

> Pten-OT.2 (chr10:11506417+11506902) (486bp)

AGATGAAACCCTGGAGCAGCTTGGAGTCTGATTGAAGAGCCAAGATAATGAAAATG
ACTCTGCTTTTATTGGTCCCTGTAGGGCATGATCATGTGGTGGGAAATCCTGCCTGA
GGTCTAAGGAGAAGAACTAAGGGATAATTGAATTGCCCTTGACAGTAGCTTTAGAAA
ACCTGAATAGATCTTCCACATATAGCCATCTAAAAATACCGGTCTCTTTGATGATGGA
GGGCCCAATGTTAATTGTTTCCTATCTAGTCATTTCATTTATTAGTTCTTAGTAAAACC
AAAGAGCTGTCATACAATCAAATTTGTATGAAGGGCTCCTATCAGGCACCGAAGTCA
CTGCCTTTTTGAGCTCTAAATGAAAAATAAACTGTCTTCAAGACCCTTTGCAGTGAG
AACTGTGCTGGCAATGGAGCTTATTTTATGAAAACATCTACTCAGAGAAATGCTATTT
CTTTTATAAGGAGTCCCCATCCCTTCA

> Pten-OT.3 (chr6:110090414+110090889) (476bp)

GAAGCCCGGTCTTTGGTGTAGAGGTATTGATGATATAGGAACATTAAGAGTGAAACA
GCTTAGCTGGAAAAATTAAGCCTTTGTAGGTAGTGATGTTTAAAGGGATTAAAGCTTC
TCTTCTAACATCAGATCTTTCAAGTACAAGTTGTCATCAAAGAGGGAACCAGTTCCT
CCATAGTCTCTGCCTTGTCATCTCACTGTGAGATCCTCCCTTCTTCCTTGTATCCCT
GTCACGATGTCTTTGATGAAGGTAAAATGCAGATGGGAGGACCCTCATCAAAGGCT
GTGCAGTGAAACTACACAACCTTACCTTGGAGTTTCATTATCCAGAATGGTTAGGAA
ATAAACTCTTTTTAAATGATTATTATTGGTAAAACCACAGATTTAAATATTTAGTTAGCA
ACAGAAAATGTAATTATGCATTAACCTATTAAAATGATTTTGTTATAAGTACATATGTTA
CCCGTTCTGCCACC

> Apc-OT.1 (chr10:60270103+60270580) (478bp)

AGCTAAGCTGCTTCTGCTCCAGACCACCCAGAAGGACAACTCAAACCCCCCAAGC
TTCTGTGTCCTGCCTACGCCCAGGCTGATGCTGCAAGGCTCTGCCCCACTGGGCT
CCTGGTTCCATCCACAAGGCAGAGCTGAGCTCTGAACGTGAAGGGAGGTAAGACC
GTCTGGGAGCTTCTGGCTTGAGTCTGGGTAGCACTGGGAATCCACACTGTGAAAA
AGCTTACAGGTGACAAGGACAGAGATGCTCTAAGGAGGTCCCAGCTGAATTCCAGC
GTCCCTCCGCTTTCCCAAGCACCAGCTTCCCAGAGAGCTAGAGTCCTGGACCGCA
AGCACATACACACCATGTGGTATAAAAAGAACTCTTTTCAAAAGCAGTCCCCAAGGG

TGGGGGGAAGGGAGGTCGATGAGAAAGGAGCCAGGGGGACAATAGTGAATAGAG
TTAGGATTATTTCCTGAACTCTGACTCCTCGTGGGA

> Apc-OT.2 (chr8:44148887+44149371) (485bp)

GCTTTACATGAGGGTGGCCTCCAGGTTCCTTCCCTGTGTGAGTTCCTGTCCCGACT
TCCTTCAATGATGAACTGTGATGTGAAAGCATAAGCCACATAAACCCTTTCCTCTCC
GAGTTGGTTTGTTTGGTCATGGTGCTTCATCAGAGCAATATTAACCCTAACTCAGATA
AGACCCAATCAATATGCAAAAAAGTCTGCCATGCCTTCACTTTGGGATTTGAAAACA
GGGTAAACTAGCATACTACATGGCTTGATATCCTAGCCCAGAGCAATTTCAAATGTAA
TCATTCTTCATTCATAAAGAAAATATTATATTGAATTTTATTAAGTATTGCTATTTCATCA
TGTTTGACCTATTAAAATAGCTATTTCTAAGATTTACCCTTATGACTCTGTGTTTCTACA
AAGCCATGGCAAATATATTAATAAATAACTTTGTACTGCATAGTATTTGTGGCGAAATG
GAGCCACTTCCAATGAGCA

> Apc-OT.3 (chr14:30797100+30797625) (526bp)

GATGCTACTTGCCGATGGGAGCTATCTATCCATTGATGGATGCATTGTAGCATGCCTT
GTTGGTTTCTCTTTTCACTAAATAGTTTGTACTAATGATCTCTCATTTCTCTCTCCAGA
AACCTGAAAGTAATGAACAAAGACAAAATGGCCTTAAAATGGTGGACCAGGCAATAT
GGTCTAAAAAGATGTCAAACGGTACAAGGCGGATCCCTGTGTCACCAGAACAGGTT
CGGTCAGCATTTCCTTCACGTTGAGTATTTTTCCACTTCTCTCGATATAGCTGTGTGT
TGCGTGATGATGGGATGTGTCATACATCCCTGGAGGAGTTTGTGGTCCATACAGTA
GTGCCTATGACAGAAGACAAACTACTGCATCTTCTGCACACCTGGGCTCTATGGCCT
AGCCCTATTGCTCCAGCCTCCACAAGTAAACAGTATGTCACTATGCAGAATTCTGTA
CCCAGCATTGTCACAAATATGTTCGTATACGTAACACTTAAAACAAAATAGCGTAGGG
ACTGCAGAGT

REFERENCES

Abeshouse, A., Ahn, J., Akbani, R., Ally, A., Amin, S., Andry, C.D., Annala, M., Aprikian, A., Armenia, J., Arora, A., et al. (2015). The Molecular Taxonomy of Primary Prostate Cancer. Cell 163, 1011–1025.

Akbani, R., Akdemir, K.C., Aksoy, B.A., Albert, M., Ally, A., Amin, S.B., Arachchi, H., Arora, A., Auman, J.T., Ayala, B., et al. (2015). Genomic Classification of Cutaneous Melanoma. Cell 161, 1681–1696.

Andersson, A.K., Ma, J., Wang, J., Chen, X., Gedman, A.L., Dang, J., Nakitandwe, J., Holmfeldt, L., Parker, M., Easton, J., et al. (2015). The landscape of somatic mutations in infant MLL-rearranged acute lymphoblastic leukemias. Nat Genet 47, 330–337.

Breese, M.R., Breese, M.R., Liu, Y., and Liu, Y. (2013). NGSUtils: a software suite for analyzing and manipulating next-generation sequencing datasets. Bioinformatics 29, 494–496.

Cancer Genome Atlas Network (2012a). Comprehensive molecular characterization of human colon and rectal cancer. Nature 487, 330–337.

Cancer Genome Atlas Network (2012b). Comprehensive molecular portraits of human breast tumours. Nature 490, 61–70.

Cancer Genome Atlas Research Network (2013). Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. N Engl J Med 368, 2059–2074.

Cancer Genome Atlas Research Network (2014a). Comprehensive molecular characterization of gastric adenocarcinoma. Nature 513, 202–209.

Cancer Genome Atlas Research Network (2014b). Comprehensive molecular profiling of lung adenocarcinoma. Nature 511, 543–550.

Cerami, E., Gao, J., Dogrusoz, U., Gross, B.E., Sumer, S.O., Aksoy, B.A., Jacobsen, A., Byrne, C.J., Heuer, M.L., Larsson, E., et al. (2012). The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. Cancer Discovery 2, 401–404.

Cheung, A.F., Carter, A.M., Kostova, K.K., Woodruff, J.F., Crowley, D., Bronson, R.T., Haigis, K.M., and Jacks, T. (2010). Complete deletion of Apc results in severe polyposis in mice. Oncogene 29, 1857–1864.

Curry, N.L., Curry, N.L., Mino-Kenudson, M., Mino-Kenudson, M., Oliver, T.G., Oliver, T.G., Yilmaz, O.H., Yilmaz, O.H., Yilmaz, V.O., Yilmaz, V.O., et al. (2013). Pten-null tumors cohabiting the same lung display differential AKT activation and sensitivity to dietary restriction. Cancer Discovery 3, 908–921.

Dupage, M., Dooley, A.L., and Jacks, T. (2009). Conditional mouse lung cancer models using adenoviral or lentiviral delivery of Cre recombinase. Nat Protoc 4, 1064–1072.

Farago, A.F., Snyder, E.L., and Jacks, T. (2012). SnapShot: Lung Cancer Models. Cell 149, 246–246.e1.

Feldser, D.M., Kostova, K.K., Winslow, M.M., Taylor, S.E., Cashman, C., Whittaker, C.A., Sanchez-Rivera, F.J., Resnick, R., Bronson, R., Hemann, M.T., et al. (2010). Stage-specific sensitivity to p53 restoration during lung cancer progression. Nature 468, 572–575.

Frese, K.K., and Tuveson, D.A. (2007). Maximizing mouse cancer models. Nat Rev Cancer 7, 645–658.

Fu, Y., Fu, Y., Foden, J.A., Foden, J.A., Khayter, C., Khayter, C., Maeder, M.L., Maeder, M.L., Reyon, D., Reyon, D., et al. (2013). High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. Nat Biotechnol 31, 822–826.

Gao, J., Aksoy, B.A., Dogrusoz, U., Dresdner, G., Gross, B., Sumer, S.O., Sun, Y., Jacobsen, A., Sinha, R., Larsson, E., et al. (2013). Integrative Analysis of Complex Cancer Genomics and Clinical Profiles Using the cBioPortal. Sci. Signal. 6, pl1–pl1.

George, J., Lim, J.S., Jang, S.J., Cun, Y., Ozretić, L., Kong, G., Leenders, F., Lu, X., Fernández-Cuesta, L., Bosco, G., et al. (2015). Comprehensive genomic profiles of small cell lung cancer. Nature 524, 47–53.

Hammerman, P.S., Lawrence, M.S., Voet, D., Jing, R., Cibulskis, K., Sivachenko, A., Stojanov, P., McKenna, A., Lander, E.S., Gabriel, S., et al. (2012). Comprehensive genomic characterization of squamous cell lung cancers. Nature 489, 519–525.

Heigwer, F., Kerr, G., and Boutros, M. (2014). E-CRISP: fast CRISPR target site identification. Nat Methods 11, 122–123.

Hogan, B.L.M., Barkauskas, C.E., Chapman, H.A., Epstein, J.A., Jain, R., Hsia, C.C.W., Niklason, L., Calle, E., Le, A., Randell, S.H., et al. (2014). Repair and Regeneration of the Respiratory System: Complexity, Plasticity, and Mechanisms of Lung Stem Cell Function. Stem Cell 15, 123–138.

Hsu, P.D., Scott, D.A., Weinstein, J.A., Ran, F.A., Konermann, S., Agarwala, V., Li, Y., Fine, E.J., Wu, X., Shalem, O., et al. (2013). DNA targeting specificity of RNA-guided Cas9 nucleases. Nat Biotechnol 31, 827–832.

Jackson, E.L., Willis, N., Mercer, K., Bronson, R.T., Crowley, D., Montoya, R., Jacks, T., and Tuveson, D.A. (2001). Analysis of lung tumor initiation and progression using conditional expression of oncogenic K-ras. Genes Dev 15, 3243–3248.

Jackson, E., Olive, K., Tuveson, D., Bronson, R., Crowley, D., Brown, M., and Jacks, T. (2005). The Differential Effects of Mutant p53 Alleles on Advanced Murine Lung Cancer. Cancer Res 65, 10280–10288.

Juan, J., Muraguchi, T., Iezza, G., Sears, R.C., and Mcmahon, M. (2014). Diminished WNT → β-catenin → c-MYC signaling is a barrier for malignant progression of BRAFV600E-induced lung tumors. Genes Dev 28, 561–575.

Kormish, J.D., Kormish, J.D., Sinner, D., Sinner, D., Zorn, A.M., and Zorn, A.M. (2010). Interactions between SOX factors and Wnt/beta-catenin signaling in development and disease. Dev. Dyn. 239, 56–68.

Kuscu, C., Arslan, S., Singh, R., Thorpe, J., and Adli, M. (2014). Genome-wide analysis reveals characteristics of off-target sites bound by the Cas9 endonuclease. Nat Biotechnol 32, 677–683.

Kwon, M.-C., and Berns, A. (2013). Mouse models for lung cancer. Mol Oncol 7, 165–177.

Li, H., Li, H., Durbin, R., and Durbin, R. (2010). Fast and accurate long-read alignment with Burrows-Wheeler transform. Bioinformatics 26, 589–595.

Li, H., Li, H., Handsaker, B., Handsaker, B., Wysoker, A., Wysoker, A., Fennell, T., Fennell, T., Ruan, J., Ruan, J., et al. (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics 25, 2078–2079.

Lohr, J.G., Stojanov, P., Carter, S.L., Cruz-Gordillo, P., Lawrence, M.S., Auclair, D., Sougnez, C., Knoechel, B., Gould, J., Saksena, G., et al. (2014). Widespread genetic heterogeneity in multiple myeloma: implications for targeted therapy. Cancer Cell 25, 91–101.

Madisen, L., Zwingman, T.A., Sunkin, S.M., Oh, S.W., Zariwala, H.A., Gu, H., Ng, L.L., Palmiter, R.D., Hawrylycz, M.J., Jones, A.R., et al. (2010). A robust and high-throughput Cre reporting and characterization system for the whole mouse brain. Nature Neuroscience 13, 133–140.

McKenna, A., McKenna, A., Hanna, M., Hanna, M., Banks, E., Banks, E., Sivachenko, A., Sivachenko, A., Cibulskis, K., Cibulskis, K., et al. (2010). The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Research 20, 1297–1303.

Moon, R.T., Moon, R.T., Kohn, A.D., Kohn, A.D., De Ferrari, G.V., De Ferrari, G.V., Kaykas, A., and Kaykas, A. (2004). WNT and beta-catenin signalling: diseases and therapies. Nature Publishing Group 5, 691–701.

Pacheco-Pinedo, E.C., Durham, A.C., Stewart, K.M., Goss, A.M., Lu, M.M., Demayo, F.J., and Morrisey, E.E. (2011). Wnt/β-catenin signaling accelerates mouse lung tumorigenesis by imposing an embryonic distal progenitor phenotype on lung epithelium. J Clin Invest 121, 1935–1945.

Pasqualucci, L., Trifonov, V., Fabbri, G., Ma, J., Rossi, D., Chiarenza, A., Wells, V.A., Grunn, A., Messina, M., Elliot, O., et al. Analysis of the coding genome of diffuse large B-cell lymphoma. Nat Genet 43, 830–837.

Patch, A.-M., Christie, E.L., Etemadmoghadam, D., Garsed, D.W., George, J., Fereday, S., Nones, K., Cowin, P., Alsop, K., Bailey, P.J., et al. (2015). Whole–genome characterization of chemoresistant ovarian cancer. Nature 521, 489–494.

Psarras, S., Karagianni, N., Kellendonk, C., Tronche, F., Cosset, F.-L., Stocking, C., Schirrmacher, V., Boehmer Hv, H.V., and Khazaie, K. (2004). Gene transfer and genetic modification of embryonic stem cells by Cre- and Cre-PR-expressing MESV-based retroviral vectors. J Gene Med 6, 32–42.

Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 26, 841–842.

Robinson, D., Van Allen, E.M., Wu, Y.-M., Schultz, N., Lonigro, R.J., Mosquera, J.-M., Montgomery, B., Taplin, M.-E., Pritchard, C.C., Attard, G., et al. (2015). Integrative clinical genomics of advanced prostate cancer. Cell 161, 1215–1228.

Rock, J.R., and Hogan, B.L.M. (2011). Epithelial Progenitor Cells in Lung Development, Maintenance, Repair, and Disease. Annu Rev Cell Dev Biol 27, 493–512.

Sanchez-Rivera, F.J., and Jacks, T. (2015). Applications of the CRISPR-Cas9 system in cancer biology. Nat Rev Cancer 15, 387–395.

Schwank, G., Koo, B.-K., Sasselli, V., Dekkers, J.F., Heo, I., Demircan, T., Sasaki, N., Boymans, S., Cuppen, E., van der Ent, C.K., et al. (2013). Functional repair of CFTR by CRISPR/Cas9 in intestinal stem cell organoids of cystic fibrosis patients. Cell Stem Cell 13, 653–658.

Shalem, O., Sanjana, N.E., Hartenian, E., Shi, X., Scott, D.A., Mikkelsen, T.S., Heckl, D., Ebert, B.L., Root, D.E., Doench, J.G., et al. (2014). Genome-scale CRISPR-Cas9 knockout screening in human cells. Science 343, 84–87.

Snyder, E.L., Watanabe, H., Magendantz, M., Hoersch, S., Chen, T.A., Wang, D.G., Crowley, D., Whittaker, C.A., Meyerson, M., Kimura, S., et al. (2013). Nkx2-1 represses a latent gastric differentiation program in lung adenocarcinoma. Mol Cell 50, 185–199.

Song, M.S., Salmena, L., and Pandolfi, P.P. (2012). The functions and regulation of the PTEN tumour suppressor. Nat Rev Mol Cell Biol 13, 283–296.

Sweet-Cordero, A., Mukherjee, S., Subramanian, A., You, H., Roix, J., Ladd-Acosta, C., Mesirov, J., Golub, T., and Jacks, T. (2005). An oncogenic KRAS2 expression signature identified by cross-species gene-expression analysis. Nat Genet 37, 48–55.

Sweet-Cordero, A., Tseng, G.C., You, H., Douglass, M., Huey, B., Albertson, D., and Jacks, T. (2006). Comparison of gene expression and DNA copy number changes in a murine model of lung cancer. Genes Chromosom. Cancer 45, 338–348.

The Cancer Genome Atlas Research Network (2015). Comprehensive, Integrative Genomic Analysis of Diffuse Lower-Grade Gliomas. N Engl J Med 372, 2481–2498.

Thorvaldsdottir, H., Thorvaldsdottir, H., Robinson, J.T., Robinson, J.T., Mesirov, J.P., and Mesirov, J.P. (2013). Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. Briefings in Bioinformatics 14, 178–192.

Waddell, N., Pajic, M., Patch, A.-M., Chang, D.K., Kassahn, K.S., Bailey, P., Johns, A.L., Miller, D., Nones, K., Quek, K., et al. (2015). Whole genomes redefine the mutational landscape of pancreatic cancer. Nature 518, 495–501.

Wang, K., Wang, K., Li, M., Li, M., Hakonarson, H., and Hakonarson, H. (2010). ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res 38, e164–e164.

Winslow, M.M., Dayton, T.L., Verhaak, R.G.W., Kim-Kiselak, C., Snyder, E.L., Feldser, D.M., Hubbard, D.D., DuPage, M.J., Whittaker, C.A., Hoersch, S., et al. (2011). Suppression of lung adenocarcinoma progression by Nkx2-1. Nature 1–7.

Wu, X., Scott, D.A., Kriz, A.J., Chiu, A.C., Hsu, P.D., Dadon, D.B., Cheng, A.W., Trevino, A.E., Konermann, S., Chen, S., et al. (2014). Genome-wide binding of the CRISPR endonuclease Cas9 in mammalian cells. Nat Biotechnol 32, 670–676.

Xue, W., Chen, S., Yin, H., Tammela, T., Papagiannakopoulos, T., Joshi, N.S., Cai, W., Yang, G., Bronson, R., Crowley, D.G., et al. (2014). CRISPR-mediated direct mutation of cancer genes in the mouse liver. Nature 514, 380–384.

Chapter 3

Systematic identification of *Keap1* mutant-specific vulnerabilities in lung adenocarcinoma through CRISPR-Cas9 genetic screening

Francisco J. Sánchez-Rivera[1,2*], Yadira M. Soto-Feliciano[1,2*], Rodrigo Romero[1,2], Peter M. K. Westcott[1], Arjun Bhutkar[1], Michael T. Hemann[1,2], and Tyler Jacks[1,2,3].

[1] David H. Koch Institute for Integrative Cancer Research, Massachusetts Institute of Technology, Cambridge, MA 02142

[2] Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02142

[3] Howard Hughes Medical Institute, Massachusetts Institute of Technology, Cambridge, MA 02139

[*] These authors contributed equally to this work.

Author Contributions

F.J.S.R., Y.M.S.F., M.T.H. and T.J. designed the study; F.J.S.R. and Y.M.S.F. designed, cloned, sequenced, and screened CRISPR libraries, as well as deconvoluted and analyzed screen data; R.R. generated and validated *Keap1* mutant cell lines, performed initial validation of screen candidates with F.J.S.R. and P.M.K.W., and is actively pursuing gene candidates from the screens; A.B. conducted bioinformatic analyses.

ABSTRACT

The goal of precision oncology is to tailor present and future cancer therapies to specific patients based on the systematic genomic assessment of their tumors. Therefore, large-scale cancer genome sequencing efforts represent an important first step towards achieving this goal. Such studies have unveiled recurrent mutations in *KEAP1* - the main negative regulator of the NRF2 transcription factor - in approximately 20-30% of all lung cancer patients, and in ~ 20% of *KRAS*-mutant lung adenocarcinoma patients. Despite the fact that there is ample evidence suggesting that hyperactivation of the NRF2 pathway plays an important tumor-supporting role in multiple cancer types, including lung cancer and pancreatic cancer, targeted therapies that specifically target this patient subpopulation are currently lacking. To discover novel genetic dependencies that are uniquely essential for cancers harboring *Keap1* mutations, we established and employed a novel CRISPR-Cas9-based experimental pipeline that allows for the systematic interrogation of cancer cell lines harboring mutations frequently observed in human cancer. These studies have led to the identification of multiple *Keap1*-mutant synthetic lethal candidates, including several canonical transcriptional targets of NRF2 and multiple genes involved in glutathione synthesis and utilization, as well as NADPH production. Our results demonstrate the power of CRISPR-based genetic screens for uncovering novel genetic dependencies, such as synthetic lethal interactions, in the context of clinically relevant cancer-associated genotypes. Furthermore, our data lends strong support to efforts aimed at directly or indirectly targeting the NRF2 antioxidant pathway in the context of *KRAS*-mutant lung adenocarcinoma - a lethal subtype for which new targeted therapies are desperately needed.

INTRODUCTION

The goal of precision oncology is to tailor present and future cancer therapies to specific patients based on the systematic genomic assessment of their tumors (Roychowdhury and Chinnaiyan 2014). Therefore, large-scale cancer genome sequencing efforts, such as the ones described extensively in **Part III** of the introduction, represent an important first step towards achieving this goal. Such studies have identified *KEAP1*, the main negative regulator of *NFE2L2* (which encodes for the NRF2 transcription factor), to be mutated in approximately 20-30% of all lung cancers, with loss of function mutations located predominantly within the NRF2 binding domains (see **Figures 3 and 5** of **Part IV** of the introduction) (LUAD TCGA 2014, Hammerman et al., 2012). The pleiotropic NRF2 transcription factor is the master regulator of the cellular antioxidant response, regulating a large repertoire of genes that coordinate multiple cytoprotective pathways, including glutathione (GSH) synthesis and conjugation, thioredoxin (TXN) synthesis and utilization, drug detoxification, and NADPH synthesis, among others (reviewed in Suzuki and Yamamoto 2015). Inactivating mutations in *KEAP1* have been shown to result in constitutive NRF2 activity and activation of the oxidative stress response pathway in cancer cells (Padmanabhan et al., 2006, Singh et al., 2006, Ohta et al., 2008, LUAD TCGA 2014, Hammerman et al., 2012). In addition to *KEAP1* mutations, *NFE2L2* has been shown to be mutated in a significant fraction of all lung cancers, including squamous cancers of the lung (~ 19%) and lung adenocarcinomas (~ 3%) (LUAD TCGA 2014, Hammerman et al., 2012). Remarkably, these mutations are almost invariably found in functionally important domains of the NRF2 protein, including the KEAP1-interacting DLG and ETGE domains, through which KEAP1 exerts its negative

regulation (see **Figures 3-5** of **Part IV** of the introduction) (Shibata et al., 2008a, LUAD TCGA 2014, Hammerman et al., 2012). These studies strongly suggest that hyperactivation of the NRF2 antioxidant program via loss of function mutations in *KEAP1* or gain of function mutations in *NFE2L2* can have major oncogenic properties in the context of lung cancer tumorigenesis and progression. Indeed, multiple studies over the last few years have demonstrated that hyperactivation of the NRF2 antioxidant program (in the presence or absence of somatic mutations) can promote several aspects of the cancer phenotype. For example, DeNicola *et al.* recently demonstrated that potent oncogenes, including $Kras^{G12D}$, $Braf^{V619E}$, and *Myc* can trigger a substantial transcriptional induction of the *Nfe2l2* gene, which in turn promotes ROS detoxification and cellular survival in conditions of high oxidative stress (DeNicola et al., 2011). Furthermore, DeNicola *et al.* demonstrated that genetic inactivation of *Nfe2l2* in GEMMs of pancreatic adenocarcinoma and lung adenocarcinoma substantially impairs the malignant phenotype (DeNicola et al., 2011). NRF2 was also recently shown to play a critical role in promoting multiple anabolic pathways in cancer cells by transcriptionally activating a repertoire of genes involved in the pentose phosphate pathway (PPP), de novo nucleotide biosynthesis pathway, and NADPH production (Mitsuishi et al., 2012b). In addition, recent evidence indicates that the GSH and TXN antioxidant pathways (which are controlled by NRF2) play a critical role in supporting tumor initiation and progression in GEMMs of breast cancer and sarcoma (Harris et al., 2015). Collectively, these and other contemporary studies (Sayin et al., 2014, DeNicola et al., 2015, and Piskounova et al., 2015) have convincingly demonstrated that the NRF2 antioxidant program and cellular antioxidants in general can play a major role in promoting several

aspects of the cancer phenotype. Of direct relevance to this thesis, approximately 20% of lung adenocarcinoma patients that harbor oncogenic *KRAS* mutations concomitantly harbor loss of function mutations in *KEAP1*, implicating the importance of the oxidative stress response pathway in the initiation and/or maintenance of this tumor type (see **Figure 2** of **Part III** of the introduction) (LUAD TCGA 2014). The high frequency of mutation of genes in the Keap1-Nrf2 pathway makes it an attractive target for therapy in lung cancer and other cancer types that exhibit hyperactivation of NRF2. Despite the fact that this is well acknowledged by the cancer research community, effective therapies that specifically target the subpopulation of human patients whose tumors harbor hyperactivation of the NRF2 pathway are currently lacking.

To functionally discover novel genetic dependencies that are uniquely important for cancers harboring hyperactivation of the NRF2 pathway, we have established a novel CRISPR-Cas9-based experimental pipeline that allows for the systematic interrogation of cancer cell lines harboring mutations frequently observed in human cancer (**Figure 1A**). We have employed this pipeline for the discovery of novel genetic dependencies in *KEAP1*-mutant lung adenocarcinoma by performing focused CRISPR-based synthetic lethal screens in murine isogenic $Kras^{G12D/+};p53^{\Delta/\Delta}$ (KP) and $Kras^{G12D/+};p53^{\Delta/\Delta};Keap1^{\Delta/\Delta}$ (KPK) cell lines. These studies have led to the identification of multiple *Keap1*-mutant synthetic lethal candidates, including several *bona fide* transcriptional targets of Nrf2 and genes involved in glutathione synthesis and utilization, as well as NADPH production. Our results demonstrate the power of CRISPR-based genetic screens for uncovering novel genetic dependencies, such as synthetic lethal interactions, in the context of clinically relevant cancer-associated genotypes. Furthermore, our data lends

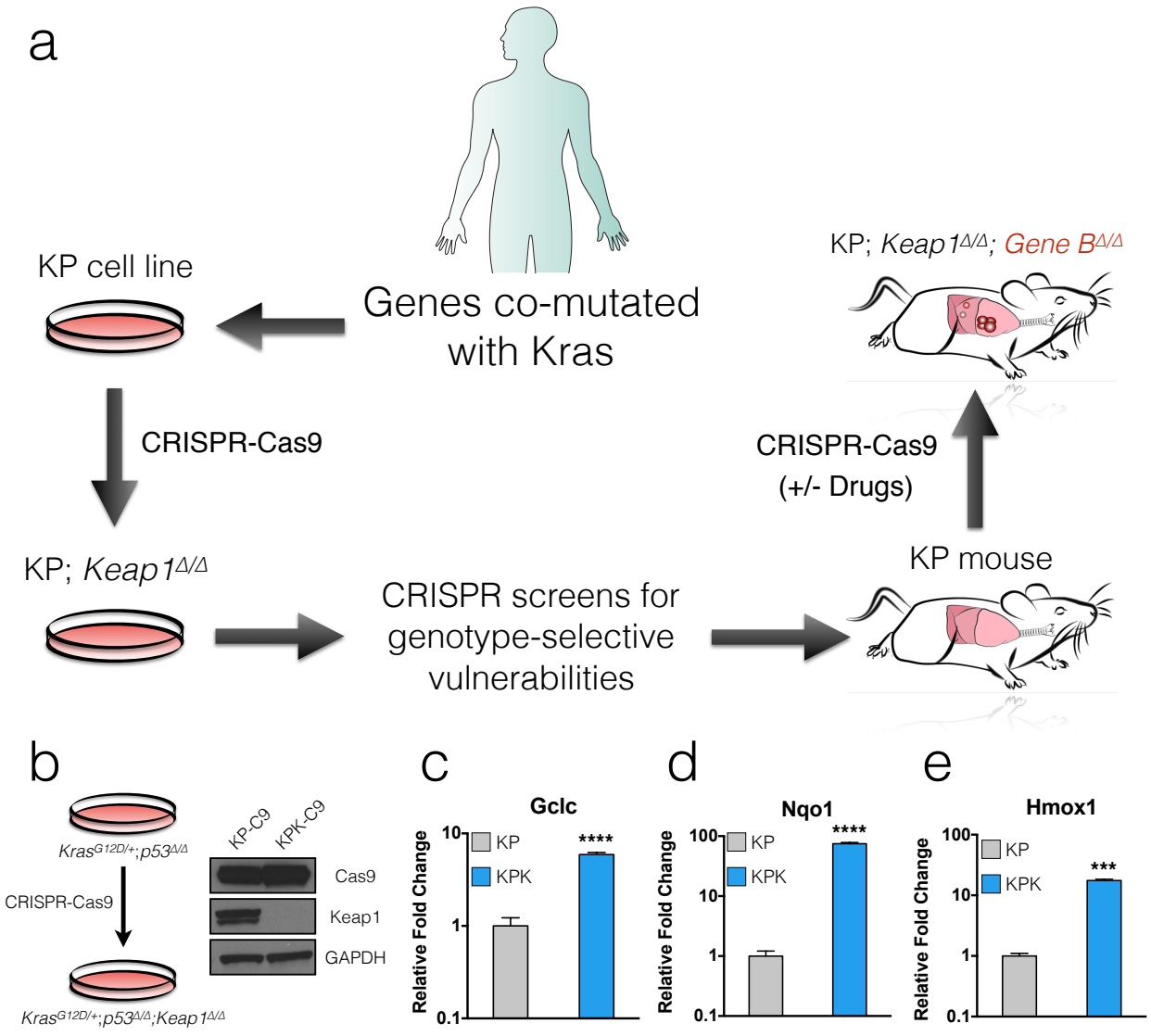


Figure 1: Uncovering genotype-specific vulnerabilities using CRISPR-Cas9. (a) Scheme for our experimental approach. Isogenic KP cell lines engineered to contain clinically-relevant mutations are systematically interrogated using CRISPR-Cas9 to uncover genotype-specific vulnerabilities, which can subsequently be validated *in vivo* using pre-clinical GEMMs of lung adenocarcinoma. (b) Generation of isogenic KP and KPK cell lines using CRISPR-Cas9. Engineered isogenic cell lines were subsequently transduced with a lentiviral vector that expresses Cas9. (c-e) Validation of KPK cell lines. KPK cell lines show evidence of hyperactivation of the Nrf2 pathway, as gauged by the substantial transcriptional induction of (c) *Gclc,* (d) *Nqo1*, and (e) *Hmox1*. Data represent the mean ± S.E.M, n=3. Statistics were calculated with two-sided Student's t-test: ***P < 0.001, ****P < 0.0001.

strong support to efforts aimed at directly or indirectly targeting the Nrf2 antioxidant pathway in the context of *Kras*-mutant lung adenocarcinoma - a lethal subtype for which new targeted therapies are desperately needed.

RESULTS

Uncovering genotype-specific vulnerabilities using CRISPR-Cas9

We have taken a two-tiered CRISPR-based approach for systematically identifying genes that are selectively required for the proliferation and survival of Keap1-mutant lung adenocarcinoma cell lines (**Figure 1A**). Building on cell lines derived from the well studied KP GEMM of lung adenocarcinoma, we have engineered a panel of isogenic tumor cell lines harboring either $Kras^{G12D/+};p53^{\Delta/\Delta}$ (KP) or $Kras^{G12D/+};p53^{\Delta/\Delta}$;$Keap1^{\Delta/\Delta}$ (KPK) mutant genotypes using the CRISPR-Cas9 system (**Figure 1B**). Importantly, we have demonstrated that the Nrf2 pathway is hyperactivated in these cells, as gauged by the substantial transcriptional induction of *Gclc*, *Nqo1*, and *Hmox1* - all of which are canonical Nrf2 transcriptional targets (Suzuki and Yamamoto 2015) (**Figure 1C-E**). In addition, these KPK cell lines have increased Nrf2 levels and are highly resistant to multiple inducers of oxidative stress, including dimethyl fumarate (DMF) (R.R., unpublished observations). Collectively, these results indicate that this pair of KP and KPK isogenic cell lines represent an excellent model in which to study the consequences of hyperactivation of the Nrf2 pathway, as well as to uncover potential unique vulnerabilities in this biological setting.

A novel sgRNA library targeting the Druggable Genome

To identify potential genetic dependencies in KPK cell lines, we designed a custom library of ~ 12,000 sgRNAs targeting ~ 2,000 "druggable" genes (at ~ 6 sgRNAs/gene), as well as 500 non-targeting control sgRNAs, which we refer to as the "Druggable Genome" library (**Figure 2A**). This library was designed by implementing the recently described Broad Institute sgRNA Designer algorithm (Doench et al., 2014; also refer to the **Methods** section). We employed Gibson cloning to assemble the pool of ~ 12,500 sgRNAs into pUSCG, a novel lentiviral vector developed in-house that co-expresses an sgRNA and GFP (**Figure 2B**). To ensure that our batch-cloning protocol of sgRNAs into pUSCG was efficient, we performed high-throughput sequencing of the plasmid libraries (**Figure 2C**). Notably, 100% of the sgRNA sequences were present at an adequate representation, and > 98% of the sgRNAs were within a 5-fold range of the mean (log_2 counts), demonstrating that the ~ 12,000 sgRNAs were similarly represented in the library (**Figure 2C**). To carry out CRISPR-based screens in isogenic KP and KPK cell lines, we infected Cas9-expressing cells with the Druggable Genome library at a low MOI (~ 0.3; titrated using the GFP reporter built into pUSCG; see **Methods**) to ensure a single viral integration per cell (**Figure 2D**).

