TECHNISCHE UNIVERSITÄT DRESDEN

Universitätsklinikum Carl Gustav Carus

Identification and inactivation of cancer driver mutations using the CRISPR-Cas9 system

Shady Ahmed Atteya Sayed

Aus dem Universitäts KrebsCentrum (UCC), Medizinische Systembiologie
Leiter: Prof. Dr. Frank Buchholz

Identification and inactivation of cancer driver mutations using the CRISPR-Cas9 system

DISSERTATIONSSCHRIFT

zur Erlangung des akademischen Grades

Doctor of Philosophy (Ph.D.)

vorgelegt

der Medizinischen Fakultät Carl Gustav Carus

der Technischen Universität Dresden

von

Shady Sayed, M.Sc.

aus Kairo

Dresden 2020

1. Gutachter: Prof. Dr. Frank Buchholz, Technische Universität Dresden

2. Gutachter: Prof. Dr. Dirk Lindemann, Technische Universität Dresden

Tag der mündlichen Prüfung: June 9th, 2021

Vorsitzender der Promotionskommission: Prof. Dr. Henning Morawietz

To my parents

Zusammenfassung

Somatische Mutationen sind eine Hauptursache für die Entstehung von Krebs. Allerdings tragen nicht alle Mutationen gleichermaßen zur Tumorentstehung bei. Ein wichtiges Ziel der personalisierten Medizin ist es daher, die für das Wachstum und Überleben des Tumors wesentlichen (sogenannte „Treiber"-Mutationen) von den zahlreichen biologisch neutralen Mutationen (sogenannte „Passagier"-Mutationen) zu unterscheiden. In der vorliegenden Studie etablierte ich einen CRISPR-basierten, genetischen Screen mit dessen Hilfe die funktionelle Rolle von Mutationen bei Krebs untersucht werden kann. Ich konnte nachweißen, dass diese mutationsselektive Strategie geeignet ist, um neue Krebstreibermutationen in der Kolorektalkarzinom- Zelllinie RKO zu identifizieren. Dazu verwendete ich 100 unterschiedliche sgRNAs, welche jeweils eine Krebsmutationssequenz spezifisch schneiden während die Wildtyp-Sequenz nicht verändert wird. Als Kontrolle nutzte ich die Kolorektalkarzinom- Zelllinie HCT116, welche die Zielmutationen nicht trägt. Interessanterweise ergab die Datenanalyse, dass zwei sgRNAs, welche die gleiche Mutation (*UTP14A: S99del*) schneiden, besonders rasch und ausschließlich in RKO-Zellen verloren gingen. Im Einklang mit den Screening-Ergebnissen führte die individuelle Infektion der Zellen mit diesen sgRNAs zu einem selektiven Verlust in RKO-, nicht aber HCT-Zellen, wodurch *UTP14A: S99del* als mutmaßliche Treiber-Mutation in RKO-Zellen identifiziert werden konnte. Die weitere Validierung und Charakterisierung dieser mutmaßlichen Treiber-Mutation wird diskutiert. Insgesamt zeigt dieser Ansatz, dass ein solches CRISPR-basiertes System ein leistungsfähiges Werkzeug auch für umfangreichere Untersuchungen von Krebsmutationen darstellt.

Parallel dazu setzte ich die CRISPR-Cas-Technologie ein, um bekannte und bisher nicht therapierbare Treiber-Mutationen, wie z.B. innerhalb der Ras-Onkogen-Familie, zu untersuchen. Bemerkenswert ist in diesem Zusammenhang, dass jeder dritte Krebspatient ein durch Mutationen aktiviertes *KRAS* exprimiert, welches damit das am häufigsten mutierte Onkogen in menschlichen Tumorzellen ist. Im Gegensatz zu anderen Molekülen des *MAPK*-Signalweges konnte *KRAS* bisher nicht mittels kleiner, inhibitorischer Moleküle inaktiviert werden. Unter diesen Voraussetzungen birgt ein genomischer, CRISPR-basierter Ansatz das Potenzial, eine dringend benötigte therapeutische Alternative zur *KRAS*-Inaktivierung zu liefern. Ich entwarf daher drei mutationsselektive sgRNAs abzielend auf die häufigsten KRAS-Mutationen. Obwohl diese Strategie geeignet war, um *KRAS*-mutierte Tumorzellen in 3 unterschiedlichen Krebszelllinien effizient und spezifisch zu entfernen, führte die langfristige Cas9-Expression zur Bildung von onkogenen, resistenten Klonen. Dieses Phänomen wird durch DNA-Doppelstrangbrüche und die nachfolgend einsetzende, endogene DNA-Reparaturmaschinerie begünstigt. Ich konnte zeigen, dass der Adenin-Basen-Editor im

Gegensatz dazu nicht nur in der Lage ist, die *KRAS*-Mutation ohne Doppelstrangbruch zu inaktivieren, sondern diese auch zur Wildtyp-Sequenz reparieren kann. Mit Hilfe dieses Ansatzes erreichte ich insbesondere bei Vorliegen der G12D-Mutation, einen fast vollständigen Abbau der *KRAS*-korrigierten Zellen. Die Validierung in patienten-abgeleiteten *KRAS*-G12D-Organoiden bestätigte die effiziente Korrektur sowie die daraus resultierende erhöhte Sensitivität, wenn auch in einem geringeren Maße als in Zelllinien. Somit konnte in dieser Studie erstmals gezeigt werden, dass Basen-Editierung sowohl in Zelllinien als auch in Organoiden, welche aus Tumorzellen der Patienten stammen, erfolgreich eingesetzt werden kann. Darüber hinaus ist dieses System gut verträglich und induziert weder in Zelllinien noch in Organoiden bei Vorliegen des *KRAS* Wildtyps unerwünschte Nebeneffekte (sogenannte „Off-target-Effekte“). Langfristig kann die Anwendung von CRISPR-basierten- und Basen-Editierungs-technologien zum Ausschalten von *KRAS*-Mutationen nicht nur zu einem besseren Verständnis der *RAS*-Biologie führen, sondern zusammen mit neuen Verabreichungsformen und Technologien die Grundlage für eine dringend benötigte *KRAS*-Therapie bilden.

Summary

DNA mutations are a major cause of cancer development. Identification of all mutations present in cancer cells has been enabled by modern advances in genome sequencing technologies. However, not all mutations contribute equally to tumorigenesis. Distinguishing the mutations essential for tumor growth and/or survival (so- called ‘‘driver’’ mutations) from the numerous biologically neutral mutations (so-called “passenger” mutations) is an important goal for precision medicine with substantial impact on disease diagnosis, prognosis, and treatment regimens. A classic knock-out experiment to study mutations in cancer cells was often thwarted by the complexity of the systems used. The CRISPR-Cas9 system emerged as a rather simple and programmable gene editing tool. Adapting the system to functionally test the relevance of mutations for cancer cell growth and viability would therefore offer a significant advance in the field. Furthermore, sharpening the tool to disable cancer driver mutations, especially those which have been notoriously “undruggable”, shall provide a stepping-stone into an era of therapeutic genome surgery. In this study, I harnessed the power of genetic screens through CRISPR-based system to interrogate the functional role of mutations in cancer. As a proof-of-concept, I employed a mutation-selective strategy to identify novel cancer driver mutations in the colorectal carcinoma (CRC) cell line RKO, using 100 high-quality sgRNAs to specifically cleave cancer mutation sequences, while sparing the wild-type sequence. As a control, the CRC cell line HCT116 that does not carry the targeted mutations was also infected. Interestingly, the analysis revealed two sgRNAs targeting the same mutation (*UTP14A*: S99del) to be rapidly depleted only in RKO cells. In line with screen results, individual infection of the two sgRNAs resulted in the selective depletion in RKO cells, but not in HCT116, nominating *UTP14A*: S99del as a putative driver mutation in RKO cells. Further validation and characterization of this putative RKO driver mutation is discussed. Overall, our approach demonstrates that the CRISPR-Cas9 system is a powerful tool to functionally dissect cancer mutations at large-scale.

In parallel, I employed the CRISPR-Cas9 system to target known cancer driver mutations that are notoriously “undruggable” such as the Ras oncogene family. One in every three cancer patients has a mutationally-activated *KRAS* making it the most frequently mutated oncogene in human cancers. However, in contrast to other members of the *MAPK* pathway, small molecules directed at targeting *KRAS* have not been successful. Therefore, a genomic CRISPR-based approach holds the potential to offer a much-needed therapeutic alternative to inactivate *KRAS*. To this aim, I designed three mutation-selective sgRNAs targeting three of the most common *KRAS* mutations, and infected cells in conjunction with Cas9 nuclease and GFP. Although targeting mutant *KRAS* depleted cancer cells efficiently and specifically, tested

in three different cancer cell lines, long-term Cas9 expression led to the formation of oncogenic escape clones, a caveat fostered by DNA double-stranded break and endogenous DNA repair machinery. In contrast to CRISPR-Cas9 nucleases, a recently developed CRISPR base editor system enables predefined nucleotide exchanges in genomic sequences without generating DNA double stranded breaks. Using the cleavage-deficient Adenine Base Editor, we not only inactivated *KRAS* driver mutations, but also repaired it back to wild-type sequence. Intriguingly, *KRAS*-corrected cells were rapidly depleted, almost in its entirety especially in G12D mutant cells. Validation in Patient-derived *KRAS* G12D mutant organoids revealed efficient base editing and sensitivity, albeit, to lesser extent than that of cell lines. Conclusively, this study demonstrates- for the first time- that Base Editing could be used for the correction of cancer driver mutations in established cell lines and in patient-derived organoids. Moreover, the system is well-tolerated and did not induce off-target toxicities in the *KRAS* wild-type background, both in cancer cell lines and in patient-derived organoids. On the long run, the application of CRISPR and base editing technology for targeting *KRAS* mutations can be used not only for a better understanding of RAS biology but, together with new delivery approaches and technologies, might lay the foundation for a much-needed anti-*KRAS* CRISPR-based therapy.

Table of Contents

List of Figures

List of Tables

Abbreviations

ABE	Adenine Base Editing
Amp	Ampicillin
BAC	Bacterial Artificial Chromosome
bp	Base pair of DNA
Cas9	CRISPR-associated protein
CBE	Cytosine Base Editing
Cm	Chloramphenicol
COSMIC	Catalogue on Somatic Mutations in Cancer
CRC	Colorectal Cancer
CRISPR	Clustered regularly interspaced short palindromic repeats
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleoside triphosphate
dsDSB	double- stranded DNA break
E. coli	*Escherichia coli*
FACS	Fluorescence-activated cell sorting
gDNA	Genomic DNA
GFP	Green fluorescent protein
ICGC	International Cancer Genome Consortium
Kb	Kilobase pairs of DNA
KRAS	Kirsten RAt Sarcoma viral oncogene
LTR	Long terminal repeat
MOI	Multiplicity of infection
NGS	Next generation sequencing
PAM	Protospacer adjacent motif
PCR	Polymerase chain reaction
PDAC	Pancreatic ductal adenocarcinoma
sgRNA	Single guide RNA
siRNAs	small interfering RNAs
TCGA	The Cancer Genome Atlas
UTP14A	U3 Small Nucleolar RNA-associated Protein 14 Homolog A

1 Introduction

Human beings and other animals have had cancer throughout recorded history. Perhaps the oldest description of cancer (although the word cancer was not used) was found in the Edwin Smith Papyrus, discovered in Egypt around 3000 BC. The writing describes 8 cases of tumors or ulcers of the breast that were removed by cauterization and says about the disease, “There is no treatment”. Today, Cancer presents a real global threat accounting for more than 8.5 million deaths per year, a death toll that is more than AIDS, tuberculosis and malaria combined (IARC). In 2018, the global burden stood at 17.0 million new cancer cases and 9.6 million deaths and is expected to grow by 2040, to 27.5 million new cancer cases and 16.3 million deaths due to the growth and aging of the population. In Germany, nearly a quarter of men and 20% of women die of cancer, and it is estimated that around 51% men and 43% women will develop cancer during their lifetime (Quante et al., 2016).

From a biological perspective, all cancers share a common pathogenesis as they develop as a result of changes that have occurred in the DNA sequence of the genomes of cancer cells. Early insights into the central role of the genome in cancer development emerged in the late nineteenth and early twentieth centuries from studies by David von Hansemann (Hansemann, 1890) and Theodor Boveri (Boveri, 1914) They observed the presence of bizarre chromosomal aberrations while examining dividing cancer cells. This led to the proposal that cancers are abnormal clones of cells characterized and caused by abnormalities of hereditary material. Following the discovery of DNA as the molecular substance of inheritance (Avery et al., 1944) and determination of its structure (Watson & Crick, 1953), earlier speculations was supported by the demonstration that agents which damage DNA and generate mutations also cause cancer (reviewed in Loeb & Harris, 2008). Subsequently, increasingly refined analyses of cancer cell chromosomes showed that specific and recurrent genomic abnormalities, such as the translocation between chromosomes 9 and 22 (known as the ‘Philadelphia’ translocation) are associated with chronic myeloid leukaemia (Rowley, 1973). A few years later, it was demonstrated that introduction of total genomic DNA from human cancers into phenotypically normal NIH3T3 cells could transform them into cancer cells (Krontiris & Cooper, 1981; Shih et al., 1981). Isolation of the specific DNA fragment responsible for this transforming activity led to the identification of the first naturally occurring, human cancer-causing sequence mutation—the single base G>T transversion that causes a glycine to valine substitution in codon 12 of the HRAS gene (Reddy et al., 1982; Tabin et al., 1982). This seminal discovery in 1982 inaugurated an era of vigorous searching for the abnormal genes underlying the development of human cancer that continues today.

A turning point in cancer research came about with the Human genome project completion in April 2003. Sequencing the human genome enabled researchers to determine, with base pair resolution, the order of nucleotides which make up the human genome, paving the way to identify genetic variants critical in the development of human cancers. Several cancer genomics programs such as The Cancer Genome Atlas (TCGA), the Cancer Genome Project (CGP), and the International Cancer Genome Consortium (ICGC) have boosted the production of cancer genomics data through massively parallel next-generation sequencing (NGS) of cancer genomes. To name a few of these accomplishments, in the past decade, the entire repertoire of human exons has been sequenced in glioblastoma (Parsons et al., 2008), pancreatic (Sondka et al., 2018), which offers a valuable resource for studying driver genes in the cancer of individual patients, having an eye for the application of precision medicine (Garay & Gray, 2012).

1.1 Driver vs Passenger mutations

A major challenge in cancer research is the identification of cellular drug targets whose inhibition leads to the selective killing of cancer cells, while sparing healthy cells. The number of driver mutations required to generate a tumor has been a long-standing question in cancer (Armitage & Doll, 1954; Martincorena & Campbell, 2015; Nordling, 1953). By definition, mutations that provide a selective growth advantage, or promote cancer cell survival, are termed driver mutations, while those that are merely accompanying mutations with no effect on cancer development are termed passenger mutations (Stratton et al., 2009). Often times, the terms driver and passenger are used to refer to the genes harboring driver mutations. Genes that have been identified as drivers in at least one cancer type are described as cancer genes (Vogelstein et al., 2013). Of note, oncogenes are defined as driver genes in which driver mutations are activating or result in new functions. Meanwhile, tumor suppressors are driver genes in which driver mutations are typically inactivating. Oncogenes tend to be affected by copy number amplifications or missense mutations at a limited number of codons, whereas tumor suppressors tend to be affected by focal deletions, or nonsense, frameshift, and splice-site mutations dispersed across the gene (Gonzalez-Perez et al., 2013).

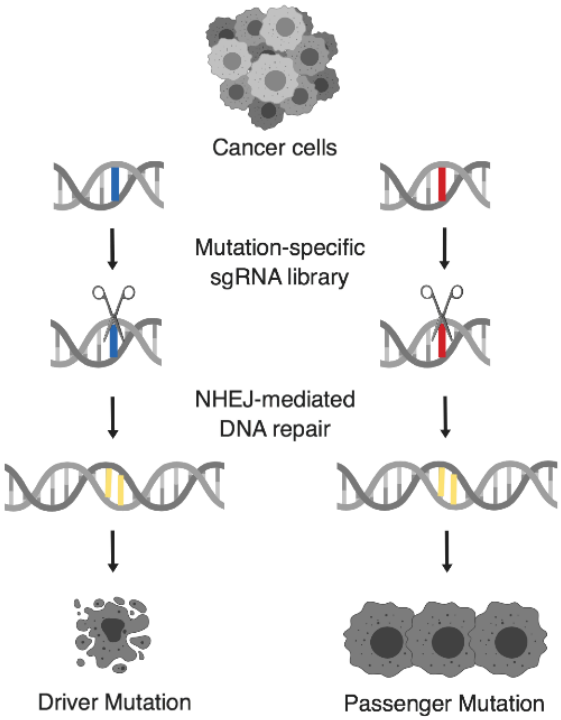


Figure 1 Functional profiling of cancer mutations. While driver mutations are those essential for cancer cell survival, passenger mutations are of no biological relevance. Therefore, mutation knock-out is only detrimental to targeted cells bearing a driver mutation. This method disregards mutations in tumor suppressor genes.

While mathematical modelling estimates that 5 to 8 driver mutations are required for cancer development (Stratton et al., 2009), recent sequencing consortia reveals 3 to 4 driver mutations to be sufficient (Brown et al., 2019; Iranzo et al., 2018; Tomasetti et al., 2015). However, the number of passenger mutations far exceeds the number of driver mutations in any given tumor sample and scientists and clinicians alike go by the assumption that mutations in non-cancer genes are passenger mutations. Moreover, cancers which harbor mutations in DNA repair genes can tolerate extreme levels of hypermutation, evidenced by many hundreds of mutations with every cell division (Shlien et al., 2015), which further complicates the search for driver mutations. Conventionally, there are two categories of approaches for identifying driver mutations: those that assess the frequency of mutations and those that predict the functional impact of mutations (Getz et al., 2007). Both approaches can be applied to find novel driver mutations, genes or pathways, however, both have their limitations. For the former, given the variability in mutation rates between cancer types (>1000-fold variation), between samples of the same cancer type (~1000-fold variation), and across the genome itself (>5-fold variation), it is often difficult to accurately estimate background mutation frequencies (Lawrence et al., 2013). Moreover, frequency-based approaches, naturally, fail to identify driver genes infrequently mutated (Gonzalez-Perez & Lopez-Bigas, 2012). Unlike frequency-

based approaches, function-based approaches can identify candidate driver mutations using data from a single sample. Moreover, a few methods consider mutations in specific functional domains (Florian et al., 2013), while others consider features specific to certain protein families such as kinases (Torkamani & Schork, 2008). However, one must keep caution with generalizations drawn from these methods as not all mutations in well-conserved domains are drivers and not all in poorly conserved domains are passengers. Notably, out of eight functional impact prediction methods, none had greater than 81% accuracy (Florian et al., 2013). Sensitivities ranged from 40% to 79% and specificities from 57% to 99%.

Given the different advantages and drawbacks of frequency- versus function-based approaches, it is not surprising that candidate driver lists are overlapping only to a limited extent. For instance, when the frequency-based algorithm MuSiC (Dees et al., 2012) and the functional impact tool OncodriveFM (Gonzalez-Perez & Lopez-Bigas, 2012) were used to identify novel drivers in 3,205 samples form 12 cancer types, both methods identified similar number of candidate driver genes (232 and 259, respectively), but only 68 of those candidate driver genes were in common (Tamborero et al., 2013). More recently, the IntOGen-mutations platform was developed to run multiple frequency- and function-based methods in a single pipeline and produced high-confidence candidate drivers (Gonzalez-Perez et al., 2013). Nevertheless, the resulting high-confidence list contained 165 candidate driver genes that were not in the Cancer Gene Census, demonstrating that integrating frequency- and function-based approaches can suggest novel drivers. Furthermore, the list contained a total of 291 genes, far fewer than the 522 in COSMIC, suggesting it may also contain less false positives.

Recently, studies incorporating quantitative measures of selection in cancer evolution have provided robust statistical approaches to evaluating the impact of various mutations on the development of cancer using dN/dS ratio (Greenman et al., 2006). The dN/dS ratio was originally developed for the analysis of genetic sequences from divergent species. It quantifies the mode and strength of selection by comparing nonsynonymous substitution rates (dN), which are exposed to selection as they change the amino acid composition of a protein with synonymous substitution rates (dS)—assumed to be neutral. Adapting a dN/dS model to identify novel cancer driver mutations has the advantage that it accounts for the context-dependent mutation spectrum in addition to estimating the background mutation rate of each gene separately using synonymous mutations. Moreover, the mutation rate is estimated locally, thus inherently correcting for the variation in mutation rate across the genome, and differences in read depth across the genes examined. Genes under positive selection can be identified, and the number of driver mutations can be quantified from the excess of non-synonymous mutations (Martincorena & Campbell, 2015). These approaches revealed a

distinct pattern of selection universally shared across cancer types suggesting on average ~4 coding mutations per tumor under positive selection and highlighted a limited impact of negative selection on cancer genome, unlike species evolution (Martincorena et al., 2017). Moreover, dN/dS ratios analysis showed that half of the driver substitutions reside in yet-to-be-discovered cancer genes (Anandakrishnan et al., 2019; Darbyshire et al., 2019; L. Jiang et al., 2019; Martincorena et al., 2017).

1.2 CRISPR-Cas9 as a genome editing tool

Bioinformatic methods cannot provide definitive classification of mutations as drivers or passengers but can prioritize them for functional testing (Gonzalez-Perez et al., 2013). The gold standard of evidence that a mutation is a driver is that the mutation produces a cellular phenotype that contributes a selective advantage to the cells harboring it. Large-scale genomic screening is a powerful technology capable of detecting driver mutations. Therefore, reagents such as short interfering RNAs (siRNAs), short hairpin RNAs (shRNAs) and exogenous plasmids have been widely used to identify and characterize targeted genes through altering their expression. However, it is not always feasible to achieve site-specific modulations at the precision of base pairs using these techniques. The CRISPR-Cas9 technology provides a rapid approach to modify endogenous loci to overcome these limitations (Fig. 2).

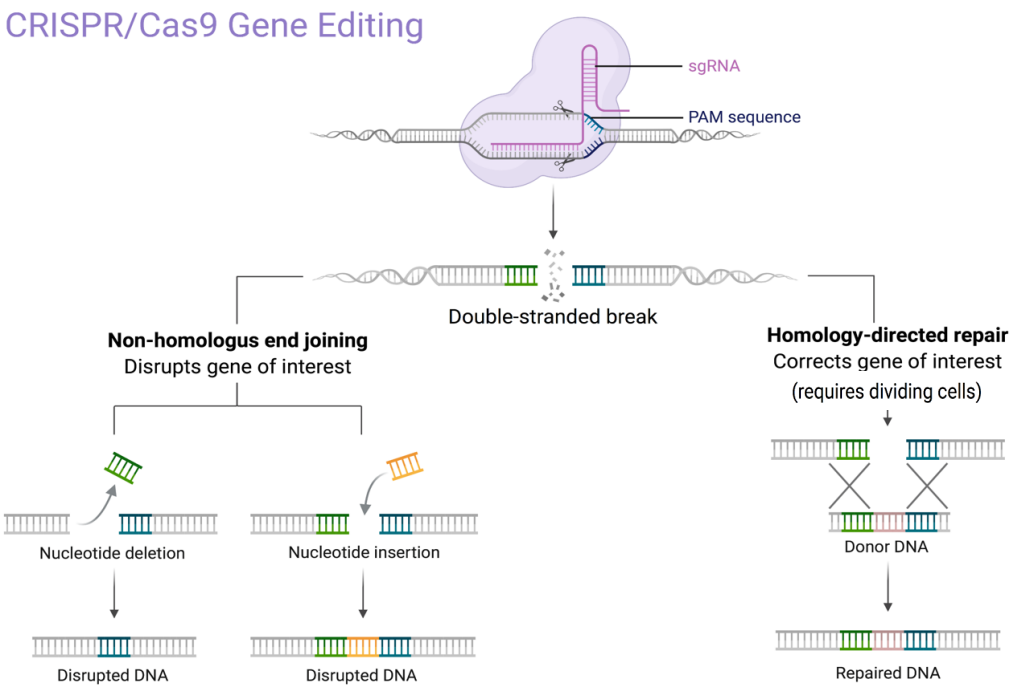


Figure 2 Schematic representation of CRISPR/Cas9 Gene Editing approaches.

The emergence of the Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR)-Cas9 system as a genome editing tool has revolutionized studies of the human genome (Hsu et al., 2014; Nowak et al., 2016; Sander & Joung, 2014; X. Wang et al., 2016). Originally discovered as a prokaryotic RNA-guided adaptive immune system, the CRISPR/Cas9 system has been repurposed for eukaryotic genome editing (Jinek et al., 2012; Mojica et al., 2005), a fantastic achievement that secured the Nobel prize in Chemistry this year. The mechanism of targeting by all DNA-targeting CRISPR-Cas systems described to date requires that a short sequence known as a protospacer-adjacent motif, or PAM, occur near the target DNA site. A CRISPR RNA (crRNA) molecule pairs with a *trans*-activating crRNA (tracrRNA), or a fusion of both crRNA and tracrRNA into a single guide RNA molecule (sgRNA), to guide Cas9 nuclease via sequence complementarity to a specific site, adjacent to a nuclease-specific PAM sequence (Cong et al., 2013; Mali et al., 2013). The PAM sequence for all Cas9 effectors is located directly at 3' end of the protospacer on the non-complementary DNA strand. Cas9 nuclease predominantly makes a blunt-ended Double-Stranded Break (DSB) 3 bp upstream of the PAM (Jinek et al., 2013). For all Cas9 effectors from type-II CRISPR systems, target-site recognition begins with binding of the Cas9-guide RNA ribonucleoprotein complex to the native PAM sequence, followed by DNA unwinding and concurrent formation of an RNA-DNA heteroduplex between the sgRNA spacer sequence and the target DNA strand which leaves the non-target DNA strand exposed and accessible to other molecules, forming a single-stranded DNA called "R-loop" (Szczelkun et al., 2014).

After R-loop formation, Cas9 undergoes conformational changes that result in the activation of its nuclease domains (F. Jiang et al., 2016; Jinek et al., 2014; Sternberg et al., 2014). These conformational changes are obstructed by mismatches between the target strand and guide RNA spacer, thus limiting nuclease activation to sequences that are of high complementarity to the guide RNA spacer (Sternberg et al., 2019). After nuclease activation, the DNA phosphodiester backbone is hydrolyzed by Cas9's two distinct nuclease domains: the HNH nuclease domain, which cleaves the guide RNA–bound target DNA strand; and the RuvC-like nuclease domain, which cleaves the PAM-containing non-target DNA strand. Importantly, mutating either of the nuclease domains produces a Cas9 nickase (an enzyme that cleaves only one of the DNA strands), while inactivation of both nuclease domains generates a catalytically-dead Cas9 (dCas9) (Cong et al., 2013; Nishimasu et al., 2014). Nickases are particularly useful for base editors and prime editors, both CRISPR-based advancements that precisely edit DNA without requiring the formation of DSBs or homology-directed repair, and will be discussed later (Anzalone et al., 2019; Gaudelli et al., 2017; Komor et al., 2016; Rees & Liu, 2018). Meanwhile, dCas9 can be exploited for various applications without permanent

DNA alteration ranging from transcriptional regulation (Dominguez et al., 2016; Shalem et al., 2015) to epigenetic modifications (Adli, 2018; Pickar-Oliver & Gersbach, 2019; Thakore et al., 2016).

Since the initial reports of programmed DNA cleavage by Cas9 nuclease came from *Streptococcus pyogenes* (*Sp*Cas9) *in vitro* (Jinek et al., 2012) and in mammalian cells (Cho et al., 2019; Cong et al., 2013; Jinek et al., 2013; Mali et al., 2013), several Cas9 variants have been discovered and adapted for genome editing, including orthologs from *Staphylococcus aureus* (Ran et al., 2015), *Streptococcus thermophilus* (Esvelt et al., 2013; Müller et al., 2016), *Neisseria meningitidis* (Edraki et al., 2019; Hou et al., 2013), *Campylobacter jejuni* (E. Kim et al., 2019) and many other organisms (Chatterjee et al., 2018; Harrington et al., 2017; Hirano et al., 2016). These Cas9 effectors differ in their overall size, native PAM sequences, guide RNA architecture, optimal spacer length, editing efficiency and specificity. For example, *Sp*Cas9, currently the most widely used CRISPR-Cas nuclease and the variant primarily used throughout this study, contains 1,368 amino acids, recognizes a relatively common NGG PAM, can be used with either an sgRNA or crRNA/tracrRNA fusion pair, functions optimally with 20-nt spacers, has robust DNA targeting and cleavage activity, and, depending on the spacer composition, supports relatively high levels of off-target editing (Adli, 2018; D. Kim et al., 2019; Komor et al., 2017). Some Cas9 variants offer particular advantages over *Sp*Cas9, such as smaller sizes (for example, SaCas9 is 1,053 amino acids) (Ran et al., 2015) or pyrimidine-rich PAMs (for example, Nme2Cas9) (Edraki et al., 2019; Hou et al., 2013).

CRISPR-Cas nucleases are most commonly used to efficiently and selectively disrupt target gene sequences (Sternberg et al., 2015). In most mammalian cells, Cas9-induced DSBs are most often repaired by two major DNA repair pathways, namely, End-joining including the canonical Non-Homologous End Joining (NHEJ), and Microhomology-Mediated End Joining (MMEJ), both of which typically create small insertion or deletion (indel) products at the cut site leading to disruption of the coding sequence (Ciccia & Elledge, 2010). On the other hand, a more conservative, albeit less efficient repair process is Homology-Directed Repair (HDR), which occurs mostly in dividing cells as it requires proteins that are expressed predominantly in the S and G2 cell-cycle phases. It is based on using DNA templates for repair at the cut site, or through using the sister chromatid as a template thus allowing the correction of dysfunctional elements (Chapman et al., 2012). If a perfect end-joining repair regenerates the starting sequence, it will nevertheless remain a substrate for subsequent nuclease cleavage, however, end-joining resulting in the insertion or deletion of nucleotides at the break site will prevent subsequent recognition and re-cutting by the nuclease (Brinkman et al., 2018). Therefore, if

nucleases are targeted to open reading frames, indel products after end-joining usually generate frameshift mutations in coding sequences that abrogate protein function (Shalem et al., 2014; T. Wang et al., 2014). Importantly, the mixture of insertion and deletion (indel) products that results from DSBs cannot (as of yet) be controlled, but they are not entirely random, and recent studies have shown that it can be predicted using machine-learning modelling (Allen et al., 2019; W. Chen et al., 2019; Leenay & Beisel, 2017; Shen et al., 2018; van Overbeek et al., 2016). In some cases, end-joining produces high yields of a single desired product (Iyer et al., 2019), particularly when the break site lies in a region of microhomology.

Leveraging endogenous DNA repair machinery to generate gene knockouts on a genomic scale is the cornerstone of CRISPR-based genetic engineering and has led to a myriad of applications for the CRISPR-Cas9 system, where HDR is exploited to insert exogenous DNA sequences into the genome, such as markers and corrective sequences. However, this approach is often hampered by the low efficiency of HDR (D. B. T. Cox et al., 2015; DeWitt et al., 2016). Another application strategy of CRISPR-Cas9 technology comprises Loss-of function studies with indel-prone NHEJ repair with a typical deletion spectrum of 1–10 base pairs leading to frameshift mutations to induce gene inactivation (Hsu et al., 2013; Ran et al., 2013a). Given the simplicity and high versatility of the system, numerous laboratories have adopted the NHEJ-mediated pathway to screen essential genes where all the sgRNAs comprising a library are synthesized, cloned and delivered to a pooled population of cells. Subsequently, the edited cells with selected phenotypic change can be separated from the entire pooled population using either positive or negative selection. Consequently, several genome-scale libraries for pooled CRISPR screens have been generated and allow for high-throughput and/or multi-loci studies (Miles et al., 2016; Peng et al., 2015; Sanjana, 2017). In the past few years, several pooled CRISPR screens have focused on gene inactivation for *in vitro* and *in vivo* applications) (S. Chen et al., 2015; Hart et al., 2015; Koike-Yusa et al., 2014; Kiessling et al., 2016; Parnas et al., 2015; J. Shi et al., 2015; Tzelepis et al., 2016; Wallace et al., 2016; T. Wang et al., 2014; Wong et al., 2016; R. Zhang et al., 2016).

1.3 Genome editing with base editors

Base editors are a recent addition to the CRISPR toolbox and are particularly useful when targeted point mutations without DSBs or donor DNA templates are required (Gaudelli et al., 2017; Komor et al., 2016; Molla & Yang, 2019). Guided by a sgRNA molecule, Cas9 directs the deaminase to install a transition mutation in a ~5-nucleotide region within the protospacer target sequence. Two classes of base editors have been developed to date: cytosine base editors (CBEs), which catalyze the conversion of C•G to T•A base pairs; and adenine base editors (ABEs), which catalyze A•T to G•C conversions. CBEs and ABEs can efficiently mediate all four possible transition mutations (C→T, A→G, T→C, G→A), which represent approximately 30% of currently annotated human pathogenic variants (Landrum et al., 2016). For ABE, TadA, an adenine deaminase from *Escherichia coli* that works on single-stranded RNA, was evolved to function on the single-stranded DNA loop created by Cas9, while avoiding double-stranded off-target sites. Current base editors consist of a catalytically impaired Cas9 nuclease fused to a single-stranded DNA deaminase enzyme and, in some cases, to proteins that manipulate DNA repair machinery. Several base editors have been described since the initial two reports, showing altered editing efficiencies with different Cas variants (T. P. Huang et al., 2019; Gaudelli et al., 2020; Koblan et al., 2018; Zafra et al., 2018). The first described ABE, ABE7.10, showed improved base editing efficiency when a cleavage-deficient Cas nickase was used, rather than a dCas DNA-binding protein, to nick the non-deaminated DNA strand. The resulting nick stimulates cellular repair machinery to use the deaminated strand as a template for resynthesizing the nicked strand. Deamination of one strand and re-synthesis of the complementary strand therefore results in editing of both target DNA strands to yield stable conversion of the target base pair (Komor et al., 2016). Additional improvements include linker optimization, Nuclear localization sequence and codon optimization yielded the ABEmax variants, which offer increased editing efficiencies in mammalian cells and *in vivo* (Koblan et

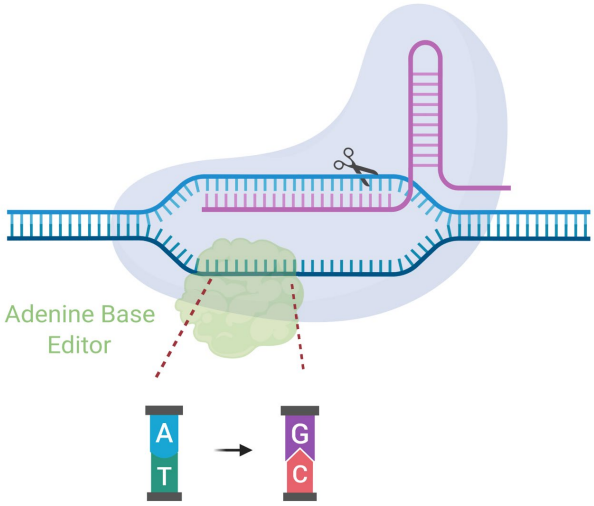


Figure 3 Adenine Base Editor (ABE) A•T to G•C base editing strategy. ABEs contain a deoxyadenosine deaminase, and a catalytically impaired Cas9. They bind target DNA in a guide RNA-programmed manner, exposing a small bubble of single-stranded DNA. The deoxyadenose domain catalyses conversion of adenine to inosine within this bubble. Following DNA repair or replication, inosine is subsequently corrected to guanine thereby the original A•T base pair is replaced with a G•C base pair at the target site.

al., 2018; Zafra et al., 2018). Mechanistically, The R-loop that is formed after the Cas domain binds its target provides a single-stranded DNA substrate for the deaminase domain of a base editor. Base editing efficiency within this R-loop is determined by productive interactions between R-loop substrate nucleotides and the deaminase enzyme (Thuronyi et al., 2019). The efficiency with which target bases are edited typically peaks around the most accessible nucleotides within the R-loop (Thuronyi et al., 2019). The location of the editing window can change when a different Cas domain is used or when the deaminase domain changes (T. P. Huang et al., 2019; Y. B. Kim et al., 2017; Richter et al., 2020). However, it is clear that broader editing windows also come with an increased risk of bystander editing.

