Aus dem Universitäts KrebsCentrum
Direktor: Herr Prof. Dr. med. Gerhard Ehninger

Developing the CRISPR/Cas-system for Inactivation of Proto-oncogenes in Human Cancer Cells

Dissertationsschrift

zur Erlangung des akademischen Grades
Doktor der Medizin
Doctor medicinae (Dr. med.)
vorgelegt
der Medizinischen Fakultät Carl Gustav Carus
der Technischen Universität Dresden

von

Christina Gebler
aus Lübben (Spreewald)

Dresden 2017

1. Gutachter: Prof. Dr. rer. nat. Frank Buchholz

2. Gutachter: Prof. Dr. med. Axel Roers

Tag der mündlichen Prüfung: 19.06.2018

gez.: Prof. Dr. rer. nat. habil. Henning Morawietz
Vorsitzender der Promotionskommission

Table of contents

List of tables

List of figures

List of abbreviations

AA	Amino acid
AAV	Adeno-associated virus
ABL	Abelson murine leukemia viral oncogene homolog 1
AKT	Protein kinase B
AML	Acute myeloid leukemia
ARF	Alternate reading frame tumor suppressor
Bcl-2	B-cell lymphoma 2
BCR	Breakpoint cluster region
bp	Base pair
BRAF	V-Raf murine sarcoma viral oncogene homolog B1
Cas	CRISPR associated proteins
CCND1	Cyclin D1
CCNE1	Cyclin E1
CDS	Coding DNA sequence
CEBPA	CCAAT/Enhancer-Binding Protein, Alpha
CjCas9	*Campylobacter jejuni* Cas9
Cpf1	Centromere and Promoter Factor 1
CRC	Colorectal cancer
CRISPR	Clustered regularly interspaced short palindromic repeats
crRNA	CRISPR RNA
DAPI	4',6-diamidino-2-phenylindole
DMEM	Dulbecco's modified eagle medium
DNA	Deoxyribonucleic acid
DNMT3A	DNA Methyltransferase 3A
DSB	Double strand break
EGFR	Epidermal growth factor receptor
FACS	Fluorescence-activated cell sorting
FBS	Fetal bovine serum
FGF	Fibroblast growth factor
FLT3	FMS-related tyrosine kinase 3
FLT3-ITD	FLT3 internal tandem duplication
GFP	Green fluorescent protein
gRNA	Guide RNA
h	Hour
HDR	Homology directed repair
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HER2/Neu	Human epidermal growth factor receptor 2
HRAS	Harvey rat sarcoma viral oncogene homolog
IDH1	Isocitrate dehydrogenase 1

IgG	Immunglobuline G
KDR	Kinase insert domain receptor
KIT	V-Kit Hardy-Zuckerman 4 Feline Sarcoma Viral Oncogene Homolog
KRAS	Kirsten rat sarcoma viral oncogene homolog
MEK	Mitogen-activated protein kinase kinase
MET	Hepatocyte growth factor receptor
min	Minute
mL	Milliliter
MOI	Multiplicity of infection
mRNA	Messenger RNA
mTOR	Mechanistic target of rapamycin
mut	Mutant
NES	Nuclear export signal
NF2	Neurofibromin 2
ng	Nanogram
NGS	Next generation sequencing
NHEJ	Non-homologous end joining
NLS	Nuclear localization signal
NmCas9	*Neisseria meningitidis* Cas9
NoLS	Nucleolar localization signal
NPM1	Nucleophosmin/nucleoplasmin family member 1
NSCLC	Non-small-cell lung carcinoma
p53	Tumor protein p53
PAM	Protospacer adjacent motif
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PDGF	Platelet-derived growth factor
PDGF	Platelet-derived growth factor, alpha polypeptide
PEI	Polyethylenimine
PFA	Paraformaldehyde
PI	Propidium iodide
PI3K	Phosphatidylinositol 3-kinase
PIK3CA	Phosphatidylinositol 3-kinase, catalytic alpha
PML	Inducer of Acute Promyelocytic Leukemia
pRb	Retinoblastoma protein
ps	Protospacer
PTEN	Phosphatase and tensin homolog
PTPN11	Protein-Tyrosine Phosphatase, Nonreceptor-Type, 11
RAF	Rapidly accelerated fibrosarcoma
RAR	Retinoic acid receptor

Ras	Ras protein family (HRas, KRas,NRas)
RFP	Red fluorescent protein
RNA	Ribonucleic acid
rpm	Rounds per minute
RPMI	Roswell Park Memorial Institute medium
RT-qPCR	Reverse transcription quantitative PCR
SaCas9	*Staphylococcus aureus* Cas9
SCR7	5,6-bis((E)-benzylideneamino)-2-mercaptopyrimidin-4-ol
SD	Standard deviation
sgRNA	Single gRNA
siRNA	Small interfering RNA
SpCas9	*Streptococcus pyogenes* Cas9
SRC	V-Src Avian Sarcoma (Schmidt-Ruppin A-2) Viral Oncogene
TGFα	Transforming growth factor alpha
TGFβ	Transforming growth factor beta
tracrRNA	Trans-activating crRNA
U	Unit
U2AF1	U2 Small Nuclear Rna Auxiliary Factor 2
V	Volt
VEGF	Vascular endothelial growth factor
WT	Wild type
WT1	Wilms tumor gene
αMEM	Alpha minimal essential medium
°C	Degree Celsius
µg	Microgram
µL	Microliter
µm	Micrometer
µM	Micromolar

1 Introduction

1.1 Cancer

The lifetime risk of being diagnosed with cancer is about 40-50% in developed countries like Great Britain or the United States of America (Ahmad et al., 2015; Howlader et al., 2016). In 2014 more than 200,000 people died from cancer in Germany, making the disease the second most common cause of mortality (Statistisches Bundesamt, 2015). Research on cancer treatment is one of the fastest growing branches of medicine, consequently survival-rates significantly improved over the past years. In the USA, for example, five-year-survival for all cancer types improved from approximately 49% in the late 1970s to 67% in 2012 (Howlader et al., 2016).

Albeit the term cancer comprises a broad variety of entities, they all have distinct characteristics in common. Tumorigenesis is based on various changes in the cells occurring stepwise over a longer time period. Transformed cells employ different mechanisms to achieve their essential characteristics. For growth self-sufficiency, cells use autocrine and paracrine growth factor stimulation, e.g. the production of platelet-derived growth factor (PDGF) and transforming growth factor alpha (TGFα). Cancer cells may also employ altered growth signal receptors and pathways, resulting in hyperactivated and deregulated responses, like Human epidermal growth factor receptor 2 (*HER2/Neu*) overexpression. The insensitivity to growth-suppressors represents the second characteristic of transformed cells. The cell cycle gatekeeper circuits, regulated by retinoblastoma protein (pRb) and tumor protein 53 (p53), are often disrupted. Transformed cells are also able to circumvent tumor suppressor effects and the inhibitory effects of transforming growth factor beta (TGFβ) and cell-cell-contact on cell growth. The third hallmark is the resistance to cell death. Again, p53 plays a central role through the induction of apoptosis after deoxyribonucleic acid (DNA) damage. p53 loss-of-function mutations mediate insensitivity to DNA damage. Furthermore, anti-apoptotic B-cell lymphoma 2 (Bcl-2), involved in mitochondrial death signaling, is often altered in transformed cells. To achieve the fourth characteristic, replicative immortality, cells have to overcome senescence and to conserve telomeres. While mechanisms of the former are not completely understood yet, the latter is promoted by telomerase expression in about 90% of malignant cells. As the supply of oxygen and nutrient limits tissue expansion, the induction of angiogenesis is another necessary characteristic. Possible trigger factors are the increased expression of vascular endothelial growth factor (VEGF) or fibroblast growth factor (FGF), downregulation of angiogenesis inhibitors and immune cell caused inflammation, leading to a formation of abnormal vessels. Further characteristics of transformed cells include genome instability, deregulation of the metabolism, tumor-promoting inflammation and evasion of

immune destruction. The last characteristic comprises invasion and metastasis, the main reasons for dying from cancer. These processes require complex changes in cell-environment interactions and cell motility and can be promoted by the alteration of adherence proteins, for example by the loss of E-cadherin or the permutation of integrins, the increased expression of proteases and the epithelial-mesenchymal transition. Even provided with all characteristics, the great majority of resulting transformed cells die. But analogously to evolutionary processes, single cell clones survive, start a massive and invasive cell proliferation and bypass control mechanisms (Hanahan and Weinberg, 2000; Hanahan and Weinberg 2011).

1.2 Oncogenes

1.2.1 Role in cancer

Oncogenes are genes that arise from altered proto-oncogenes. The gene products of wild type (WT) proto-oncogenes affect proliferation, differentiation, angiogenesis and cell death. Thus, they play a crucial role in the development of tissues, organogenesis, embryogenesis and growth. After termination of these processes, proto-oncogenes are physiologically turned off or down-regulated. Proto-oncogenes include membrane receptors for growth factors, e.g. Epidermal growth factor receptor (*EGFR)* and Kinase insert domain receptor (*KDR)*, and proteins downstream the signaling pathway of these receptors, e.g. Kirsten rat sarcoma viral oncogene homolog (*KRAS)* and Harvey rat sarcoma viral oncogene homolog (*HRAS)*. Moreover, some cell cycle regulators as cyclin D1 (*CCND1)* and cyclin E1 (*CCNE1)* are classified as proto-oncogenes. Mechanisms of the proto-oncogene into oncogene transformation comprise various mutations on gene and chromosome level, namely point mutations, deletions, insertions, inversions, gene amplifications and chromosomal translocations. The mutations can either directly affect the coding regions of the gene or the regulating elements like the promoter. As a result, the former proto-oncogene, now oncogene, shows an increased expression, is translated into a hyperactivated gene product and/or into a fusion protein. Furthermore, these alterations lead to the circumvention of regulation processes and in the end to an uncontrolled proliferation. Taken together these consequences of unregulated oncogene expression play a key role in achieving the hallmarks of cancer mentioned above (Chial, 2008).

Studies in human cancer cell lines, as well as clinical trials, show evidence for cancer cells being dependent on the expression of distinct oncogenes. Targeted therapies that inhibit the products of these genes lead to loss of the malignant phenotype. These findings are the basis for the “oncogene addiction theory” (Weinstein, 2002).

In recent years, big sets of cancer types have been analyzed with next generation sequencing (NGS), uncovering certain patterns. Solid tumors harbor numerous DNA changes,

consequently altering the protein product of 33 to 66 genes per tumor in average. Only two to eight mutations per tumor are driver mutations, driving the initial cell transformation and providing a growth advantage, whereas the others are supposed to be passenger mutations and do not confer growth advantage. The vast majority of driver mutations are point mutations that either activate oncogenes or inactivate tumor suppressor genes. Comprising all types of solid tumors, more than 50 genes, involved in 12 signaling pathways, have been identified to harbor activating driver mutations. Furthermore, about 70 genes have been discovered to contain inactivating driver mutations. However, the discrimination between driver and passenger mutations is mainly based on *in silico* analysis, namely the analysis of mutation patterns and frequencies (Vogelstein et al., 2013).

Taken together, "[genetic] mutations are a hallmark of cancer development, and more than 140 cancer driver genes have been described to date ([Kandoth et al., 2013; Vogelstein et al., 2013]). [But although identification] of all mutations in an actual tumor of a patient by whole-genome sequencing is rapidly emerging as the method of choice for precision diagnostics ([Dewey et al., 2014]) […], detailed knowledge of the functional roles and relevance of most mutations arising during tumorigenesis [is] still lacking." (Gebler et al., 2017)

1.2.2 Targeted therapies

Mutational status of a tumor does not only influence a patient's prognosis but therapeutic strategies as well. Based on genetic alterations, monoclonal antibodies and small molecule inhibitors targeting these alterations show good efficacy. Next to cytotoxic chemotherapy and hormonal therapy, the targeted therapy has become the third keystone in pharmacological cancer treatment with the number of available drugs continuously increasing (Stuart and Sellers, 2009).

As targeted therapies are selective against their targets – the altered proteins in cancer cells (Supplementary Table 1) – they do not affect healthy tissue. Consequently, the molecular diagnostics of cancer tissue becomes more and more important in personalized cancer treatment. For instance, only non-small-cell lung carcinoma (NSCLC) patients with mutated *EGFR* show a positive clinical response to EGFR tyrosine kinase inhibitor gefitinib, but not the patients with WT *EGFR* (Paez et al., 2004).

Supplementary Table 1: Targeted therapies.

Examples of diseases with available targeted therapies and corresponding genetic alterations (Stuart and Sellers, 2009; Zhao and Adjei, 2014).

PDGFA – Platelet-derived growth factor, alpha polypeptide; KIT – V-Kit Hardy-Zuckerman 4 Feline Sarcoma Viral Oncogene Homolog; PML – Inducer of Acute Promyelocytic Leukemia; RAR – Retinoic acid receptor; BRAF – V-Raf murine sarcoma viral oncogene homolog B1

Disease	Genetic alteration	Target	Drug(s)	Category
Gastric cancer	*HER2/Neu* overexpression	HER2/Neu	Trastuzumab	Antibody
Breast cancer	*HER2/Neu* overexpression	HER2/Neu	Trastuzumab	Antibody
Colorectal cancer	*EGFR* amplification	EGFR	Cetuximab	Antibody
Squamous cell carcinoma of the head and neck	*EGFR* amplification	EGFR	Cetuximab	Antibody
Lung cancer	*EGFR* mutation	EGFR	Erlotinib, Gefitinib	Kinase inhibitor
Gastrointestinal stromal tumors	*PDGFA* mutation	PDGF-R	Imatinib	Kinase inhibitor
	KIT mutation	c-kit	Imatinib	Kinase inhibitor
Chronic myelogenous leukemia	*BCR-ABL* (t9;22)	abl	Imatinib	Kinase inhibitor
Melanoma	*BRAF* mutation	B-Raf	Vemurafenib	Kinase inhibitor
Acute promyelocytic leukemia	*PML–RAR* (t15;17)	unknown	All-*trans*-retinoic acid	Other

Escape mechanisms from targeted therapies have been reported, often through the reactivation of a particular oncogene or pathway, thus emphasizing the dependency of cancer cells on that pathway (Weinstein, 2002). One mechanism of cancer cell resistance is caused by additional mutations in the target protein, producing proteins insensitive to inhibitors or antibodies. For instance, additional *BCR-ABL* (fusion gene of Breakpoint cluster region and Abelson murine leukemia viral oncogene homolog 1) and *EGFR* mutations cause imatinib and gefitinib resistance, respectively (Gorre et al., 2001). Another mechanism includes alterations of proteins downstream the target pathway. The resistance to gefitinib in NSCLC is mediated through Hepatocyte growth factor receptor (MET) amplification, rescuing Phosphatidylinositol-3-kinase-Protein-kinase-B-Mechanistic-target-of-rapamycin (PI3K-AKT-mTOR) pathway

(Turke et al., 2010). Those alterations are present in pre-treatment cancer cell populations and cells harboring them expand during treatment (Pao et al., 2005).

Only about 30 of the identified oncogenic driver genes can theoretically be targeted by inhibiting kinase domain or by small molecules. Some of the existing drugs specifically target the altered protein, like Vemurafenib inhibits only B-Raf^{V600E}. Other drugs inhibit the WT and mutant protein, like the currently existing *PIK3CA* inhibitors, and therefore also affect healthy tissue to a certain extent. For the remaining oncogenes without a targetable enzymatic activity, e.g. for the potential oncogene *NPM1*, targeted therapeutic strategies need to be yet established (Lee et al., 2010; Vogelstein et al., 2013; Fritsch et al., 2014).

The majority of mutations occurring in cancer cells is presently not studied in respect to their function in tumorigenesis or only partially understood, like the role of *NPM1* mutations. Hence, a huge amount of data about cancer genomics has the potential to be converted into personalized medicine (Grisendi et al., 2006; Chin et al., 2011).

1.2.3 *NPM1*

Nucleophosmin/nucleoplasmin family member 1 *(NPM1)* encodes the 294-amino acid phosphoprotein nucleophosmin, located in the nucleolus (Chan et al., 1989). Nucleophosmin has various functions and contains several functional protein domains. The oligomerization domain is required to form pentamers and decamers. The resulting complex acts as a histone chaperone. Furthermore, it contains a domain for binding nucleic acids. Nucleophosmin is active in centrosome duplication during mitosis, in ribosome biogenesis as well as in the repair, replication and transcription of DNA. Two nuclear export signals (NES), two nuclear localization signals (NLS) and one nucleolar localization signals (NoLS) in nucleophosmin reflect the ability of nucleophosmin to shuttle between cytoplasm, nucleoplasm and nucleolus during cell cycle (Federici and Falini, 2013). Besides interactions with many other proteins, nucleophosmin activates p53 together with Alternate reading frame tumor suppressor (ARF). Genetic alterations of *NPM1*, as translocations, mutations and deletions, occur in cancer cells and hematological diseases. Considering its multiple interactions, *NPM1* is thought to either be a proto-oncogene or a tumor suppressor, depending on circumstances (Grisendi et al., 2006). *NPM1* is found to be mutated and localized in the cytoplasm in about 35% of adult primary acute myeloid leukemia (AML) patients' samples, particularly in normal karyotype AML. *NPM1* mutations lead to frameshifts in the nucleic acid binding domain, whereby nucleophosmin gains an additional NES and loses NoLS (tryptophan 288 and 290), crucial for nucleolar localization. The most frequent *NPM1* mutation is type A (c.863_864insTCTG). In heterozygous mutated cells, mutated nucleophosmin interacts with WT nucleophosmin in a

dominant-negative way, resulting in WT and mutant protein being translocated into cytoplasm together (Falini et al., 2005; Falini et al., 2006).

Currently, there is no nucleophosmin inhibitor approved for the treatment of patients, but research is being conducted on several drugs. Knock-down with small interfering ribonucleic acid (siRNA) specific for mutated *NPM1* inhibits growth of AML cell lines, indicating oncogenic potential of *NPM1* in AML. Deguelin is a rotenoid showing anti-cancer effects. Although detailed interactions and target molecules are a current topic of research, compared to WT cells, deguelin significantly enhances apoptosis in *NPM1* mutant OCI-AML3 cells and selectively diminishes mutant nucleophosmin expression (Yi et al., 2015). Another potential nucleophosmin targeting drug, NSC348884, disrupts the oligomerization of WT and mutated nucleophosmin. In the OCI-AML3 cell line, it induces apoptosis at lower concentrations than in WT cell lines (Qi et al., 2008).

1.2.4 *BRAF*

The *BRAF* gene encodes a 651-amino acid serine/threonine kinase from Rapidly accelerated fibrosarcoma (Raf) kinase family (Sithanandam et al., 1990). As part of Ras-Raf-MEK pathway (MEK – Mitogen-activated protein kinase kinase), the gene product B-Raf plays a role in the regulation of cell proliferation. *BRAF* is known to be a proto-oncogene in human cells and is mutated in many types of cancer with the highest frequencies in melanoma. The substitution c.1799T>A (p.V600E) is the most common mutation in *BRAF*, generating higher kinase activity and increasing malignant transformation of cells (Davies et al., 2002). As this mutation changes conformation in the activation loop, imitating activating phosphorylation, B-Raf becomes constitutively active and promotes proliferation (Wan et al., 2004). The role of *BRAF*^{V600E} mutation in melanoma has been confirmed by RNA interference (RNAi) experiments. Lentiviral delivery of *BRAF*^{V600E} specific siRNA inhibits melanoma cell proliferation, increases cell death and decreases invasion *in vitro*. Furthermore, it reduces tumorigenicity *in vivo*, rendering *BRAF* a putative target for molecular treatment (Sumimoto et al., 2004).

Targeted therapy of *BRAF* has first been established with Vemurafenib (PLX4032), a selective B-Raf^{V600E} inhibitor, showing therapeutic effects only in melanoma with mutated *BRAF* (Lee et al., 2010; Chapman et al., 2011). More B-Raf^{V600E} inhibitors are being currently tested in clinical trials.

In colorectal cancer (CRC), *BRAF*^{V600E} mutation is found as well. In contrast to melanoma, inhibition of only mutated B-Raf reduces cell viability and tumor growth, but does not induce apoptosis. Rather a combination of B-Raf^{V600E} and PI3K/mTOR inhibition is needed to avoid

compensational PI3K-AKT-mTOR over-activation. Under this treatment, apoptosis is enhanced and tumor size is reduced *in vivo* (Coffee et al., 2013).

1.2.5 *PIK3CA*

Class I PI3K are heterodimeric protein complexes and contain regulatory p85 and catalytic p110 subunit. In the case of the latter, three isoforms (α, β and δ) exist. Phosphatidylinositol 3-kinase, catalytic alpha (*PIK3CA)* gene encodes for p110α subunit (Carpenter et al., 1990; Volinia et al., 1994). This isoform promotes proliferation, survival, and metabolism via PI3K-AKT-mTOR pathway. Tumor suppressor Phosphatase and tensin homolog (PTEN) inhibits these effects (Li et al., 2014). Mutations and gene amplifications of *PIK3CA* occur with high frequencies in many types of cancer. In fact, *PIK3CA* is the most common mutated oncogene with an average of 17.8% across all cancer entities (Kandoth et al., 2013). Hot spot mutation c.3140A>G (p.H1047R) is located in the kinase domain close to activation loop, thus changing interaction with the substrate phosphatidylinositide. As a result, lipid kinase activity of p100α is enhanced (Huang et al., 2007). *PIK3CA*^{H1047R} mutation leads to tumor transformation *in vivo*, emphasizing its role as an oncogene (Bader et al., 2006).
As downstream component in tyrosine kinase signaling pathway, *PIK3CA* has been proposed as a therapeutic target (Whyte and Holbeck, 2006). Unlike pan-PI3K inhibitors, PI3Kα inhibitors like NVP-BYL719 show specific activity against this PI3K isoform. Cell lines from a cancer panel harboring *PIK3CA* mutations or amplifications are more sensitive to this drug than WT cell lines are (Fritsch et al., 2014). This drug is currently under clinical evaluation with breast cancer patients with first results indicating a better clinical response in *PIK3CA* mutant patients (Fritsch et al., 2014; Mayer et al., 2017).