CRISPR screens to uncover novel drug targets in *Keap1*-mutant lung adenocarcinoma

To ensure that our screening conditions would allow for accurate detection of consistently depleted sgRNAs over time with high sensitivity, we maintained an optimal coverage of ~ 1,000X at every step of the screen. Therefore, each replicate of the

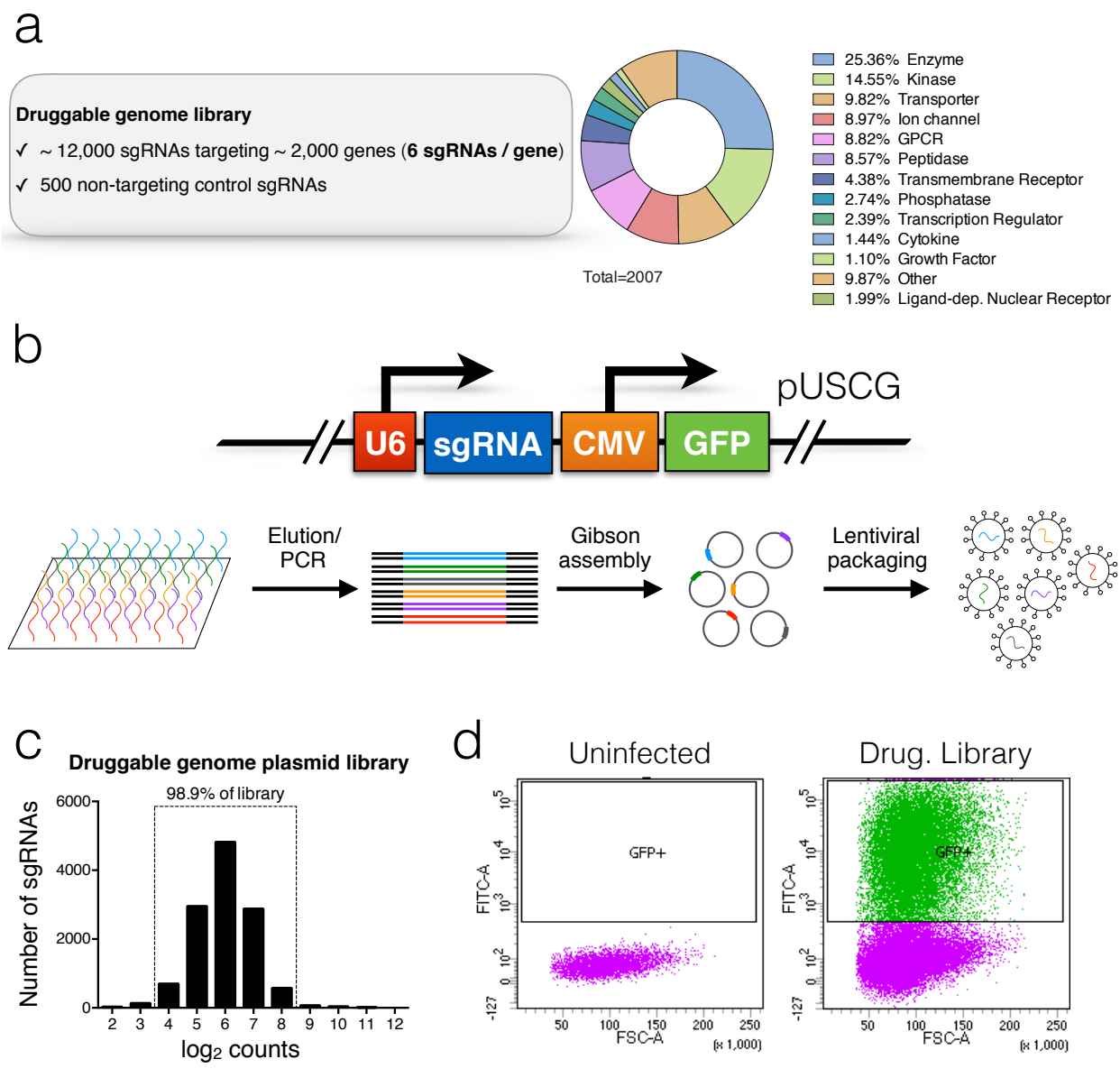


Figure 2: A novel sgRNA library targeting the Druggable Genome. (a) Left: sgRNA content and coverage of the Druggable Genome library. Right: Distribution of protein categories within the Druggable Genome library. (b) Top: pUSCG (U6-sgRNA-CMV-GFP) lentiviral vector. Bottom: Strategy of sgRNA library construction and cloning into pUSCG. (c) Histogram of sgRNA representation in the Druggable Genome plasmid library, showing that > 98% of the constructs are within 5-fold of the mean (log_2 counts). (d) Representative example of flow cytometric analysis of cells infected with the Druggable Genome sgRNA library.

screen consisted of a minimum of ~ 12 million cells at every step of the screen (12,000 sgRNAs vs 1,000X coverage). Importantly, to address the possibility that a subset of targets could be selectively required in the context of whole organism physiology and the presence of an intact immune system, we also carried out pooled screens *in vivo* utilizing subcutaneous transplantation into immunocompetent C57BL/6JN mice. Therefore, the Druggable Genome screens presented here consisted of a total of 18 samples: 6 early *in vitro* samples (3 independent infections per isogenic cell line), 6 final *in vitro* samples (harvested after ~ 20 cumulative population doublings), and 6 final *in vivo* samples (3 per isogenic cell line, subcutaneously injected on Day 0 of the screen and harvested on the same day as the late/final *in vitro* samples) (see **Figure 3A** for experimental scheme of the screen). For screen deconvolution by high-throughput sequencing, we employed a nested-PCR strategy to amplify the integrated sgRNAs from the genomic DNA isolated from all samples (**Figure 3B** and **Methods** section).

Independent screen replicates are very well correlated

To assess the quality of the screens and determine whether any sgRNAs were enriching or depleting throughout the screen, we analyzed the correlation between the $\log_2$ normalized sequencing counts among all independent replicate samples from both *in vitro* and *in vivo* screens (**Figures 4-5**). Importantly, the correlation between multiple independent early/input *in vitro* replicates was excellent, with correlation values ranging from r = 0.94 - 0.96 (**Figure 4**). This indicates that all independent viral transductions were highly consistent, which suggests that the sgRNA library was consistently and appropriately represented among all independent technical replicates from both KP and

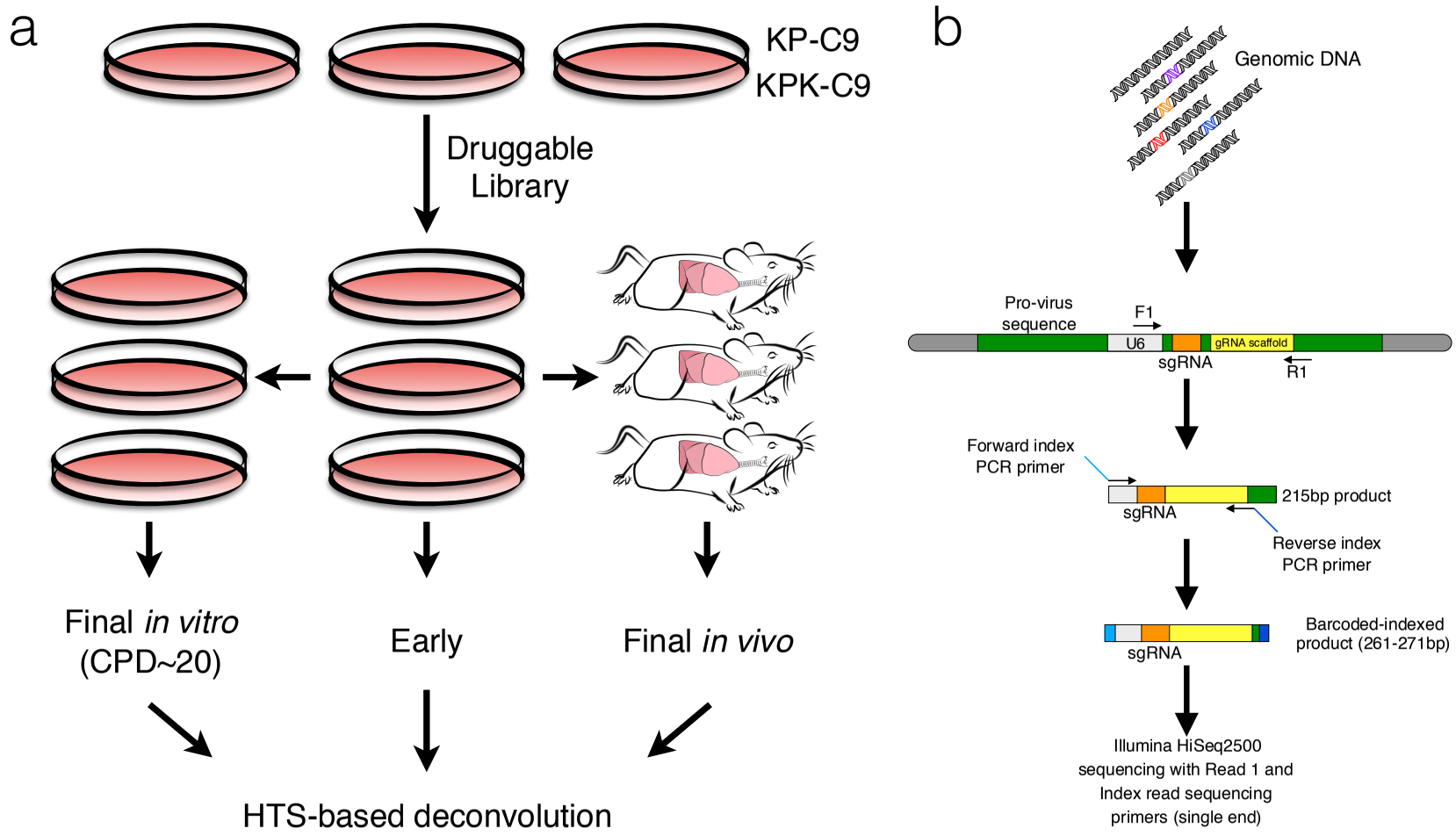


Figure 3: CRISPR screens to uncover novel drug targets in Keap1-mutant lung adenocarcinoma. (a) Screen setup. Briefly, three independent infection replicates per KP or KPK cell line are screened in parallel *in vitro* and *in vivo.* The end point of both screens corresponds to the day when the *in vitro* samples have accumulated ~ 20 cumulative population doublings (CPDs). See text and methods section for more details. (b) Screen deconvolution. See text and methods section for more details.

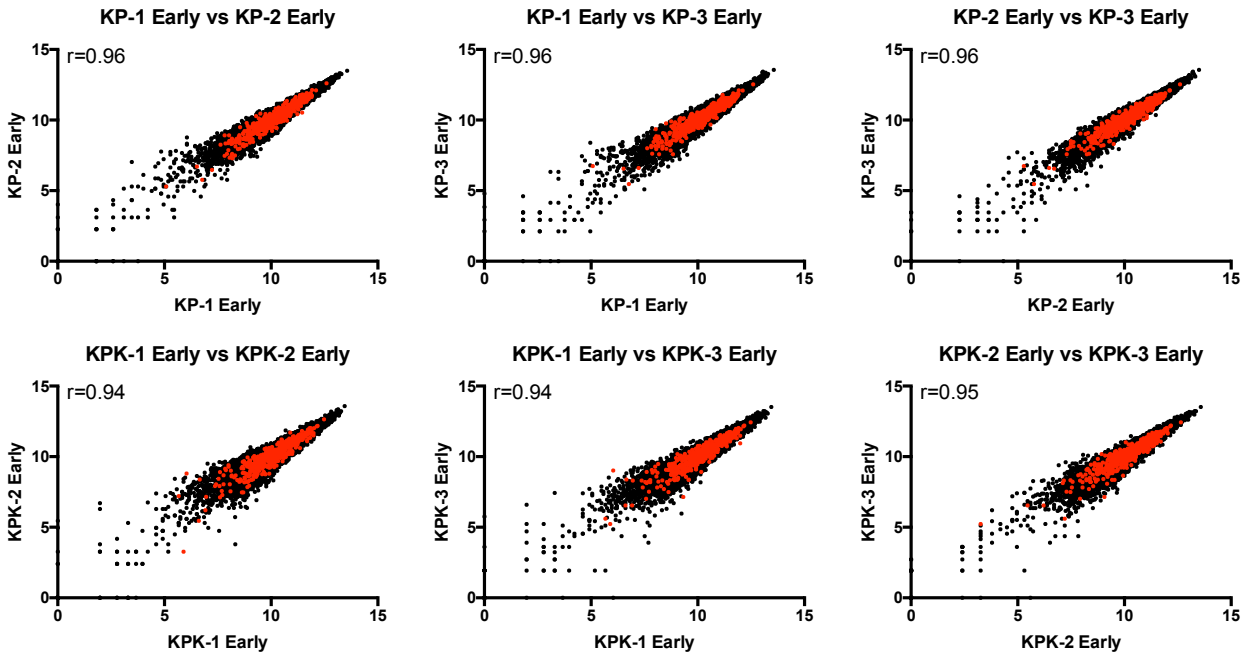


Figure 4: Early independent technical replicates are very well correlated. Top: Scatter plots of $\log_2$ normalized read counts between multiple independent early/input KP infection replicates. Bottom: Scatter plots of $\log_2$ normalized read counts between multiple independent early/input KPK infection replicates. This data suggests that the Druggable Genome library was equally and consistently represented among multiple independent infection replicates. Red dots correspond to the internal set of 500 control sgRNAs.

KPK arms of the screen. Similarly, the correlation between multiple independent late/final *in vitro* samples was very good, with correlation values ranging from r = 0.67 - 0.76 (**Figure 5A-B**). Perhaps not surprisingly, the correlation between multiple independent final *in vivo* samples was lower than their *in vitro* counterparts, with correlation values ranging from r = 0.52 - 0.71 (**Figure 5C-D**). Nevertheless, the fact that some of the *in vivo* replicates had correlation values in a similar range to their *in vitro* counterparts suggests that complex libraries of up to ~ 12,000 reagents can be moderately represented in the *in vivo* setting.

Interestingly, when we compared multiple early/input *in vitro* replicates with their corresponding late/final replicates, we observed that the correlation values were remarkably decent (but still markedly lower than the early/input vs early/input correlation values), with values ranging from r = 0.81 - 0.84 (**Figure 6A-B**), which suggests at least four things: 1) the experimental variability *within* a given technical replicate set was lower than *between* different technical replicates; 2) sgRNA molecules did not just randomly drift throughout the screen (if this was the case, the correlation values would be very poor); 3) the Druggable Genome library does not merely consist of sgRNAs targeting essential genes; and 4) a relatively small number of sgRNAs are undergoing enrichment or depletion during the course of the experiment. Evidence for the latter comes from the fact that the set of 500 internal control sgRNAs remained tightly correlated (red dots in **Figures 4-6**), demonstrating that the changes observed in the distribution of non-control sgRNAs were not due to random sgRNA fluctuations due to experimental artifacts (such as improper passaging of the cells *in vitro*). Indeed, populations of cells from both arms of the screen proliferated identically within technical

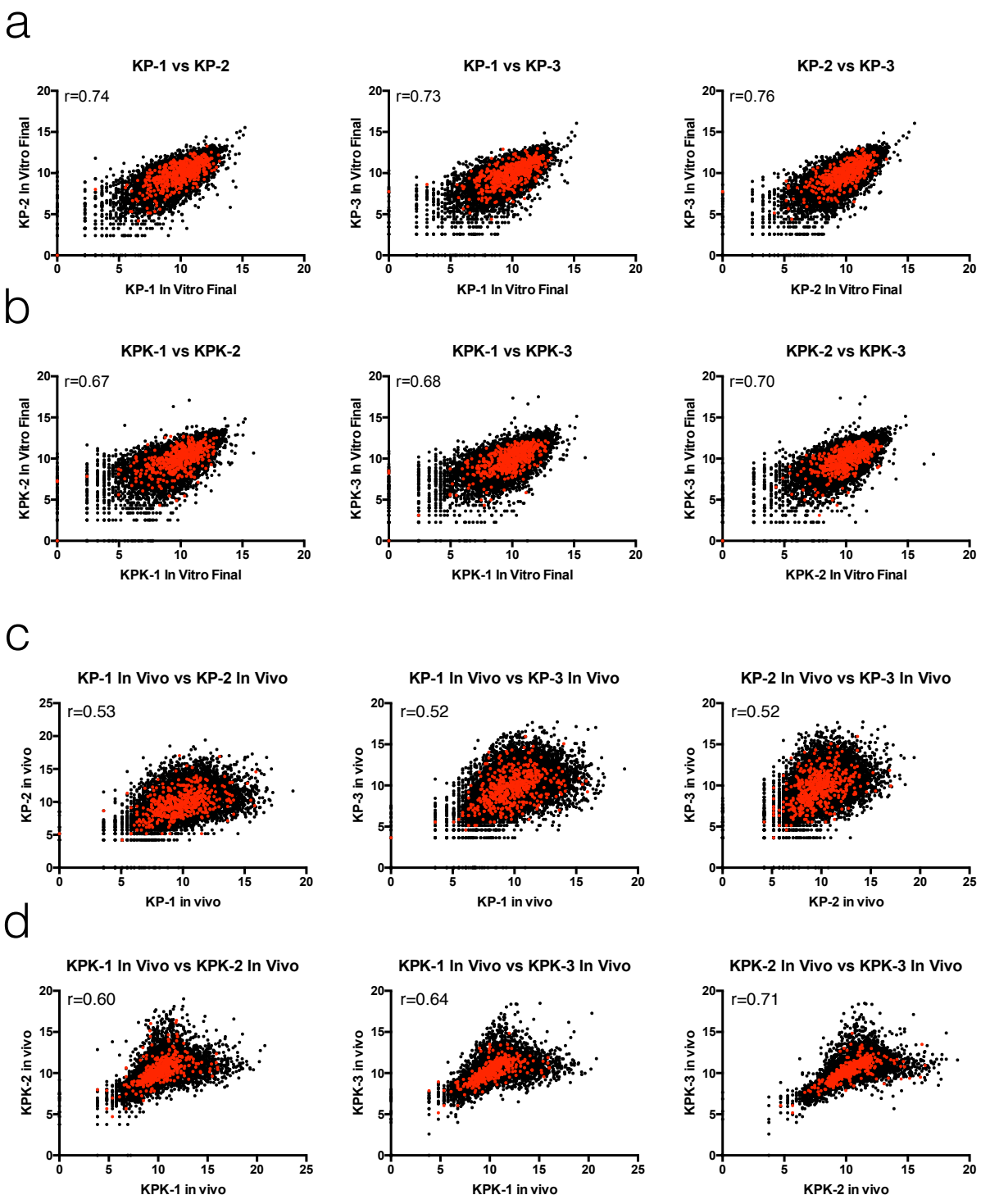


Figure 5: Late/final *in vitro* samples and *in vivo* samples are very well correlated. (a-b) Scatter plots of $\log_2$ normalized read counts between multiple independent late/final KP (a) and KPK (b) *in vitro* samples. (c-d) Scatter plots of $\log_2$ normalized read counts between multiple independent KP (c) and KPK (d) *in vivo* samples. Red dots correspond to the internal set of 500 control sgRNAs.

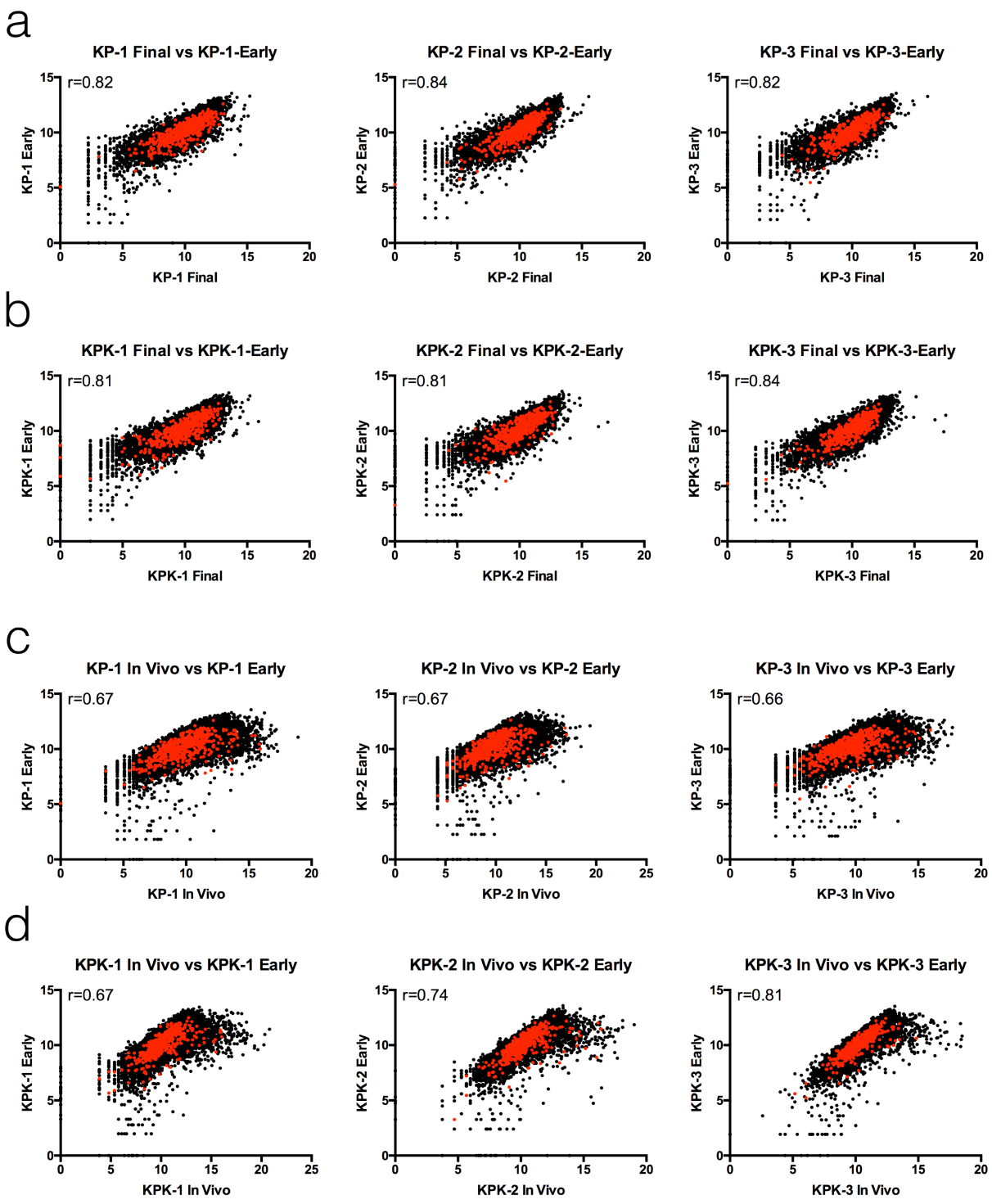


Figure 6: Control sgRNAs remain largely unchanged throughout the screens. (a-b) Scatter plots of $\log_2$ normalized read counts between matched (Final vs Early) KP (a) and KPK (b) *in vitro* samples. (c-d) Scatter plots of $\log_2$ normalized read counts between matched (Final *in vivo* vs Early) KP (c) and KPK (d) *in vivo* samples. Red dots correspond to the internal set of 500 control sgRNAs.

replicates (see **Methods**). Similarly, when we compared multiple *in vivo* replicates with their corresponding early/input *in vitro* replicates, we observed that the correlation values were also decent, with values ranging from r = 0.66 - 0.81, suggesting, again, that intra-replicate variability was significantly lower than inter-replicate variability (**Figure 6C-D**). Remarkably, control sgRNAs also remained largely unchanged in the *in vivo* setting, confirming once again that complex libraries can be represented in this *in vivo* context. As an additional quality control metric, correlation plots generated using the median $\log_2$ sgRNA values calculated across independent technical replicates from both KP and KPK *in vitro* screens have excellent correlation values (r = 0.88 and 0.89, respectively) and clearly show minimal variation of all 500 control sgRNAs (**Figure 7**). Collectively, this data indicates that the technical quality of these genetic screens was excellent.

CRISPR screens reveal multiple known tumor suppressor genes and growth promoting genes

We employed a multi-pronged bioinformatic approach to identify gene candidates that were acting as tumor suppressors or growth promoters in both KP and KPK cells. To account for any possible biological noise inherent to these CRISPR screens, we utilized a form of Blind-Source Separation commonly known as Independent Component Analysis (ICA) (Bhutkar et al., 2015). The main advantage of this analysis is that it has the ability to distinguish the true signal among the noise, and it can also isolate gene signatures that reflect significant changes between conditions. When we applied ICA to

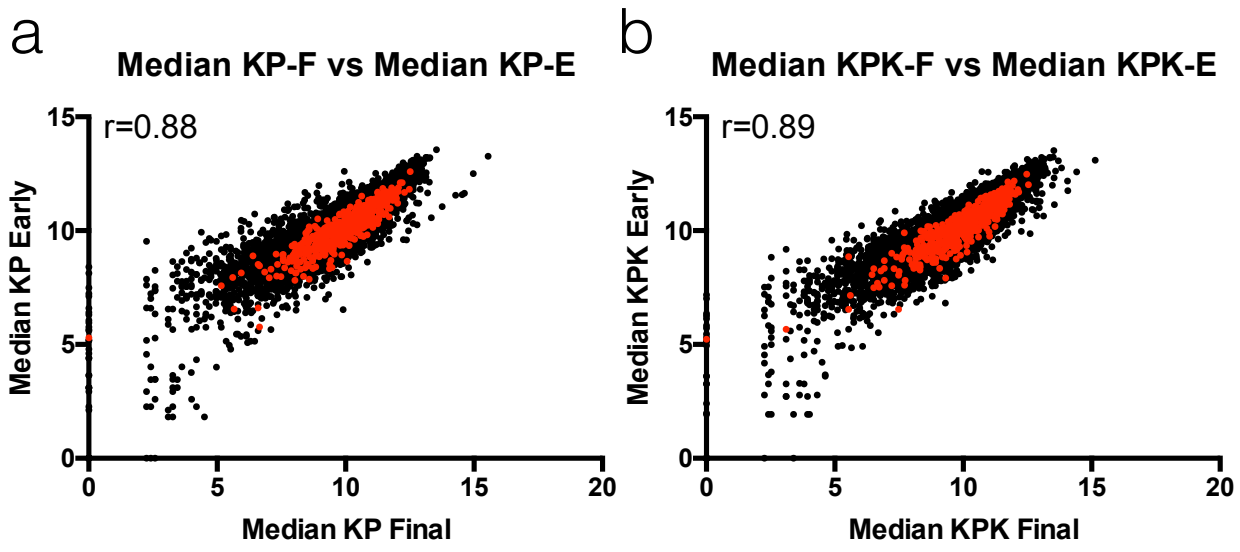


Figure 7: A subset of sgRNAs enrich or deplete throughout the course of the screen. Scatter plots of median $\log_2$ normalized read counts between Final and Early time points from KP (a) and KPK (b) screens demonstrate that control sgRNAs (red dots) remain largely unchanged throughout the screen whereas a subset of sgRNAs (black dots) appear to enrich or deplete throughout the screen. This data suggests that the technical quality of the screens was excellent, and that there is a small subset of potential biologically-relevant sgRNAs that are enriching or depleting during the course of both KP and KPK screens.

the *in vitro* screens (by comparing the final representation of sgRNAs with the initial representation in the early/input samples), we uncovered one major independent component (IC1) for each independent screen that effectively separated the early/input samples from the final *in vitro* samples (**Figure 8A, C**). Remarkably, plotting the IC1 signature z-scores derived from either KP or KPK screens (which reflects the extent of enrichment and depletion for all sgRNAs targeting a given gene when compared to its respective early/input time points) readily identified multiple known tumor suppressor genes and growth promoters (**Figure 8B, D**). sgRNAs targeting Ephrin type-A receptor 2 (*Epha2*) and Tyrosine-protein phosphatase non-receptor type 12 (*Ptpn12*) were consistently enriched across multiple independent technical replicates in both KP and KPK screens (**Figure 8E-F**). On the other hand, sgRNAs targeting the mechanistic target of rapamycin (*mTOR*) serine/threonine-protein kinase were consistently depleted across multiple independent replicates in both KP and KPK screens (**Figure 8G**). Interestingly, *Epha2* was recently demonstrated to be a potent tumor suppressor gene *in vivo* in $Kras^{G12D}$-driven lung adenocarcinoma mouse models by Inder Verma's group (Yeddula et al., 2015). Similarly, *Ptpn12* was demonstrated to be a potent tumor suppressor gene in triple negative breast cancer by Thomas Westbrook, Stephen Elledge, and colleagues back in 2011 (Sun et al., 2011). On the other hand, *mTOR* is a well established positive regulator of cellular growth and it is known to play a major tumor supporting role across multiple types of human cancers (reviewed in Zoncu et al., 2010). As an additional set of quality control metrics for each independent CRISPR screen, we performed gene set enrichment analysis (GSEA) (Subramanian et al., 2005) on both KP and KPK IC1 signatures and observed that gene sets composed of well

a
KP Screen
Early Samples
Final Samples
KP-1
KP-2
KP-3
KP-1
KP-2
KP-3
b
Sig. z-score
Epha2
Ptpn12
Mtor
c
KPK Screen
Early Samples
Final Samples
KPK-1
KPK-2
KPK-3
KPK-1
KPK-2
KPK-3
d
Sig. z-score
Epha2
Rb1
Ptpn12
Mtor
e
Epha2 sgRNAs
KPK
KP
Change in sgRNA Abundance (log2)
f
Ptpn12 sgRNAs
KPK
KP
Change in sgRNA Abundance (log2)
g
Mtor sgRNAs
KPK
KP
Change in sgRNA Abundance (log2)

Figure 8: CRISPR screens reveal multiple known tumor suppressor genes and growth promoting genes. (a) Hinton diagram representation of ICA-derived signatures from the KP DNA sequencing dataset comprised of late/final *in vitro* and early/input samples. Columns denote signatures and rows denote samples. Colors denote relative directionality of gene representation (red = enrichment, green = depletion) and the size of each square represents the magnitude of the contribution of each sample to the respective individual component (IC). Each signature is two-sided. Vertical boxes denote statistically significant (p=0.01, Mann-Whitney test) ICs. IC1 identified a gene signature that significantly distinguishes the early/input samples from the final *in vitro* samples. (b) Waterfall plot of IC1 z-scores. Blue = significantly enriched genes. Red = significantly depleted genes. (c) Hinton diagram representation of ICA-derived signatures from the KPK DNA sequencing dataset comprised of late/final *in vitro* and early/input samples. IC1 identified a gene signature that significantly distinguishes the early/input samples from the final *in vitro* samples. (d) Waterfall plot of IC1 z-scores. Blue = significantly enriched genes. Red = significantly depleted genes. (e-g) Bar graphs denoting the change in the abundance of individual sgRNAs targeting (e) *Epha2*, (f) *Ptpn12*, and (g) *Mtor* in both KP and KPK screens.

described essential genes, such as those involved in tRNA biosynthesis, tRNA aminoacylation, and other classes of essential genes, were significantly down-regulated in both datasets (**Figure 9**). The fact that we successfully identified a number of well described tumor suppressor genes and growth promoters in a relatively unbiased way, and that they scored among the top gene candidates in both screens strongly supports the validity of our platform for uncovering novel cancer-associated genes, including those that counteract or promote multiple hallmarks of the cancer phenotype.

CRISPR screens uncover multiple putative *Keap1*-mutant synthetic lethal candidates

We employed two bioinformatic approaches - one at the level of individual genes and one at the level of gene sets - to determine whether our screens had uncovered any putative *Keap1* mutant-specific genetic vulnerabilities.

For gene-level analyses, we first computed a gene-level score (which we refer to as the Gene Score) by calculating the median $\log_2$ fold change in the abundance of all sgRNAs targeting a given gene during the culture period (early/input samples vs final samples) (Wang et al., 2014). We then computed a delta gene-level score between KPK and KP samples (which we refer to as the Differential Gene Score) by subtracting the KP Gene Score from the KPK Gene Score (Birsoy et al., 2015). A negative Differential Gene Score for a given gene suggests that the gene is a potential *Keap1* mutant-specific genetic dependency. In other words, a negative KPK-KP Differential Gene Score would predict that genetic or pharmacological inactivation of the gene candidate would preferentially *decrease* the fitness of KPK cells. On the other hand, a positive

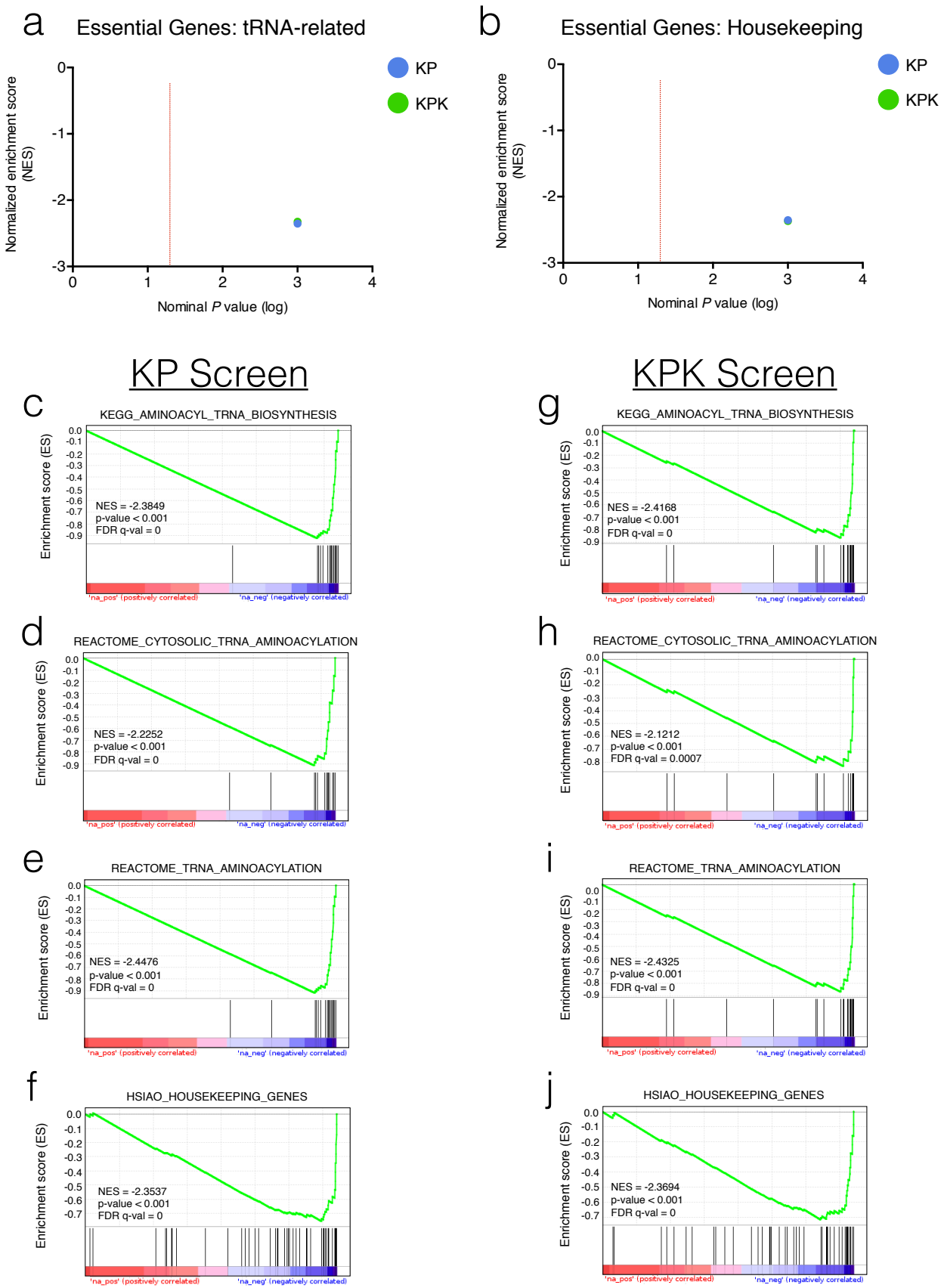

a
Essential Genes: tRNA-related
KP
KPK
Normalized enrichment score (NES)
Nominal P value (log)
b
Essential Genes: Housekeeping
KP Screen
KPK Screen
c
KEGG_AMINOACYL_TRNA_BIOSYNTHESIS
NES = -2.3849
p-value < 0.001
FDR q-val = 0
g
KEGG_AMINOACYL_TRNA_BIOSYNTHESIS
NES = -2.4168
p-value < 0.001
FDR q-val = 0
d
REACTOME_CYTOSOLIC_TRNA_AMINOACYLATION
NES = -2.2252
p-value < 0.001
FDR q-val = 0
h
REACTOME_CYTOSOLIC_TRNA_AMINOACYLATION
NES = -2.1212
p-value < 0.001
FDR q-val = 0.0007
e
REACTOME_TRNA_AMINOACYLATION
NES = -2.4476
p-value < 0.001
FDR q-val = 0
i
REACTOME_TRNA_AMINOACYLATION
NES = -2.4325
p-value < 0.001
FDR q-val = 0
f
HSIAO_HOUSEKEEPING_GENES
NES = -2.3537
p-value < 0.001
FDR q-val = 0
j
HSIAO_HOUSEKEEPING_GENES
NES = -2.3694
p-value < 0.001
FDR q-val = 0
Enrichment score (ES)
'na_pos' (positively correlated)
'na_neg' (negatively correlated)

Figure 9: Gene set enrichment analysis (GSEA) of IC1 signatures from both KP and KPK screens corroborate the role of multiple known essential gene sets in both settings. (a-b) Multiple gene sets corresponding to (a) tRNA-related genes (n=3) and (b) house keeping genes (n=1) are substantially and equally depleted in both KP and KPK screens. The plot represents the normalized enrichment score (NES) against nominal *p* values (log) and the red line denotes threshold for significantly ($p < 0.05$) depleted gene sets. (c-j) Individual gene sets corresponding to tRNA-related gene sets (n=3) and house keeping genes (n=1) from KP (c-f) and KPK (g-j) screens used for generating the plots above. NES: normalized enrichment score. FDR: false discovery rate.