Just as CRISPR-Cas9 system, base editing requires a particular set of criteria: i) An editing window (range of editable 4-5 Target Adenine or Cytosine nucleobases) which requires ii) a suitable PAM that is roughly 12-18 bp distant. Moreover, off-target potential is influenced by several factors, such as the presence of another cytosine or adenine within the editing window, nucleotide sequence context, and R-loop accessibility by deaminases, that are not satisfied for all Cas nuclease-dependent off-target sites. Of note, the transient formation of R-loops by Cas9 melting of off-target DNA may also lead to off-target base editing at sites that are not bona fide nuclease off-target loci (D. Kim et al., 2019; Liang et al., 2019). This fact has been exploited by detection methods such as CIRCLE-seq (Akcakaya et al., 2018; Tsai et al., 2017), an effective assay for identifying Cas-dependent off-target base editing candidate loci. A few approaches have already been trialed to curb base editors off-target effects such as using Cas domains with enhanced DNA specificity and sgRNA truncations as well have been reported to limit off-target base editing (J. K. Lee et al., 2018). Moreover, limiting exposure of cells to base editors by delivering base editor RNPs or mRNA, also decreases off-target base editing while maintaining on-target editing, likely due to more rapid action of base editors at on-target loci than off-target loci (Rees & Liu, 2018). ABE7.10 has proven to be less tolerant of Cas domains beyond wild-type SpCas9 than CBEs, although several ABE7.10 Cas variants have been reported with at least mild activity in mammalian cells. Recently, the incompatibility of ABEs with some Cas domains was overcome through the evolution of ABE8 deaminases. Phage assisted non-continuous and continuous evolution (PANCE and PACE) has enabled the development of TadA variants with increased deaminase activity, namely TadA-8e (Richter et al., 2020). ABE8e kinetic and structural data suggest that it catalyzes DNA deamination up to ~1100-fold faster than earlier ABEs because of eight new mutations that stabilize DNA substrates in a constrained, transfer RNA–like conformation (Lapinaite et al., 2020). Furthermore, ABE8e supports a much broader set of Cas domains compared to ABE7.10, although the higher activity of ABE8 variants may necessitate the use of Cas domain variants

that decrease Cas-dependent off-target DNA editing or the use of deaminase variants described that reduce Cas-independent off-target DNA and RNA editing (for example, ABE8e V106W or ABE8.17-m V106W) (Gaudelli et al., 2020; Richter et al., 2020). Base editors have been applied in a variety of cell types and organisms including cell lines, organoids, and animal models of human genetic diseases, to install or revert transition point mutations using CBEs (Billon et al., 2017; Chadwick et al., 2018; Kuscu et al., 2017; Levy et al., 2020; Shimatani et al., 2017; B. Ren et al., 2019; Rossidis et al., 2018; Tang & Liu, 2018; Yeh et al., 2018; Zafra et al., 2018) and ABEs (C. Lee et al., 2019; Z. Liu et al., 2018; Sürün et al., 2020; M. Song et al., 2020; Suh et al., 2021).

1.4 CRISPRing *KRAS*: New hope to an old problem

In 1982, the single base G>T substitution that causes a glycine to valine substitution in codon 12 of the *HRAS* gene was identified as the first somatic point mutation in human cancer (Reddy et al., 1982; Tabin et al., 1982). This pivotal work launched a new era of molecular cancer genetics research. Later studies have demonstrated that removal of a *RAS* oncogene from human tumor cell lines or mouse models resulted in reversal of transformation (Chin et al., 1999; Collins et al., 2012; Fisher et al., 2001; Scolnick et al., 1973; Shirasawa et al., 1993). suggesting that these tumors can show *RAS* oncogene addiction (Sharma et al., 2006). RAS genes family (*HRAS*, *KRAS*, and *NRAS*) is the most mutated gene family in cancers, with a mutation frequency of roughly one third of all human cancers and accounting for up to one million deaths per year worldwide (A. D. Cox & Der, 2014). Most of these missense mutations occur in *KRAS* (85%), and less frequently in NRAS (12%) and HRAS (3%) therefore, in this study we will focus on *KRAS*. In mammalian cells, there are two protein products of the *KRAS* gene that result from the alternative use of exon 4 (exon 4A and 4B, respectively): *KRas*4B is ~5 times more abundant than *KRas*4A according to TCGA datasets, however, these proteins differ only in the structure in their C-terminal region and use different mechanisms to localize to the plasma membrane (Welman et al., 2000). Interestingly, out of the 134 different missense mutations identified in *KRAS* in all cancers, 98% of them are found at one of three hotspots, amino acids G12, G13 or Q61 and the frequency of mutation at these three residues, the resulting mutant residue, and the RAS isoform mutated, varies amongst different cancer types (Simanshu et al., 2017). The vast majority of *KRAS* mutations occur in exon2 at residue 12 which is normally occupied by a glycine residue. The mutation of this glycine to anything other than proline results in steric hindrance which prevents GTP hydrolysis and thereby increases levels of the GTP-bound active form (Haigis, 2017). Activating mutations in *KRAS* are capable of inducing transformation and lead to a constitutively active signaling for proliferation, differentiation and metastasis (Fig. 4), making it an attractive therapeutic target (Haigis, 2017).

However, despite more than three decades of intense research efforts, no clinically effective RAS inhibitor has reached a cancer patient and the reasons are manifold. To summarize some of the significant challenges hampering traditional drug discovery approaches: (i) *KRAS* proteins are small proteins (~21kDa, 189 amino acids) and lack pockets large enough for small molecules to bind; (ii) *KRAS* exhibit exceptionally high affinity for GTP leading to an extremely strong binding (the dissociation constants are in the picomolar range) which prevents direct targeting of the nucleotide-binding pocket; (iii) The high intracellular GTP concentrations (>500μM) which impedes competition; and (iv) potential toxicity arising from the off-target inhibition of wild-type *KRAS* proteins (A. D. Cox & Der, 2014; J. M. L. Ostrem & Shokat, 2016). Therefore, current views have shifted away from a pan-*KRAS* inhibitor to an approach supporting a mutation-selective strategy to be a more successful strategy. Although renewed hope has been inspired by the recent development of *KRAS*G12C-specific inhibitors (Christensen et al., 2020; J. M. Ostrem et al., 2013; Patricelli et al., 2016), other *KRAS* mutant alleles such as *KRAS*G12D or *KRAS*G13D, which are more common mutations than *KRAS*G12C, remain “undruggable”.

Functional genomic studies based on shRNA or CRISPR technologies to disrupt the expression of thousands of gene products in hundreds of cancer cell lines have revealed that *KRAS*-mutant cancer cells depend on *KRAS* function for growth and survival (McDonald et al., 2017; Tsherniak et al., 2017). The results highlight a central role for mutant *KRAS* as a causal driving force in a large number of cancers and present an exciting opportunity for therapeutic agents that target *KRAS* to make a profound impact in treating these cancers (Christensen et al., 2020). Moreover, siRNAs that selectively inhibit mutant *KRAS* mRNAs (T. L. Yuan et al., 2014), lead to impeded growth of cancer cells both *in vitro* and *in vivo*, however, their continuous expression/delivery is required to maintain target RNA suppression. Therefore, a CRISPR-Cas9 mutation-selective targeting could be employed for permanent gene disruption, in principle, even when transiently expressed.

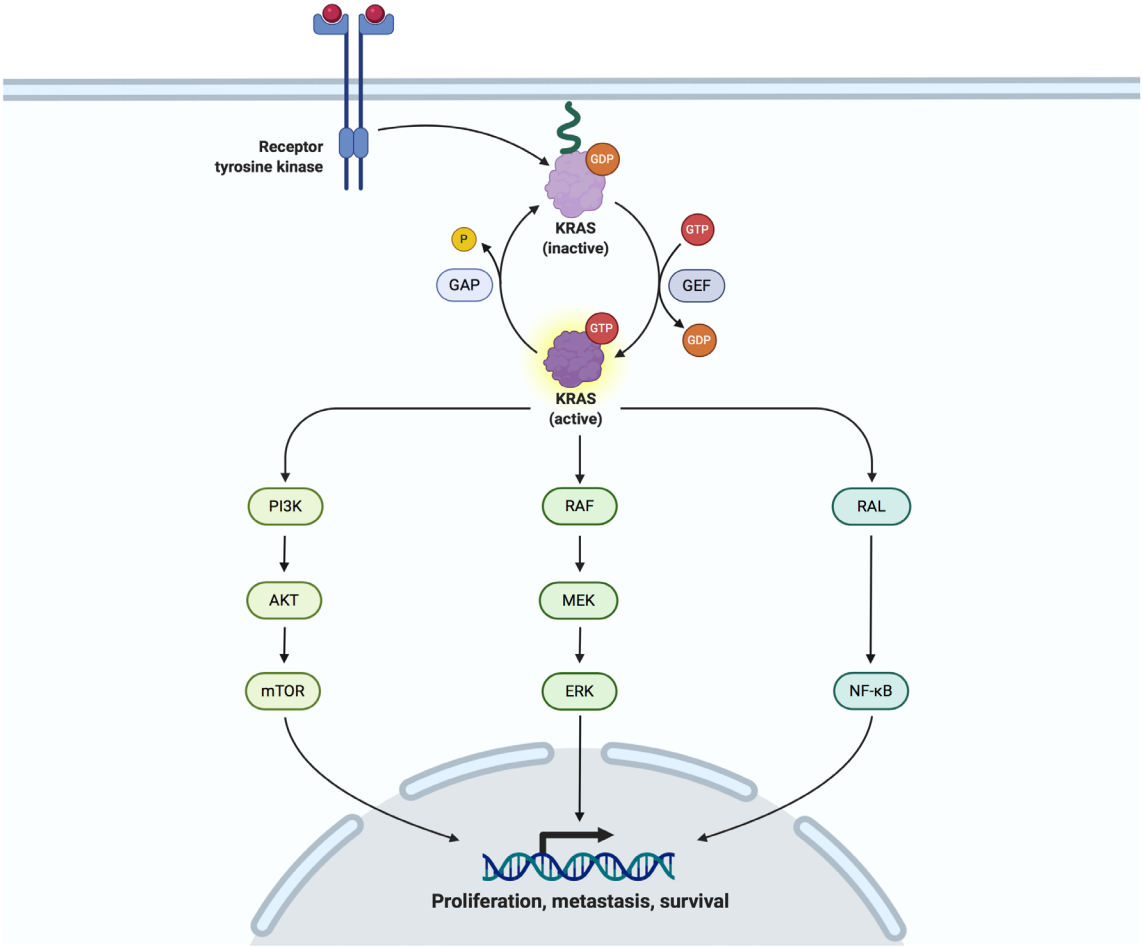


Figure 4 Key players in RAS/MAPK signalling pathway. Upon ligand binding to the extracellular domain of a receptor tyrosine kinase (e.g. EGFR), dimerized activated receptor is phosphorylated and leads to activation of the oncogene KRAS. RAS-selective guanine nucleotide exchange factors (GEFs) and GTP-activating proteins (GAPs) regulate GDP– GTP cycling. Activating mutations, found primarily (>99%) at residues G12, G13, or Q61 disrupt GDP– GTP cycling by impairing intrinsic and GAP-stimulated GTP hydrolysis. Oncogenic KRAS in turn activates intracellular PI3K, RAF or RAL-GEF pathways to promote cell survival, proliferation and metastasis.

Pancreatic ductal adenocarcinoma (PDAC) is the most prevalent neoplastic disease of the pancreas accounting for more than 90% of all pancreatic malignancies (Orth et al., 2019). With above 95% *KRAS* mutation frequency, PDAC is considered the most RAS-addicted of all cancers (Orth et al., 2019). Owing to the late detection and a particularly aggressive biology, PDAC remains a highly devastating disease with no effective therapeutic options and a dismal 5-year survival rate of 6%. *KRAS* mutant allele-specific imbalance have been correlated with poor prognosis in PDAC patients (Krasinskas et al., 2013). Interestingly, *KRAS* mutations (predominantly G12 variants) are detected in the earliest lesions in PDACs and are retained in all metastases (Prior et al., 2012). Therefore, an effective anti-*KRAS* therapeutic strategy is

anticipated to make a significant impact on the treatment of PDAC. Similarly, *KRAS* appears to be the initiating event in lung adenocarcinoma and the common G12C mutation is a hallmark of exposure to tobacco smoke (Fisher et al., 2001; McCormick, 2015). *KRAS* mutations are found in 25%-30% of lung adenocarcinomas, and are of particular importance in disease prognosis and response to targeted therapies (Yu et al., 2015). Likewise, anti-*KRAS* therapy is particularly relevant for colorectal cancer (CRC), the second most common cause of cancer-related death, although *KRAS* mutations are probably not primary initiating events (Haigis et al., 2008) the vast majority of which are thought to be initiated by loss of *APC* or, in mismatch-repair deficient tumors, by mutations in β-catenin signaling. In CRC, other factors such as microsatellite instability, cancer subsite within the colon versus rectum, and age at diagnosis contribute to the disease course and therapeutic response (Domingo et al., 2013; Serebriiskii et al., 2019). Nevertheless, *KRAS* activating mutations have been reported in ~40% of CRCs and studies on cell lines have shown a substantial dependency on *KRAS* mutations for survival (Dwane et al., 2020; Tsherniak et al., 2017). Moreover, preclinical and clinical data show that the specific codon mutated, and the resulting mutant residue, differentially impact the disease prognosis (Arrington et al., 2012), and the response to epidermal growth factor receptor EGFR-directed therapy in patients with CRC (Fiala et al., 2016; Imamura et al., 2012). Of interest, residue G13 mutations are poorly studied and the least understood mutation hotspot compared with G12 and Q61 mutations (Kennedy et al., 2020). Notably, G13D is the third most common *KRAS* mutation in all cancers (13%), and the second most frequent in CRC and is associated with aggressive behavior and poor clinical outcomes (Hobbs et al., 2016).

1.5 Aim of the thesis

Recent advances in genome sequencing technologies provide the means to identify all mutations present in tumor cells. This mutational profile can predict disease prognosis and aid to design the best treatment regimen for each patient. However, in many cases the most important driver mutations remain elusive, rendering decisions on the optimal treatment option ineffective. Having a system that can functionally test the significance of particular mutations for cancer cell growth and viability would therefore be an important advance in the field of precision oncology. In this study, I harness the versatility of CRISPR-Cas9 system for a mutation-selective strategy to identify novel cancer driver mutations in established cancer cell lines. As a readout, I used cellular fitness assayed through a cell competition assay, after mutations knockout. In addition, and in parallel, I employed the system to target known cancer driver mutations that are infamously “undruggable” such as the *KRAS* oncogene. The application of CRISPR-Cas9 for targeting *KRAS* mutations can be used not only for a better understanding of *RAS* biology but, together with new delivery approaches and technologies, may lay a foundation for a much needed anti-*KRAS* CRISPR-based therapy. To sum it up, Discovery of novel cancer driver mutations and inactivating known driver mutations are the two main objectives of my thesis.

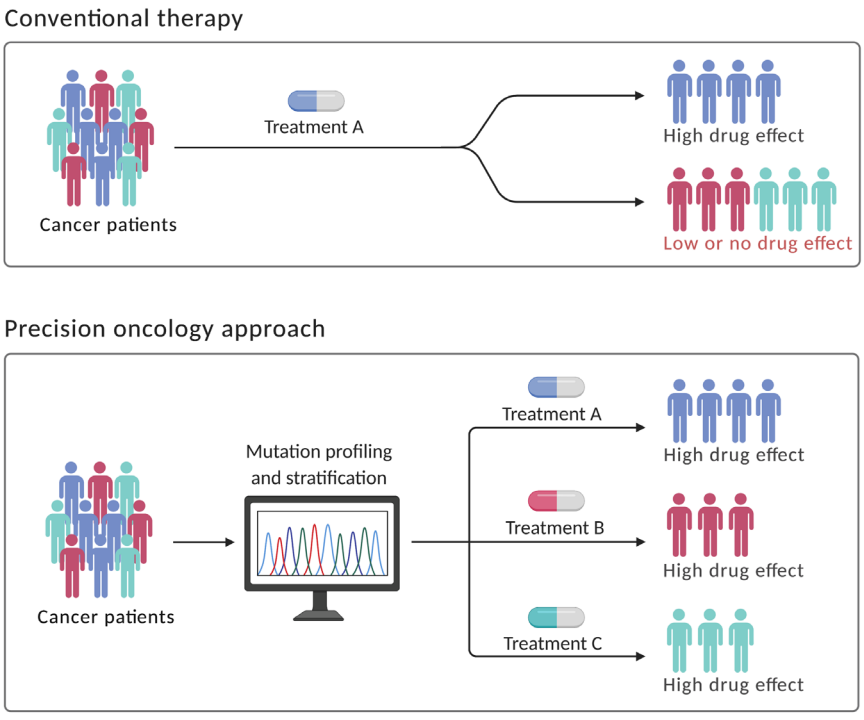


Figure 5 The Vision of Cancer Precision Medicine. Functional profiling of cancer mutations through next-generation sequencing holds the promise to guide medical decisions, and provide personalised treatments to subgroups of patients, instead of a suboptimal one-drug-fits-all model.

2 Methods

2.1 Synthetic oligonucleotides

All Primers/Oligonucleotides used were synthesized by Sigma-Aldrich for standard cloning purposes.

2.2 Recombinant DNA techniques

2.2.1 DNA purification

Column purification: DNA clean-up was performed using ISOLATE II PCR and Gel Kit (Bioline) according to manufacturer's instructions. Purification from agarose gel: The DNA band of interest was excised under a Safe Imager 2.0 (Invitrogen) and transferred to a 1.5 ml reaction tube. DNA was extracted using the ISOLATE II PCR and Gel Kit (Bioline) according to manufacturer's instructions. After the wash and centrifugation, the column was incubated at 65°C open-capped for 5 min to remove residual traces of ethanol. The water used for elution was similarly preheated to 65°C.

2.2.2 Competent cells preparation

On the previous day, 5 ml of LB medium was inoculated by picking a single *E. coli* XL1-Blue colony growing on a stock plate. On the next day, 100 ml LB medium was inoculated with 1 ml of overnight culture and grown for 3.5 hours at 37°C in a standard shaker. Cells were then transferred to 50 ml reaction tubes and incubated on ice for 20 min. All subsequent steps were carried out on ice. Cells were centrifuged for 10 mins at 3500 g, then the pellet was resuspended in 2x 50 ml ice-cold water. Cells were centrifuged for 10 mins at 3500 g, then the pellet was resuspended in 2x 25 ml cold water. Afterwards, cells were centrifuged for 10 min at 3500 g, then the pellet was resuspended in 1x 25 ml cold water. Finally, cells were centrifuged for 10 mins at 3500 g, then the competent cells were resuspended in 300 µl cold water.

2.2.3 DNA ligation

For DNA ligation, roughly 1:3 vector to insert copy number ratio of DNA was used according to NEB ligation calculator (https://nebiocalculator.neb.com/#!/ligation). Typically, 50 ng of the backbone was mixed with appropriate amounts of insert DNA and 2 µl 10x T4 ligase buffer and 1 µl T4 DNA ligase in a total volume of 20 µl. Reactions were incubated for 30 mins at room temperature and inactivated for 10 mins at 65°C. 3 µl of this reaction were transformed

into One Shot MAX Efficiency *DH5α* Competent Cells by heat shock according to manufacturer's instruction. If high transformation efficiency was required, samples were purified by microdialysis where the appropriate number of wells of a 6-well cell culture plate were filled with water and a single MF membrane filter (0.025 µm VSWP, Millipore) added to each well. The sample was carefully added to the floating membrane, left for 20-30 mins, and then transferred to a new 1.5 ml reaction tube for transformation and plating on the relevant antibiotic plate.

2.2.4 Transformation

For transformations by heat-shock, One Shot MAX Efficiency *DH5α* Competent Cells were used for according to the manufacturer's instructions. Typically, 50 µl of competent cells were mixed with not more than 100 ng DNA, incubated on ice for 30 mins, moved to the 42ºC for 90 secs and transferred back on ice for 2 mins. Then, 500 µl SOC medium was added and cells incubated at 37ºC for 1 hour with 600 rpm shaking. After this, all cells were plated on agar plates substituted with suitable antibiotic. Single clones for each construct were transferred into liquid culture and used for DNA extraction. For transformations by electroporation, *E. coli* XL1-Blue competent cells, ligation reactions and Gene Pulser cuvettes (0.1 cm, Bio-Rad) were kept on ice at all times. In a 1.5 ml reaction tube, 50 µl competent cells and 2 µl ligation reaction were mixed. The sample was transferred to the electroporation cuvette and electroporated at 1700 V using an Electroporator 2510 (Eppendorf). When recovery is needed, the sample was resuspended in 1 ml SOC medium, transferred to a 14 ml round-bottom tube (Corning), and incubated for 1 hour in a shaker. In the meantime, LB agar plates containing the appropriate antibiotics were warmed to 37°C. Part or all of the transformation was plated on the plates and incubated at 37°C in a HeraTherm incubator (Thermo Fisher).

2.2.5 PCR

Unless otherwise stated, Standard PCR for was performed with MyTaq Red DNA polymerase (Bioline) in a 50 µl reaction with 16.7 µl water, 10µl 5x MyTaq Red buffer, 1µl of each primer (20µM), 0.5µl MyTaq Red, and 1 µl template. The standard PCR conditions were: 3mins of 95°C, 35 cycles of 15 s 94°C, 20 s 55°C, 30 s 72°C, and finally 5 min 72°C. For DNA cloning and other applications requiring high-fidelity amplification, PCR was performed with Herculase II Fusion DNA polymerase (Agilent) in a 50 µl reaction with 35.5 µl water, 10 µl 5x Herculase buffer, 0.5 µl dNTP mix (100 mM), 1.25 µl each primer (10 mM), 0.5 µl Herculase, and 1 µl template. The standard PCR conditions were: 2 min of 95°C, 30 cycles of 10 s 95°C, 20 s 55°C, 60 s 72°C, and finally 5 min 72°C.

2.2.6 Plasmid DNA preparation

Mini prep: Plasmid DNA was isolated from up to 5 ml of bacterial overnight culture using GeneJET Plasmid Miniprep Kit (Thermo Fisher) according to manufacturer's instructions. DNA was eluted in 30 µl water. Maxi prep: Plasmid DNA was isolated from up to 200 ml of bacterial overnight culture using Plasmid Maxi Kit (QIAGEN) according to manufacturer's instructions. DNA was eluted in 1 ml TE buffer or Nuclease-free water.

2.2.7 Preparation of genomic DNA from mammalian cells

Genomic DNA was isolated using QIAamp DNA Blood Mini Kit (QIAGEN) according to manufacturer's instructions.

2.2.8 Restriction enzyme digestion

Typically, PCR product or plasmid DNA (up to 5 µg) were mixed with 8µl of the enzyme-compatible buffer e.g. 10x CutSmart buffer (New England Biolabs, NEB) and 2µl of the restriction enzyme (unless otherwise stated, purchased from NEB) in a total volume of 80 µl. Samples were incubated at 37°C for 3 hours and subsequently column/gel purified.

2.2.9 DNA gel electrophoresis

PCR products and plasmids were generally resolved on an 0.5%-2% agarose gel supplemented with RedSafe Nucleic Acid Staining Solution (iNtRON Biotechnolobgy) in 1x TBE. Samples which were later gel purified were loaded on gels prepared with Ultra-Pure Agarose. To determine the size of the bands, 1 kb Plus DNA Ladder (NEB) and 6x Orange DNA Loading Dye (NEB) were used. Images of the gels were acquired with a BioVision Imager Gel Documentation System (Vilber Lourmat). If necessary, the desired bands were excised from the gel and DNA was isolated using the LATE II PCR and Gel Kit (Bioline) according to the manufacturer's instructions.

2.2.10 DNA sequence analysis software

For in silico assembly of DNA constructs and analysis of Sanger sequencing reads, the programmes GenomeCompiler (Twist Bioscience) and SnapGene were used. DNA samples for Sanger sequencing were run by Microsynth Seqlab GmbH (Göttingen, Germany), or Eurofins genomics GmbH (Ebersberg, Germany).

2.2.11 TA cloning

TA cloning was performed using the TA Cloning® Kit with pCR™2.1 vector (Thermofisher). The amplicons were separated using agarose gel electrophoresis and the target bands were extracted and purified using the Gel extraction Purification Kit (Bioline). The DNA concentration was measured with Nanodrop spectrophotometer. For TA cloning ligation, 0.1–0.3 pmol of DNA fragments was mixed with pCR™2.1 vector according to manufacturer's protocol, incubated at 16 °C for 3 hours. The ligation product was transformed into competent *E. Coli DH5α* employing blue/white screening. White Colonies were picked the next day and plasmid DNA was extracted via Mini-prep. Finally, colony PCR using MyTaq Red Polymerase and M13 forward and reverse primers followed by Sanger sequencing. Visualisation and alignment of reads using SnapGene® Version 4.3.2 software (GSL BiotechLLC, San Diego, California, United States).

2.2.12 T7 Endonuclease 1 Assay

T7E1 assay for NHEJ mediated mutation rate was performed as previously described (Sayed et al., 2019). Briefly, genomic DNA isolated from infected cells was used as substrate for a targeted UTP14A or *KRAS* PCR. After purification and quantification of PCR product, fragments were then hybridized using a thermocycler, followed by T7 endonuclease digestion and loaded on a 1% agarose gel containing 0,01% RedSafe. The gel was run at 90 V for 30 minutes and visualised BioVision Imager Gel Documentation System (Vilber Lourmat).

2.3 sgRNA Design

sgRNA design tools to generate a knockout and/or to predict the frameshift efficiency of a certain sgRNA have greatly improved in the past few years, I have used some of the online tools such as CHOPCHOP (Montague et al., 2014), https://chopchop.cbu.uib.no), CCTop (Stemmer et al., 2015), https://cctop.cos.uni-heidelberg.de), GPP sgRNA Designer (Doench et al., 2016), https://portals.broadinstitute.org/gpp/public/analysis-tools/sgrna-design) or InDelphi (Shen et al., 2018), http://indelphi.giffordlab.mit.edu), or a combination of all of them. These tools have been instrumental for planning a gene knockout experiment. However, since we adopt a mutation-selective approach targeting point mutations, most of the sgRNAs used target mutations. Therefore, sgRNAs were manually modifed to match the mutant sequence in question and subsequently ranked. Often times, such as for all base editing experiments, sgRNAs differ in only one nucleotide from the wild-type allele sequence, making the wild-type allele the most likely off-target site. Therefore, for each experiment, mutation-selective sgRNAs are delivered to cells which do not carry the mutation, to control for off-target toxicities.

2.4 RKO Screen

To test the feasibility of using the CRISPR-Cas9 system to functionally profile mutations in a cancer cell line. I utilized the CRC cell line RKO and first downloaded the reported 4,762 mutations, of which 4,410 are coding sequence mutations, from the largest catalogue to date on somatic mutations in cell lines is reported on the COSMIC database (http://www.sanger.ac.uk/cosmic/), we used version v80 (from 13th of February 2017) to extract mutations in RKO cells publicly available (ftp://ftp.sanger.ac.uk/pub/CGP/cosmic/) for our analyses. The sequences were then modified to contain the mutated sequences, and searched for potential Cas9 target sites using the Azimuth 2.0 algorithm implemented by (Doench et al., 2014; Doench et al., 2016) We then extracted 30 bp of sequence upstream and downstream of the reported mutations and identified potential sgRNAs in sense and antisense orientation. All sgRNAs with a perfect match in the wild-type hg19 genome were discarded from further analysis steps. An off-target search for obtained 20-bp long sgRNAs was done using an exhaustive short sequence aligner GEM ver. 1.376, with parameters set to report all alignments with up to 4 mismatches. Most of the designed sgRNAs target point mutations, whereas 190 sgRNAs target small indel mutations. Of note, depending on the mismatch position and sequence complexity, Cas9–sgRNA complexes can potentially tolerate 1–5 base pairs mismatches between the protospacer and the target sequence. Therefore, we applied more stringent criteria through filtering out sgRNAs with up to 2 mismatches to the reference genome, especially those residing outside the seed region, distal to the PAM.

Moreover, excluding sgRNAs bearing stretches of at least 4 nt of the same identity, or those containing too high (>70%) or too low (<20%) %GC content (T. Wang et al., 2014). This filtering led to 100 high quality sgRNAs targeting RKO mutations. In addition, I included 5 sgRNAs targeting the MYC proto-oncogene, an RNA polymerase subunit and essential ribosomal proteins as positive controls, as well as 2 sgRNAs targeting LUC as non-targeting negative controls. Transduction of both RKO (target) and HCT116 (control) cell lines was done at low multiplicity of infection (MOI=0,3) to ensure that most cells receive one copy of Cas9-sgRNA complex targeting the mutant sequence, along with GFP as a transduction marker. Sorting GFP positive cells three days post infection was followed by genomic DNA isolation for baseline time point calling. Cell pellets were collected from the sorted population every three days throughout the course of the screen. The number of cells in culture were kept at 1 million cells to fully represent the diversity of the library, which exceeds the recommended 100x coverage (in our case 10,000x), owing to the small library size.

2.5 Plasmids

For expression of Cas9 and sgRNAs from lentiviral vectors we optimized the sequence of pL-CRISPR.EFS.GFP (Addgene plasmid #57818, (Heckl et al., 2014)), where the tracr sequence was modified to increase sgRNA stability and enhance its assembly with Cas9 protein (B. Chen et al., 2013). *Streptococcus pyogenes* Cas9 and GFP were linked via P2A and were expressed from the EFS promoter, whereas the sgRNAs were expressed from the human *U6 pol III* promoter. Protospacers targeting mutations were cloned into pL-CRISPR.EFS.GFP by cloning complementary oligonucleotides into the vector. Unless a guanine was the first base in the protospacer, a guanine was added to the 5' end of the protospacer before cloning to boost the expression of the gRNA from the human U6 promoter. Likewise, for ABE8e timecourses, protospacers were cloned into LRT2B vector (Addgene plasmid #110854, (Zafra et al., 2018)), using BsmBI/BbsI sites following the standard protocol (Ran et al., 2013b). Briefly, oligos for gRNA were phosphorylated and annealed in a 10 µl reaction containing 1 µl of each sgRNA oligo (100 µM), 1 µl 10X T4 ligation buffer and 0.5 µl polynucleotide kinase. The reaction was run in thermocycler at 37°C for 30 mins, 95°C for 5 mins and then slowly brought down to 25°C at 5°C/min rate. The reaction was then diluted 250 times in water and 2 µl was used for digestion/ligation reaction. In this reaction, 100 ng backbone plasmid (e.g. pL-CRISPR.EFS.GFP) was added to annealed oligos together with 2 µl Tango buffer, 1 µl 10 mM DTT, 1 µl BsmBI/BbsI and 0.5 µl T4 DNA ligase in 20 µl final reaction. To digest the vector and ligate oligos into it, the following temperature profile was repeated 6 times in thermocycler: 37°C for 5 mins, 23°C for 5 mins. 2 µl of this final reaction was used to transform *DH5α E. coli* bacteria. Occasionally, incubation at 37°C and a colony PCR was performed, with MyTaq Red Polymerase according to the manufacturer's instructions. Plasmids were sequenced through sequencing service provider Microsynth Seqlab GmbH (Göttingen, Germany), or Eurofins genomics GmbH (Ebersberg, Germany) with forward primer to U6 promoter: 5' GAGGGCCTATTTCCCATGATTCC 3'.

2.6 Cloning of adenine base editors in Lentiviral vectors

The backbone used is pLenti-FNLS-P2A-GFP-PGK-Puro (Addgene #110869, referred to as pL-Backbone), which is a lentiviral vector bearing a codon optimized version of a CBE (BE3) in conjunction with GFP and puromycin resistance gene. The pipeline established here is used for all three adenine base editor variants where CBE in the backbone is exchanged with an ABE variant of choice. Due to unavailable restriction sites, cloning required two subsequent steps where two restriction sites for Acc65I and BstZ17I-HF were first introduced replacing CBE, and on a second step the ABE of choice was cloned in through intermediary Acc65I and BstZ17I restriction sites. First, the CBE cassette was excised, through digestion with BamHI-

HF and NsiI-HF in CutSmart buffer at 37°C for 3 hours. The digestion reaction was run on a 0.8% Ultrapure agarose gel (0,01% RedSafe™) for 30 mins at 90 V and the DNA band corresponding to the linearized pL-Backbone was cut out from the gel and purified with ISOLATE II PCR and Gel kit, following the manufacturer's instructions. To generate the insert with cloning sites for Acc65I and BstZ17I-HF, oligos bearing Acc65I and BstZ17I restriction sites were phosphorylated and annealed, then the ligation was performed at a 5:1 molar ratio of insert to vector using T4 DNA ligase following the manufacturer's instructions. The ligation product was transformed and the bacteria was plated for overnight incubation. Single colonies were picked and sequenced using the *E. coli* NightSeq® service from Microsynth Seqlab GmbH (Göttingen, Germany) with the standard primer EGFP-N-Rev provided by the service. Clones containing the correct insert were grown overnight in 5 mL LB-Ampicillin (50 µg/mL) and were miniprepped according to manufacturer's instructions. To generate the insert, miniprepped and linearized DNA of xCas9-ABE7.10 (Addgene #108382), or NG-Cas9ABEmax (Addgene #124163), or NG-ABE8e plasmid (Addgene #138491) was used as a substrate for PCR using Herculase II Fusion proof-reading DNA Polymerase according to the manufacturer's instructions, using primer pairs bearing restriction sites in their overhangs (Table.4), with an extension time of 3 mins and 30 secs, for 30 cycles. After sequence confirmation of an intermediate new vector bearing a multiple cloning site in place of CBE, miniprepped plasmid DNA together with PCR product of ABEs were sequentially digested with Acc65I and BstZ17I-HF in 10X NEBuffer3.1 and 10X Cutsmart Buffer, respectively. Each digestion took place for for 3 hours at 37°C. Then, ligation was performed at a 5:1 molar ratio of insert to vector using T4 DNA ligase following the manufacturer's suggestions. The ligation product was transformed and the bacteria was plated for overnight incubation. Single clones were picked, grown and miniprepped followed by sequencing by Microsynth Seqlab GmbH (Göttingen, Germany) employing multiple primers aligning the full sequence of the base editor in addition to diagnostic test digests confirming the correct integration. Finally, Plasmid DNA were transformed into *E. coli* DH5α and cells were grown overnight in at 37°C with constant shaking. Plasmid DNA was purified using QIAGEN-tip 20 Maxiprep Kit (Qiagen 10023), according to manufacturer's protocol. The DNA was re-suspended at 1 µg/µl, validated through sequencing and then used for virus production or stored at − 20 °C.

2.7 Cell culture

Cell lines were maintained in the following media: Dulbecco's modified Eagle's medium DMEM (high glucose, GlutaMAX™, Invitrogen, Carlsbad, CA) for HEK293T, RKO, HeLa, HCT116, PANC-1, and A549. All of the media were supplemented with 10% fetal bovine serum (FBS, Gibco) and 1% Penicillin-Streptomycin (Gibco) and kept at 37°C, 5% CO_2. If selection was

necessary, Media was prepared with the following concentrations of antibiotics: 400 µg/ml Geneticin (Gibco), 2 µg/ml Puromycin (Thermofischer), 200 µg/ml Hygromycin B (Life Technologies), Blasticidin HCl 20 µg/ml (Sigma-Aldrich). For all cell lines used, cells were allowed to recover after thawing for two passages, before performing experiments. Most cells lines used needed comparable time for population doublings and were passaged every 2–3 days (1:8 splitting ratio).

2.8 Lentivirus Production and Transduction

Lentiviral particle production and infection were performed as previously published (Sayed et al., 2019). Briefly, fifteen milion HEK293T cells were seeded in T-175 flasks (or 7 million in 10cm dishes) and transfected on the next day at ~80% confluency with pMD2.G (Addgene plasmid #12259), psPAX2 (Addgene plasmid #12260) and pL-CRISPR.EFS.GFP.gRNA using PEI (1 mg/mL). After ~20 hours the medium was changed to complete DMEM and 72 hours post transfection the viral supernatant was collected, filtered through 0.45 µm filter and centrifuged for 2 hours at 100,000xg at 4°C. The supernatant was decanted and the viral pellets resuspended in PBS overnight at 4°C on a shaker. For long-term storage the virus particles were kept in cryovials at -80°C. Where indicated, the virus particles where concentrated using Amicon Ultra-15 Centrifugal Filter Devices according to the manufacturer's instructions. When a new cell line is used for the first time, the amount of virus needed to infect 10%–30% of the cells was determined by titration. Transduction was performed in 96 well plates in the presence of protamine sulfate (final concentration 5 µg/mL, Sigma-Aldrich) and spin-infected for 1 hour at 1,000 g and 37° C. As with any transduction, the quality of the plasmid DNA is paramount. Plasmid maxiprep kits provided transfection-level DNA of a higher quality compared to miniprep or midiprep kits. The amount of DNA used for virus production is also an important parameter, since higher transfection efficiency and dosage usually yield better quality virus and higher genome-targeting efficiency.

2.9 Flow Cytometry

Throughout the study I aimed a moderate infection rate to create an internal competition between infected and non-infected cells. Transduced cells were further cultured and analyzed by MACSQuant VYB Analyzer (Miltenyi Biotec and FlowJO software) for GFP/TdTomato expression. Reduction of the percentage of GFP/TdTomato positive cells indicates that the infected cells expressing the nuclease-sgRNA complex have a growth disadvantage in comparison to the non-infected cells.