1.3 CRISPR/Cas-system

Bacteria and archaea employ a defense mechanism against phages and plasmids, analogous to the adaptive immune system of higher animals, called the CRISPR/Cas-system (CRISPR - Clustered regularly interspaced short palindromic repeats; Cas - CRISPR associated proteins). It consists of repetitive sequences in the prokaryotic genome with 24-48 base pairs (bp) in length, separated by spacers of similar length. Spacer sequences match phage DNA sequences and new spacers are integrated into the genome during the infection with phages, thus permitting adaptive defense (Barrangou et al., 2007). A guide RNA (gRNA), transcribed from spacers and repeats, complexes with Cas endonuclease and binds to a sequence of complementary phage DNA, the protospacer. If a protospacer adjacent motif (PAM), a short DNA sequence specific for each species' CRISPR/Cas-system, is present downstream of the protospacer, the endonuclease generates double strand breaks (DSB) in the phage DNA,

which is degraded afterwards (Supplementary Figure 1). Today, the most utilized Cas endonuclease is class 2, type II *Streptococcus pyogenes* Cas9 (SpCas9). The target sequence for the corresponding gRNA is 20 bp in length, followed by 5'-NGG-3' PAM (N stands for any nucleotide). A blunt DSB is generated 3 bp upstream the PAM. The type II Cas9 proteins require a trans-activating crRNA (tracrRNA) to first process the CRISPR RNA (crRNA), which contains the sequence complementary to the protospacer. Second, both crRNA and tracrRNA stay complexed with the Cas9 for the DNA binding and cleavage. For the use as genome editing tool, crRNA and tracrRNA have been fused to a chimeric single gRNA (sgRNA) (Jinek et al., 2012).

Numerous alternatives for SpCas9 can bypass limits of SpCas9. Cas9 orthologues from different bacteria species, such as from *Neisseria meningitidis* (NmCas9), *Campylobacter jejuni* (CjCas9) or *Staphylococcus aureus* (SaCas9), have been discovered. They extend the spectrum of targetable sites. Their protospacers are about 20 bp long, similar to SpCas9. However, their PAM sequences vary in length and sequence, SaCas9, NmCas9 and CjCas9 recognize 5'-NNGRRT-3', 5'-NNGATT-3' and 5'-NNNNRYAC-3' PAMs, respectively (R stands for purine, Y stands for pyrimidine). Furthermore, these orthologue Cas9 proteins are about 300 amino acids shorter compared to SpCas9, thus allowing delivery applications with size limitations, e.g. the use of adeno-associated viruses (AAV) as vectors (Esvelt et al., 2013; Ran et al., 2015; Kim et al., 2017). Most recently, two SpCas9 recognizing the PAMs 5'-NGA-3' and 5'-NAG-3', have been engineered by directed evolution and thereby also expanding the number of targets (Kleinstiver et al., 2015).

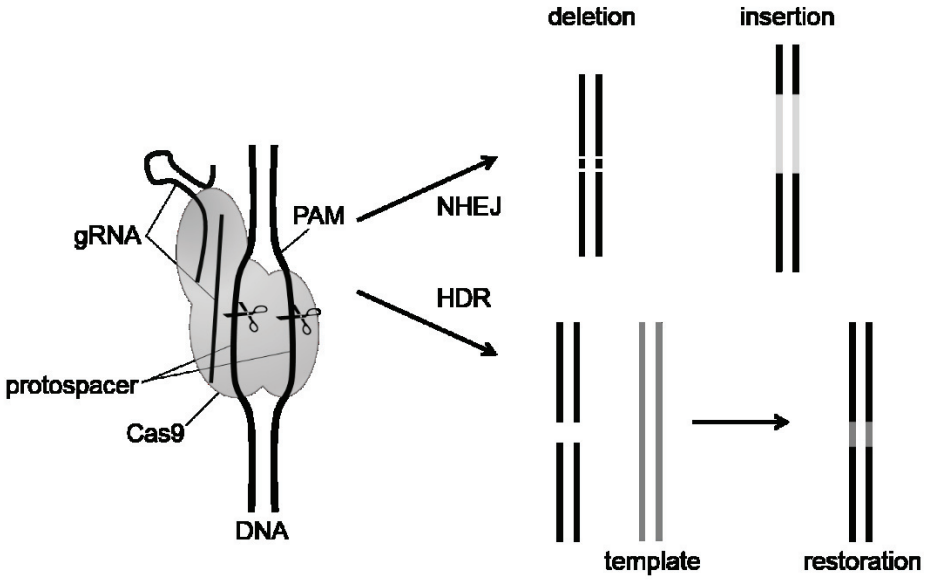
Centromere and Promoter Factor 1 (Cpf1) are class 2, type V Cas proteins found in many species, showing several differences to SpCas9. They do not require a tracrRNA, but only crRNA and they recognize T-rich PAMs. Different than type II Cas9 proteins, they create DSB with overhangs, thus facilitating many applications, for instance gene insertions (Zetsche et al., 2015).

In mammalian cells, DSB are mainly repaired by non-homologous end joining (NHEJ) and homology-directed repair (HDR) (Supplementary Figure 1). The former process joins DNA ends back together, often leading to small insertions or deletions and consequently frame shifts in coding regions. The latter uses a homologous donor, phsysiologically the homolog chromosome, as a template and restores the original DNA sequence completely (Iliakis et al., 2004).

Re-engineered CRISPR/Cas-systems are used as programmable genome editing tools in numerous organisms. The targets are chosen by designing the sequence of sgRNA. The first application in human cells was shown in 2013. Furthermore, multiple site editing is feasible (Cong et al., 2013; Mali et al., 2013). Artificial sgRNAs can be delivered as synthesized RNA or via plasmids, Cas endonucleases as proteins, messenger RNA (mRNA) or plasmids. Other

purposes of CRISPR/Cas-system include deleting genomic regions, repairing mutated genes, delivering transgenes and modifying transcription via nuclease-inactivated Cas coupled to transcription factors or inhibitors (Ran et al., 2013a).

The ability of designed sgRNAs to induce DSB together with Cas9 varies between different targets. The factors influencing sgRNA activity are not completely understood yet, but the sequence of the sgRNA target, as well as the sequences upstream and downstream of it, play a crucial role. Based on large screens, several prediction algorithms of SpCas9 sgRNA activity have been developed (Doench et al., 2014; Chari et al., 2015; Moreno-Mateos et al., 2015; Doench et al., 2016). Testing sgRNA activity can either be conducted by deep sequencing, with a red- and green-fluorescent-protein-reporter (RFP-GFP-reporter) or with endonuclease assays, the latter detecting small insertions or deletions as the result of DSB (Ran et al., 2013a; Karimova et al., 2015).



Supplementary Figure 1: Overview of CRISPR/Cas-system.

In a *Streptococcus pyogenes* CRISPR/Cas-system, a single Cas9 protein is guided by two RNA molecules, activating tracrRNA and protospacer binding crRNA. Cas9-gRNA-complex binds PAMs in DNA. If the DNA sequence upstream the PAM matches the crRNA, Cas9 cleaves DNA 3 bp upstream the PAM. After cleavage, DNA is repaired. Mammalian cells normally fix double strand breaks mainly via error-prone NHEJ in contrast to more precise HDR.

1.4 Aim and motivation

“We set out to test whether the CRISPR/Cas9 system ([Jinek et al., 2012]) can aid the functional investigation of mutations detected in cancer cells.” (Gebler et al., 2017)

Considering the increasing number of therapies targeting oncogene products, the differentiation between the few driver mutations and the numerous passenger mutations occurring in cancer cells is of growing therapeutic interest. Previous gene knock-out methods require great efforts or, in the case of siRNA, show only a temporary and incomplete knock down. Thus, it has been a major purpose of this work to demonstrate CRISPR/Cas-system-driven DNA cleavage as a feasible and fast forward functional investigation tool for cancer mutations.

If used as an oncogene knock-out tool, the important condition, which CRIPS/Cas-system must fulfill, is successful discrimination between mutated and WT oncogenes. In the case of point mutations, this difference is only a single base bond between DNA and sgRNA. Therefore, a system measuring the activity of sgRNAs towards oncogenes in comparison to WT alleles of the corresponding proto-oncogenes had to be established first.

The next step would then be to verify the oncogene inactivation on genomic and protein expression level and to investigate the phenotypic effects of this inactivation. To procure proof of concept, two already known oncogenes, *BRAF* and *PIK3CA*, and one potential oncogene, *NPM1*, have been inactivated with CRISPR/Cas-system in cancer cells lines.

It has been my motivation to work on this topic to pave out the way for advanced, combined targeted therapies and a more personalized cancer treatment.

2 Material and Methods

2.1 Design of sgRNAs

“Using a complete set of somatic mutations reported by Kandoth *et al*. ([Kandoth et al., 2013]) we designed sgRNAs targeting the mutations employing bedtools ver. 2.17.0 using hg19 as a reference genome assembly. Each mutation site was flanked by 30 bp on both sides. The sequences were then subjected to an sgRNA search and scoring script, an in-house programmed implementation of an algorithm proposed by Doench *et al*. ([Doench et al., 2014; Doench et al., 2016]). An off-target search for obtained 20-bp long sgRNAs was done using an exhaustive short sequence aligner GEM ver. 1.376 ([Marco-Sola et al., 2012]), with parameters set to report all alignments with up to 4 mismatches. All alignments not followed by a PAM sequence (NGG) were filtered out using a trivial Python script and all sgRNAs with a perfect match in a wild-type hg19 genome were discarded from further analysis steps. The 13 mutations analyzed with the traffic light reporter assay were searched for in the Catalogue of somatic mutations in cancer (COSMIC) to cover different mutation types (insertions, deletions, point mutations) and several cancer types. Furthermore the selection was biased towards frequently mutated genes that are common in cancer cells. Another criterion for inclusion was presence of a PAM sequence close to the mutation, so that sgRNAs span the mutation site.“ (Gebler et al., 2017)

2.2 Plasmids

“*Streptococcus pyogenes* Cas9 was expressed from pSpCas9(BB)-2A-Puro (PX459, ([Ran et al., 2013a]) V2.0 from the CAG promoter and sgRNAs were expressed from the human U6 pol III promoter. Protospacers for each target were cloned into PX459 by cloning complementary oligonucleotides into the vector. If necessary an additional G on the 5’-end was added. Oligos (Metabion, Planegg/Steinkirchen, Germany) were annealed, phosphorylated and ligated into PX459 using BbsI sites.

Neisseria meningitidis Cas9 was expressed from pSimpleII-U6-tracr-U6-BsmBI-NLS-NmCas9-HA-NLS(s) (Addgene plasmid # 47868, ([Hou et al., 2013]) from the EF1a promoter and sgRNA from the human U6 pol III promoter. The protospacer for the FLT3-ITD [FMS-related tyrosine kinase 3 internal tandem duplication] target was cloned into pSimpleII-U6-tracr-U6-BsmBI-NLS-NmCas9-HA-NLS(s) with an additional G on the 5’-end using BsmBI sites.

RFP-GFP traffic light reporter plasmids were constructed based on pRG-HBV-S Karimova *et al*. ([Karimova et al., 2015]). Wildtype and oncogenic targets of 13 genes were cloned into

pRG-HBV-S via EcoRI and BamHI sites. In-frame stop codons within target sites were anticipated by changing orientation or shifting target inserts.

For expression of Cas9 and sgRNAs from lentiviral plasmids the pL-CRISPR.EFS.GFP was used (Addgene plasmid # 57818, ([Heckl et al., 2014]). *Streptococcus pyogenes* Cas9 and eGFP are linked via P2A and expressed from the EFS promoter, whereas the sgRNAs were again expressed from the human U6 pol III promoter. Protospacers, either targeting *NPM1* mutation or non-targeting, were cloned into pL-CRISPR.EFS.GFP using BsmBI sites, analog to cloning into PX459.“ (Gebler et al., 2017)

Furthermore, sgRNAs either targeting *BRAF* or *PIK3CA* mutation where cloned into pL-CRISPR.EFS.GFP using BsmBI sites.

For simultaneous expression of sgRNAs targeting *BRAF* and *PIK3CA* mutations, a second sgRNA expression cassette with H1 promoter was cloned into pL-CRISPR.EFS.GFP, using the PrecisionX^{TM} Multiplex gRNA Cloning Kit (System Biosciences, Palo Alto, CA). Therefore, an overhang polymerase chain reaction (PCR) was performed with H1 promoter and RNA scaffold as template. The forward and reverse primer sequences were 5’-AAAGGACGAAACACCGTAGCTACAGAGAAATCTCGAGTTTTAGAGCTAGAAATAGCA AG-3’ and 5’-TTCTAGCTCTAAAACCCATGACGTGCATCATTCATGGATCCAAGGTGTCT CATAC-3’, containing the *BRAF* and *PIK3CA* protospacer sequences, respectively (Metabion). Then, the PCR product was used for insertion into BsmBI digested pL-CRISPR.EFS.GFP plasmid by fusion, following the manufacturer’s instructions.

2.3 Cell culture

“HeLa and HEK293 cells were kept in DMEM (high glucose, GlutaMAX^{TM}, pyruvate, Gibco, Carlsbad, CA) with 10% (v/v) FBS (Gibco) and antibiotics (100 U/mL penicillin, 100 µg/mL Streptomycin, Gibco). For the RFP-GFP traffic light reporter assays HeLa cells were plated at a density of 2.6 x 10^4 cells per well in 24-well dishes. After 24 h cells were co-transfected with 200 ng of plasmid DNA (100 ng of px459-sgRNA and 100 ng of corresponding RFP-GFP-reporter) using Effectene transfection reagent according to manufacturer’s instructions (1.6 µL Enhancer and 2 µL Effectene per well, Qiagen, Hilden, Germany). For the *FLT3-ITD* target and *Neisseria meningitidis* Cas9 the total amount of co-transfected DNA was 250 µg (200 ng of pSimpleII-U6-tracr-U6-BsmBI-NLS-NmCas9-HA-NLS(s)-gRNA and 50 ng of corresponding RFP-GFP-reporter). Cells were analyzed by flow cytometry at 48 h post transfection.

Lentiviral particles were produced in HEK293 cells, plated at a density of 6 x 10^6 cells in a 10 cm cell culture dish. After 24 h cells were transfected with 2 µg of pMD2.G, 6 µg of psPAX2 (Addgene plasmids # 12259 and # 12260) and 10 µg of pL-CRISPR.EFS.GFP-sgRNA using 45 µL PEI (1 mg/ml in 1x PBS, sterile-filtered, Sigma-Aldrich, St. Louis, MO) per dish. Fresh

medium was added 24 h post transfection and viral supernatant were collected 48 h post transfection, passed through 0.22 µm filters (Merck Millipore, Darmstadt, Germany) and centrifuged for 2 h at 50,000 g and 4 ºC. Supernatants were decanted and viral pellets were resuspended in 300 mL PBS at 4 ºC overnight.

OCI-AML3 cells (carrying the c.863_864insTCTG mutation) were kept in αMEM (Biochrom, Berlin, Germany), supplemented with 20% (v/v) heat inactivated FBS (Gibco), antibiotics (100 U/mL penicillin, 100 µg/mL Streptomycin, Gibco) and 2 mM L-Glutamine (Sigma-Aldrich). MV4-11 cells were kept in RPMI 1640 (L-Glutamine, Gibco), supplemented with 10% (v/v) heat inactivated FBS (Gibco), antibiotics (100 U/mL penicillin, 100 µg/mL Streptomycin, Gibco) and 2 mM L-Glutamine (Sigma-Aldrich).

Prior to transduction 24 well plates were coated with RetroNectin® (Takara Clontech, Mountain View, CA) following the manufacturer's protocol. $1x10^6$ of OCI-AML3 and MV4-11 cells were transduced with 500 µL of 117-fold concentrated viral supernatant in the presence of 100 µL protamine sulfate (5 mg/mL, Sigma-Aldrich) in 1 mL total volume and spinoculated for 1 h at 1,000 g and 37 ºC. Transduced OCI-AML3 and MV4-11 were sorted for GFP expression 96 h post transduction using a BD Aria™ system and FACSDiva™ software (Becton Dickinson, Heidelberg, Germany). One fraction of sorted cells was mixed 1:1 with non-transduced cells. Another fraction was expanded for further 96 h and used for immunostaining, deep sequencing, RT-qPCR and cell cycle analysis. To inhibit NHEJ, 5 µM of the DNA ligase IV inhibitor SCR7 (Sigma-Aldrich) was added to the culture media." (Gebler et al., 2017)

A-375, SK-MEL-28, HCT116 and RKO cells were kept in DMEM (high glucose, GlutaMAX™, pyruvate, Gibco) with 10% (v/v) FBS (Gibco) and antibiotics (100 U/mL penicillin, 100 µg/mL Streptomycin, Gibco).

For lentiviral transduction of A-375 and SK-MEL-28, $1.3x10^5$ cells were seeded in 6 well plates and transduced with a multiplicity of infection (MOI) of approximately 0.5 in the presence of 1% polybrene (4 mg/mL, Sigma-Aldrich) and 1% HEPES (1 M, pH 7.25, Carl Roth, Karlsruhe, Germany) in 2.5 mL total volume. Subsequently, cells were spinoculated for 30 minutes at 2,500 rpm and 37 ºC and sorted for GFP expression 6 days post transduction using a BD Aria™ system and FACSDiva™ software (Becton Dickinson). One fraction of the sorted cells were mixed 1:1 with non-transduced cells, another fraction was used for the T7 assay.

$3x10^5$ of HEK293, HCT116 and RKO cells were seeded in 6 well plates and transduced with 25 µL of 117-fold concentrated viral supernatant in the presence of 25 µL protamine sulfate (5 mg/mL, Sigma-Aldrich) in 2.5 mL total volume and spinoculated for 30 minutes at 2,500 rpm and 37º C. Transduced cells were sorted for GFP expression 48 h post transduction using a BD Aria™ system and FACSDiva™ software (Becton Dickinson). One fraction of sorted cells was mixed 1:1 with non-transduced cells. Another fraction was expanded for further 5 days and used for deep sequencing.

2.4 FACS analysis

“For the traffic light reporter assay, HeLa cells were transfected and analyzed for percentage of RFP and GFP using a BD FACSCanto II (Becton Dickinson) and FlowJO software. Assays were performed in [technical] triplicates. OCI-AML3 and MV4-11 cells in 1:1 mixture of transduced and non-transduced cells were analyzed for GFP expression every 48 h by flow cytometry, using a MACS Quant® Analyzer (Miltenyi Biotech, Bergisch Gladbach, Germany) and FlowJO software.“ (Gebler et al., 2017)

A-375, SK-MEL-28, HEK293, HCT116 and RKO cells in 1:1 mixture of transduced and non-transduced cells were analyzed for GFP expression every 48 to 72 h by flow cytometry, using a BD FACSCalibur (Becton Dickinson) and FlowJO software.

2.5 T7 assay

A-375 and SK-MEL-28 cells were collected in PBS prior to genomic DNA extraction using the QIAamp DNA Blood Mini Kit, following the manufacturer’s protocol (Qiagen). The genomic DNA was used as a template for PCR analysis, following the Phusion High-Fidelity DNA Polymerase protocol with GC buffer (New England Biolab, Ipswich, MA), using the PCR program (Supplementary Table 2) with *BRAF* specific primers (Supplementary Table 3).

Subsequently, the PCR products were purified using a GeneJET PCR purification Kit (Thermo Scientific), following the manufacturer’s protocol. 400 ng purified PCR product, 1 µl NEBuffer 2 (New England Biolabs) and water were mixed to a total volume of 10 µl. PCR products were hybridized in a thermocycler with the protocol shown (Supplementary Table 4**)**.

1 µl T7 Endonuclease (New England Biolabs), 1 µL NEBuffer 2 and 8 µL water were added to the reaction and incubated at 37 ˚C for 20 min. Subsequently, 10 µL of the reaction mix was mixed with 2 µL of 6x NEB orange loading dye (New England Biolabs), loaded on a 2% agarose gel and run for 30 min at 90 V.

Supplementary Table 2: PCR program for the T7 assay of *BRAF* locus.

	Temperature	**Time**	**Cycles**
Lid	98°C	∞	
Denaturation	98°C	30 seconds	1x
Denaturation	98°C	10 seconds	35x
Annealing	66°C	30 seconds	
Elongation	72°C	30 seconds	
Elongation	72°C	8 minutes	1x
Storage	8°C	∞	1x

Supplementary Table 3: Primer list for the T7 assay of *BRAF* locus.

Primer	Sequence
***BRAF* - Forward**	5'-CCTTCAATGACTTTCTAGTAACTCAGC-3'
***BRAF* - Reverse**	5'-CCAATGAAGAGCCTTTACTGCTC-3'

Supplementary Table 4: Hybridization program for the T7 assay of *BRAF* locus.