Differential Gene Score for a given gene suggests that the gene acts as a potential *Keap1* mutant-specific tumor suppressor gene. In other words, a positive KPK-KP Differential Gene Score would predict that genetic or pharmacological inactivation of the gene candidate would preferentially *increase* the fitness of KPK cells. Neutral Differential Gene Scores suggest that the gene plays a similar role in both KP and KPK cells. To maximize the chances of identifying true synthetic lethal interactions with *Keap1* mutation, we performed an additional filtering step to exclusively focus on genes for which the distribution of sgRNAs remained highly constant throughout the screen (see **Methods**). This analysis revealed multiple putative *Keap1* mutant-specific genetic vulnerabilities (**Figure 10-11**). Remarkably, our gene-level approach identified at least two direct Nrf2 transcriptional targets (*Fth1* and *Gstp2*) and at least two genes involved in the synthesis and utilization of glutathione (*Gstp2* and *Slc13a3*) (Pietsch et al., 2003, Chanas et al., 2002, and Schorbach et al., 2013). In addition to these candidates, our analysis uncovered at least one protein tyrosine phosphatase (*Ptprh*) that is known to undergo oxidation-dependent conformational changes via reactive cysteine residues, as well as a metabolic enzyme (*Pcx*) whose suppression has been demonstrated to result in markedly reduced intracellular NADPH levels (Wälchli et al., 2005, Xu et al., 2008, Singh et al., 2008). Collectively, these results indicate that multiple putative *Keap1* mutant synthetic lethal candidates directly or indirectly converge on the Nrf2 antioxidant program and cellular antioxidant pathways in general, underscoring the validity of our approach for uncovering *Keap1* mutant-specific genetic vulnerabilities.

For gene set-level analyses, we performed GSEA on both KP and KPK IC1 signatures in order to determine whether any gene sets were preferentially depleted in one

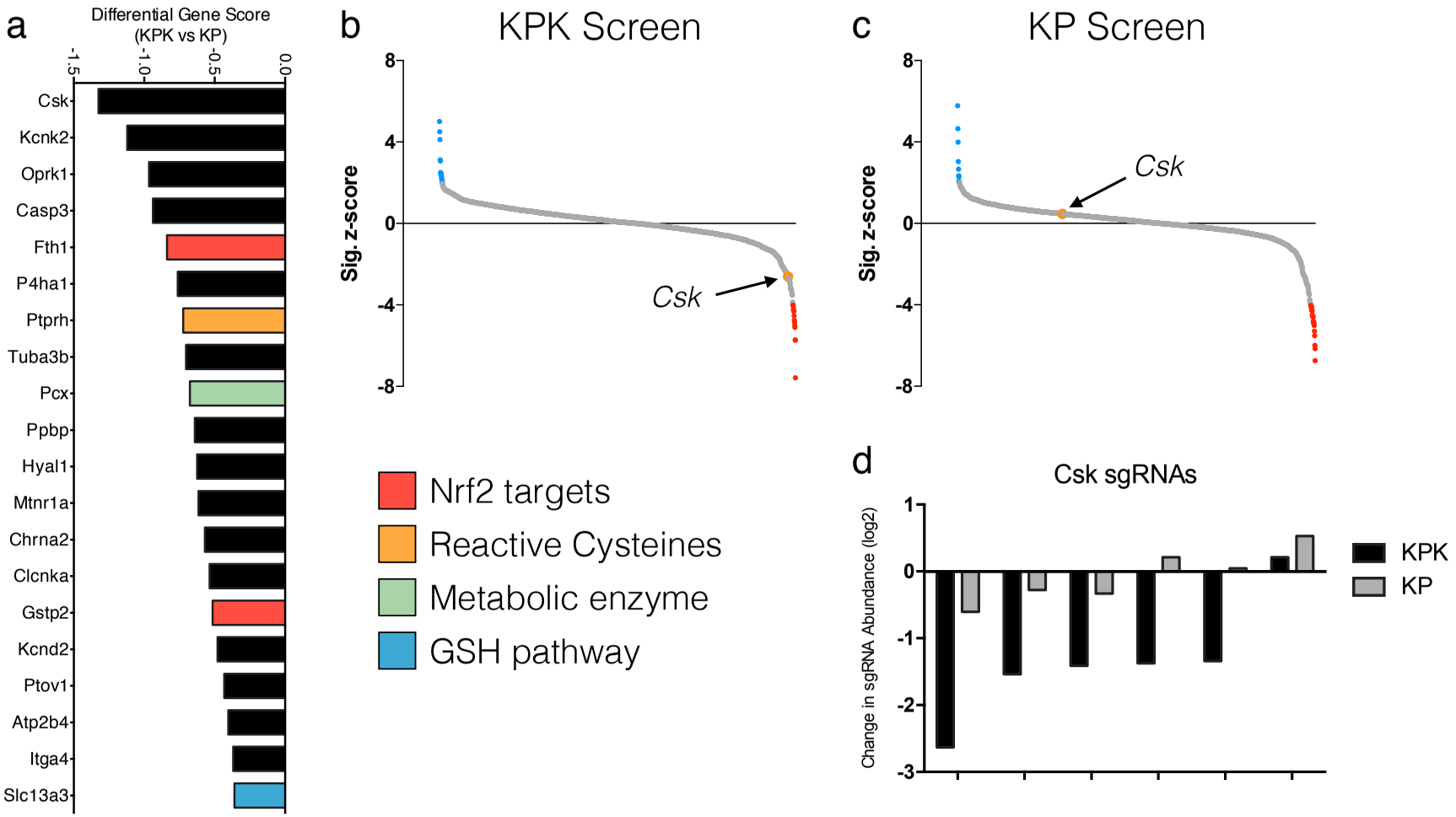


Figure 10: CRISPR screens uncover multiple putative *Keap1*-mutant synthetic lethal candidates. (a) Differential Gene Score values for the top 20 genes with the lowest KPK-KP scores. (b-c) Waterfall plots of KPK (b) and KP (c) IC1 z-scores. The gene with the lowest KPK-KP Differential Gene Score, *Csk* (orange dot in each of the waterfall plots), substantially depletes in KPK cells but not in KP cells. (d) Bar graphs denoting the change in the abundance of individual sgRNAs targeting *Csk* in both KP and KPK screens.

particular genotype. Strikingly, while gene sets related to mTOR and cell cycle/ proliferation were equally depleted in both settings, as expected (**Figure 12A-B**), gene sets related to amino acid metabolism and the oncogenic Myc transcriptional network were significantly more depleted in KPK cells (**Figure 12C-D**). Notably, multiple studies over the last few years have demonstrated that the Nrf2 transcription factor can promote metabolic reprogramming to support anabolic pathways by redirecting glucose and glutamine metabolism via the concerted action of multiple direct Nrf2 transcriptional targets (Mitsuishi et al., 2012, Singh et al., 2013). Moreover, a recent study demonstrated that Nrf2 regulates genes encoding for key serine and glycine biosynthesis enzymes to support GSH and nucleotide biosynthesis (DeNicola., 2015). On the other hand, multiple studies have implicated the Myc transcription factor in promoting *de novo* synthesis of GSH and indirectly contributing to NADPH generation (David et al., 2010, Le et al., 2010). These results indicate that multiple gene sets, including those that contain genes involved in amino acid metabolism, as well as direct and indirect Myc transcriptional targets, might be highly informative for uncovering additional putative *Keap1* mutant synthetic lethal candidates.

DISCUSSION

By employing a novel CRISPR-Cas9-based experimental pipeline that allows for the systematic interrogation of cancer cell lines harboring clinically relevant mutations, we have uncovered multiple putative *Keap1*-mutant synthetic lethal candidates and candidate gene sets, including multiple canonical transcriptional targets of Nrf2, as well as genes involved in glutathione synthesis and utilization, NADPH production, amino

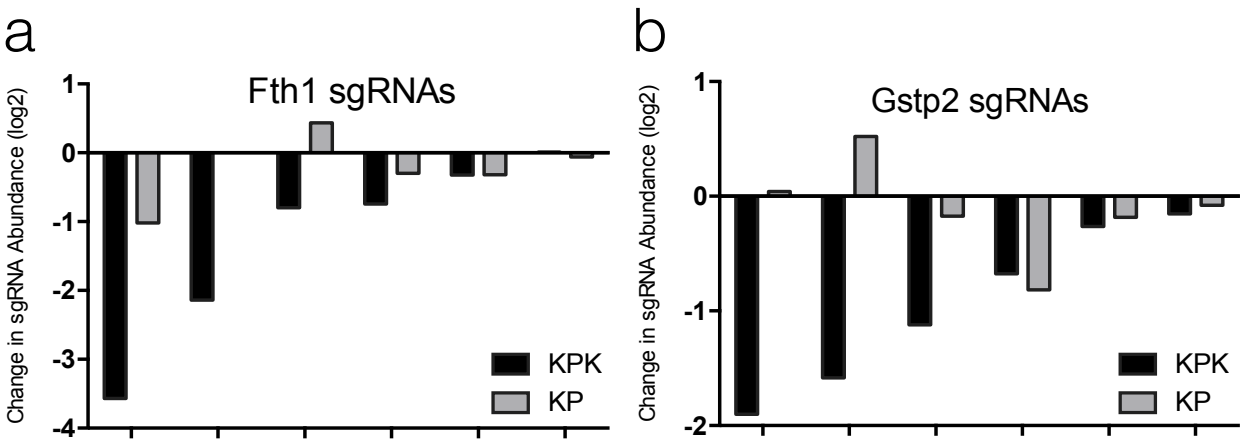


Figure 11: *Fth1* and *Gstp2* are two *Keap1*-mutant synthetic lethal candidates. (a-b) Bar graphs denoting the change in the abundance of individual sgRNAs targeting (a) *Fth1* and (b) *Gstp2* in both KP and KPK screens.

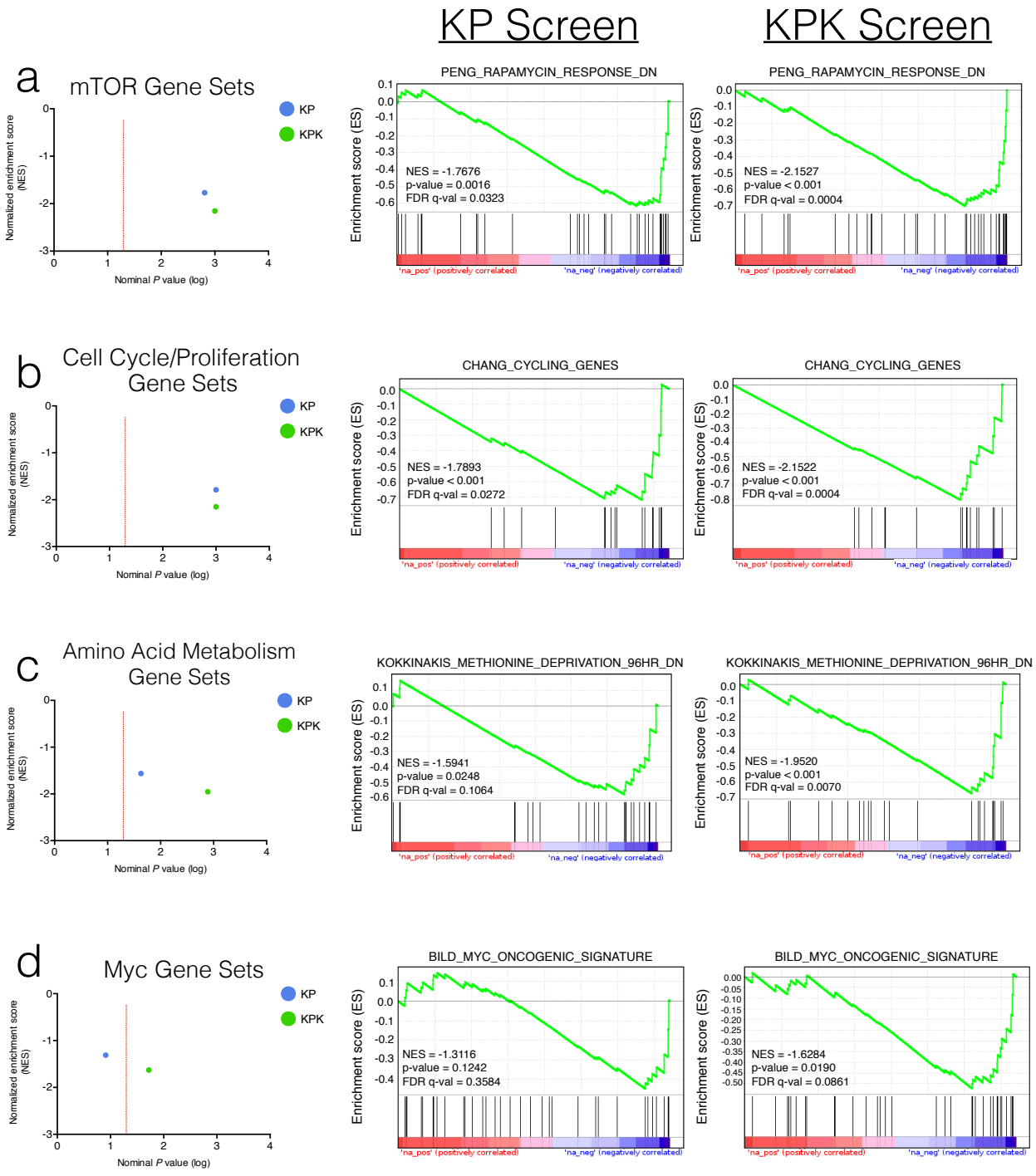

KP Screen
KPK Screen
a
mTOR Gene Sets
KP
KPK
Normalized enrichment score (NES)
Nominal P value (log)
PENG_RAPAMYCIN_RESPONSE_DN
Enrichment score (ES)
NES = -1.7676
p-value = 0.0016
FDR q-val = 0.0323
NES = -2.1527
p-value < 0.001
FDR q-val = 0.0004
'na_pos' (positively correlated)
'na_neg' (negatively correlated)
b
Cell Cycle/Proliferation Gene Sets
CHANG_CYCLING_GENES
NES = -1.7893
p-value < 0.001
FDR q-val = 0.0272
NES = -2.1522
p-value < 0.001
FDR q-val = 0.0004
c
Amino Acid Metabolism Gene Sets
KOKKINAKIS_METHIONINE_DEPRIVATION_96HR_DN
NES = -1.5941
p-value = 0.0248
FDR q-val = 0.1064
NES = -1.9520
p-value < 0.001
FDR q-val = 0.0070
d
Myc Gene Sets
BILD_MYC_ONCOGENIC_SIGNATURE
NES = -1.3116
p-value = 0.1242
FDR q-val = 0.3584
NES = -1.6284
p-value = 0.0190
FDR q-val = 0.0861

Figure 12: Gene set enrichment analysis (GSEA) of IC1 signatures from both KP and KPK screens uncovers multiple gene sets with *Keap1*-mutant synthetic lethal potential. (a-b) Gene sets corresponding to (a) mTOR-related genes (n=1) and (b) cell cycle/proliferation (n=1) are equally depleted in both KP and KPK screens. The plots on the left represent the normalized enrichment score (NES) against nominal *p* values (log) and the red line denotes threshold for significantly ($p < 0.05$) depleted gene sets. The plots on the right denote the individual gene sets used for the plots on the left (c-d) Multiple gene sets related to (c) amino acid metabolism (n=4) and (d) the oncogenic Myc transcriptional network (n=1) are significantly more depleted in KPK cells. The plots on the left represent the normalized enrichment score (NES) against nominal *p* values (log) and the red line denotes threshold for significantly ($p < 0.05$) depleted gene sets. The plots on the right denote representative examples of the individual gene sets used for the plots on the left. NES: normalized enrichment score. FDR: false discovery rate.

acid metabolism, and the Myc transcriptional network. I will describe below the therapeutic implications of a select few of these gene candidates, emphasizing any known connections to the Nrf2 antioxidant program and cellular antioxidant pathways in general.

Gstp2 and glutathione-mediated ROS detoxification

As discussed in **Part IV** of the introduction, GSH-mediated ROS detoxification plays a very important cytoprotective role in conditions of high oxidative stress. *Gstp2*, which is also a direct transcriptional target of Nrf2*,* encodes for glutathione S-transferase, pi 2 (Chanas et al., 2002). GSTs play a critical role in this pathway by means of directly conjugating GSH to ROS for the purposes of ROS detoxification (see **Figure 1** from **Part IV** of the introduction). Therefore, loss of function of a critical GST is expected to have a negative impact in GSH-mediated ROS detoxification and the overall ability of a cell to deal with high levels of oxidative stress. That *Gstp2* is within the top 20 *Keap1*-mutant synthetic lethal candidates certainly warrants further investigation, given recent work from Tak Mak's group demonstrating the importance of the GSH antioxidant pathway in the initiation of breast and sarcoma tumors *in vivo* (Harris et al., 2015).

Fth1, the ferritin complex, and iron sequestration

As discussed in **Part IV** of the introduction, the ferritin complex plays a critical cytoprotective role in conditions of high oxidative stress. *Fth1*, which is a direct transcriptional target of Nrf2, encodes for the ferritin heavy chain of the ferritin complex (Pietsch et al., 2003). The ferritin complex catalyzes the conversion of highly reactive Fe(II) molecules into non-reactive Fe(III) molecules accompanied by the incorporation of

Fe(III) molecules within its own protein structure (Gozzelino et al., 2010). Under homeostatic conditions, Fe(II)-containing molecules are inserted into the heme pockets of hemoproteins; however, in conditions of high oxidative stress, hemoproteins release these Fe(II)-containing molecules, which are subsequently broken down by another direct transcriptional target of Nrf2 - heme oxygenase (HMOX, also called HO-1) (Gozzelino et al., 2010). Unabated HMOX-mediated release of highly reactive Fe(II) molecules is highly dangerous to the cell, as these molecules catalyze the production of highly reactive hydroxyl free radicals, which can trigger massive cellular toxicity by directly damaging multiple organelles and macromolecules, such as DNA, proteins, and lipids (Gutteridge 1986, Gozzelino et al., 2010). Remarkably, cells have evolved a tightly regulated built-in cytoprotective mechanism, whereby HMOX-mediated release of highly reactive Fe(II) molecules is coupled to ferritin complex-mediated Fe(II) → Fe(III) detoxification (Gozzelino et al., 2010, Gorrini et al., 2013). Therefore, Nrf2 appears to orchestrate what looks to be a futile cycle during conditions of high oxidative stress. How does this connect to our finding that *Fth1* appears to play a critical supporting role in the survival of *Keap1*-mutant cell lines? *Keap1* mutation triggers constitutive activation of the Nrf2 transcription factor, which in turn leads to a substantial transcriptional induction of a large repertoire of genes, including *Hmox1* and *Fth1*. Consequently, *Keap1* mutant cells are constitutively releasing highly reactive Fe(II) molecules in an Hmox-dependent manner, which in turn get detoxified to Fe(III) molecules in an Fth1-dependent manner. In this scenario, genetic or pharmacological inactivation of Fth1 would lead to a substantial increase in the intracellular amount of Fe(II) molecules, which in turn would lead to a massive buildup of highly cytotoxic

hydroxyl free radicals. Therefore, hyperactivation of the Nrf2 transcription factor via *Keap1* loss of function mutations, *Nfe2l2* gain of function mutations, or other non-mutational mechanisms, creates a unique Achilles heel in cells that are highly dependent on the Nrf2 antioxidant pathway, such as *Keap1*-mutant cells (**Figure 13**). Interestingly, *Fth1* was recently implicated in the suppression of ferroptosis, which is a form of regulated cell death triggered by a massive, iron-dependent accumulation of lipid reactive oxygen species (Dixon et al., 2012, Sun et al., 2015). Whether *Fth1* inactivation in *Keap1* mutant cells triggers ferroptosis remains to be determined. Mechanistically, this acquired dependency on *Fth1* resonates with previous work from Eyal Gottlieb's group, demonstrating that fumarate hydratase (FH) deficient murine kidney cells (which are characterized by the abnormal accumulation of the fumarate and succinate metabolites) are exquisitely sensitive to Hmox1 inactivation (Frezza et al., 2011). In this context, FH-deficient cells become completely dependent on the heme biosynthesis and degradation pathway in order to regenerate NADH. Consequently, Hmox1 activity becomes critical in order for the cells to utilize the built up metabolites for subsequent NADH production (Frezza et al., 2011).

We believe that *Fth1* represents a novel synthetic lethal interaction with *Keap1* mutation (or otherwise constitutive Nrf2 activation) that offers a unique and attractive opportunity for selectively targeting tumors that are highly dependent on this antioxidant pathway. Indeed, successfully exploiting this synthetic lethal interaction has the potential to benefit over 30% of all lung cancer patients, and may even offer hope for other lethal cancers characterized by Nrf2 pathway hyperactivation, such as pancreatic adenocarcinoma and hepatocellular carcinoma (LUAD TCGA 2014, Hammerman et al.,

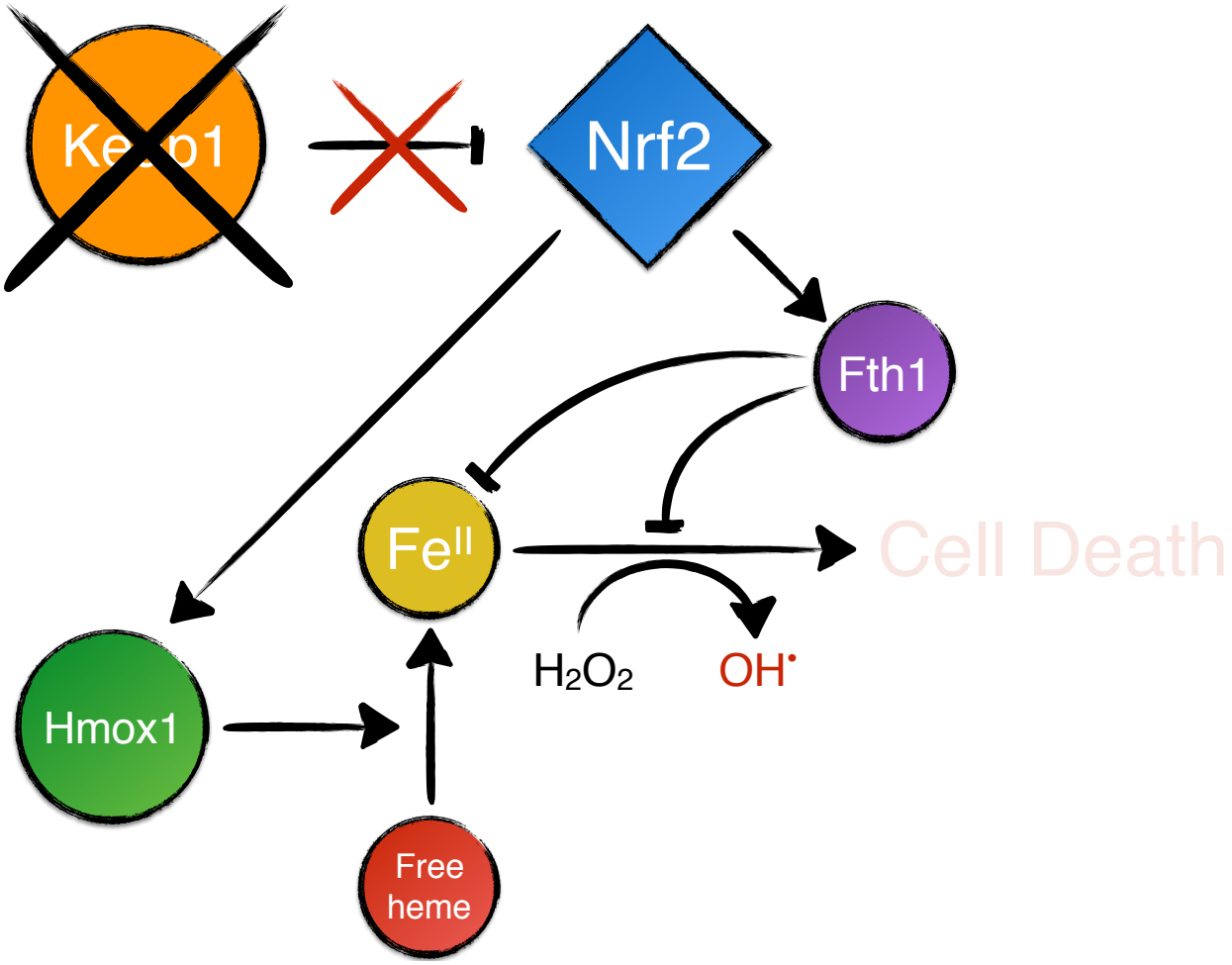


Figure 13: A model explaining the potential synthetic lethal relationship between *Keap1* mutation and *Fth1* inactivation. Hyperactivation of the Nrf2 pathway concomitantly activates *Fth1* and *Hmox1*. The Hmox1 enzyme catalyzes the production of highly reactive Fe(II) molecules, which can promote the generation of highly reactive hydroxyl free radicals from hydrogen peroxide (H_2O_2) through the Fenton reaction. Hydroxyl free radicals (OH·) are highly cytotoxic. Fth1 - as part of the ferritin complex - promotes ferritin-mediated Fe(II) → Fe(III) detoxification. In this scenario, pharmacological or genetic inactivation of Fth1 would result in the massive accumulation of highly reactive and cytotoxic hydroxyl free radicals. Therefore, hyperactivation of the Nrf2 pathway via loss of function mutations in *Keap1* creates a unique Achilles heel that could be exploited for the benefit of patients harboring these types of mutations. Refer to the text and **Chapter 4** for more details.

2012, DeNicola et al., 2011, Guichard et al., 2012, Totoki et al., 2014, Schulze et al., 2015).

Additional synthetic lethal candidates that impinge on antioxidant pathways

Beyond the two canonical Nrf2 transcriptional targets described above, our screen uncovered multiple other *Keap1* mutant synthetic lethal candidates that might directly or indirectly impinge in one or more cellular antioxidant pathways. For example, the metabolic enzyme pyruvate carboxylase (encoded by the *Pcx* gene), which catalyzes the carboxylation of pyruvate to generate oxaloacetate, indirectly contributes to the generation of NADPH from malate (Sugden and Holness 2011). Interestingly, recent oxidative stress studies in *Pseudomonas fluorescens* have suggested the existence of alternative metabolic pathways that can convert NADH to NADPH, thereby contributing to one or more antioxidant pathways (Singh et al., 2008). Therefore, *Pcx* inactivation could conceivably lead to a marked selective decrease in the cellular fitness of *Keap1*-mutant cells by lowering the levels of NADPH, which would in turn affect the regeneration of the GSH and TXN cellular antioxidants (discussed in **Part IV** of the introduction). On the other hand, the solute carrier family 13 member 3 (encoded by the *Slc13a3* gene) was recently shown to be a low-affinity GSH transporter in the kidneys (Schorbach et al., 2013), suggesting a possible connection between Slc13a3 and the glutathione antioxidant pathway.

Other synthetic lethal candidates

As shown in **Figure 10A**, our screen uncovered multiple *Keap1* synthetic lethal candidates with unknown connections to antioxidant responses. One notable example

was C-terminal Src kinase (encoded by the *Csk* gene), which was the top synthetic lethal candidate (lowest KPK-KP Differential Gene Score) (**Figure 10A-D**). Interestingly, this non-receptor tyrosine kinase is known to phosphorylate and inactivate multiple Src-family kinases, including c-Src, Fyn, c-Yes, Hck, c-Fgr, Blk, Lyn, and Lck (Okada et al., 2012). The fact that Csk exerts an inhibitory effect over Src-family kinases, which are known to promote several aspects of the cancer phenotype (Yeatman 2004), presents a major scientific conundrum due to the fact that our genetic screen suggests that Csk is required for the growth of *Keap1* mutant cells. Remarkably, a multitude of bioinformatic iterations accompanied by highly stringent data filtering approaches (data not shown) consistently unveiled Csk as the most significant *Keap1* mutant synthetic lethal candidate gene. Despite the fact that the biological reason behind this particular result remains unknown, many genes are known to play context-specific roles. Indeed, the aforementioned *Epha2* gene, which was recently determined to be a potent suppressor of *Kras*^{G12D}-driven lung adenocarcinoma, is part of a family of receptor tyrosine kinases that are widely considered to be oncogenes in most biological contexts (Pasquale et al., 2010, Yeddula et al., 2015). In this setting, *Epha2* inactivation was shown to relieve feedback inhibition of Kras, which resulted in potent activation of the MAPK pathway and subsequent enhancement of cellular proliferation (Yeddula et al., 2015). Notably, the tumor suppressive function of *Epha2* was unveiled via unbiased shRNA screening, underscoring the potential of genetic screens to identify context-specific tumor suppressor genes and growth promoters. Therefore, we hypothesize that Csk might be playing a potent growth promoting role exclusively in the context of *Keap1* mutation/Nrf2 pathway hyperactivation. Mechanistically, perhaps Csk is targeting a non-oncogenic

Src-related tyrosine kinase, such as Rak/Frk, which has been shown to have potent tumor suppressive activity in breast cancer through the phosphorylation-dependent activation of the *Pten* tumor suppressor gene (Yim et al., 2009), as well as antagonistic activity towards oncogenic EGFR mutant proteins (Jin and Craven 2014).

Our results underscore the power of CRISPR-based genetic screens for uncovering novel genetic dependencies in the context of clinically relevant cancer-associated genotypes. Furthermore, our data predicts multiple *Keap1* mutant synthetic lethal interactions that, if validated, could be readily exploited for the benefit of a large percentage of cancer patients suffering from tumors characterized by hyperactivation of the Nrf2 pathway. Experimental approaches, such as the one described here, are essential in order to fulfill the promise of precision oncology.

MATERIALS AND METHODS

Vectors

Cas9-expressing cells were generated with lentiCas9-Blast (Zhang Lab, Addgene, 52962). Gibson compatible parts used to assembled pUSCG were generated from lentiCRISPRv1 (Zhang lab, Addgene, 49535) and pCMV-GFP (unpublished; available upon request). Lentiviral packaging plasmids psPax2 (Addgene, 12260) and pMD2G (Addgene, 12259) were obtained from Addgene.

The U6-sgRNA-CMV-EGFP (pUSCG) vector was generated by Gibson Assembly (GA) (NEB) of the following parts (modified to contain compatible Gibson overhangs): a 1.4 kb part corresponding to the CMV promoter and E-GFP fluorescent protein, a 2.3 kb

part corresponding to the human U6 promoter and sgRNA scaffold, and a lentiviral backbone with compatible Gibson overhangs.

Antibodies

The following antibodies were used for western blotting experiments: anti-Cas9 (Clone 7A9-3A3, Active Motif, 1:1000), anti-GAPDH (Santa Cruz, sc-25778, 1:500 dilution), and anti-Keap1 (CST, 8047S, 1:1000 dilution).

Quantitative PCR (qPCR)

RNA was prepared using TRIzol (Thermo Fisher Scientific, 15596-026). Synthesis of cDNA was performed using the High-Capacity cDNA Reverse Transcription Kit (Thermo Fisher Scientific, 4368814). qPCR was performed in Roche Lightcycler 480 II. Data were analyzed using the comparative ΔC_T method, and were normalized to the levels of Gapdh.

Primers used were the following:

Target	**Forward Oligo (5' → 3')**	**Reverse Oligo (5' → 3')**
Gapdh	TTTGATGTTAGTGGGGTCTCG	AGCTTGTCATCAACGGGAAG
Gclc	AGATGATAGAACACGGGAGGAG	TGATCCTAAAGCGATTGTTCTTC
Nqo1	AGCGTTCGGTATTACGATCC	AGTACAATCAGGGCTCTTCTCG
Hmox1	AGGCTAAGACCGCCTTCCT	TGTGTTCCTCTGTCAGCATCA

Generation of cell lines and cell culture

KPK cells (an isogenic derivative from a KP-derived cell line; Jackson et al., 2005) were generated using pX458 (Addgene, 48138) essentially as described in Ran et al., 2013

using sgKeap1.4 (5' - GTGTTCCACGCGTGCATCGAC - 3'). All cells were maintained in DMEM supplemented with 10% Fetal Bovine Serum and gentamicin. Lentiviral particles were generated and packaged using HEK293T cells. Briefly, cells were plated the day before transfection. HEK293T cells were co-transfected with lentiviral constructs from sgRNA libraries and packaging plasmids (psPAX2 and pMD2.G) using TransIT-LT1 (Mirus Bio LLC). Viral supernatant was collected 48 and 72 hours post-transfection. Cells infected with lentiCas9-Blast were allowed to recover for 48 hours post-infection and were subsequently selected with 20 μg/mL Blasticidin-S (Life Technologies) for a period of 5-7 days.

Animal experiments

All animal studies described in this study were approved by the MIT Institutional Animal Care and Use Committee. A total of 12 million cells (KP or KPK) divided over 4 subcutaneous sites per mouse at 3 million cells/site were implanted in a 1:1(v/v) mixture with Matrigel (Corning) into syngeneic C57BL6/6J (The Jackson Laboratory) recipient mice at the age of 6-8 weeks (see below for more details).

Western blotting

Cells were lysed with RIPA buffer (BP-115, Boston BioProducts) supplemented with 1X protease inhibitor solution (cOmplete EDTA-free, 11873580001, Roche). Protein concentration of cell lysates was determined by Pierce BCA protein assay (23225, Thermo Fisher Scientific). Total protein (50μg) was separated on 4-12% Bis-Tris gradient SDS-PAGE gels (Life Technologies) and then transferred to PVDF membranes (IPVH00010, EMD Millipore) for blotting.

sgRNA oligo design/structure/synthesis

Gene lists for each target library were translated to mouse gene symbols using HUGO gene nomenclature (http://www.genenames.org/) and Mouse Genome Informatics (http://www.informatics.jax.org/) orthology mapping resources. Transcript sequences for all isoforms of each gene were obtained from the UCSC mouse build mm9 annotation (https://genome.ucsc.edu/). Candidate 20mer guide RNA sequences upstream potential PAM sites ([20mer]NGG) were identified in sense and antisense directions for each transcript. Corresponding 30mer were also extracted (NNNN[20mer]NGGNNN) and scored using an implementation of the Doench et al. sgRNA scoring algorithm (Doench et al., 2014). For each gene, unique 20mer sgRNAs were retained across all isoforms and sequences disrupted by genomic splice sites were dropped from further consideration. The *n* top scoring sgRNAs (n=6, 10) were retained for each gene. A 5'-G (necessary for U6 transcriptional initiation) was added to any sequence that did not start with a G and adapter sequences (in **Supplementary Tables Section** below) were added to generate the final library sequences. Each library included 500 control sgRNAs, adapted from Shalem et al., 2014. DNA oligonucleotides in the form of [Pool F adapter – sgRNA (20-21 nt) – Pool R adapter], for a total length of 74-75nt, were synthesized on the CustomArray 12K (Epigenome library) and 90K (Druggable genome library) arrays (CustomArray, Inc.) as separate pools.

Oligo pool PCR amplification strategy

Oligonucleotides were PCR amplified and GA-compatible overhangs were added using Herculase II Fusion DNA Polymerase (Agilent) and Array F and Array R primers.

Specifically, a 50µL reaction contained 5ng of template (oligo pool from array), 10µL of 5X Herculase II reaction buffer, 0.25µM of each dNTP, 2mM of $MgCl_2$, 0.25µM of each primer and 1µL of Herculase II Fusion DNA polymerase. For this PCR, the thermocycling parameters were: 95°C for 2 mins, 15 cycles of (95°C for 20 s, 55°C for 30 s, 72°C for 30 s), and 72°C for 3 mins. PCR products were size selected (140 nt) using a 2% agarose gel (Affymetrix).

Gibson assembly into pUSCG and amplification of plasmid in bacteria

The pUSCG vector was linearized with BsbmBI (NEB) at 55°C and gel purified on a 1% agarose gel (Affymetrix). A total of 10 x 20µL GA reactions (NEB) were performed per library using 1.66ng of insert (pooled sgRNA oligos) and 100ng of vector (pUSCG). Ligation reactions were combined, purified and 2.5µL were transformed into 25µL of electrocompetent cells (Endura Lucigen), according to manufacturer's protocol using a MicroPulser Electroporator (BioRad). To maintain library representation, we performed 8 bacterial transformations per library using the same ligation reaction, yielding 2,000X and 175X for Epigenome and Druggable genome library, respectively. All the transformations per library were pooled and entirely plated onto 10-cm Carbenicillin/LB plates (40 plates per library). Bacterial colonies were scraped off the plates and combined before plasmid isolation using PureLink HiPure Plasmid Maxiprep Kit (Invitrogen).

sgRNA plasmid library sequencing

sgRNA inserts were PCR amplified from plasmid library using Herculase II Fusion DNA Polymerase (Agilent) and Array F and Array R primers. Specifically, a 50µL reaction

contained 100ng of template (pooled sgRNA plasmid), 10µL of 5X Herculase II reaction buffer, 0.25µM of each dNTP, 2mM of $MgCl_2$, 0.25µM of each primer and 1µL of Herculase II Fusion DNA polymerase. For this PCR, the thermocycling parameters were: 95°C for 2 mins, 15 cycles of (95°C for 20 s, 55°C for 30 s, 72°C for 30 s), and 72°C for 3 mins. PCR products were size selected (140 nt) using a 2% agarose gel (Affymetrix).