RKO, PANC-1, A549, and HCT116 were transduced cells in 96 well plates and analyzed for GFP or TdTomato expression. 24 hours post transduction, infected cells were washed and fresh medium was replenished. 72 hours post transduction, cells were trypsinized and collected for flow cytometry analysis. Viable single cells were first gated using the forward and side scatter and the GFP fluorescence signal was measured using a 488 nm laser while TdTomato using a 561 nm yellow laser and the log area of the signal is collected. FACS sorting was carried out using a BD FACS Aria cell sorter (BD Biosciences). For time course experiments adherent cells were processed every 48-72 hours by flow cytometry in technical triplicates. Briefly, for each time point, medium was aspirated completely using a multichannel pipette. Then, cells were washed 200µl sterile PBS and treated with 30µl/well trypsin enough to cover the adherent layer of cells. Cells were incubated for 5 mins at 37 °C. Next, 170µl full medium was added directly on cells, to obtain a homogenized cells solution by pipetting up and down vigorously. Then, 30µl was transferred to a new 96 well plate for later acquisitions, topped up to 200µl with prewarmed full medium. Of the remaining 170µl cells solution, 50µl was used for FACS acquisition. The rest was used for gDNA isolation and/or freezing.

2.10 Nanopore sequencing

MinION sequencing was performed on a R9.4.1 MinION Flow Cell following the manufacturer's protocol for 1D Amplicon by ligation (Oxford Nanopore technologies, SQK-LSK108, version: ADE_9003_v108_revT_18Oct2016). Briefly, genomic DNA from RKO cells treated with sgRNA targeting mutant *UTP14A* was isolated five days after infection. Next, the *UTP14A* locus flanking the mutation was amplified using Phusion® High-Fidelity DNA Polymerase (M0530), followed by PCR purification using a GeneJET PCR purification Kit (Thermo Scientific), following the manufacturer's protocol. End repair and dA-tailing (NEBNext Ultra II End-Repair/dA-tailing Module) was then performed on 200 fmol input DNA by adding 7 µl Ultra II End-Prep buffer, 3 µl Ultra II End-Prep enzyme mix, and 5 µl Nuclease-free water. The mixture was incubated at 20 °C for 5 min and 65 °C for 5 min. A volume (60 µl) AMPure XP clean-up was performed and the DNA was eluted in 31 µl Nuclease-free water. A 1-µl aliquot was quantified by fluorometry (Qubit) to ensure ≥70% input DNA was retained. Afterwards, ligation was performed by adding 20 µl Adapter Mix (SQK-LSK108 Ligation Sequencing Kit 1D, Oxford Nanopore Technologies) and 50 µl NEB Blunt/TA Master Mix (NEB, cat. no. M0367) to the 30 µl dA-tailed DNA, mixing gently and incubating for 10 min at room temperature. The adaptor-ligated DNA was cleaned up by adding 60 µl of AMPure XP beads, incubating for 5 min at room temperature and resuspending the pellet twice in 140 µl ABB (SQKLSK108). The purified-ligated DNA was resuspended by adding 25 µl Elution Buffer (SQK-LSK108) and resuspending the beads, incubating at room temperature for 10 min,

pelleting the beads again, and transferring the supernatant (pre-sequencing mix) to a new tube. A 1-µl aliquot was quantified by fluorometry (Qubit) to ensure ≥43% input DNA was retained. Samples were primed and loaded into the SpotON flowcell containing up to 512 nanopore channels for sequencing DNA in real time.

The Nanopore sequencing data of sg*UTP14A* was acquired with MinKNOW 1.15.1 (Oxford Nanopore Technology) and base-called with Albacore 2.2.7 (Oxford Nanopore Technology). For the sequence data visualization, the sequences were aligned on the *UTP14A* transcript reference (Ensemble ID: ENST00000394422.7) with minimap2 version 2.12 (H. Li, 2018) and the alignment displayed with Tablet 1.17.08.17 (Milne et al., 2013). To quantify the deletion and insertion sizes caused by the Cas9-sgRNA treatment, the sequence data were aligned on a modified *UTP14A* transcript containing the TCT deletion on position X:129911061-129911063. From the resulting sam alignment file, CIGAR values for deletions and insertions on the sgRNA target site X:129911061 from 100,000 forward aligned reads were read out using (Bunn, 2010) and plotted using the R package ggplot2.

2.11 Deep Sequencing

RKO and HCT116 cells were collected in PBS prior to genomic DNA extraction using the QIAamp® DNA Mini and Blood kit (Qiagen), following the manufacturer's protocol (Qiagen). The genomic DNA was used as a template for PCR analysis, following the MyTaq Red DNA Polymerase protocol using the PCR program depicted below employing U6-sgRNA specific primers. Subsequently, the PCR products were purified using a GeneJET PCR purification Kit (Thermo Scientific), following the manufacturer's protocol. PCR primers were designed to amplify the U6-sgRNA on the lentiviral backbone containing sgRNA (Table.4). Primers for second PCR contain universal ‘‘tails” that allow for a second amplification step to incorporate Illumina adapter sequences as well as sample-specific barcodes. 6 µg/µl in 10µl were submitted for next generation sequencing. sgRNAs counts were determined through alignment of paired-end reads and programmatic evaluation of variants relative to our baseline sample.

2.12 Organoids culture

2.12.1 Generation of Human PDAC Organoids (This section was performed by our collaborator PD Dr. Dr. med. Daniel E. Stange)

Tumor specimens were cut into pieces smaller than 1 mm3 and digested with dispase II (2.5mg/ml, Roche) and collagenase II (0.625 mg/ml, Sigma-Aldrich) in DMEM/F12+++ medium (DMEM/F12 (Invitrogen) supplemented with 1x HEPES (Invitrogen), 1x Pen/Strep (Invitrogen), and 1x GlutaMAX (Invitrogen)) at 37°C for 30-120 minutes depending on sample size. After

several washing steps with DMEM/F12+++ medium, the remaining cell pellet was resuspended in GFR Matrigel (Corning) and cultivated in human PDAC organoid medium DMEM/F12+++ supplemented with Wnt3a-conditioned medium (50% v/v), noggin-conditioned medium (10% v/v), RSPO1-conditioned medium (10% v/v), B27 (1x, Invitrogen), nicotinamide (10mM, Sigma-Aldrich), gastrin (1 nM, Sigma-Aldrich), N-acetyl-L-cysteine (1 mM, Sigma Aldrich), primocin (1 mg/ml, InvivoGen), recombinant human fibroblast growth factor 10 (hFGF10, 100 ng/ml, PeproTech), A-83-01 (0.5 µM, Tocris Bioscience), and N2 (1x, Invitrogen).

2.12.2 Human Organoids Culture

Human PDAC and gastric cancer tissues were obtained from patients who underwent surgery at the Department of Visceral, Thoracic and Vascular Surgery at the University Hospital Carl Gustav Carus of the TU Dresden. Human PDAC and gastric cancer organoids were cultured as described earlier and were passaged twice a week with a split ratio of 1:2/1:3 (Bartfeld et al., 2015). Briefly, for 48 well plates, 1 mL cold Advanced-DMEM/F12 was added per well, using micropipette and transferred to 15 mL falcon tube. The end of a glass Pasteur pipette was narrowed using fire so that this pipette takes up medium slower than an un-narrowed pipette. Organoids were pipetted up in Advanced DMEM/F12 using the narrowed pipette 10x up and down, breaking up the organoids. Then after centrifugation for 5 min 300 g, the supernatant was carefully discarded and the resulting pellet was resuspended thoroughly in Matrigel while avoiding air bubbles. In each well of a pre-warmed 48 well plate, a 20 µL drop of the Matrigel-cell mixture was placed. Then, after solidifying for 10 mins at 37°C, the Matrigel drop was overlaid with 250 µL medium per well. Finally, the plate was transferred back to the incubator and media was replenished every 2-3 days. For long term storage, Organoids were disrupted as for passaging using the Pasteur pipette. Then, the fragments were dispensed in cold Recovery Cell Culture Freezing Medium (500 µL /well) and placed in 1 mL cryotubes. Organoids were frozen down over night in a -80°C freezer in a cryo-freezing container (Mr. Frosty, Nalgene). Cells were transferred few days later to liquid nitrogen. For thawing, cryotube was warmed at 37°C and cells were resuspended in 10 mL Advanced DMEM/F12. Then, cells were centrifuged for 5 mins at 300 g. Pelleted cells were resuspended in 20 µL Matrigel and placed in the center of a well of a pre-warmed 48 well plate. The plate was carefully placed in the incubator to let the Matrigel solidify at 37°C for 10 mins. Finally, the Matrigel was overlaid with 250 µL of prewarmed medium containing all growth factors including RHOKi.

2.12.3 PDAC Organoids Time course

Wells of the same condition were pooled together (up to four wells/falcon tube) by scraping off the well bottom with a 1ml pipette to collect Matrigel-containing organoids into a 15ml falcon, followed by dissociation by mechanical disruption using a narrow-tip glass pipette and a pipette boy. Next, 4ml F+++ was added, followed by mixing and a spin down, 5 mins at 300g. Organoid pellets were resuspended in 1ml F+++, splitted into two equal aliquots and spun down again, 5 mins 300g. One aliquot is passaged in culture, resuspended in an appropriate amount (20ul per well) of Matrigel (2 wells/condition, at least), Matrigel was left to polymerize for 10 mins at 37°C, then 250µl of Pancreatic media containing Puromycin (0,3 µg/ml) or Pancreatic medium for WT cells was added to each well and incubated at 37°C. The second aliquot is resuspended in prewarmed Tryple Express at 37°C, pellet is gently mixed and dispensed into 48 well plate, followed by 5 mins incubation at 37°C. Tryple Express was diluted by adding 5 ml F+++ and spun down, 5 mins at 300g. For FACS measurement, cells were resuspended in 200µl 1%BSA in PBS, pipette into FACS tubes with filter caps.

2.13 Genotyping of base edited cells

Genomic DNA was isolated using QIAamp® DNA Blood Mini kit according to manufacturer's instructions. Targeted *KRAS* exon2 PCR amplification was performed using high-fidelity Phusion polymerase using the following protocol:

10 µl HF Buffer
1 µl dNTP mix (10 mM each)
1,25 µl 20 µM (*KRAS*-FwdPrimer)
1,25 µl 20 µM (*KRAS*- RevPrimer)
250ng gDNA or 0.5 µl cell lysate
0.5 µl Phusion DNA polymerase
nuclease-free water up to 50 µl

Cycler conditions	Temperature	Time	Cycles
Lid	110°C	Hold	
Initial Denaturation	98°C	30 seconds	1x
Denaturation	98°C	10 seconds	35x
Annealing	65°C	20 seconds	
Elongation	72°C	30 seconds	
Elongation	72°C	5 minutes	1x
Storage	8°C	Hold	1x

A 5-µl aliquot of each PCR reaction was run on agarose gel, confirming primer specificities and correctly sized bands. Next, PCR purification was performed using Bioline kit following the manufacturer's instructions and DNA concentrations was quantified using a Nanodrop spectrophotometer. Then, the appropriate amount of DNA together with the sequencing primer (*KRAS* forward or reverse primer were used according to company protocol) was submitted for Sanger sequencing, a cost-effective alternative to next generation sequencing.

2.14 EditR to quantify base editing efficiency

EditR is a free online tool (www.moriaritylab.shinyapps.io/editr_v10/, (Kluesner et al., 2018)) to quantify sequencing reads from raw ab1 files. We amplified a 544bp PCR product spanning *KRAS* residues 12 and 13 of the sgRNA-treated cells, while using the cells treating with the no sgRNA vector as a control followed by Sanger sequencing of the PCR products. The obtained ab1 files of the potentially edited region was uploaded to EditR together with the sgRNA protospacer sequence (~20 bp). EditR generates a plot displaying editing efficiencies at each base within the protospacer (Fig. 25B, C).

2.15 Statistical analysis

Data were analyzed using GraphPad Prism version 6 (GraphPad Software). Unless otherwise stated, timepoints in time-course experiments are presented as the standard deviations (SD, presented as error bars) of three independent experiments, performed in biological triplicates. The raw FACS points were processed using functions implemented in RStudio Version 1.2.1335 (RStudio Inc., Boston, Massachusetts, United States) and the statistical difference between the mean percentage at end point of experimental gRNA and that of non-targeting gRNA/no gRNA was determined using unpaired two-tailed Student t test. $P<0.05$ was considered to be statistically significant.

2.16 Time-lapse microscopy

PANC-1-ABE8e-GFP cells expressing the base editor were infected with pLenti.sgG12D-1-TdTomato, at a MOI=1, resulting in ~60% infection rate to create an internal competition between cells bearing the base editing sgRNA and cells expressing only the base editor. Four days after infection, cells were seeded into a µ-Slide 8 well-chambered coverslip slide (Ibidi) containing 300 µl of complete DMEM media. The media was replaced with 300 µl of FluoroBrite DMEM (Gibco) supplemented with 10% FBS, 4 mM GlutaMAX, and 1% Pen/Strep. Next, 0,2 µM SiR-Hoechst far-red live cell DNA stain with 2,5 µM efflux pump inhibitor verapamil following authors' suggestion (Lukinavičius et al., 2015). Cells were incubated in SiR-verapamil in supplemented Fluorobrite for one hour before I started the experiment. Time-lapse microscopy-based imaging was performed using the Deltavision Elite deconvolution microscope. Images were acquired on FITC, TRITC and Cy5 channels every 1 hour for 72 hours using a 60x/1.42 plan-Apochromat objective with immersion oil as an imaging medium, at 37°C with 5% CO_2. Subsequently, images were deconvolved and z-projected using Deltavision image processing and analysis software, SoftWoRx.

2.17 Immunofluorescence

RKO Cells were grown on coverslips overnight, fixed in ice-cold methanol at −20°C for 10 mins, quenched in acetone for 1 min, and blocked with 0.2% gelatin from cold-water fish skin (Sigma-Aldrich) in PBS (PBS/FSG) for 20 mins. Cells were stained by incubation with primary antibody for *UTP14A* was diluted in blocking solution (1:1000) for 1 hour in PBS/FSG and washed with PBS/FSG. The cells were then incubated with fluorescent-dye donkey anti rabbit-IgG secondary antibody for 1 hour at room temperature. After washing with PBS/FSG, coverslips were mounted on glass slides containing 4′,6-diamidino-2-phenylindole (DAPI; ProLong Gold anti-fade; Thermo Fisher Scientific). Images were acquired on an Olympus IX71 equipped with the DeltaVision Elite imaging system using 60×/0.95 plan apo objective, deconvolved, and projected using softWoRx software (Applied Precision). Acquired images were cropped and contrast adjusted using Fiji (Schindelin et al, 2012).

2.18 BAC transfection (RKO-*UTP14A* BAC line)

The LAP-tagged h*UTP14A* BAC was a kind gift from Mihail Sarov, MPI-CBG, Dresden. Stable BAC transfection in RKO cells was carried out using Effectene transfection reagent and freshly-purified BAC DNA. The transfection mix preparation and transfections were performed following the manufacturers protocol (Qiagen) in 12 well plate format. Three days after transfection, Geneticin (G418, Invitrogen 10131–019) selection was applied (200 µg/ml). One week after transfection, GFP+ cells were visible and G418 concentration was increased to 400 µg/ml and maintained while changing media every 2-3 days to remove dead cells and cell debris. Geneticin-resistant GFP+ clones were pooled into a T-75 flask for expansion, and kept at 400 µg/ml geneticin selection throughout all subsequent experiments.

2.19 BAC recombineering

For generation of a *UTP14A* BAC variant that carries the *UTP14A* mutation while being resistant to mutation-targeting sgRNA, I utilized the power of recombineering via *ccdB* counterselection and single strand oligonucleotide repair (ssOR) on wild-type *UTP14A* BAC, as previously described (Fu et al., 2010; H. Wang et al., 2014). Briefly, freshly purified WT *UTP14A* BAC was transformed into *ccdB*-resistant *E. coli* GBred *gyrA462* and plated on chloramphenicol plates (15 µg/mL). Next, clones were grown and made electrocompetent while separately, the ccdB-Hygromycin cassette was PCR amplified from the linearized and DpnI-digested plasmid p15A-*ccdB*-Hygromycin bearing the necessary homology arms to an internal locus flanking *UTP14A*:S99delS mutation (Table.3). The ccdB-Hygromycin PCR gel-extract was used for transformation into electrocompetent GBred *gyrA462*/*UTP14A*-BAC, after

the expression of recombination-mediating genes, induced by L-arabinose and cells were plated on Hygromycin plates. Afterwards, the dual inducible expression plasmid pSC101-ccdA-gbaA was electroporated into *E. coli GB05* and plated on tetracyclin plates (5 μg/mL) followed by growing colonies and making these cells electrocompetent. Of note, the temperature-sensitive pSC101-ccdA-gbaA carries the ccdA gene under the control of the arabinose-inducible P_{BAD} promoter, as well as the λ phage redα, redβ and redγ genes together with the *E. coli recA gene (redαβγA)* in a polycistronic operon under the control of the rhamnose-inducible P_{RhaB} promoter. Induction with L-rhamnose promotes homologous recombination, whereas induction with L-arabinose promotes *CcdA* expression to confer *CcdB* resistance. Next, BAC-*UTP14A*/p15A-*ccdB*-Hyg was transformed into *E. coli GB05*/*pSC101BADccdARhaγβ* at 30°C. Finally, ssOR (Table.3) to replace the ccdB-Hygromycin-cassette with the *UTP14A* mutant oligonucleotide; the oligonucleotides for the insertion are flanked by 40 nt homology on each side. In order to target either the leading or the lagging strand in respect to the direction of replication both complementary oligonucleotides were used. The recombinants were counter selected on chloramphenicol plates at 37 °C. Only clones growing are the ones in which the counterselection cassette was replaced by the oligonucleotide. Targeted *UTP14A* PCR and sanger sequencing validated the correct insert integration.

Original sequence on *UTP14A* BAC: CTTGAGC**CTGTTAAAACTTCATCTTCTTTGG**CCACT
Desired mutant BAC sequence: CTTGAGC**CCGTCAAGACCAGTAGCCTTG**CCACT

2.20 Protein Immunoblotting

The rabbit polyclonal antibody against human *UTP14A* (11474-1-AP, Proteintech) was used for detecting 88-kDa *UTP14A* protein, protein is fused to GFP of Mol wt ~27 kDa). The Anti-GAPDH polyclonal antibody (AP16240PU-N, Acris Antibodies) for 36-kDa GAPDH protein.

For sample preparation, four million cells of RKO and *UTP14A*-BAC line were collected and washed in PBS (Pellet can be stored at −80°C). Cells were resuspended in 200 µl Lysis buffer mixed with Protease Inhibitor Cocktail (ThermoFisher, 100x) and 5 µl benzonase 25U/µl (Merckmillipore, 70664-3), and incubated for 10 min on ice. Then, 9 µl of sample were mixed with 3 µl 4x loading dye (Invitrogen, NP0007) and heated for 5 min at 50 °C, before centrifugation for 1 min at 14,000 rpm. 12 µl of sample was loaded to 4%−12% Bis-Tris Nu-Page (Invitrogen, NP0322BOX) with 3 µl Broad Range Marker; Running buffer: MOPS Nu-Page20x (Invitrogen, NP0001); run at 150 V for 1.15h. Next, blotting was done onto Amersham Protran 0.45 NC nitrocellulose membrane using a semi-dry blotting method. Briefly, wet nitrocellulose membrane overlaid by gel, were placed between two 1x thick Whatman papers moistened in blot buffer (100 ml 10xBlot Buffer+100 ml methanol+800 ml water). Blotting was

done through 50 mA (per gel) for 60 min (voltage should be less than 10 V). Afterwards, membrane was washed three times with PBST (PBS/0.1% Tween-20), 5 min each and incubated for 1 hour at room temperature in blocking buffer (here: milk powder). Blocking buffer is always 5% w/v BSA or milk powder in PBST. Then, nitrocellulose membrane was washed for 5 min with PBST prior to incubating overnight with primary antibodies: a) anti-*UTP14A* rabbit (11474–1-AP) in 5 ml blocking buffer (5% milk) (1: 1000) and b) anti-GAPDH mouse (Sigma, cat. no. G8795) in 5 ml blocking buffer (5% milk) (1:10000). Membrane was incubated overnight at 4°C with gentle shaking. Then, nitrocellulose membrane was washed for 5 min with PBST before incubating with secondary antibodies, donkey anti-rabbit and anti-goat antibody for *UTP14A* (15,000) and GAPDH (1: 20,000), respectively in 5 ml blocking buffer (5% milk). Nitrocellulose membrane was incubated for 30 mins at room temperature, following washing 3x for 5 min and scanning on an Odyssey LI-COR Infrared Imaging System.

2.21 Online resources

Collaborative efforts on cancer genome sequencing have resulted in enormous datasets which are compiled on free-access online portals. These expanding databases facilitated information exchange, helped me gain knowledge on mutations reported in patients and in cell lines as well as giving hints for experimental validation. To list a few of the resources I heavily used through this thesis:

cBioPortal for Cancer Genomics,

https://cancer.sanger.ac.uk/cosmic,

https://score.depmap.sanger.ac.uk,

International Cancer Genome Consortium Data Portal,

TumourPortal - Broad Institute,

Cancer Cell Line Encyclopedia (CCLE),

canSAR knowledgebase,

The Cancer Genome Atlas Data Portal,

International Cancer Genome Consortium.

2.22 Materials, Reagents, Equipment and Buffers

Reagents & kits

- MyTaq DNA polymerase with 5x MyTaq buffer (Bioline, BIO-21105)
- Phusion High Fidelity DNA Polymerase with 5x Phusion HF Reaction Buffer (NEB, M0530S)
- dNTP solution mix, 25 mM each (Enzymatics, cat. no. N205L)
- 1-kb Plus DNA ladder (Life Technologies, cat. no. 10787-018)
- Gel Loading Dye, Purple (6x) (New England BioLabs, cat. no. B7024S)
- BsmBI (New England BioLabs, cat. no. R0580S)
- EcoRI-HF (NEB)
- BbsI (NEB)
- CutSmart buffer (NEB)
- NEBuffer3.1 (NEB)
- 10x Tango buffer (Thermo Scientific, cat. no. BY5)
- DTT (Thermo Scientific, cat. no. R0862)
- RedSafe Nucleic Acid Staining Solution (intron Biotechnology)
- T4 DNA ligase with 10x ligation buffer (New England BioLabs, cat. no. B0202S
- T4 polynucleotide kinase (New England BioLabs, cat. no. M0201S)
- Isopropyl-D-thiogalactopyranoside (IPTG) solution (Sigma-Aldrich)
- Adenosine 5′-triphosphate, 10 mM (New England BioLabs, cat. no. P0756S)
- *Stbl3* chemically competent *E. coli* (Life Technologies, cat. no. C7373-03)
- One Shot MAX Efficiency DH5α-T1 Competent Cells (Invitrogen)
- Library Efficiency *DH5 alpha* Competent Cells (Invitrogen 18263012).
- Ampicillin, 100 mg ml − 1, sterile filtered (Sigma, cat. no. A5354)
- Bioline PCR Purification Kit (BIO-52060).
- QIAprep Spin Miniprep Kit (Qiagen 27106).
- QIAGEN-tip 20 Maxiprep Kit (Qiagen 10023)
- QIAquick gel extraction kit (Qiagen, cat. no. 28704)
- QIAamp® DNA Blood Mini kit (Qiagen, 51104)
- TA Cloning™ Kit, with pCR™2.1 Vector (Invitrogen)
- Protease Inhibitor Cocktail (Cell Signaling Technology)
- Bovine Serum Albumin (BSA) (NEB)
- Ultra-Pure Agarose (Invitrogen)

Mammalian cell culture

- HEK 293T, HCT116, RKO and A549 cells (Life Technologies, cat. no. R700-07)
- PANC-1 were a kind gift from Max Heiduk, Seifert AG, Department of Surgery at the Univeristy Hospital Carl Gustav Carus Dresden
- DMEM, Glutamax (Life Technologies, cat. no. GIBCO 10829018)
- Fetal Bovine Serum (Invitrogen 10108165)
- 0.05% Trypsin–EDTA (GIBCO 25300054).
- Penicillin-streptomycin, 100x (GIBCO 15070063)
- Protamine sulfate (Sigma-Aldrich)
- Puromycin dihydrochloride (Life Technologies, cat. no. A11138-03)
- Dimethyl Sulfoxide (DMSO, Sigma D2650)
- 10 cm diameter tissue culture dishes (Corning 430167).
- 96-well round bottom tissue culture plate (Corning 3790).
- 24-well tissue culture plate (Corning 3524).
- Amicon® Ultra-15 Centrifugal Filter Devices (Merck Millipore)
- Filtropur S 0.45 µm (Sarstedt)
- Open-Top Thickwall Polycarbonate Tube (Beckman Coulter)
- Falcon tubes, polypropylene, 15 ml (BD Falcon, cat. no. 35209)
- Falcon tubes, polypropylene, 50 ml (BD Falcon, cat. no. 352070)

Organoids culture

- Advanced Dulbecco's Modified Eagle Medium/Ham's F12 (Gibco, 12634028)
- B27 (Thermofischer, 17504001)
- Nicotinamide (Sigma, N0636-100G)
- TrypLE Express (Thermofischer, 12604-021)
- Gastrin (Sigma, G945)
- N-acetyl-L-cysteine (Sigma, A9165-5G)
- Primocin (InvivoGen, ant-pm-1)
- FGF-10 (Peprotech, 100-26)
- A-83-01 (Tocris Bioscience, 2939)
- ROCK inhibitor (Sigma, Y0503-1MG)
- N2 (Invitrogen, 17502-048)

Equipment:

- Thermocycler with programmable temperature stepping functionality.

- Desktop microcentrifuges (e.g., Eppendorf, cat. nos. 5424 and 5804). Centrifuge (Eppendorf 5810).
- GelDoc Infinity 3000 (Vilber Lourmat Sté, Collégien, France).
- MACSQuant® VYB (Miltenyi Biotec, Bergisch Gladbach, Germany)
- Odyssey® CLx Imaging System (Li-cor Biosciences GmbH, Bad Homburg von der Höhe, Germany)
- Avanti JXN-30 centrifuge (Beckman Coulter).
- FlowJo™ Software FlowJo, LLC, Ashland, Oregon, United States.
- Avanti JS-24.38 swinging-bucket rotor (Beckman Coulter).
- Blue-light transilluminator and orange goggles (Safe Imager 2.0; Invitrogen, cat. no. G6600).
- UV spectrophotometer (NanoDrop 2000c, Thermo Scientific).

Buffers Recipes:

10X Phosphate Buffered Saline (10X PBS)

For 1000 ml

43 mM Na_2HPO_4 ($Na_2HPO_4 \cdot H_2O$; MW = 178.0 g/mol)	**7.6 g**
14 mM KH_2PO_4 (MW = 136.1 g/mol)	1.9 g
27 mM KCl (MW = 74.6 g/mol)	2.0 g
1.47 M NaCl (MW = 58.4 g/mol)	85.9 g

All components were dissolved in distilled water. The pH of this solution was adjusted to 7.4 using NaOH. If PBS were used for cell culture it was autoclaved for 20 min at 121°C prior to use.

Protein Lysis buffer (1x): 100 ml

280mM NaCl (58.4 g/mol)	**1.635 g**
0.5% Igepal	500 µl
5mM MgCl2-hexahydrate (203.30 g/mol)	101.65 mg
10% Glycerol	10 ml
50mM Tris, (121.14 g/mol) pH 8.0	605.7 mg

All components were diluted/dissolved in distilled water.

0.5 M EDTA (pH 8.0)

For 1000 ml

0.5 M $C_{10}H_{14}N_2Na_2O_8 \cdot 2\ H_2O$ (MW = 372.2 g/mol)	**186.1 g**

EDTA was dissolved in distilled water and pH was adjusted with NaOH to 8.0.

1 M Tris buffer (1 M Tris-HCl)

For 1000 ml

1 M Tris (MW = 121.1 g/mol)	**121.1 g**

Tris was dissolved in distilled water and pH was adjusted with HCl depending on the exact experimental requirements.

1X Tris-Borate-EDTA buffer (1X TBE)

For 1000 ml

89 mM Tris (MW = 121.1 g/mol)	**10.8 g**
89 mM H_3BO_3 (MW = 61.8 g/mol)	5.5 g
2 mM EDTA (0.5 M stock solution, pH 8.0)	4 ml

All components were diluted/dissolved in distilled water.

Protein Immunoblotting buffer (Blot buffer - 10x):

For 500 ml

Tris	**15.4 g**
Glycine	72.05 g
SDS	1 g

All components were diluted/dissolved in 500ml distilled water.

Bacterial media recipes

Lysogeny broth (LB-Lennox version)

For 1000 ml

Bacto tryptone	**10 g**
Yeast extract	5 g
NaCl	5 g

All components were dissolved in filter sterilized, distilled water and the pH was adjusted to using NaOH to 7.5, or adjusted to 8 for LS-LB. The volume was the filed up to 1000 ml and the medium was autoclaved for 20 min at 121°C prior to use. Medium was supplemented with ampicillin when necessary at working concentration 100 μg/ml.

3 Results

3.1 Pooled lentiviral screen to identify novel cancer driver mutations

I set out to test the feasibility of using the CRISPR-Cas9 system to functionally profile mutations in a cancer cell line with the aim to nominate novel cancer driver mutations. As a proof-of-concept, I utilized the CRC cell line RKO and first downloaded the reported 4,762 mutations, of which 4,410 are coding sequence mutations, from the COSMIC database (Forbes et al., 2010) followed by mutation-specific sgRNAs design in sense and antisense orientation using the sgRNA Designer algorithm by Doench et al. (Doench et al., 2014; Doench et al., 2016). Most of the designed sgRNAs targeted point mutations, whereas 190 sgRNAs targeted small indel mutations. Targeting point mutations with CRISPR-Cas9 is possible (Gebler et al., 2016), but more challenging, and enriches false positive hits especially in a pooled screen setting. Therefore, in favor of optimal specificity I restricted the sgRNAs in the library to indel mutations, regardless of the genes it resides in. Moreover, applying further criteria such as PAM proximity, >2 mismatches to off-target sites in the genome in addition to %GC content and stretches of same sequence identity nominated 100 sgRNAs with the highest score, targeting indel mutations for the screen (Table.1). In addition, 5 sgRNAs were designed targeting the MYC proto-oncogene, an RNA polymerase subunit and essential ribosomal proteins as positive controls, as well as 2 sgRNAs targeting LUC as non-targeting negative controls. The sgRNAs were cloned into an all-in-one lentiviral Cas9-expression vector also carrying a GFP-expression cassette. The GFP cassette allows for simple evaluation of the infection rate and can be used as a straightforward phenotypic readout.

To perform the screen to identify possible vulnerability mutations, I infected RKO cells as well as another CRC cell line (HCT116) with the lentiviral sgRNA library at MOI of 0,3. This low infection rate was chosen to ensure that most cells were infected with a single copy Cas9-sgRNA virus, making it possible to investigate the inactivation of individual mutations. HCT116 cells do not carry the same mutations present in RKO cells and therefore served as the negative control. Three days post infection, cells were split and genomic DNA was collected from 50% of the cells. These samples were used as the baseline to quantify sgRNA abundance at the beginning of the experiment. After 15 days of continuous culture, genomic DNA was collected again from both cell lines, the sgRNA cassettes were PCR amplified and sgRNA counts were quantified by means of deep sequencing (Fig. 6).

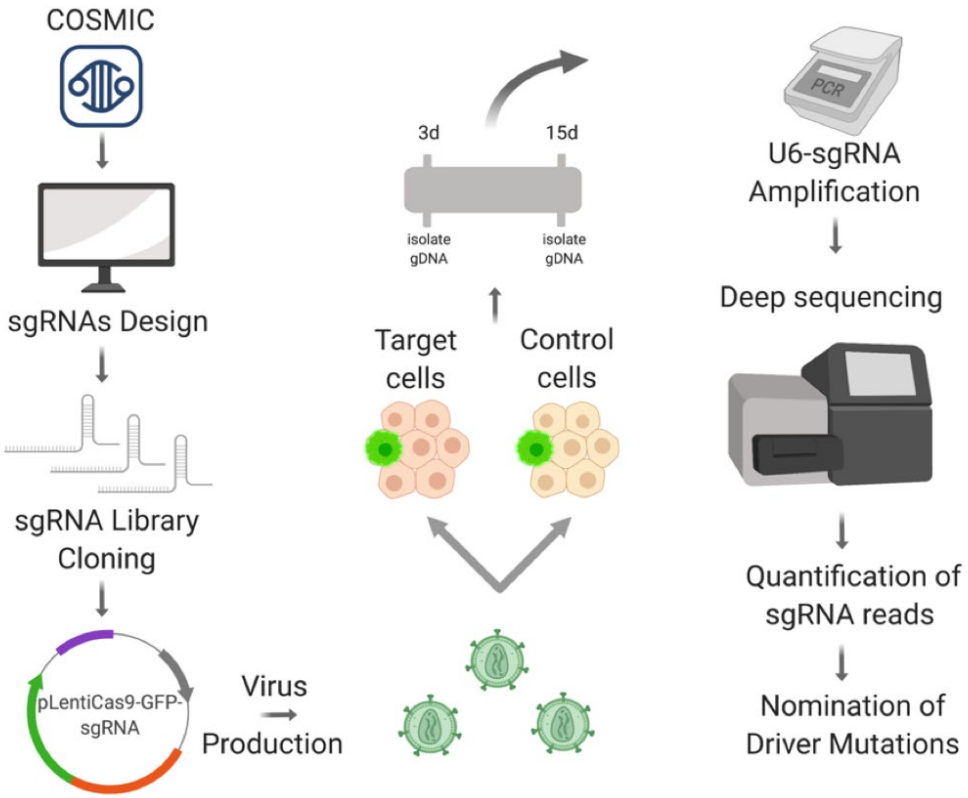


Figure 6 RKO Screen pipeline. Important steps are highlighted by arrows.

As expected, the comparative analysis showed that the positive control sgRNAs targeting essential genes were depleted over time in both cell lines, whereas the relative abundance of the negative controls did not change significantly over time (Fig. 7). In addition to the positive controls, two sgRNAs targeting the same mutation were depleted in RKO cells, suggesting that cleavage of the mutation impairs RKO cells growth (Fig. 7). Importantly, reads of these sgRNAs were not depleted in HCT116 cells, demonstrating the selectivity of the sgRNAs to the mutant sequence in RKO cells.

The analysis revealed depleted counts of two sgRNAs targeting a mutation found in the U3 small nucleolar RNA-associated protein 14 homolog A gene (*UTP14A*:S99delS) selectively in RKO cells. This mutation reflects a three nucleotide in-frame deletion in the N-terminal part of the gene. *UTP14A* is found on the X chromosome and encodes for a component of a large ribonucleoprotein complex implicated in 18S rRNA biogenesis that is involved in pathways such as rRNA processing in the nucleus and the cytosol, ribosome biogenesis and gene expression (L. Hu et al., 2011; Zhu et al., 2016). Interestingly, *UTP14A* has recently been identified as a critical hub gene in colorectal carcinoma cells (Sun & Qian, 2018) suggesting that the gene plays an important role during the transformation process of this tumor type. I therefore decided to investigate the *UTP14A*:S99delS mutation in more detail.

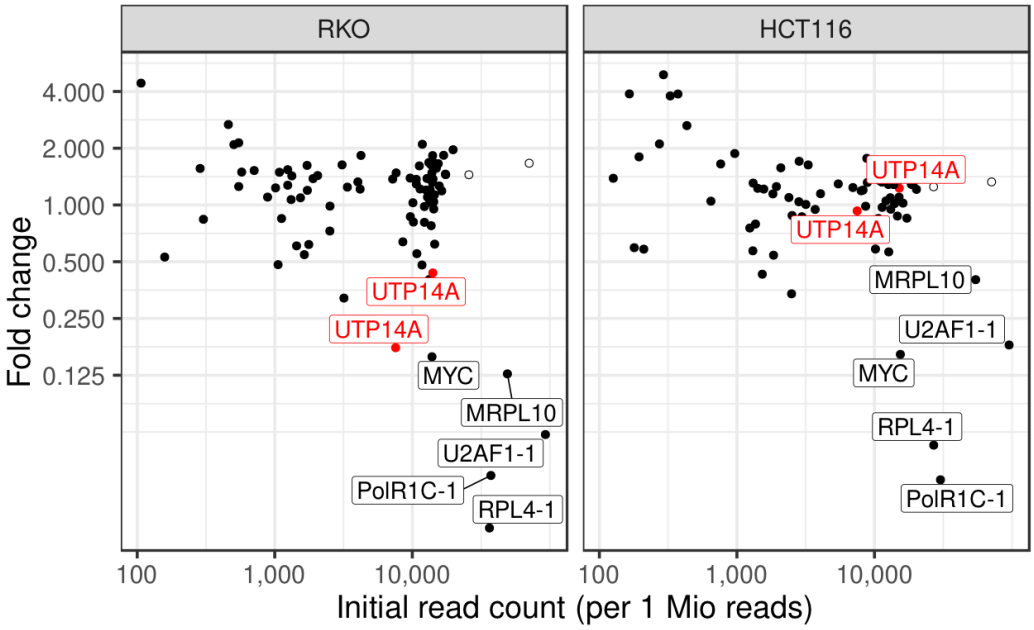


Figure 7 sgRNA screen results in RKO and HCT116 cells. The plots show sgRNAs initial read count vs Fold-change after two weeks of infection. Results are shown for RKO (Left) target cells and HCT116 control cells (right) for 100 selected sgRNAs and positive control sgRNAs targeting essential genes (filled circles). Non-targeting control sgRNAs are shown as non-filled circles. Two independent sgRNAs targeting UTP14A:S99delS are shown in red.