	Temperature	Time
Lid	110°C	∞
Denaturation	95°C	5 minutes
Ramp	95-85°C	-2°C/second
Ramp	85-25°C	-0.1°C/second
Storage	4°C	∞

2.6 Cell cycle analysis

"OCI-AML3 and MV4-11 cells were collected at day 8 post transduction and fixed with ethanol. Cells were washed with PBS and resuspended in PI staining solution (0.1 % Triton X-100, Serva, 10 µg/mL PI, Molecular Probes, Carlsbad, CA, 100 µg/mL DNase-free RNase A, Invitrogen, Carlsbad, CA, in 1x PBS). FACS analysis was performed using a BD FACSCalibur (Becton Dickinson) and FlowJO software." (Gebler et al., 2017)

2.7 Immunostaining

"Microscopy slides were coated with poly-L-lysine P8920 (Sigma-Aldrich), following the manufacturer's protocol. OCI-AML3 and MV4-11 cells were collected in PBS at day 8 post transduction and were spun on poly-L-lysine coated coverslips at room temperature. Afterwards cells were fixed with 4 % PFA and again centrifuged at 800 rpm for 8 min. Slides were washed twice with 30 mM glycine (Sigma-Aldrich) in PBS. Cells were then permeabilized with 0.3% Triton X-100 (Sigma-Aldrich) in PBS for 5-10 min at room temperature, followed by two washes with 30mM glycine in 1x PBS. Subsequently cells were blocked in blocking solution (1x PBS, 0.2% fish skin gelatin Sigma-Aldrich) for 15-30 min at room temperature. Primary antibodies were diluted 1:150 in blocking solution and incubated over night at 4 ºC (mouse anti-NPM1, Abcam, Cambridge, UK). Cells were washed 3 times for 10 min each in blocking solution followed by 1 h incubation at room temperature with secondary antibodies in the presence of DAPI (1:1000, 1 µg/ml, Sigma-Aldrich), diluted in blocking solution. As secondary antibodies fluorescently labeled donkey anti mouse-IgG antibodies (1:500, Life Technologies, Carlsbad, CA, AlexaFluor ® 594) were used. After 3 final washes for 5 min each

in blocking solution slides were mounted using ProLong® Gold Antifade Mountant (Thermo-Fisher, Carlsbad, CA). Images were acquired using a Delta Vision Elite microscopy system (GE Healthcare, Little Chalfont, UK) and finally processed in Fiji.“ (Gebler et al., 2017)

2.8 Apoptosis assay

“RKO cells were transfected with plasmids encoding *S. Pyogenes* Cas9 and a [sgRNA] for the mutated gene *BRAF* (1799 T>A) or a mock control plasmid (empty) using Effectene (Qiagen) following the manufacturer’s instructions. Cells were harvested 24 hours post-transfection and the number of apoptotic cells was determined by measuring phosphotidylserine exposure utilizing an Annexin V-FITC kit (Miltenyi Biotec) according to manufacturer’s instructions. The percentage of Annexin V positive and propidium-iodide negative cells was normalized to the percentage of mock transfected control cells.“ (Gebler et al., 2017)

2.9 Quantification of mutant *NPM1* transcripts

“RNA-expression of mutant *NPM1* was measured as described ([Shayegi et al., 2013]). Briefly, total RNA was extracted, reverse transcribed and subjected to quantitative PCR using primers specific for *NPM1*-mutation A and a generic probe. Quantitation was performed using a dilution of cloned mutant fragments. Results are expressed as %mutant NPM1-molecules compared to the ABL1-reference gene.“ (Gebler et al., 2017)

2.10 Deep sequencing

“The [OCI-AML3] cells were collected in PBS prior to genomic DNA extraction using the QIAamp DNA Blood Mini Kit, following the manufacturer’s protocol (Qiagen). The genomic DNA was used as a template for PCR analysis, following the NEB Phusion Polymerase Q5 protocol using the PCR program depicted below employing *NPM1* specific primers [(Table 1,Table 2)]. Subsequently, the PCR products were purified using a GeneJET PCR purification Kit (Thermo Scientific), following the manufacturer’s protocol.

100 ng of each PCR amplicon were applied for end repair, following the Ion Plus Fragment Kit (Life Technologies). The end-repaired PCR amplicons were purified using the Agencourt AMPure Kit (Beckman Coulter). Next, adapters and barcodes were ligated to the amplicons, followed by a nick repair using the Ion Plus Fragment Library and Ion Xpress™ Barcode Adapters Kit (Life Technologies). The adapter ligated and nick translated DNA was purified using the Agencourt AMPure Kit (Beckman Coulter, Brea, CA). Samples were pooled into a 1.5 ml Eppendorf LoBind tube and amplified following the Ion Plus Fragment Library Kit (Life Technologies). The Ion Library Quantification Kit (Life Technologies) was used to quantify the library. Subsequently, template-positive ion PGM template OT2 200 ion sphere particles were

established and enriched, following the Ion PGM Template OT2 Kit (Life Technologies). Samples were sequenced with the Ion PGM Sequencing 200 Kit v2 (Life Technologies) on an Ion 318 v2 Chip using the Ion Personal Genome Machine (PGM) system following the manufacturer's protocol.

To limit the analysis to sequencing reads fully enclosing the *NPM1* mutation site, only reads containing the following flanking sequences were considered: GTGTTGTGGTTCCTTAACCA and ATGCAGAGTGAGAACTTTCC. This in-silico amplification procedure was performed with an aid of PatMaN sequence aligner ver. 1.2.2, allowing up to 2 mismatches in each flanking sequence. Parts of reads enclosed by the flanks were aligned to the reference amplicon using needleall tool from EMBOSS package, ver. 6.6.0. Analysis of the reported multiple pairwise alignments was done with a Python script developed to detect insertions and deletions spanning the mutation site. Visualization of the results was done with R ver. 3.2.1, using a ggplot2 package."(Gebler et al., 2017)

For deep sequencing of HEK293, HCT116 and RKO cells, cells were predominantly processed in the same way as OCI-AML3 cells. *BRAF* and *PIK3CA* specific primers were employed as listed in Supplementary Table 5, the PCR program is shown in Supplementary Table 6.

Only reads including the flanking sequences for *BRAF* and *PIK3CA* were considered: 5'-AGCATCTCAGGGCCAAAAAT-3', 5'-CATTTTCCTATCAGAGCAAGCA-3' (*BRAF*), and 5'-ATGATGCTTGGCTCTGGAAT-3', 5'-CAGCATGCATTGAACTGAAA-3' (*PIK3CA*).

Table 1: Primer List [for the deep sequencing of *NPM1* and *DNMT3A* loci.]

The primer design was performed using the Primer3 online tool (Version 4.0.0.)

Primer	Sequence
***NPM1* - Forward**	5'-TGTCTATGAAGTGTTGTGGTTCC-3'
***NPM1* - Reverse**	5'-AACACGGTAGGGAAAGTTCTCA-3'
***DNMT3A* – Forward**	5'-CCATGTCCCTTACACACACG-3'
***DNMT3A* – Reverse**	5'-TCCTGCTGTGTGGTTAGACG-3'

Table 2: PCR program [for the deep sequencing of *NPM1* and *DNMT3A* loci.]

	Temperature	Time	Cycles
Lid	98°C	∞	
Denaturation	98°C	3 minutes	1x
Denaturation	98°C	30 seconds	35x
Annealing	66°C	30 seconds	
Elongation	72°C	30 seconds	
Elongation	72°C	8 minutes	1x
Storage	8°C	∞	1x

Supplementary Table 5: Primer List for the deep sequencing of *BRAF* and *PIK3CA* loci.

The primer design was performed using the Primer3 online tool (Version 4.0.0.)

Primer	Sequence
***BRAF* - Forward**	5'-AGCATCTCAGGGCCAAAAAT-3'
***BRAF* - Reverse**	5'-TGCTTGCTCTGATAGGAAAATG-3'
***PIK3CA* – Forward**	5'-ATGATGCTTGGCTCTGGAAT-3'
***PIK3CA* – Reverse**	5'-TTTCAGTTCAATGCATGCTG-3'

Supplementary Table 6: PCR program for the deep sequencing of *BRAF* and *PIK3CA* loci.

	Temperature	Time	Cycles
Lid	98°C	∞	
Denaturation	98°C	30 seconds	1x
Denaturation	98°C	10 seconds	40x
Annealing	63°C	30 seconds	
Elongation	72°C	15 seconds	
Elongation	72°C	5 minutes	1x
Storage	8°C	∞	1x

2.11 Statistical Analysis

“To assess biological statistical significance of differences between observed cleavages of sgRNAs targeting wild type or mutated sequences in the "traffic light" reporter assay, we performed double-sided Student's t-tests without an assumption of equal variances (Welch's t-test). Correlations between sgRNA activity predictions and measured effects were calculated using Pearson's method. A P-value of less than 0.05 was considered statistically significant.“ (Gebler et al., 2017)

3 Results

3.1 Design of sgRNAs targeting oncogenes

“To first investigate how many cancer mutations could theoretically be targeted by *Streptococcus pyogenes* (sp)Cas9, we performed a comprehensive bioinformatics analysis of published cancer mutations ([Kandoth et al., 2013]). From the reported 608 671 unique mutations, we were able to design 1 909 172 sgRNAs that cover 554 069 mutations (91.0%) and 20 756 out of 20 948 mutated genes. We then performed an analysis to avoid off-target cleavage and discarded all sgRNAs having additional perfect matches to sequences in the reference genome, in addition to prioritizing sgRNAs with the highest divergence to homologous sequences elsewhere in the genome ([Hsu et al., 2013]). Based on these criteria, we nominated 1 701 813 sgRNAs that could theoretically target 535 327 (88.0%) of known cancer mutations encompassing 10 349 (85.0%) of known cancer driver mutations ([Table 3, Figure 1A]).” (Gebler et al., 2017)

Table 3: Design of cancer mutation targeting sgRNAs.

[Table 3 is available on the compact disc and at the following website:]

http://www.buchholzlab.org/index.php/supplementary-data/

Abbreviations used in the table

Gene	HGNC symbol of a gene affected by the mutation
Chromosome Start Position	Genomic coordinates of the mutation (GRCh37/hg19 genome version)
Variant Classification	Classification of a mutation according to Kandoth et al.
Reference Allele	Refeference genome sequence
Variant Allele	Mutated sequence
Transcript ID	Ensembl ID of an affected transcript (database version 69)
AA Change	Amino acid change
Frequently mutated gene (FMG)	"Yes" if the gene belongs to a list of frequently mutated genes
FMG & More than 1 sample mutated	"Yes" if the mutation was also identified in more than one sample
# Tumor Samples	Count of samples with the mutation
Tumor Samples	List of samples with the mutation (from Kandoth et al.)
Designed sgRNAs	Count of designed sgRNAs
[Doench] Score ≥ .75	sgRNAs with a score (according to [Doench] et al.) above 0.75
[Doench] Score ≥ .5	sgRNAs with a score (according to [Doench] et al.) between 0.5 and 0.75
[Doench] Score ≥ .25	sgRNAs with a score (according to [Doench] et al.) between 0.25 and 0.5
[Doench] Score < .25	sgRNAs with a score (according to [Doench] et al.) lower than 0.25

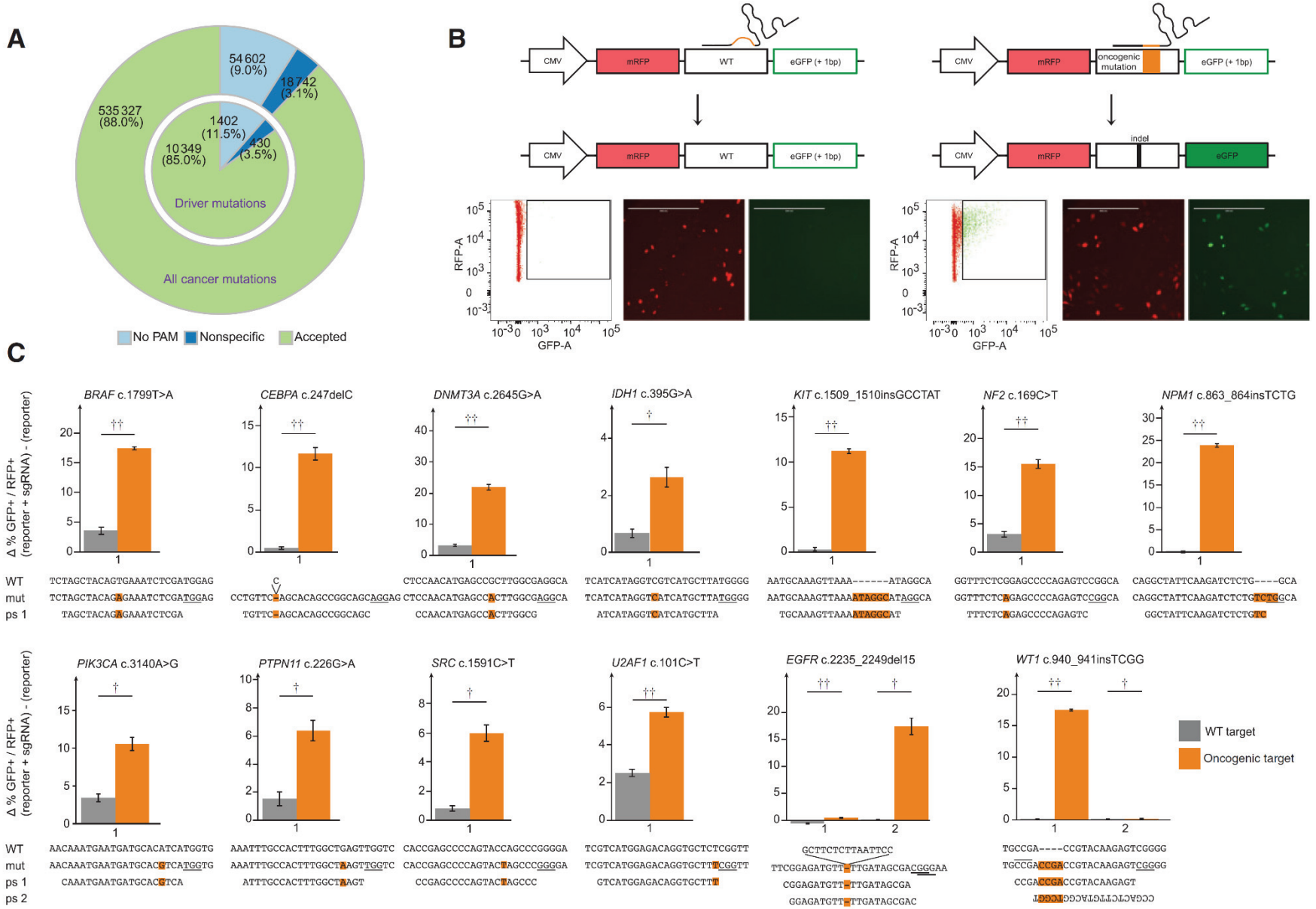

A
54 602 (9.0%)
18 742 (3.1%)
535 327 (88.0%)
1402 (11.5%)
430 (3.5%)
10 349 (85.0%)
Driver mutations
All cancer mutations
No PAM
Nonspecific
Accepted
B
CMV
mRFP
WT
eGFP (+ 1bp)
oncogenic mutation
indel
eGFP
RFP-A
GFP-A
C
Δ % GFP+ / RFP+ (reporter + sgRNA) - (reporter)
BRAF c.1799T>A
CEBPA c.247delC
DNMT3A c.2645G>A
IDH1 c.395G>A
KIT c.1509_1510insGCCTAT
NF2 c.169C>T
NPM1 c.863_864insTCTG
PIK3CA c.3140A>G
PTPN11 c.226G>A
SRC c.1591C>T
U2AF1 c.101C>T
EGFR c.2235_2249del15
WT1 c.940_941insTCGG
WT target
Oncogenic target
WT
mut
ps 1
ps 2

Figure 1: sgRNA design and evaluation of sgRNA efficacy and selectivity

A) Bioinformatics analysis and sgRNA design for cancer mutations. A pie chart for reported cancer mutations and for cancer driver mutations for *Streptococcus pyogenes* sgRNAs is shown. B) Overview of "traffic light" reporter assay. Important elements are indicated. Representative examples of fluorescence-activated cell sorting plots and microscopy images are shown (scale bars = 400 mm). C) Activity and selectivity of employed sgRNAs. The targeted mutations are indicated above each graph, with the wild-type, mutant, and protospacer sequences illustrated below each graph. Error bars represent [standard deviation (SD)] from experiments performed in [technical] triplicates. Two-sided Student's t test *P<.05; **P<.01; ***P<.001. mut – mutant; ps – protospacer; WT – wild-type.

[CEBPA – CCAAT/Enhancer-Binding Protein, Alpha; DNMT3A – DNA Methyltransferase 3A; IDH1 – Isocitrate dehydrogenase 1; NF2 – Neurofibromin 2; PTPN11 – Protein-Tyrosine Phosphatase, Nonreceptor-Type, 11; SRC – V-Src Avian Sarcoma (Schmidt-Ruppin A-2) Viral Oncogene; U2AF1 – U2 Small Nuclear RNA Auxiliary Factor 2; WT1 – Wilms tumor gene]

3.2 Evaluation of sgRNA efficacy and selectivity

“We next established a “traffic-light” reporter system ([Karimova et al., 2015]), where Cas9 cleavage activates GFP expression in transiently transfected mammalian cells in culture to rapidly evaluate efficacy and selectivity of designed sgRNAs (Figure 1B). Sequences bearing 13 different cancer mutations or the corresponding WT sequences were cloned into the reporter construct and subsequently cotransfected into HeLa cells with a Cas9 expression plasmid ([Ran et al., 2013a]) that also expressed the cancer mutation–specific sgRNA (Figure 1, B and C). Efficient cleavage was observed for most constructs bearing cancer mutations, with 10 out of 13 sgRNAs also showing a higher than 4-fold target site selectivity over the wild-type (WT) sequence, with the remaining three still showing 2.7- to 3.8-fold selectivity of mutant over WT. In particular, insertion and deletion mutations reported in the genes *KIT*, *NPM1*, *CEBPA*, *EGFR*, and *WT1* showed little to no appearance of green cells when combined with the WT reporters (Figure 1C), reflecting that the WT sequences were not cleaved efficiently. In contrast, 10% to 25% of GFP-positive cells were detected when the cancer mutation reporters were used in combination with matching sgRNAs. Hence, these sgRNAs created indels in the reporter plasmids that brought the GFP sequence into the correct reading frame, demonstrating their potency to cleave the cancer mutation sequence. Overall, we observed a descent correlation between the sgRNA prediction score and the actual activity in the traffic light reporter assay. However, we detected considerable differences in cleavage efficacy for some sgRNAs targeting the identical cancer mutation, despite the fact that their prediction scores ([Doench et al., 2014; Doench et al., 2016]) were similar ([Table 4]). For instance, sgRNA#1 with a score of 0.42 for the *EGFR* 2235_2249del15 mutation only produced 0.5% (+/-0.1%) of GFP-positive cells, whereas the related sgRNA#2 with a score of 0.33 that is only shifted by one base pair was highly efficient and resulted in more than 17.4% (+/-1.5%) of GFP-positive cells. […]

Remarkably, many point mutations, such as the *DNMT3A* c.2645G>A mutation, were efficiently cleaved by the cancer mutation sgRNA (21.9% [+/-0.8%] GFP-positive cells) without appreciably cleaving the WT sequence (3% [+/-0.2%] GFP-positive cells), demonstrating that the CRISPR/Cas9 system can be sensitive enough to distinguish single base pair alterations.” (Gebler et al., 2017)

As a small fraction of cancer mutations lack neighboring NGG-PAMs, we tested the applicability of the traffic-light reporter for orthogonal CRISPR/Cas-systems. A gRNA targeting *FLT3-ITD* was cloned into a NmCas9 expressing plasmid and co-transfected with the reporter construct into HeLa cells. The measured activity revealed 2.3% GFP-positive cells and was lower than the activity of any tested SpCas9 sgRNA. The specificity was comparable to SpCas9 with no GFP-positive cells detectable (Figure 3).