MOI determination

To determine the optimal viral supernatant volume to achieve an MOI of 0.3-0.5 (ensuring one integrant per cell), Cas9-expressing cells were tested by spin-infection. Briefly, a total of 12 million cells were seeded at 1 million cells / well in a 12-well plate in a final volume of 2mL of media / well with increasing volumes of viral supernatant (0, 25, 50, 100, 250, 500, 750, 1000µL) and supplemented with polybrene (Millipore) at 10µg/mL. Subsequently, cells were spin-infected at 1,500RPM for 2 hours at room temperature. After the spin-infection, we added 1mL of warm media and incubated the cells overnight at 37°C. Infection efficiency was determined by flow cytometric analysis at 24-48 hrs post infection. To calculate viral titer, we used the following equation: **Titer = [(F x C_n) / V] x DF***,* where F, frequency of GFP^+ cells determined by flow cytometry; C_n, number of target cells infected; V, volume of virus; DF, virus dilution factor.

Infection of target cells with Druggable Genome lentiviral library

To ensure a library representation of >1,000X at every step of the screen, we set up multiple large-scale spin-infections in both KP and KPK cells aimed at achieving very low transduction efficiencies (which are conditions that are predicted to predominantly

lead to a single lentiviral integration per cell; see above for MOI calculations) exactly as follows:

8 x 12-well plates per **independent** infection replicate
x
3 replicates per cell line
x
2 cell lines (KP and KPK)

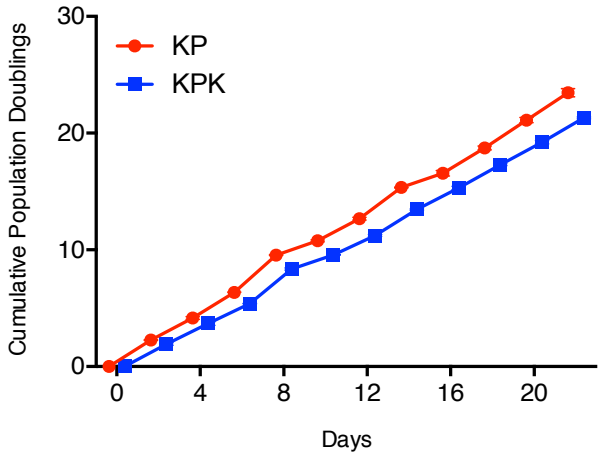
Total: 48 x 12-well plates

Subsequently, cells were spin-infected at 1,500RPM for 2 hours at room temperature. After the spin-infection, we added 1mL of warm media and incubated the cells overnight at 37°C. Next day, cells were trypsinized and corresponding wells from 12-well plates were pooled and re-plated in 15-cm plates (for a total of 48 independent 15-cm plates). The day after (48hrs post-infection), cells were sorted based on GFP expression and re-plated appropriately. Importantly, we consistently kept the coverage at >1,000X during and after sorting. We expanded the sorted cells for a total of ~7 days post-sorting, ensuring, again, that we consistently maintained the coverage at >1,000X.

In vitro **screen**

After expanding the cells for ~7 days, we set up the *in vitro* screen as follows: for each *in vitro* replicate, we plated 2 independent 15-cm plates at 6 million cells per 15-cm plate for a total of 12 million cells per replicate at all times. This corresponds to a ~1,000X coverage for the Druggable Genome library. At the time of plating, a total of 12 million cells were collected as the early/input sample. To maintain a ~1,000X representation of the library during the extent of the screen, we trypsinized cells,

counted, and re-plated a minimum of 12 million cells every 2 days. Importantly, throughout the screen, both populations of cells from both arms of the screen proliferated identically within technical replicates (see below for cumulative population doublings; each individual line represents three independent KP or KPK replicates):



Cells were cultured for a total of 20 cumulative population doublings, after which genomic DNA (gDNA) was extracted.

***In vivo* screen**

A total of 12 million KP or KPK cells transduced with the Druggable Genome library (same cells used for the *in vitro* screen - i.e. from the same transduced population described above), divided over 4 subcutaneous sites per mouse at 3 million cells/site were implanted in a 1:1 (v/v) mixture with Matrigel (Corning) into syngeneic C57BL6/6J (The Jackson Laboratory) recipient mice at the age of 6-8 weeks. We injected a total of 6 mice per independent infection replicate (for a total of n = 36 mice for the KP-KPK Druggable Genome screen). Importantly, these injections were performed the same day as the initial day of the *in vitro* screen (therefore, the early/input sample taken at Day 0

of the *in vitro* screen corresponds to the early/input sample for both *in vitro* and *in vivo* screens). When the *in vitro* cells reached 20 cumulative population doublings, mice were sacrificed by CO_2 asphyxiation and all 4 subcutaneous tumors per mouse were pooled, snap-frozen and considered a single technical replicate (i.e. 1 mouse = 1 replicate within each independent technical replicate). Frozen subcutaneous tumors were finely grounded over liquid nitrogen and gDNA isolation was performed as described below.

Genomic DNA isolation

gDNA was isolated using an already described salt precipitation method (Chen et al., 2015). In a 15mL conical tube, 6mL of NK lysis buffer (50mM Tris, 50mM EDTA, 1% SDS, pH 8.0) supplemented with 30µL of 20mg/mL Proteinase K (19131, QIAGEN) were added to tissue (up to 200mg) or cells (up to 30 million cells) and was incubated at 55°C overnight. Then, 30mL of RNAse A (19101, QIAGEN, resuspended in NK lysis buffer to 10mg/mL) was added to the lysate, inverted 25 times and incubated at 37°C for 30 minutes. Samples were cooled down on ice before addition of 2mL of pre-chilled 7.5M ammonium acetate (NH_4OAc) (A1542, Sigma-Aldrich) to precipitate proteins. After addition of NH_4OAc, the samples were vortexed at high speed for 20s and then centrifuged at 4,000rpm for 10 minutes. After centrifugation, a white pellet was visible in the tube and supernatant was transferred to a new tube. Then 6mL of 100% isopropanol was added to the supernatant, tube was inverted 50 times and centrifuged at 4,000rpm for 10 minutes. At this stage, gDNA was visible as a small pellet. The supernatant was discarded, 6mL of 70% ethanol was added, tube was inverted 10 times, and then

centrifuged at 4,000rpm for 1 minute. Supernatant was discarded by pouring, tube was briefly spun down and remaining ethanol was remove by using a pipette. DNA was allowed to air-dry for approximately 30 minutes, and then resuspended in 200µL of water. Tube was incubated at 65°C for an hour and at room temperature overnight. gDNA concentration was determined by Nanodrop (Thermo Scientific). When processing higher amounts of tissue or cells, quantities were scaled up proportionally.

Deconvolution of genetic screen

High-throughput sequencing libraries were generated by PCR amplification of the integrated sgRNAs using a nested PCR strategy. For each *in vitro* sample, gDNA from at least 12.2 million cells (1000X) was used as a template in multiple parallel 50µL reactions (when assuming 6.6pg of gDNA per cell, this corresponded to ~ 80µg of gDNA; Chen et al., 2015). Similarly, for each *in vivo* sample, ~ 80µg of gDNA were used as a template. The first PCR reaction (PCR #1 – amplification of sgRNA from gDNA), sgRNA sequences were amplified by using Herculase II Fusion DNA Polymerase (Agilent) and F1 and F2 primers. Specifically, a 50µL reaction contained 2µg of template (gDNA), 10µL of 5X Herculase II reaction buffer, 0.25µM of each dNTP, 2mM of $MgCl_2$, 0.25µM of each primer and 1µL of Herculase II Fusion DNA polymerase. For this PCR, the thermocycling condition parameters were: 95°C for 2 mins, 25 cycles of (95°C for 30 s, 65°C for 30 s, 72°C for 30 s), and 72°C for 3 mins.

PCR products were combined for each sample and concentrated using 3M sodium acetate solution (NaOAc) (pH 5.5, Ambion). A NaOAc volume corresponding to 1/10 of the total volume of combined PCR products was added followed by 2 volumes of ice-

cold 100% ethanol. Sample was incubated at -80°C for 10 minutes and subsequently centrifuged at 4°C at 3500rpm for 10 minutes. The DNA pellet was washed with 5mL of 70% ethanol. Sample was centrifuged at room temperature at 3500rpm for 10 minutes. DNA was air-dried for approximately 30 minutes. Once completely dried, it was resuspended in water and size selected (215 nt) using a 2% agarose gel (Affymetrix). In the second PCR reaction (PCR #2 – barcoding and indexing), sgRNAs amplified in PCR#1 were barcoded and indexed by using Herculase II Fusion DNA Polymerase (Agilent) and F01-F12 and R01-R02 primers. Specifically, a 50µL reaction contained 100ng of template (PCR#1 product), 10µL of 5X Herculase II reaction buffer, 0.25µM of each dNTP, 2mM of $MgCl_2$, 0.25µM of each primer and 1µL of Herculase II Fusion DNA polymerase. For this PCR, the thermocycling condition parameters were: 95°C for 2 mins, 25 cycles of (95°C for 30 s, 65°C for 30 s, 72°C for 30 s), and 72°C for 3 mins. PCR products were combined for each sample and concentrated using 3M sodium acetate solution (NaOAc) (pH 5.5, Ambion) as described above. Samples were sized selected (261-271nt) by 2% agarose gel (Affymetrix) electrophoresis. Libraries were quantified by qPCR using the Illumina Library Quantification Kit (KAPA Biosystems) according to manufacturer guidelines. Libraries were sequenced on the Illumina HiSeq2500 for 100-nt in single read mode (100 SE).

Data analysis and hit selection

Illumina 100mer single-ended reads were analyzed for 10mer adapter sequences flanking library-specific sgRNA sequences. Sequences with exact matches to sgRNA library sequences were tabulated to generate high confidence raw counts for each

sgRNA within each sample, using custom scripts. To reduce uncertainty introduced by off-target effects, sgRNAs with greater than 10 exact matches to the mouse genome sequence (mm9 build) were excluded from downstream analyses. Raw sgRNA counts per sample were upper-quartile normalized to a count of 2000 (Bullard et al., 2010) and normalized counts were $\log_2$ transformed. Fold changes for sgRNA representation (in $\log_2$ space) were calculated using these values for all required contrasts. For gene-level representation, the median normalized count of all sgRNAs targeting that gene was used for further analysis. Gene-level fold-changes between conditions were similarly derived. For all "treated" conditions (i.e., KPK cells when comparing KP vs KPK), fold-change values with respect to corresponding basal conditions were compared to estimate depletion/enrichment of gene-level representation (Differential Gene Score analysis).

Apart from ranking sgRNAs and genes based on fold-change values, unbiased signature analysis was performed to accommodate for noise and heterogeneity across conditions. Independent Component Analysis (ICA), an unsupervised blind source separation technique, was used on this discrete count-based expression dataset to elucidate statistically independent and biologically relevant signatures (Bhutkar et al., 2015). ICA is a signal processing and multivariate data analysis technique in the category of unsupervised matrix factorization methods. Conceptually, ICA decomposes the overall count dataset into independent signals (sgRNA representation patterns) that represent distinct signatures. High-ranking positively and negatively correlated sgRNA levels in each signature represent sgRNA sets that drive the corresponding expression pattern (in either direction). Signatures were visualized using the sample-to-signature

correspondence schematic afforded by Hinton plots where colors represent directionality of sgRNA representation (red enriched, green depleted) and the size of each rectangle quantifies the strength of a signature (column) in a given sample (row). Each signature is two-sided, allowing for identification of enriched and depleted sgRNAs for each signature within each sample.

Formally, utilizing input data consisting of a genes-samples matrix, ICA uses higher order moments to characterize the dataset as a linear combination of statistically independent latent variables. These latent variables represent independent components based on maximizing non-gaussianity, and can be interpreted as independent source signals that have been mixed together to form the dataset under consideration. Each component includes a weight assignment to each gene that quantifies its contribution to that component. Additionally, ICA derives a mixing matrix that describes the contribution of each sample towards the signal embodied in each component. This mixing matrix can be used to select signatures among components with distinct sgRNA representation profiles across the set of samples. The R implementation of the core JADE algorithm (Joint Approximate Diagonalization of Eigenmatrices) (Nordhausen et al., 2015, Rutledge & Bouveresse, 2013, Biton et al., 2013) was used along with custom R utilities.

All data analyses were conducted in the R Statistical Programming language (http://www.r-project.org/). Heatmaps were generated using the Heatplus package in R.

ACKNOWLEDGEMENTS

We thank Isaac S. Harris and Thales Papagiannakopoulos for scientific discussions regarding Keap1-Nrf2 biology. We thank T. Volkert and S. Gupta for DNA Sequencing support, and G. Paradis, M. Griffin, M. Jennings and M. Saturno-Condón for flow cytometry and FACS support. This work was supported by the Howard Hughes Medical Institute and in part by Cancer Center Support (core) grant P30-CA14051 from the National Cancer Institute. Y.M.S.F. is supported by the National Cancer Institute of the National Institutes of Health under Award Number F31-CA183405. M.T.H is the Eisen and Chang Associate Professor of Biology. T.J. is a Howard Hughes Medical Institute Investigator, the David H. Koch Professor of Biology, and a Daniel K. Ludwig Scholar.

SUPPLEMENTARY TABLES

Oligo pool adapters

Sequence Name	Sequence (5' → 3')
Upstream	TATATCTTGTGGAAAGGACGAAACACC
Downstream	GTTTTAGAGCTAGAAATAGCAAGTTAA

Oligo pool PCR primers

Sequence Name	Sequence (5' → 3')
Array F	GTAACTTGAAAGTATTTCGATTTCTTGGCTTTATATATCTTGTGGAAAGGACGAAACACC
Array R	ACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAAC

PCR #1 primers

Sequence Name	Sequence (5' → 3')
F1 PCR #1	AATGGACTATCATATGCTTACCGTAACTTGAAAGTATTTCG
R1 PCR #1	AGGCCTCGGGATTCCTAGGAACAGCGGTTTAAAAAAGCACC

PCR #2 forward primers

Sequence Name	Sequence (5' → 3')
F01	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTtA AGTAGAGTCTTGTGGAAAGGACGAAACACCG
F02	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTat ACACGATCTCTTGTGGAAAGGACGAAACACCG
F03	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTg atCGCGCGGTTCTTGTGGAAAGGACGAAACACCG
F04	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTcg atCATGATCGTCTTGTGGAAAGGACGAAACACCG
F05	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTtc gatCGTTACCATCTTGTGGAAAGGACGAAACACCG
F06	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTat cgatTCCTTGGTTCTTGTGGAAAGGACGAAACACCG
F07	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTg atcgatAACGCATTTCTTGTGGAAAGGACGAAACACCG
F08	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTcg atcgatACAGGTATTCTTGTGGAAAGGACGAAACACCG
F09	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTac gatcgatAGGTAAGGTCTTGTGGAAAGGACGAAACACCG
F10	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTtA ACAATGGTCTTGTGGAAAGGACGAAACACCG
F11	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTat ACTGTATCTCTTGTGGAAAGGACGAAACACCG
F12	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTg atAGGTCGCATCTTGTGGAAAGGACGAAACACCG

PCR #2 reverse primers

Sequence Name	Sequence (5' → 3')
R01	CAAGCAGAAGACGGCATACGAGATAAGTAGAGGTGACTGGAGTTCAGACGTGTGCTCT TCCGATCTtGTTTAAAAAAGCACCGACTCGGTGC
R02	CAAGCAGAAGACGGCATACGAGATACACGATCGTGACTGGAGTTCAGACGTGTGCTCT TCCGATCTatGTTTAAAAAAGCACCGACTCGGTGC
R03	CAAGCAGAAGACGGCATACGAGATCGCGCGGTGTGACTGGAGTTCAGACGTGTGCTC TTCCGATCTgatGTTTAAAAAAGCACCGACTCGGTGC
R04	CAAGCAGAAGACGGCATACGAGATCATGATCGGTGACTGGAGTTCAGACGTGTGCTCT TCCGATCTcgatGTTTAAAAAAGCACCGACTCGGTGC

R05	CAAGCAGAAGACGGCATACGAGATCGTTACCAGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTtcgatGTTTAAAAAAGCACCGACTCGGTGC
R06	CAAGCAGAAGACGGCATACGAGATTCCTTGGTGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTatcgatGTTTAAAAAAGCACCGACTCGGTGC
R07	CAAGCAGAAGACGGCATACGAGATAACGCATTGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTgatcgatGTTTAAAAAAGCACCGACTCGGTGC
R08	CAAGCAGAAGACGGCATACGAGATACAGGTATGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTcgatcgatGTTTAAAAAAGCACCGACTCGGTGC
R09	CAAGCAGAAGACGGCATACGAGATAGGTAAGGGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTacgatcgatGTTTAAAAAAGCACCGACTCGGTGC
R10	CAAGCAGAAGACGGCATACGAGATAACAATGGGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTtGTTTAAAAAAGCACCGACTCGGTGC
R11	CAAGCAGAAGACGGCATACGAGATACTGTATCGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTatGTTTAAAAAAGCACCGACTCGGTGC
R12	CAAGCAGAAGACGGCATACGAGATAGGTCGCAGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTgatGTTTAAAAAAGCACCGACTCGGTGC

Illumina Sequencing primers

Sequence Name	**Sequence (5’ → 3’)**
Multiplexing Read 1 Sequencing Primer 5'	ACACTCTTTCCCTACACGACGCTCTTCCGATCT
Multiplexing Index Read Sequencing Primer 5'	GATCGGAAGAGCACACGTCTGAACTCCAGTCAC
Multiplexing Read 2 Sequencing Primer 5’	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATC

REFERENCES

Bhutkar, A., Blat, I., Boutz, P.L., Cameron, E.R., Chen, P.-Y., Chen, S., Ferretti, R., Gurtan, A.M., Ianari, A., Muzumdar, M.D., et al. (2015). High resolution signature discovery in NGS expression datasets using Blind Source Separation, *In preparation*.

Birsoy, K., Wang, T., Chen, W.W., Freinkman, E., Abu-Remaileh, M., and Sabatini, D.M. (2015). An Essential Role of the Mitochondrial Electron Transport Chain in Cell Proliferation Is to Enable Aspartate Synthesis. Cell 162, 540–551.

Biton, A., Zinovyev, A., Barillot, E., and Radvanyi, F. (2013). MineICA: Independent component analysis of transcriptomic data.

Bullard, J.H., Purdom, E., Hansen, K.D., and Dudoit, S. (2010). Evaluation of statistical methods for normalization and differential expression in mRNA-Seq experiments. BMC Bioinformatics 11, 94.

Cancer Genome Atlas Research Network (2014). Comprehensive molecular profiling of lung adenocarcinoma. Nature 511, 543–550.

Chanas, S.A., Jiang, Q., McMahon, M., McWalter, G.K., McLellan, L.I., Elcombe, C.R., Henderson, C.J., Wolf, C.R., Moffat, G.J., Itoh, K., et al. (2002). Loss of the Nrf2 transcription factor causes a marked reduction in constitutive and inducible expression of the glutathione S-transferase Gsta1, Gsta2, Gstm1, Gstm2, Gstm3 and Gstm4 genes in the livers of male and female mice. Biochem. J. 365, 405–416.

Chen, S., Sanjana, N.E., Zheng, K., Shalem, O., Lee, K., Shi, X., Scott, D.A., Song, J., Pan, J.Q., Weissleder, R., et al. (2015). Genome-wide CRISPR Screen in a Mouse Model of Tumor Growth and Metastasis. Cell.

David, C.J., Chen, M., Assanah, M., Canoll, P., and Manley, J.L. (2010). HnRNP proteins controlled by c-Myc deregulate pyruvate kinase mRNA splicing in cancer. Nature 463, 364–368.

Denicola, G.M., Chen, P.-H., Mullarky, E., Sudderth, J.A., Hu, Z., Wu, D., Tang, H., Xie, Y., Asara, J.M., Huffman, K.E., et al. (2015). NRF2 regulates serine biosynthesis in non-small cell lung cancer. Nat Genet.

Denicola, G.M., Karreth, F.A., Humpton, T.J., Gopinathan, A., Wei, C., Frese, K., Mangal, D., Yu, K.H., Yeo, C.J., Calhoun, E.S., et al. (2011). Oncogene-induced Nrf2 transcription promotes ROS detoxification and tumorigenesis. Nature 475, 106–109.

Dixon, S.J., Lemberg, K.M., Lamprecht, M.R., Skouta, R., Zaitsev, E.M., Gleason, C.E., Patel, D.N., Bauer, A.J., Cantley, A.M., Yang, W.S., et al. (2012). Ferroptosis: An Iron-Dependent Form of Nonapoptotic Cell Death. Cell 149, 1060–1072.

Doench, J.G., Hartenian, E., Graham, D.B., Tothova, Z., Hegde, M., Smith, I., Sullender, M., Ebert, B.L., Xavier, R.J., and Root, D.E. (2014). Rational design of highly active sgRNAs for CRISPR-Cas9-mediated gene inactivation. Nat Biotechnol 32, 1262–1267.

Frezza, C., Zheng, L., Folger, O., Rajagopalan, K.N., MacKenzie, E.D., Jerby, L., Micaroni, M., Chaneton, B., Adam, J., Hedley, A., et al. (2011). Haem oxygenase is synthetically lethal with the tumour suppressor

fumarate hydratase. Nature 477, 225–228.

Gozzelino, R., Jeney, V., and Soares, M.P. (2010). Mechanisms of Cell Protection by Heme Oxygenase-1. Annu. Rev. Pharmacol. Toxicol. 50, 323–354.

Guichard, C., Amaddeo, G., Imbeaud, S., Ladeiro, Y., Pelletier, L., Maad, I.B., Calderaro, J., Bioulac-Sage, P., Letexier, M., Degos, F., et al. (2012). Integrated analysis of somatic mutations and focal copy-number changes identifies key genes and pathways in hepatocellular carcinoma. Nat Genet 44, 694–698.

Gutteridge, J.M. (1986). Iron promoters of the Fenton reaction and lipid peroxidation can be released from haemoglobin by peroxides. FEBS Lett 201, 291–295.

Hammerman, P.S., Lawrence, M.S., Voet, D., Jing, R., Cibulskis, K., Sivachenko, A., Stojanov, P., McKenna, A., Lander, E.S., Gabriel, S., et al. (2012). Comprehensive genomic characterization of squamous cell lung cancers. Nature 489, 519–525.

Harris, I.S., Treloar, A.E., Inoue, S., Sasaki, M., Gorrini, C., Lee, K.C., Yung, K.Y., Brenner, D., Knobbe-Thomsen, C.B., Cox, M.A., et al. (2015). Glutathione and Thioredoxin Antioxidant Pathways Synergize to Drive Cancer Initiation and Progression. Cancer Cell.

Jin, L., and Craven, R.J. (2014). The Rak/Frk tyrosine kinase associates with and internalizes the epidermal growth factor receptor. Oncogene 33, 326–335.

Le, A., Lane, A.N., Hamaker, M., Bose, S., Gouw, A., Barbi, J., Tsukamoto, T., Rojas, C.J., Slusher, B.S., Zhang, H., et al. (2012). Glucose-independent glutamine metabolism via TCA cycling for proliferation and survival in B cells. Cell Metab. 15, 110–121.

Mitsuishi, Y., Taguchi, K., Kawatani, Y., Shibata, T., Nukiwa, T., Aburatani, H., Yamamoto, M., and Motohashi, H. (2012). Nrf2 redirects glucose and glutamine into anabolic pathways in metabolic reprogramming. Cancer Cell 22, 66–79.

Nordhausen, K., Cardoso, J.-F., Miettinen, J., Oja, H., Ollila, E., and Taskinen, S. (2015). JADE: Blind Source Separation Methods Based on Joint Diagonalization and Some BSS Performance Criteria.

Ohta, T., Iijima, K., Miyamoto, M., Nakahara, I., Tanaka, H., Ohtsuji, M., Suzuki, T., Kobayashi, A., Yokota, J., Sakiyama, T., et al. (2008). Loss of Keap1 function activates Nrf2 and provides advantages for lung cancer cell growth. Cancer Research 68, 1303–1309.

Okada, M. (2012). Regulation of the SRC family kinases by Csk. Int. J. Biol. Sci. 8, 1385–1397.

Padmanabhan, B., Tong, K.I., Ohta, T., Nakamura, Y., Scharlock, M., Ohtsuji, M., Kang, M.-I., Kobayashi, A., Yokoyama, S., and Yamamoto, M. (2006). Structural basis for defects of Keap1 activity provoked by its point mutations in lung cancer. Mol Cell 21, 689–700.

Pasquale, E.B. (2010). Eph receptors and ephrins in cancer: bidirectional signalling and beyond. Nat Rev Cancer 10, 165–180.

Pietsch, E.C., Chan, J.Y., Torti, F.M., and Torti, S.V. (2003). Nrf2 mediates the induction of ferritin H in response to xenobiotics and cancer chemopreventive dithiolethiones. Journal of Biological Chemistry

278, 2361–2369.

Piskounova, E., Agathocleous, M., Murphy, M.M., Hu, Z., Huddlestun, S.E., Zhao, Z., Leitch, A.M., Johnson, T.M., DeBerardinis, R.J., and Morrison, S.J. (2015). Oxidative stress inhibits distant metastasis by human melanoma cells. Nature advance online publication SP - EP .

Ran, F.A., Hsu, P.D., Wright, J., Agarwala, V., Scott, D.A., and Zhang, F. (2013). Genome engineering using the CRISPR-Cas9 system. Nat Protoc 8, 2281–2308.

Roychowdhury, S., and Chinnaiyan, A.M. (2015). Translating cancer genomes and transcriptomes for precision oncology. CA Cancer J Clin.

Rutledge, D.N., and Bouveresse, D. (2013). Independent components analysis with the JADE algorithm. TrAC Trends in Analytical Chemistry.

Sayin, V.I., Ibrahim, M.X., Larsson, E., Nilsson, J.A., Lindahl, P., and Bergo, M.O. (2014). Antioxidants Accelerate Lung Cancer Progression in Mice. Science Translational Medicine 6, 221ra15–221ra15.

Schorbach, L., Krick, W., Burckhardt, G., and Burckhardt, B.C. (2013). Glutathione is a low-affinity substrate of the human sodium-dependent dicarboxylate transporter. Nephron Physiol 124, 1–5.

Schulze, K., Imbeaud, S., Letouze, E., Alexandrov, L.B., Calderaro, J., Rebouissou, S., Couchy, G., Meiller, C., Shinde, J., Soysouvanh, F., et al. (2015). Exome sequencing of hepatocellular carcinomas identifies new mutational signatures and potential therapeutic targets. Nat Genet 47, 505–511.

Shalem, O., Sanjana, N.E., Hartenian, E., Shi, X., Scott, D.A., Mikkelsen, T.S., Heckl, D., Ebert, B.L., Root, D.E., Doench, J.G., et al. (2014). Genome-scale CRISPR-Cas9 knockout screening in human cells. Science 343, 84–87.

Shibata, T., Ohta, T., Tong, K.I., Kokubu, A., Odogawa, R., Tsuta, K., Asamura, H., Yamamoto, M., and Hirohashi, S. (2008). Cancer related mutations in NRF2 impair its recognition by Keap1-Cul3 E3 ligase and promote malignancy. Proceedings of the National Academy of Sciences 105, 13568–13573.

Singh, A., Happel, C., Manna, S.K., Acquaah-Mensah, G., Carrerero, J., Kumar, S., Nasipuri, P., Krausz, K.W., Wakabayashi, N., Dewi, R., et al. (2013). Transcription factor NRF2 regulates miR-1 and miR-206 to drive tumorigenesis. J Clin Invest 123, 2921–2934.

Singh, A., Misra, V., Thimmulappa, R.K., Lee, H., Ames, S., Hoque, M.O., Herman, J.G., Baylin, S.B., Sidransky, D., Gabrielson, E., et al. (2006). Dysfunctional KEAP1–NRF2 Interaction in Non-Small-Cell Lung Cancer. PLoS Med 3, e420.

Singh, R., Lemire, J., Mailloux, R.J., and Appanna, V.D. (2008). A novel strategy involved in anti-oxidative defense: the conversion of NADH into NADPH by a metabolic network. PLoS ONE 3, e2682.

Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S., et al. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci USA 102, 15545–15550.

Sugden, M.C., and Holness, M.J. (2011). The pyruvate carboxylase-pyruvate dehydrogenase axis in islet

pyruvate metabolism: Going round in circles? Islets 3, 302–319.

Sun, T., Aceto, N., Meerbrey, K.L., Kessler, J.D., Zhou, C., Migliaccio, I., Nguyen, D.X., Pavlova, N.N., Botero, M., Huang, J., et al. (2011). Activation of Multiple Proto-oncogenic Tyrosine Kinases in Breast Cancer via Loss of the PTPN12 Phosphatase. Cell 144, 703–718.

Sun, X., Ou, Z., Chen, R., Niu, X., Chen, D., Kang, R., and Tang, D. (2015). Activation of the p62-Keap1-NRF2 Pathway Protects against Ferroptosis in Hepatocellular Carcinoma Cells. Hepatology.

Suzuki, T., and Yamamoto, M. (2015). Molecular basis of the Keap1–Nrf2 system. Free Radic Biol Med 88, Part B IS -, 93–100.

Totoki, Y., Tatsuno, K., Covington, K.R., Ueda, H., Creighton, C.J., Kato, M., Tsuji, S., Donehower, L.A., Slagle, B.L., Nakamura, H., et al. (2014). Trans-ancestry mutational landscape of hepatocellular carcinoma genomes. Nat Genet 46, 1267–1273.

Wang, T., Wei, J.J., Sabatini, D.M., and Lander, E.S. (2014). Genetic screens in human cells using the CRISPR-Cas9 system. Science 343, 80–84.

Wälchli, S., Espanel, X., and Hooft van Huijsduijnen, R. (2005). Sap-1/PTPRH activity is regulated by reversible dimerization. Biochem Biophys Res Commun 331, 497–502.

Xu, J., Han, J., Long, Y.S., Epstein, P.N., and Liu, Y.Q. (2008). The role of pyruvate carboxylase in insulin secretion and proliferation in rat pancreatic beta cells. Diabetologia 51, 2022–2030.

Yeatman, T.J. (2004). A renaissance for SRC. Nat Rev Cancer 4, 470–480.

Yeddula, N., Xia, Y., Ke, E., Beumer, J., and Verma, I.M. (2015). Screening for tumor suppressors: Loss of ephrin receptor A2 cooperates with oncogenic KRas in promoting lung adenocarcinoma. Proceedings of the National Academy of Sciences.

Zoncu, R., Efeyan, A., and Sabatini, D.M. (2010). mTOR: from growth signal integration to cancer, diabetes and ageing. Nature Publishing Group 12, 21–35.

Chapter 4

Discussion and future directions

Part I - Constructing cancer using CRISPR-Cas9

The need for faithful experimental platforms to construct and deconstruct cancer

The motivation behind a large portion of the work that I have carried out throughout my PhD studies directly stems from the desire to model and study human cancer as faithfully as possible. Cancer is indeed a genetic disease defined by the sequential acquisition of mutations in tumor suppressor genes and oncogenes. However, cancer is also a highly heterogeneous disease at the genetic and epigenetic level. Cancers do not just evolve through a crystal clear genetic path, whereby all patients suffering from a given cancer harbor the same exact complement of mutations. It is just not that simple.

Despite the fact that multiple seminal studies over the last few decades have elucidated the main drivers (or so-called mountains) behind most cancer types, it is becoming increasingly clear that weaker drivers (or so-called hills) can also play fundamental roles in several aspects of the cancer phenotype. Indeed, the notion that no two cancer genomes are alike has sparked a scientific movement that aims to achieve a more comprehensive view of the cancer genomes of individual patients. This so-called personalized oncology approach aims to tailor present and future cancer therapies to specific patients based on the systematic and comprehensive genomic assessment of their tumors (Garraway et al., 2013). To this end, the creation of large, multi-institutional

consortia, such as The Cancer Genome Atlas, and the subsequent execution of large-scale cancer genome sequencing studies, is an important first step towards achieving this goal. Nevertheless, these invaluable efforts, and the highly curated catalogs of recurrent and non-recurrent mutations they generate can only take the field of cancer research so far. We need more than just cataloguing.

Indeed, one of the main limitations of these studies has been the inability to rapidly and systematically determine which of these mutations play a fundamental role in the context of tumorigenesis, cancer progression, and/or therapeutic response. This limitation is more evident in the context of faithful GEMM-based models of the disease, due to the fact that genetically engineering novel mutations into these models has traditionally been very costly and time consuming. Nevertheless, I would argue that studying the role of novel mutations should employ appropriate models that closely recapitulate multiple aspects of the disease, including whole-organism physiology and metabolism, as well as the presence of an intact immune system. GEMMs and non-germline GEMMs of cancer effectively fulfill all of these criteria and many others, which makes them an invaluable gold standard in the field of cancer genetics (for a recent excellent review, see Gould et al., 2015). Therefore, I would strongly argue that just as the enzymologist spends countless hours developing and benchmarking a highly sensitive and reproducible biochemical assay with which to dissect the role of a particular enzyme, the cancer biologist should also ensure that the experimental system employed to study the role of a putative cancer gene or to perform comprehensive pre-clinical therapeutic studies is as faithful and physiologically relevant as possible.

Otherwise, are we actually studying relevant aspects of the disease? Will we achieve the required clinical impact?

Constructing cancer *in vivo* using CRISPR-Cas9

To tackle the aforementioned challenges, I have devoted a significant portion of my PhD career to developing, validating, and applying novel experimental approaches based on emerging genome engineering technologies for studying cancer in faithful GEMMs of the disease. In **Chapter 2**, I described the development and application of pSECC, a novel lentiviral platform that leverages the power of faithful pre-clinical GEMMs of cancer in combination with the highly precise and versatile CRISPR-Cas9 genome editing system. This novel platform allowed us to assess the molecular and phenotypic consequences of somatically mutating three tumor suppressor genes (*Nkx2.1*, *Pten*, and *Apc*) *in vivo* in the context of autochthonous, immunocompetent, *Kras*-driven GEMMs of lung adenocarcinoma, without having to generate additional mouse models expressing Cre/loxP-conditional alleles or perform extensive mouse husbandry to obtain compound mutant mice harboring conditional alleles in each of these three genes. The unmatched flexibility that this platform offers has already allowed our group and others to assess the consequences of mutating several putative cancer genes *in vivo*, all within the context of autochthonous, immunocompetent GEMMs of lung cancer (see **Figure 1** in the next page and **Appendix 2**; other data not shown). This platform has also been successfully adapted for generating novel GEMMs of other epithelial malignancies, including colorectal adenocarcinoma (data not shown) and pancreatic adenocarcinoma (Mazur et al., 2015). Therefore, this system has effectively filled a major gap between

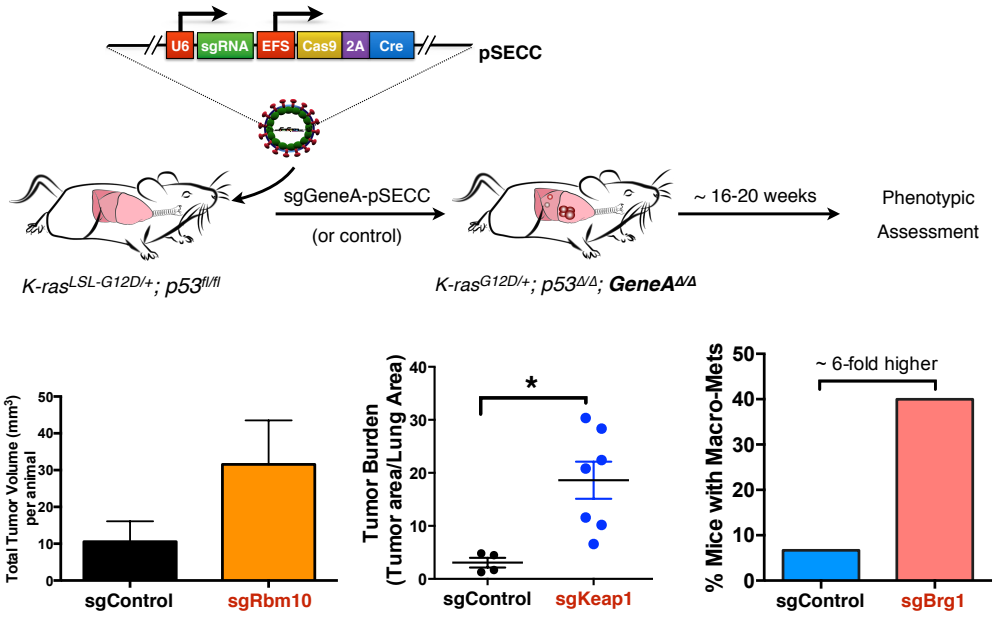


Figure 1: pSECC-mediated validation of *Rbm10*, *Keap1*, and *Brg1* as lung cancer tumor suppressor genes.

the identification of recurrent mutations from large-scale cancer genome sequencing studies and the streamlined validation and characterization of these putative cancer genes *in vivo* using well established GEMMs of cancer. We can now - for the first time - obviate the need for time consuming manipulations of the mouse germline and extensive mouse husbandry in order to generate appropriately sized cohorts for assessing whether a given putative cancer gene is indeed a driver (or a relevant genetic modifier of the cancer phenotype) or a mere passenger. I would argue that we have finally entered an era where cancer biologists can study the disease with unprecedented precision, and within the context of appropriate, physiologically relevant animal models. In addition to some of the broad cancer-centric applications of the CRISPR-Cas9 system for modeling and deconstructing cancer *in vivo*, which I extensively discussed in **Part II** of the introduction, what's next?