3.2 Validation and characterization of *UTP14A*:S99delS in RKO cells

For validation, the two sgRNAs "hits" targeting the *UTP14A*:S99delS mutation and control sgRNAs were picked for fresh virus production and transduction of RKO and HCT116 cells. No significant change of GFP positive cells was observed over time in HCT116 cells transduced with the negative control or with the sgRNAs targeting the *UTP14A*:S99delS mutation (Fig. 8A). In contrast and in line with screen results, the percentage of GFP positive cells progressively declined over time in RKO cells transduced with the sgRNAs targeting the *UTP14A* mutation (Fig. 8A), supporting the idea that *UTP14A*:S99delS might be a driver mutation in RKO. To demonstrate selective DNA cleavage in RKO cells, I performed a T7 endonuclease I assay showing efficient mutation-specific cleavage of *UTP14A* evident only in RKO cells (Fig.8B), demonstrating that indels are introduced at the site of the *UTP14A*:S99delS mutation after cleavage. Nanopore sequencing further confirmed the presence of numerous indels in the sample that was treated with the sgRNA targeting *UTP14A*:S99delS (Fig. 8C, D).

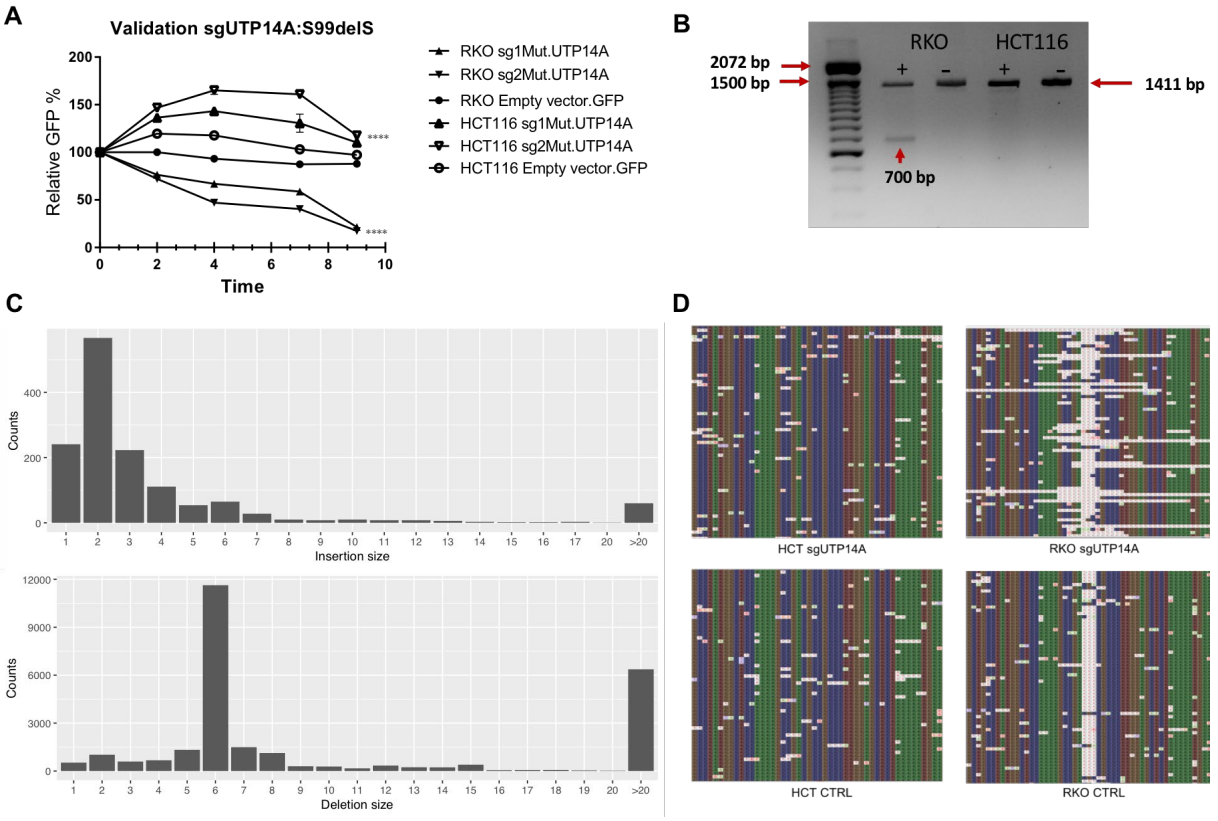


Figure 8 Validation of UTP14A:S99delS mutation as a putative driver mutation in RKO. A) Relative abundance of cells treated with sgRNAs targeting mutant UTP14A in RKO (UTP14A mutant) and HCT116 (UTP14A wild-type) cells over time, as compared to infection with Empty vector bearing no sgRNA. Day zero timepoint refers to three days post infection (dpi), the initial timepoint for measuring GFP signal. In each condition, half of the cells are infected and percentage of GFP is measured over time. Error bars presents means ± SD from experiments performed in technical triplicates. Significance was assessed by means of Student t test; ****P < 0.0001. B) T7 Endonuclease I (T7E1) assay to investigate cleavage of sgRNA targeting mutant UTP14A. Genomic DNA used is isolated from 5dpi from indicated cell lines with (+) or without (−) expression of the sgRNA. The arrow points to the cleavage product. C) Distribution of the deletion (bottom) and insertion (top) sizes of sgUTP14A. A total of 100,000 aligned nanopore sequencing reads were quantified. A total of 26,974 deletions and 1,410 insertions were detected. D) Alignment visualization of Nanopore sequencing data of sgMutUTP14A and UTP14A transcript in RKO and HCT116 cells. Analysis shows the TCT deletion on X:129911061-129911063 in the WT RKO cells (bottom right panel) and deletions of varying size in the treated RKO cell line (top right panel), whereas no changes in HCT116 between sgMutUTP14A treated and untreated cells. Deletions are marked with * on grey background. Sequence deviations to the reference are clearly visible through highlighted bases. Most deviations apart from the TCT deletions are caused by low accuracy of the Nanopore sequencing technology.

To further characterize the *UTP14A*:S99delS mutation in RKO cells, I PCR amplified a DNA fragment spanning the mutation, cloned the fragments into a TA-cloning vector and sequenced the insert of 40 independent clones. While RKO cell line is reported as male (Kong et al., 2019), COSMIC database lists this mutation as heterozygous. By contrast, I found that all 40 clones sequenced carried the in-frame delTCT mutation (Supplementary figure.1).

Furthermore, sequencing of the purified PCR product also revealed that all reads contain the delTCT mutation (Fig. 9A), suggesting that this allele is indeed present in a homozygous or hemizygous state in RKO cells.

To support the finding that *UTP14A*:S99delS is indeed a cancer mutation and not a natural human polymorphism, I searched for this sequence in several Human Genetic Variation databases (accessed July 2018, (Higasa et al., 2016)), http://www.hgvd.genome.med.kyoto-u.ac.jp/) as well as in recently published large-scale human genome sequencing data (Bycroft et al., 2018) (Auton et al., 2015). *UTP14A*:S99delS has not been reported in any of these reports, not even in a heterozygous state, making it highly unlikely that it is a natural human polymorphic alteration. Next, I questioned the role of wild-type *UTP14A* in cell proliferation. To test whether *UTP14A* is dispensable for growth and/or viability I designed two different sgRNAs targeting wild-type *UTP14A* and infected RKO, HCT116 and Hela cells using lentiviral vectors expressing these sgRNAs in conjunction with Cas9 and GFP. Surprisingly, the percentage of

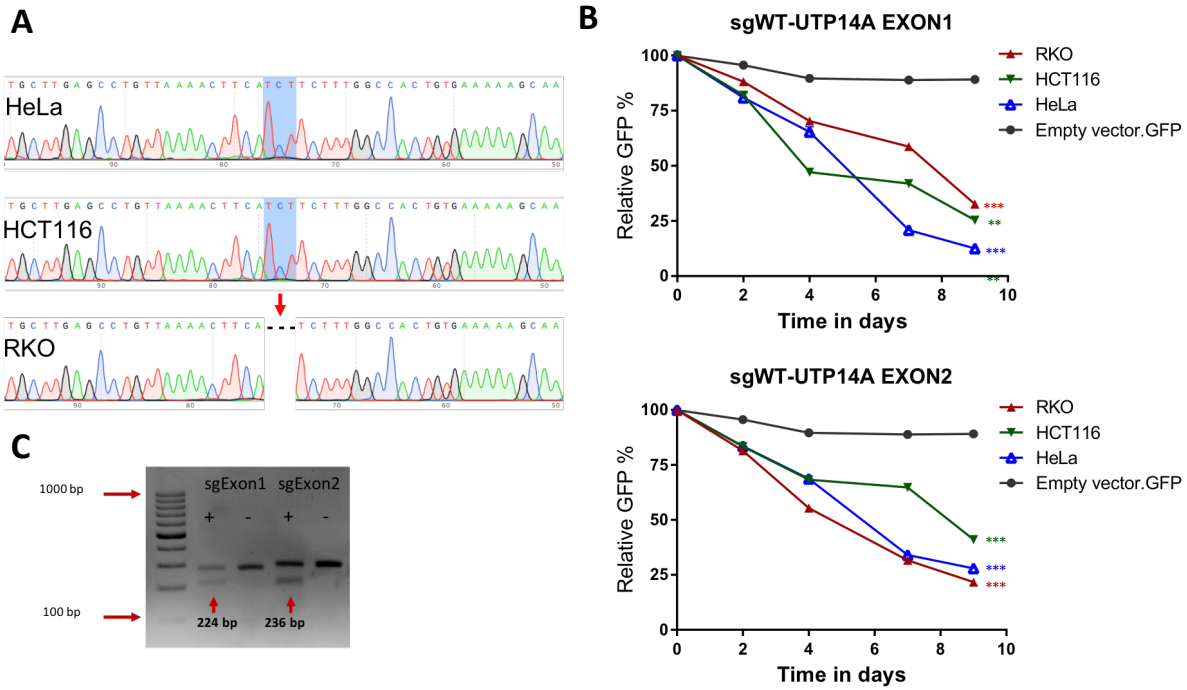


Figure 9 Characterization of UTP14A:S99delS mutation. A) Sanger sequencing of the purified PCR product of UTP14A locus in various cells lines. In contrast to HCT116 and HeLa cells, RKO cells show a TCT deletion as highlighted by an arrow in the chromatogram. In HCT116 and HeLa cells the TCT sequence is highlighted by a blue box. This chromatogram was generated using Snapgene viewer. B) Relative abundance of cells treated with two different sgRNAs targeting exon1 (top) and exon2 (bottom) of WT UTP14A in RKO, HCT116 and HeLa cells over time. Day zero timepoint refers to three days post infection (dpi), the initial timepoint for measuring GFP signal. In each condition, half of the cells are infected and percentage of GFP is measured over time. Error bars presents means ± SD from experiments performed in technical triplicates. Significance was assessed by means of Student t test; **P < 0.005 ***P < 0.0005. C) T7 Endonuclease I (T7E1) assay to investigate cleavage of two sgRNAs targeting exon1 and exon2. The arrow points to the cleavage product.

GFP-positive cells declined over time in all three cell lines for both sgRNAs (Fig. 9B). Moreover, T7 endonuclease I assay showing efficient cleavage of *UTP14A* (Fig. 9C), demonstrating that sgRNAs are efficient and indels are introduced after cleavage. Therefore, we conclude that *UTP14A* knockout cells have a growth disadvantage in cultured cell lines, suggestive of gene essentiality in the three tested cell lines.

3.3 Rescue experiment using BAC transgenomics

The fact that cells require *UTP14A* for proliferation makes it more difficult to conclude that the *UTP14A*:S99delS is a cancer driver mutation. Towards testing this possibility, I introduced wild-type *UTP14A* into the mutant background of RKO cells by means of BAC transgenomics (Poser et al., 2008) (Fig. 10A). I transfected RKO cells with a human BAC construct carrying the GFP-tagged wild-type *UTP14A* gene and confirmed stable and comparable expression levels to wild-type *UTP14A* in these cells (Fig. 10B, C, D). Next, I infected the original and the BAC-modified RKO cells with the Cas-sgRNA vectors cleaving the *UTP14A*:S99delS mutation.

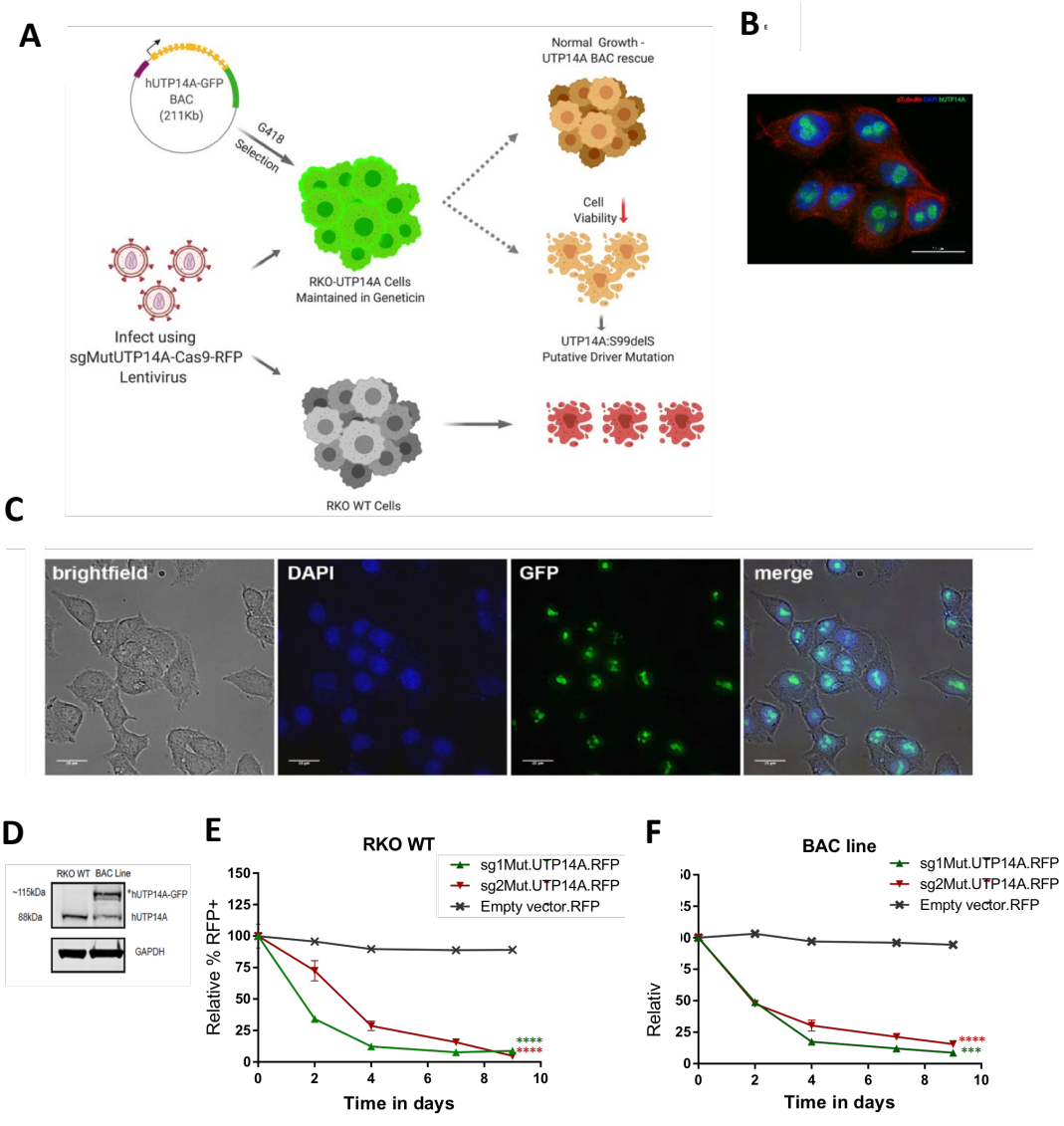


Figure 10 Phenotype rescue via BAC transgenomics. A) Scheme of UTP14A-BAC rescue experiment. In parallel to vulnerable RKO WT cells, UTP14A-GFP BAC-transgenic cells are infected with mutation-targeting sgRNAs. B) Immunofluorescent image of fixed RKO WT cells, using UTP14A antibody GFP-labeled secondary goat anti-rabbit IgG (Green). Red staining: αTubulin. Blue staining: DAPI. Cells nuclei are visualized by DAPI staining whereas, UTP14A-GFP signal could be seen with nucleolar localization. (scale bar= 15μm). C) Images (brightfield, DAPI, GFP and merged) of UTP14A-GFP BAC-transgenic RKO cells (scale bar= 25μm). Note the nucleolar localization of the fusion protein. WT – wild-type cells. G418: Geneticin. D) E) Relative abundance of cells treated with indicated sgRNAs in RKO WT and BAC transgenic line over time. Day zero timepoint refers to three days post infection, the initial timepoint for measuring GFP signal. In each condition, half of the cells are infected and percentage of GFP is measured over time. Error bars presents means ± SD from biological triplicates performed in three independent experiments. Significance was assessed by means of Student t test; ***P < 0.0005 ****P < 0.0001.

If *UTP14A*:S99delS is not a driver mutation, I expected the expression of the wild-type *UTP14A* to rescue the growth deficiency specifically in the BAC-transgenic cells. Surprisingly, both cell lines were equally sensitive to the inactivation of the *UTP14A*:S99delS mutation (Fig. 10E, F), inconsistent with the idea that *UTP14A*:S99delS is a mere passenger mutation. Importantly, GFP signal was expressed at a comparable level over the course of the experiment ensuring the employed sgRNAs do not cut the rescue construct (Supplementary Figure.2).

Finally, the definitive proof that *UTP14A*:S99delS is a driver mutation is the ability to rescue the phenotype resulting from targeting the mutation, using the mutant *UTP14A* protein carrying the S99delS mutation (Fig. 11A). I attempted this second rescue experiment using a mutant *UTP14A* BAC that is resistant to the mutation-targeting sgRNA, which I generated using recombineering technology (Fu et al., 2010) through counterselection principle (H. Wang et al., 2014). Firstly, I deleted the in-frame TCT trinucleotides to generate the S99delS mutation on the BAC. Moreover, for one sgRNA, I abrogated the PAM sequence required for Cas9 binding, in addition to introducing silent mutations that would impede sgRNA binding of either sgRNAs to the BAC without changing the resulting amino acids, using codons of comparable usage frequency (Fig. 11B). Then, after validation of correct recombination, I transfected RKO cells with the newly generated BAC construct carrying mutant-*UTP14A*, maintained cells under Geneticin selection and confirmed stable and comparable expression levels to wild-type *UTP14A* in these cells (Fig. 11C). Next, the mutant BAC-RKO cells were infected with the sgRNA targeting the *UTP14A*:S99delS mutation. Importantly, GFP signal was stable over the course of the experiment ensuring the employed sgRNAs do not cut the rescue construct (Supplementary Figure.3). However, it seems that this transgenic line was still sensitive to the inactivation of the *UTP14A*:S99delS mutation (Fig. 11D), albeit depletion was seen to a lesser extent compared to RKO WT or the BAC transgenic cells. Further experiments for this project were not pursued due to time restrictions.

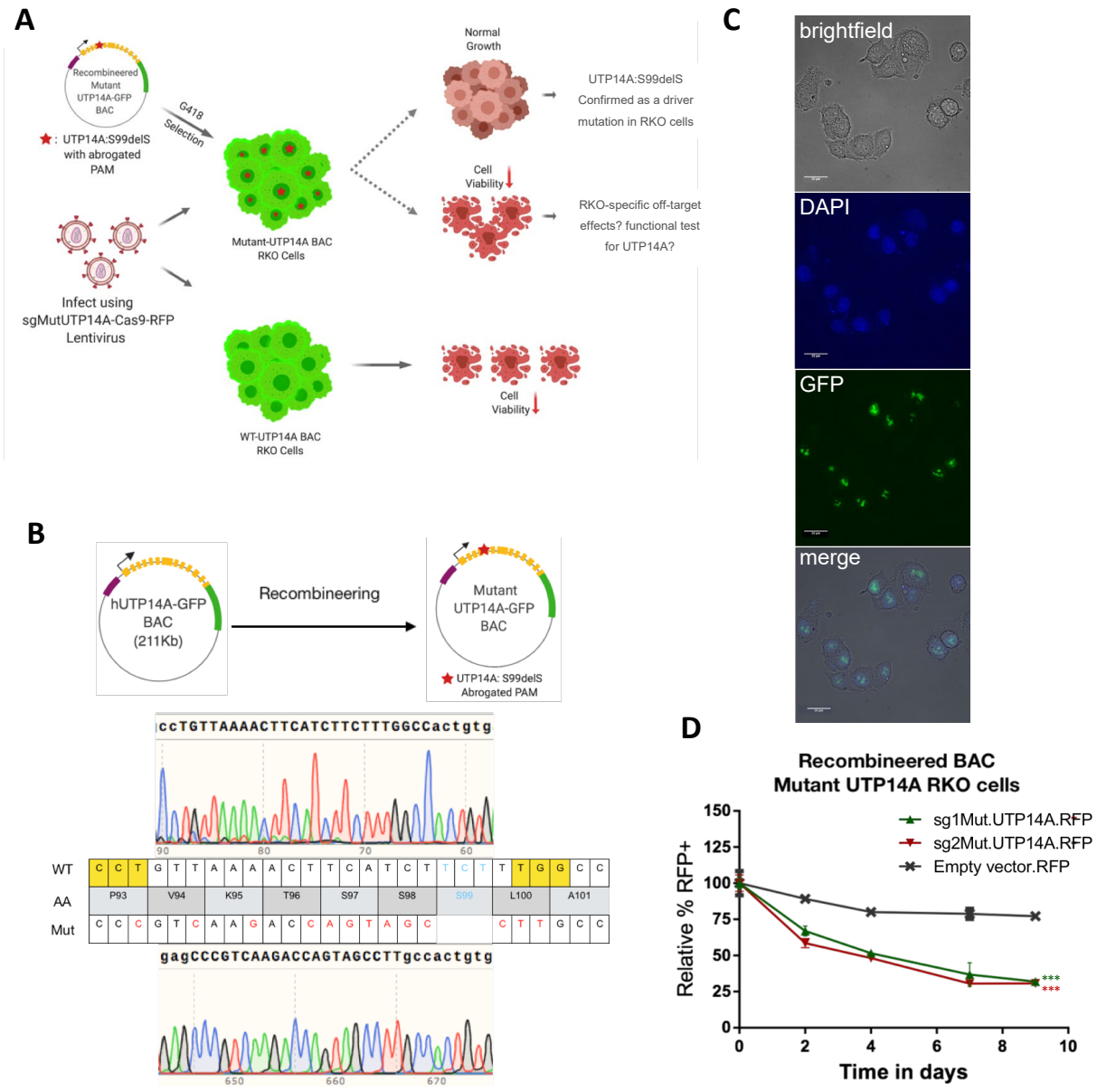


Figure 11 Phenotype rescue using mutant UTP14A recombineered BAC. A) The recombineering principle involves generating a mutation-expressing BAC while evading being targeted by mutation-selective sgRNA. In red are the mutated nucleotides, while blue depicts the in frame TCT deletion. Highlighted in yellow are the two PAMs used by each sgRNA targeting the mutation. WT: wild-type BAC. AA: amino acid. Mut: Mutant BAC. Exemplary chromatograms of wild-type BAC (top) and mutant BAC (bottom) are shown. B) Images (brightfield, DAPI, GFP and merged) of mutant UTP14A-GFP recombineered-BAC RKO cells (scale bar= 25µm). Note the expected nucleolar localization of the fusion protein. C) Scheme of mutant UTP14A-BAC rescue experiment. In parallel to vulnerable UTP14A-GFP BAC-transgenic cells, both transgenic lines are infected with mutation-targeting sgRNAs. G418: Geneticin. D) Relative abundance of cells treated with indicated sgRNAs in RKO WT and BAC transgenic line over time. Day zero timepoint refers to three days post infection (dpi), the initial timepoint for measuring GFP signal. In each condition, half of the cells are infected and percentage of GFP is measured over time. Error bars presents means ± SD from biological triplicates performed in three independent experiments. Significance was assessed by means of Student t test; ***P < 0.0005.

3.4 Arrayed Screen for potential driver mutations in HCT116

Utilizing the precision and versatility of CRISPR-Cas9 system to identify novel driver mutations is scalable to other cell lines. In parallel to the sgRNAs screen in RKO, I sought to screen for driver mutations in CRC cell line HCT116. Assisted by a bachelor rotation student, I collected mutational information of HCT116 from the COSMIC database, followed by gRNAs design and ranking. Mutations were mainly selected based on how specific a sgRNA could be designed. However, for this sceen, we adopted a different approach to the unbiased approach in RKO: whereas the latter was for identification of novel driver mutations, the former is designed to test suspect driver mutations. Therefore, I selected mutations in genes which are implicated in cancer (cancer census genes), and thus, are likely to display sensitivity upon inactivation followed by gRNAs design and ranking. Briefly, *ZNHIT6* (Zinc Finger HIT-Type Containing 6) is believed to play a role in invasive breast cancer (W. Shi et al., 2019). Whereas *MTR*, Methionine Reductase, is associated with increased risk for malignant lymphoma (Shrubsole et al., 2006). The gene encodes an enzyme catalyzing a reaction to produce Methionine, the precursor for the universal methyl donor, S -adenosylmethionine (*SAM*). Similarly, Mutations in *FLT3* (FMS-like tyrosine kinase 3) which encodes a receptor tyrosine kinase are the most common genetic alterations and well-established driver mutations, in actue myeloid leukemia (AML), although, the mutation found in HCT116 cells has not been described in AML (Daver et al., 2019; Stirewalt & Radich, 2003). *EVL*, a transcriptional target of estrogen receptor, plays a role in remodelling of the actin cytoskeleton, which is controlled by a multitude of actin regulators whose differential expression in cancer leads to distinct architectures that impact invasion (Mouneimne et al., 2012; Padilla-Rodriguez et al., 2018).

Interestingly, the transcriptional regulation of *ZNF20*, was highly correlated with cancer-related long non-coding RNA in patients with cervical cancer (L. Yuan et al., 2019). Previous studies have confirmed that ryanodine receptor 2, *RyR2*, is associated with several types of cancers, including melanoma (Carpi et al., 2018), breast cancer (Lu et al., 2017), lymphoma (McCarthy et al., 2003) and prostate cancer (Mariot et al., 2000). Interestingly, Telomerase reverse transcriptase (*TERT*) is relatively well-studied with known mutations across several tumor types (Pestana et al., 2017; X. Yuan et al., 2019), however, in contrast to its frequently reported promoter mutations influencing transcription, *TERT* mutation in HCT116 cells was missense mutation in protein, which could potentially reveal a novel driver mutation. For *ZNF198*, it has been shown that fusion of *ZNF198* with *FGFR1* kinase yields a powerful transforming oncogene and alters the transcription profile of the cells expressing it (Kunapuli et al., 2003). Additionally, expression of the oncoprotein *CAPZB*, F-actin-capping protein subunit beta, has been strongly linked to some types of sarcoma (Mukaihara et al., 2016).

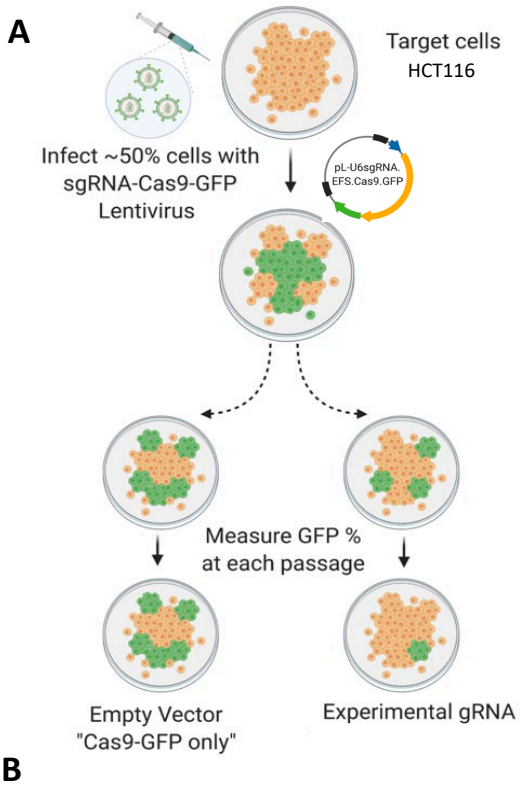
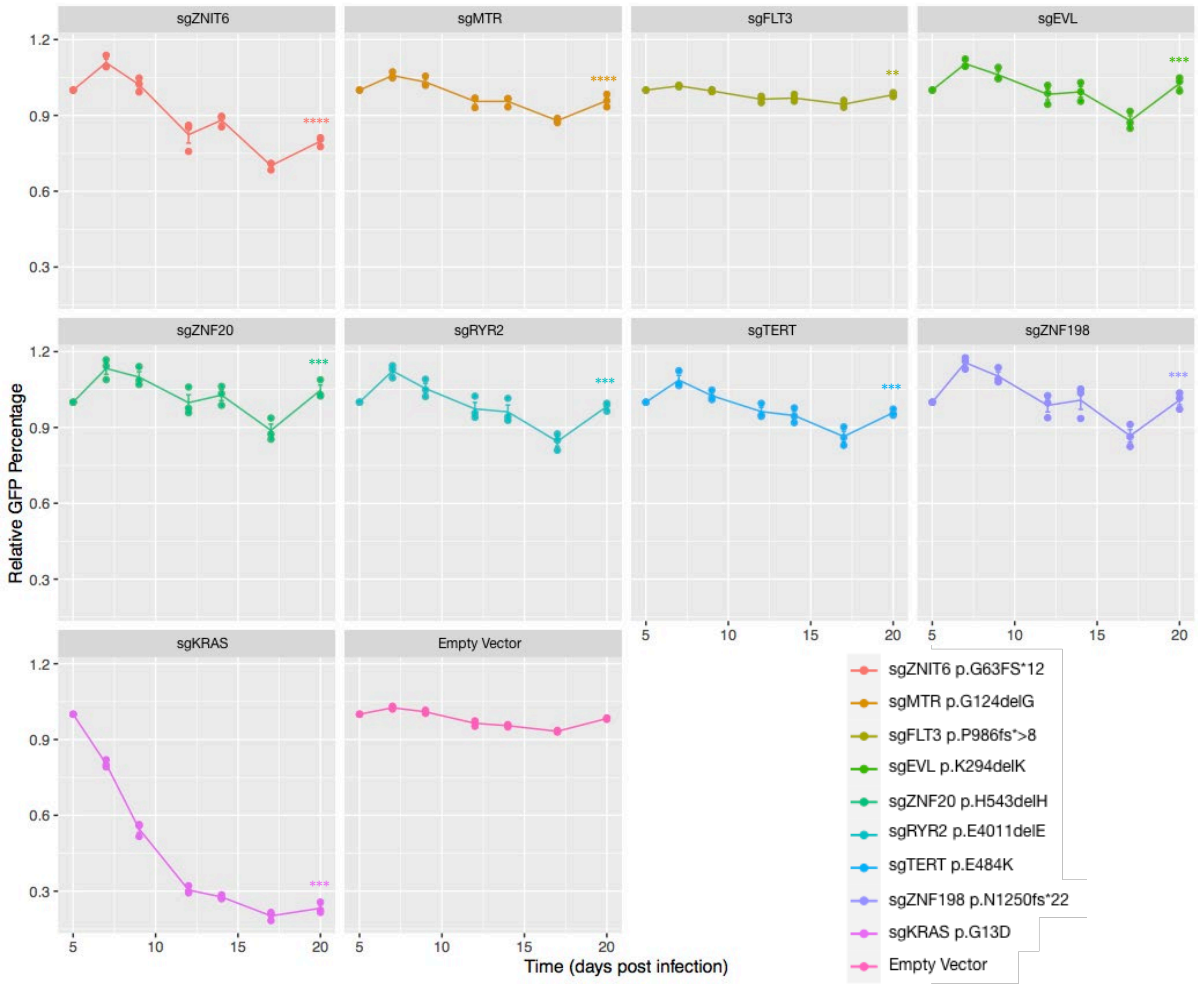


Figure 12 Arrayed screen in HCT116 cells. A) Overview of the competition assay. Mutant cells are transduced with an all-in-one sgRNA-Cas9-GFP lentivirus designed to target the mutation, infected at a multiplicity-of-infection of nearly 1 and the relative abundance of the transduced population (GFP+ cells) is measured over the course of time via flow cytometry. B) Nine different mutations were tested and the relative abundance of cells treated with indicated sgRNAs in HCT116 cells is shown over time as function of GFP% relative to day zero. Day zero timepoint refers to three days post infection (dpi), the initial timepoint for measuring GFP signal. Error bars presents means ± SD from experiments performed in technical triplicates. Significance was assessed by means of Student t test; **P < 0.005; ***P < 0.0005; ****P < 0.0001.

B



Finally, mutations in Kirsten-RAt Sarcoma (*KRAS*) remain the most frequently reported oncogenic mutations in cancer, present in more than a quarter of cancer patients. As a readout, I used cellular fitness after mutation knockout, assayed through a cell competition assay (Fig. 12A). Most mutations showed mild alterations over time and infected cells seemed rather healthy, except for the cells infected with sgRNA targeting *KRAS* mutation. The percentage of GFP positive cells progressively declined over time in cells bearing the G13D-targeting sgRNA (Fig. 12B), suggesting that disruption of mutant *KRAS* impaired cancer cell proliferation. Although the mini-screen in HCT116 cells did not nominate novel oncogenic drivers, the resulting phenotype of targeting mutant *KRAS* emerged as an attractive phenotype. Given its remarkable significance in various cancers and profound clinical relevance, we decided to explore CRISPR-based targeting of mutant KRAS in more detail.

3.5 CRISPR inactivation of mutant *KRAS*

To prove that the effect seen is due to mutation disruption in HCT116 cells, I repeated the experiment and performed a T7 endonuclease I assay showing efficient cleavage in HCT116 cells in both sgRNAs against mutant *KRAS* as well as mutant *FLT3* used here as a control for cutting a passenger mutation (Fig. 13A, B), demonstrating that indels are introduced after mutation cleavage. Next, to test whether this effect is limited to HCT116 cell line, I expanded the scope to include some of the most prevalent *KRAS* mutations in other cancer entities, namely G12D in pancreatic cancer and the G12S in lung cancer. Interestingly, in both tested cell line models, PANC-1 and A549 harboring G12D and G12S mutations, respectively, I noticed a rapid progressive depletion of GFP positive cells over time (Fig. 13C, D), suggesting that these cells too depend on mutant *KRAS* for proliferation.

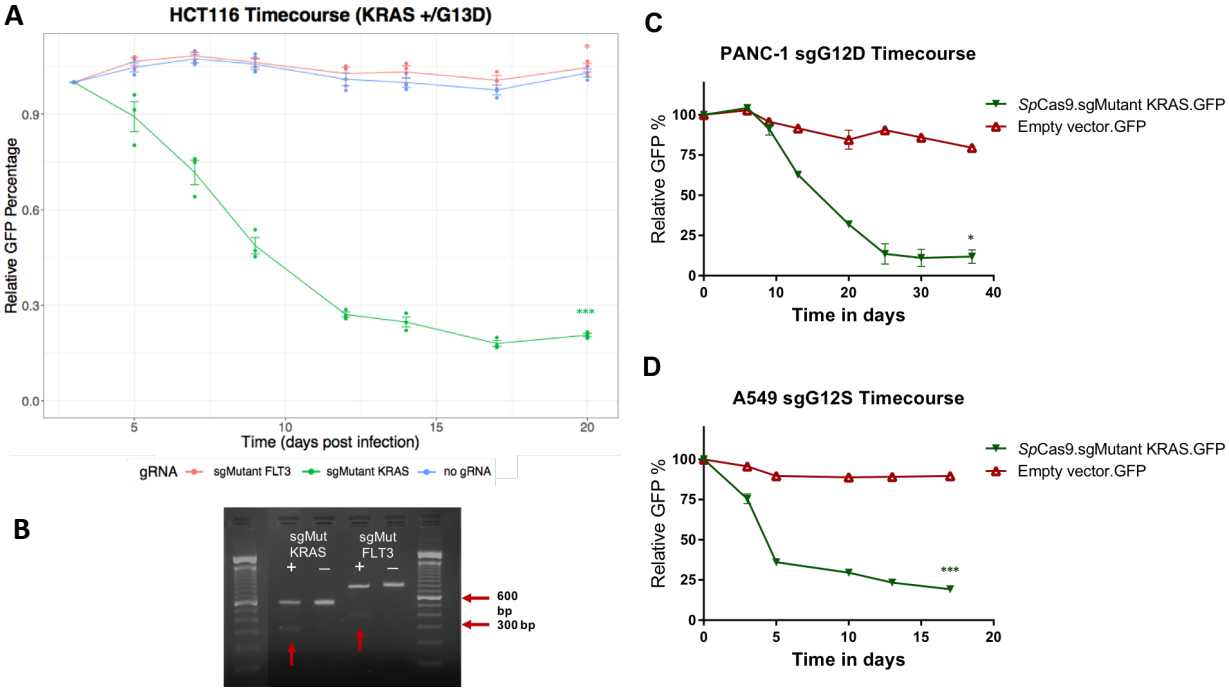


Figure 13 CRISPR/Cas9 targeting Mutant KRAS in cancer cell lines. A) The relative abundance of cells treated with indicated sgRNAs in HCT116 cells is shown over time as function of GFP% relative to day zero. B) T7 Endonuclease I (T7E1) assay to investigate cleavage of sgRNA targeting mutant. Genomic DNA used is isolated from 5dpi. The arrow points to the cleavage product. C) The relative abundance of cells treated with sgRNA targeting KRASG12D mutation in PANC-1 cells is shown over time as function of GFP% relative to day zero. D) The relative abundance of cells treated with sgRNA targeting KRASG12S mutation in A549 cells is shown over time as function of GFP% relative to day zero. In all three experiments, Day zero timepoint refers to three days post infection, the initial timepoint for measuring GFP signal. Error bars Error bars presents means ± SD from experiments performed in technical triplicates. Significance was assessed by means of Student t test; *P < 0.05; ***P < 0.0005.