Table 4: Experimentally tested sgRNAs and their prediction scores*

#	CDS mutation	AA mutation	sgRNA	Doench score (2014)	Doench score (2016)	Measured activity	Genomic sequence ([wt/mut])
1	c.1799T>A	V600E	TAGCTACAGAGAAATCTCGA	0.23	0,67	0,1734	CTGATGGGACCCACTCCATCGAGATTTC[A/T]CTGTAGCTAGACCAAAATCACCTATTTTTAC
2	c.247delC	p.Q83fs*77	TGTTCAGCACAGCCGGCAGC	0,13	0,49	0,1166	CACGGCCGCCTTGGCCTTCTCCTGCTGCCGGCTGTGCT[G/-]GAACAGGTCGGCCAGGAACTCGTCGTTGAAGGCGGCCGGGTCGATGTAGGCGCTGATG
3	c.2645G>A	R882H	CCAACATGAGCCACTTGGCG	0,53	0,64	0,2190	ATTTGGTTTCCCAGTCCACTATACTGACGTCTCCAACATGAGCC[G/A]CTTGGCGAGGCAGAGACTGCTGGGCCGGTCATGGAGCGTGCCAGTCATCCGCCACC
4	p.E746_A750delELREA	p.E746_750del	CGGAGATGTTTTGATAGCGA	0,42	0,57	0,0045	GATCCCAGAAGGTGAGAAAGTTAAAATTCCCGTCGCTATCAA[GGAATTAAGAGAAGC/-]AACATCTCCGAAAGCCAACAAGGAAATCCTCGATGTGAGTTTCTGCTTTG
5	p.E746_A750delELREA	p.E746_750del	GGAGATGTTTTGATAGCGAC	0,33	0,52	0,1742	GATCCCAGAAGGTGAGAAAGTTAAAATTCCCGTCGCTATCAA[GGAATTAAGAGAAGC/-]AACATCTCCGAAAGCCAACAAGGAAATCCTCGATGTGAGTTTCTGCTTTG
6	c.395G>A	p.R132H	ATCATAGGTCATCATGCTTA	0,03	0,21	0,0271	TATTGCCAACATGACTTACTTGATCCCCATAAGCATGA[C/T]GACCTATGATGATAGGTTTTACCCATCCACTCACAAGCCGGGGGATATTTTTGCAGATAA
7	c.1509_1510insGCCTAT	p.Y503_F504insAY	TGCAAAGTTAAAATAGGCAT	0,33	0,59	0,1120	AAGGCTTACAACGATGTGGGCAAGACTTCTGCCTAT[-/GCCTAT]TTTAACTTTGCATTTAAAGGTAACA
8	c.169C>T	p.R57*	TTTCTCAGAGCCCCAGAGTC	0,12	0,47	0,1555	GGACCTCTTTGATTTGGTGTGCCGGACTCTGGGGCTC[C/T]GAGAAACCTGGTTCTTTGGACTGCAGTACACAAT

9	c.863_864insTCTG	p.W288fs*12	GGCTATTCAAGATCTCTGTC	0,22	0,51	0,2387	ATTTCTTTTTTTTTTTTTCCAGGCTATTCAAGATCTCTG[-/TCTG]GCAGTGGAGGAAGTCTCTTTAAGAAAATAGTTT
10	c.3140A>G	p.H1047R	CAAATGAATGATGCACGTCA	0,06	0,67	0,1058	TTCATGAAACAAATGAATGATGCAC[A/G]TCATGGTGGCTGGACAACAAAAATGGA
11	c.226G>A	p.E76K	ATTTGCCACTTTGGCTAAGT	0,04	0,41	0,0638	TACTATGACCTGTATGGAGGGGAGAAATTTGCCACTTTGGCT[G/A]AGTTGGTCCAGTATTACATGGAACATCACGGGCAATTAAAAGAGAA
12	c.1591C>T	p.Q531*	CGAGCCCCAGTACTAGCCCG	0,75	0,67	0,0599	CTTCCTGGAGGACTACTTCACGTCCACCGAGCCCCAGTAC[C/T]AGCCCGGGGAGAACCTCTAGGCACAGGCGGGCCCAGACCGGCTTCTCGGCTTGG
13	c.101C>T	p.S34F	GTCATGGAGACAGGTGCTTT	0,23	0,43	0,0575	CAAACAAACCTGGCTAAACGTCGGTTTATTGTGCAACCGA[G/A]AGCACCTGTCTCCATGACGACATGCTCCAATTTTGAAATAAAATGAACAGTTGA
14	c.940_941insTCGG	p.A314fs*4	CCGACCGACCGTACAAGAGT	0,14	0,61	0,1748	GCACACATGAAGGGGCGTTTCTCACTGGTCTCAGATG[-/CCGA]CCGACCGTACAAGAGTCGGGGCTACTCCAGGCACACGTCGCACATCCTGCAGGCAGAGAG
15	c.940_941insTCGG	p.A314fs*4	CCGACTCTTGTACGGTCGGT	0,31	0,60	0,0017	GCACACATGAAGGGGCGTTTCTCACTGGTCTCAGATG[-/CCGA]CCGACCGTACAAGAGTCGGGGCTACTCCAGGCACACGTCGCACATCCTGCAGGCAGAGAG

*CDS – coding DNA sequence – AA, Amino acid

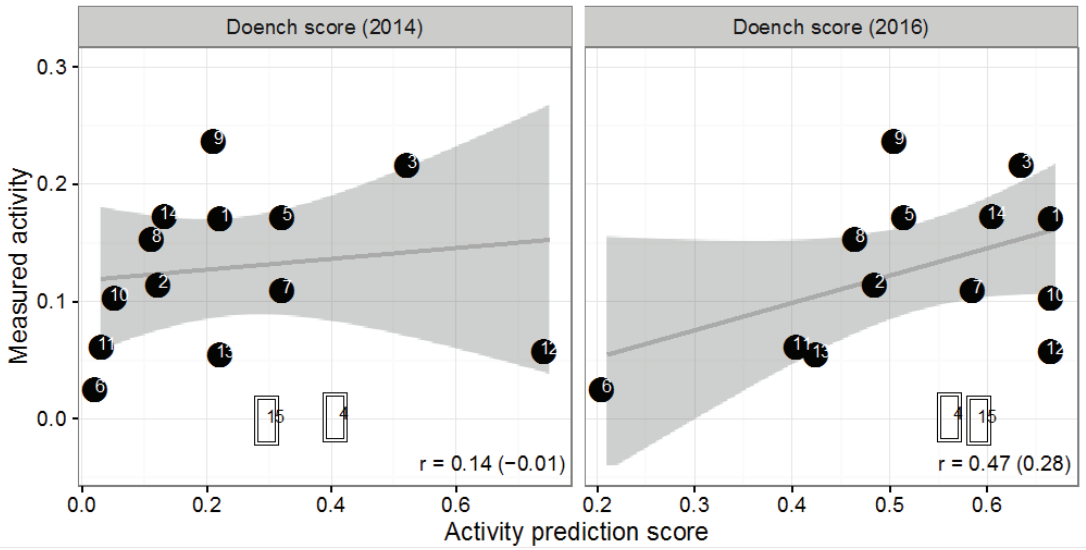


Figure 2: Evaluation of sgRNA prediction score.

Both scatter plots show relationships between predicted (according to Doench 2014 and Doench 2016 scoring systems) and measured activities of sgRNAs (depicted as circles with numbers corresponding to identifiers from [Table 4]). Linear regression lines (dark gray) and surrounding 95% confidence interval areas (light gray) represent correlations between predicted and measured activities of selected sgRNAs (shown as black-filled circles). White-filled circles represent sgRNAs excluded from the regression analysis due to minimal measured activity (sgRNAs: #4 and #15). Corresponding Pearson's correlation coefficients are indicated below, with numbers in brackets showing r-values based on all sgRNAs (black- and white-filled circles).

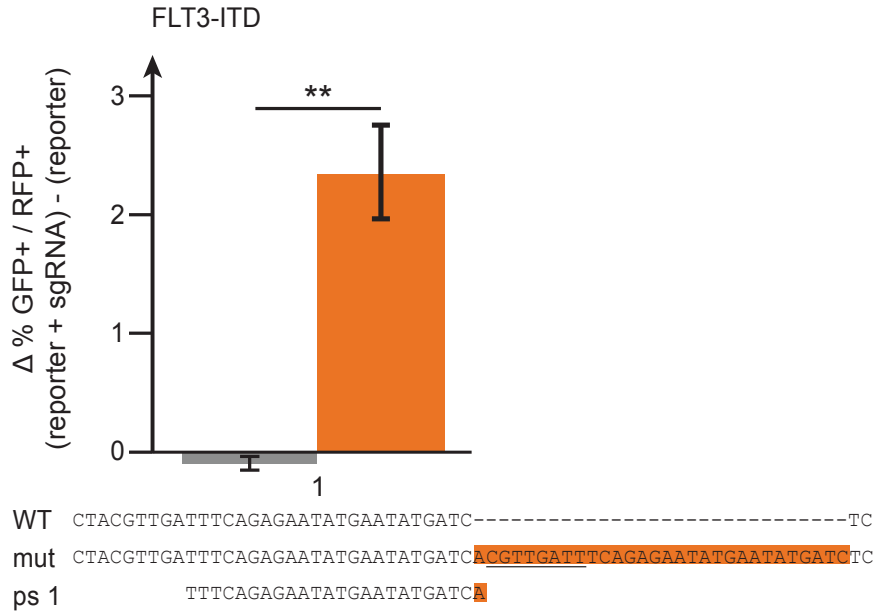


Figure 3: Inactivation of cancer mutation with an orthogonal CRISPR/Cas-system.

Efficiency and selectivity utilizing the *Neisseria meningitidis* CRISPR/Cas-system on the *FLT3-ITD* mutation is shown. The wild type (WT), mutant (mut) and protospacer (ps) sequences are illustrated. The PAM sequence is underlined. [Error bars represent SD from experiments performed in technical triplicates.] Two-sided Student's t-test, **P = 0.008.

3.3 Effects of oncogene knock-out in cancer cell lines

3.3.1 Targeting *NPM1* in AML cells

"We next investigated the functional relevance of [three] common cancer mutations in tumor cells. The nucleophosmin gene (*NPM1*) is mutated in about 30% of patients suffering from acute myeloid leukemia (AML) ([Falini et al., 2005]). Mutant *NPM1* is thought to play an important role in AML proliferation, indicating that a direct way to inactivate the mutation could affect malignant growth ([Federici and Falini, 2013]).

We cloned the tested sgRNA sequence targeting mutant *NPM1* (Figure 1C) into a lentiviral vector ([Heckl et al., 2014]) expressing Cas9 in conjunction with EGFP and transduced *NPM1* mutant OCI-AML3 cells and *NPM1* WT MV4-11 cells with the virus. Efficient cleavage of mutant *NPM1* in OCI-AML3 cells was evident in employing multiple assays ([Figure 4]). Strikingly, transduced OCI-AML3, but not the MV4-11 cells, were successively depleted over time ([Figure 4C]), signifying that the mutant NPM1 protein is required for efficient cell proliferation in OCI-AML3. Cell cycle analyses revealed that OCI-AML3 cells treated with the *NPM1* sgRNA arrested in G1 without markedly altering the subG1 fraction ([Figure 4D]), suggesting that mutant *NPM1* expression in these cells is required for cell cycle progression. To investigate the mutational spectrum at the site of cleavage, we performed deep sequencing of the *NPM1* locus in control- and sgNPM1-treated cells. As expected, cells treated with a control sgRNA revealed a 50:50 ratio for the WT and mutant allele, reflecting the heterozygous nature of the *NPM1* mutation. In contrast, cells treated with the sgRNA-targeting mutant *NPM1* showed efficient cleavage and repair of the mutant allele. Remarkably, the WT:*NPM1* ratio in this sample increased to around 70:30, indicating that a substantial fraction (34.2%) of the cells had repaired the mutation back to the WT sequence ([Figure 4E]), likely through homologous recombination utilizing the WT allele as a template. The fraction of indel mutations was further reduced (from 62.8% to 55.7%) when the cells were treated with the DNA ligase IV inhibitor SCR7 ([Srivastava et al., 2012]), indicating that enhanced HR-mediated repair can be achieved when the NHEJ pathway is inhibited." (Gebler et al., 2017)

Consequently, the mutant allele was almost completely inactivated, either by indel mutations (62.8%) leading to frameshifts and inactivation, or by WT recovery (34.2%).

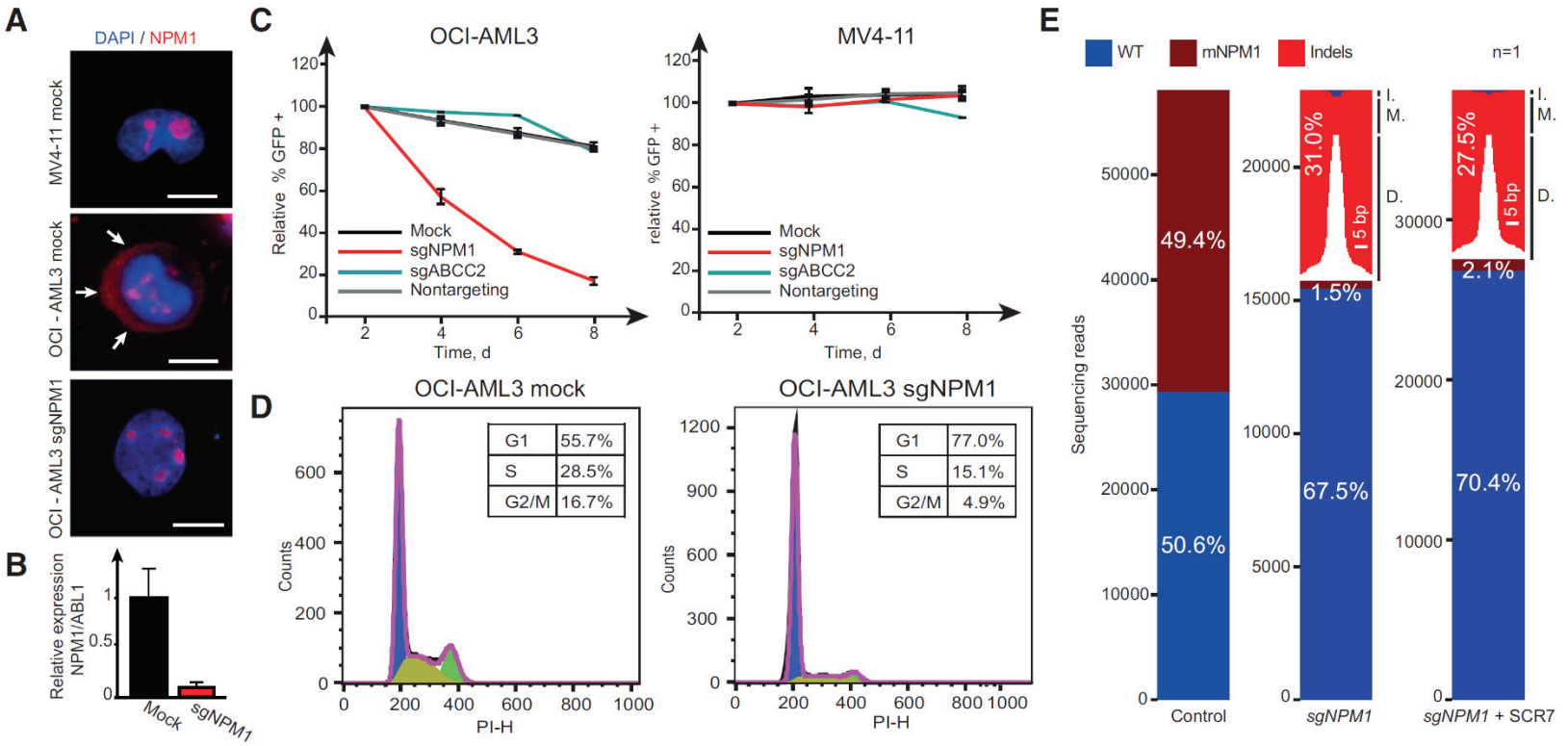


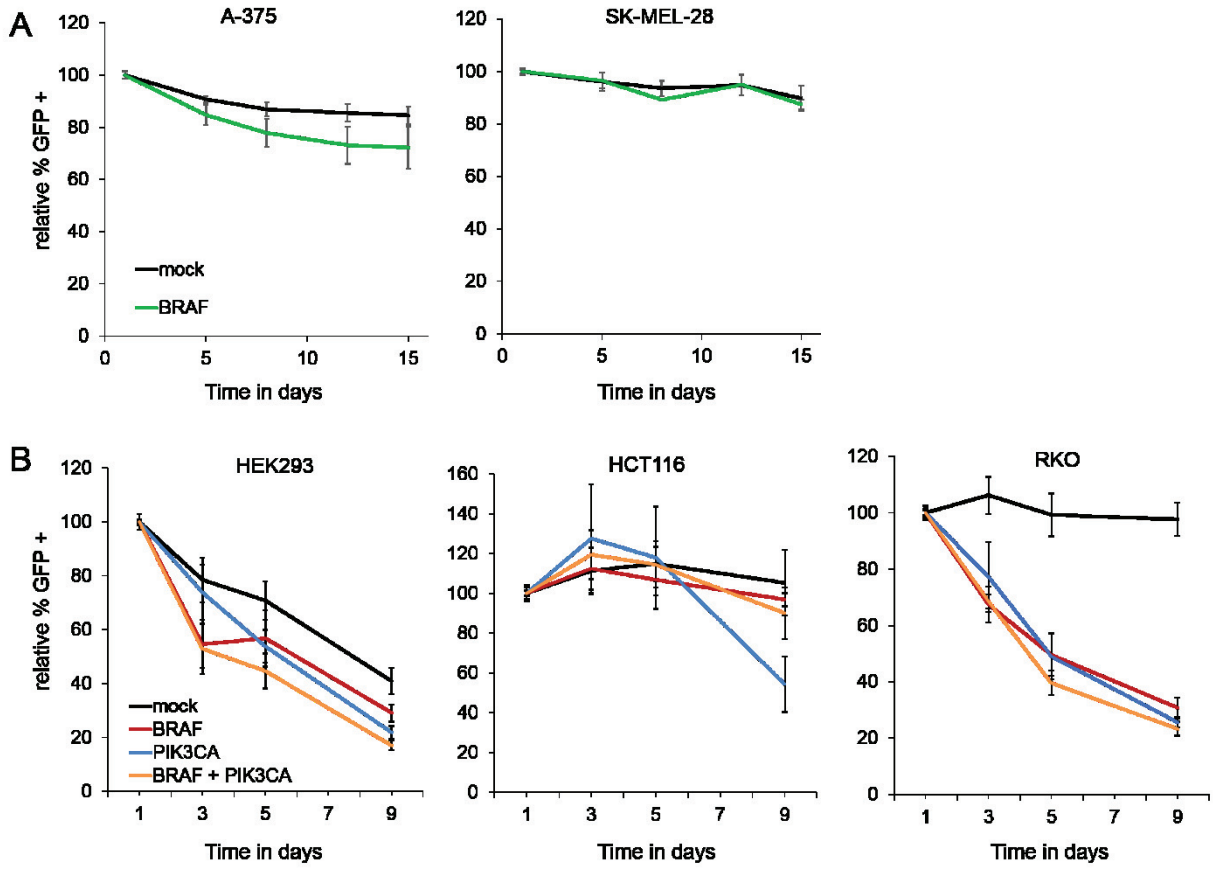
Figure 4: Effects of mutant *NPM1* inactivation.

A) Localization of NPM1 in MV4-11 and OCI-AML3 cells under indicated treatment conditions (scale bars = 10 µm). Arrows highlight the cytoplasmatic localization of mutant NPM1 in mock-treated OCI-AML3 cells. OCI-AML3 sgNPM1 images were taken 96 hours after sgRNA treatment. B) Relative mRNA expression level of mutated NPM1 after sgRNA treatment. mRNA was extracted 96 hours after sgRNA treatment. C) Relative abundance of cells treated with indicated sgRNAs in MV4-11 and OCI-AML3 over time. Error bars show SD from experiments performed in [technical] triplicates. D) Cell-cycle profile of OCI-AML3 cells after indicated treatments. Fluorescence-activated cell sorting analyses were performed eight days after sgRNA treatment. E) Graphical representation of *NPM1* sequencing reads under indicated conditions (one

biological replicate each). Size of deletions and insertions are indicated in white and [dark blue], respectively. sgNPM1–OCI-AML3 cells treated with sgRNA-targeting mutant *NPM1* (mNPM1). sgNPM1 + SCR7–OCI-AML3 cells treated with sgRNA-targeting mutant *NPM1* in the presence of the ligase IV inhibitor SCR7. Sequencing was performed on genomic DNA isolated eight days after sgRNA treatment. Sixty-four point six percent of the indels resulted in reading frame shifts. D. – deletions; I. – insertions; M. – mutations; WT – wild-type.

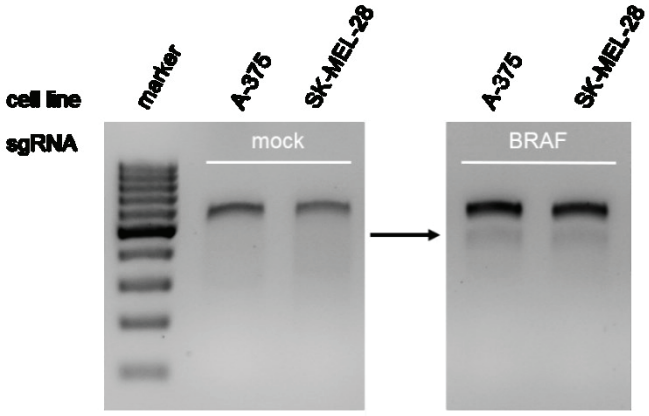
3.3.2 Targeting *BRAF* in melanoma cells

Single *BRAF* inactivation with Vemurafenib showed an anti-proliferative effect in melanoma cell lines (Lee et al., 2010). Therefore, we used the lentivirus, expressing the *BRAF* targeting sgRNA, to transduce the homozygous mutated A-375 and SK-MEL-28 cell lines. Even after 15 days of follow-up, GFP positive cells expressing the sgRNA did not diminish in comparison to mock treatment (Supplementary Figure 2A), although indels could be detected in the T7 assay (Supplementary Figure 3).



Supplementary Figure 2: Effects of mutant *BRAF* and *PIK3CA* inactivation on relative abundance of cells.

(A) Melanoma cell lines A-375 and SK-MEL-28 cells (both $BRAF^{V600E}$ homozygous), were transduced with lentiviruses containing SpCas9 and either a sgRNA targeting mutant *BRAF* or a mock sgRNA. GFP expressing cell fraction was measured by FACS analysis over time.
(B) HCT116 ($BRAF^{WT}$ and $PIK3CA^{H1047R/WT}$), HEK293 ($BRAF^{WT}$ and $PIK3CA^{WT}$) and RKO cells ($BRAF^{V600E/V600E/WT}$ and $PIK3CA^{H1047R/WT}$) were transduced with lentiviruses containing Cas9 and either sgRNA targeting mutant *BRAF*, *PIK3CA*, both or a mock sgRNA. GFP expressing cell fraction was measured by FACS analysis over time. Error bars show SD from experiments performed in technical triplicates.



Supplementary Figure 3: T7 assay of *BRAF* locus in melanoma.
A-375 and SK-MEL-28 cells were transduced with lentivirus containing either a mock sgRNA or a sgRNA targeting *BRAF*^{V600E}. The arrow points to the cleavage product.