Interrogating low frequency drivers - do they actually matter?

As briefly discussed in **Part III** of the introduction, the current consensus is that the cancer genome sequencing studies carried out so far have identified ~ 125 driver genes (of which ~ 71 appear to be tumor suppressor genes and ~ 54 appear to be oncogenes) (Vogelstein et al., 2013). Is the catalog of driver genes complete? Functionally and statistically speaking, the answer appears to be no. First, the identification of these genes has mostly relied on defining them based on mutation frequency, an approach that is inherently biased towards the identification of only the most potent driver genes (reviewed in Pon and Marra 2015). In fact, cancer genome mutational landscapes across different tumor types tend to be dominated by genes that are mutated in <5% of the samples (Wood et al., 2007). Moreover, there is solid experimental evidence suggesting that low-frequency drivers (such as *CUX1*, which is mutated only in ~ 1-5% of all cancers) can play important roles in multiple aspects of the cancer phenotype (Wong et al., 2014). Furthermore, recent stringent bioinformatic analyses have proposed that there are ~ 320 tumor suppressor genes and ~ 250 oncogenes, suggesting that many more drivers remain to be discovered (Davoli et al., 2013). The fact that some of these unidentified drivers may be weaker drivers than, say *KRAS*, *PIK3CA*, *EGFR*, *RB*, or *TP53*, does not necessarily mean that their mutation will not have profound consequences in the right context and, more importantly, in the right patient. Indeed, it is conceivable that mutation of these low frequency drivers might certainly modify certain aspects of the cancer phenotype independent of other potent drivers, including adaptation to certain metabolic conditions through the rewiring of metabolic pathways, adaptation to foreign tissue micro-environments during the

metastatic cascade, and perhaps even response and resistance to certain broad and targeted therapies, among others. The truth is that we will not know if any of these low frequency drivers are playing a functional role in cancer until they are rigorously tested in appropriate experimental models. CRISPR-based systems, such as pSECC, will play a very important role in the quest to interrogate these low-frequency drivers. Furthermore, multiplexable CRISPR-based approaches will allow for testing whether any of these low frequency drivers can functionally cooperate with other low or high frequency drivers in multiple aspects of cancer. For example, low-frequency drivers that are found to be co-mutated with oncogenic *KRAS* (see for example those in **Figure 2A** in **Part III** of the introduction) might functionally cooperate to promote certain aspects of the malignant phenotype. I expect that many more low-frequency drivers (both oncogenes and tumor suppressor genes) will be rigorously validated in the next several years, leading to a substantial expansion of the catalog of true cancer drivers across multiple types of cancers. Furthermore, I expect that many of these low-frequency drivers will bring unexpected surprises to the field of cancer genetics, just as the so-called “exceptional responders” have brought to light fundamental aspects of the cancer phenotype that could only be elucidated when focusing on the smaller percentage of unusual cases (for example, $EGFR^{L858R}$ mutations and response to gefitinib (Lynch et al., 2004, Paez et al., 2004, Pao et al., 2004), *MTOR* mutations and response to everolimus (Wagle et al., 2014), and many others; for a review discussing exceptional responders, see Chau and Lorch 2015). This is an essential step to fulfill the promise of the post cancer-genome era and the promise of precision oncology approaches.

Somatic copy number alterations, haploinsufficiency, and triplosensitivity

Beyond somatic point mutations and indel mutations, cancer genomes are frequently characterized by harboring multiple somatic copy number alterations, including focal and large hemizygous deletions and amplifications (Beroukhim et al., 2010, Bignell et al., 2010, Zack et al., 2013). Mechanistically, hemizygous deletions can simultaneously inactivate multiple linked tumor suppressor genes, whereas amplifications can simultaneously lead to overexpression of multiple growth promoting genes, including many *bona fide* oncogenes. There is a substantial amount of functional evidence demonstrating that genomic regions that frequently undergo hemizygous deletions are enriched for linked tumor suppressor genes (Scuoppo et al., 2012, Xue et al., 2012, Solimini et al., 2012, Davoli et al., 2013, Chen et al., 2014, Zhao et al., 2015). However, beyond a handful of tumor suppressor genes present in commonly deleted regions, such as *TP53* in 17p13.1 (Soenen et al., 1998), the precise location and identity of cooperating tumor suppressor genes in most of these genomic regions remain largely unknown. Traditional approaches for *functionally* studying large chromosomal regions have mainly relied on two methodologies: (1) chromosome engineering (Bagchi et al., 2007, Barlow et al., 2009, and Wong et al., 2015) for deleting large genomic regions, or (2) RNAi-based interrogation of candidate genes within the genomic regions (Ebert et al., 2008, Scuoppo et al., 2012, Xue et al., 2012). Chromosome engineering, which allows for deletion of large genomic regions encompassing both protein-coding and non-protein coding genes, is costly and labor-intensive. RNAi-based approaches have inherent limitations with multiplexability, which usually restrains the methodology to the co-inactivation of a maximum of 2-3 genes and is mostly limited to protein coding

genes. As I extensively described in **Part II** of the introduction, the CRISPR-Cas9 system readily overcomes these and other limitations associated with modeling and studying large cancer-associated genomic aberrations. Therefore, researchers are now finally able to properly interrogate these large genomic regions, which should lead to the identification and functional characterization of multiple haploinsufficient tumor suppressor genes whose combined inactivation via hemizygous deletions can have major consequences during tumorigenesis (Solimini et al., 2012, Davoli et al., 2013). Beyond tumor suppressor genes, the implementation of multiplexed CRISPR-based gene activation approaches (see **Figure 8** in **Part II** of the introduction) should also allow for modeling and studying focal genomic amplifications, which are presumably enriched for growth promoting genes and oncogenes. It will be possible to systematically identify and characterize the specific protein and non-protein coding growth promoting genes (including oncogenes) within these recurrently amplified genomic regions, whose combined overrepresentation can have a major cumulative effect in the context of tumorigenesis through the phenomenon of triplosensitivity (Davoli et al., 2013). I expect that the next few years will witness multiple mechanistic studies involving the systematic dissection of genomic regions frequently lost in human cancers, such as 9p21.3 in lung adenocarcinoma (see **Figure 1B** in **Part III** of the introduction) and the 5q, 7q and 17p regions in karyotypically complex acute leukemias (Mrozek 2008), as well as those frequently gained in human cancers, such as 5p15.33 in lung adenocarcinoma (see **Figure 1C** in **Part III** of the introduction). These systematic experiments were just not technically feasible 2 years ago. CRISPR-based

systems, such as pSECC, will play a crucial role in dissecting these largely uncharacterized cancer-associated genomic regions.

Part II - Deconstructing cancer using CRISPR-Cas9

CRISPR screens for functionally deconstructing cancer

As I mentioned at the beginning of this chapter, one of the main motivations behind the work described in this thesis is the desire to develop and analyze models of cancer that recapitulate several important aspects of the disease. In addition to serving a critical role for studying some of the basic aspects of tumorigenesis, progression, and therapeutic response, these models can also be functionally deconstructed through genetic or small molecule screens in order to identify potential vulnerabilities that could be exploited for therapeutic purposes. Despite the fact that cell culture-based models do not fully recapitulate the organismal physiology of a sophisticated immunocompetent GEMM of cancer, and do not faithfully mimic the sporadic nature of human cancers, they have a proven track record for uncovering multiple genetic dependencies, some of which end up successfully validating in appropriate mouse models *in vivo* and are subsequently pursued for therapeutic purposes (see for example Zuber et al., 2011 and Crystal et al., 2014). The combination of GEMM-derived cell lines (such as those derived from the KP model of lung adenocarcinoma; Jackson et al., 2005) with CRISPR-Cas9 technologies allows one to create customized experimental platforms harboring one or more mutations of interest within the context of a relatively homogeneous genetic background. These experimental platforms offer a unique opportunity for engineering clinically relevant mutations with the goal of creating

isogenic cell lines carrying specific genotypes frequently observed in cancer patients, which can be subsequently deconstructed using multiple approaches, including CRISPR screens and small molecule screens.

In **Chapter 3**, I described the generation of an isogenic pair of *Kras*$^{G12D/+}$;*p53*$^{\Delta/\Delta}$ (KP) and *Kras*$^{G12D/+}$;*p53*$^{\Delta/\Delta}$;*Keap1*$^{\Delta/\Delta}$ (KPK) cell lines using CRISPR-Cas9. To discover novel genetic dependencies that are uniquely essential for cancers harboring *Keap1* mutations, I established and employed a novel CRISPR-Cas9-based experimental pipeline that allows for the systematic interrogation of cancer cell lines harboring clinically relevant mutations frequently observed in human cancers. I subjected the KP and KPK set of isogenic cell lines to a CRISPR-based genetic screen aimed at uncovering novel druggable genetic dependencies that are uniquely essential for cancers harboring *Keap1* mutations. These studies led to the identification of multiple *Keap1*-mutant synthetic lethal candidates, including several well described canonical transcriptional targets of Nrf2, as well as genes involved in GSH synthesis and utilization, NADPH production, amino acid metabolism, and the oncogenic Myc transcriptional network. The data described and discussed extensively in **Chapter 3** has multiple therapeutic implications, as there are currently no known synthetic lethal interactions with either *Keap1* loss of function or *Nrf2* gain of function mutations. Indeed, if any of these putative *Keap1* mutant synthetic lethal interactions were to be extensively validated, they could be readily exploited for the benefit of a large percentage of cancer patients suffering from tumors characterized by hyperactivation of the Nrf2 pathway, including lung cancer patients, pancreatic cancer patients, and hepatocellular carcinoma patients (LUAD TCGA 2014, Hammerman et al., 2012, DeNicola et al., 2011,

Guichard et al., 2012, Totoki et al., 2014, Schulze et al., 2015). Experimental approaches, such as the one described in **Chapter 3**, are essential in order to fulfill the promise of precision oncology. I will discuss below the therapeutic implications of one of these genes and the future work that needs to be carried out to confirm and extend these screening results in multiple other settings.

Fth1 - a promising *Keap1*-mutant synthetic lethal target

As discussed in **Part IV** of the introduction and in **Chapter 3**, the ferritin complex plays a critical cytoprotective role in conditions of high oxidative stress via the conversion of highly reactive Fe(II) molecules into non-reactive Fe(III) molecules. Free Fe(II) molecules, which are released from Fe(II)-containing molecules via the action of Hmox1, promote the generation of highly reactive hydroxyl free radicals ($OH\cdot$) from hydrogen peroxide (H_2O_2) through the Fenton reaction (Gozzelino et al., 2010, Gorrini et al., 2013). Hydroxyl free radicals are highly cytotoxic, and can trigger massive cellular toxicity by directly damaging multiple organelles and macromolecules, such as DNA, proteins, and lipids (Gutteridge 1986, Gozzelino et al., 2010). Fth1 - as part of the ferritin complex - promotes ferritin-mediated Fe(II) → Fe(III) detoxification. In this scenario, pharmacological or genetic inactivation of Fth1 would result in the massive accumulation of highly reactive and cytotoxic hydroxyl free radicals, which would result in cell death (**Figure 2**). Therefore, hyperactivation of the Nrf2 pathway via loss of function mutations in *Keap1*, which is accompanied by hyperactivation of both *Hmox1* and *Fth1*, potentially creates a unique Achilles heel that could be exploited for the benefit of patients harboring these types of mutations. I believe that *Fth1* represents a

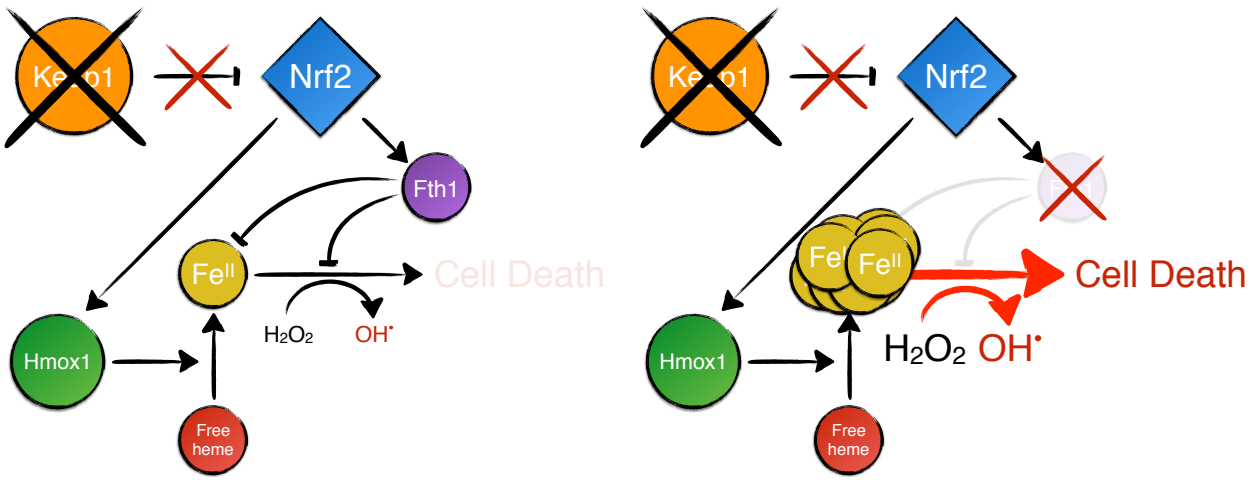


Figure 2: *Fth1* is a promising Keap1-mutant synthetic lethal target. See text for mechanistic details.

novel synthetic lethal interaction with *Keap1* mutation (or otherwise constitutive Nrf2 activation) that might offer a unique and attractive opportunity for selectively targeting tumors that are highly dependent on this antioxidant pathway. The fact that a large percentage of lung cancer, pancreatic cancer, and liver cancer patients might benefit from such a therapeutic strategy makes it warrant further investigation.

Future work: validating and extending synthetic lethal screens

The work described in **Chapter 3** and the identification of multiple putative *Keap1*-mutant synthetic lethal targets, such as *Fth1*, has great potential to lead to the discovery of clinically actionable drug targets for this particular population of cancer patients. Nevertheless, these specific results, as well as the results from both KP and KPK CRISPR screens in general, need to be extensively validated by performing secondary screens in both murine and human cell lines. To do this, I would propose to perform secondary CRISPR-based screens focusing on the top 50 genes with the most negative KPK-KP Differential Gene Scores (for *Keap1*-mutant synthetic lethal screens) or the top

50 genes that deplete the most in both KP and KPK cells. These secondary sgRNA libraries could consist of ~ 1,000 sgRNAs (at 20 sgRNAs/gene) targeting either mouse or human genes and could be screened across (1) a panel of murine KP and KPK cell lines, and (2) a panel composed of human lung cancer cell lines with known *KEAP1* mutational status (such as H460, A549, H1435, and H838) (Singh et al., 2006), as well as CRISPR-engineered isogenic human *KEAP1* mutant lung adenocarcinoma cell lines. Importantly, the level of saturation obtained by testing 20 sgRNAs/gene will ensure that only the most promising gene candidates are pursued further. After identifying the most promising putative *Keap1* mutant-specific drug targets, one could systematically test their requirement through a series of *in vitro* and *in vivo* experiments in both murine and human lung adenocarcinoma cell lines. To rapidly test the functional requirement of gene candidates for proliferation of KPK murine cell lines *in vitro*, one can carry out multi-color competition assays using Cas9-expressing cell populations partially transduced with either control vectors or vectors expressing sgRNAs targeting putative *Keap1*-mutant synthetic lethal targets. Importantly, to rule out the possibility that these synthetic lethal targets represent the idiosyncrasies of one or a few particular cell lines, one should seek to extend validation experiments beyond the aforementioned cell culture systems to include, if possible, primary patient-derived xenografts with known *KEAP1* loss of function or *NFE2L2* gain of function mutations.

Beyond these *in vitro* experiments, one should perform stringent validation of these putative synthetic lethal targets by testing the requirement of these candidate genes for tumor initiation and tumor maintenance *in vivo*. To do this, one could clone the most potent sgRNAs targeting the candidate gene into either constitutive or inducible sgRNA

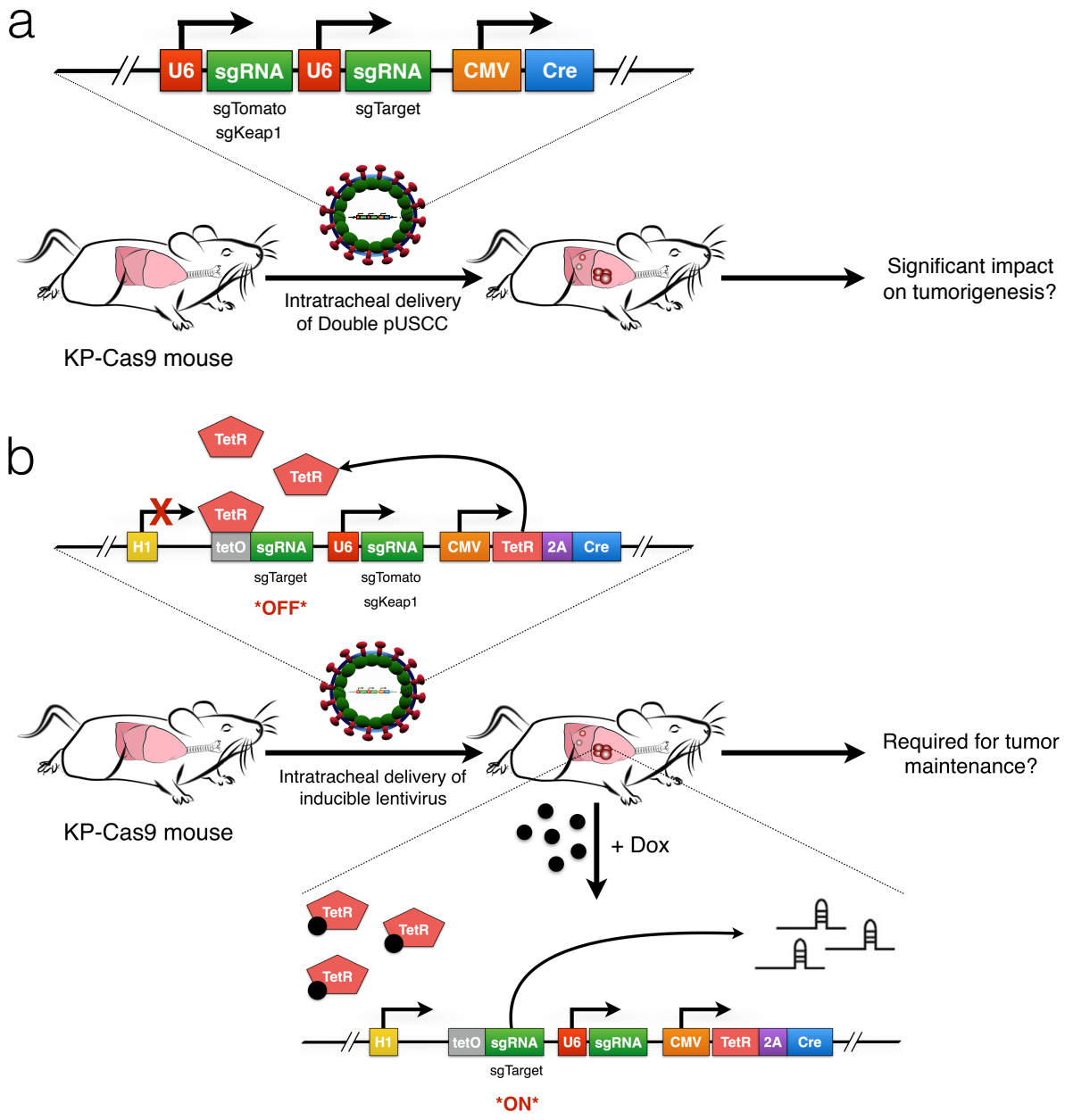


Figure 3: Functional characterization of putative *Keap1*-mutant synthetic lethal targets in GEMMs of lung adenocarcinoma. (a) Intratracheal delivery of dual sgRNA-expressing vectors concomitantly expressing sgTomato-sgTarget (control arm) or sgKeap1-sgTarget (experimental arm) combinations and Cre recombinase into KP-Cas9 mice would allow for functional validation of candidate Keap1-mutant synthetic lethal targets *in vivo*. (b) Intratracheal delivery of inducible lentiviral vectors concomitantly expressing a doxycycline-inducible sgRNA (targeting either Tomato as control or *Keap1*), the Tet repressor, and Cre recombinase would allow for determining whether the candidate *Keap1*-mutant synthetic lethal target is required for tumor maintenance. See text for additional details. Figure 1b was generated with Chris Ellis.

vectors, such as those based on the Tet-On system (Aubrey et al., 2015). The former would test whether the putative synthetic lethal interaction is functionally important in the context of transformation and tumor initiation, whereas the latter would test whether it is functionally important for tumor progression and tumor maintenance (see **Figure 3**). Similar experiments could also be performed with small molecule inhibitors of target proteins, where available.

The studies described in **Chapter 3** employed a custom sgRNA library targeting the Druggable Genome, which is of course a relatively biased approach if one were interested in thoroughly deconstructing the biology of the Keap1-Nrf2 pathway. To extend these efforts, one can perform genome-wide scale CRISPR screens in a collection of murine and human lines with known mutational alterations in the Keap1-Nrf2 pathway (Wang et al., 2014, Shalem et al., 2014, Sanjana et al., 2014, Wang et al., 2015). By employing genome-wide sgRNA libraries, one could unbiasedly and comprehensively identify additional novel genetic vulnerabilities in the context of *Keap1*-mutant cancer, as well as potentially identify novel components of the Keap1-Nrf2 pathway. Notably, the latter discoveries would contribute to a greater mechanistic understanding of this important antioxidant pathway. For example, given the fact that *Keap1*-mutant cells have increased Nrf2 levels compared to *Keap1*-wild type cells, disruption of a functionally important gene upstream or downstream of Nrf2 should sensitize the cells to conditions of high oxidative stress, such as those induce by dimethyl fumarate (DMF). Therefore, the significant expansion of the screening strategies described in **Chapter 3** combined with the implementation of validation experiments beyond traditional *in vitro* proliferation assays and *in vivo* tumor growth and

tumor maintenance assays could allow one to achieve a deeper mechanistic understanding of the role of candidate genes in the biology of this pathway. It is also conceivable that genome-wide screens could lead to the discovery of novel direct Nrf2 transcriptional targets. Beyond nuclease-based CRISPR screens, one could conceivably extend the spectrum of synthetic lethal targets (as well as the spectrum of genes directly involved in multiple mechanistic aspects of the biology of the Keap1-Nrf2 pathway) by implementing complementary CRISPR-activator screens (described in **Part II** of the introduction; see **Figures 8 and 9** in that specific section). Indeed, approaches based on CRISPR-Cas9 technologies offer an endless amount of experimental possibilities that can lead to the comprehensive deconstruction of biological pathways that are highly relevant to cancer biology, such as the Keap1-Nrf2 pathway.

It is hard to believe that the Jinek *et al.* manuscript was published just a little over 3 years ago (Jinek et al., 2012). Now in 2015, it is hard to imagine an area of cancer research - from basic biology to clinical oncology - that is not positively impacted by the CRISPR-Cas9 system. The next several years of cancer research will be truly transformative.

REFERENCES

Aubrey, B.J., Kelly, G.L., Kueh, A.J., Brennan, M.S., O'Connor, L., Milla, L., Wilcox, S., Tai, L., Strasser, A., and Herold, M.J. (2015). An Inducible Lentiviral Guide RNA Platform Enables the Identification of Tumor-Essential Genes and Tumor-Promoting Mutations In Vivo. Cell Rep 10, 1422–1432.

Bagchi, A., Papazoglu, C., Wu, Y., Capurso, D., Brodt, M., Francis, D., Bredel, M., Vogel, H., and Mills, A.A. (2007). CHD5 is a tumor suppressor at human 1p36. Cell 128, 459–475.

Barlow, J.L., Drynan, L.F., Hewett, D.R., Holmes, L.R., Lorenzo-Abalde, S., Lane, A.L., Jolin, H.E., Pannell, R., Middleton, A.J., Wong, S.H., et al. (2009). A p53-dependent mechanism underlies macrocytic anemia in a mouse model of human 5q- syndrome. Nat Med 16, 59–66.

Beroukhim, R., Mermel, C.H., Porter, D., Wei, G., Raychaudhuri, S., Donovan, J., Barretina, J., Boehm, J.S., Dobson, J., Urashima, M., et al. (2010). The landscape of somatic copy-number alteration across human cancers. Nature 463, 899–905.

Bignell, G.R., Greenman, C.D., Davies, H., Butler, A.P., Edkins, S., Andrews, J.M., Buck, G., Chen, L., Beare, D., Latimer, C., et al. (2010). Signatures of mutation and selection in the cancer genome. Nature 463, 893–898.

Cancer Genome Atlas Research Network (2014). Comprehensive molecular profiling of lung adenocarcinoma. Nature 511, 543–550.

Chau, N.G., and Lorch, J.H. (2015). Exceptional Responders Inspire Change: Lessons for Drug Development From the Bedside to the Bench and Back. The Oncologist 20, 699–701.

Chen, C., Liu, Y., Rappaport, A.R., Kitzing, T., Schultz, N., Zhao, Z., Shroff, A.S., Dickins, R.A., Vakoc, C.R., Bradner, J.E., et al. (2014). MLL3 is a haploinsufficient 7q tumor suppressor in acute myeloid leukemia. Cancer Cell 25, 652–665.

Crystal, A.S., Shaw, A.T., Sequist, L.V., Friboulet, L., Niederst, M.J., Lockerman, E.L., Frias, R.L., Gainor, J.F., Amzallag, A., Greninger, P., et al. (2014). Patient-derived models of acquired resistance can identify effective drug combinations for cancer. Science.

Davoli, T., Xu, A.W., Mengwasser, K.E., Sack, L.M., Yoon, J.C., Park, P.J., and Elledge, S.J. (2013). Cumulative Haploinsufficiency and Triplosensitivity Drive Aneuploidy Patterns and Shape the Cancer Genome. Cell 155, 948–962.

Denicola, G.M., Karreth, F.A., Humpton, T.J., Gopinathan, A., Wei, C., Frese, K., Mangal, D., Yu, K.H., Yeo, C.J., Calhoun, E.S., et al. (2011). Oncogene-induced Nrf2 transcription promotes ROS detoxification and tumorigenesis. Nature 475, 106–109.

Ebert, B.L., Pretz, J., Bosco, J., Chang, C.Y., Tamayo, P., Galili, N., Raza, A., Root, D.E., Attar, E., Ellis, S.R., et al. (2008). Identification of RPS14 as a 5q- syndrome gene by RNA interference screen. Nature 451, 335–339.

Garraway, L.A., Verweij, J., and Ballman, K.V. (2013). Precision oncology: an overview. J Clin Oncol 31, 1803–1805.

Gorrini, C., Harris, I.S., and Mak, T.W. (2013). Modulation of oxidative stress as an anticancer strategy. Nature Reviews Drug Discovery 12, 931–947.

Gould, S.E., Junttila, M.R., and de Sauvage, F.J. (2015). Translational value of mouse models in oncology drug development. Nature Medicine 21, 431–439.

Gozzelino, R., Jeney, V., and Soares, M.P. (2010). Mechanisms of Cell Protection by Heme Oxygenase-1. Annu. Rev. Pharmacol. Toxicol. 50, 323–354.

Guichard, C., Amaddeo, G., Imbeaud, S., Ladeiro, Y., Pelletier, L., Maad, I.B., Calderaro, J., Bioulac-Sage, P., Letexier, M., Degos, F., et al. (2012). Integrated analysis of somatic mutations and focal copy-number changes identifies key genes and pathways in hepatocellular carcinoma. Nat Genet 44, 694–698.

Gutteridge, J.M. (1986). Iron promoters of the Fenton reaction and lipid peroxidation can be released from haemoglobin by peroxides. FEBS Lett 201, 291–295.

Hammerman, P.S., Lawrence, M.S., Voet, D., Jing, R., Cibulskis, K., Sivachenko, A., Stojanov, P., McKenna, A., Lander, E.S., Gabriel, S., et al. (2012). Comprehensive genomic characterization of squamous cell lung cancers. Nature 489, 519–525.

Jackson, E., Olive, K., Tuveson, D., Bronson, R., Crowley, D., Brown, M., and Jacks, T. (2005). The Differential Effects of Mutant p53 Alleles on Advanced Murine Lung Cancer. Cancer Res 65, 10280–10288.

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A., and Charpentier, E. (2012). A Programmable Dual-RNA-Guided DNA Endonuclease in Adaptive Bacterial Immunity. Science 337, 816–821.

Lynch, T.J., Bell, D.W., Sordella, R., Gurubhagavatula, S., Okimoto, R.A., Brannigan, B.W., Harris, P.L., Haserlat, S.M., Supko, J.G., Haluska, F.G., et al. (2004). Activating Mutations in the Epidermal Growth Factor Receptor Underlying Responsiveness of Non–Small-Cell Lung Cancer to Gefitinib. N Engl J Med 350, 2129–2139.

Mazur, P.K., Herner, A., Mello, S.S., Wirth, M., Hausmann, S., Sanchez-Rivera, F.J., Lofgren, S.M., Kuschma, T., Hahn, S.A., Vangala, D., et al. (2015). Combined inhibition of BET family proteins and histone deacetylases as a potential epigenetics-based therapy for pancreatic ductal adenocarcinoma. Nature Medicine 21, 1163–1171.

Mrózek, K. (2008). Cytogenetic, Molecular Genetic, and Clinical Characteristics of Acute Myeloid Leukemia With a Complex Karyotype. Acute Myeloid Leukemia 35, 365–377.

Paez, J.G., Jänne, P.A., Lee, J.C., Tracy, S., Greulich, H., Gabriel, S., Herman, P., Kaye, F.J., Lindeman, N., Boggon, T.J., et al. (2004). EGFR mutations in lung cancer: correlation with clinical response to gefitinib therapy. Science 304, 1497–1500.

Pao, W., Miller, V., Zakowski, M., Doherty, J., Politi, K., Sarkaria, I., Singh, B., Heelan, R., Rusch, V., Fulton, L., et al. (2004). EGF receptor gene mutations are common in lung cancers from “never smokers”

and are associated with sensitivity of tumors to gefitinib and erlotinib. Proc Natl Acad Sci USA 101, 13306–13311.

Pon, J.R., and Marra, M.A. (2015). Driver and Passenger Mutations in Cancer. Annu. Rev. Pathol. Mech. Dis. 10, 25–50.

Sanjana, N.E., Shalem, O., and Zhang, F. (2014). Improved vectors and genome-wide libraries for CRISPR screening. Nat Methods 11, 783–784.

Schulze, K., Imbeaud, S., Letouze, E., Alexandrov, L.B., Calderaro, J., Rebouissou, S., Couchy, G., Meiller, C., Shinde, J., Soysouvanh, F., et al. (2015). Exome sequencing of hepatocellular carcinomas identifies new mutational signatures and potential therapeutic targets. Nat Genet 47, 505–511.

Scuoppo, C., Miething, C., Lindqvist, L., Reyes, J., Ruse, C., Appelmann, I., Yoon, S., Krasnitz, A., Teruya-Feldstein, J., Pappin, D., et al. (2012). A tumour suppressor network relying on the polyamine-hypusine axis. Nature 487, 244–248.

Shalem, O., Sanjana, N.E., Hartenian, E., Shi, X., Scott, D.A., Mikkelsen, T.S., Heckl, D., Ebert, B.L., Root, D.E., Doench, J.G., et al. (2014). Genome-scale CRISPR-Cas9 knockout screening in human cells. Science 343, 84–87.

Singh, A., Misra, V., Thimmulappa, R.K., Lee, H., Ames, S., Hoque, M.O., Herman, J.G., Baylin, S.B., Sidransky, D., Gabrielson, E., et al. (2006). Dysfunctional KEAP1–NRF2 Interaction in Non-Small-Cell Lung Cancer. PLoS Med 3, e420.

Soenen, V., Preudhomme, C., Roumier, C., Daudignon, A., Luc Laï, J., and Fenaux, P. (1998). 17p Deletion in Acute Myeloid Leukemia and Myelodysplastic Syndrome. Analysis of Breakpoints and Deleted Segments by Fluorescence In Situ. Blood 91, 1008–1015.

Solimini, N.L., Xu, Q., Mermel, C.H., Liang, A.C., Schlabach, M.R., Luo, J., Burrows, A.E., Anselmo, A.N., Bredemeyer, A.L., Li, M.Z., et al. (2012). Recurrent Hemizygous Deletions in Cancers May Optimize Proliferative Potential. Science 337, 104–109.

Totoki, Y., Tatsuno, K., Covington, K.R., Ueda, H., Creighton, C.J., Kato, M., Tsuji, S., Donehower, L.A., Slagle, B.L., Nakamura, H., et al. (2014). Trans-ancestry mutational landscape of hepatocellular carcinoma genomes. Nat Genet 46, 1267–1273.

Vogelstein, B., Papadopoulos, N., Velculescu, V.E., Zhou, S., Diaz, L.A., and Kinzler, K.W. (2013). Cancer genome landscapes. Science 339, 1546–1558.

Wagle, N., Grabiner, B.C., Van Allen, E.M., Hodis, E., Jacobus, S., Supko, J.G., Stewart, M., Choueiri, T.K., Gandhi, L., Cleary, J.M., et al. (2014). Activating mTOR mutations in a patient with an extraordinary response on a phase I trial of everolimus and pazopanib. Cancer Discovery 4, 546–553.

Wang, T., Birsoy, K., Hughes, N.W., Krupczak, K.M., Post, Y., Wei, J.J., Lander, E.S., and Sabatini, D.M. (2015). Identification and characterization of essential genes in the human genome. Science 350, 1096–1101.

Wang, T., Wei, J.J., Sabatini, D.M., and Lander, E.S. (2014). Genetic screens in human cells using the CRISPR-Cas9 system. Science 343, 80–84.

Wong, C.C., Martincorena, I., Rust, A.G., Rashid, M., Alifrangis, C., Alexandrov, L.B., Tiffen, J.C., Kober, C., Chronic Myeloid Disorders Working Group of the International Cancer Genome Consortium, Green, A.R., et al. (2014). Inactivating CUX1 mutations promote tumorigenesis. Nat Genet 46, 33–38.

Wong, J.C., Weinfurtner, K.M., Alzamora, M.D.P., Kogan, S.C., Burgess, M.R., Zhang, Y., Nakitandwe, J., Ma, J., Cheng, J., Chen, S.-C., et al. (2015). Functional evidence implicating chromosome 7q22 haploinsufficiency in myelodysplastic syndrome pathogenesis. eLife Sciences 4, e07839.

Wood, L.D., Parsons, D.W., Jones, S., Lin, J., Sjöblom, T., Leary, R.J., Shen, D., Boca, S.M., Barber, T., Ptak, J., et al. (2007). The genomic landscapes of human breast and colorectal cancers. Science 318, 1108–1113.

Xue, W., Kitzing, T., Roessler, S., Zuber, J., Krasnitz, A., Schultz, N., Revill, K., Weissmueller, S., Rappaport, A.R., Simon, J., et al. (2012). A cluster of cooperating tumor-suppressor gene candidates in chromosomal deletions. Proceedings of the National Academy of Sciences 109, 8212–8217.

Zack, T.I., Schumacher, S.E., Carter, S.L., Cherniack, A.D., Saksena, G., Tabak, B., Lawrence, M.S., Zhsng, C.-Z., Wala, J., Mermel, C.H., et al. (2013). Pan-cancer patterns of somatic copy number alteration. Nat Genet 45, 1134–1140.

Zhao, Z., Chen, C.-C., Rillahan, C.D., Shen, R., Kitzing, T., McNerney, M.E., Diaz-Flores, E., Zuber, J., Shannon, K., Le Beau, M.M., et al. (2015). Cooperative loss of RAS feedback regulation drives myeloid leukemogenesis. Nat Genet.

Zuber, J., Shi, J., Wang, E., Rappaport, A.R., Herrmann, H., Sison, E.A., Magoon, D., Qi, J., Blatt, K., Wunderlich, M., et al. (2011). RNAi screen identifies Brd4 as a therapeutic target in acute myeloid leukaemia. Nature 478, 524–528.