Although the non-infected cells outcompeted the targeted cells in the pool, I noticed that the depletion curve flattened at a low percentage of GFP positive cells, an observation seen in all three cell lines. Of note, these persistent cells appeared healthy, sustained in culture for several weeks despite carrying the mutation-targeting sgRNA and Cas9. Therefore, I sought to validate if these cells have been targeted by the Cas9 nuclease through sorting the residual GFP-positive HCT116 cells and sequencing ~50 independent clones (Fig. 14A). Interestingly, half of all the reads matched the WT allele sequence, reflecting the heterozygous nature of the G13D mutation and suggesting high specificity of mutation targeting. Surprisingly, the rest of the clones showed several indels, as well as other oncogenic mutations (Fig. 14B, C). Importantly, these mutant alleles are resistant to targeting by the initially designed sgRNA.

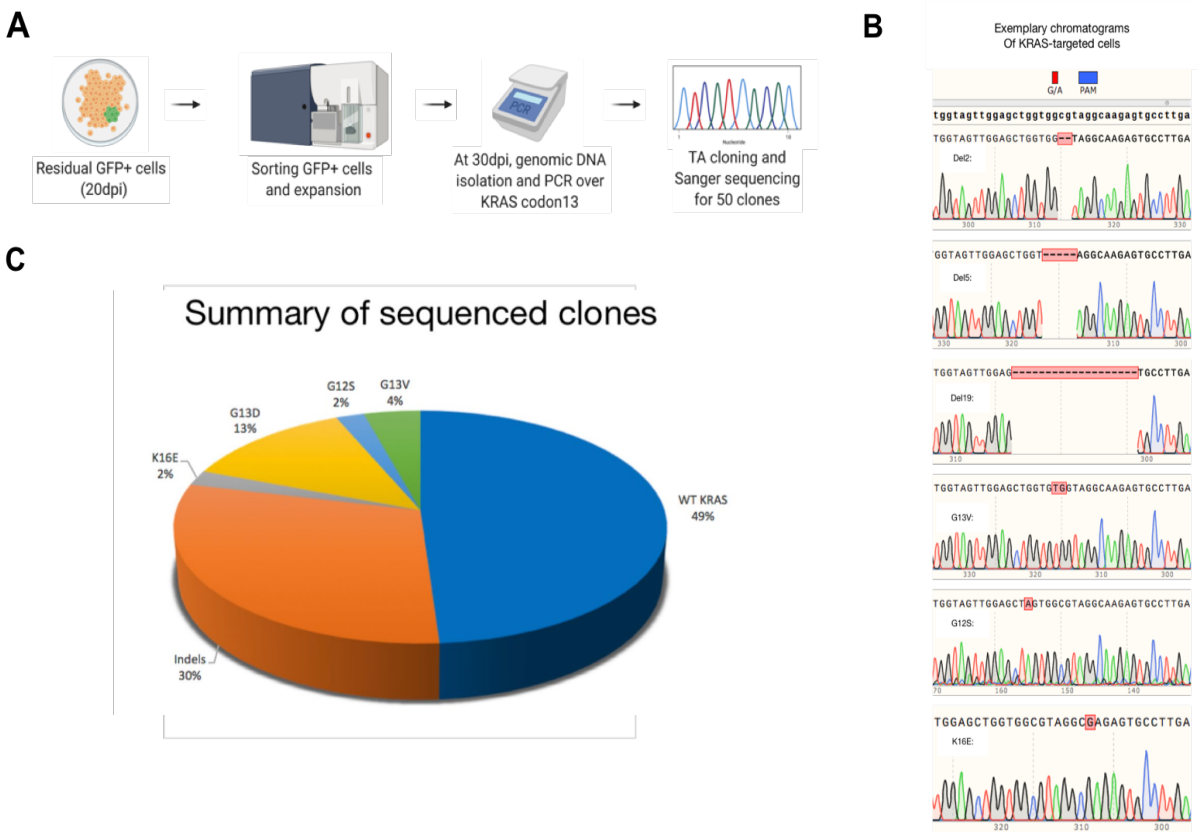


Figure 14 Genetic makeup of persisting GFP+ cells at KRAS cleavage site. A) Scheme of the experiment in HCT116 cells. B) Representative chromatograms of the escape clones. C) Pie chart representing the sum of mutations at codon 13 cut site.

Next, I sought to improve our Cas9-based approach which relies on double-stranded DNA breaks (DSBs) to leverage the endogenous DNA repair machinery to inactivate mutations. Although the system is efficient and specific, new oncogenic variants at cleavage site present a problem, especially for a therapeutic approach. If the approach is to be adapted for a clinical setting, we have to circumvent DSBs. Since, its conception, base editing has enabled the community to perform targeted chemistry on specific DNA nucleobases, importantly without introducing DSBs (Anzalone et al., 2020). This is of particular relevance to my thesis work as not only we shall avoid generation of resistant oncogenic clones, but also the adenine base editor bodes well to correct the majority of *KRAS* point mutations e.g. G12D, G13D and G12S which are all G>A point mutations. Therefore, I decided to test *KRAS* mutation inactivation using base editing. For *KRAS* base editing, owing to the lack of a well-positioned *Sp*Cas9 NGG PAM, I used a less restrictive Cas9 variants that favor NG PAM, while testing different sgRNAs. The three generations of base editors used were: xCas9-3.7-ABE7.10, NGCas9-ABEmax and NG-ABE8e. The lab of David R. Liu has used directed evolution to generate Cas9 proteins

with broadened PAM compatibility. Phage-assisted continuous evolution (PACE) and selection in bacteria that links phage replication to broader PAM compatibility generated xCas9-3.7 variant, which displays higher activity on non-NGG PAM sequences (especially NGT and NGA PAMs) than that of wild-type *Sp*Cas9 and was tested with the newly developed ABE7.10 (J. H. Hu et al., 2018). *KRAS* mutant cell lines were first infected with a lentiviral vector encoding the base editor, in addition to Puromycin resistance gene and GFP (Fig. 15). The base editor expressing cells were maintained in Puromycin and subjected to a cell competition assay, reminiscent of the prior *Sp*Cas9-infections, after infecting ~50% of the cells with a sgRNA-TdTomato encoding lentivirus and monitoring over time.

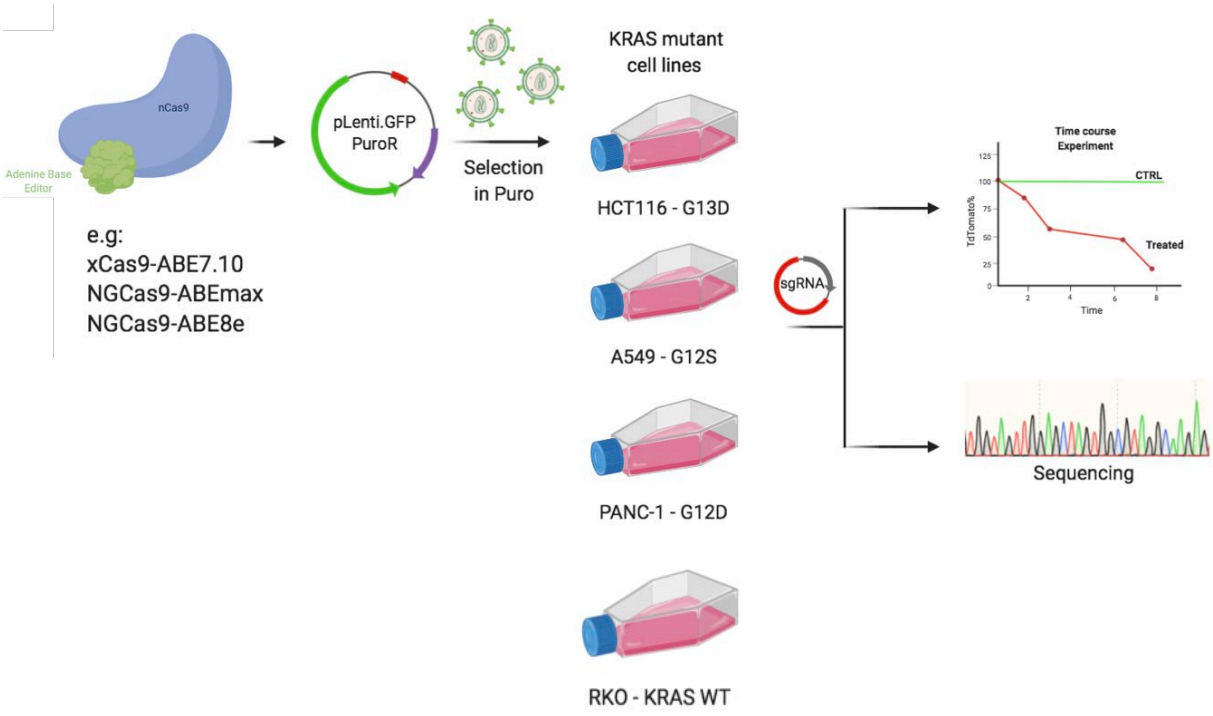
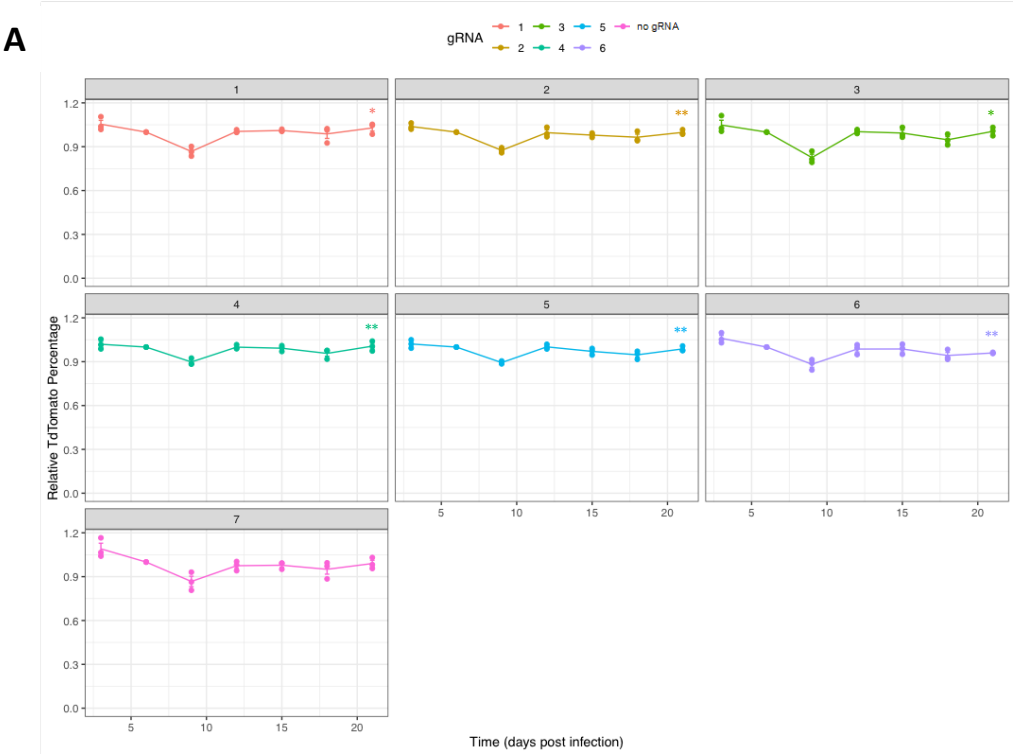


Figure 15 Adenine Base Editing Experimental Strategy. Key steps are highlighted.

First off, in both HCT116 and PANC-1 cells expressing the xCas9-3.7-ABE7.10 base editor, all the tested sgRNAs did not result in any apparent compromise of viability, demonstrated by the comparable growth rate of the cells bearing the ABE7.10-sgRNA complex targeting mutation relative to the non-infected cells (Fig. 16A, B).



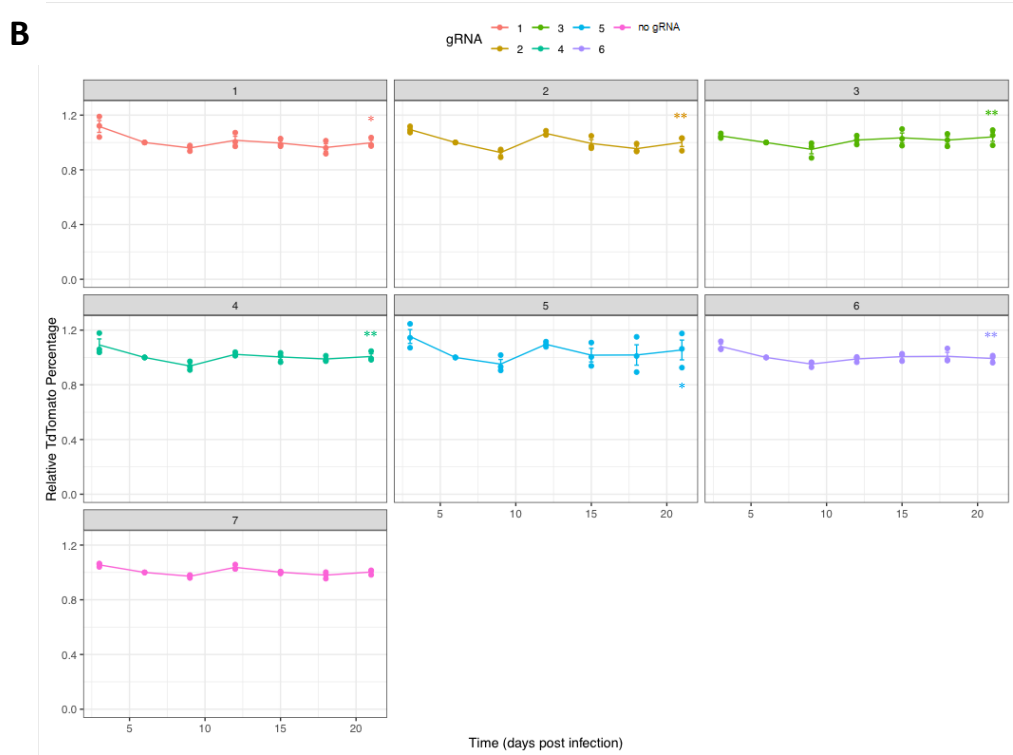


Figure 16 Adenine Base Editing in HCT116 (A) and PANC-1 (B) cells using xCas9(3.7)-ABE (7.10). Cells expressing the xCas9.ABE7.10 Base editor were infected at ~50% infection rate and monitored over time. In each cell line, in addition to the no gRNA negative control (panel 7), six sgRNAs were tested (panels 1-6) targeting KRAS G13D and G12D in HCT116 cells and PANC-1, respectively. The relative abundance of cells treated with each sgRNA is shown over time as function of TdTomato% relative to day zero. Day zero timepoint refers to three days post infection, the initial timepoint for measuring TdTomato signal. Error bars presents means ± SD from biological triplicates performed in three independent experiments. Significance was assessed by means of Student t test; *P < 0.05; **P < 0.005.

Next, Nureki and co-workers used structure-guided rational design to develop *Sp*Cas9-NG, a Cas9 variant that can target all NG PAM sequences with varying activities, with higher efficiency than xCas9-3.7 (Nishimasu et al., 2018). In parallel, the Liu lab optimized the ABE7.10 variant through codon usage, nuclear localization signals (NLS), and ancestral reconstruction of the deaminase to yield ABEmax (Koblan et al., 2018). I tested the newly developed *Sp*Cas9-NG fused to the codon optimized ABEmax for base editing *KRAS*. However, all the tested sgRNAs did not result in any apparent decrease in viability, demonstrated by the comparable growth rate of the cells bearing the ABEmax-sgRNA complex targeting mutation relative to the non-infected cells (Fig. 17A, B).

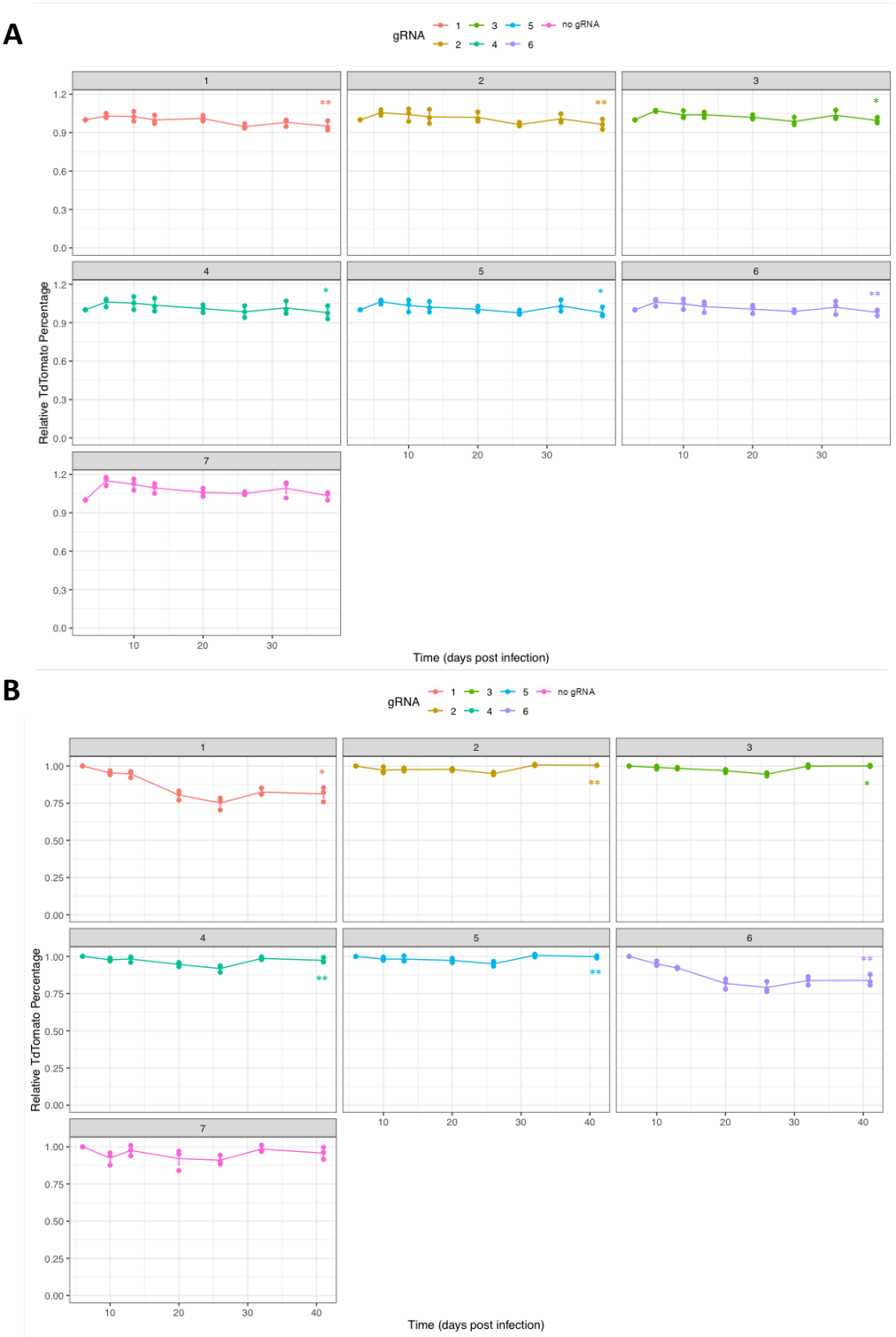
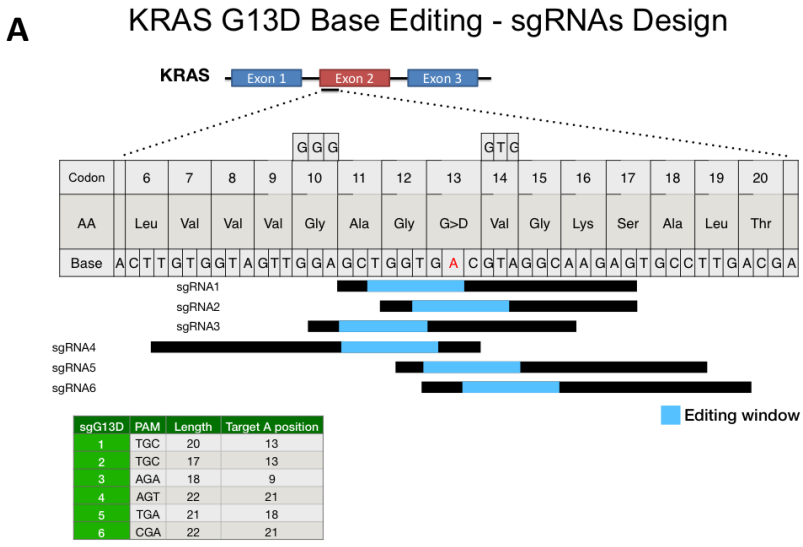
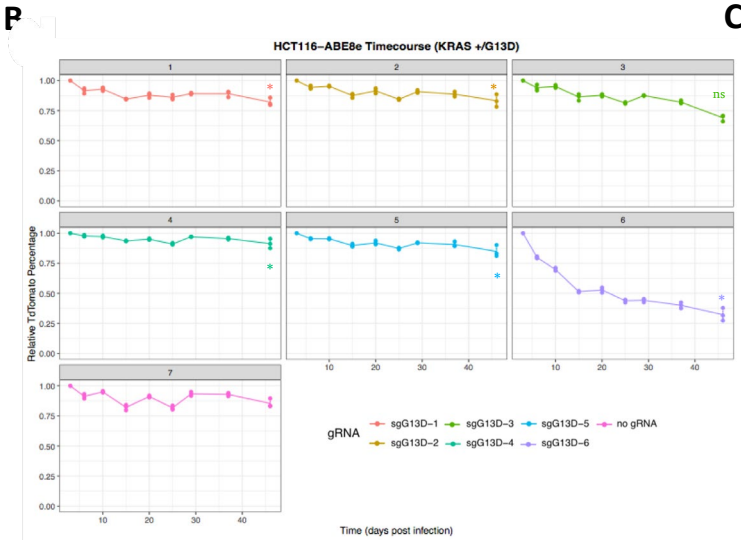


Figure 17 Adenine Base Editing in HCT116 (A) and PANC-1 (B) cells using ABEmax. Cells expressing the NGCas9.ABEmax Base editor were infected at ~50% infection rate and monitored over time. In each cell line, in addition to the no gRNA negative control (panel 7), six sgRNAs were tested (panels 1-6) targeting KRAS G13D and G12D in HCT116 cells and PANC-1, respectively. The relative abundance of cells treated with each sgRNA is shown over time as function of TdTomato% relative to day zero. Day zero timepoint refers to three days post infection, the initial timepoint for measuring TdTomato signal. Error bars presents means ± SD from biological triplicates performed in three independent experiments. Significance was assessed by means of Student t test; *P < 0.05; **P < 0.005.

In the third and last editor tested, I sought to test most recent addition to the adenine base editing toolbox, ABE8e. Interestingly, ABE8e enabled efficient A•T-to-G•C editing with all Cas9 and Cas12 domains tested, including LbCas12a and *Sp*Cas9-NG (Gaudelli et al., 2020; Richter et al., 2020). I set out to test the newly developed *Sp*Cas9-NG fused to ABE8e for *KRAS* base editing. In HCT116 cells, out of six different sgRNAs tested (Fig. 18A), only sgG13D-6 showed a sustained depletion of TdTomato cells over time, suggesting that infected cells are experiencing a growth disadvantage (Fig. 18B).



sgG13D	PAM	Length	Target A position
1	TGC	20	13
2	TGC	17	13
3	AGA	18	9
4	AGT	22	21
5	TGA	21	18
6	CGA	22	21



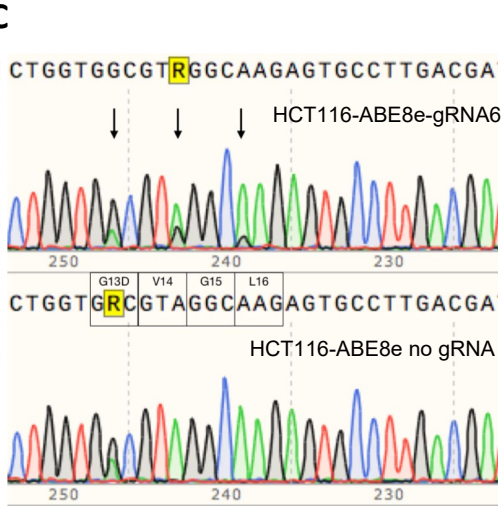


Figure 18 Adenine Base Editing in HCT116 cells using ABE8e. A) Overview of sgRNA design strategy targeting G13D, depicting sgRNAs position with respect to mutation, highlighting the hypothetical editing window, in addition to the sgRNAs length and PAM sequence exploited. B) Cells expressing the NGCas9-ABE8e Base editor were infected at ~50% infection rate and passaged over time. The relative abundance of cells treated with six sgRNAs in addition to an empty vector control is shown over time as function of TdTomato% relative to three days post infection, the initial timepoint for measuring TdTomato signal. Error bars presents means ± SD from biological triplicates performed in three independent experiments. Significance was assessed by means of Student t test; *P < 0.05. C) Representative chromatogram of cells treated with sgRNA6 (top panel) in comparison to cells treated with empty vector (bottom panel), at 15dpi. No-gRNA control cells show sequence identity analogous to HCT116 WT cells with G13D heterozygous mutation (highlighted in box) depicted as overlapping “G” and “A” peaks. On the top panel, arrows highlight editing activity shown as “G” peaks in black.

To prove that this effect is due to base editing, we isolated gDNA from the pool of cells infected with sgG13D-6-TdTomato, did PCR amplification of *KRAS* codon 13 followed by sanger sequencing. Although, the target A seems only slightly edited, we noticed neighboring “A”s are edited as bystander mutations, seen through small “G” peaks on the chromatogram (Fig. 18C). In PANC1 cells, as six different sgRNAs were tested (Fig. 19A) fast depletion of TdTomato signal occurred over time, remarkably down to 1% as seen with gG12D-1 and sgG12D-6, suggesting that these infected cells are completely outcompeted by wild-type cells (Fig. 19B). To prove that this effect is due to base editing, we isolated gDNA from the pool of cells infected with sgG12D-1 six days after infection, did PCR amplification of *KRAS* codon 12 followed by sanger sequencing. Interestingly, we could see a high ratio change of A>G conversions, signifying that target A, and seemingly target A only, is edited and corrected back to wild-type *KRAS* with high efficiency (Fig. 19C).

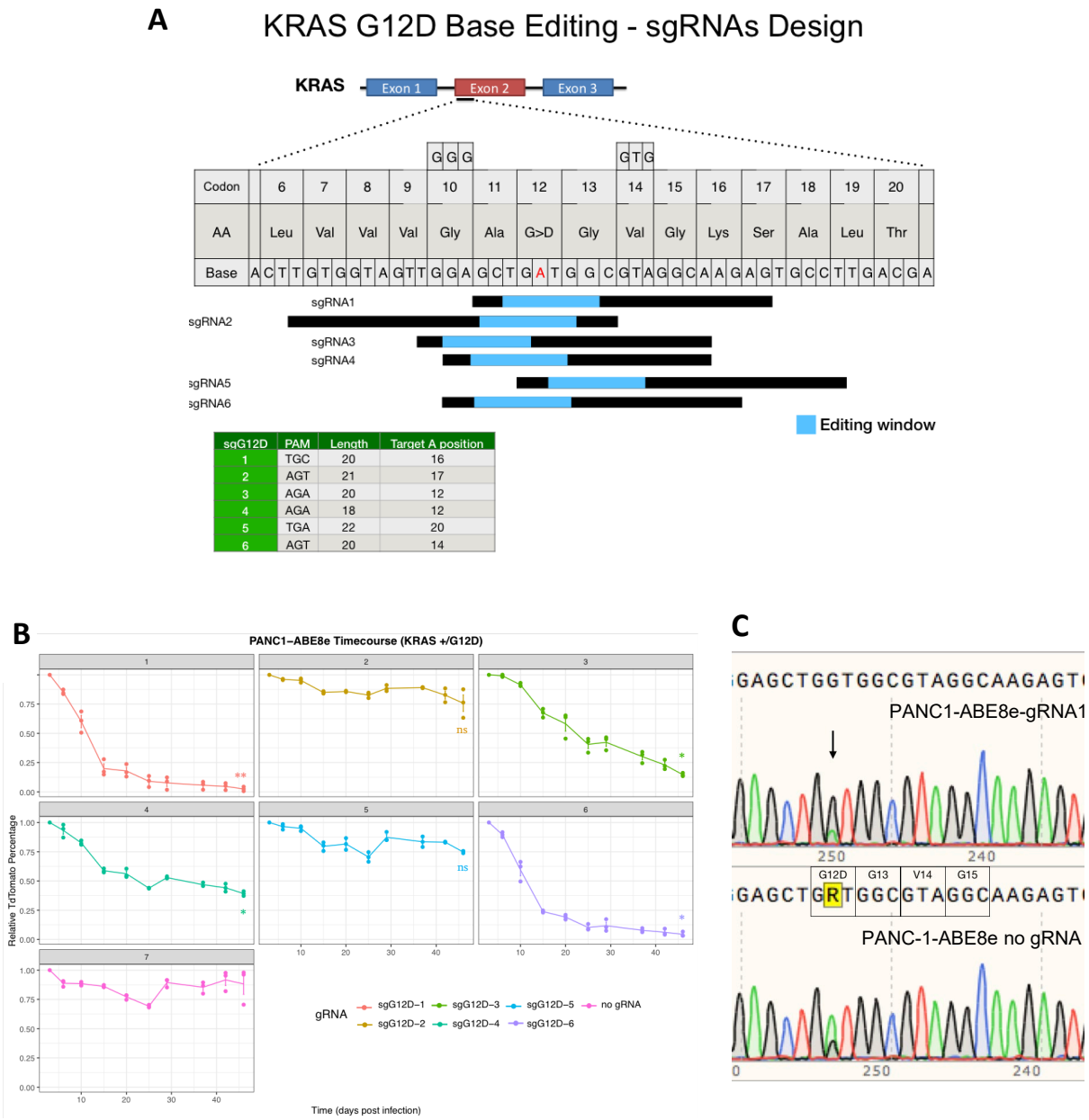


Figure 19 Adenine Base Editing in PANC-1 cells using ABE8e. A) Overview of sgRNA design strategy targeting G12D, depicting sgRNAs position with respect to mutation, highlighting the hypothetical editing window, in addition to the sgRNAs length and PAM sequence used. B) Cells expressing the NGCas9-ABE8e Base editor were infected at ~50% infection rate and passaged over time. The relative abundance of cells treated with six sgRNAs in addition to an empty vector control is shown as a function of TdTomato% relative to three days post infection, the initial timepoint for measuring TdTomato signal. Error bars presents means ± SD from biological triplicates performed in three independent experiments. Significance was assessed by means of Student t test; *P < 0.05; **P < 0.005. C) Representative chromatogram of cells treated with sgRNA1 (top panel) in comparison to cells treated with empty vector (bottom panel), at 6dpi. No-gRNA control cells show sequence identity analogous to PANC1 WT cells with G12D heterozygous mutation (highlighted in box) depicted as overlapping “G” and “A” peaks. On the top panel, arrows highlight editing activity shown as “G” peaks in black.

Importantly, ABE8e-expressing RKO cells (*KRAS* WT) were infected with all 12 sgRNAs tested in HCT116 and PANC1 as a test for off-target effects, no toxicity was observed in RKO cells (Fig. 20).

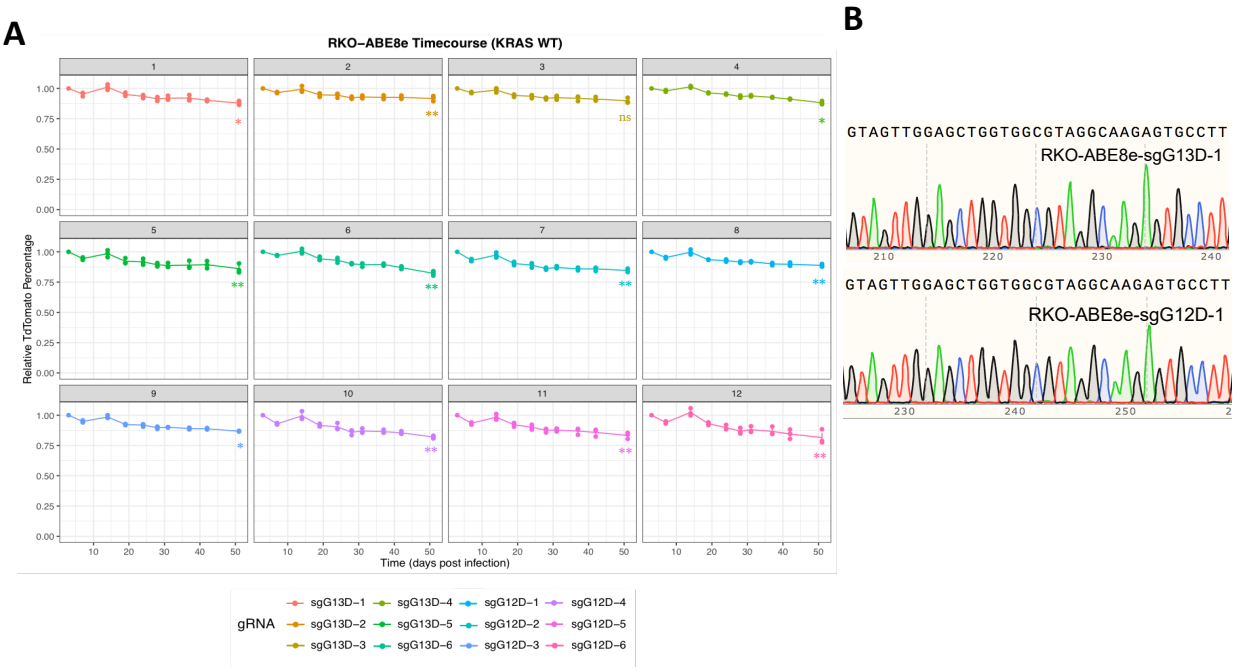


Figure 20 Off-target toxicities of ABE8e- sgG13D and sgG12D in RKO cells. A) RKO Cells expressing the NGCas9-ABE8e Base editor were infected at ~50% infection rate and passaged over time. The relative abundance of cells treated with all twelve sgRNAs used in HCT116 and PANC-1 cells is shown as function of TdTomato% relative to three days post infection, the initial timepoint for measuring TdTomato signal. Error bars presents means ± SD from biological triplicates performed in three independent experiments. Significance was assessed by means of Student t test; *P < 0.05; **P < 0.005. B) Representative chromatogram of RKO cells treated with sgRNA1 against G13D (top panel) as well as cells treated with sgRNA1 against G12D (bottom panel), at 29dpi. No off-target activity could be seen at KRAS codon 12 and 13.

Following the rapid loss of edited PANC-1 cells, especially those treated with sgG12D-1 and 6, I thought to investigate whether the corrected cells arrest in culture or undergo apoptosis in these cells in order to probe the consequences of correcting *KRAS*, I opted for fluorescent time-lapse microscopy of the base-edited cells, where I incubated the live cells while acquiring images every half an hour for a total duration of over three days. Intriguingly, the base edited PANC-1 cells seem to undergo apoptosis as observed in time-lapse microscopy (Fig. 21). Full movies of the time-lapse can be found on the following weblink: https://cloudstore.zih.tu-dresden.de/index.php/s/cYJEQscNaFLzqTi.