3.3.3 Targeting *BRAF* and *PIK3CA* in colorectal carcinoma cells

As in the melanoma cell line the *BRAF*^{V600E} knock-out did not show anti-proliferative effects, we infected the heterozygous mutated RKO cell line, originated from CRC, with the lentivirus containing *BRAF*^{V600E} sgRNA. The sgRNA expressing cell populations were depleted over time and apoptosis increased 1.4-fold (Figure 5, B and C).
Previous investigations emphasized a concomitant inhibition of B-Raf and PI3K/mTOR as more efficient to reduce tumor size (Coffee et al., 2013). To achieve a high *PIK3CA*^{H1047R} inactivation, we tested two different sgRNAs targeting *PIK3CA*^{H1047R} and chose sgRNA#2 to target the mutation in CRC cells, albeit we observed a lower target specificity than for sgRNA#1 (Supplementary Figure 4). Next, we transduced RKO, HCT116 and HEK293 cells with lentiviruses harboring either *BRAF* or *PIK3CA* sgRNA or both sgRNAs. Toxicity was not observed in any of the cell lines after transduction. In double WT HEK293 cells, the lentivirus expressing cell fractions were depleted over time with no difference between mock and sgRNA lentiviruses. In *PIK3CA*^{H1047R} heterozygous HCT116 cells, only the fraction with stable sgRNA expression targeting mutant *PIK3CA*, was reduced to 54.4% after 9 days. In double mutant RKO cells, the mock lentivirus cell population remained constant, whereas all sgRNA expressing cell populations were depleted over time. However, no difference could be detected between single sgRNA and double sgRNA expressing population (Supplementary Figure 2B). The cleavage of *PIK3CA* and *BRAF* in HEK293, HCT116 and RKO cells was further analyzed on genome level by deep sequencing. Confirming the results of the RFP-GFP-reporter system, the sgRNA targeting *BRAF*^{V600E} induced very little DSB in *BRAF* WT HEK293 and HCT116 cells. RKO cells harbor a triplication of *BRAF* (Kleivi et al., 2004). Deep sequencing results showed a WT to mutated allele ratio of 32:68 without sgRNA treatment, what concurs with the

triplication. With single *BRAF* sgRNA expression, the ratio shifted to 51:49. 14.8% of the mutated allele was altered by indels, 27.0% was replaced by the WT allele. Confirming the RFP-GFP-reporter results of sgRNA#2 targeting *PIK3CA*^{H1047R}, 7.8% DSB occurred in *PIK3CA* WT HEK293 cells. The ratio of *PIK3CA* WT to mutated allele was 50:50 in HCT116 and RKO samples, confirming the heterozygosity of the mutation. In RKO cells, 48.6% of the mutated allele was altered by indels after sgRNA treatment, and 31.1% in HCT116 cells. After single *PIK3CA* sgRNA treatment, the *PIK3CA* WT allele fraction increased slightly and replaced 10.2% and 12.2% of the mutated allele in RKO and HCT116 cells, respectively, thus also indicating homologous donor repair with the WT allele as template. The combination of two sgRNAs did not induce DSB as efficiently as single sgRNA treatment. For double knock-out in RKO cells, the ratio of *BRAF* WT:mut was 42:58 compared to 51:49 after single sgRNA treatment and the induction of indels dropped to 7.2% compared to 14.7% after single sgRNA treatment. For *PIK3CA*, the double knock-out was even less efficient. The ratio of WT:mut allele did not shift in HCT116 nor in RKO cells. The indel alteration in the mutant allele dropped to 12.8% and 12.5% in HCT116 and RKO cells, respectively (Supplementary Figure 5).

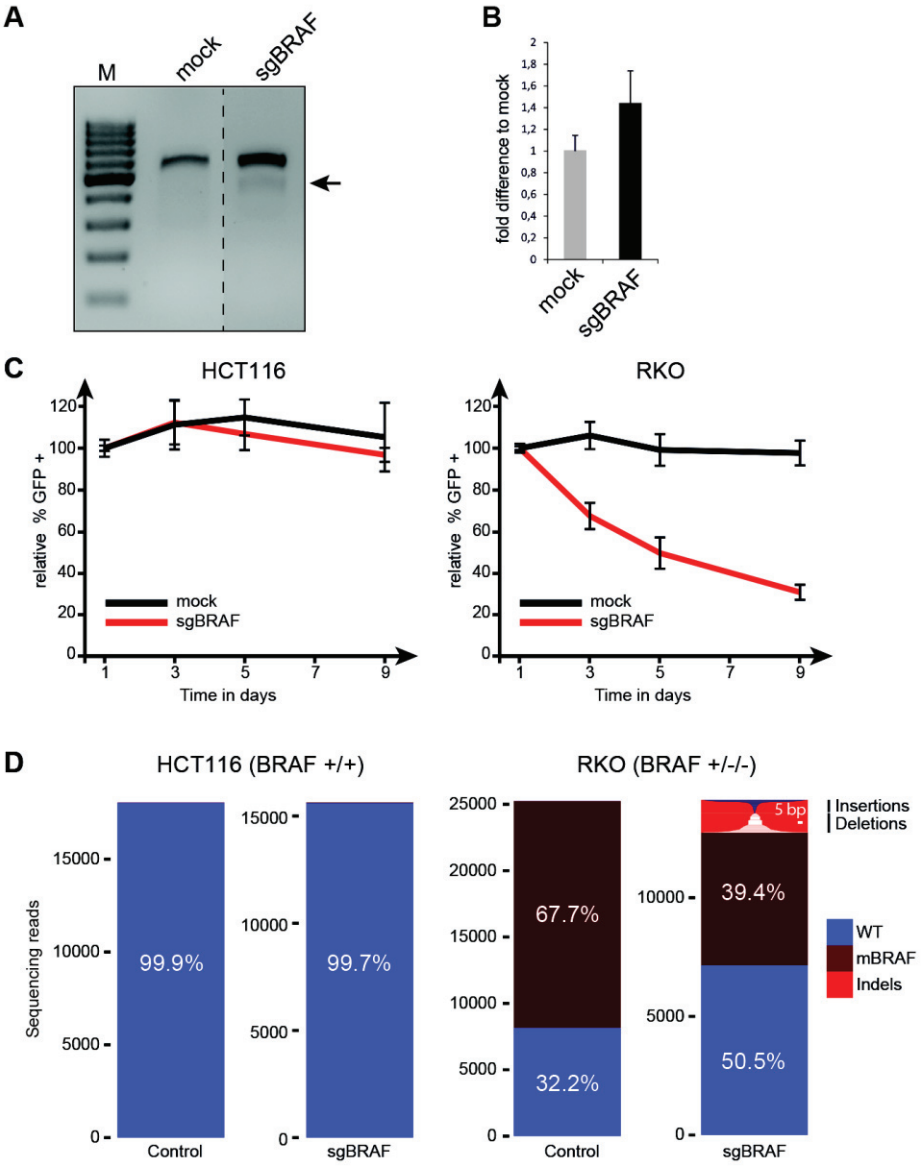
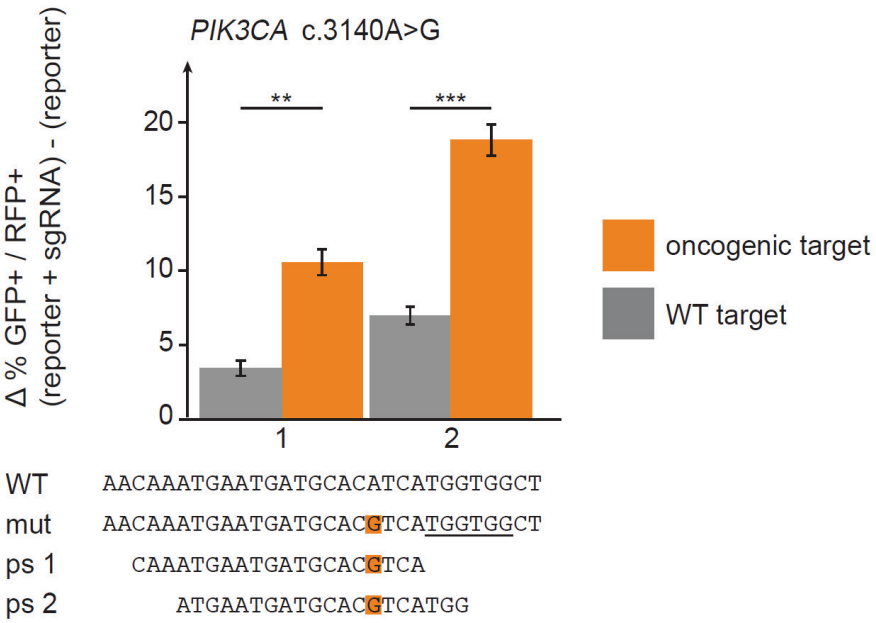


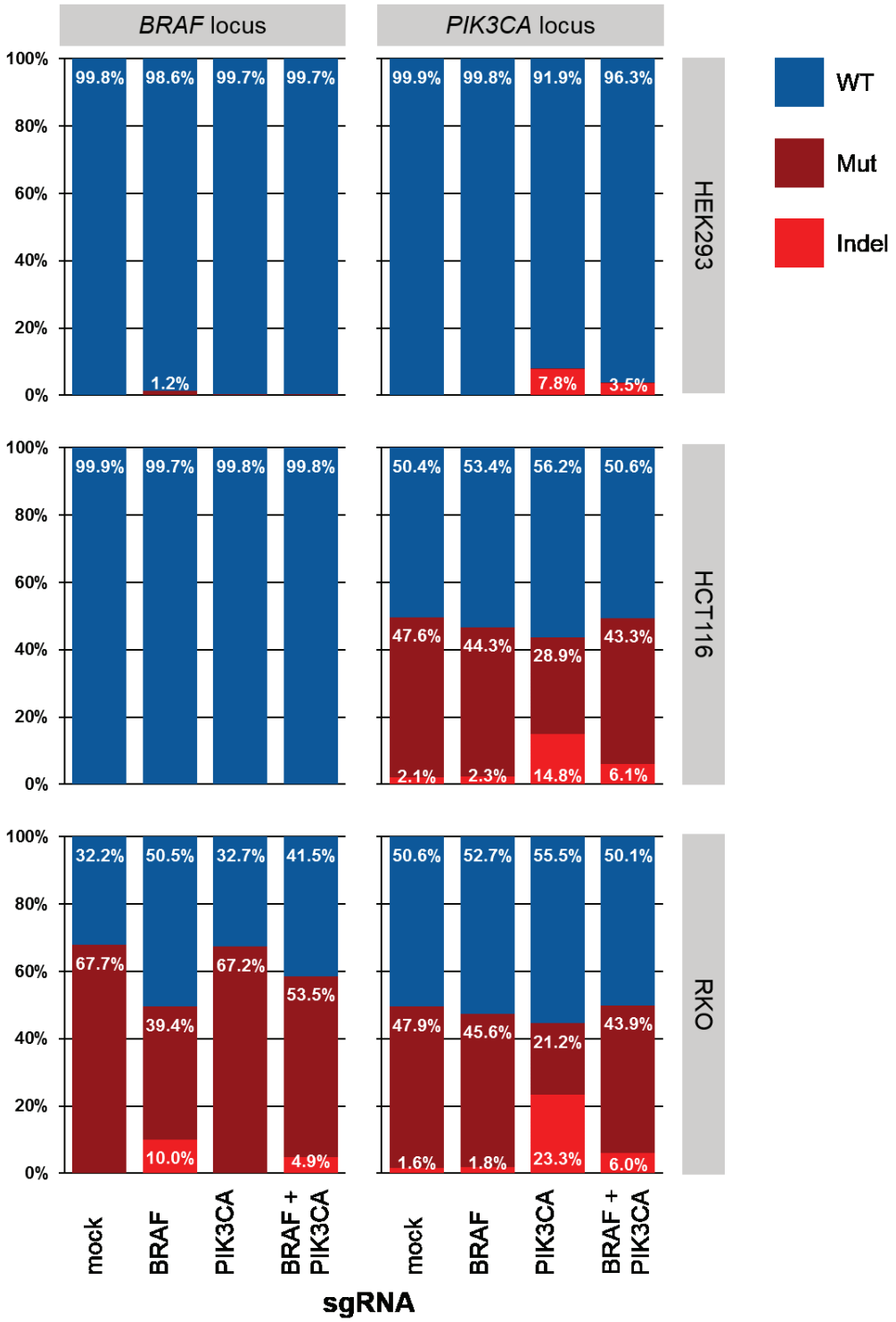
Figure 5: CRISPR-Cas9 mediated inactivation of mutant *BRAF* in RKO cells.

A) T7 Surveyor assay to investigate cleavage of the sgRNA targeting the mutant form of *BRAF*. The arrow points to the cleavage product. B) Quantification of Annexin positive cells treated with indicated sgRNAs. Error bars show SD from experiments performed in [technical] triplicates. C) Relative abundance of cells treated with indicated sgRNAs in HCT116 (no *BRAF* mutation; *BRAF* +/+) and RKO (carrying 2 *BRAF* mutant alleles and one *BRAF* WT allele; *BRAF* +/-/-) over time. Error bars show SD from experiments performed in [technical] triplicates. D) Graphical representation of *BRAF* sequencing reads under indicated conditions. Size of deletions and insertions are indicated in white and [dark blue], respectively.



Supplementary Figure 4: Activity and selectivity of sgRNAs targeting *PIK3CA*.

The wild type (WT), mutant (mut), and protospacer (ps) sequences are illustrated below the graph. Error bars represent SD from experiments performed in technical triplicates. Two-sided Student's t test **P<.01; ***P<.001.



Supplementary Figure 5: Graphical representation of *BRAF* and *PIK3CA* sequencing reads after sgRNA treatment.

HCT116 (*BRAF*^{WT} and *PIK3CA*$^{H1047R/WT}$), HEK293 (*BRAF*^{WT} and *PIK3CA*^{WT}) and RKO cells (*BRAF*$^{V600E/V600E/WT}$ and *PIK3CA*$^{H1047R/WT}$) were transduced with lentiviruses containing Cas9 and either sgRNA targeting mutant *BRAF*, *PIK3CA*, both or a mock sgRNA. Deep sequencing was performed on genomic DNA isolated seven days after sgRNA treatment. WT – wild type;

Mut – mutations; Indel – insertions or deletions. Proportions with less than 1.0% not labelled. One biological replicate each.

4 Discussion

4.1 The design of sgRNAs is possible for most cancer mutations

“With the spCas9, we were able to design sgRNAs for 88% of reported cancer mutations. Orthogonal CRISPR/Cas9 systems ([Esvelt et al., 2013]) and/or the engineering of Cas9 proteins to recognize alternative PAMs ([Kleinstiver et al., 2015]) will increase the spectrum of cancer mutations that can be targeted ([Figure 3]). By generating cell line–specific sgRNA libraries, it might be possible to rapidly identify the most important driver mutations in this setting. […]

Limitations to our study include that even with additional orthogonal CRISPR/Cas systems it will be impossible to design unique sgRNAs for every cancer mutation.” (Gebler et al., 2017)

Furthermore, we observed anti-proliferative results in cancer cell lines only for heterozygous mutation state (see 4.3.2 *BRAF* in melanoma cells). Further investigations in cell lines with homozygous mutations are required to evaluate the efficiency of the CRISPR/Cas-system in these cases.

4.2 sgRNAs targeting oncogenes have to be tested

Although we observed a correlation between the prediction scores and the measured sgRNA activity, we had two outlier sgRNA with almost no activity despite a good prediction score.

“Hence, the current prediction algorithms provide a guideline for the design of efficient sgRNAs, but experimental testing of the actual sequences seems recommendable ([Figure 2]). Interestingly, it was recently shown that nucleosome occupancy impedes Cas9 function ([Horlbeck et al., 2016]), possibly explaining the discrepancy between score and activity for some sgRNAs. […]

Taken together, these results show that the CRISPR/Cas9 traffic-light reporter system is a valuable method to classify efficient and selective sgRNAs that can cleave cancer mutations.” (Gebler et al., 2017)

4.3 Oncogenes can be knocked out with the CRISPR/Cas-system

4.3.1 *NPM1* in AML cells

Although the role of mutant *NPM1* in AML is not completely understood yet, some results support the oncogene theory. Former investigations showed, that a knock-down of wild type nucleophosmin with siRNA decreased the colony growth, reduced mRNA expression and in consequence cytosolic nucleophosmin levels and lead to a G1 arrest in OCI-AML3 cells (Balusu et al., 2011). Confirming these results, we observed a strong growth inhibition, lower

mRNA expression levels and a depletion of cytosolic nucleophosmin in some OCI-AML3 cells after transduction with lentivirus expressing Cas9 and sgRNA targeting mutant *NPM1*. Although the baseline fractions of cells in each cell cycle phase were different in our experiments, we could also observe an increase in G1 phase and a decrease in S phase. Remarkably, the mutated allele was almost completely replaced by either inactivating indels or the WT allele. The fraction of HDR was higher than expected, as NHEJ is the main DSB repair mechanism in mammalian cells (Iliakis et al., 2004).
"Hence, expression of a cancer-specific sgRNA can act in a gene drive fashion to push selection toward the WT sequence." (Gebler et al., 2017)

4.3.2 *BRAF* in melanoma cells

Previous studies emphasized the dependency of *BRAF* mutant melanoma cells on this oncogene. *BRAF* inactivation with Vemurafenib selectively inhibits proliferation in mutant, but not in WT melanoma cell lines *in vitro* (Lee et al., 2010). Confirming these results, a shRNA mediated knock-down showed good anti-proliferative and apoptotic effects in several melanoma cell lines, particularly with a shRNA targeting the mRNA region around the V600E mutation (Hingorani et al., 2003). Therefore, it was surprising that the CRISPR/Cas9 driven knock-out of $BRAF^{V600E}$ did not show anti-proliferative results, although indels could be detected after sgRNA treatment. One explanation could be the homozygous $BRAF^{V600E}$ status of the examined cell lines A-375 and SK-MEL-28. The permanent expression of Cas9 and sgRNA continuously cleaves $BRAF^{V600E}$, until the sequence is altered and the sgRNA fails to bind. In case of HDR, the second mutant allele is the template for repair. Consequently, the WT allele cannot be restored, but only the mutant allele. In the case of NHEJ, the indels may cause mutations in *BRAF* gene, which have unpredictable effects on the protein activity, even a stronger self-activation of B-Raf is imaginable.

4.3.3 *BRAF* and *PIK3CA* in CRC cells

In former studies, a shRNA mediated inactivation of WT *BRAF* mRNA decreased proliferation in $BRAF^{V600E}$ heterozygous RKO cells, but not in $BRAF^{WT}$ HCT116 cells (Preto et al., 2008). These results were confirmed with the selective $BRAF^{V600E}$ inhibitor GDC-0879, which reduces cell viability *in vitro* and tumor growth *in vivo* in nude mice. Apoptosis was slightly, but not significantly enhanced by GDC-0879 (Coffee et al., 2013). In concordance with these results, me and my colleagues showed that CRISPR/Cas9 mediated knock-out of $BRAF^{V600E}$ inhibits cell growth only in the *BRAF* mutant cell line RKO, but not in WT cell lines. Furthermore, the induction of apoptosis was similarly enhanced as with GDC-0879, although we used a different apoptosis assay.

The deep sequencing in RKO cells revealed 27.0% repair of the mutant *BRAF* allele back to the WT allele. Compared with the *NPM1* knockout, “[similar] results were obtained targeting a second common cancer mutation (*BRAF* c.1799T>A) in the colon carcinoma cell line RKO ([Figure 5]), demonstrating that the approach can pinpoint cancer mutation dependencies in cell lines of different origins.” (Gebler et al., 2017)

Previous studies showed that a silencing of p100α with siRNA or shRNA resulted in cell viability and growth reduction, independent on the *PIK3CA* mutation status (Fernandes et al., 2016). Currently, there is no *PIK3CA*^{H1047R} specific p100α inhibitor. The existing ones are either dual p100α/mTor inhibitors (e.g. NVP-BEZ235) or single p100α inhibitors (e.g. GDC0941, NVP-BKM120) with effects on WT as well as mutant p100α, thus producing growth inhibition in *PIK3CA*^{WT} and *PIK3CA*^{H1047R} cell lines (Kong et al., 2014). Another p100α specific inhibitor, NVP-BYL719, is more likely to inhibit mutant *PIK3CA* cell lines from different tissues, but its half maximal inhibitory concentration (IC_{50}) on WT and mutant p100α is the same (Fritsch et al., 2014). Therefore, a comparison with the data from my *PIK3CA*^{H1047R} specific knock-out is limited, although an anti-proliferative effect was observed as well.

In contrast to *NPM1* and *BRAF* knock-out, the deep sequencing results revealed a high fraction of NHEJ and less HDR of the mutant *PIK3CA* allele. Nevertheless, *PIK3CA* sgRNA expression decreased RKO and HCT116 growth in my experiments. NHEJ might result in activating mutations of *PIK3CA*, but more probably many of the frameshift indels will inactivate and knock out the mutant *PIK3CA* allele. Indeed, the tested *PIK3CA* sgRNA also leads to DSB in the wild type allele, what limits conclusions on mutant specific *PIK3CA*^{H1047R} knock-out.

There are results indicating a combined BRAF and p100α inhibition have synergistic effects in CRC treatment in xenograft experiments (Mao et al., 2013). My results did not show a stronger growth inhibition after a double knock-out of *BRAF*^{V600E} and *PIK3CA*^{H1047R} in double mutant RKO cells and deep sequencing results revealed a less effective DSB induction for both, *BRAF*^{V600E} and *PIK3CA*^{H1047R} locus. There are several possible explanations for this result. The sgRNAs in the double knock-out experiment were expressed from the same lentiviral vector, which has only a single Cas9 gene. Therefore, the stoichiometry of Cas9 and sgRNA is different than in single sgRNA expression from this vector. This problem could be solved by the transduction with two lentiviruses, each containing Cas9 gene and sgRNA expression cassette. The *PIK3CA* sgRNA cassette has a H1 promoter. Compared to the hU6 promoter, H1 promoter induces less sgRNA transcription (Ranganathan et al., 2014). In addition, the transcription start site from the H1 promoter is more variable, resulting in production of sgRNAs of different lengths (Ma et al., 2014). These prolongated sgRNAs are less active and less specific than 20 bp long sgRNAs (Ran et al., 2013b). Furthermore, multiplex sgRNA targeting seems to be less effective than single sgRNA targeting (Sakuma et al., 2014). These findings

together could explain the difference of my work to the former described synergistic effect of *PIK3CA* and *BRAF* inhibition.

4.4 Advantages and disadvantages to target oncogenes with the CRISPR/Cas-system

A disadvantage of targeting drugs is the presence of escape mechanisms, for instance through additional mutations. Attacking oncogenic driver mutations before they are expressed into gene products could help to prevent those escape mechanisms. This goal could be achieved with gene knock-out, for example upstream of a mutated and constitutively activated domain, causing a loss-of-function mutation. With the development of effective delivery possibilities, gene therapy of cancer might be a future goal of cancer research.

However, there are certain disadvantages of gene editing in general. Due to the observed HDR effects after DSB induction and the statistical smaller chance to cleave two alleles than one, gene therapy in homozygous mutant or aneuploid cells is supposed to be less efficient than in heterozygous mutant cells. But if the knock-out or correction works, it is permanent and inheritable to the daughter cells. Another disadvantage is the use of HDR only in proliferating cells, whereas resting cells employ NHEJ to fix a DSB (Unniyampurath et al., 2016).