Appendix 1

The level of mitochondrial apoptotic priming is a critical determinant of cell fate upon p53 restoration

Francisco J. Sánchez-Rivera[1,2*], Jeremy Ryan[4*], Mary Clare Beytagh[1,2], Yadira M. Soto-Feliciano[1,2], Lucius Xuan[1], David M. Feldser[5], Michael T. Hemann[1,2], Anthony Letai[4], and Tyler Jacks[1,2,3].

[1] David H. Koch Institute for Integrative Cancer Research, Massachusetts Institute of Technology, Cambridge, MA 02142

[2] Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02142

[3] Howard Hughes Medical Institute, Massachusetts Institute of Technology, Cambridge, MA 02139

[4] Dana-Farber Cancer Institute, Harvard Medical School, Boston, MA 02215, USA

[5] Department of Cancer Biology, Abramson Family Cancer Research Institute, and Perelman School of Medicine at the University of Pennsylvania, Philadelphia, PA 19104

[*] These authors contributed equally to this work.

Author Contributions

F.J.S.R., J.R., A.L., and T.J. designed the study; J.R. performed BH3 profiling experiments and ABT-263 experiments; F.J.S.R. performed all other experiments, with assistance from M.C.B., Y.M.S.F., and L.X.; D.M.F. and M.T.H. provided conceptual advice.

ABSTRACT

Using genetically engineered mouse models of cancer, we and others have recently shown that a variety of established tumors require sustained inactivation of the p53 pathway. Reactivation of p53 in established tumors typically results in one of two cell fate decisions: cell cycle arrest or cell death via apoptosis. Since p53 is a transcription factor, one hypothesis that may explain these tumor-specific outcomes is that p53 transcriptionally activates a particular set of genes that is specific to each tumor type, in addition to genes involved in a core p53 pathway. Another possibility is that the level of mitochondrial apoptotic priming in a cell might dictate whether p53 activation promotes cell death or cell cycle arrest. To test this this hypothesis, we derived murine p53-restorable cell lines from three different tumor types: lung adenocarcinoma, sarcoma, and lymphoma. Upon restoration of the endogenous *Trp53* gene, these cell lines undergo cell cycle arrest (lung adenocarcinoma and sarcoma) or apoptosis (lymphoma). We hypothesized that this tumor-specific response to p53 restoration could be explained by the level of mitochondrial apoptotic priming on each specific tumor type, where lymphoma cell lines would be highly primed relative to the lung adenocarcinoma and sarcoma cell lines. Indeed, mitochondrial BH3 profiling demonstrated that lymphoma cell lines were highly primed and this correlated with their apoptotic response to p53 restoration. On the other hand, lung adenocarcinoma and sarcoma cell lines were poorly primed and this was consistent with their cell cycle arrest phenotype upon p53 restoration. Modulating the expression levels of Bcl2 family members to modulate mitochondrial apoptotic priming was sufficient to change the fate of these cells. Lymphoma cell lines were forced to undergo cell cycle arrest upon p53 restoration by

overexpression of the anti-apoptotic Bcl2 family member *Bcl-xL*. In contrast, CRISPR-mediated genetic ablation of *Bcl-xL*, which increased the levels of mitochondrial apoptotic priming, forced lung adenocarcinoma and sarcoma cell lines to undergo cell death upon p53 restoration. Similarly, sarcoma cell lines were forced to undergo cell death upon p53 restoration when priming was increased by overexpression of the pro-apoptotic Bcl2 family member *Bad*. Moreover, established tumors derived from *Bad*-overexpressing cells were specifically sensitized to undergo cell death upon p53 restoration *in vivo*, whereas cells expressing empty vector underwent cell cycle arrest. These findings indicate that the level of mitochondrial apoptotic priming is a critical determinant of cell fate upon p53 restoration. Moreover, our data suggests that the combination of a p53 restoring drug with an agent that increases mitochondrial apoptotic priming will strongly synergize and trigger cell death in tumors that commonly undergo cell cycle arrest upon p53 restoration.

INTRODUCTION

As described in **Part III** of the introduction to this thesis, *TP53* is the most commonly mutated tumor suppressor gene, with approximately 50% of human cancers harboring mutations in this locus (Polager and Ginsberg 2009). The remaining 50% of human cancers typically harbor mutations that directly inactivate the p53 pathway, usually through downregulation or mutation of its main upstream activator $p14^{Arf}$ ($p19^{Arf}$ in mice), or genomic amplification and/or overexpression of its main negative regulator *Mdm2* (Momand et al., 1998). As a consequence, great effort has been placed on elucidating p53's tumor suppressive mechanisms.

Upon activation by a variety of stress signals, such as DNA damage, oncogenic stress, and hypoxia, p53 transcriptionally activates a large repertoire of genes (**Figure 1**).

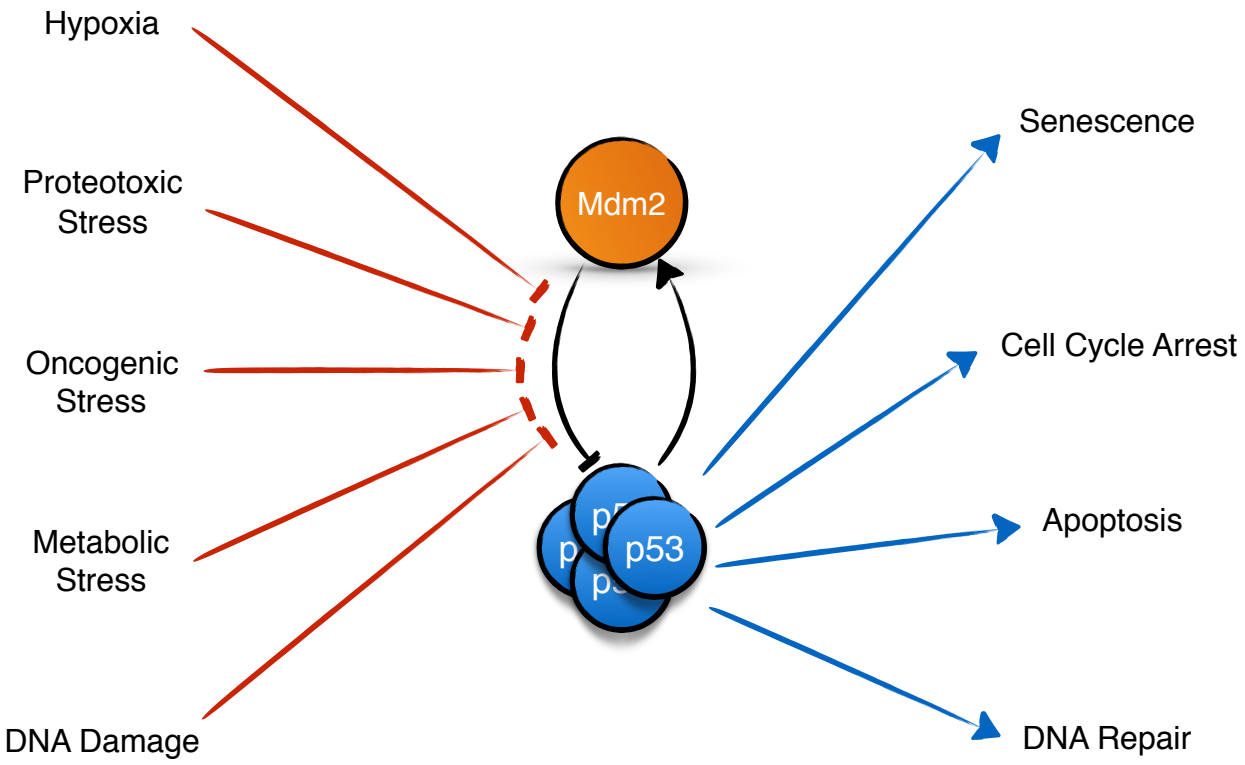


Figure 1: The p53 pathway. Multiple cellular stresses, such as the ones listed on the left, can inhibit Mdm2-mediated negative regulation of p53, which leads to p53 activation and tumor suppression.

This can lead to a variety of outcomes, ranging from cell cycle arrest, senescence, or apoptosis, to changes in metabolism and/or autophagy (Lu, 2010; Lane and Levine, 2010). Due to the prevalence of inactivating mutations in p53, recent efforts have focused on strategies to pharmacologically re-activate mutant p53 as a possible option for treating human cancers that harbor p53 mutations (reviewed in Khoo et al., 2014). This concept has been supported by experiments in GEMMs of cancer (Martins et al., 2006, Ventura et al., 2007, Xue et al., 2007, Feldser et al., 2010). For example, our lab

constructed a mouse model that allows temporal control of p53 expression in established, autochthonous tumors (**Figure 2A**). This mouse model expresses a *LSL-p53* allele in the germline (Ventura et al., 2007). Homozygous mutant mice ($p53^{LSL/LSL}$) are phenotypically identical to homozygous knockout mice ($p53^{-/-}$) (Donehower et al., 1992, Jacks et al., 1994). However, upon induction of Cre recombinase (which in this case is expressed from the ubiquitously expressed *Rosa26* locus (Zambrowicz et al., 1997) as an estrogen receptor fusion protein called Cre-ER^{T2}) and excision of the LSL cassette, p53 transcription is restored. Utilizing this mouse model, our group demonstrated that established tumors that spontaneously arise in the context of p53 deficiency (mostly soft tissue sarcomas and lymphomas) (Donehower et al., 1992, Jacks et al., 1994) remain exquisitely sensitive to the reactivation of p53 (Ventura et al., 2007). Notably, two contemporaneous studies by the groups of Gerard Evan and Scott Lowe reached the same conclusion using estrogen receptor-regulatable systems and inducible-reversible RNAi technologies, respectively (Martins et al., 2006, Xue et al., 2007). In a subsequent study, our group crossed the p53 restorable mouse to the $Kras^{LA2}$ mouse model, which bears a latent allele of oncogenic $Kras^{G12D}$ that gets spontaneously activated *in vivo* (Johnson et al., 2001), for the purposes of studying the effects of p53 reactivation in established lung tumors (Feldser et al., 2010). Surprisingly, we found that the tumor suppressive effects of p53 reactivation in established lung tumors are stage-specific, whereby p53 triggers the elimination of highly advanced lesions but spares low grade lesions. Mechanistically, high grade lesions harbor hyperactive MAPK signaling, which in turn leads to potent induction of the $p19^{Arf}$ tumor suppressor gene, leading to stabilization and activation of the p53 tumor suppressor

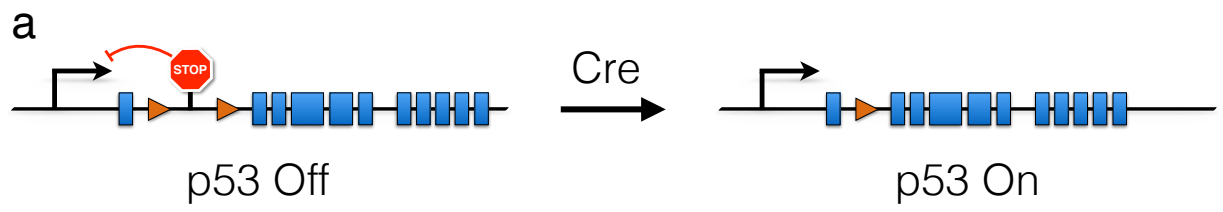
(Feldser et al., 2010). A contemporaneous study by Gerard Evan and colleagues using an estrogen receptor-regulatable allele of *p53* reached the same conclusion (Junttila et al., 2010). Collectively, these GEMM-based studies have convincingly demonstrated that p53 can mediate different tumor suppressive responses in different tumor types and even at different stages of tumor progression, and lend strong support to efforts aimed at pharmacologically reactivating p53 as a possible cancer therapeutic strategy (Khoo et al., 2014).

Despite the fact that p53 has been extensively studied for over 30 years, the molecular mechanisms behind its highly context-specific tumor suppressive responses remain widely unknown (Jackson et al., 2010). For example, it has been suggested that p53-mediated induction of a particular tumor suppressive outcome depends on a combination of promoter selectivity mechanisms and the magnitude and duration of the particular response that induces p53 at that given time (Purvis et al., 2012, and others). Moreover, it has been suggested that the absolute levels of p53 or certain p53-interacting proteins can dictate whether a given response will dominate over the other (Jackson et al., 2010). Another possibility is that the level of mitochondrial apoptotic priming in a cell dictates whether p53 activation promotes cell death or cell cycle arrest (Letai 2008, Ni Chonghaile et al., 2011). Here, I present data that support a model whereby the level of mitochondrial apoptotic priming is a critical determinant of cell fate upon p53 restoration.

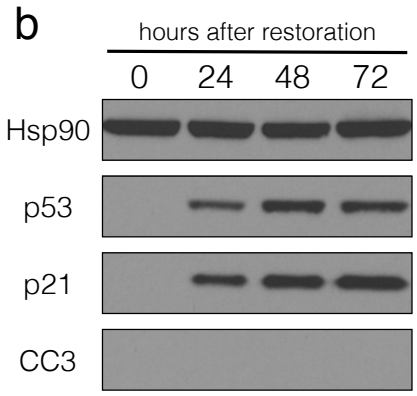
RESULTS

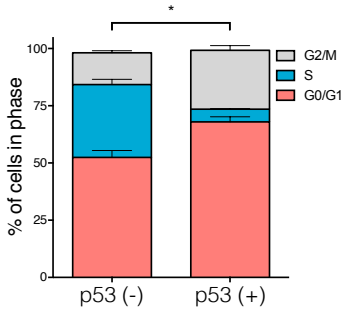
Tumor-specific responses to p53 restoration

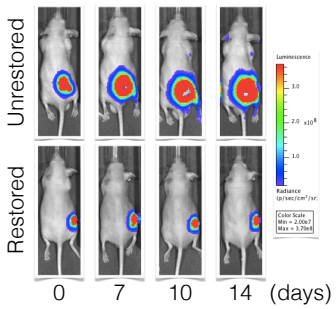
We hypothesized that all tumor-specific responses to p53 restoration could be explained at least in part by the level of mitochondrial priming on each specific tumor type. If this were true, highly primed cells would readily undergo cell death upon p53 restoration, but poorly primed cells would instead undergo cell cycle arrest upon p53 restoration. To test this, we derived *Kras*$^{LA2/+}$*;Trp53*$^{LSL/LSL}$*;Rosa26*$^{CreERT2/CreERT2}$ p53-restorable cell lines (hereafter referred to as just p53 restorable cell lines) from three different tumor types that arise in the background of spontaneous *Kras*^{LA2} mutation (G12D mutation) and p53 deficiency: lung adenocarcinoma, sarcoma, and lymphoma. These cells experience markedly different tumor suppressive responses upon restoration of the endogenous p53 gene (**Figure 2B-J)**. Lung adenocarcinoma and sarcoma cell lines predominantly undergo cell cycle arrest, as shown by the potent induction of p21 (**Figure 2B, E**), by the significant accumulation of p53-restored cells in the G0/G1 stage of the cell cycle (**Figure 2C, F**), by the pronounced tumor stasis observed upon restoration of p53 in established tumors *in vivo* (**Figure D, G**), and by the marked absence of cleaved caspase 3 (CC3), even after prolonged restoration of p53 for 72 hours (**Figure 2B, E**). On the other hand, lymphoma cell lines predominantly undergo apoptosis, as shown by a substantial increase in CC3 levels (**Figure 2H**), and by a significant increase in the percentage of Annexin V-7AAD double positive cells and a substantial increase in the percentage of cells that stain with propidium iodide over time (**Figure 2I-J)**. Remarkably, p21 is also substantially induced upon p53 restoration in lymphoma cells, despite the



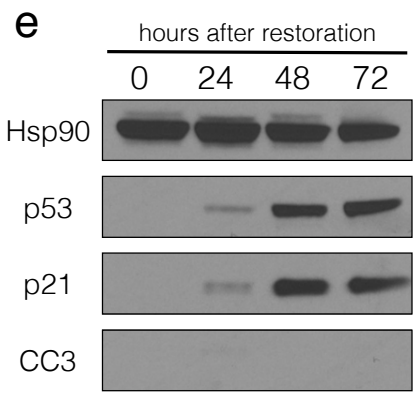
Lung adenocarcinoma

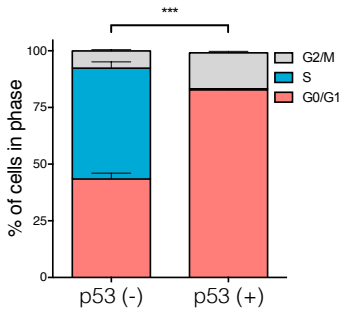


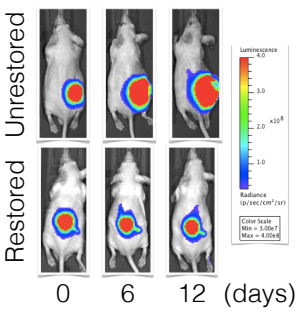




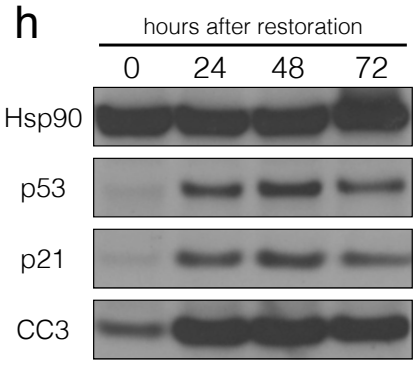
Sarcoma

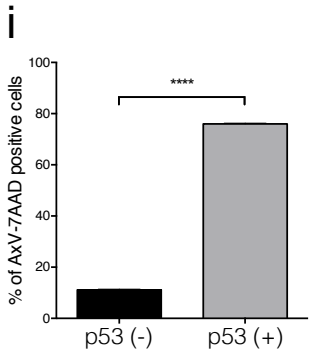






Lymphoma





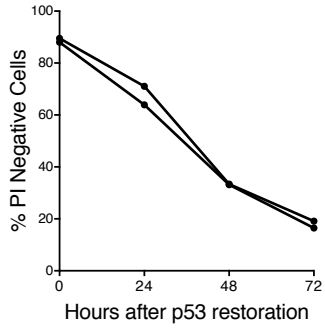


Figure 2: The LSL-p53 allele and tumor-specific responses to p53 restoration. (a) LSL-p53 allele. Upon administration of tamoxifen, Cre is shuttled to the nucleus where it catalyzes the recombination between loxP sites, thereby restoring p53 expression from its endogenous locus. (b) p53 restoration time course in lung adenocarcinoma cell lines. (c) Cell cycle profiling of lung adenocarcinoma cell lines 72hrs after p53 restoration. (d) Lung adenocarcinoma tumors arrest upon p53 restoration *in vivo*. (e) p53 restoration time course in sarcoma cell lines. (f) Cell cycle profiling of sarcoma cell lines 72hrs after p53 restoration. (g) Sarcoma tumors arrest upon p53 restoration *in vivo*. (h) p53 restoration time course in lymphoma cell lines. (i) Percentage of Annexin V-7AAD double positive lymphoma cells 72hrs after p53 restoration. (j) Viability of two independent lymphoma cell lines 72hrs after p53 restoration measured using propidium iodide. Data in (c), (f) and (i) represent the mean ± S.E.M, n=3. Statistics were calculated with two-sided Student's t-test: *P < 0.05, **P < 0.01, ***P < 0.001, ****P < 0.0001. CC3=cleaved caspase 3

fact that these cells are readily undergoing apoptosis (**Figure 2H**). The importance of this observation is two-fold. First, it rules out a potential promoter selectivity-based mechanism that suggests that p53 is preferentially (or exclusively) inducing the expression of pro-apoptotic genes in lymphomas, since p21 is also being substantially induced upon restoration of p53. Therefore, both tumor suppressive programs are being concomitantly activated upon restoration of p53 in lymphomas. Regarding the latter point, these data suggests that the cell death pathway is overriding the cell cycle arrest pathway in lymphomas, due to the fact that p53 restoration readily induces cell death in these cells.

The level of mitochondrial apoptotic priming explains tumor-specific responses

To test whether these tumor-specific responses to p53 restoration are dictated by the level of basal mitochondrial apoptotic priming in each of these different tumor types, we employed a technique known as BH3 profiling, which measures the mitochondrial response to peptides derived from the BH3 domains of pro-apoptotic BH3-only proteins of the Bcl2 family (Ni Chonghaile et al., 2011). In this assay (**Figure 3A**), cells are exposed to BH3 peptides and the resulting mitochondrial outer membrane permeabilization (MOMP) is indirectly measured using the potential-sensitive fluorescent dye JC-1 (Ryan and Letai 2013). This indirect MOMP measurement serves as a readout that describes the proximity of the cells to the apoptotic threshold. Therefore, high MOMP, or mitochondrial depolarization values, indicate that cells are highly primed. Conversely, low values indicate that cells are poorly primed (Ni Chonghaile et al., 2011). Strikingly, BH3 profiling performed on p53-restorable lung

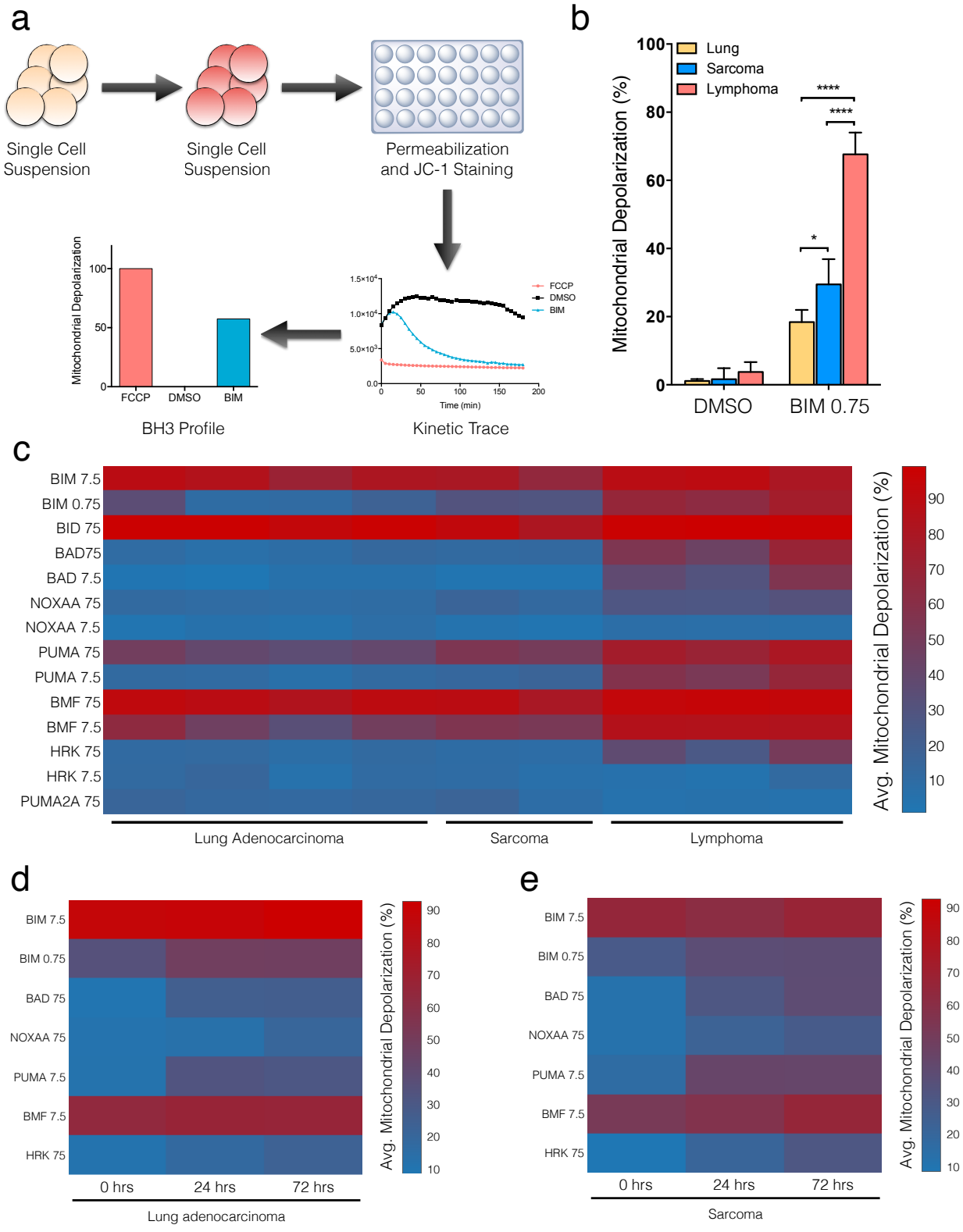

a
Single Cell Suspension
Single Cell Suspension
Permeabilization and JC-1 Staining
Mitochondrial Depolarization
100
50
0
FCCP
DMSO
BIM
BH3 Profile
1.5×10^4
1.0×10^4
5.0×10^3
0.0
0
50
100
150
200
Time (min)
FCCP
DMSO
BIM
Kinetic Trace
b
Lung
Sarcoma
Lymphoma
Mitochondrial Depolarization (%)
100
80
60
40
20
0
****
****
*
DMSO
BIM 0.75
c
BIM 7.5
BIM 0.75
BID 75
BAD75
BAD 7.5
NOXAA 75
NOXAA 7.5
PUMA 75
PUMA 7.5
BMF 75
BMF 7.5
HRK 75
HRK 7.5
PUMA2A 75
Lung Adenocarcinoma
Sarcoma
Lymphoma
Avg. Mitochondrial Depolarization (%)
90
80
70
60
50
40
30
20
10
d
BIM 7.5
BIM 0.75
BAD 75
NOXAA 75
PUMA 7.5
BMF 7.5
HRK 75
0 hrs
24 hrs
72 hrs
Lung adenocarcinoma
Avg. Mitochondrial Depolarization (%)
90
80
70
60
50
40
30
20
10
e
BIM 7.5
BIM 0.75
BAD 75
NOXAA 75
PUMA 7.5
BMF 7.5
HRK 75
0 hrs
24 hrs
72 hrs
Sarcoma
Avg. Mitochondrial Depolarization (%)
90
80
70
60
50
40
30
20
10

Figure 3: The level of mitochondrial apoptotic priming explains tumor-specific responses. (a) Schematic of the BH3 profiling assay (refer to Ni Chonghaile et al., 2011 and the text for more details). (b) Individual BIM-only BH3 profile of lung adenocarcinoma (n=4), sarcoma (n=2), and lymphoma (n=3) cell lines, demonstrating that lung adenocarcinoma and sarcoma cell lines are poorly primed whereas lymphoma cell lines are highly primed. (c) Full BH3 profile of cell lines from (b) represented in heatmap form (note scale on the right). (d) BH3 profile of a lung adenocarcinoma cell line, demonstrating that p53 restoration slightly increases mitochondrial apoptotic priming. (e) BH3 profile of a sarcoma cell line, demonstrating that p53 restoration slightly increases mitochondrial apoptotic priming. Data in (b) represent the mean ± S.E.M, n=3 or more. Statistics were calculated with two-sided Student's t-test: *P < 0.05, **P < 0.01, ***P < 0.001, ****P < 0.0001.

adenocarcinoma, sarcoma, and lymphoma cell lines unveiled that lung adenocarcinoma and sarcoma cell lines are poorly primed, whereas lymphoma cell lines are highly primed (**Figure 3B-C**). As shown in **Figure 3B**, lung adenocarcinoma cell lines tend to exhibit the lowest mitochondrial depolarization values when compared to sarcoma cell lines and lymphoma cell lines. Perhaps not surprisingly, p53 restoration by itself was sufficient to cause a minor increase in the levels of mitochondrial apoptotic priming in both lung adenocarcinoma and sarcoma cell lines (**Figure 3D-E**). This result is consistent with the fact that p53 transcriptionally activates multiple pro-apoptotic genes, including the BH3-only proteins *Puma* and *Noxa* (Hemann and Lowe 2006). Nevertheless, it appears that this minor increase in the levels of mitochondrial apoptotic priming is not enough to induce cell death in this context, as demonstrated by the lack of cell death in both lung adenocarcinoma and sarcoma cell lines upon p53 restoration (**Figure 2B, E**).

Genetic manipulation of mitochondrial apoptotic priming levels changes cell fate

The results described above prompted us to test the hypothesis that manipulating the levels of mitochondrial apoptotic priming would be sufficient to switch the observed tumor-specific responses to p53 restoration. We tackled this hypothesis in a few different ways. First, we genetically ablated *Bcl-xL* in both lung adenocarcinoma and sarcoma cell lines using CRISPR-Cas9 to test whether its deletion would increase the basal level of mitochondrial apoptotic priming in these poorly primed cells (**Figure 4A, 5A**). Remarkably, CRISPR-mediated deletion of *Bcl-xL* in either lung adenocarcinoma (**Figure 4B-C**) or sarcoma (**Figure 5B-C**) cell lines substantially increased the basal

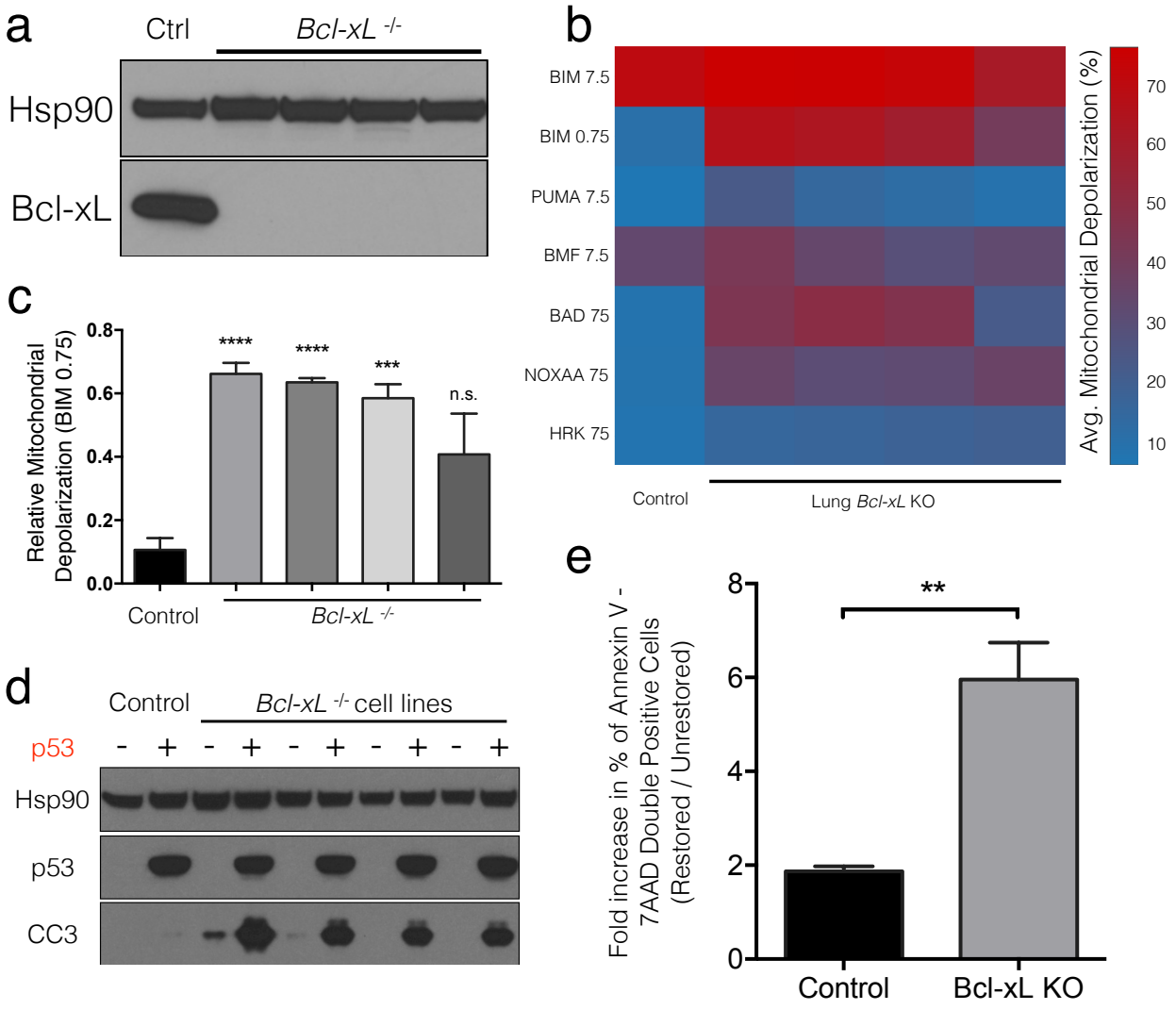


Figure 4: Genetic deletion of *Bcl-xL* increases mitochondrial apoptotic priming in lung adenocarcinoma cell lines and is sufficient to trigger cell death in the context of p53 restoration.

(a) CRISPR-mediated deletion of *Bcl-xL* in lung adenocarcinoma cell lines. (b) BH3 profile of four *Bcl-xL* knockout lung adenocarcinoma cell lines, demonstrating that *Bcl-xL* deletion increases mitochondrial apoptotic priming. (c) Individual BIM-only BH3 profile of four *Bcl-xL* knockout lung adenocarcinoma cell lines. (d) *Bcl-xL* deletion increases mitochondrial apoptotic priming and this is sufficient to sensitize lung adenocarcinoma cells to undergo cell death upon restoration of p53. Data shown is from cells harvested 72hrs after p53 restoration. (e) Percentage of Annexin V-7AAD double positive *Bcl-xL* knockout cells 72hrs after p53 restoration. Data in (c) and (e) represent the mean ± S.E.M, n=3 or more. Statistics were calculated with two-sided Student's t-test: *P < 0.05, **P < 0.01, ***P < 0.001, ****P < 0.0001.

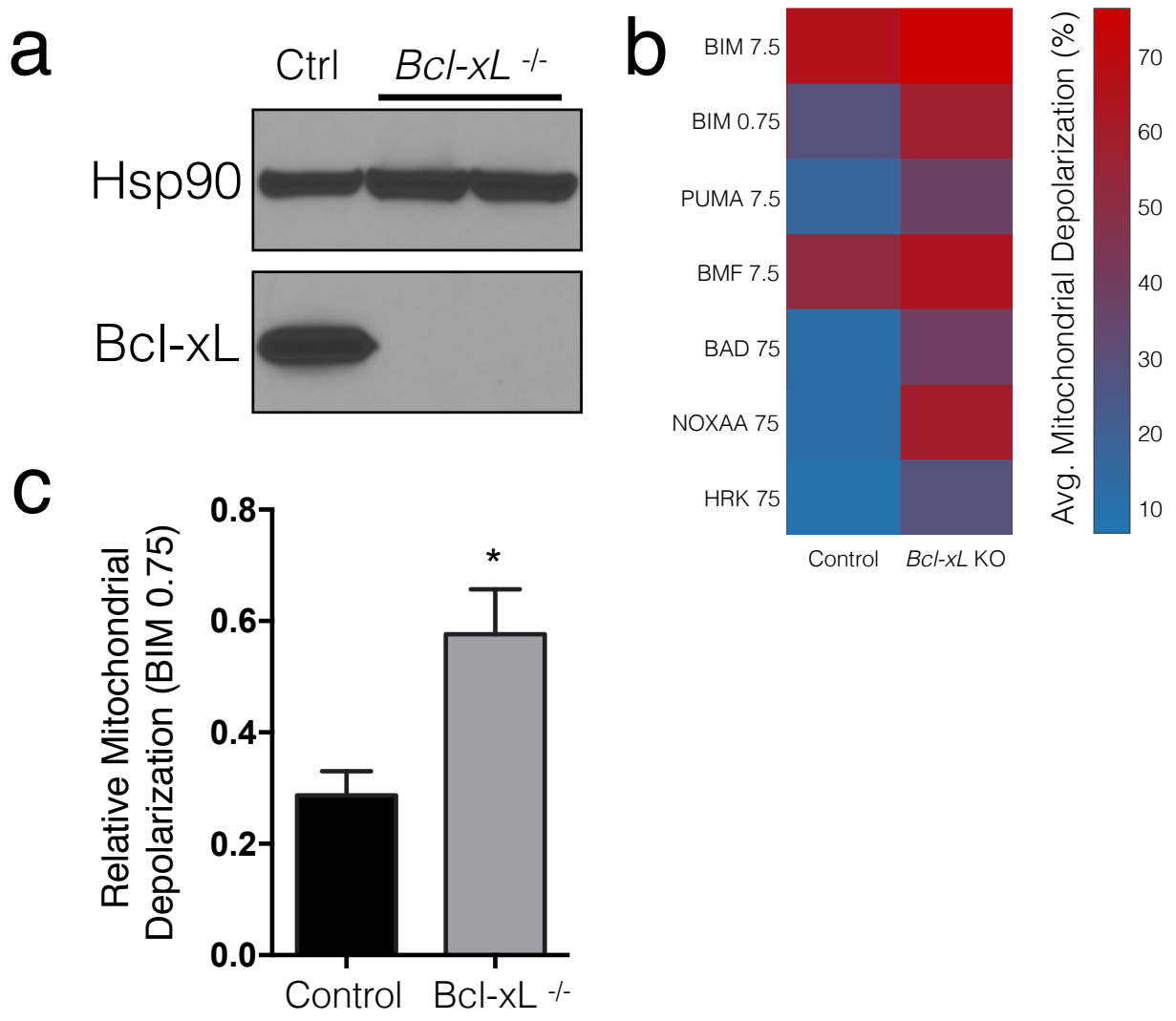


Figure 5: Genetic deletion of *Bcl-xL* increases mitochondrial apoptotic priming in sarcoma cell lines. (a) CRISPR-mediated deletion of *Bcl-xL* in sarcoma cell lines. (b) BH3 profile of one *Bcl-xL* knockout sarcoma cell line, demonstrating that *Bcl-xL* deletion increases mitochondrial apoptotic priming. (c) Individual BIM-only BH3 profile of one *Bcl-xL* knockout sarcoma cell line, demonstrating increased mitochondrial apoptotic priming. Data in (c) represent the mean ± S.E.M, n=3 or more. Statistics were calculated with two-sided Student's t-test: *P < 0.05

levels of mitochondrial apoptotic priming. This increase in mitochondrial apoptotic priming was sufficient to switch the fate of lung adenocarcinoma cells upon p53 restoration. Notably, restoration of p53 in lung adenocarcinoma *Bcl-xL* knockout cells triggered a substantial increase in the levels of CC3 (**Figure 4D**) and in the percentage of Annexin V-7AAD double positive cells (**Figure 4E**). To test whether decreasing the level of mitochondrial apoptotic priming in lymphoma cells would switch their tumor-specific response from apoptosis to cell cycle arrest, we overexpressed *Bcl-xL* using a retrovirus (**Figure 6A**). Remarkably, *Bcl-xL* overexpression led to a significant decrease in the level of mitochondrial apoptotic priming (**Figure 6B-C**). This decrease in mitochondrial apoptotic priming was sufficient to switch the fate of lymphoma cells upon p53 restoration. Notably, *Bcl-xL*-overexpression was able to block p53-mediated cell death, as demonstrated by the absence of CC3 (**Figure 6A**) and a significant decrease in the percentage of Annexin V-7AAD double positive cells (**Figure 6D**). This *Bcl-xL*-overexpression-mediated decrease in mitochondrial apoptotic priming triggered a change in cell fate - from cell death to cell cycle arrest - in the context of p53 restoration, as demonstrated by the significant accumulation of lymphoma cells in the G0/G1 stage of the cell cycle (**Figure 6E**).