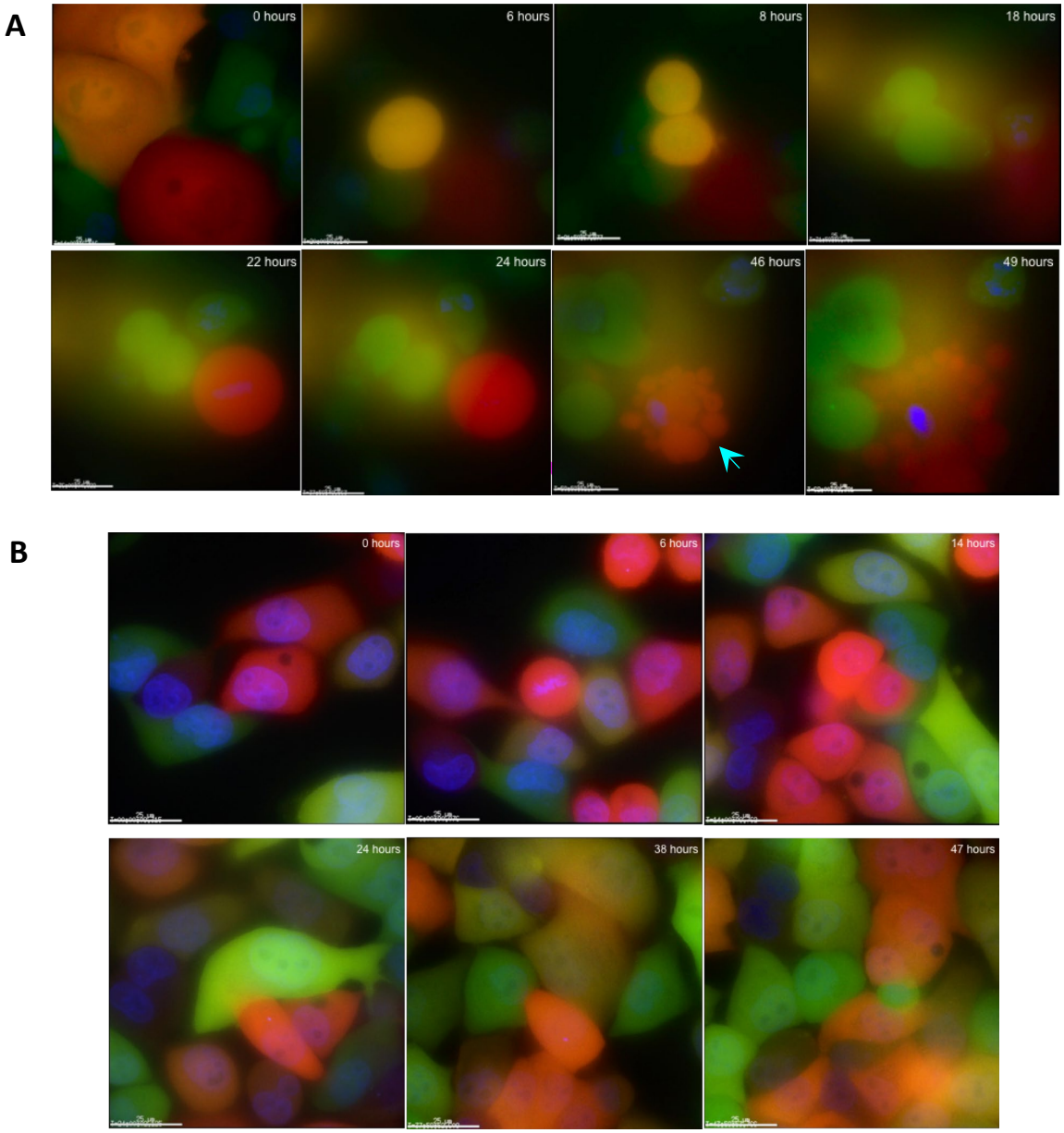
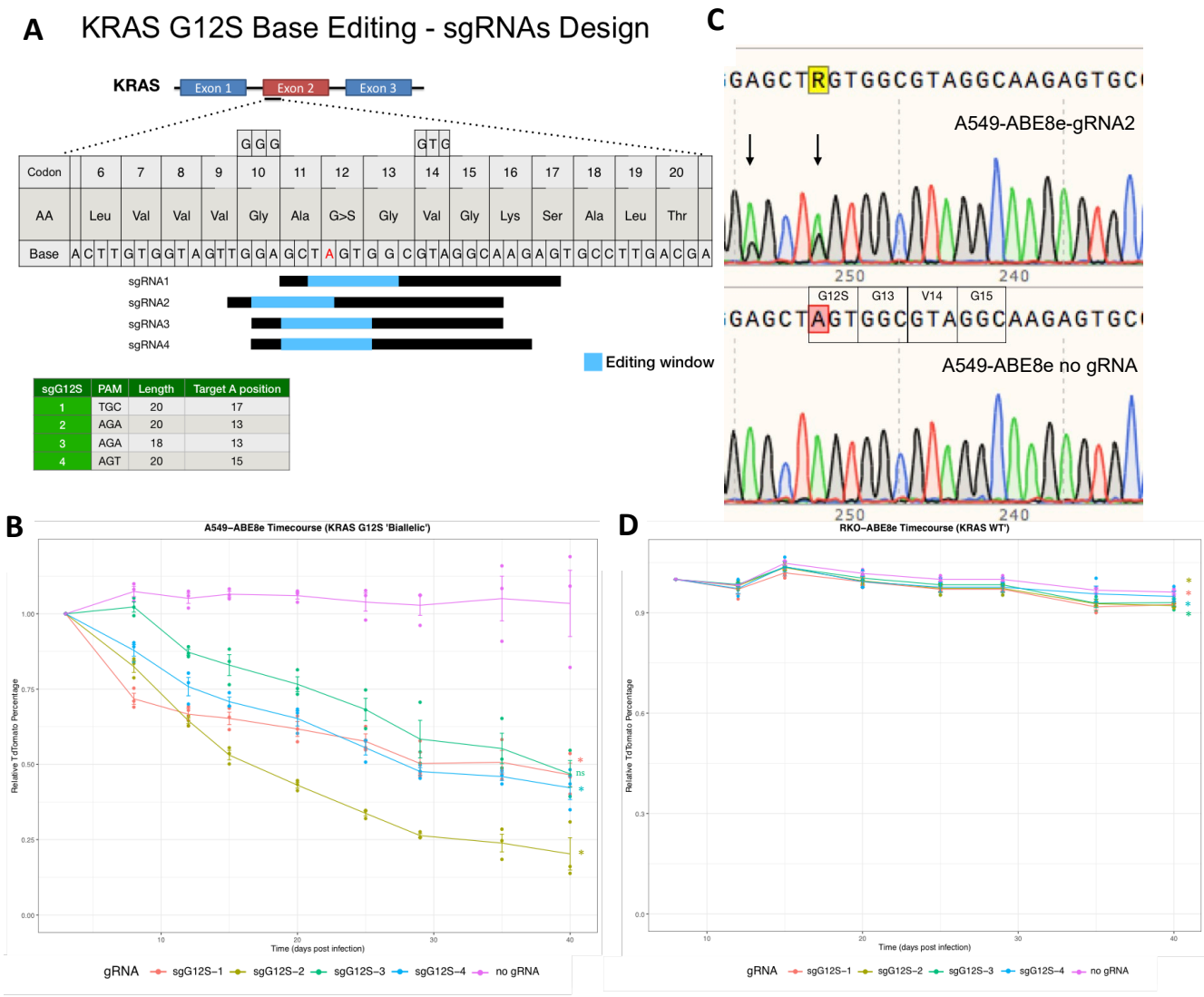


Figure 21 Time-lapse microscopy of PANC-1-ABE8e cells. *Representative time-lapse snapshots* of PANC-1 cells stably expressing ABE8e (green), after infection with sgRNA vector (red) with G12D-targeting sgRNA1 (panel A) or without gRNA (panel B). Infected cells in panel A seem to undergo apoptosis (*arrow in cyan*) whereas infected cells in panel B are proliferating at a regular rate, comparable to non-infected cells in green. *Cells were stained with SiR-Hoechst live-cell stain (blue). Scale bar= 25μm.*

In the third *KRAS* mutant line, A549 cells harboring *KRAS* G12S homozygous mutation, we tested four sgRNAs (Fig. 22A) Interestingly, all four infections, but not Empty vector, exhibited a progressive loss of TdTomato cells (Fig. 22B). Sequencing of the PCR product revealed a substantial rate of editing at twelve days after infection and to some extent to bystander A although it yields a silent mutation (Fig. 22C). Importantly, RKO-ABE8e cells were not affected by infections of the same four viruses (Fig. 22D).



sgG12S	PAM	Length	Target A position
1	TGC	20	17
2	AGA	20	13
3	AGA	18	13
4	AGT	20	15

Figure 22 Adenine Base Editing in A549 cell line using ABE8e. A) Overview of sgRNA design strategy targeting G12S, depicting sgRNAs position with respect to mutation, highlighting the hypothetical editing window, sgRNAs length and PAM used. B) Cells expressing the NGCas9-ABE8e Base editor were infected at ~50% infection rate and passaged over time. The relative abundance of cells treated with four sgRNAs in addition to an empty vector control is shown over time as function of TdTomato% relative to three days post infection, the initial timepoint for measuring TdTomato signal. Error bars presents means ± SD from biological triplicates performed in three independent experiments. Significance was assessed by means of Student t test; *P < 0.05. C) Representative chromatogram of cells treated with sgG12S-2 (top panel) in comparison to cells treated with empty vector (bottom panel), at 12dpi. No-gRNA control cells show sequence identity analogous to A549 WT cells with G12S homozygous mutation (highlighted in box) depicted as “A” peak. On the top panel, arrows highlight editing activity shown as “G” peaks in black, at target A as well as neighboring “A” as bystander editing. D) RKO Cells expressing the NGCas9-ABE8e Base editor were infected at ~50% infection rate and passaged over time. The relative abundance of cells treated with all four sgRNAs used in A549 cells is shown as function of TdTomato% relative to three days post infection, the initial timepoint for measuring TdTomato signal. Error bars presents means ± SD from experiments performed in technical triplicates. No off-target toxicities of ABE8e-sgG12S were observed in RKO cells. Significance was assessed by means of Student t test; *P < 0.05.

Cancer organoids have proven to constitute valuable hybrid systems that combine the experimental tractability of traditional two-dimensional cell lines with the cellular attributes of *in vivo* model systems (Lo et al., 2020). Seminal characteristics of these systems include a three-dimensional structure which recapitulates characteristics of self-renewal, differentiation and disease pathology, offering a valuable intermediary model between 2D cell lines and patient trials, that is amenable to drug screens and genetic manipulations (Driehuis & Clevers, 2017).

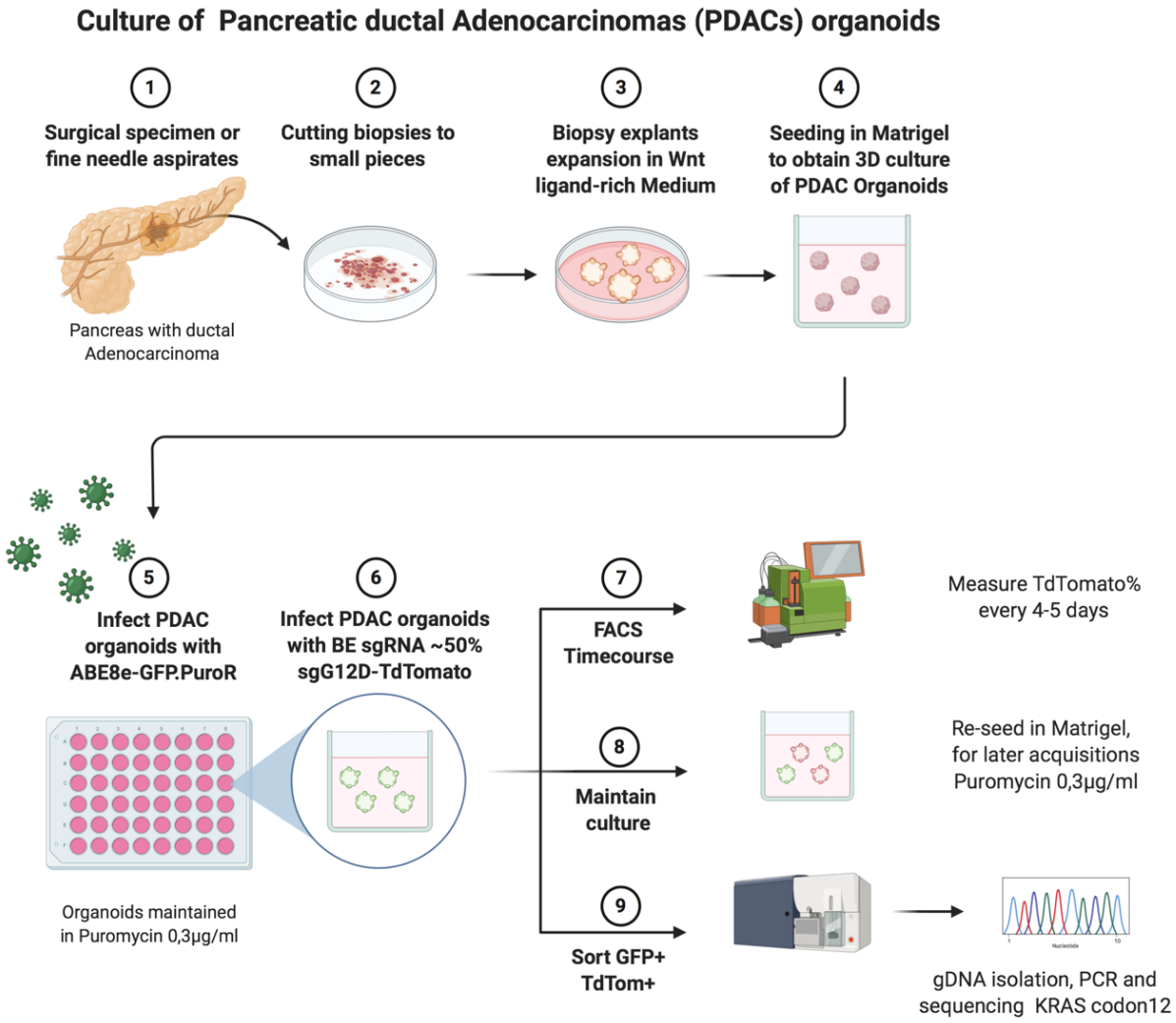


Figure 23 Graphical presentation of validation experiments in organoids. Important steps are indicated by arrows.

Therefore, I set out to validate our findings in patient-derived organoids of the highly *KRAS*-dependent Pancreatic ductal adenocarcinoma (PDAC). The experimental scheme above summarizes the procedure (Fig. 23). I adopted two *KRAS* G12D mutant organoid lines for testing base editing in addition to a *KRAS* WT control: DD442 (+/G12D), DD314 (*KRAS* +/G12D) and DD107 (*KRAS* WT) as a control for sgRNA-specific toxicity. Importantly, in

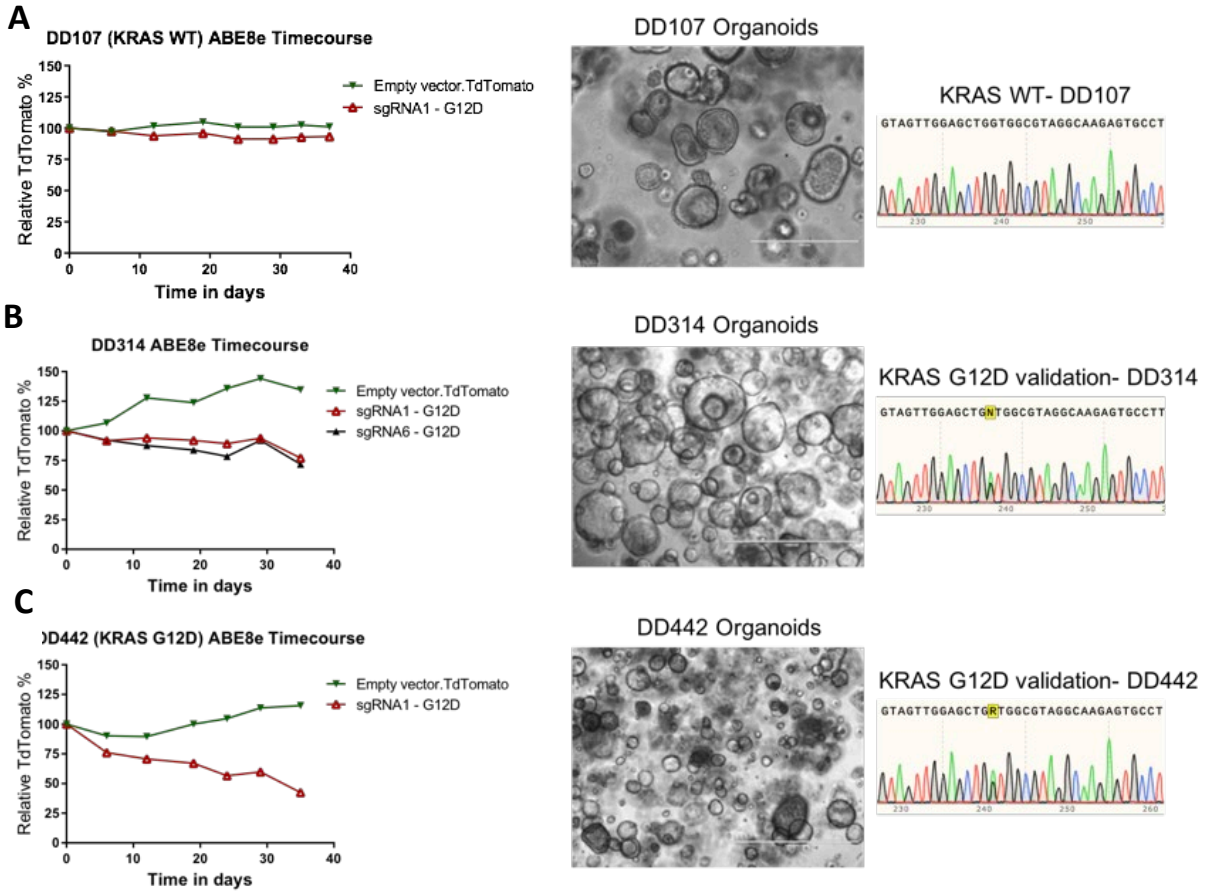


Figure 24 Adenine Base Editing in patient-derived organoids using ABE8e. Three organoids lines expressing the NGCas9-ABE8e Base editor were infected with selected G12D-targeting base editing sgRNAs in addition to an empty vector control at ~50% infection rate and passaged over time. The relative abundance of cells treated is shown over time as function of TdTomato% relative to five days post infection, the initial timepoint for measuring TdTomato signal. Presented Time course in *KRAS* WT DD107 organoids as a measure of off-target toxicities (A), and two experimental *KRAS* G12D mutant organoid lines DD314 (B) and DD442 (C) n=1. Representative bright field images highlight organoids morphology (Scale bar=1000 µm). Representative Chromatograms of KRAS codon 12 in DD107 (*KRAS* WT) and the heterzygous G12D mutant allele in lines DD314 and DD442 with peculiar overlapping peaks.

DD107 time course I could see virtually no change between infection using the lentivirus with or without the sgRNA1- G12D (Fig. 24A), demonstrating that the ABE8e-sgG12D-1 complex is well tolerated in organoids in *KRAS* WT background not inducing off-target toxicities. Using sgG12D-1, I could detect a decline in TdTomato signal over time in DD442 line (Fig. 24C), and to a lesser extent in DD314 line (Fig. 24B). Notably, repeating the experiment using biological replicates (Fig. 25A), after sorting for organoids carrying the base editor and sgG12D-1 (GFP+TdTomato+), followed by gDNA isolation and sequencing, I could detect >30% correct A>G editing 10 days after infection (Fig. 25B, C), indicating that slow regression of infected cells seems to reflect the slow base editing reaction.

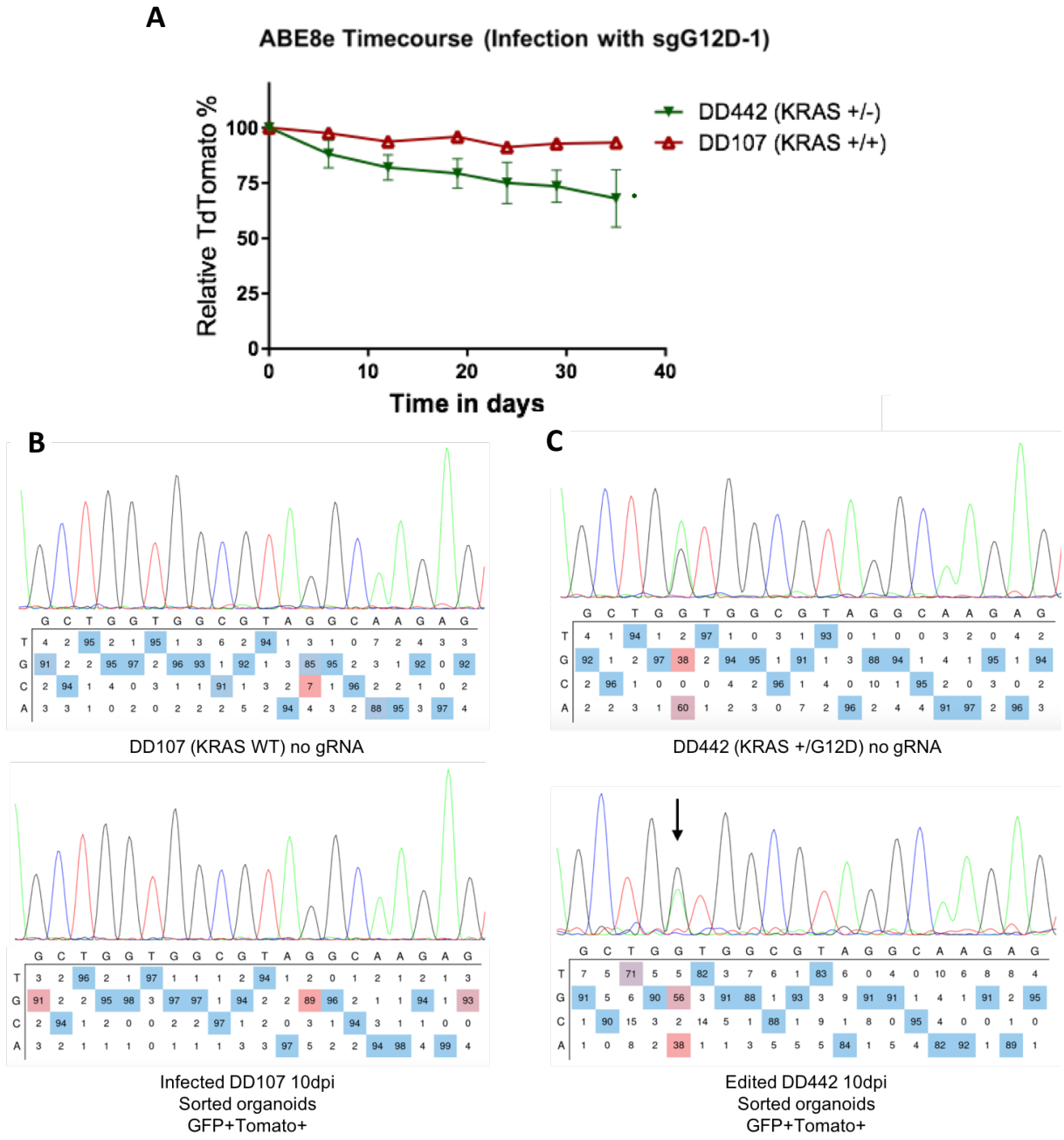


Figure 25 Time course of ABE8e and editing efficiency in patient-derived organoids. A) KRAS G12D mutant and KRAS WT organoids expressing the NG.Cas9-ABE8e base editor were infected at ~50% infection rate and passaged over time. The relative abundance of cells treated with G12D targeting sgRNA (sgG12D-1) in addition to an empty vector control in both lines is shown as a function of TdTomato% relative to five days post infection, the initial timepoint for measuring TdTomato signal. Error bars presents means ± SD from biological triplicates performed in three independent experiments. Significance was assessed by means of Student t test; *P < 0.05. B) Quantification of sanger sequencing reads of KRAS codon 12 in DD107 treated with an empty vector (top panel) and with sgG12D-1 (bottom panel), 10 days after infection. C) Quantification of sanger sequencing reads of KRAS codon 12 in DD442 treated with an empty vector (top panel) and with sgG12D-1 (bottom panel), 10 days after infection.

4 Discussion

The focus of this work was primarily to exploit the CRISPR-Cas9 system against cancer mutations. From one side the aim is to identify novel driver mutations, which has enormous implications for disease diagnosis and management. On the other hand, I sought to sharpen the tool to efficiently inactivate known oncogenic mutations, especially those which have been notoriously "undruggable", with the ultimate goal of inhibiting tumor growth *in vivo*. Of particular interest, in 2019 over twenty CRISPR-based clinical trials have started recruiting patients for conditions such as esophageal cancer, metastatic non-small cell lung cancer, refractory leukemias and blood disorders (www.clinicaltrials.gov), paving the way for therapeutic genome editing (Bolukbasi et al., 2016; Mullard, 2020).

For driver mutations identification the premise is that only a small proportion of the somatic mutations found in tumor cells drive tumor development (Stratton et al., 2009). One of the key questions in the field is to pinpoint the few phenotypically causal variants among the many variants present in cancer genomes, a daunting challenge analogous to finding needles in stacks of needles (Cooper & Shendure, 2011). With over 30 million coding mutations reported to date (COSMIC v92 – 27th August 2020) (Tate et al., 2018), scientists have resorted to computational methods to nominate novel driver candidates. A number of highly sophisticated bioinformatic tools have been developed to date to identify driver mutations (Buljan et al., 2018; Dietlein et al., 2020; Gonzalez-Perez & Lopez-Bigas, 2012; Lawrence et al., 2013; Martincorena et al., 2017; Mularoni et al., 2016; Weghorn & Sunyaev, 2017). However, computational approaches can only prioritize functional testing (Gonzalez-Perez et al., 2013) and doesn't provide definitive classification of mutations as drivers vs passengers. Programmable nucleases enable the targeting of specific genomic sequences including cancer mutations. In this study, I demonstrated a powerful method to identify candidate oncogenic cancer driver mutations amongst the background of passenger mutations utilizing the CRISPR-Cas9 system to efficiently and specifically target cancer mutations. In addition, I harnessed CRISPR and base editing systems to selectively abolish mutant *KRAS* in three cancer cell lines representing three different tumor entities, in addition to patient-derived organoids. Importantly, correction of disease-causing *KRAS* mutations has major implications on a smooth transition to the gene therapy arena. Simultaneously inactivating a driver mutation while precisely restoring the wild-type allele without any further modifications or DNA integrations is arguably the least invasive strategy. Moreover, the proposed approach is potentially scalable to other oncogenic mutations.

In this chapter I will examine the findings of the mini-screen in RKO, potential applications of a mutation-selective strategy and its limitations. In addition, I will review the promise of CRISPR targeting of mutant *KRAS* and further employing base editing. Finally, I will discuss limitations of the approaches used with prospective resolutions guiding future directions.

4.1 Pooled Screen in RKO cells nominates a putative driver mutation

The screen results identified *UTP14A*:S99delS as a putative driver mutation in RKO cells. Interestingly, while the COSMIC database reports that most *UTP14A* mutations are missense mutations, this 3-nt deletion mutation has not been found in any other cancer cell line and it has also not been found in the human population. Interestingly, a more comprehensive understanding of the role of h*UTP14A* in human cancers is emerging: h*UTP14A* has been recently demonstrated to promote angiogenesis in CRC through upregulation of transcription and secretion of *PDGFA* (P. Ren et al., 2019). Moreover, other studies have revealed that overexpression of the nucleolar protein h*UTP14A* is associated with poor prognosis in hepatocellular carcinoma (J.-Y. Zhang et al., 2017) and in esophageal squamous cell carcinoma (K. Li et al., 2019). Collectively, these studies suggest h*UTP14A* a novel prognostic biomarker that is associated with tumor invasiveness. Moreover, studies in yeast have shown that *UTP14A* is a direct interactor with *WIG-1* which encodes an RNA zinc finger protein and a known *TP53* target gene. These results come in line with previous studies that showed that h*UTP14a* caused *TP53* protein degradation (L. Hu et al., 2011), and forms positive regulation loop with *c-Myc* (J. Zhang et al., 2019), indicating its oncogenic potential. This emerging body of evidence implicates *UTP14A* in several cancers including CRC, thus motivating me to study RKO cells further to elucidate whether *UTP14A*:S99delS is a bona fide driver mutation, or rather is a passenger mutation that occurred by chance in RKO cells in an essential gene with no effect of protein functionality. In this study, for the rescue experiments I opted to use bacterial artificial chromosomes (BAC) rather than cDNA constructs, whose expression is typically driven by viral promoters, and as a consequence, the GFP-tagged genes are typically expressed at non-physiological levels, which can lead to phenotypes and mislocalization (Garrido et al., 2003; Lisenbee et al., 2003). Furthermore, the expression of cDNAs precludes the visualization of alternatively spliced variants, which may differ in their subcellular localization. Thus, for these reasons, the use of modified BACs may overcome these limitations.

Surprisingly, the growth defect due to targeting the mutation could not be rescued by the expression wild-type h*UTP14A*, in its genomic context using a BAC tagged with a GFP marker.

Importantly, protein expression and the characteristic *UTP14A* nucleolar localization highlight the efficiency of BAC transgenomics. However, the similar sensitivity in RKO WT cells and RKO-BAC-*UTP14A*-expressing cells to mutation targeting suggests that *UTP14A*:S99delS mutation is indeed an essential mutation, irrelevant of the wild-type *UTP14A* expression. To corroborate this hypothesis, the ultimate experiment is to express the *UTP14A*:S99delS mutation by BAC transgenomics, ensuring physiological expression of the mutant rescue construct while shielding it from Cas9 cleavage with the employed gRNAs. To this aim, I successfully recombineered the original BAC sequence to carry the *UTP14A*:S99delS mutation while being resistant to Cas9 cleavage, verified and established RKO line bearing BAC-Mutant-*UTP14A*. However, stable expression of this construct in RKO cells did not rescue the phenotype resulting from the mutation targeting, we saw only a milder depletion, compared to RKO WT cells upon infection with the sgmut*UTP14A*.RFP lentivirus. Importantly, maintaining cells in Geneticin and stable GFP expression throughout both rescue time courses excludes the possibility that the effect seen is due to the targeting of the BAC construct (supplementary figure 2, 3). Indeed, live cell imaging illustrated the correct nucleolar localization of the protein expressed from the recombineered BAC. However, expression and localization are often misinterpreted as functionality. Fusion of *UTP14A* to a fluorescent marker might render the protein inactive or impede binding to effector proteins, which might affect both rescue experiment that I attempted. This hypothesis could be confirmed by a test for functionality or assaying binding effector proteins, e.g. quantitative interaction proteomics using both wild-type and Mutant BAC constructs (Hubner et al., 2010). A detailed analysis for functionality will provide us clear insights guiding future directions for this project. Complementary to functional assays, whole genome sequencing or targeted amplification of likely off-target sites in RKO cells after treatment with sgmut*UTP14A*.RFP shall provide a proof if the sgRNAs used have an off-target effect that is specific to RKO genome. However, due to time restrictions I decided to pause further analysis on *UTP14A* and give priority to the base editing story. Nevertheless, while the data makes it difficult to conclude that S99 deletion is a driver mutation, a mutation-selective CRISPR-based method has successfully nominated this vulnerability, a finding which has several promising applications.

4.2 Potential applications for a CRISPR mutation-selective approach

Even if *UTP14A*:S99delS is not a driver mutation, this finding highlights the power of the CRISPR-Cas9 system to inactivate mutations in an efficient and specific manner, and offers an interesting new avenue to tackle a cancer vulnerability. In case future work proves *UTP14A* to be an essential gene that acquired a passenger mutation, which could be nominated through

our CRISPR-Cas9 mutation-selective pipeline. Motivated by advances in sequencing technologies, this could be harnessed to compromise cell growth/viability, where sequencing of cancer cells is followed by mutations mapping in reference to a healthy tissue control, mutations accumulating in essential genes in cancer cells could be exploited to draw a cancer-specific targeting strategy. This will become an increasingly important question as cancer genome sequencing moves into routine clinical practice. A mutation-specific sgRNA library could be designed in a patient-specific manner has the potential to nominate the driver mutations in a patient-specific manner, i.e. including passenger mutations in essential genes. Alternatively, the delivery of Cas9 and mutation-specific sgRNAs into tumor cells could be envisioned through oncolytic viruses (Kaufman et al., 2015) as a potent, individualized therapy that could complement current treatment strategies (Gerstung et al., 2017).

Another interesting avenue which is feasible using CRISPR-Cas9 mutation-selective approach is targeting noncoding DNA mutations. So far, most efforts have focused on the characterization of cancer mutations in protein coding genes. However, diverse mutations in enhancers, promoters, and gene regulatory elements might suggest novel driver mutations, beyond the exome. Recent reports suggest that non-coding mutations can be vital for cancer growth and survival (Diederichs et al., 2016) and transcriptional regulatory architecture (W. Zhang et al., 2018). It would therefore be interesting to extend this screening approach to non-coding mutations found in cancer cells.

One explorative aspect of this method would be to simultaneously inactivate mutations with the aim to investigate potential synergies between certain mutations or synthetic lethal mutations. In line with genetic interaction screens in mammalian cells (Du et al., 2017; Rauscher et al., 2018), I believe it is possible that similar epistasis effects can also be observed between different cancer mutations as tumor phenotype typically result from multiple genetic interactions. It would therefore be interesting to inactivate these cancer mutations simultaneously and investigate if they have enhancing, or buffering effects. This approach is of high clinical relevance to feed the resulting information back to clinicians with the cellular outcome of combinatorial targeting of mutations and how it affects cellular networks through gain, loss or switch of function in a patient-specific manner.

Finally, recent functional impact prediction models include the incorporation of transcriptome data. For instance, searching for correlations between transcription factor mutations and target gene expression has identified both well-known and novel candidate drivers (Shah et al., 2012). Cancer genes tend to be expressed in the tissues in which they are mutated (D'Antonio

& Ciccarelli, 2013), a feature only recently incorporated into high-throughput driver identification (TCGA Network, 2012; William Lee et al., 2010). Comparison with protein-protein interaction networks has also helped identify candidate drivers. Genes whose protein products are central nodes in protein interaction networks (Jonsson & Bates, 2006; Rambaldi et al., 2008) or interact with proteins from known cancer genes (Babaei et al., 2013; B.-Q. Li et al., 2013) are more likely to be drivers and thereby, could be prirotized for a targeted sgRNA library.

4.3 CRISPR-Cas9 targeting mutant *KRAS* is efficient and specific

Programmable nucleases enable the targeting of specific genomic sequences including cancer driver mutations. In this study, I demonstrated the CRISPR-Cas9 system efficiency to target cancer driver mutations. In HCT116 cells, I targeted several mutations in cancer census genes including *TERT*, *FLT3* and *KRAS*. Surprisingly, the majority of these mutations seemed dispensable for cell growth, at least for the period of the experiment which lasted one month. However, Inactivating *KRAS* G13D mutation progressively depleted targeted cells from the pool suggesting the dependence of cells on this mutation. To corroborate this hypothesis, I tested whether this phenotype is limited to HCT116 cells bearing the G13D mutation. Intriguingly, targeting mutant *KRAS* in A549 (G12S) and PANC-1 (G12D) cells rapidly depleted up to 80% of targeted cells, consistent with the phenotype observed in HCT116 cells. These findings come in line with previous studies suggesting that *KRAS* is indeed a genetic dependency in cancer cells (Dwane et al., 2020; Steckel et al., 2012; Tsherniak et al., 2017). Of note, targeted *KRAS* mutations are point mutations with a single nucleotide difference from the wild-type sequence introducing the possibility that the *KRAS* WT allele is targeted as the most likely off-target site. Interestingly, TA cloning analysis revealed nearly 50% representation of wild-type *KRAS* allele after one month of culture, reflecting the heterozygous nature of the mutation. This observation suggests that the effects seen are not due to targeting of the wild-type allele, highlighting the precision of the Cas9 nuclease to the mutant sequence, even when the difference is a single nucleotide. However, I noticed some of the sequencing reads matched other known *KRAS* oncogenic variants, that were not in parental HCT116 line. This is not unexpected, owing to the continuous expression of Cas9 nuclease in long term culture, and due to the lack of control over the DNA repair machinery, it seems some of the targeted cancer cells adapted to cas9-exerted selection pressure by repairing the DSBs to give rise to new escape clones, which are resistant to the Cas9 targeting.