“Furthermore, control over the repair mechanism after Cas9-mediated DNA cleavage is limited. It is therefore likely that sgRNA-resistant clones may emerge that maintain the oncogenic phenotype. In addition, off-target cleavage has to be considered a potential risk factor in a therapeutic setting.” (Gebler et al., 2017)

However, former studies, employing siRNA or shRNA for the silencing of oncogenes, showed high off-target effects. The off-target potential of the CRISPR/Cas-system in human cells might be lower compared to siRNA and shRNA and improvements promise a further reduction (Unniyampurath et al., 2016). Nevertheless, off-target activity of the CRIPSR/Cas-system is detectable at sites with up to four base mismatches between sgRNA and DNA, particularly with mismatches at the 5’-end of the protospacer, and to a smaller extent at sites with single base insertion and deletion mismatches, thus increasing the number of potential off-targets exponentially (Hsu et al., 2013; Doench et al., 2016).

“Nevertheless, considering current cancer treatment regimes employing DNA-damaging drugs and/or radiation, the CRISPR/Cas9 system is conceivable to be less genotoxic and cause less undesired DNA lesions in cells. Given the prominent gene drive effect we observed to repair the cancer mutation back to the WT sequence, installing the CRISPR/Cas9 system as a “tumor protection system” is also worth contemplating ([Figure 6]). This way, the system would act as a “cancer mutation immune system,” eliminating or repairing malignant lesions when they occur and before cells become cancerous.” (Gebler et al., 2017)

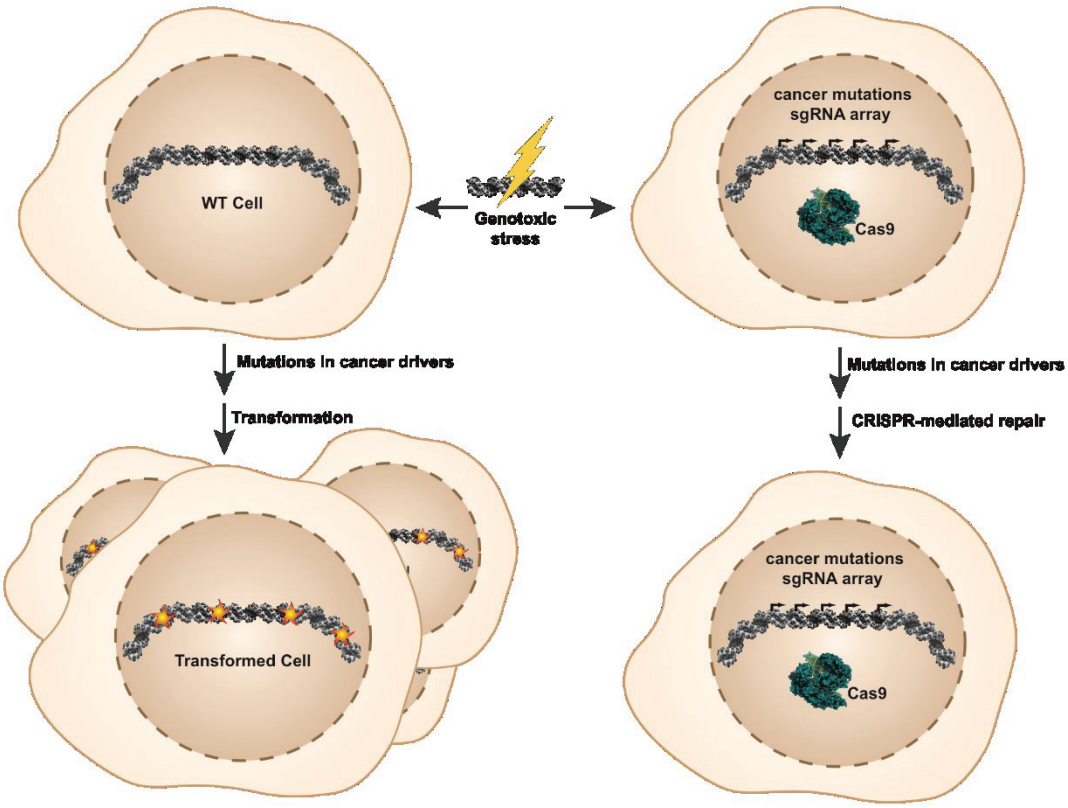


Figure 6: Model of a CRISPR/Cas9 tumor protection system.
Important steps are highlighted. WT - wild type cells.

4.5 Concluding remarks

Besides investigations on the reversion of the BCR-ABL1-translocation and the removal of oncogenic viruses, there are few studies on oncogene knock-out utilizing the CRISPR/Cas-system (Lekomtsev et al., 2016).

Currently, this work is the first one to employ the CRISPR/Cas-system for the knock-out of oncogenes with gene mutations and moreover the first work using programmable nucleases on mutant *NPM1*, *BRAF* and *PIK3CA*. Our results indicate that a knock-out of oncogenes with differentiation between the WT and mutated allele is feasible and shows anti-proliferative effects only in cells with the correspondent mutation.

“Overall, we conclude that mutant *NPM1* and mutant *BRAF* are required for OCI-AML3 and RKO proliferation, respectively, and that the CRISPR/Cas9 system is a powerful tool to dissect the relevance of cancer mutations in tumor cells. […]

Furthermore, we envision that this approach is transferable to primary patient samples ([Figure 7]), and, in the long run, CRISPR/Cas9 could potentially be considered a therapeutic approach to target patient-specific mutations in affected individuals. Delivery of Cas9 and mutation-specific sgRNAs into tumor cells by, eg, oncolytic viruses ([Kaufman et al., 2015]) could form

a potent, individualized therapy that could complement current treatment strategies. In particular, combination therapy, where two or more cancer mutations are targeted at the same time, is relatively straightforward in this setting, when several specific sgRNAs can be provided simultaneously ([Figure 7])." (Gebler et al., 2017)

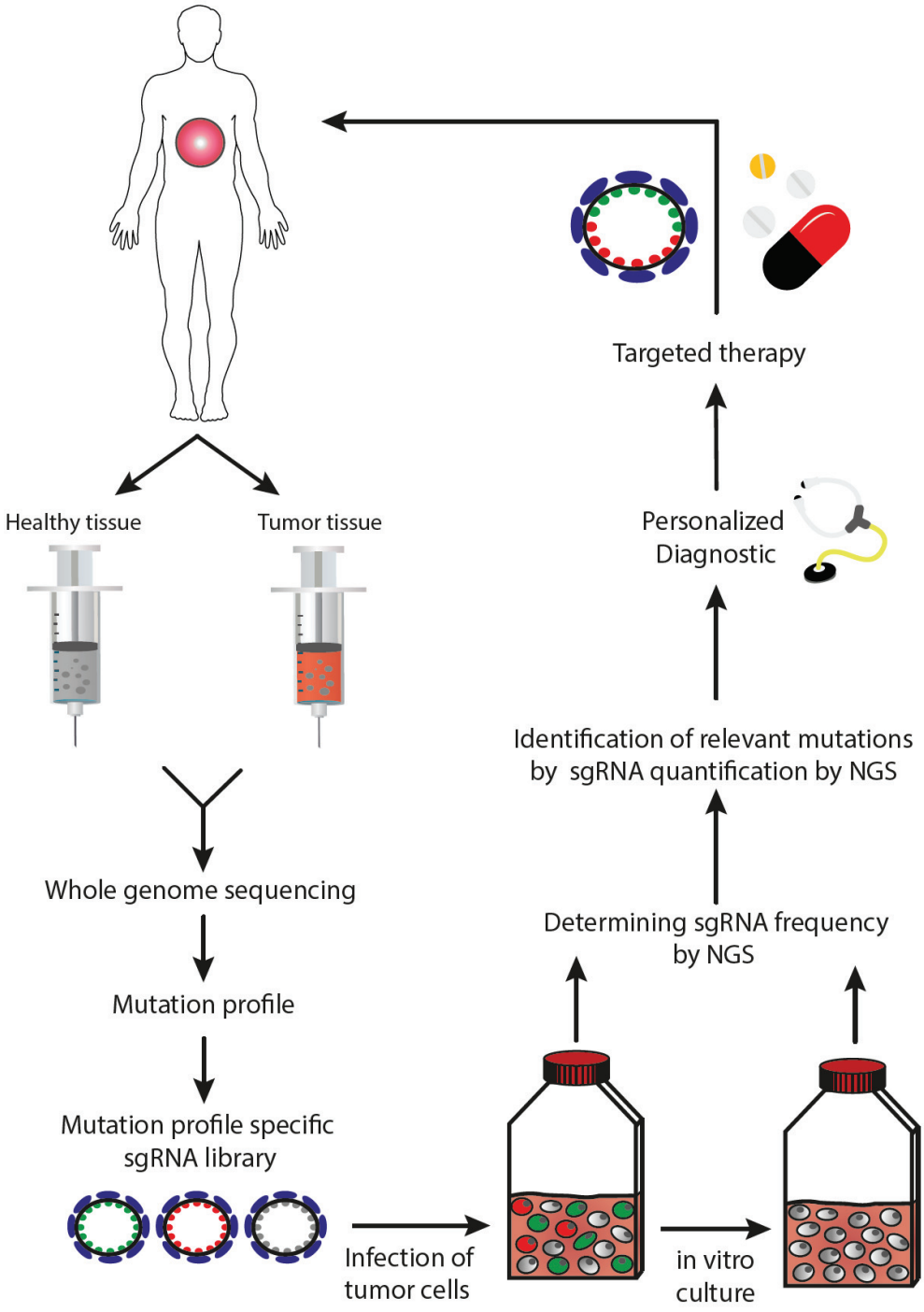


Figure 7: Scheme for functional profiling of cancer mutations in individualized therapy. Important steps are highlighted. NGS - next generation sequencing.

OXFORD

JNCI J Natl Cancer Inst (2017) 109(1): djw183

doi: 10.1093/jnci/djw183
First published online August 30, 2016
Brief Communication

BRIEF COMMUNICATION

Inactivation of Cancer Mutations Utilizing CRISPR/Cas9

Christina Gebler*, Tim Lohoff*, Maciej Paszkowski-Rogacz, Jovan Mircetic, Debojyoti Chakraborty, Aylin Camgoz, Martin V. Hamann, Mirko Theis, Christian Thiede, Frank Buchholz

Affiliations of author: Medical Systems Biology, Medical Faculty Carl Gustav Carus, University cancer center (UCC), Technische Universität (TU) Dresden, Dresden, Germany (CG, TL, MPR, JM, DC, AC, MVH, MT, FB); Medical Faculty Carl Gustav Carus, Medizinische Klinik und Poliklinik I, Dresden, Germany (CT); Max Planck Institute of Molecular Cell Biology and Genetics, Dresden, Germany (FB); German Cancer Research Center (DKFZ), Heidelberg and German Cancer Consortium (DKTK) partner site Dresden, Germany (FB, CT); National Center for Tumor Diseases (NCT), University Hospital Carl Gustav Carus, TU Dresden, Dresden, Germany (FB)

Correspondence to: Frank Buchholz, PhD, Medical Systems Biology, UCC, Medical Faculty Carl Gustav Carus, TU Dresden, Fetscherstrasse 74, 01307 Dresden, Germany (e-mail: frank.buchholz@tu-dresden.de).
*Authors contributed equally to this work.

Abstract

Although whole-genome sequencing has uncovered a large number of mutations that drive tumorigenesis, functional ratification for most mutations remains sparse. Here, we present an approach to test functional relevance of tumor mutations employing CRISPR/Cas9. Combining comprehensive sgRNA design and an efficient reporter assay to nominate efficient and selective sgRNAs, we establish a pipeline to dissect roles of cancer mutations with potential applicability to personalized medicine and future therapeutic use.

Genetic mutations are a hallmark of cancer development, and more than 140 cancer driver genes have been described to date (1,2). Identification of all mutations in an actual tumor of a patient by whole-genome sequencing is rapidly emerging as the method of choice for precision diagnostics (3). However, detailed knowledge of the functional roles and relevance of most mutations arising during tumorigenesis are still lacking.

We set out to test whether the CRISPR/Cas9 system (4) can aid the functional investigation of mutations detected in cancer cells. To first investigate how many cancer mutations could theoretically be targeted by *Streptococcus pyogenes (sp)*Cas9, we performed a comprehensive bioinformatics analysis of published cancer mutations (2). From the reported 608 671 unique mutations, we were able to design 1 909 172 sgRNAs that cover 554069 mutations (91.0%) and 20 756 out of 20 948 mutated genes. We then performed an analysis to avoid off-target cleavage and discarded all sgRNAs having additional perfect matches to sequences in the reference genome, in addition to prioritizing sgRNAs with the highest divergence to homologous sequences elsewhere in the genome (5). Based on these criteria, we nominated 1 701 813 sgRNAs that could theoretically target 535 327 (88.0%) of known cancer mutations encompassing 10 349 (85.0%) of known cancer driver mutations (Figure 1A; Supplementary Table 1, available online).

We next established a "traffic-light" reporter system (6), where Cas9 cleavage activates GFP expression in transiently transfected mammalian cells in culture to rapidly evaluate efficacy and selectivity of designed sgRNAs (Figure 1B). Sequences bearing 13 different cancer mutations or the corresponding WT sequences were cloned into the reporter construct and subsequently cotransfected into HeLa cells with a Cas9 expression plasmid (7) that also expressed the cancer mutation–specific sgRNA (Figure 1, B and C). Efficient cleavage was observed for most constructs bearing cancer mutations, with 10 out of 13 sgRNAs also showing a higher than 4-fold target site selectivity over the wild-type (WT) sequence, with the remaining three still showing 2.7- to 3.8-fold selectivity of mutant over WT. In particular, insertion and deletion mutations reported in the genes *KIT*, *NPM1*, *CEBPA*, *EGFR*, and *WT1* showed little to no appearance of green cells when combined with the WT reporters (Figure 1C), reflecting that the WT sequences were not cleaved efficiently. In contrast, 10% to 25% of GFP-positive cells were

Received: December 4, 2015; **Revised:** May 6, 2016; **Accepted:** June 24, 2016

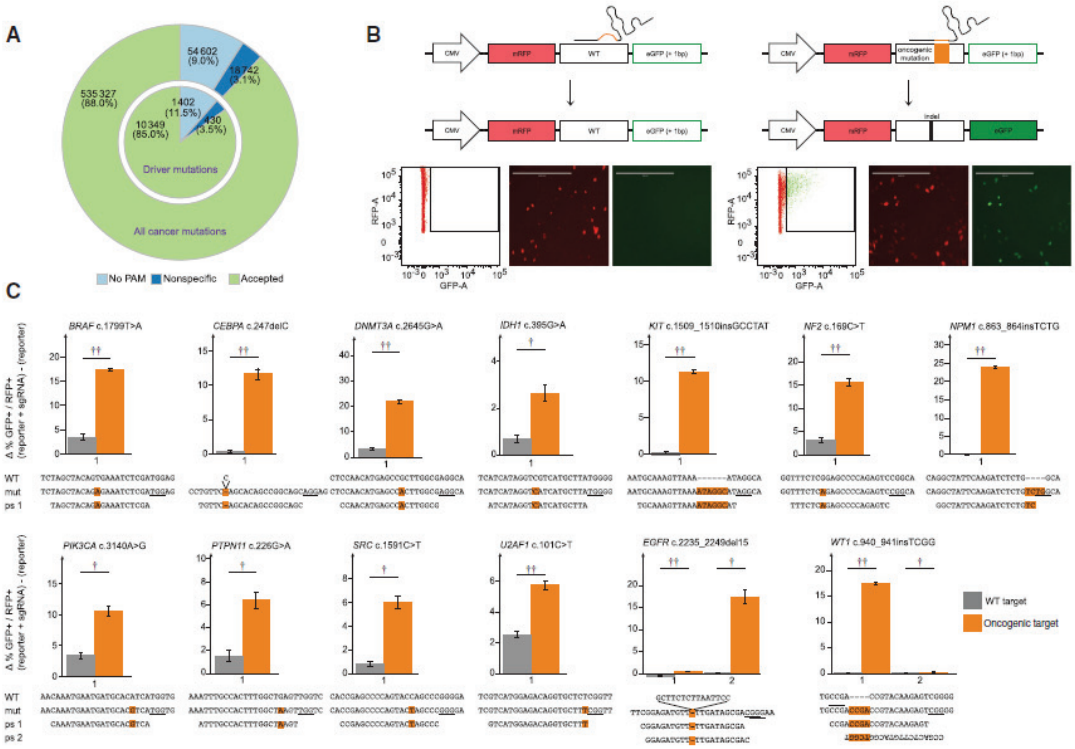


Figure 1. sgRNA design and evaluation of sgRNA efficacy and selectivity. **A)** Bioinformatics analysis and sgRNA design for cancer mutations. A **pie chart** for reported cancer mutations and for cancer driver mutations for *Streptococcus pyogenes* sgRNAs is shown. **B)** Overview of "traffic light" reporter assay. Important elements are indicated. Representative examples of fluorescence-activated cell sorting plots and microscopy images are shown (**scale bars** = 400 μm). **C)** Activity and selectivity of employed sgRNAs. The targeted mutations are indicated **above each graph**, with the wild-type, mutant, and protospacer sequences illustrated **below each graph. Error bars** represent SD from experiments performed in triplicates. Two-sided Student's t test *P<.05; **P<.01; ***P<.001. mut = mutant; ps = protospacer; WT = wild-type.

detected when the cancer mutation reporters were used in combination with matching sgRNAs. Hence, these sgRNAs created indels in the reporter plasmids that brought the GFP sequence into the correct reading frame, demonstrating their potency to cleave the cancer mutation sequence. Overall, we observed a descent correlation between the sgRNA prediction score and the actual activity in the traffic light reporter assay. However, we detected considerable differences in cleavage efficacy for some sgRNAs targeting the identical cancer mutation, despite the fact that their prediction scores (8,9) were similar (Supplementary Table 2, available online). For instance, sgRNA#1 with a score of 0.42 for the *EGFR* 2235_2249del15 mutation only produced 0.5% (+/-0.1%) of GFP-positive cells, whereas the related sgRNA#2 with a score of 0.33 that is only shifted by one base pair was highly efficient and resulted in more than 17.4% (+/-1.5%) of GFP-positive cells. Hence, the current prediction algorithms provide a guideline for the design of efficient sgRNAs, but experimental testing of the actual sequences seems recommendable (Supplementary Figure 1, available online). Interestingly, it was recently shown that nucleosome occupancy impedes Cas9 function (10), possibly explaining the discrepancy between score and activity for some sgRNAs. Remarkably, many point mutations, such as the *DNMT3A* c.2645G>A mutation, were efficiently cleaved by the cancer mutation sgRNA (21.9% [+/-0.8%] GFP-positive cells) without appreciably cleaving the WT sequence (3% [+/-0.2%] GFP-positive cells), demonstrating that the CRISPR/Cas9 system can be sensitive enough to distinguish single base pair alterations. Taken together, these results show that the CRISPR/Cas9 traffic-light reporter system is a valuable method to classify efficient and selective sgRNAs that can cleave cancer mutations.

We next investigated the functional relevance of two common cancer mutations in tumor cells. The nucleophosmin gene (*NPM1*) is mutated in about 30% of patients suffering from acute myeloid leukemia (AML) (11). Mutant *NPM1* is thought to play an important role in AML proliferation, indicating that a direct way to inactivate the mutation could affect malignant growth (12). We cloned the tested sgRNA sequence targeting mutant *NPM1* (Figure 1C) into a lentiviral vector (13) expressing Cas9 in conjunction with EGFP and transduced *NPM1* mutant OCI-AML3 cells and *NPM1* WT MV4-11 cells with the virus. Efficient cleavage of mutant *NPM1* in OCI-AML3 cells was evident in employing multiple assays (Figure 2). Strikingly, transduced OCI-AML3, but not the MV4-11 cells, were successively depleted over time (Figure 2C), signifying that the mutant NPM1 protein is required for efficient cell proliferation in OCI-AML3. Cell cycle analyses revealed that OCI-AML3 cells treated with the *NPM1* sgRNA arrested in G1 without markedly altering the subG1 fraction (Figure 2D), suggesting that mutant *NPM1* expression in these cells is required for cell cycle progression. To investigate the

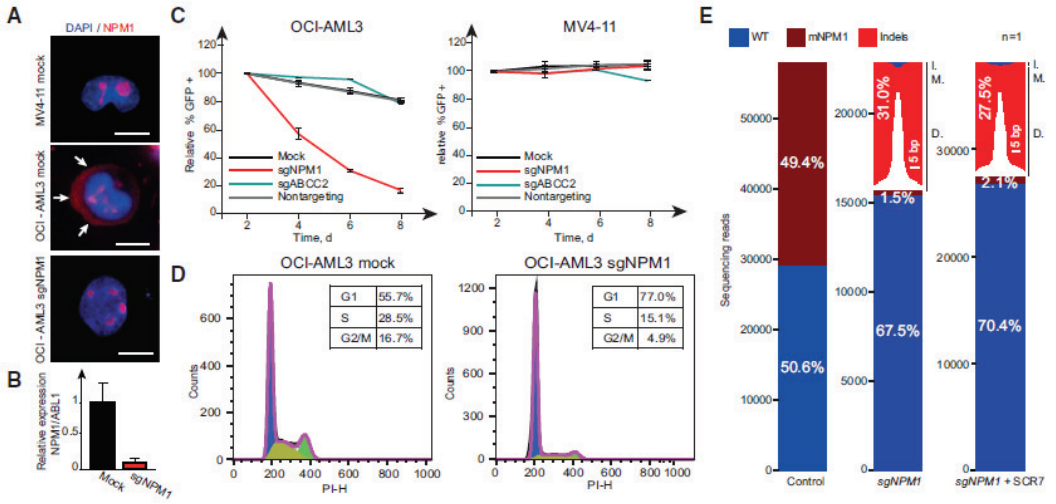


Figure 2. Effects of mutant *NPM1* inactivation. **A)** Localization of NPM1 in MV4-11 and OCI-AML3 cells under indicated treatment conditions (**scale bars** = 10 μm). **Arrows** highlight the cytoplasmatic localization of mutant NPM1 in mock-treated OCI-AML3 cells. OCI-AML3 sgNPM1 images were taken 96 hours after sgRNA treatment. **B)** Relative mRNA expression level of mutated *NPM1* after sgRNA treatment. mRNA was extracted 96 hours after sgRNA treatment. **C)** Relative abundance of cells treated with indicated sgRNAs in MV4-11 and OCI-AML3 over time. **Error bars** show SD from experiments performed in triplicates. **D)** Cell-cycle profile of OCI-AML3 cells after indicated treatments. Fluorescence-activated cell sorting analyses were performed eight days after sgRNA treatment. **E)** Graphical representation of *NPM1* sequencing reads under indicated conditions (one biological replicate each). Size of deletions and insertions are indicated in **white** and **black**, respectively. sgNPM1–OCI-AML3 cells treated with sgRNA-targeting mutant *NPM1* (*mNPM1*). sgNPM1 + SCR7–OCI-AML3 cells treated with sgRNA-targeting mutant *NPM1* in the presence of the ligase IV inhibitor SCR7. Sequencing was performed on genomic DNA isolated eight days after sgRNA treatment. Sixty-four point six percent of the indels resulted in reading frame shifts. D. = deletions; I. = insertions; M. = mutations; WT = wild-type.

mutational spectrum at the site of cleavage, we performed deep sequencing of the *NPM1* locus in control- and sgNPM1-treated cells. As expected, cells treated with a control sgRNA revealed a 50:50 ratio for the WT and mutant allele, reflecting the heterozygous nature of the *NPM1* mutation. In contrast, cells treated with the sgRNA-targeting mutant *NPM1* showed efficient cleavage and repair of the mutant allele. Remarkably, the WT:*NPM1* ratio in this sample increased to around 70:30, indicating that a substantial fraction (34.2%) of the cells had repaired the mutation back to the WT sequence (Figure 2E), likely through homologous recombination utilizing the WT allele as a template. The fraction of indel mutations was further reduced (from 62.8% to 55.7%) when the cells were treated with the DNA ligase IV inhibitor SCR7 (14), indicating that enhanced HR-mediated repair can be achieved when the NHEJ pathway is inhibited. Hence, expression of a cancer-specific sgRNA can act in a gene drive fashion to push selection toward the WT sequence. Similar results were obtained targeting a second common cancer mutation (*BRAF* c.1799T>A) in the colon carcinoma cell line RKO (Supplementary Figure 2, available online), demonstrating that the approach can pinpoint cancer mutation dependencies in cell lines of different origins. Overall, we conclude that mutant *NPM1* and mutant *BRAF* are required for OCI-AML3 and RKO proliferation, respectively, and that the CRISPR/Cas9 system is a powerful tool to dissect the relevance of cancer mutations in tumor cells.