To assess whether increasing mitochondrial apoptotic priming would be sufficient to trigger p53-dependent apoptosis in established tumors *in vivo*, we overexpressed the BH3-only sensitizer protein *Bad* in sarcoma cells using a retrovirus (**Figure 7A**). Importantly, *Bad* overexpression by itself was insufficient to cause cell death, despite the fact that its overexpression led to a measurable increase in the levels of mitochondrial apoptotic priming (**Figure 7B**). Remarkably, restoration of p53 in *Bad*-

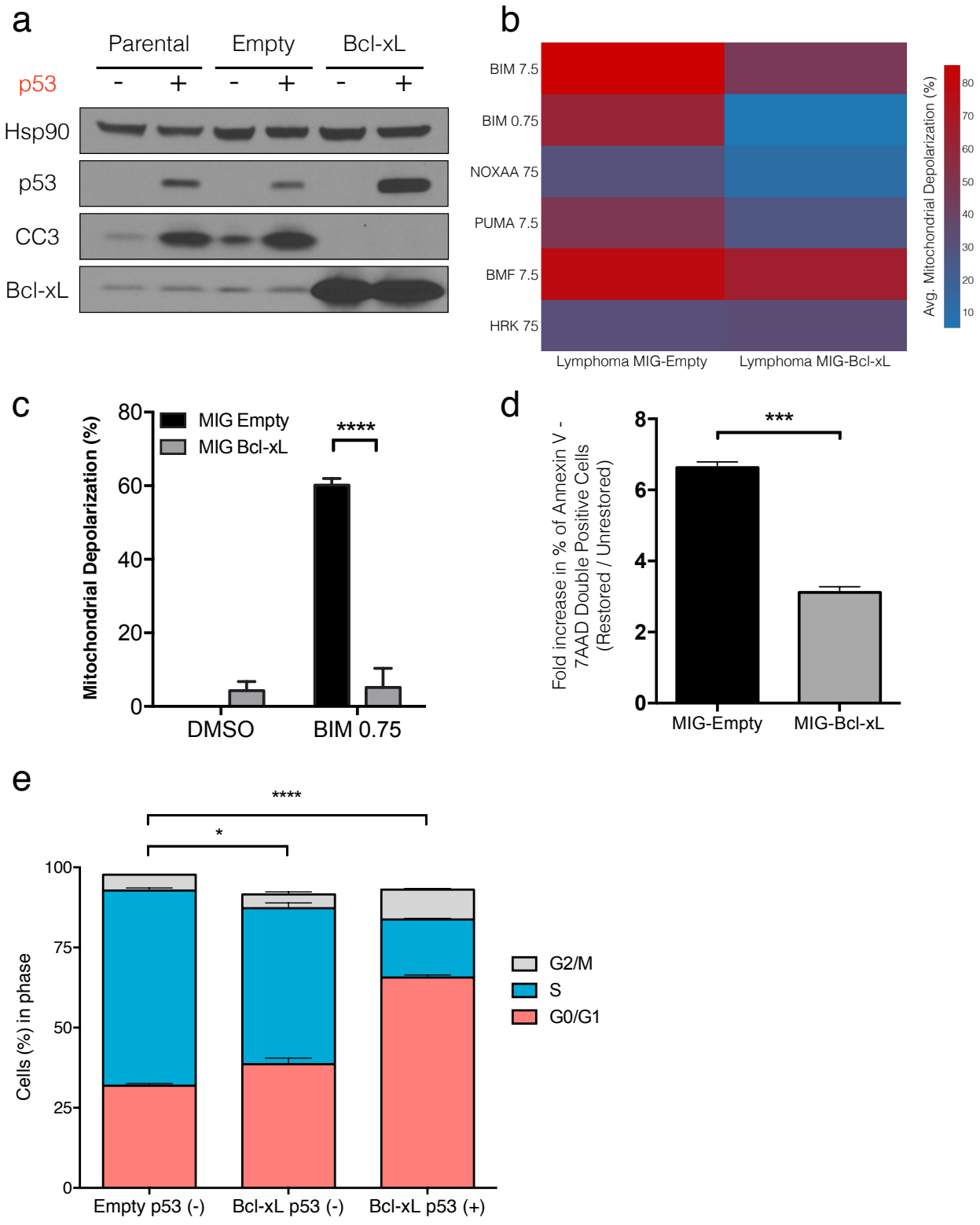

a
Parental
Empty
Bcl-xL
p53
- + - + - +
Hsp90
p53
CC3
Bcl-xL
b
BIM 7.5
BIM 0.75
NOXAA 75
PUMA 7.5
BMF 7.5
HRK 75
Lymphoma MIG-Empty
Lymphoma MIG-Bcl-xL
Avg. Mitochondrial Depolarization (%)
80
70
60
50
40
30
20
10
c
Mitochondrial Depolarization (%)
MIG Empty
MIG Bcl-xL
****
80
60
40
20
0
DMSO
BIM 0.75
d
Fold increase in % of Annexin V - 7AAD Double Positive Cells (Restored / Unrestored)
***
8
6
4
2
0
MIG-Empty
MIG-Bcl-xL
e
****
*
Cells (%) in phase
100
75
50
25
0
G2/M
S
G0/G1
Empty p53 (-)
Bcl-xL p53 (-)
Bcl-xL p53 (+)

Figure 6: *Bcl-xL* overexpression decreases mitochondrial apoptotic priming in lymphoma cell lines and is sufficient to trigger cell cycle arrest in the context of p53 restoration. (a) *Bcl-xL* overexpression blocks p53-mediated cell death. (b) BH3 profile of either control or *Bcl-xL* overexpressing lymphoma cells, demonstrating that *Bcl-xL* overexpression decreases mitochondrial apoptotic priming. (c) Individual BIM-only BH3 profile of *Bcl-xL* overexpressing lymphoma cells. (d) Percentage of Annexin V-7AAD double positive *Bcl-xL* overexpressing lymphoma cells 72hrs after p53 restoration. (e) Cell cycle profiling of *Bcl-xL* overexpressing lymphoma cells 72hrs after p53 restoration. Data in (c-e) represent the mean ± S.E.M, n=3 or more. Statistics were calculated with two-sided Student's t-test: *P < 0.05, **P < 0.01, ***P < 0.001, ****P < 0.0001.

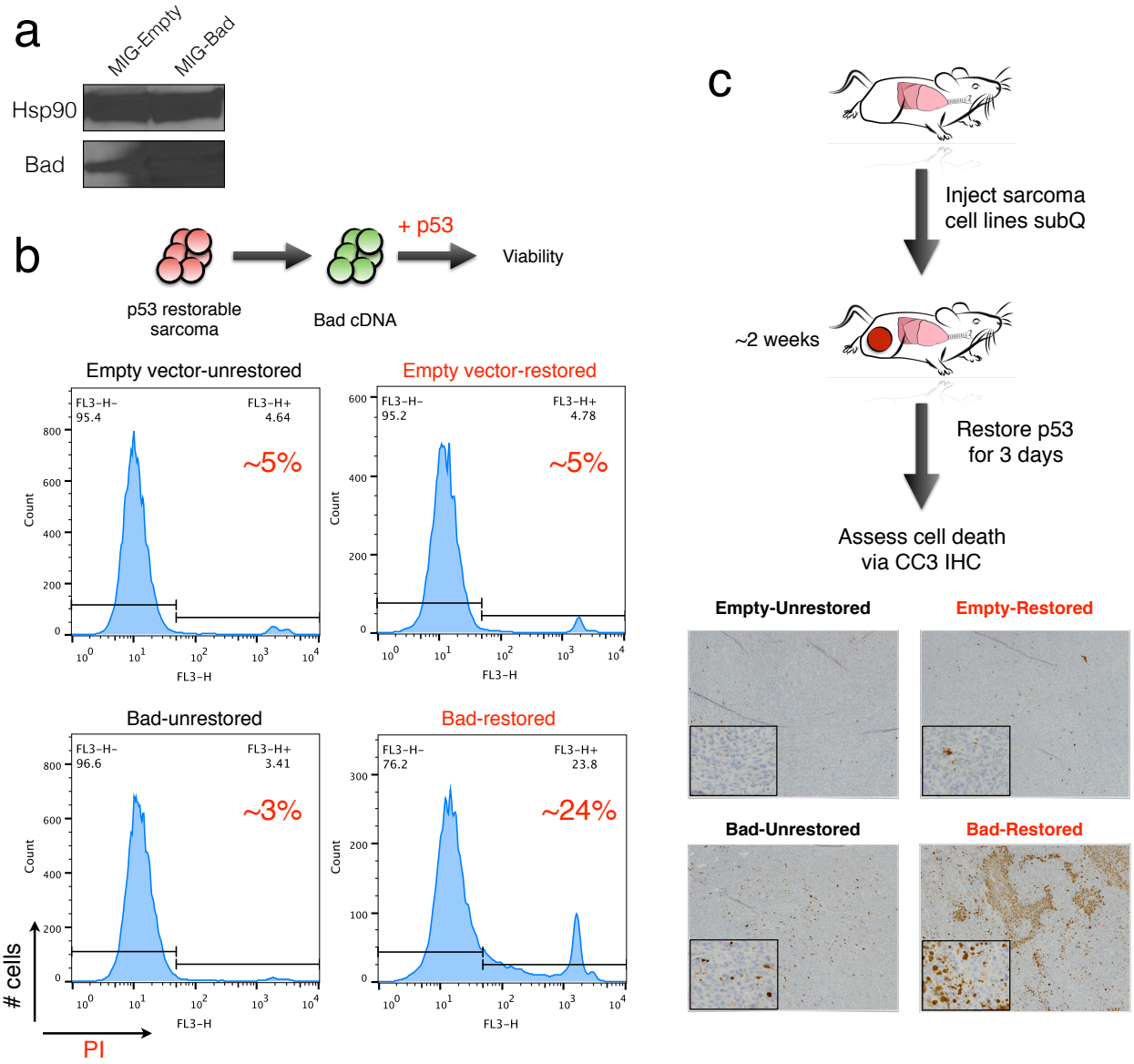


Figure 7: *Bad* overexpression in sarcoma cell lines increases mitochondrial apoptotic priming and is sufficient to trigger cell death *in vitro* and *in vivo* in the context of p53 restoration. (a) Western blot showing *Bad* overexpression. (b) Viability analysis using propidium iodide incorporation. Note that *Bad* overexpressing cells are as viable as empty-vector cells before p53 restoration, but that their viability substantially decreases upon p53 restoration, as demonstrated by an increase in the percentage of cells incorporating propidium iodide. (c) *Bad* overexpression sensitize sarcoma tumors to p53-mediated cell death *in vivo*. CC3=cleaved caspase 3.

overexpressing cells resulted in a substantial increase in the level of cell death *in vitro* (**Figure 7B**), as well as in established sarcoma tumors *in vivo* (**Figure 7C**).

Pharmacological manipulation of mitochondrial apoptotic priming levels changes cell fate

The data above convincingly demonstrated that genetic inactivation of *Bcl-xL* or overexpression of *Bad* leads to a substantial increase in mitochondrial apoptotic priming, which consequently synergizes with p53 restoration to induce cell death in lung adenocarcinoma and sarcoma cell lines. To determine whether pharmacological manipulation of mitochondrial apoptotic priming could trigger a similar synergistic effect in the context of p53 restoration, we treated lung adenocarcinoma and sarcoma cells with the BH3 mimetic ABT-263, which is a potent inhibitor of Bcl-xL (as well as Bcl-2 and Bcl-w) (Tse et al., 2008, Lessene et al., 2008). Remarkably, treatment with ABT-263 strongly synergized with p53 restoration and induced a substantial amount of cell death in both lung adenocarcinoma and sarcoma cell lines, as demonstrated by a significant increase in the percentage of Annexin V-7AAD double positive cells (**Figure 8**). These results strongly suggest that pharmacological induction of mitochondrial priming via the use of BH3 mimetic compounds could strongly synergize with drugs that restore p53, such as PRIMA-1 (Lambert et al., 2009), Nutlin-3 (Vassilev et al., 2004), and others (Khoo et al., 2014).

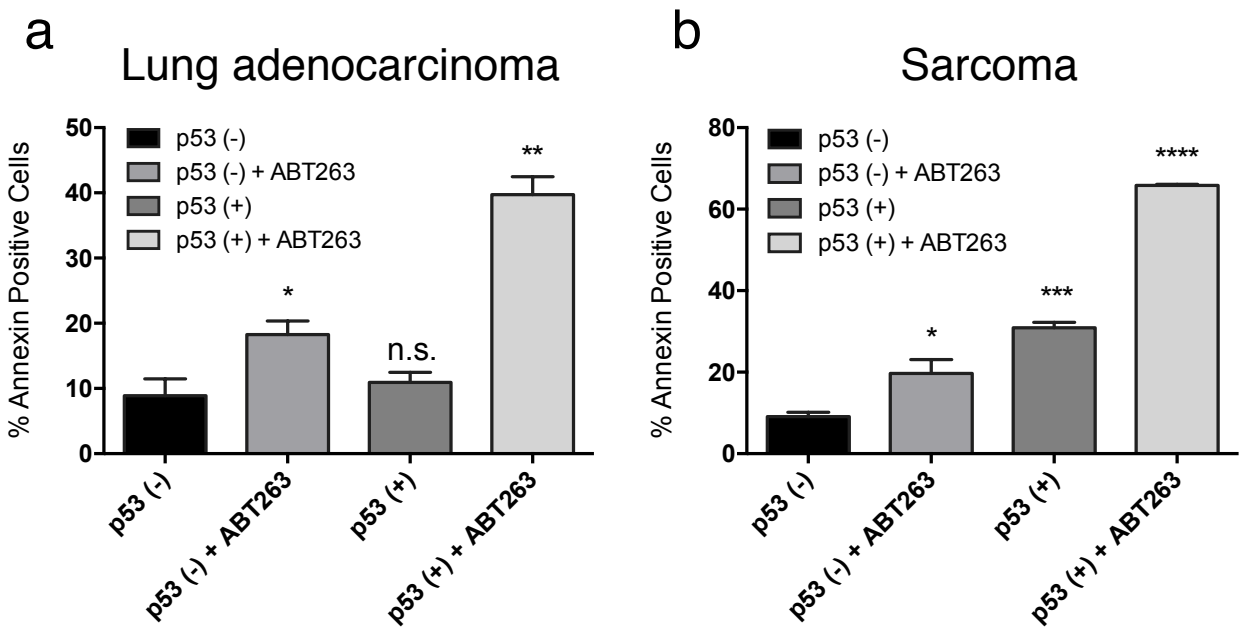


Figure 8: Pharmacological manipulation of mitochondrial apoptotic priming levels using BH3 mimetic compounds sensitizes cells to undergo cell death upon p53 restoration. These graphs represent the mean ± S.E.M, n=3 or more technical replicates. Statistics were calculated with two-sided Student's t-test: *P < 0.05, **P < 0.01, ***P < 0.001, ****P < 0.0001.

DISCUSSION

Using a set of defined GEMM-derived p53 restorable lung adenocarcinoma, sarcoma, and lymphoma cell lines in combination with BH3 profiling, we have demonstrated that the level of mitochondrial apoptotic priming is a critical determinant of cell fate in the context of p53 restoration. We discovered that lung adenocarcinoma and sarcoma tumors - which normally undergo cell cycle arrest upon p53 restoration - are poorly primed. Consequently, reactivation of the p53 tumor suppressor gene in this context is unable to trigger cell death, presumably because the cells are below a certain apoptotic threshold that needs to be overcome in order to elicit MOMP and subsequent cell death. Interestingly, we demonstrated that p53 by itself is able to increase the level of mitochondrial apoptotic priming, presumably via the transcriptional activation of pro-apoptotic BH3-family members, including *Noxa* and *Puma*. However, p53 by itself is unable to trigger cell death in these two tumor types; therefore, the p53-dependent increase in mitochondrial apoptotic priming does not appear to be sufficient to bring the cells over the apoptotic threshold. On the other hand, we discovered that lymphoma cell lines - which normally undergo cell death upon p53 restoration - are highly primed. Consequently, reactivation of the p53 tumor suppressor gene in this context readily triggers cell death, presumably because the cells are very close to the apoptotic threshold. In this context, p53 by itself is able to overcome the apoptotic threshold; perhaps in this setting, the slight aforementioned p53-dependent increase in mitochondrial priming is enough to overcome the apoptotic threshold and elicit MOMP and subsequent cell death.

By employing a combination of CRISPR-based loss of function approaches and cDNA-based gain of function approaches, we demonstrated that genetic manipulation of mitochondrial apoptotic priming levels is sufficient to change cell fate upon p53 restoration. We demonstrated that CRISPR-mediated deletion of *Bcl-xL* triggers a substantial increase in the levels of mitochondrial apoptotic priming, which are measurable via BH3 profiling. The fact that we could measure this dynamic property in a highly sensitive way allowed us to functionally explain how a single genetic manipulation could cooperate with p53 reactivation to induce cell death in otherwise recalcitrant tumor type. Indeed, *Bcl-xL* inactivation by itself did not trigger any detectable cell death; nevertheless, we were able to conclude that it was having a major functional effect in the levels of mitochondrial apoptotic priming through the use of BH3 profiling. This increase in mitochondrial apoptotic priming synergized with p53 reactivation to induce cell death. Similarly, increasing the levels of mitochondrial apoptotic priming via overexpression of the BH3-only sensitizer *Bad* also synergized with p53 reactivation to induce cell death. On the other hand, *Bcl-xL* overexpression triggered a substantial decrease in the levels of mitochondrial apoptotic priming, as measured by BH3 profiling. This increase in mitochondrial apoptotic priming completely blocked p53-mediated cell death and instead allowed for cell cycle arrest. Importantly, *Bcl-xL* overexpression by itself did not trigger any significant cell cycle arrest; nevertheless, we were able to conclude that it was having a major functional effect in the levels of mitochondrial apoptotic priming through the use of BH3 profiling.

Our results support a model whereby lung and sarcoma cell lines are far from the apoptotic threshold due to the fact that they have inherently low levels of mitochondrial apoptotic priming, whereas lymphoma cell lines are close to the apoptotic threshold due to the fact that they have inherently high levels of mitochondrial apoptotic priming (**Figure 9**).

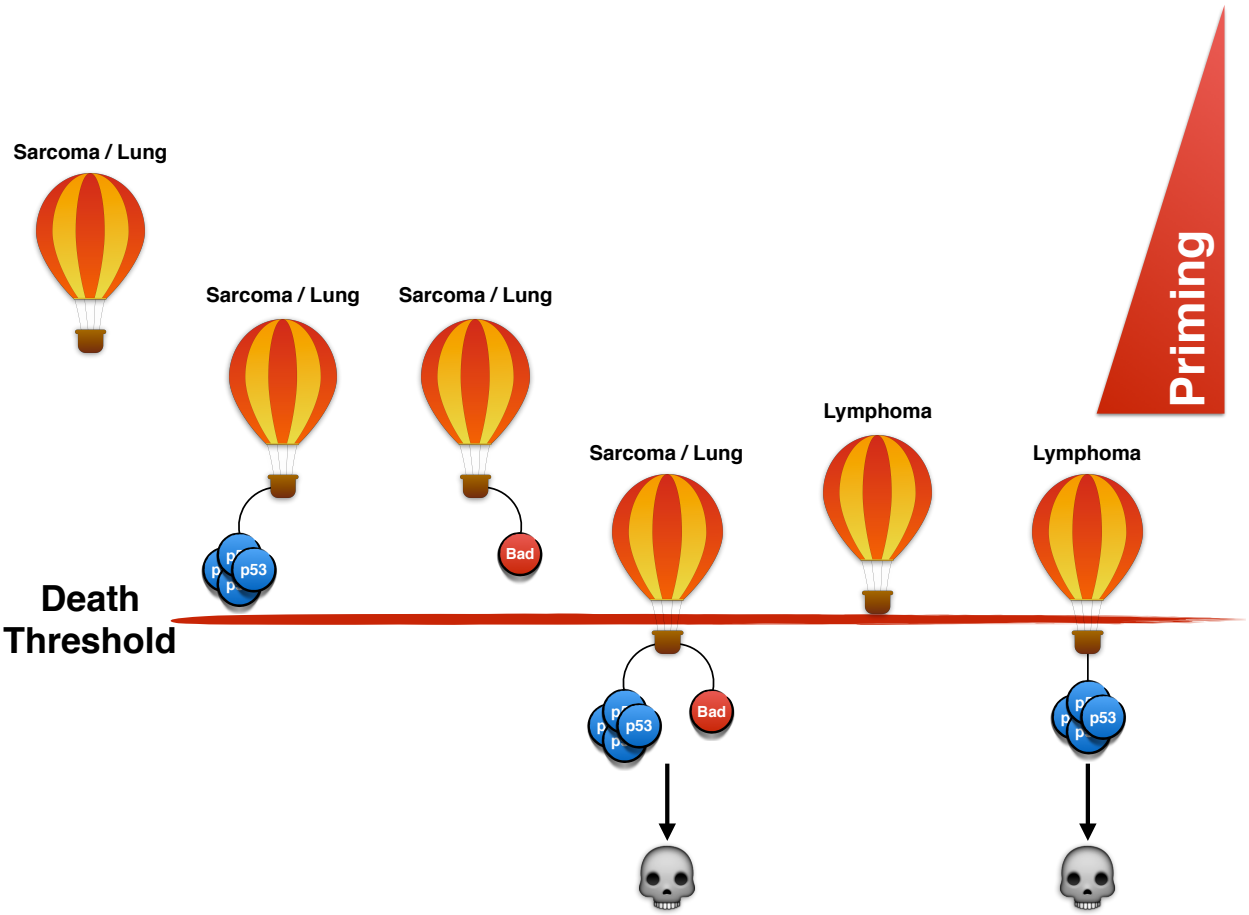


Figure 9: The apoptotic threshold model. Lung adenocarcinoma and sarcoma cell lines are far from the apoptotic threshold due to the fact that they have low levels of mitochondrial apoptotic priming. Despite the fact that p53 brings them closer to the apoptotic threshold, their inherently low level of basal mitochondrial apoptotic priming does not allow them to cross through the apoptotic threshold. On the other hand, lymphoma cell lines are highly primed, and p53 restoration readily triggers their passage through the apoptotic threshold. Manipulating the levels of mitochondrial apoptotic priming through genetic deletion of *Bcl-xL* or *Bad* overexpression (shown here), as well as through the use of BH3 mimetic drugs (such as ABT-263), can dramatically increase the levels of mitochondrial priming and strongly synergize with p53 to induce cell death in poorly primed tumor types, such as lung adenocarcinoma and sarcoma. Adapted from Lowe et al., 2004.

Remarkably, we were also able to mimic our genetic results using the ABT-263 BH3 mimetic compound, suggesting that the combination of a p53 restoring drug with an agent that increases mitochondrial apoptotic priming could strongly synergize and trigger cell death in tumors that commonly undergo cell cycle arrest upon therapeutic p53 reactivation.

MATERIALS AND METHODS

Generation of cell lines and cell culture

Bcl-xL knockout cells were generated by transient transfection using the pX458 vector (Addgene, 48138) essentially as described in the protocol by Ran et al., 2013 using sgBcl-xL.2 (5' - **G** CCCAGCTTCACATAACCCCA - 3'), where the **G** in bold was added for ensuring appropriate U6 transcription. *Bcl-xL* overexpressing cell lines were generated using the pMIG Bcl-XL retroviral construct (Addgene, 3541). The pMIG-Empty retroviral construct (Addgene, 9044) was used as control. Retroviral particles were generated and packaged using Phoenix cell system (G. Nolan, Stanford University). Cells were maintained in DMEM (lung adenocarcinoma and sarcoma cell lines) or IMDM (lymphoma cell lines) supplemented with 10% Fetal Bovine Serum and gentamicin. For p53 restoration experiments, 4-hydroxy-tamoxifen was added at a final concentration of 250-500nM for a total of 72hrs.

Antibodies

The following antibodies were used for western blotting experiments: anti-Bcl-xL (CST, #2764, 1:1000), anti-p53 (Novocastra, NCL-p53-505 1:1000), anti-Hsp90 (BD, 610418,

1:10000 to 1:15000), anti-cleaved caspase 3 (CST, #9661, 1:1000), anti-Bad (CST, #9292, 1:1000), and anti-p21 (Santa Cruz, sc-6246).

Western blotting

Cells were lysed with RIPA buffer (BP-115, Boston BioProducts) supplemented with 1X protease inhibitor solution (cOmplete EDTA-free, 11873580001, Roche). Protein concentration of cell lysates was determined by Pierce BCA protein assay (23225, Thermo Fisher Scientific). Total protein (50µg) was separated on 4-12% Bis-Tris gradient SDS-PAGE gels (Life Technologies) and then transferred to PVDF membranes (IPVH00010, EMD Millipore) for blotting.

Cell cycle analysis

For cell cycle analysis, we used a FITC BrdU Flow Kit (559619, BD Biosciences). Briefly, 1 million cells were plated in triplicate with or without p53 restoration for a total of 72 hrs. Cells were then labeled with 10µM BrdU for 30-60 minutes, and subsequently fixed and stained with anti-BrdU and 7-AAD and analyzed by flow cytometry.

Annexin V - 7AAD analysis

Cells were cultured identically as described for cell cycle experiments. 72 hrs after p53 restoration, cells were analyzed using the BioLegend assay kit (640919) following manufacturer guidelines.

BH3 profiling

BH3 profiling was performed essentially as described in Ryan and Letai 2013.

Animal experiments

All animal studies described in this study were approved by the MIT Institutional Animal Care and Use Committee. A total of 500,000 sarcoma cells were injected subcutaneously into Nude mice from Taconic at the age of 6-8 weeks. Tamoxifen (Sigma) was dissolved in corn oil (Sigma) and a single injection was administered intraperitoneally at 200 μg per gram of total body weight.

Immunohistochemistry

Immunohistochemistry was performed identically as described in **Chapter 2** of this thesis, using anti-cleaved caspase 3 (CST, #9661, 1:1000).

ACKNOWLEDGEMENTS

We thank members of the Jacks, Letai, and Hemann laboratories for multiple scientific discussions, as well as K. Cormier and C. Condon from the Hope Babette Tang (1983) Histology Facility for technical support, and G. Paradis, M. Griffin, M. Jennings and M. Saturno-Condón for flow cytometry and FACS support. This work was supported by the Howard Hughes Medical Institute and in part by Cancer Center Support (core) grant P30-CA14051 from the National Cancer Institute. Y.M.S.F. is supported by the National Cancer Institute of the National Institutes of Health under Award Number F31-CA183405. M.T.H is the Eisen and Chang Associate Professor of Biology. T.J. is a Howard Hughes Medical Institute Investigator, the David H. Koch Professor of Biology, and a Daniel K. Ludwig Scholar.

REFERENCES

Donehower, L.A., Harvey, M., Slagle, B.L., McArthur, M.J., Montgomery, C.A., Butel, J.S., and Bradley, A. (1992). Mice deficient for p53 are developmentally normal but susceptible to spontaneous tumours. Nature 356, 215–221.

Feldser, D.M., Kostova, K.K., Winslow, M.M., Taylor, S.E., Cashman, C., Whittaker, C.A., Sanchez-Rivera, F.J., Resnick, R., Bronson, R., Hemann, M.T., et al. (2010). Stage-specific sensitivity to p53 restoration during lung cancer progression. Nature 468, 572–575.

Hemann, M.T., and Lowe, S.W. (2006). The p53-Bcl-2 connection. Cell Death Differ 13, 1256–1259.

Jacks, T., Remington, L., Williams, B.O., Schmitt, E.M., Halachmi, S., Bronson, R.T., and Weinberg, R.A. (1994). Tumor spectrum analysis in p53-mutant mice. Curr Biol 4, 1–7.

Jackson, J.G., Post, S.M., and Lozano, G. (2010). Regulation of tissue- and stimulus-specific cell fate decisions by p53 in vivo. The Journal of Pathology.

Johnson, L., Mercer, K., Greenbaum, D., Bronson, R.T., Crowley, D., Tuveson, D.A., and Jacks, T. (2001). Somatic activation of the K-ras oncogene causes early onset lung cancer in mice. Nature 410, 1111–1116.

Junttila, M.R., Karnezis, A.N., Garcia, D., Madriles, F., Kortlever, R.M., Rostker, F., Swigart, L.B., Pham, D.M., Seo, Y., Evan, G.I., et al. (2010). Selective activation of p53-mediated tumour suppression in high-grade tumours. Nature 468, 567–571.

Khoo, K.H., Hoe, K.K., Verma, C.S., and Lane, D.P. (2014). Drugging the p53 pathway: understanding the route to clinical efficacy. Nature Reviews Drug Discovery 13, 217–236.

Lambert, J.M.R., Gorzov, P., Veprintsev, D.B., Soderqvist, M., Segerback, D., Bergman, J., Fersht, A.R., Hainaut, P., Wiman, K.G., and Bykov, V.J.N. (2009). PRIMA-1 reactivates mutant p53 by covalent binding to the core domain. Cancer Cell 15, 376–388.

Lane, D., and Levine, A. (2010). P53 Research: The Past Thirty Years and the Next Thirty Years. Cold Spring Harbor Perspectives in Biology.

Lessene, G., Czabotar, P.E., and Colman, P.M. (2008). BCL-2 family antagonists for cancer therapy. Nat Rev Drug Discov 7, 989–1000.

Letai, A.G. (2008). Diagnosing and exploiting cancer's addiction to blocks in apoptosis. Nat Rev Cancer 8, 121–132.

Lowe, S.W., Cepero, E., and Evan, G. (2004). Intrinsic tumour suppression. Nature 432, 307–315.

Lu, X. (2010). Tied up in loops: positive and negative autoregulation of p53. Cold Spring Harbor Perspectives in Biology 2, a000984.

Martins, C.P., Brown-Swigart, L., and Evan, G.I. (2006). Modeling the therapeutic efficacy of p53 restoration in tumors. Cell 127, 1323–1334.

Momand, J., Jung, D., Wilczynski, S., and Niland, J. (1998). The MDM2 gene amplification database. Nucleic Acids Res 26, 3453–3459.

Ni Chonghaile, T., Sarosiek, K.A., Vo, T.-T., Ryan, J.A., Tammareddi, A., Moore, V.D.G., Deng, J., Anderson, K.C., Richardson, P., Tai, Y.-T., et al. (2011). Pretreatment mitochondrial priming correlates with clinical response to cytotoxic chemotherapy. Science 334, 1129–1133.

Polager, S., and Ginsberg, D. (2009). p53 and E2f: partners in life and death. Nat Rev Cancer 9, 738–748.

Purvis, J.E., Karhohs, K.W., Mock, C., Batchelor, E., Loewer, A., and Lahav, G. (2012). p53 dynamics control cell fate. Science 336, 1440–1444.

Ran, F.A., Hsu, P.D., Wright, J., Agarwala, V., Scott, D.A., and Zhang, F. (2013). Genome engineering using the CRISPR-Cas9 system. Nat Protoc 8, 2281–2308.

Ryan, J., and Letai, A. (2013). BH3 profiling in whole cells by fluorimeter or FACS. Methods 61, 156–164.

Tse, C., Shoemaker, A.R., Adickes, J., Anderson, M.G., Chen, J., Jin, S., Johnson, E.F., Marsh, K.C., Mitten, M.J., Nimmer, P., et al. (2008). ABT-263: a potent and orally bioavailable Bcl-2 family inhibitor. Cancer Research 68, 3421–3428.

Vassilev, L.T., Vu, B.T., Graves, B., Carvajal, D., Podlaski, F., Filipovic, Z., Kong, N., Kammlott, U., Lukacs, C., Klein, C., et al. (2004). In vivo activation of the p53 pathway by small-molecule antagonists of MDM2. Science 303, 844–848.

Ventura, A., Kirsch, D., McLaughlin, M., Tuveson, D., Grimm, J., Lintault, L., Newman, J., Reczek, E., Weissleder, R., and Jacks, T. (2007). Restoration of p53 function leads to tumour regression in vivo. Nature 445, 661–665.

Xue, W., Zender, L., Miething, C., Dickins, R.A., Hernando, E., Krizhanovsky, V., Cordon-Cardo, C., and Lowe, S.W. (2007). Senescence and tumour clearance is triggered by p53 restoration in murine liver carcinomas. Nature 445, 656–660.

Zambrowicz, B.P., Imamoto, A., Fiering, S., Herzenberg, L.A., Kerr, W.G., and Soriano, P. (1997). Disruption of overlapping transcripts in the ROSA beta geo 26 gene trap strain leads to widespread expression of beta-galactosidase in mouse embryos and hematopoietic cells. Proc Natl Acad Sci USA 94, 3789–3794.

Appendix 2

Brg1 suppresses several aspects of lung cancer and its inactivation creates unique genetic dependencies

Francisco J. Sánchez-Rivera[1,2*], Christine M. Fillmore[4,5,6*], Mary Clare Beytagh[1,2], Tuomas Tammela[1,2], Isaac S. Harris[7], Jose Maria Amich[7], Arjun Bhutkar[1], Matthew R. Bauer[1], Joan S. Brugge[7], Carla F. Kim[6,7], and Tyler Jacks[1,2,3].

[1] David H. Koch Institute for Integrative Cancer Research, Massachusetts Institute of Technology, Cambridge, MA 02142

[2] Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02142

[3] Howard Hughes Medical Institute, Massachusetts Institute of Technology, Cambridge, MA 02139

[4] Stem Cell Program, Boston Children's Hospital, Boston, Massachusetts 02115, USA.

[5] Department of Genetics, Harvard Medical School, Boston, Massachusetts 02115, USA.

[6] Harvard Stem Cell Institute, Cambridge, Massachusetts 02138, USA.

[7] Department of Cell Biology and the Ludwig Center at Harvard, Harvard Medical School, Boston, Massachusetts 02115, USA.

[*] These authors contributed equally to this work.

Author Contributions

F.J.S.R., C.M.F., C.F.K, and T.J. designed the study; C.M.F. performed drug studies and other experiments not shown here; F.J.S.R. and M.C.B. generated *Brg1* knockout cell lines and performed *in vivo* studies; I.S.H., J.M.A., and J.S.B. performed small molecule screens, as well as some bioinformatic analyses (data not shown here); T.T. generated preliminary Wnt pathway data (not shown); A.B. conducted bioinformatic analyses.