Contemporary to my work, in the past three years CRISPR approaches have gained momentum towards a mutation-selective therapeutic approach. Several studies have attempted, successfully, CRISPR-based targeting of activated oncogenes, such as inactivating EGFR *in vitro* and *in vivo* (Koo et al., 2017), or harnessing PAMs to distinguish mutant *EGFR* from wild-type (Cheung et al., 2018), or another study disrupting BRAF (V600E) driver mutation to control tumor growth (Yang et al., 2017). However, *EGFR* and *BRAF* inhibitors are widely used in clinical practice with more than 15 inhibitors on the market for various indications. Meanwhile, targeting *KRAS* has been a long-standing goal in cancer research (McCormick, 2015). The fact that it is mutationally-activated in one of every three cancer patients, with no protein inhibitor in the clinic as of yet, prompted several groups to take on a CRISPR-based approach for genomic disruption of mutant *KRAS* with the aim to develop a CRISPR-based therapy for cancer patients. For instance, one study showed targeting *KRAS* G12V and G12D PDAC cell lines inhibited cell proliferation (Wookjae Lee et al., 2018), while another group reported *in vitro* and *in vivo* disruption of *KRAS*, suppressing tumor growth in immunodeficient mice (W. Kim et al., 2018), and most recently *KRAS* targeting was combined with novel delivery approaches such as polymer-based delivery of Cas9 RNP (Wan et al., 2020). However, these studies have not investigated the potential occurrence of oncogenic mutant clones following Cas9-based DSBs. Consequently, these studies have not offered any means of protection against functional Cas9 nuclease, capable of initiating stochastic oncogenic mutations after treatment as I have presented, a caveat which clearly obstructs progress towards a therapeutic approach.

4.4 CRISPR base editing corrects *KRAS* recurrent mutations

To overcome the occurrence of escape clones, I took advantage of the recently developed adenine base editing (ABE) system using a cleavage-deficient Cas9 to install targeted A>G point mutations. Using this approach, we not only inactivated *KRAS* G13D driver mutation, but also repaired it back to wild-type sequence, disabling cancer cells. Throughout this study, I tested three different versions of ABEs namely ABE7.10 fused to xCas-3.7, ABEmax fused to NG-*Sp*Cas9 and ABE8e fused to NG-*Sp*Cas9. Of note, HCT116 and PANC-1 cells treated with the ABE7.10 and ABEmax base editors continued to proliferate at a normal growth rate, suggesting the inefficient base editing (Fig. 16 and 17). Several early generation ABEs applications have been hampered by the broad variation in editing efficiencies as well as the variable editing range especially when a different Cas domain other than *Sp*Cas9 is used (T. P. Huang et al., 2019; Y. B. Kim et al., 2017). This can be influenced by several parameters, including the binding affinity of the sgRNA to the nuclease (T. Wang et al., 2014), and the

sequence preference of the deaminase (Anzalone et al., 2020). Recently, Phage-assisted evolution has enabled the development of ssDNA Adenine deaminase TadA-8e variants with increased deaminase activity (Richter et al., 2020). importantly, TadA-8e, appears more compatible with a wide range of Cas domain variants (NG-*Sp*Cas9) compared to earlier versions of ABE (Gaudelli et al., 2020). Applying the ABE8e system to correct *KRAS* driver mutations, all three cell lines tested showed some degree of editing. In the microsatellite instable colorectal cell line HCT116, only one sgRNA (sgG13D-6) showed a progressive depletion, signifying editing (Fig. 18). This is surprising, given that i) Target A is out of the editing range, which could be explained by the broader ABE8e editing range (Anzalone et al., 2020), and ii) that the sgRNA is 22 nt long. Moreover, we could detect nearby A editing activity in V14 which is benign, as it is a silent mutation, however editing K16 could be problematic (Fig. 18C). A few strategies to circumvent bystander editing include sliding the protospacer sequence to shift Cas9 binding up- or downstream, as well as optimization of alternative base editor systems using Cas9 orthologs compatible with ABE8e (Richter et al., 2020), which comes with different PAM preferences, or using alternative ABE architectures, such as rigid linkers (Tan et al., 2019). Of note, the six sgRNAs employed for HCT116 editing showed no effect on proliferation in RKO cells expressing the base editor, proving the effects seen are not due to off-target editing (Fig. 20).

In the lung adenocarcinoma A549 cells, sensitivity to base editing was observed with all four sgRNAs tested, with varying efficiencies (Fig. 21). Exemplary editing shown on chromatogram with sgRNA2 showing most efficient depletion (Fig. 21C). It is worth mentioning, cells with enriched mutant *KRAS* status, such as A549 homozygous G12S cells, exhibit a distinctive metabolic and redox management profiles compared to heterozygous cells (Kerr et al., 2016). Therefore, mutant biallelic nature in A549 might explain why we observe a consistent sensitivity to all sgRNAs employed. Importantly, RKO cells expressing the A549-G12S-specific base editor complex have shown virtually no signs of sensitivity (Fig. 21D) demonstrating that the ABE8e-sgG12S expression is well-tolerated and doesn't induce off-target toxicities. Cells with enriched mutant (homozygous) *KRAS* such as A549 G12S cells, exhibit enhanced glycolysis and sensitivity to glucose deprivation. It would therefore be interesting to challenge the base edited cells with a glycolytic inhibitor 2-deoxyglucose and the glutathione synthesis inhibitor buthionine sulfoximine which have shown a higher sensitivity in mice bearing homozygous *KRAS*-mutant tumours (Ferrer et al., 2018). This would offer us another method to test base editing efficiency employing different sgRNAs.

Remarkably, in PANC-1 cells bearing G12D, the most common *KRAS* mutation, the edited cells were progressively lost from the pool, depleted almost entirely, after treatment with

sgG12D-1 or sgG12D-6 (Fig. 19). Moreover, sequencing of the pool of cells after treatment with sgG12D-1 showed substantial levels of editing at six days after infection (Fig. 19C), suggesting more than 70% editing efficiency (supplementary Fig. 13A). Of note, these sgRNAs employed for PANC-1 editing showed virtually no signs of sensitivity in RKO cells expressing the base editor proving the effects seen are due to correcting the mutation and not due to off-target editing (Fig. 20). Furthermore, the base-edited cells seem to undergo apoptosis as observed in fluorescent live cell time-lapse microscopy. Motivated by the extraordinary and selective editing efficiency of G12D mutation, I sought to step closer to clinical application and validate these findings in patient-derived *KRAS* G12D-mutant organoids.

The organoid culture system has recently emerged as a powerful technology to propagate patient-derived tissues as 3D structures (Boj et al., 2015). Using an adapted artificial stem cell niche environment patient-derived tissues could be propagated *ex vivo* and do recapitulate the histological traits, transcriptomic and mutational profiles of the parental tumor, offering a valuable resource for precision oncology (Tiriac et al., 2019; Seino et al., 2018; L. Huang et al., 2015) including studying *KRAS* biology (Cheng & Tuveson, 2018). I tested base editing in two PDAC *KRAS* G12D mutant organoid lines, DD442 and DD314 which have been recently characterized for *CFTR* expression (Hennig et al., 2019). Of note, DD314 carries a truncating mutation in *TP53*, whereas DD442 is *TP53* WT. Importantly, DD107 organoid line, which has been recently characterized (Seidlitz et al., 2019), is *KRAS* WT and has shown virtually no signs of sensitivity to ABE8e-sgG12D-1 complex, pointing towards no apparent off-target toxicities, one month after treatment. Moreover, sequencing reads of infected and sorted DD107 organoids represents *KRAS* WT at the *KRAS* codon 12 locus, indicative of no editing of nearby “A” bases (Fig. 24A). Interestingly, DD442 demonstrated a progressive depletion, albeit milder than PANC-1 cells using the same sgG12D-1 (Fig. 24C). Sequencing of DD442 organoid cells carrying the ABE8e complex revealed moderate editing of ~30% A>G conversions, ten days after infection, which demonstrates that slow regression of infected cells reflects the slow base editing reaction. Interestingly, in contrast to DD442, *TP53* mutant DD314 line infection with both sgG12D-1 and sgG12D-6 did not result in depletion of infected cells (Figure 24B). Although it could be explained by a lower editing efficiency (supplementary figure. 15A), it could be a first hint that *TP53* activity is required to observe a proliferative disadvantage following mutant *KRAS* inactivation, especially as all tested cell lines are *TP53* WT. Intriguingly, base editing in DD442 organoid line seems slower compared to PANC-1 cells even employing the same sgRNA. This comes in contrast with existing knowledge in organoids research being permissive to genetic manipulations with both NHEJ and HDR repair pathways active. It would therefore be interesting to investigate the kinetics and fidelity of DNA repair pathways employed to resolve Cas9-induced nicks and the subsequent correction of the

mutation in DD442 organoids. Nominating the pathway(s) influencing base editing in these cells, and importantly develop ways to modulate it, shall boost the translational aspect of this project. However, even though the use of base editing is not directly amenable for clinical application, this study did demonstrate- for the first time- the promise base editing holds for the correction of cancer driver mutations in established cell lines and in patient-derived organoids.

4.5 Addressing Limitations

4.5.1 Screen for driver mutations

Although the COSMIC database offers a powerful resource, it is not without caveats, and caution should be taken when designing the libraries. In the case described in this study, we found that the *UTP14A*:S99delS mutation was present in the homozygous or hemizygous state, rather than the reported heterozygous state in the database. In another example (a mutation in the gene MRPS30 (E134_P135delEP)), we found that this mutation is not a cancer mutation, but rather a frequent polymorphism in the human population (data not shown). This issue requires particular attention for already established cancer cell lines, because for these cells a “somatic control” cannot be retrieved, thus mutations are frequently called by comparison to a reference genome. Hence, one issue to keep in mind is the human genetic variation from the “reference genome”.

An important consideration for CRISPR-based mutation-inactivation screens is the specificity of Cas9 cleavage. Point mutations are more difficult to target as they only differ in a single nucleotide from the wild-type sequence. Hence, the wild-type sequence might be cleaved as well, making it difficult to interpret obtained data. Therefore, for the screen in RKO cells we used exclusively insertion or deletion mutations. Nonetheless, the specificity of individual sgRNAs can be predicted a priori. For instance, it is known that nucleotide differences in the seed region close to the PAM sequence improve selective cleavage (Anderson et al., 2015; Kuscu et al., 2014; Zheng et al., 2017). Hence, choosing sgRNA sequences that contain the PAM sequence close to the mutation might be preferred over sgRNAs where the mutation is distal the PAM sequence. Furthermore, recent protein engineering of Cas9 has revealed mutant enzymes with higher specificities (Kleinstiver et al., 2016). These enzymes might be advantageous in screens where single nucleotide differences have to be distinguished. Finally, to test the selectivity of a given sgRNA experimentally, the so-called traffic-light reporter

system (Gebler et al., 2016) can be employed to test sgRNAs on the wild-type versus the mutant sequence. It is noteworthy, mutations for which the design of a sgRNA was not possible might still be addressable by other programmable nucleases. For instance, CRISPR systems from other bacterial strains have been identified that include Cas proteins recognizing different PAM sequences (e.g. saCas9, Cpf1, CasX) (Esvelt et al., 2013; Kleinstiver et al., 2015a; Kleinstiver et al., 2019; J.-J. Liu et al., 2019; Shmakov et al., 2015). Furthermore, mutants of *Sp*Cas9 have been generated that recognize alternative PAM sequences (J. H. Hu et al., 2018; Kleinstiver et al., 2015b) and recently, the almost PAM-less SpG and SpRY variants have been developed. These are two *Sp*Cas9-based variants that recognize NGN and NRN/NYN (where R is A or G and Y is C or T) PAMs, respectively. Although NRN is recognized more efficiently than NYN (Walton et al., 2020).

4.5.2 ABE Targeting oncogenic *KRAS* mutations

One obvious limitation to the Adenine base editor system is its inherent restriction to correct G>A mutations only. However, Prime Editing, in principle, could be harnessed to correct other *KRAS* mutations (e.g. G>T, G12V) as well as other oncogenic insertion or deletion mutations (Anzalone et al., 2019). Prime Editing uses a fusion protein, consisting of a catalytically impaired Cas9 fused to an engineered reverse transcriptase enzyme, and a prime editing guide RNA, capable of identifying a target site and providing new genetic information to replace the target DNA sequence. It mediates targeted insertions, deletions, and base-to-base conversions without the need for double strand breaks (DSBs) or donor DNA templates.

The high activity of ABE8e variants may necessitate the use of Cas domain variants that decrease Cas-dependent off-target DNA editing or the use of ABE8e variants described that reduce Cas-independent off-target DNA and RNA editing (for example, ABE8e V106W or ABE8.17-m V106W) (Gaudelli et al., 2020; Richter et al., 2020). Moreover, because the on-target efficiencies of all genome editing tools vary by site (including exact protospacer and PAM position), cell type and cell state, a diverse toolbox that provides multiple base editor options for a given target of interest will facilitate many applications, with a potential to target cancer driver mutations (Billon et al., 2017; Kuscu et al., 2017; Jun et al., 2020; J. Yuan et al., 2018). Lastly, one factor that influence the editing outcomes is the unpredictable sequence preference of the deaminase. However, addressing this limitation, a machine learning model (BE-Hive) has been recently described capable of predicting editing outcomes and efficiencies of several base editors across different target sites (Arbab et al., 2020). Based on such

predictions, one can envision a base editing drop-out screening platform aiming to reveal novel cancer G>A driver mutations or study splice variants dependency through abrogating splice sites in cancer cell lines.

While one key focus of this study at large and base editing approach in particular was to disable oncogenic driver mutations, the system does not explore mutations in tumor suppressor genes, which could also be harnessed for cancer therapy (Y. Liu et al., 2015). To address this issue, ABE could be used to “rescue” inactivated tumor suppressor genes, particularly those inactivated through G>A mutations. Furthermore, such approach bodes well to exploit tumor suppressor genes inactivated through mutations in splice donor/acceptor sites. These are largely GT/AG sites, and thereby, many of which might be abrogated through G>A mutations. Consequently, these mutations are an ideal target for base editing correction and reversing to the wild-type sequence in tumor suppressors, while exploiting NG-Cas9 for a wider targetting scope. Intriguingly, nearly 75% of the mutations in *TP53,* reputed as the Guardian of the Genome, the gene most mutated in cancer*,* are splice site variants/indel mutations of unclear biological significance (Kastenhuber & Lowe, 2017). While *TP53* loss typically occurs through a two-hit mechanism involving a missense mutation in one allele and a 'loss of heterozygosity' (LOH) by larger deletion encompassing the other allele, presumably correction of one allele should be sufficient to revert the phenotype. In principle, adapting the NG-ABE8e, used successfully in this study for correcting *KRAS*, to correct the top three recurrent *TP53* G>A mutations, is feasible.

Strategies that directly or indirectly target *KRAS* have also been evaluated and include inhibiting the docking of *KRAS* to the cell membrane, inhibiting synthetic lethal interactions and targeting downstream signaling molecules including RAF, MEK, ERK and PI3K family proteins (Papke & Der, 2017). These indirect targeting strategies have also faced dire challenges including (i) low therapeutic index which arises from the targeting of essential cell growth and survival pathways (Papke & Der, 2017); (ii) compensatory escape mechanisms (Xue et al., 2020); and (iii) Gene or mutation redundancy due to the feedback regulation of these signaling pathways (Muzumdar et al., 2017). However, future avenues could investigate a multiplexed CRISPR base editing approach to simultaneously inactivate oncogenic *KRAS* in addition to cellular responses arising in response to it, such as upregulation of PIK3CA (Muzumdar et al., 2017). Moreover, mechanisms of resistance to *KRAS* inhibition could be investigated to modulate the therapeutic response.

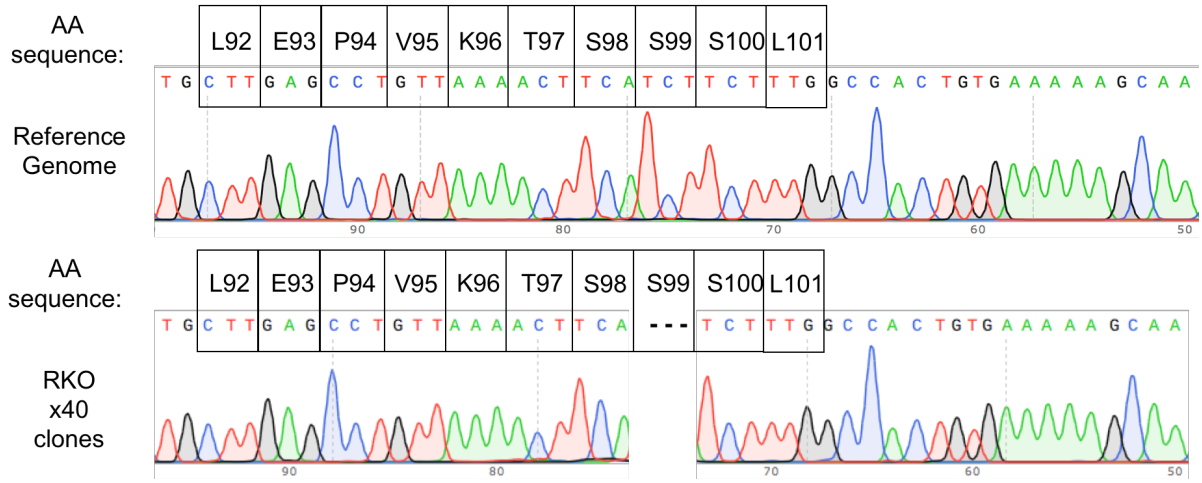
A few interesting aspects that are out of the scope of this thesis are the epidemiological correlations of different *KRAS* allele variants due to underlying biological differences between

mutant forms of *KRAS*. In particular, recently, Dow and colleagues have characterized the differential impact different *KRAS* mutant residues at codon 12 and 13 possess *in vivo* with respect to impact on tumor initiation and maintenance in the pancreas. Moreover, organoids harboring *KRAS* alterations seem to be differentially sensitive to targeted therapies (Zafra et al., 2020). These findings pose an interesting question of whether different *KRAS* alleles vary in their abilities to promote carcinogenesis, due to their inherent biological differences, or is it rather due to allele-specific co-segregating mutations in other genes, or perhaps this diverse allelic spectrum someway is linked to the underlying biology of the tissue of origin. Furthermore, an intriguing perspective is questioning whether the signaling downstream from *KRAS* alleles is qualitatively different or quantitatively different and, more clinically relevant, whether the differences in downstream signaling pathways are sufficient to design patient-tailored allele-specific therapies. Another related question, assuming equal CRISPR/Base-editing efficiencies, is whether the dependency on different *KRAS* mutations is linked to certain downstream pathways, or is it more associated with *KRAS* preferential engagement to distinct effectors GEFs and GAPs? Studies in mice suggest that genetic background plays a role in allele choice. It is still debatable whether the human genome carries modifiers of *KRAS* allele selection, for example single nucleotide polymorphisms that directly or indirectly affect *KRAS* expression.

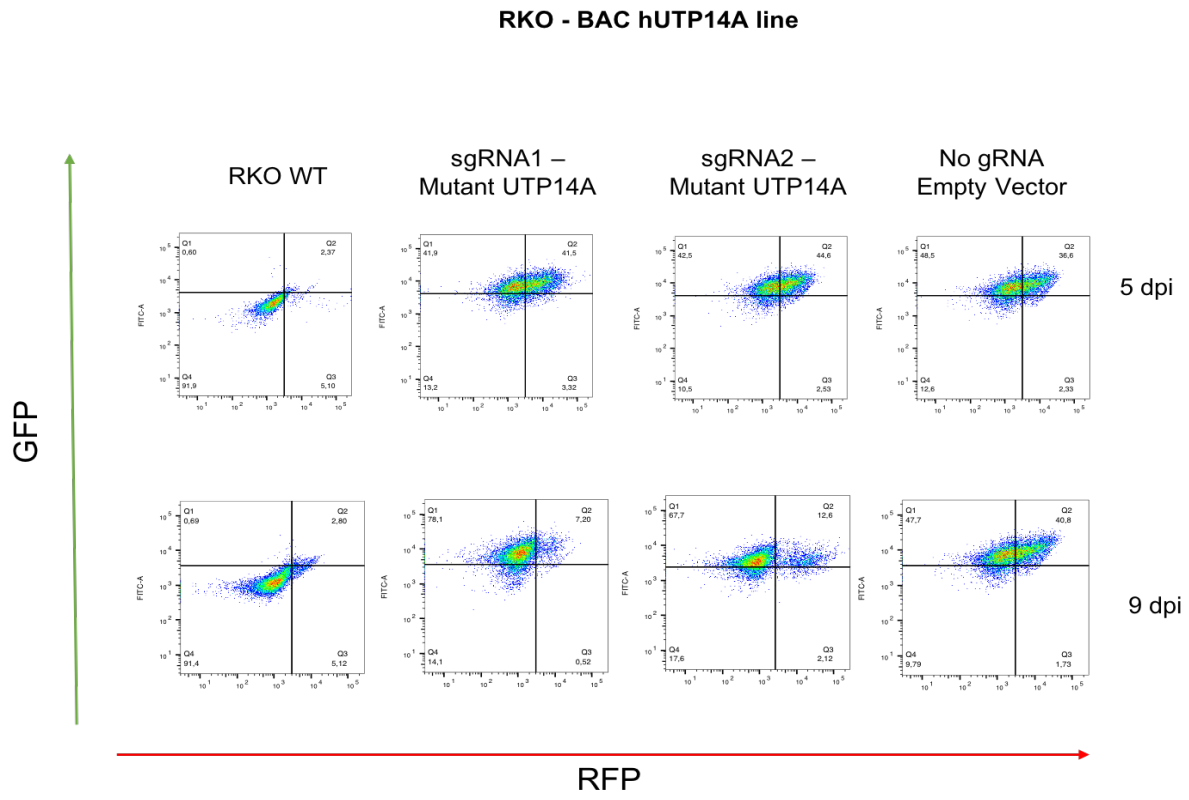
4.6 Conclusion

The presented data demonstrate that CRISPR-Cas9-mediated inactivation screens of cancer mutations in established cell lines are possible and cancer driver mutations can be nominated using this pipeline. The latest release of the COSMIC database (COSMIC v92 – 27th August 2020) reports more than 30 million coding mutations and more than 15 million Non-coding variants providing a rich resource to initiate screens to probe the vulnerability of cancer cells upon inactivation of cancer mutations. Applying the method to primary patient samples would have the advantage that the cancer mutations are frequently validated by sequencing a somatic control from the same patient as a control, which is impossible in cell lines. However, obtaining enough material and establishing efficient transduction protocols are likely challenges to transfer this method to primary cancer cells, although rapid progress in culturing primary cancer cells as organoids (Drost & Clevers, 2018; Neal & Kuo, 2015) might facilitate this process. Additionally, on a different yet complementary approach, this thesis provides a proof-of-concept that base editing technology can be harnessed to correct cancer driver mutations. To my knowledge, no one has shown a targeted efficient oncogenic correction using base editing to date, especially for *KRAS* mutations which has a huge translational potential. Furthermore, *KRAS*-corrected patient-derived organoids demonstrate a growth disadvantage, although strategies to enhance base editing efficiency are currently being investigated. The delivery of the ABE8e-sgG12D complex to treat a *KRAS* G12D mutant mouse model would be my next endeavour. Recently, base editing has been harnessed therapeutically for hematopoietic stem cells editing (Zeng et al., 2020), treating a model of tyrosinemia (C.-Q. Song et al., 2020) as well as restoring visual function in a form of inherited retinal disease (Suh et al., 2021) proving the system scalable *in vivo,* although none of the aforementioned studies used ABE8e base editor. More generally, the system can be utilized to investigate the functional pathways or identify synthetic lethal interactions with oncogenic *KRAS*, as well as various *KRAS* allele dependencies. As base editors continue to advance towards clinical applications, their continued optimization to maximize their efficiency, specificity and ability to be delivered *in vivo* remains an important priority. Recent work has reported dual AAV split base editor systems that enable efficient *in vivo* base editing (Levy et al., 2020). Further optimization of Cas domains and deaminases may enable single-AAV editing approaches. Finally, motivated by the great success signal-transduction-based therapies have achieved against mutationally activated targets (Iqbal & Iqbal, 2014; A. Kim & Cohen, 2016), there has been a renewed hope that new approaches and technologies such as base editing, together with recent advances in understanding *RAS* function, may finally bring cancer's Achilles heel within range (Berndt et al., 2011).

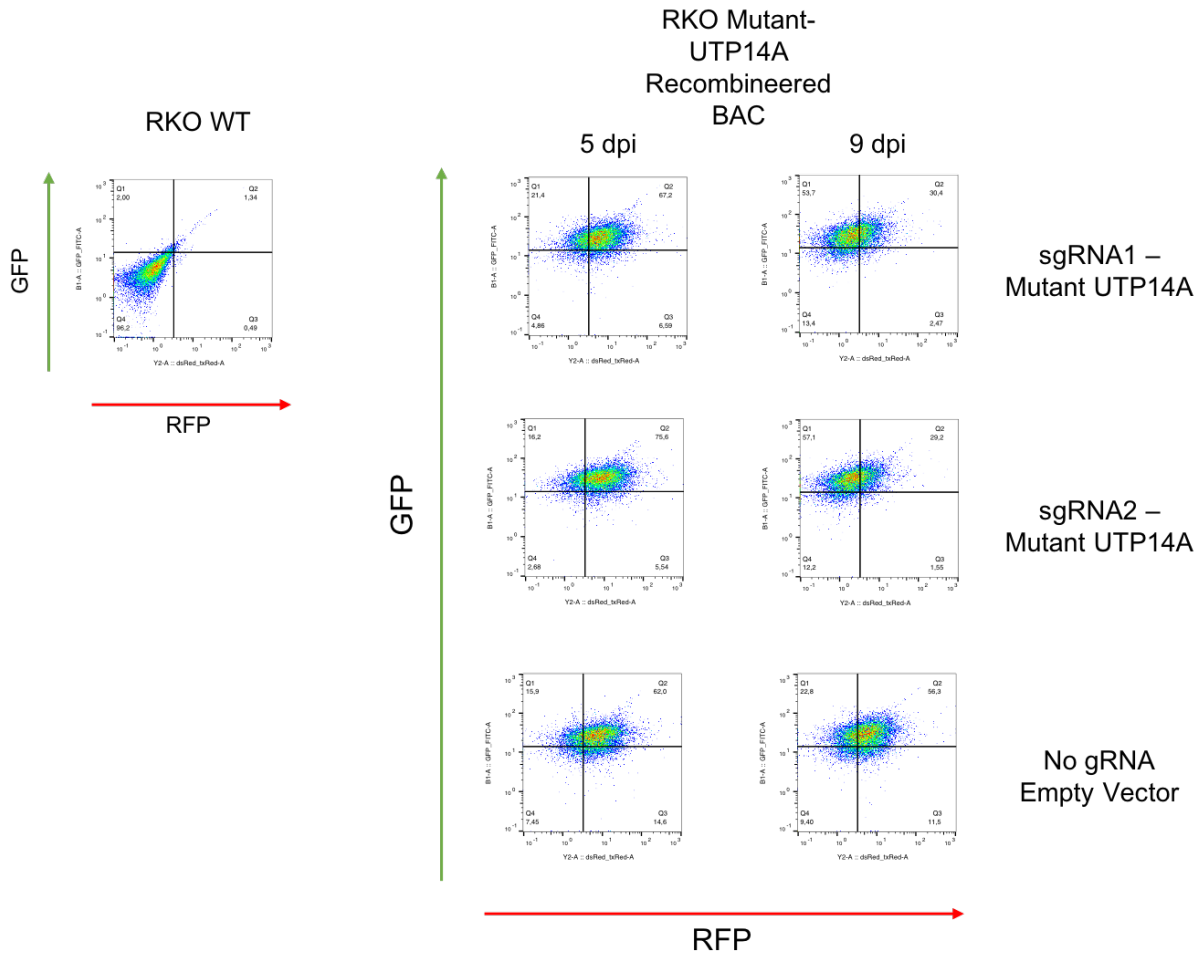
5 Supplementary Data



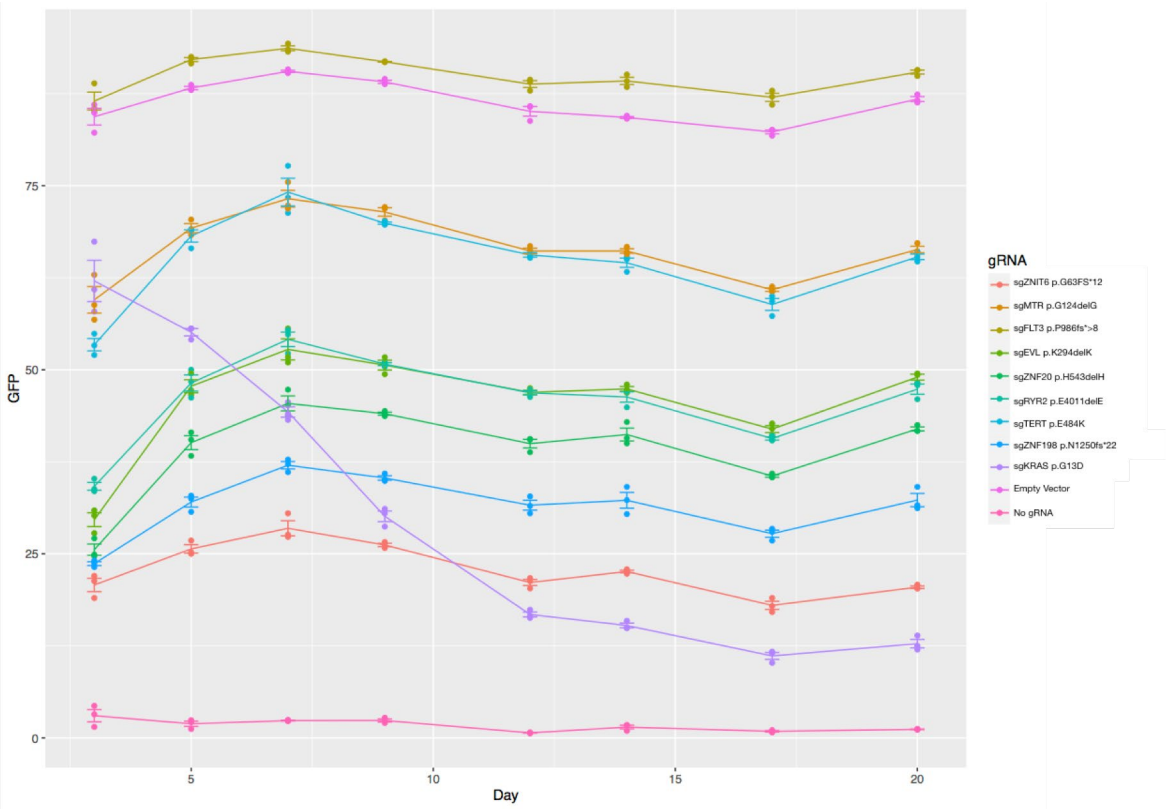
Supplementary Figure 1 Sanger sequencing of TA cloned UTP14A PCR amplicon in RKO WT cells (Bottom) compared to WT UTP14A cells (Top). A representative sequencing chromatogram shows the in-frame TCT deletion coding for Serine residue 99, seen in all 40 sequenced individual clones.



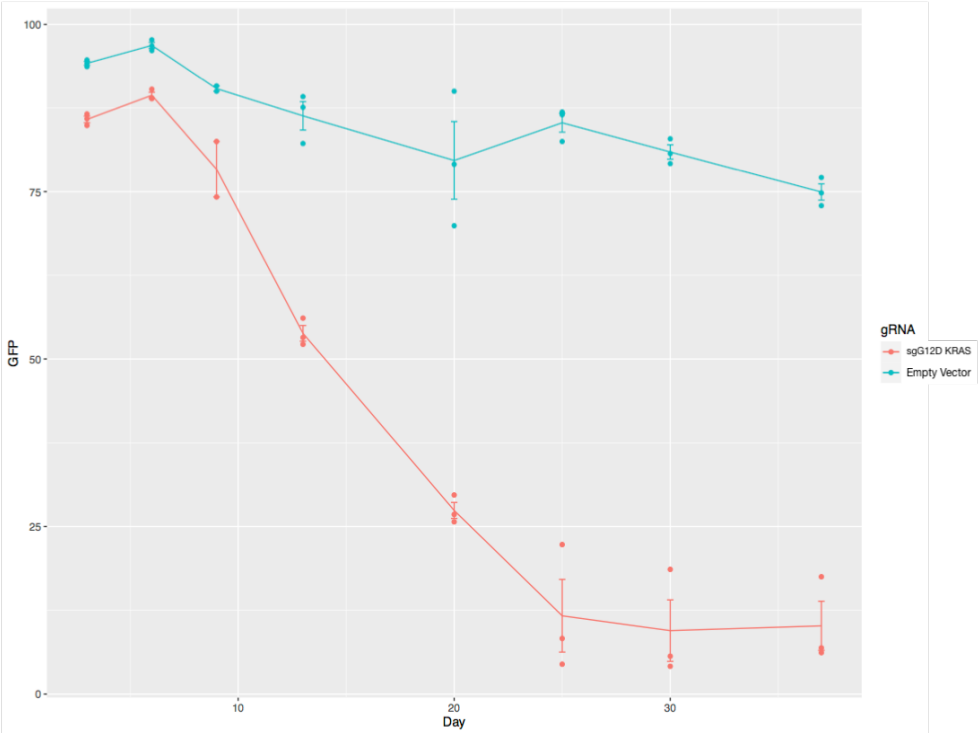
Supplementary Figure 2 Representative FACS data visualising data acquisition for UTP14A BAC rescue experiment. Gating strategy were set based on RKO wild-type cells. BAC-expressing cells were maintained in Geneticin throughout the experiment and double positive (GFP+RFP+) cells were quantified over time.



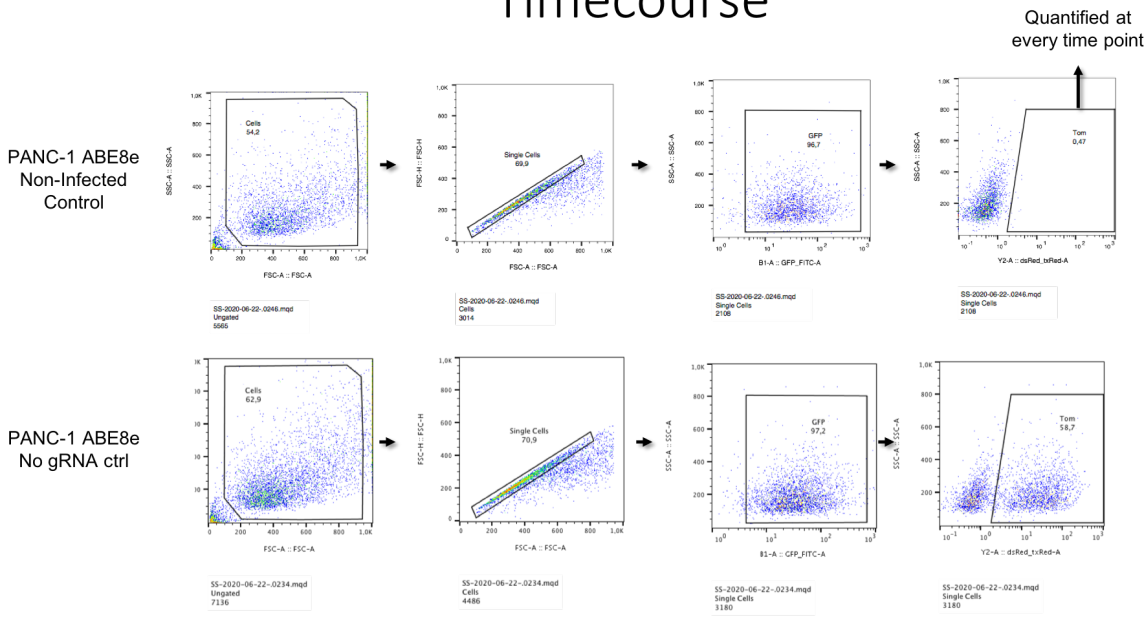
Supplementary Figure 3 Representative FACS data visualising data acquisition for mutant BAC UTP14A rescue experiment. Gating strategy were set based on RKO wild-type cells. BAC-expressing cells were maintained in Geneticin throughout the experiment and double positive (GFP+RFP+) cells were quantified over time.