With the *sp*Cas9, we were able to design sgRNAs for 88% of reported cancer mutations. Orthogonal CRISPR/Cas9 systems (15) and/or the engineering of Cas9 proteins to recognize alternative PAMs (16) will increase the spectrum of cancer mutations that can be targeted (Supplementary Figure 3, available online). By generating cell line–specific sgRNA libraries, it might be possible to rapidly identify the most important driver mutations in this setting. Furthermore, we envision that this approach is transferable to primary patient samples (Supplementary Figure 4, available online), and, in the long run, CRISPR/Cas9 could potentially be considered a therapeutic approach to target patient-specific mutations in affected individuals. Delivery of Cas9 and mutation-specific sgRNAs into tumor cells by, eg, oncolytic viruses (17) could form a potent, individualized therapy that could complement current treatment strategies. In particular, combination therapy, where two or more cancer mutations are targeted at the same time, is relatively straightforward in this setting, when several specific sgRNAs can be provided simultaneously (Supplementary Figure 4, available online).

Limitations to our study include that even with additional orthogonal CRISPR/Cas systems it will be impossible to design unique sgRNAs for every cancer mutation. Furthermore, control over the repair mechanism after Cas9-mediated DNA cleavage is limited. It is therefore likely that sgRNA-resistant clones may emerge that maintain the oncogenic phenotype. In addition, off-target cleavage has to be considered a potential risk factor in a therapeutic setting.

Nevertheless, considering current cancer treatment regimes employing DNA-damaging drugs and/or radiation, the CRISPR/Cas9 system is conceivable to be less genotoxic and cause less undesired DNA lesions in cells. Given the prominent gene drive effect we observed to repair the cancer mutation back to the WT sequence, installing the CRISPR/Cas9 system as a "tumor protection system" is also worth contemplating (Supplementary Figure 5, available online). This way, the system would act as a "cancer mutation immune system," eliminating or repairing malignant lesions when they occur and before cells become cancerous.

Funding

This work was supported by grants from the Mildred-Scheel Program of the German cancer aid, the Else Kröner-Promotionskolleg, the Excellence Initiative by the German Federal and State Governments (Institutional Strategy, measure "support the best ZUK 64"), and the German Cancer Consortium joint funding program "novel tools for dissection of oncogenic pathways."

Notes

The funders had no role in the design of the study; the collection, analysis, or interpretation of the data; the writing of the manuscript; or the decision to submit the manuscript for publication. The authors declare no competing financial interest. Sequencing data will be deposited into the National Center for Biotechnology Information Short Read Archive. The authors thank Juliane Reh and Marika Karger for technical assistance, Stanislava Popova for help with cell cycle analyses, and Martin Schneider for support with cloning and fluorescence-activated cell sorting analyses.

References

1. Vogelstein B, Papadopoulos N, Velculescu VE, et al. Cancer genome landscapes. *Science*. 2013;339(6127):1546–1558.
2. Kandoth C, McLellan MD, Vandin F, et al. Mutational landscape and significance across 12 major cancer types. *Nature*. 2013;502(7471):333–339.
3. Dewey FE, Grove ME, Pan C, et al. Clinical interpretation and implications of whole-genome sequencing. *JAMA*. 2014;311(10):1035–1045.
4. Jinek M, Chylinski K, Fonfara I, et al. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science*. 2012;337(6096):816–821.
5. Hsu PD, Scott DA, Weinstein JA, et al. DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat Biotechnol*. 2013;31(9):827–832.
6. Karimova M, Beschorner N, Dammermann W, et al. CRISPR/Cas9 nickase-mediated disruption of hepatitis B virus open reading frame S and X. *Sci Rep*. 2015;5:13734.
7. Ran FA, Hsu PD, Wright J, et al. Genome engineering using the CRISPR-Cas9 system. *Nat Protoc*. 2013;8(11):2281–2308.
8. Doench JG, Hartenian E, Graham DB, et al. Rational design of highly active sgRNAs for CRISPR-Cas9-mediated gene inactivation. *Nat Biotechnol*. 2014;32(12):1262–1267.
9. Doench JG, Fusi N, Sullender M, et al. Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. *Nat Biotechnol*. 2016;34(2):184–191.
10. Horlbeck MA, Witkowsky LB, Guglielmi B, et al. Nucleosomes impede Cas9 access to DNA in vivo and in vitro. *Elife*. 2016;5.
11. Falini B, Mecucci C, Tiacci E, et al. Cytoplasmic nucleophosmin in acute myelogenous leukemia with a normal karyotype. *N Engl J Med*. 2005;352(3):254–266.
12. Federici L, Falini B. Nucleophosmin mutations in acute myeloid leukemia: a tale of protein unfolding and mislocalization. *Protein Sci*. 2013;22(5):545–556.
13. Heckl D, Kowalczyk MS, Yudovich D, et al. Generation of mouse models of myeloid malignancy with combinatorial genetic lesions using CRISPR-Cas9 genome editing. *Nat Biotechnol*. 2014;32(9):941–946.
14. Srivastava M, Nambiar M, Sharma S, et al. An inhibitor of nonhomologous end-joining abrogates double-strand break repair and impedes cancer progression. *Cell*. 2012;151(7):1474–1487.
15. Esvelt KM, Mali P, Braff JL, et al. Orthogonal Cas9 proteins for RNA-guided gene regulation and editing. *Nat Methods*. 2013;10(11):1116–1121.
16. Kleinstiver BP, Prew MS, Tsai SQ, et al. Engineered CRISPR-Cas9 nucleases with altered PAM specificities. *Nature*. 2015;523(7561):481–485.
17. Kaufman HL, Kohlhapp FJ, Zloza A. Oncolytic viruses: a new class of immunotherapy drugs. *Nat Rev Drug Discov*. 2015;14(9):642–662.

6 Summary

Numerous mutations contribute to tumorigenesis of cancer cells. For most of them it remains unclear whether they are driver or passenger mutations. A classic knock-out to study their function in cancer cells used to take a lot of effort. The CRISPR/Cas-system can be used as a programmable “genome-editing” tool. In this work, oncogenes have been inactivated with the CRISPR/Cas-system.

Considering off-targets, *Streptococcus pyogenes* sgRNAs can be designed for 88% of the known cancer mutations. The activity of 15 sgRNAs, targeting 13 mutations in proto-oncogenes (deletions, insertions and point mutations), has been tested with a RFP-GFP-reporter plasmid. For 13 sgRNAs, activity prediction scores correlated with measured activity. Furthermore, sgRNAs have shown preferential binding to mutated versions of targeted proto-oncogene sequence and did not induce double strand breaks in the wild type sequence. For 10 sgRNAs, the activity against their target sequence has been more than 4 times higher than against the wild type sequence. Most of those sgRNAs target insertions or deletions and fewer target point mutations.

Permanent knock-out of three mutated proto-oncogenes *NPM1*, *BRAF* and *PIK3CA* has been achieved with a lentiviral expression of CRISPR/Cas. Accordingly, effects on proliferation and phenotype have been studied.

Knock-out of *NPM1* c.863_864insTCTG mutation has been studied in heterozygous mutated OCI-AML3 cell line. Proliferation was strongly inhibited by the corresponding sgRNA. Cells arrested in G0/1-phase of cell cycle (77%) compared to control cells (56%), although no difference was observed for sub-G1 phase, indicating no induction of apoptosis. Cells treated with *NPM1* sgRNA had 88% reduced expression of *NPM1* c.863_864insTCTG mRNA as well as less cytoplasmic localization of nucleophosmin as assessed by immunostaining. The activity of sgRNA has been confirmed by deep sequencing, showing a shift of wild type to mutated allele ratio from 51:49 to 68:32. This effect was enhanced by the additional treatment with the NHEJ inhibitor SCR7.

A *BRAF*^{V600E} sgRNA was tested in homozygously mutated melanoma cell lines A-375 and SK-MEL-28. No differences were detected in comparison to controls. However, in the CRC cell line RKO, heterozygous for *BRAF*^{V600E} and *PIK3CA*^{H1047R}, proliferation was inhibited through sgRNAs against either *BRAF* or *PIK3CA*. A combination of both had no synergistic effect on proliferation. Activity and specificity of the sgRNA targeting *BRAF* were confirmed by deep sequencing, while the *PIK3CA* sgRNA showed a moderate induction of double strand breaks also in the wild type allele. The relation of wild type to mutated allele of *BRAF* was changed from 32:68 before treatment to 51:49 afterwards. This effect can be explained by a

“re-mutation” to the wild type after DSB via HDR with wild type sister chromatid as template. This effect was observed for *PIK3CA* sgRNA to a lesser extent.

In conclusion, these results show the applicability of the CRISPR/Cas-system for the inactivation of mutated proto-oncogenes.

7 Zusammenfassung

In Krebszellen tritt eine Vielzahl von Mutationen auf. Für den Großteil der Mutationen ist ungeklärt, ob es sich um krebsverursachende oder passagere Mutationen handelt. Ein gezieltes Ausschalten (Knock-out) dieser Gene zur Untersuchung ihrer Funktion in Krebszellen war bisher mit großem Aufwand verbunden. Das CRISPR/Cas-System lässt sich als programmierbares „Genome-editing" Werkzeug einsetzen und wurde in der vorliegenden Arbeit verwendet, um gezielt mutierte Protoonkogene zu inaktivieren.

Für 88% der bekannten, in Krebszellen auftretenden Mutationen lassen sich, unter Berücksichtigung von off-targets, *Streptococcus pyogenes* sgRNAs entwerfen. Mit Hilfe eines RFP-GFP-Reporter-Plasmides wurde die Aktivität von 15 sgRNAs gegen 13 Mutationen (Deletionen, Insertionen und Punktmutationen) in Protoonkogenen überprüft. Für 13 der sgRNAs zeigte sich eine Aktivität, die mit der Vorhersage durch den Algorithmus korrelierte. Außerdem wurde gezeigt, dass die sgRNAs spezifisch genug binden, um zwar bei der mutierten Sequenz eines Protoonkogens, jedoch nicht bei der Wildtyp-Sequenz Doppelstrangbrüche zu erzeugen. Unter den sgRNAs waren 10 mit mehr als 4-fach höherer Aktivität bei komplett übereinstimmender Zielsequenz gegenüber der Wildtyp-Sequenz. Diese spezifischen sgRNAs waren vor allem gegen Insertions- oder Deletionsmutationen gerichtet, einige auch gegen Punktmutationen.

Durch permanente, lentivirale Expression von CRISPR/Cas wurden die Effekte eines Knock-out von drei mutierten Protoonkogenen, *NPM1*, *BRAF* und *PIK3CA*, auf das Wachstum und phänotypische Aspekte humaner Krebszelllinien untersucht.

Ein Knock-out der *NPM1* c.863_864insTCTG Mutation wurde in heterozygot mutierten OCI-AML3 Zellen untersucht, es zeigte sich eine starke Proliferationshemmung. In der Zellzyklusanalyse trat ein G0/1-Arrest dieser Zellen (77%) im Vergleich mit Kontroll-Zellen (56%) auf, jedoch keine Unterschiede in der sub-G1-Analyse, sodass nicht von einer vermehrten Apoptose auszugehen ist. Die mit sgRNA behandelten OCI-AML3 Zellen zeigten sowohl eine um 88% verminderte *NPM1* c.863_864insTCTG mRNA-Expression als auch verminderte zytoplasmatische Sublokalisation des Nucleophosmins in der Immunfärbung. Die hohe Aktivität der gRNA gegen mutiertes *NPM1* wurde durch Deep Sequencing bestätigt, außerdem hat sich das Verhältnis vom Wildtyp- zu mutiertem Allel von 51:49 zu 68:32 verschoben. Dieser Effekt wurde durch Zugabe des NHEJ-Hemmstoffes SCR7 noch verstärkt.

Die sgRNA gegen *BRAF*^{V600E} wurde in den homozygot mutierten Melanom-Zelllinien A-375 und SK-MEL-28 getestet. Bei Proliferationsversuchen zeigten sich keine Unterschiede im Vergleich zu Kontrollzellen. In der kolorektalen Krebszelllinie RKO, die heterozygot *BRAF*^{V600E} und *PIK3CA*^{H1047R} ist, zeigte sich bei der Testung von sgRNAs gegen *BRAF*, *PIK3CA* und

Kombination beider sgRNAs eine Wachstumshemmung. Jedoch lag kein synergistischer Effekt bei sgRNA-Kombination vor. Zudem bestätigten sich Aktivität und Spezifität der sgRNA gegen *BRAF* im Deep Sequencing, während die sgRNA gegen *PIK3CA* in mäßigem Umfang Doppelstrangbrüche im Wildtyp-Allel verursachte. Das Verhältnis vom Wildtyp- zu mutiertem *BRAF*-Allel verschob sich von 32:68 ohne sgRNA zu 51:49 nach sgRNA-Behandlung. Eine mögliche Erklärung dieser Beobachtung ist die Rückmutation zum Wildtyp-Allel nach Doppelstrangbruch mit Hilfe homologer Rekombination durch das Wildtyp-Schwesterchromatid. Für *PIK3CA* konnte dieser Effekt in schwächerem Ausmaß ebenfalls beobachtet werden.

Zusammengefasst zeigen diese Ergebnisse, dass das CRISPR/Cas-System zur Inaktivierung mutierter Protoonkogene genutzt werden kann.

List of references

Ahmad AS, Ormiston-Smith N, Sasieni PD. 2015. Trends in the lifetime risk of developing cancer in Great Britain: comparison of risk for those born from 1930 to 1960. Br J Cancer 112:943–947.

Bader AG, Kang S, Vogt PK. 2006. Cancer-specific mutations in PIK3CA are oncogenic in vivo. Proc Natl Acad Sci 103:1475–1479.

Balusu R, Fiskus W, Rao R, Chong DG, Nalluri S, Mudunuru U, Ma H, Chen L, Venkannagari S, Ha K, Abhyankar S, Williams C, McGuirk J, Khoury HJ, Ustun C, Bhalla KN. 2011. Targeting levels or oligomerization of nucleophosmin 1 induces differentiation and loss of survival of human AML cells with mutant NPM1. Blood 118:3096–3106.

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

Carpenter CL, Duckworth BC, Auger KR, Cohen B, Schaffhausen BS, Cantley LC. 1990. Purification and characterization of phosphoinositide 3-kinase from rat liver. J Biol Chem 265:19704–19711.

Chan WY, Liu QR, Borjigin J, Busch H, Rennert OM, Tease LA, Chan PK. 1989. Characterization of the cDNA encoding human nucleophosmin and studies of its role in normal and abnormal growth. Biochemistry (Mosc) 28:1033–1039.

Chapman PB, Hauschild A, Robert C, Haanen JB, Ascierto P, Larkin J, Dummer R, Garbe C, Testori A, Maio M, Hogg D, Lorigan P, Lebbe C, Jouary T, Schadendorf D, Ribas A, O'Day S, Sosman J, Kirkwood J, Eggermont A, Dreno B, Nolop K, Li J, Nelson B, Hou J, Lee R, Flaherty K, McArthur G. 2011. Improved Survival with Vemurafenib in Melanoma with BRAF V600E Mutation. N Engl J Med 364:2507–2516.

Chari R, Mali P, Moosburner M, Church GM. 2015. Unraveling CRISPR-Cas9 genome engineering parameters via a library-on-library approach. Nat Methods 12:823–826.

Chial H. 2008. Proto-oncogenes to oncogenes to cancer. Nature Education 1:33

Chin L, Andersen JN, Futreal PA. 2011. Cancer genomics: from discovery science to personalized medicine. Nat Med 17:297–303.

Coffee EM, Faber AC, Roper J, Sinnamon MJ, Goel G, Keung L, Wang WV, Vecchione L, Vriendt V de, Weinstein BJ, Bronson RT, Tejpar S, Xavier R, Engelman J, Martin E, Hung K. 2013. Concomitant BRAF and PI3K/mTOR Blockade Is Required for Effective Treatment of BRAFV600E Colorectal Cancer. Clin Cancer Res 19:2688–2698.

Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, Hsu PD, Wu X, Jiang W, Marraffini LA, Zhang F. 2013. Multiplex genome engineering using CRISPR/Cas systems. Science 339:819–823.

Davies H, Bignell GR, Cox C, Stephens P, Edkins S, Clegg S, Teague J, Woffendin H, Garnett MJ, Bottomley W, Davis N, Dicks E, Ewing R, Floyd Y, Gray K, Hall S, Hawes R, Hughes J, Kosmidou V, Menzies A, Mould C, Parker A, Stevens C, Watt S, Hooper S, Wilson R, Jayatilake H, Gusterson B, Cooper C, Shiplay J, Hargrave D, Pritchard-Jones K, Maitland N, Chenevix-Trench G, Riggins G, Bigner D, Palmieri G, Cossu A, Flanagan A, Nicholson A, Ho J, Leung S, Yuen S, Weber B, Seigler H, Darrow T, Paterson H, Marais R, Marhsall C, Wooster R, Stratton M, Futreal P. 2002. Mutations of the BRAF gene in human cancer. Nature 417:949–954.

Dewey FE, Grove ME, Pan C, Goldstein BA, Bernstein JA, Chaib H, Merker JD, Goldfeder RL, Enns GM, David SP, Pakdaman N, Ormond KE, Caleshu C, Kingham K, Klein TE, Whirl-Carrillo M, Sakamoto K, Wheeler MT, Butte AJ, Ford JM, Boxer L, Ioannidis JP, Yeung AC, Altman RB, Assimes TL, Snyder M, Ashley EA, Quertermous T. 2014. Clinical interpretation and implications of whole-genome sequencing. JAMA 311:1035–1045.

Doench JG, Fusi N, Sullender M, Hegde M, Vaimberg EW, Donovan KF, Smith I, Tothova Z, Wilen C, Orchard R, Virgin HW, Listgarten J, Root D. 2016. Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. Nat Biotechnol 34:184–191.

Doench JG, Hartenian E, Graham DB, Tothova Z, Hegde M, Smith I, Sullender M, Ebert BL, Xavier RJ, Root DE. 2014. Rational design of highly active sgRNAs for CRISPR-Cas9-mediated gene inactivation. Nat Biotechnol 32:1262–1267.

Esvelt KM, Mali P, Braff JL, Moosburner M, Yaung SJ, Church GM. 2013. Orthogonal Cas9 proteins for RNA-guided gene regulation and editing. Nat Methods 10:1116–1121.

Falini B, Bolli N, Shan J, Martelli MP, Liso A, Pucciarini A, Bigerna B, Pasqualucci L, Mannucci R, Rosati R, Gorello P, Diverio D, Roti G, Tiacci E, Cazzaniga G, Biondi A, Schnittger S, Haferlach T, Hiddemann W, Martelli MF, Gu W, Mecucci C, Nicoletti I. 2006. Both carboxy-terminus NES motif and mutated tryptophan(s) are crucial for aberrant nuclear export of nucleophosmin leukemic mutants in NPMc+ AML. Blood 107:4514–4523.

Falini B, Mecucci C, Tiacci E, Alcalay M, Rosati R, Pasqualucci L, La Starza R, Diverio D, Colombo E, Santucci A, Bigerna B, Pacini R, Pucciarini A, Liso A, Vignetti M, Fazi P, Meani N, Pettirossi V, Saglio G, Mandelli F, Lo-Coco F, Pelicci PG, Martelli MF; GIMEMA Acute Leukemia Working Party. 2005. Cytoplasmic nucleophosmin in acute myelogenous leukemia with a normal karyotype. N Engl J Med 352:254–266.