INTRODUCTION

A unifying theme throughout this thesis has been the motivation to fill the gap between cancer genomics and functional cancer genetics *in vivo*. This motivation partially stems from a desire to achieve a better mechanistic understanding of the biological role that putative cancer drivers (including low frequency drivers) are playing in several aspects of this disease. Large-scale cancer genome sequencing studies have identified loss of function mutations in the *SMARCA4* gene (hereafter referred to as *BRG1/Brg1*), which encodes for the BRG1 chromatin remodeling enzyme, in up to 20% of non-small cell lung cancer patients (Imielinski et al., 2012, LUAD TCGA 2014, and others). In addition, ~ 5% of *KRAS* mutant lung adenocarcinoma patients concomitantly harbor *SMARCA4* mutations (LUAD TCGA 2014), supporting efforts aimed towards the better understanding of the basic biology behind this particular cancer genotype, as well as the potential identification of genetic dependencies that could be pursued for targeting this population of patients.

Despite the fact that there is a significant amount of functional and mutational evidence suggesting that *BRG1* is a tumor suppressor in lung cancer (Medina et al., 2008, Glaros et al., 2008, Orvis et al., 2014), the exact role it is playing and the stage of cancer at which it is acting remains largely unknown. Moreover, *in vivo* studies that have examined the role of *Brg1* in the initiation and progression of other tumor types (such as pancreatic adenocarcinoma) have reported complex context-specific roles (von Figura et al., 2014, Roy et al., 2015). Indeed, studies by the group of Matthias Hebrok have suggested that *Brg1* can have both oncogenic and tumor suppressive properties

depending on the stage of tumor progression, at least in the context of cancers of the pancreas (Roy et al., 2015). Whether *Brg1* acts as a tumor suppressor in lung adenocarcinoma *in vivo,* and whether it has any context- or stage-specific roles remains largely unknown. Moreover, whether *Brg1* mutations result in unique genetic dependencies that could be exploited for therapeutically targeting this specific patient population also remains widely unknown.

Intriguingly, recent data from Christine Fillmore and Carla Kim has demonstrated that *Brg1* mutations can create unique dependencies that can be exploited for therapeutically targeting this patient population (Fillmore et al., 2015). By examining a large panel of human non-small cell lung cancer cell lines, they demonstrated that *Brg1* mutation conferred exquisite sensitivity to combined EZH2 inhibition and etoposide. Whether this combination therapy would hold for *KRAS*-mutant lung cancer patients harboring *BRG1* mutations remains to be seen. Moreover, whether *BRG1* mutations create additional genetic dependencies that could be uncovered using focused or unbiased approaches also remains to be seen.

To begin to tackle this problem, we decided to construct *Kras*-driven GEMMs of lung adenocarcinoma with concomitant mutation in *Brg1* using the pSECC system described in **Chapter 2**. In addition to *in vivo* models, we also decided to generate isogenic *in vitro* models using the strategies described in **Chapter 3**. In this Appendix, I will briefly describe a few of these efforts and discuss future directions.

RESULTS AND DISCUSSION

pSECC-mediated disruption of *Brg1 in vivo*

To dissect the role that *Brg1* might be playing in lung adenocarcinoma initiation, progression, and/or therapeutic response, we took the two-tiered approach presented below in **Figure 1**:

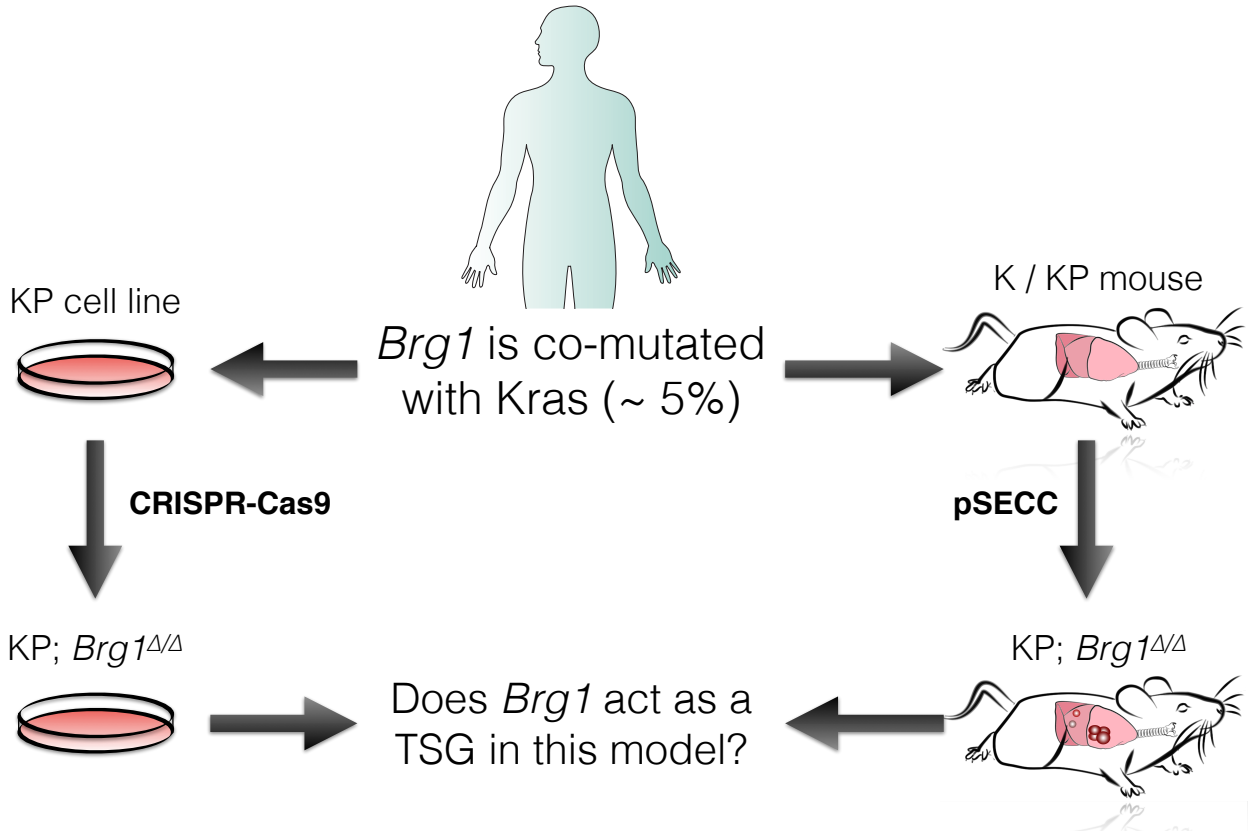


Figure 1: Two-tiered approach to functionally validate a putative tumor suppressor gene (TSG).

To generate *Kras*-driven GEMMs of lung adenocarcinoma, we employed the pSECC system described in **Chapter 2** (see below for experimental scheme; **Figure 2**). We intratracheally infected a cohort composed of 30 KP mice, of which 15 mice were infected with pSECC-Control and the remaining 15 were infected with a pre-validated

pSECC-sgBrg1. We aged the mice for a total of 20 weeks post-infection. At morbidity, we generated multiple cell lines for several purposes, including performing an initial assessment of genome editing.

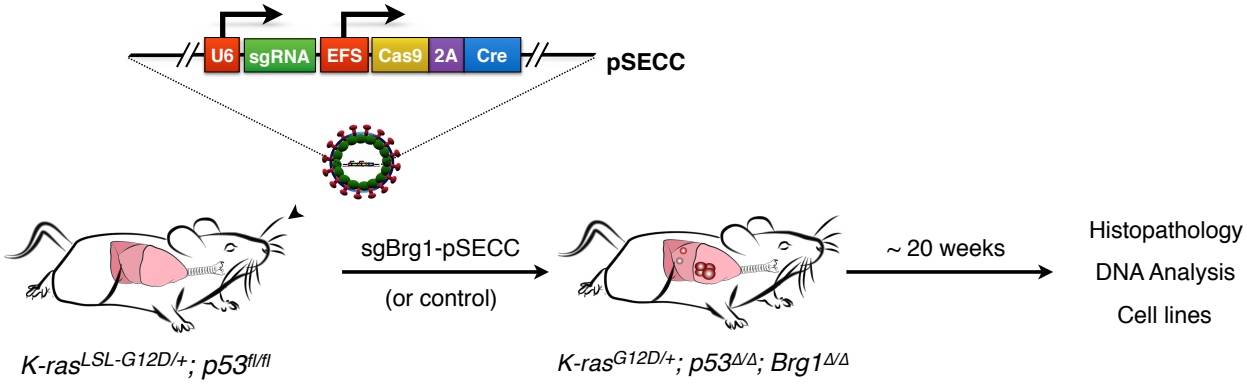


Figure 2: Experimental scheme for pSECC-mediated disruption of *Brg1 in vivo*.

To assess genome editing without the need for high-throughput sequencing, we sequenced the *Brg1* target loci in multiple tumor-derived cell lines. As expected, multiple cell lines harbored bi-allelic indel mutations in predicted areas of Cas9 cutting (~ 3bp upstream or so of the PAM site) (see below **Figure 3**).

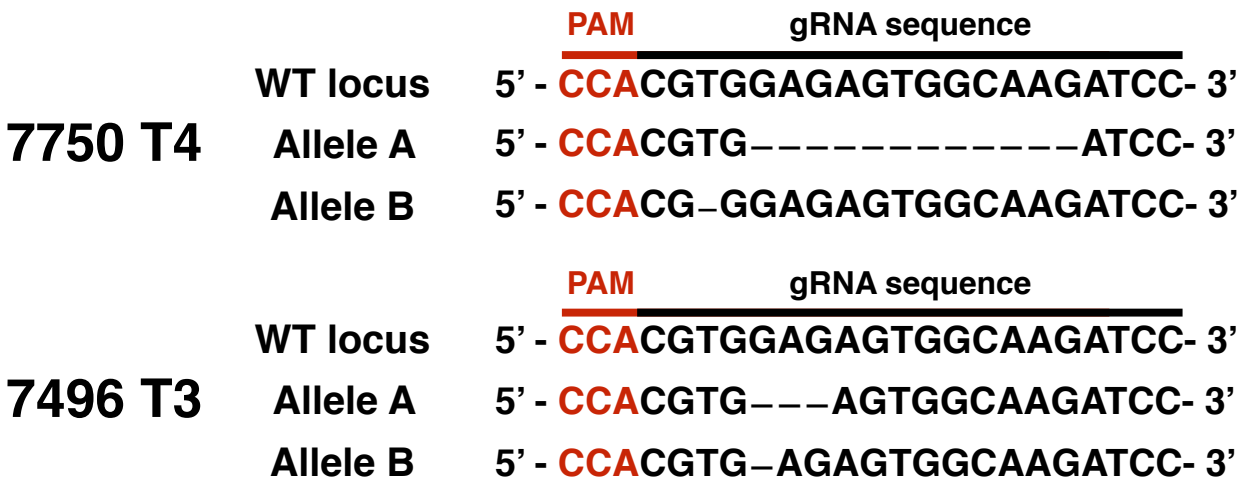


Figure 3: Evidence for bi-allelic indel mutations in the *Brg1* locus of tumor-derived cell lines.

To assess the phenotypic consequences of mutating *Brg1 in vivo*, we performed detailed histopathological analyses, including tumor burden quantitation, tumor grading, and assessment of metastatic spread (if any). The results of these phenotypic analyses are presented below in **Figure 4**.

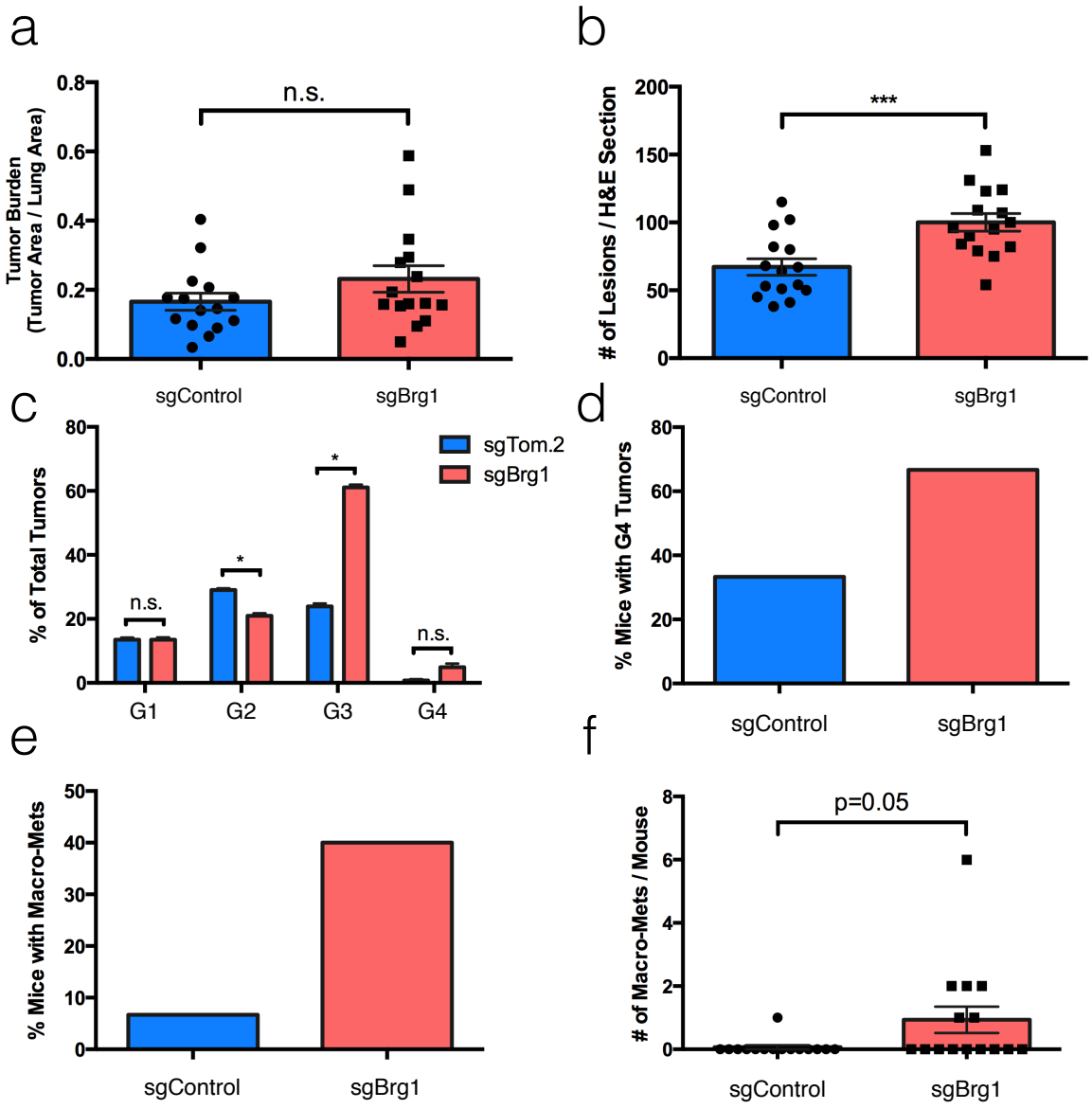


Figure 4: pSECC-mediated disruption of *Brg1 in vivo* promotes several aspects of lung cancer. (a) Overall tumor burden quantitation. (b) Overall number of lesions per H&E section. (c) Tumor Grading analysis. (d) Percentage of mice with Grade 4 tumors (G4). (e) Percentage of mice with macroscopic metastases. (f) Number of macroscopic metastases per mouse.

Interestingly, *Brg1* deletion promoted several aspects of lung cancer *in vivo*. For example *Brg1* deletion led to increased tumor initiation, as demonstrated by the statistically significant increase in the number of lesions per histological section (**Figure 4B**). Indeed, we also observed a similar trend when performing serial μCT on some of these mice, whereby pSECC-sgBrg1 mice had many more small tumors during early time points (data not shown). However, *Brg1* deletion did not significantly impact the overall tumor burden (which takes into account the area of lung covered by tumors instead of the number of individual lesions within a given lung lobe). Albeit this data is somewhat difficult to reconcile, the same unexpected phenomenon has been seen by another research group who generated $Kras^{LSL\text{-}G12D/+;}$ $Brg1^{flox/flox}$ mice and carried out very similar experiments as the ones described here (Eric Collisson & Shivani Malik, personal communication). Therefore, it appears that *Brg1* plays a critical role in suppressing oncogenic transformation and tumor initiation, but might play a less important role at middle stages of tumor progression. The only way to effectively dissect these putative context-specific and spatiotemporal functions of *Brg1* (or any other gene for that matter) is to employ sophisticated GEMMs of cancer that allow for inducible and reversible spatiotemporal control of multiple recombinases in combination with shRNAs and/or sgRNAs (these systems were extensively described in **Parts I-II** of the introduction). Nevertheless, the fact that two independent groups have observed the same biological phenomenon makes it warrant further investigation.

Beyond accelerating tumor initiation, *Brg1* deletion also shifted the tumor spectrum (**Figure 4C**), with *Brg1* deficient mice harboring a higher percentage of Grade 3 tumors and a slight (but not statistically significant) increase in the percentage of Grade 4

tumors and the percentage of individual mice harboring grade 4 tumors (**Figure 4D**). Strikingly, *Brg1* deletion led to a ~ 6-fold increase in the incidence of metastatic disease, with 6/15 pSECC-sgBrg1 mice having one or more metastases vs only 1/15 pSECC-Control mice (**Figure 4E**). Moreover, *Brg1* deletion led to a significant increase in the number of macroscopic metastases per individual mouse, with multiple mice having > 2 macroscopic metastases (**Figure 4F**). Collectively, these results indicate that *Brg1* deletion promotes several aspects of lung cancer initiation and progression *in vivo*. Thus, *Brg1* is indeed a lung adenocarcinoma tumor suppressor gene.

Generation of isogenic *Brg1* wild type and mutant cell lines using CRISPR-Cas9

Next, we decided to generate isogenic KP and KP-*Brg1*^{KO} cell lines (hereafter referred to as KPB cell lines) using CRISPR-Cas9 (**Figure 5A-B**). One of the main motivations behind generating this pair of isogenic cell lines was to test whether *Brg1* mutation would confer sensitivity to combined EZH2 inhibition and Etoposide (Fillmore et al., 2015). Remarkably, CRISPR-mediated mutation of *Brg1* in the background of oncogenic *Kras* and *p53* loss completely recapitulated the findings obtained by Carla Kim and colleagues, which were obtained by analyzing a panel human NSCLC cell lines (**Figure 5C**) (Fillmore et al., 2015). The fact that engineering a single mutation fully recapitulated these findings is truly remarkable, as it speaks to the utility of generating and applying these and other sets of engineered isogenic cell lines for uncovering new biology and systematically identifying novel genetic dependencies that arise as a consequence of a mutation in a cancer driver (for example, *Keap1* isogenic cell lines described in **Chapter 3** and *Brg1* isogenic cell lines described in this section are representative examples).

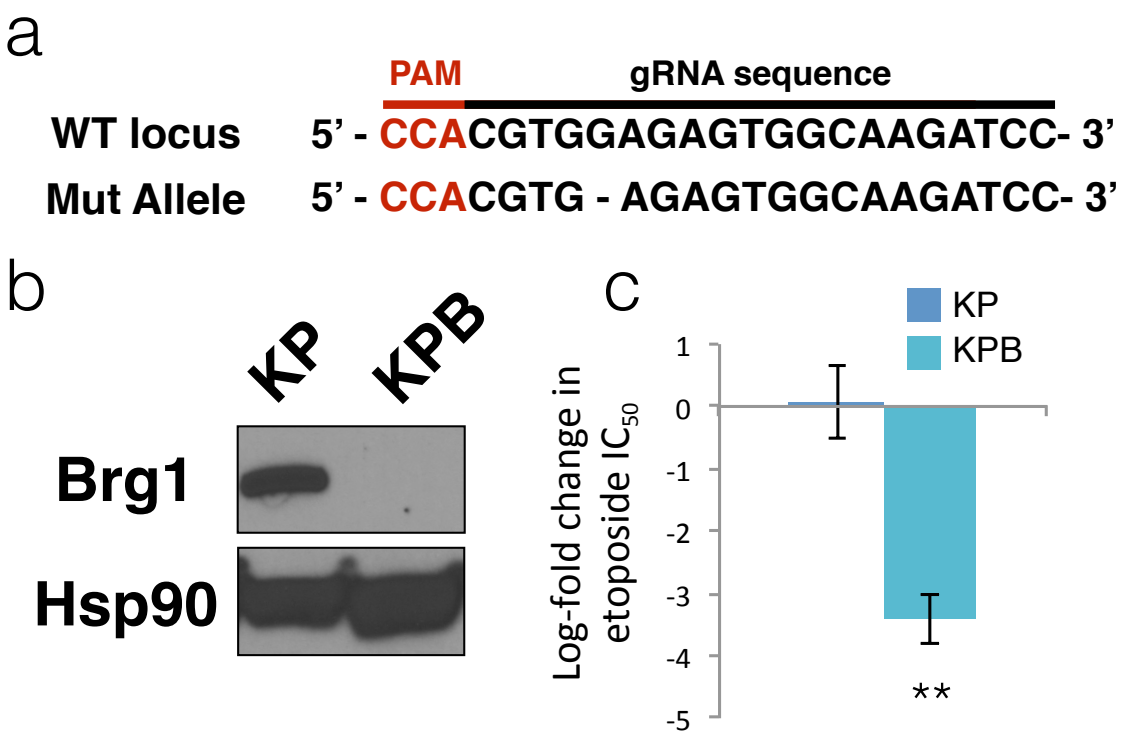


Figure 5: *Brg1* loss of function mutation confers sensitivity to combined EZH2 inhibition and etoposide. (a) A single 1bp frameshift indel is present in the KPB cell line. (b) Western blot demonstrating the complete absence of Brg1 protein. (c) KPB cell lines are exquisitely sensitive to combined EZH2 inhibition and etoposide, whereas KP cell lines are spared. See Fillmore et al., 2015 for details on the combination treatment assay.

Molecular characterization of KP and KPB cell lines via RNA sequencing

Another major motivation behind the generation of these isogenic KP and KPB cell lines was to perform whole-transcriptome analysis through RNA sequencing (RNAseq) with the goal of molecularly characterizing *Brg1* mutant cell lines. This dataset has already proven particularly powerful for dissecting the mechanisms by which *Brg1* deletion promotes several aspects of lung adenocarcinoma. Moreover, it has already stimulated multiple hypotheses with immediate therapeutic potential for this particular patient subpopulation. I will briefly describe the RNAseq data set, initial findings, and a few

hypothesis that have stemmed from this recent analysis (**Figure 6**). The first interesting observation derived from the RNAseq data is the fact that KPB cells show a massive global downregulation of gene expression (**Figure 6A-B**).

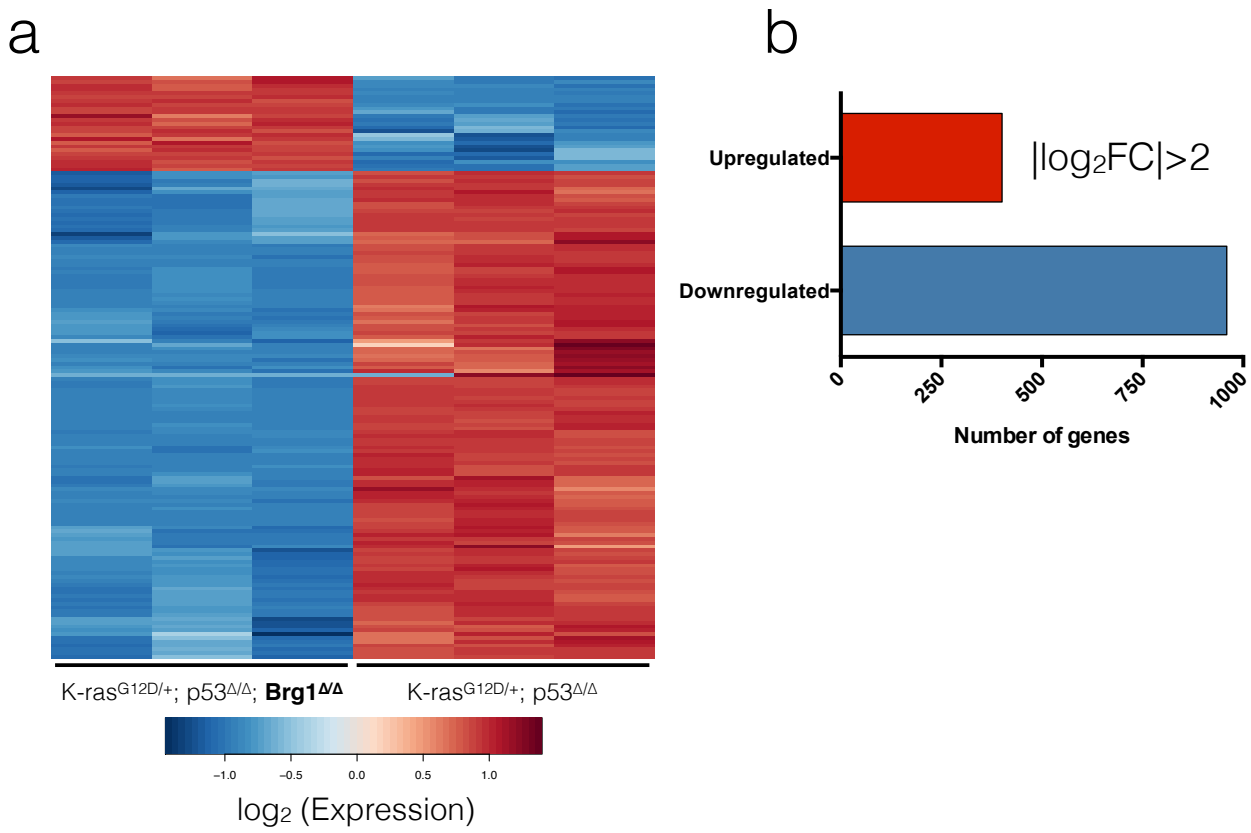


Figure 6: KPB cells exhibit major transcriptional differences and a massive downregulation of gene expression when compared to KP cells. (a) Hierarchical clustering of the differentially expressed genes (Fold Change > 2) between KP and KPB cell lines represented as a heat map. Each column represents a replicate for the indicated group. This gene expression analysis shows that KPB cells are associated with a specific gene expression signature (b) When we look at the number of genes that are significantly changing (Fold Change >2), we observe that the majority of the differentially expressed genes between KPB cells vs KP cells are downregulated, suggesting a potential transcriptional activation role for *Brg1* in lung adenocarcinoma.

This is of particular interest because it suggests that *Brg1* is playing a role in activating transcription in lung adenocarcinoma cells. If the tumor suppressive role of *Brg1* is via the transcriptional activation of a defined set of genes, then identifying those genes could potentially give insight into the tumor suppressive pathways that *Brg1* is regulating in the context of lung adenocarcinoma. Moreover, *Brg1* functions are known to be highly context dependent; for example, it is well established that *Brg1* plays a transcriptional repressive role in ES cells (Wilson and Roberts 2011). Our data suggests that *Brg1* is playing the complete opposite role in lung adenocarcinoma. Another interesting (and preliminary) observation derived from this RNAseq dataset is the fact that the gene signature associated with *Brg1* mutation is able to predict the 5-year survival of lung adenocarcinoma patients that specifically harbor mutations in *KRAS* (**Figure 7**).

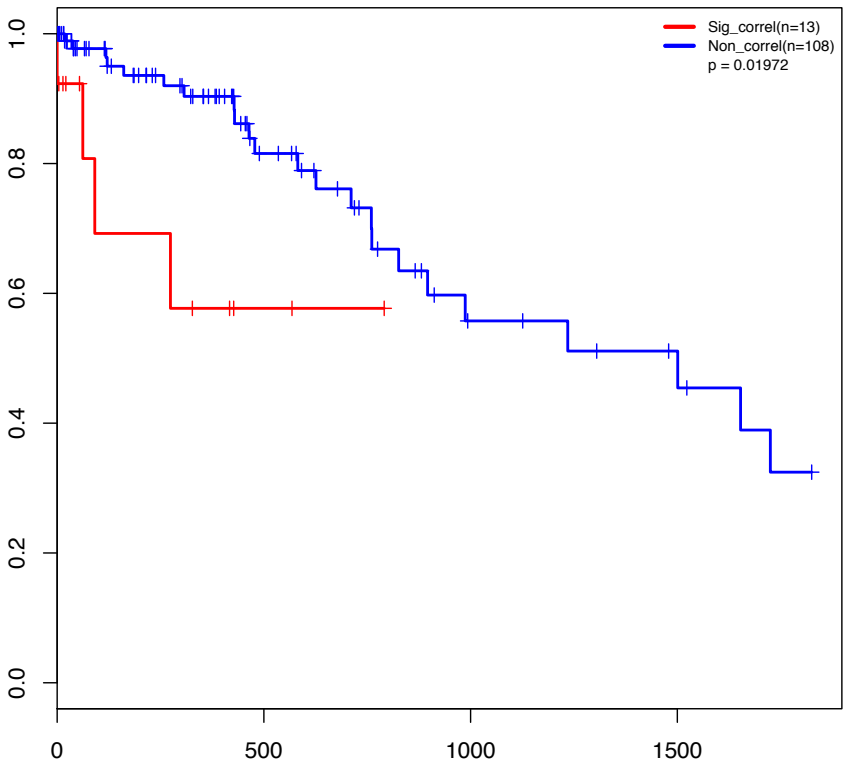


Figure 7: The *Brg1*-mutant-derived gene signature has predictive power in *KRAS* mutant patients. Red = correlation with KPB gene signature (only downregulated genes from signature here).

Another set of very interesting and exciting observations comes from our preliminary Gene Set Enrichment Analysis (GSEA) (Subramanian et al., 2005). Our analysis has revealed that *Brg1* mutant cell lines dramatically upregulate multiple gene sets involved in several aspects of the malignant phenotype, including E2F-family related gene sets (**Figure 8A**) (which are known to antagonize the *Rb* tumor suppressor gene), Wnt pathway gene sets (**Figure 8B**), and gene sets strongly associated with drug resistance (**Figure 8C**). Conversely, our analysis has revealed that *Brg1* mutant cell lines substantially downregulate multiple gene sets involved in several tumor suppressive pathways or pathways that are involved in counteracting the malignant phenotype in one way or another, including the strong repression of genes that mediate drug sensitivity (**Figure 8D**).

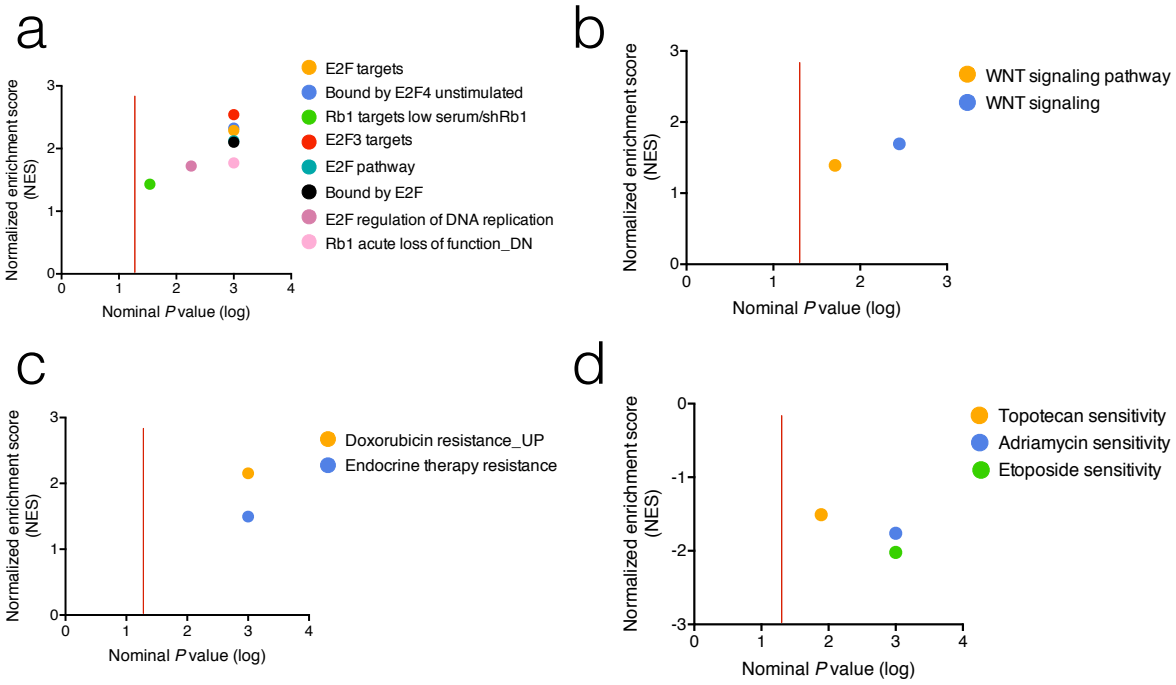


Figure 8: *Brg1*-mutant tumors upregulate and downregulate multiple gene sets that are known to be involved in the malignant phenotype. (a-c) Upregulated gene sets corresponding to (a) E2F family, (b) Wnt pathway, and (c) drug resistance. (d) Downregulated gene sets corresponding to drug sensitivity.

The results described above are provocative, and immediately suggest multiple hypotheses that can be rigorously tested using the KP-KPB set of isogenic cell lines. For example, the fact that *Brg1*-mutant tumors hyperactivate the Wnt pathway (**Figure 8B**) led us to hypothesize that these tumors might be sensitive to Wnt pathway inhibition. Indeed, our preliminary data suggest that Wnt pathway inhibition has a significant therapeutic effect in *Brg1*-mutant cells and *Brg1*-mutant tumors *in vivo*. These and other experiments will hopefully contribute to the development of novel therapeutic approaches for the benefit of this subpopulation of lung cancer patients.

ACKNOWLEDGEMENTS

We thank T. Volkert and S. Gupta for RNA Sequencing support, G. Paradis, M. Griffin, M. Jennings and M. Saturno-Condón for flow cytometry and FACS support, and K. Cormier and C. Condon from the Hope Babette Tang (1983) Histology Facility for technical support. This work was supported by the Howard Hughes Medical Institute, and in part by Cancer Center Support (core) grant P30-CA14051 from the National Cancer Institute. T.J. is a Howard Hughes Medical Institute Investigator, the David H. Koch Professor of Biology, and a Daniel K. Ludwig Scholar.

REFERENCES

Cancer Genome Atlas Research Network (2014). Comprehensive molecular profiling of lung adenocarcinoma. Nature 511, 543–550.

Figura, von, G., Fukuda, A., Roy, N., Liku, M.E., Morris, J.P., IV, Kim, G.E., Russ, H.A., Firpo, M.A., Mulvihill, S.J., Dawson, D.W., et al. (2014). The chromatin regulator Brg1 suppresses formation of intraductal papillary mucinous neoplasm and pancreatic ductal adenocarcinoma. Nat Cell Biol 16, 255–267.

Fillmore, C.M., Xu, C., Desai, P.T., Berry, J.M., Rowbotham, S.P., Lin, Y.-J., Zhang, H., Marquez, V.E., Hammerman, P.S., Wong, K.-K., et al. (2015). EZH2 inhibition sensitizes BRG1 and EGFR mutant lung tumours to TopoII inhibitors. Nature –.

Glaros, S., Cirrincione, G.M., Palanca, A., Metzger, D., and Reisman, D. (2008). Targeted Knockout of BRG1 Potentiates Lung Cancer Development. Cancer Research 68, 3689–3696.

Imielinski, M., Berger, A.H., Hammerman, P.S., Hernandez, B., Pugh, T.J., Hodis, E., Cho, J., Suh, J., Capelletti, M., Sivachenko, A., et al. (2012). Mapping the hallmarks of lung adenocarcinoma with massively parallel sequencing. Cell 150, 1107–1120.

Medina, P.P., Romero, O.A., Kohno, T., Montuenga, L.M., Pio, R., Yokota, J., and Sanchez-Cespedes, M. (2008). Frequent BRG1/SMARCA4-inactivating mutations in human lung cancer cell lines. Hum. Mutat. 29, 617–622.

Orvis, T., Hepperla, A., Walter, V., Song, S., Simon, J., Parker, J., Wilkerson, M.D., Desai, N., Major, M.B., Hayes, D.N., et al. (2014). BRG1/SMARCA4 Inactivation Promotes Non–Small Cell Lung Cancer Aggressiveness by Altering Chromatin Organization. Cancer Research.

Roy, N., Malik, S., Villanueva, K.E., Urano, A., Lu, X., Figura, von, G., Seeley, E.S., Dawson, D.W., Collisson, E.A., and Hebrok, M. (2015). Brg1 promotes both tumor-suppressive and oncogenic activities at distinct stages of pancreatic cancer formation. Genes Dev 29, 658–671.

Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S., et al. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci USA 102, 15545–15550.

Wilson, B.G., and Roberts, C.W.M. (2011). SWI/SNF nucleosome remodellers and cancer. Nat Rev Cancer 11, 481–492.