Federici L, Falini B. 2013. Nucleophosmin mutations in acute myeloid leukemia: A tale of protein unfolding and mislocalization. Protein Sci 22:545–556.

Fernandes MS, Melo S, Velho S, Carneiro P, Carneiro F, Seruca R, Sofia Fernandes M, Melo S, Velho S, Carneiro P, Carneiro F, Seruca R. 2016. Specific inhibition of p110α; subunit of PI3K: putative therapeutic strategy for KRAS mutant colorectal cancers. Oncotarget 7:68546-68558.

Fritsch C, Huang A, Chatenay-Rivauday C, Schnell C, Reddy A, Liu M, Kauffmann A, Guthy D, Erdmann D, De Pover A, Furet P, Gao H, Ferretti S, Wang Y, Trappe J, Brachmann SM, Maira SM, Wilson C, Boehm M, Garcia-Echeverria C, Chene P, Wiesmann M, Cozens R, Lehar J, Schlegel R, Caravatti G, Hofmann F, Sellers WR. 2014. Characterization of the Novel and Specific PI3Kα Inhibitor NVP-BYL719 and Development of the Patient Stratification Strategy for Clinical Trials. Mol Cancer Ther 13:1117–1129.

Gebler C*, Lohoff T*, Paszkowski-Rogacz M, Mircetic J, Chakraborty D, Camgoz A, Hamann MV, Theis M, Thiede C, Buchholz F. 2017. Inactivation of Cancer Mutations Utilizing CRISPR/Cas9. J Natl Cancer Inst 109: djw183.

Gorre ME, Mohammed M, Ellwood K, Hsu N, Paquette R, Rao PN, Sawyers CL. 2001. Clinical Resistance to STI-571 Cancer Therapy Caused by BCR-ABL Gene Mutation or Amplification. Science 293:876–880.

Grisendi S, Mecucci C, Falini B, Pandolfi PP. 2006. Nucleophosmin and cancer. Nat Rev Cancer 6:493–505.

Hanahan D, Weinberg RA. 2000. The Hallmarks of Cancer. Cell 100:57–70.

Hanahan D, Weinberg RA. 2011. Hallmarks of Cancer: The Next Generation. Cell 144:646–674.

Heckl D, Kowalczyk MS, Yudovich D, Belizaire R, Puram RV, McConkey ME, Thielke A, Aster JC, Regev A, Ebert BL. 2014. Generation of mouse models of myeloid malignancy with combinatorial genetic lesions using CRISPR-Cas9 genome editing. Nat Biotechnol 32:941–946.

Hingorani SR, Jacobetz MA, Robertson GP, Herlyn M, Tuveson DA. 2003. Suppression of BRAFV599E in Human Melanoma Abrogates Transformation. Cancer Res 63:5198–5202.

Horlbeck MA, Witkowsky LB, Guglielmi B, Replogle JM, Gilbert LA, Villalta JE, Torigoe SE, Tjian R, Weissman JS. 2016. Nucleosomes impede Cas9 access to DNA in vivo and in vitro. eLife 5:e12677, DOI: 10.7554/eLife.12677.

Hou Z, Zhang Y, Propson NE, Howden SE, Chu L-F, Sontheimer EJ, Thomson JA. 2013. Efficient genome engineering in human pluripotent stem cells using Cas9 from Neisseria meningitidis. Proc Natl Acad Sci 110:15644–15649.

Howlader N, Noone A, Krapcho M, Miller D, Bishop K, Altekruse S, Kosary C, Yu M, Ruhl J, Tatalovich Z, Mariotte A, Lewis D, Chen HS, Feuer EJ, Cronin KA (eds). 2016. SEER Cancer Statistics Review, 1975-2013, National Cancer Institute. Bethesda, MD, https://seer.cancer.gov/csr/1975_2013/, based on November 2015 SEER data submission, posted to the SEER web site, April 2016.

Hsu PD, Scott DA, Weinstein JA, Ran FA, Konermann S, Agarwala V, Li Y, Fine EJ, Wu X, Shalem O, Cradick TJ, Marraffini LA, Bao G, Zhang F. 2013. DNA targeting specificity of RNA-guided Cas9 nucleases. Nat Biotechnol 31:827–832.

Huang C-H, Mandelker D, Schmidt-Kittler O, Samuels Y, Velculescu VE, Kinzler KW, Vogelstein B, Gabelli SB, Amzel LM. 2007. The Structure of a Human p110α/p85α Complex Elucidates the Effects of Oncogenic PI3Kα Mutations. Science 318:1744–1748.

Iliakis G, Wang H, Perrault AR, Boecker W, Rosidi B, Windhofer F, Wu W, Guan J, Terzoudi G, Pantelias G. 2004. Mechanisms of DNA double strand break repair and chromosome aberration formation. Cytogenet Genome Res 104:14–20.

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

Kandoth C, McLellan MD, Vandin F, Ye K, Niu B, Lu C, Xie M, Zhang Q, McMichael JF, Wyczalkowski MA, Leiserson MDM, Miller CA, Welch J, Walter M, Wendel M, Ley T, Wilson R, Raphael B, Ding L. 2013. Mutational landscape and significance across 12 major cancer types. Nature 502:333–339.

Karimova M, Beschorner N, Dammermann W, Chemnitz J, Indenbirken D, Bockmann J-H, Grundhoff A, Lüth S, Buchholz F, Schulze Zur Wiesch J, Hauber J. 2015. CRISPR/Cas9 nickase-mediated disruption of hepatitis B virus open reading frame S and X. Sci Rep 5:13734.

Kaufman HL, Kohlhapp FJ, Zloza A. 2015. Oncolytic viruses: a new class of immunotherapy drugs. Nat Rev Drug Discov 14:642–662.

Kim E, Koo T, Park SW, Kim D, Kim K, Cho H-Y, Song DW, Lee KJ, Jung MH, Kim S, Kim JH, Kim JH, Kim JS. 2017. In vivo genome editing with a small Cas9 orthologue derived from Campylobacter jejuni. Nat Commun 8:14500.

Kleinstiver BP, Prew MS, Tsai SQ, Topkar VV, Nguyen NT, Zheng Z, Gonzales APW, Li Z, Peterson RT, Yeh J-RJ, Aryee MJ, Joung JK. 2015. Engineered CRISPR-Cas9 nucleases with altered PAM specificities. Nature: 523:481-5.

Kleivi K, Teixeira MR, Eknæs M, Diep CB, Jakobsen KS, Hamelin R, Lothe RA. 2004. Genome signatures of colon carcinoma cell lines. Cancer Genet Cytogenet 155:119–131.

Kong D, Yamori T, Yamazaki K, Dan S. 2014. In vitro multifaceted activities of a specific group of novel phosphatidylinositol 3-kinase inhibitors on hotspot mutant PIK3CA. Invest New Drugs 32:1134–1143.

Lee JT, Li L, Brafford PA, Eijnden M van den, Halloran MB, Sproesser K, Haass NK, Smalley KSM, Tsai J, Bollag G, Herlyn M. 2010. PLX4032, a potent inhibitor of the B-Raf V600E oncogene, selectively inhibits V600E-positive melanomas. Pigment Cell Melanoma Res 23:820–827.

Lekomtsev S, Aligianni S, Lapao A, Bürckstümmer T. 2016. Efficient generation and reversion of chromosomal translocations using CRISPR/Cas technology. BMC Genomics 17:739.

Li H, Zeng J, Shen K. 2014. PI3K/AKT/mTOR signaling pathway as a therapeutic target for ovarian cancer. Arch Gynecol Obstet 290:1067–1078.

Ma H, Wu Y, Dang Y, Choi J-G, Zhang J, Wu H. 2014. Pol III Promoters to Express Small RNAs: Delineation of Transcription Initiation. Mol Ther Nucleic Acids 3:e161.

Mali P, Yang L, Esvelt KM, Aach J, Guell M, DiCarlo JE, Norville JE, Church GM. 2013. RNA-guided human genome engineering via Cas9. Science 339:823–826.

Mao M, Tian F, Mariadason JM, Tsao CC, Lemos R, Dayyani F, Gopal YNV, Jiang Z-Q, Wistuba II, Tang XM, Bornman WG, Bollag G, Mills G, Powis G, Desai J, Gallick G, Davies M, Kopetz S. 2013. Resistance to BRAF Inhibition in BRAF-Mutant Colon Cancer Can Be Overcome with PI3K Inhibition or Demethylating Agents. Clin Cancer Res 19:657–667.

Marco-Sola S, Sammeth M, Guigó R, Ribeca P. 2012. The GEM mapper: fast, accurate and versatile alignment by filtration. Nat Methods 9:1185–1188.

Mayer IA, Abramson V, Formisano L, Balko JM, Estrada MV, Sanders M, Juric D, Solit D, Berger MF, Won H, Li Y, Cantley LC, Winer E, Arteaga C. 2016. A Phase Ib Study of Alpelisib (BYL719), a PI3Kα-specific Inhibitor, with Letrozole in ER+/HER2-Negative Metastatic Breast Cancer. Clin Cancer Res Clin Cancer Res 23:26-34.

Moreno-Mateos MA, Vejnar CE, Beaudoin J-D, Fernandez JP, Mis EK, Khokha MK, Giraldez AJ. 2015. CRISPRscan: designing highly efficient sgRNAs for CRISPR-Cas9 targeting in vivo. Nat Methods: 12:982-8.

Paez JG, Jänne PA, Lee JC, Tracy S, Greulich H, Gabriel S, Herman P, Kaye FJ, Lindeman N, Boggon TJ, Naoki K, Sasaki H, Fujii Y, Eck M, Sellers W, Johnson B, Meyerson M. 2004. EGFR Mutations in Lung Cancer: Correlation with Clinical Response to Gefitinib Therapy. Science 304:1497–1500.

Pao W, Miller VA, Politi KA, Riely GJ, Somwar R, Zakowski MF, Kris MG, Varmus H. 2005. Acquired Resistance of Lung Adenocarcinomas to Gefitinib or Erlotinib Is Associated with a Second Mutation in the EGFR Kinase Domain. PLOS Med 2:e73.

Preto A, Figueiredo J, Velho S, Ribeiro AS, Soares P, Oliveira C, Seruca R. 2008. BRAF provides proliferation and survival signals in MSI colorectal carcinoma cells displaying BRAFV600E but not KRAS mutations. J Pathol 214:320–327.

Qi W, Shakalya K, Stejskal A, Goldman A, Beeck S, Cooke L, Mahadevan D. 2008. NSC348884, a nucleophosmin inhibitor disrupts oligomer formation and induces apoptosis in human cancer cells. Oncogene 27:4210–4220.

Ran F, Cong L, Yan WX, Scott DA, Gootenberg JS, Kriz AJ, Zetsche B, Shalem O, Wu X, Makarova KS, Koonin EV, Sharp PA, Zhang F. 2015. In vivo genome editing using Staphylococcus aureus Cas9. Nature 520:186–191.

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

Ran F, Hsu P, Wright J, Agarwala V, Scott DA, Zhang F. 2013a. Genome engineering using the CRISPR-Cas9 system. Nat Protoc 8:2281–2308.

Ranganathan V, Wahlin K, Maruotti J, Zack DJ. 2014. Expansion of the CRISPR–Cas9 genome targeting space through the use of H1 promoter-expressed guide RNAs. Nat Commun 5:4516.

Sakuma T, Nishikawa A, Kume S, Chayama K, Yamamoto T. 2014. Multiplex genome engineering in human cells using all-in-one CRISPR/Cas9 vector system. Sci Rep 4:5400.

Shayegi N, Kramer M, Bornhäuser M, Schaich M, Schetelig J, Platzbecker U, Röllig C, Heiderich C, Landt O, Ehninger G, Thiede C, Study Alliance Leukemia (SAL). 2013. The level of residual disease based on mutant NPM1 is an independent prognostic factor for relapse and survival in AML. Blood 122:83–92.

Sithanandam G, Kolch W, Duh F, Rapp U. 1990. Complete coding sequence of a human B-raf cDNA and detection of B-raf protein kinase with isozyme specific antibodies. Oncogene 5:1775–1780.

Srivastava M, Nambiar M, Sharma S, Karki SS, Goldsmith G, Hegde M, Kumar S, Pandey M, Singh RK, Ray P, Natarajan R, Kelkar M, De A, Choudjary B, Raghavan S. 2012. An inhibitor of nonhomologous end-joining abrogates double-strand break repair and impedes cancer progression. Cell 151:1474–1487.

Statistisches Bundesamt Wiesbaden. 2015. Todesursachen in Deutschland - Fachserie 12 Reihe 4 - 2014 - Todesursachen2120400147004.pdf. Published 12/15/2015. [Updated 01/06/2016, accessed: 09/16/2016] URL: https://www.destatis.de/DE/Publikationen/Thematisch/Gesundheit/Todesursachen/Todesursachen2120400147004.pdf?__blob=publicationFile.

Stuart D, Sellers WR. 2009. Linking somatic genetic alterations in cancer to therapeutics. Curr Opin Cell Biol 21:304–310.

Sumimoto H, Miyagishi M, Miyoshi H, Yamagata S, Shimizu A, Taira K, Kawakami Y. 2004. Inhibition of growth and invasive ability of melanoma by inactivation of mutated BRAF with lentivirus-mediated RNA interference. Oncogene 23:6031–6039.

Turke AB, Zejnullahu K, Wu YL, Song Y, Dias-Santagata D, Lifshits E, Toschi L, Rogers A, Mok T, Sequist L, Lindeman NI, Murphy C, Akhavanfard S, Yeap BY, Xiao Y, Capelletti M, Iafrate AJ, Lee C, Christensen JG, Engelman JA, Jänne PA. 2010. Preexistence and Clonal Selection of MET Amplification in EGFR Mutant NSCLC. Cancer Cell 17:77–88.

Unniyampurath U, Pilankatta R, Krishnan MN. 2016. RNA Interference in the Age of CRISPR: Will CRISPR Interfere with RNAi? Int J Mol Sci 17:291.

Vogelstein B, Papadopoulos N, Velculescu VE, Zhou S, Diaz LA, Kinzler KW. 2013. Cancer Genome Landscapes. Science 339:1546–1558.

Volinia S, Hiles I, Ormondroyd E, Nizetic D, Antonacci R, Rocchi M, Waterfield MO. 1994. Molecular Cloning, cDNA Sequence, and Chromosomal Localization of the Human Phosphatidylinositol 3-Kinase p110α (PIK3CA) Gene. Genomics 24:472–477.

Wan PTC, Garnett MJ, Roe SM, Lee S, Niculescu-Duvaz D, Good VM, Project CG, Jones CM, Marshall CJ, Springer CJ, Barford D, Marais R. 2004. Mechanism of Activation of the RAF-ERK Signaling Pathway by Oncogenic Mutations of B-RAF. Cell 116:855–867.

Weinstein IB. 2002. Cancer. Addiction to oncogenes--the Achilles heal of cancer. Science 297:63–64.

Whyte DB, Holbeck SL. 2006. Correlation of PIK3Ca mutations with gene expression and drug sensitivity in NCI-60 cell lines. Biochem Biophys Res Commun 340:469–475.

Yi S, Wen L, He J, Wang Y, Zhao F, Zhao J, Zhao Z, Cui G, Chen Y. 2015. Deguelin, a selective silencer of the NPM1 mutant, potentiates apoptosis and induces differentiation in AML cells carrying the NPM1 mutation. Ann Hematol 94:201–210.

Zetsche B, Gootenberg JS, Abudayyeh OO, Slaymaker IM, Makarova KS, Essletzbichler P, Volz SE, Joung J, Oost J van der, Regev A, Koonin EV, Zhang F. 2015. Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. Cell 163:759–771.

Zhao Y, Adjei AA. 2014. Targeting Oncogenic Drivers. In: Peters S, Stahel RA, (eds). Progress in Tumor Research. Vol 41. Basel: S. KARGER AG, pp 1–14.

Appendix

Own Contribution

I designed and conducted following experiments of the presented work on my own or in collaboration with members of the laboratory.

For the traffic light reporter assay, I selected sgRNAs from the cancer mutation set. Subsequently I cloned sgRNA and reporter plasmids and tested the sgRNA activity by fluorescence-activated cell sorting in HeLa cells. Furthermore, I performed statistical analysis of the traffic light reporter results.

For the knock-out of oncogenes, I cloned sgRNAs into the lentiviral vector system and infected the cancer cell lines. Then I conducted proliferation assays and cell cycle analysis by fluorescence activated cell sorting and I performed T7 assays. Moreover, I was involved in the sample preparation for the deep sequencing experiments.

All original text from the publication (Gebler et al., 2017) is put in quotation marks. All original figures and tables from this publication are named Figure 1, 2, etc. and Table 1, 2, etc., respectively.

Additional Figures and Tables, which are not originated from the publication, are named Supplementary Figure 1, 2, etc. and Supplementary Table 1, 2, etc., respectively.

Acknowledgement

First of all, I appreciate all the present and past Buchholz Lab members for creating an inspiring and helpful atmosphere, what made me even more enthusiastic and curious about my own project and far beyond. Many thanks for the introduction to the world of science, great advice and for all the cake!

I like to express my deep gratitude to Prof. Dr. Frank Buchholz, who gave me the opportunity to work on this project. He was sometimes even more inquiring than me and had many cutting-edge ideas to advance my project.

My special thanks to Dr. Jovan Mirčetić for being a great supervisor, for being more patient than I was and for sharing his huge knowledge with me.

I am very thankful to my co-authors Tim Lohoff, Maciej Paszkowski-Rogacz, Jovan Mirčetić, Debojyoti Chakraborty, Aylin Camgöz, Martin Hamann, Mirko Theis, Christian Thiede and Frank Buchholz for their scientific contribution and cooperation to publish my first paper.

My great thanks to the members of my Thesis Advisory Committee, Prof. Dr. Frank Buchholz, Prof. Dr. Evelin Schröck and PD Dr. Sebastian Brenner for their valuable contribution of ideas throughout the whole project duration.

I am grateful to the Deutsche Krebshilfe (DKH) “Mildred-Scheel-Doktorandenprogramm” for the financial support of this project. Thanks to the Else-Kröner-Promotionskolleg for supporting me scientifically and ideationally.

I was particularly fortunate to participate in the International Biology Olympiad as a high school student and got to know the huge “Olympiad family”. My great thanks to all of you for being the initial spark to enlighten my passion for research, especially to Arne Jahn and Dr. Dennis Kappei.

I deeply thank my friends Maren, Manuel, Antje, Miriam and Max. You always had doors and ears open for my sorrows, but also made me forget them.

Endless thanks to my family. You supported me in every way I can imagine to keep me walking my way and to reach my goals.

Thank you, Philipp, for giving selfless love and encouragement and for keeping me grounded.

Anlage 1

Technische Universität Dresden
Medizinische Fakultät Carl Gustav Carus
Promotionsordnung vom 24. Juli 2011

Erklärungen zur Eröffnung des Promotionsverfahrens

1. Hiermit versichere ich, dass ich die vorliegende Arbeit ohne unzulässige Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe; die aus fremden Quellen direkt oder indirekt übernommenen Gedanken sind als solche kenntlich gemacht.

2. Bei der Auswahl und Auswertung des Materials sowie bei der Herstellung des Manuskripts habe ich Unterstützungsleistungen von folgenden Personen erhalten:

Tim Lohoff, Maciej Paszkowski-Rogacz, Jovan Mirčetić, Debojyoti Chakraborty, Aylin Camgöz, Martin Hamann, Mirko Theis, Christian Thiede, Frank Buchholz

3. Weitere Personen waren an der geistigen Herstellung der vorliegenden Arbeit nicht beteiligt. Insbesondere habe ich nicht die Hilfe eines kommerziellen Promotionsberaters in Anspruch genommen. Dritte haben von mir weder unmittelbar noch mittelbar geldwerte Leistungen für Arbeiten erhalten, die im Zusammenhang mit dem Inhalt der vorgelegten Dissertation stehen.

4. Die Arbeit wurde bisher weder im Inland noch im Ausland in gleicher oder ähnlicher Form einer anderen Prüfungsbehörde vorgelegt.

5. Die Inhalte dieser Dissertation wurden in folgender Form veröffentlicht:

Gebler*, Lohoff* *et al*. 2017. Inactivation of Cancer Mutations Utilizing CRISPR/Cas9. J Natl Cancer Inst 109(1): djw183
Journal of the National Cancer Institute (ISSN 0027-8874)
Impact Faktor (2016): 12.589
Fachgebiet Onkologie: Rang 9 von 217

6. Ich bestätige, dass es keine zurückliegenden erfolglosen Promotionsverfahren gab.

7. Ich bestätige, dass ich die Promotionsordnung der Medizinischen Fakultät der Technischen Universität Dresden anerkenne.

8. Ich habe die Zitierrichtlinien für Dissertationen an der Medizinischen Fakultät der Technischen Universität Dresden zur Kenntnis genommen und befolgt.

Ort, Datum

Unterschrift des Doktoranden

Anlage 2

Hiermit bestätige ich die Einhaltung der folgenden aktuellen gesetzlichen Vorgaben im Rahmen meiner Dissertation

 das zustimmende Votum der Ethikkommission bei Klinischen Studien, epidemiologischen Untersuchungen mit Personenbezug oder Sachverhalten, die das Medizinproduktegesetz betreffen
Aktenzeichen der zuständigen Ethikkommission

…….entfällt……..

 die Einhaltung der Bestimmungen des Tierschutzgesetzes
Aktenzeichen der Genehmigungsbehörde zum Vorhaben/zur Mitwirkung

…….entfällt……..

☒ die Einhaltung des Gentechnikgesetzes

Projektnummer 55-8811.71/197
Lab project no. 9 "Application of programmable nucleases such as ZNFs, TALENs and Cas9 to activate, disrupt and edit genes in mammalian cell lines and bacteria“

☒ die Einhaltung von Datenschutzbestimmungen der Medizinischen Fakultät und des Universitätsklinikums Carl Gustav Carus.

Ort, Datum

Unterschrift des Doktoranden