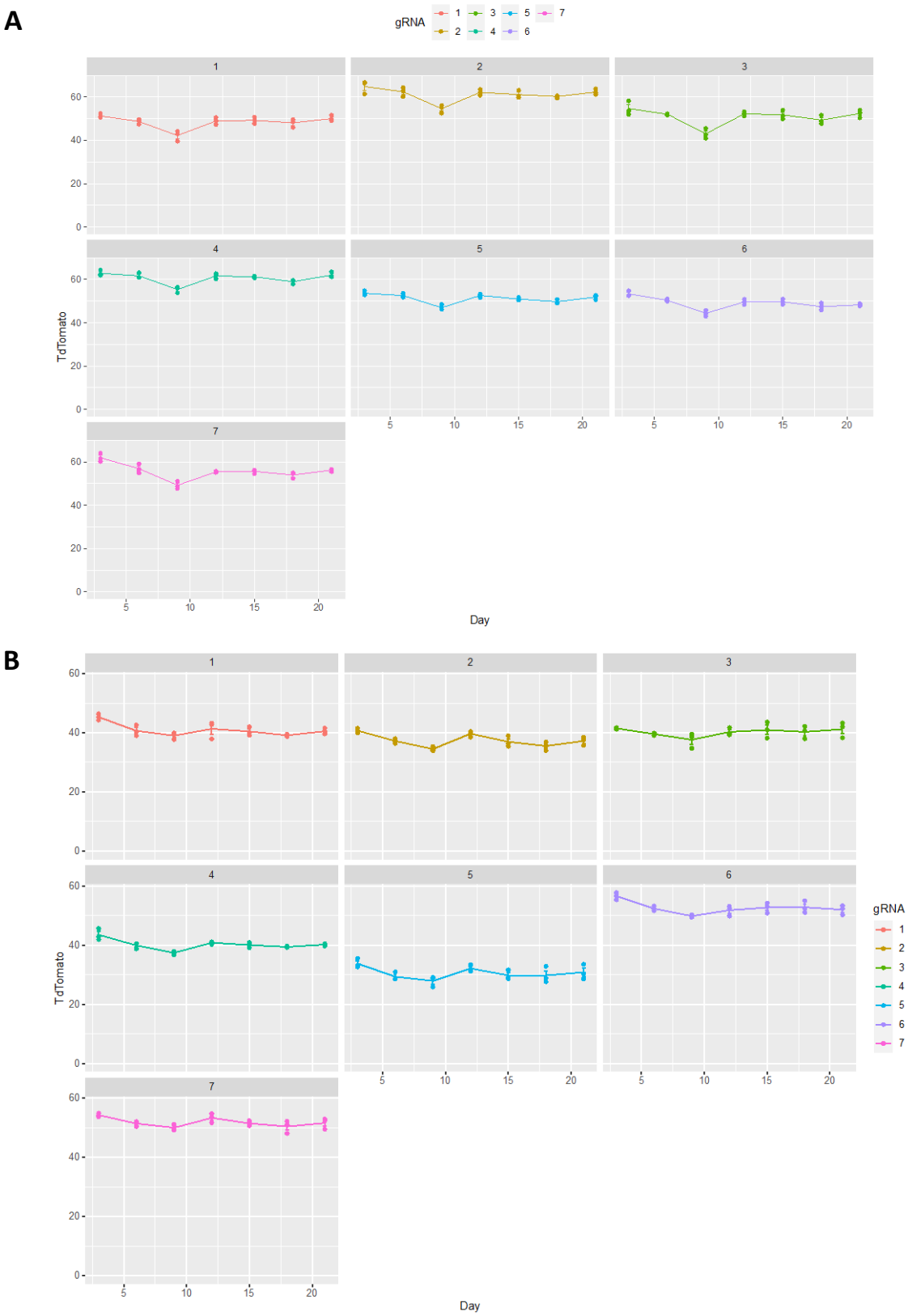
Supplementary Figure 4 Arrayed screen in HCT116. Raw FACS data points.



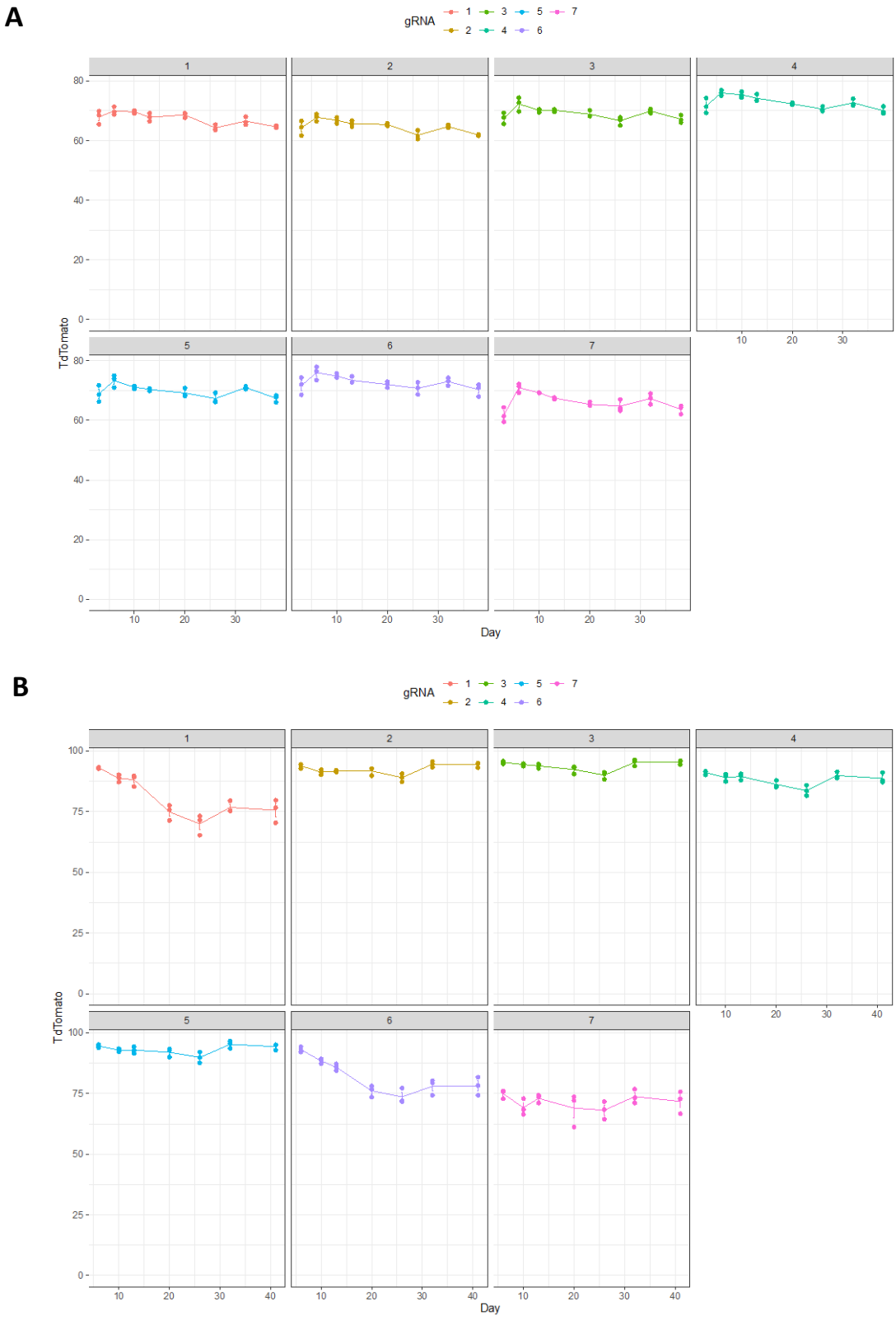
Supplementary Figure 5 PANC-1 mutant G12D KRAS Time course. Raw FACS data points.



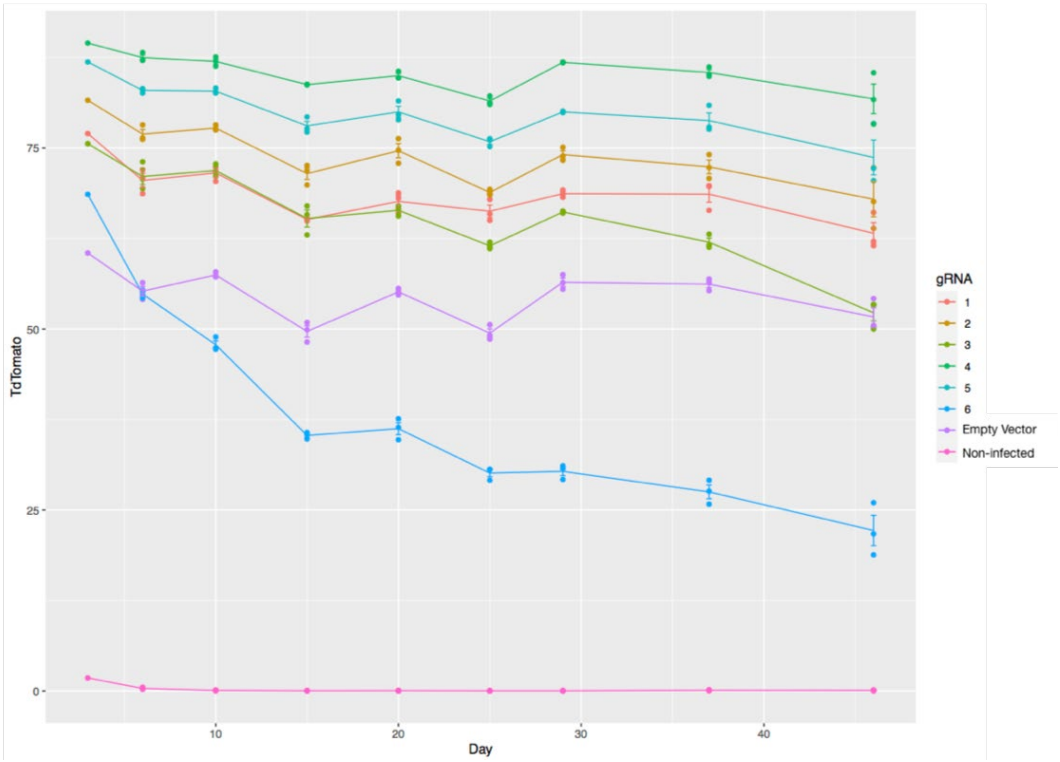
Supplementary Figure 6 Gating strategy used for Base Editing Time course experiments.



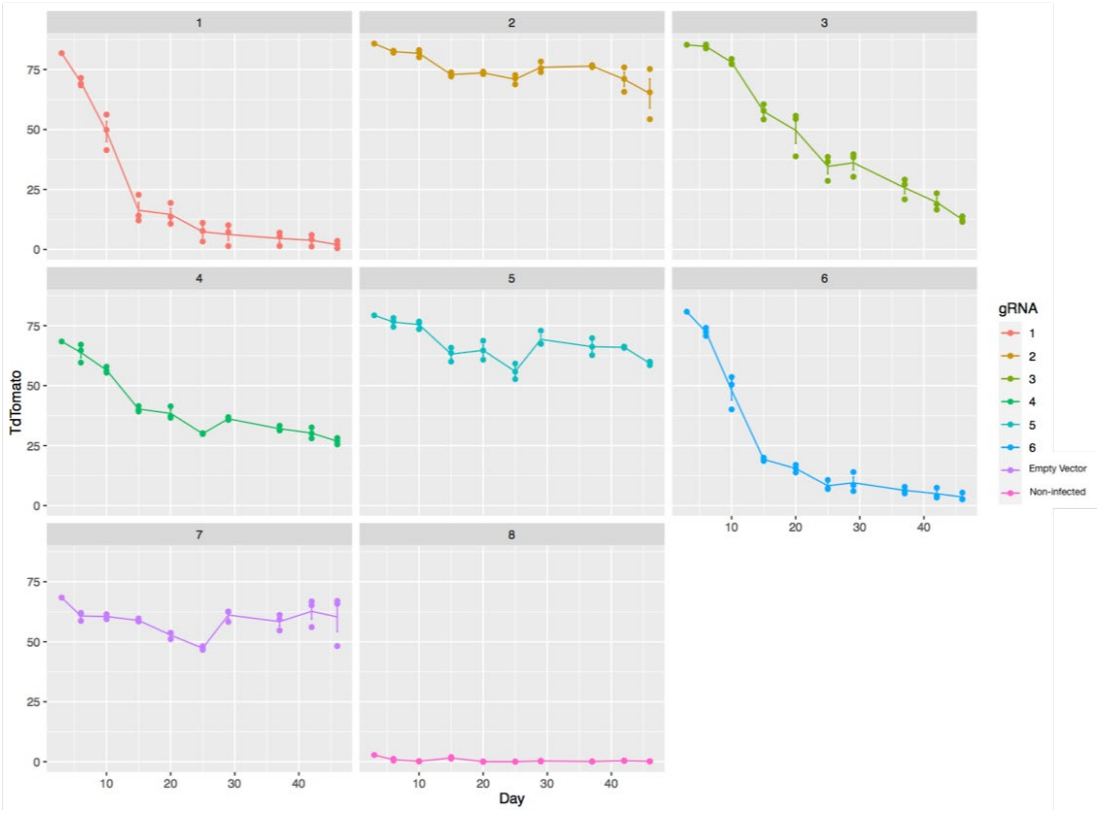
Supplementary Figure 7 xCas9.ABE7.10 Time course in HCT116 (A) and PANC-1 (B). Raw FACS data points. Infections of mutation-specific sgRNAs are Depicted here 1-6 and empty vector infection is shown in panel 7.



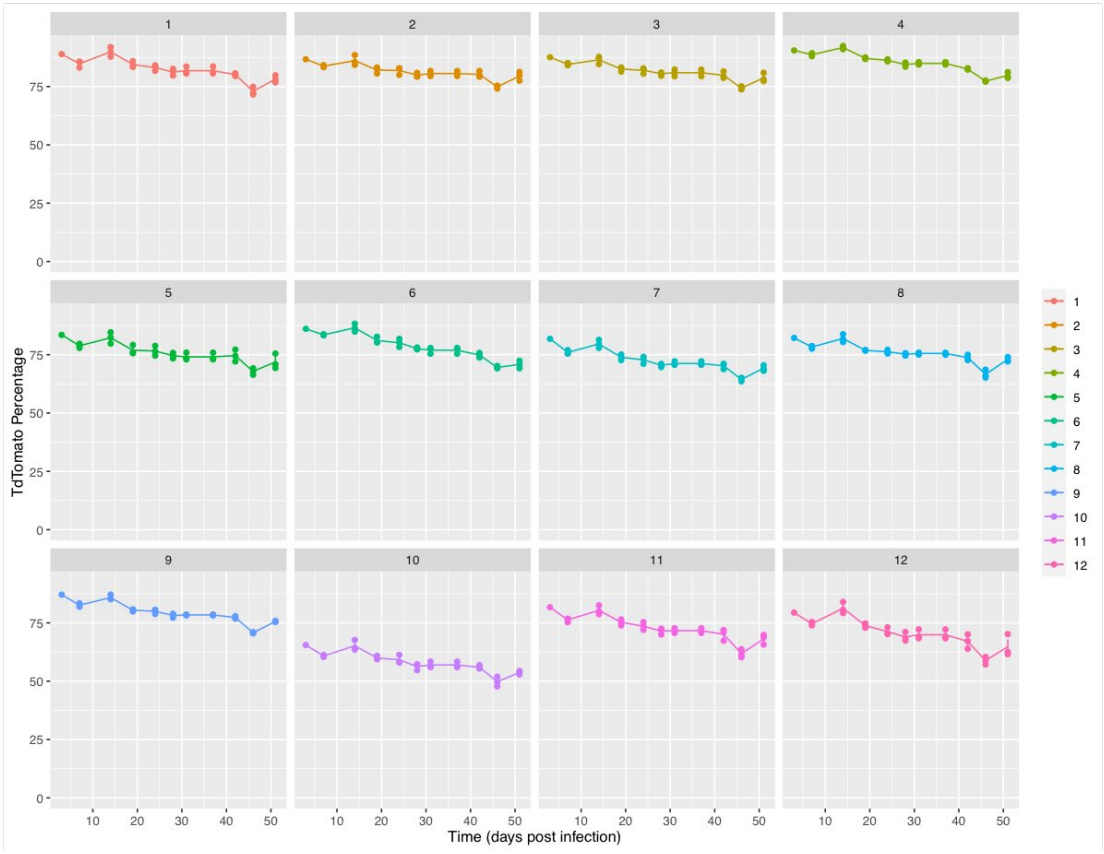
Supplementary Figure 8 NGCas9-ABEmax Time course in HCT116 (A) and PANC-1 (B). Raw FACS data points. Infections of mutation-specific sgRNAs are Depicted here 1-6 and empty vector infection is shown in panel 7.



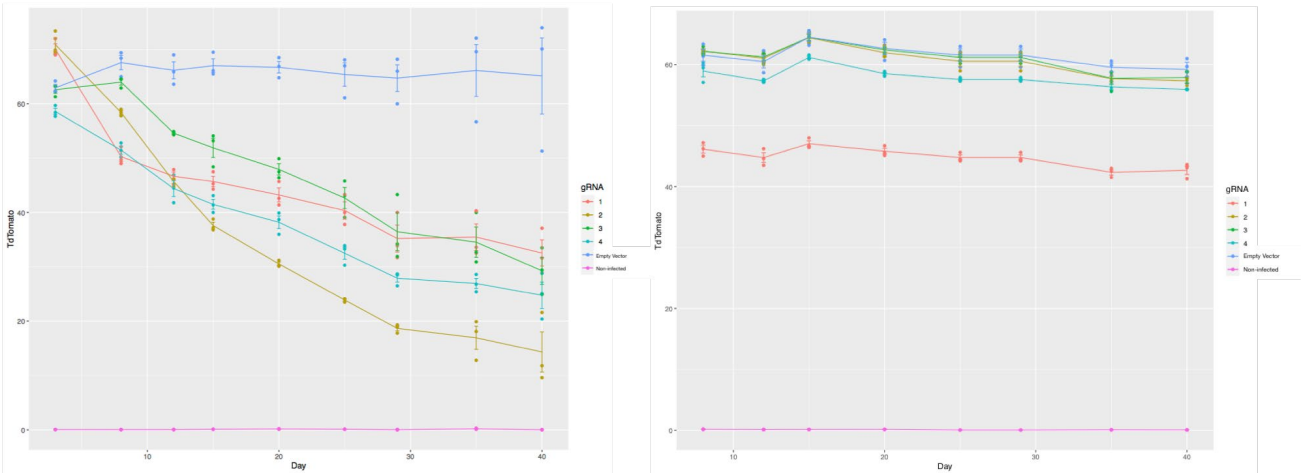
Supplementary Figure 9 HCT116 ABE8e Base Editing Time course. Raw FACS data points.



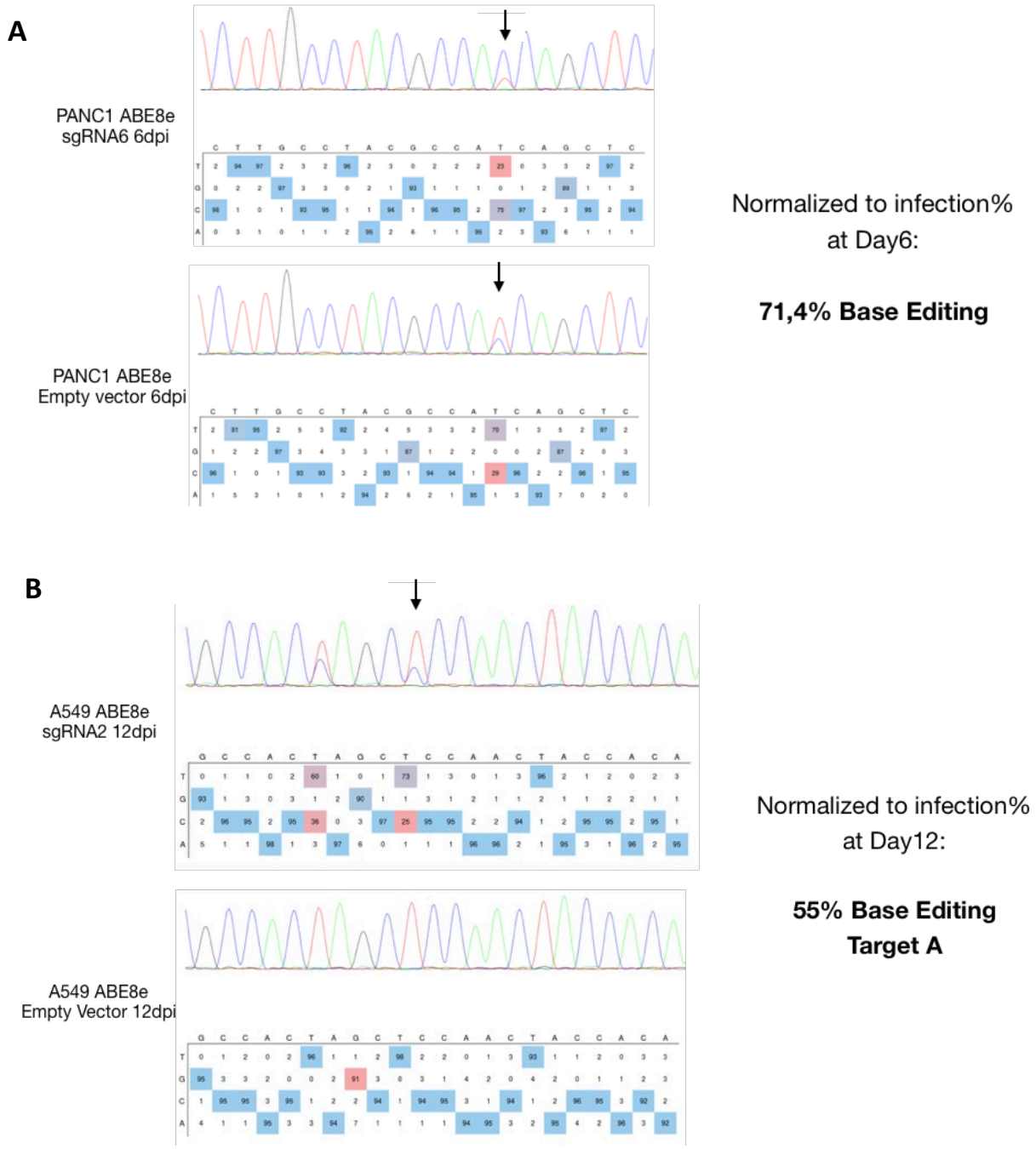
Supplementary Figure 10 PANC-1 ABE8e Base Editing Time course. Raw FACS data points.



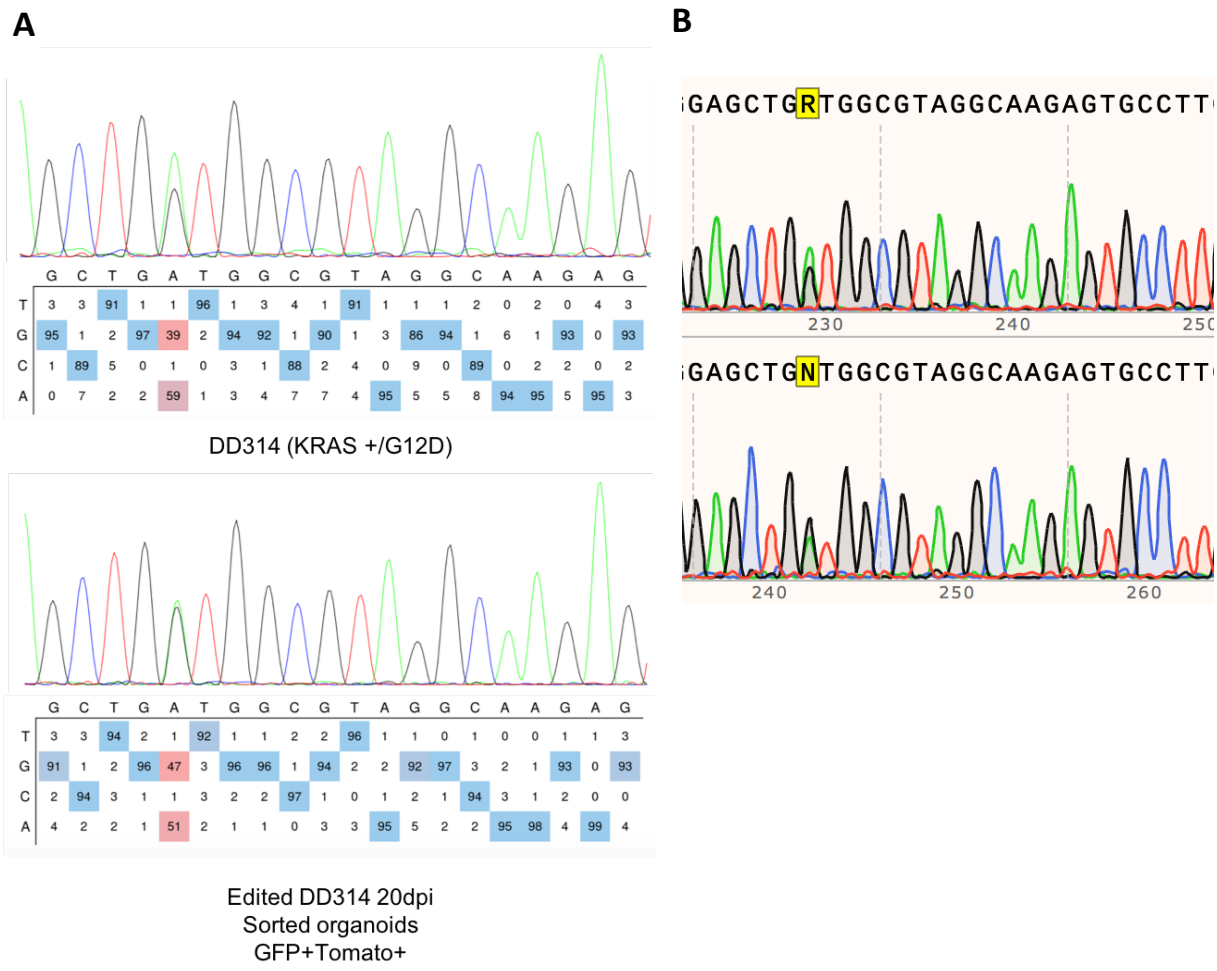
Supplementary Figure 11 Screen for off-target toxicities of ABE8e-sgG13D (1-6) and sgG12D (7-12) sgRNAs in RKO cells. Raw FACS data points.



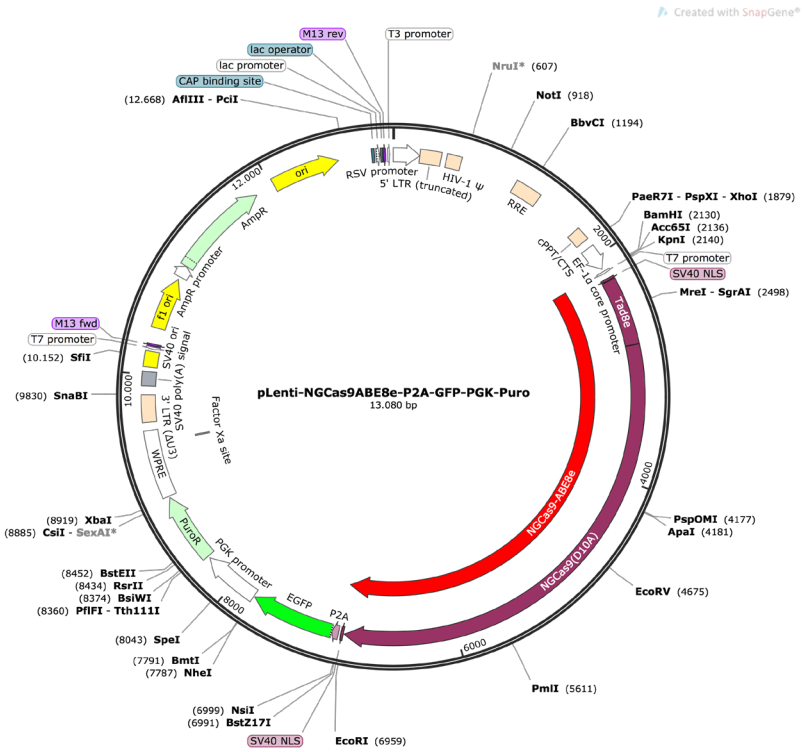
Supplementary Figure 12 A549 ABE8e (left) and RKO ABE8e (right) Time course. Raw FACS data points.



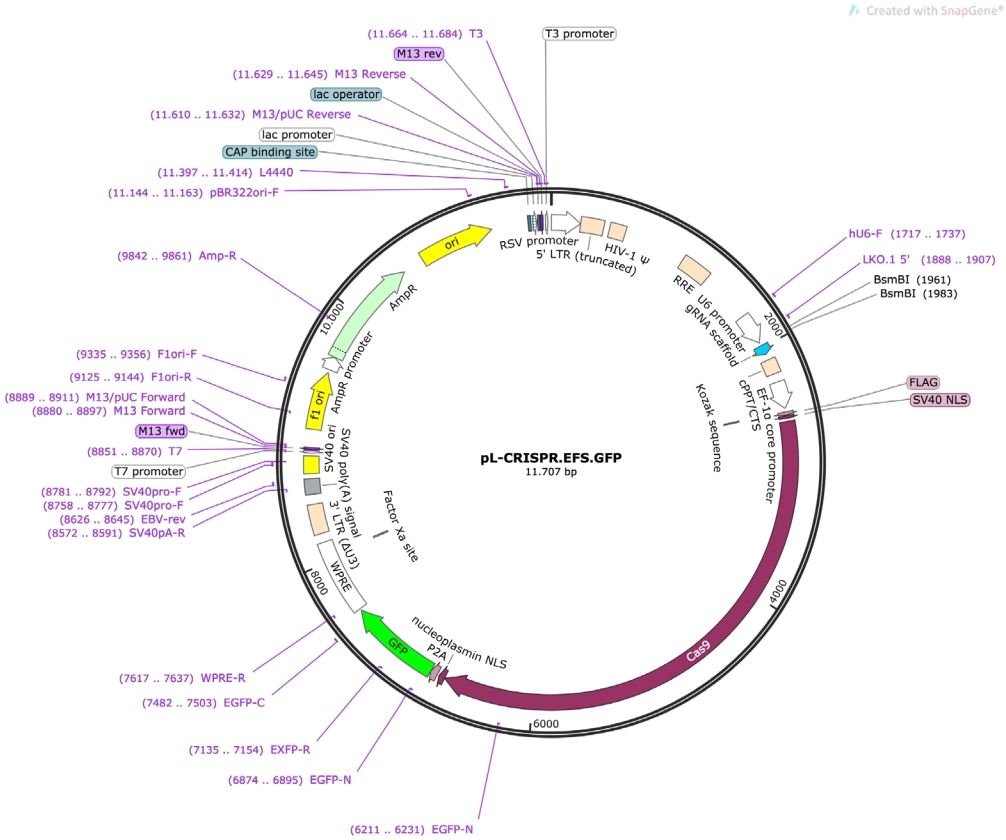
Supplementary Figure 13 ABE8e Base Editing quantification of sequencing results. A) PANC-1 cells TdTomato percentage at 6dpi timepoint stood at ~70%. Mathematical calculation of quantified Editing percentages normalised to %TdTomato+ at that timepoint to estimate the cells expressing the sgRNA, yields 71,4% Base Editing efficiency at target A (Highlighted by arrow). B) A549 cells TdTomato percentage at 12dpi timepoint stood at ~45%. Mathematical calculation of quantified Editing percentages normalised to %TdTomato+ at the same timepoint to estimate the cells expressing the sgRNA, yields 55% Base Editing Efficiency at target A (highlighted by arrow) whereas 80% Editing Efficiency at nearby bystander A.



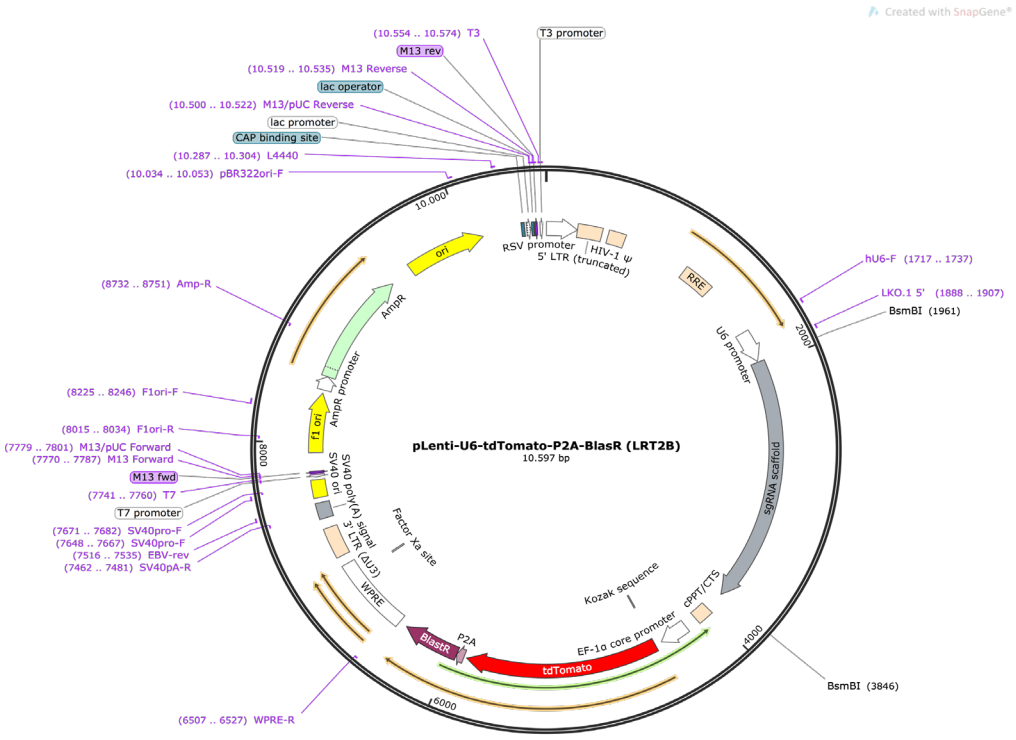
Supplementary Figure 14 Base editing in organoids Time course. A) Quantification of base editing in DD314 Time course. Cells treated with empty vector show sequence identity analogous to DD314 WT cells (top panel), however, sorted cells treated with sgG12D-1 (Bottom panel) display a moderate increase in “G” reads, verifying Editing. B) Raw Chromatogram of Base edited DD442 (bottom panel), compared to empty vector treatment (Top panel).



Supplementary Figure 15 pLenti.GFP.PGK.Puro backbone used for cloning all base editors in this study. Exemplary ABE8e vector presented.



Supplementary Figure 16 pL-CRISPR.EFS.GFP.U6sgRNA all-in-one vector used in RKO screen as well as hit validation experiments.



Supplementary Figure 17 pLenti.U6sgRNA.EFS.TdTomato.BlastR vector used for cloning Base Editing sgRNAs.

Gene	Census Gene	AA Mutation	CDS Mutation	Score (Doench 2016)	Selected sgRNA	Type	sgRNA#
MSH6	Yes	p.Y1066fs*1	c.3196_3197delTA	0,60	TGTGCCTGGCTAACTAGTCG	Deletion - Frameshift	RKO_1
HSPA14_ENST00000309584	No	p.E84_S86delETS	c.250_258delGAAACCAGT	0,63	TGACCGGATACGAGACCTTG	Deletion - In frame	RKO_2
MSH6	Yes	p.Y1066fs*1	c.3196_3197delTA	0,56	ACCCCCTCGACTAGTTAGCC	Deletion - Frameshift	RKO_3
MSH6	Yes	p.Y1066fs*1	c.3196_3197delTA	0,59	GTGCCTGGCTAACTAGTCGA	Deletion - Frameshift	RKO_4
TOP2A	No	p.T620fs*7	c.1854_1855insAGCA	0,58	GGCACCTGCTAGCACATCAA	Insertion - Frameshift	RKO_5
MSH6	Yes	p.Y1066fs*1	c.3196_3197delTA	0,64	TGCCTGGCTAACTAGTCGAG	Deletion - Frameshift	RKO_6
SOX9	No	p.Q195fs*58	c.583_584insGAGC	0,68	ATGTGCGTCTGCTCGCTCCG	Insertion - Frameshift	RKO_7
ZNF181_ENST00000492450	No	p.R537fs*3	c.1607_1608delAA	0,51	TCTCCATTATGTACTCTGAT	Deletion - Frameshift	RKO_8
S100A3	No	p.C99fs*>4	c.294_295insCC	0,49	GAGGCTACTGGGAGCACCGG	Insertion - Frameshift	RKO_9
EML2_ENST00000536630	No	p.M180delM	c.540_542delGAT	0,48	GCCAGCTCGTCTGGGATCAT	Deletion - In frame	RKO_10
TOP2A	No	p.T620fs*7	c.1854_1855insAGCA	0,45	GCTTCCTTTGATGTGCTAGC	Insertion - Frameshift	RKO_11
ZBTB48	No	p.H8fs*30	c.22_23delCA	0,63	ACCCTCACACTGCTGGACGA	Deletion - Frameshift	RKO_12
UTP14A	No	p.S99delS	c.292_294delTCT	0,26	GCCTGTTAAAACTTCATCTT	Deletion - In frame	RKO_13
RP1L1	No	p.K1578_L1579insK	c.4729_4730insAGA	0,54	AGAGCTCCTCTAGAAGCTCC	Insertion - In frame	RKO_14
RP1L1	No	p.K1578_L1579insK	c.4729_4730insAGA	0,45	CTTCTAGAGGAGCTCTCTCT	Insertion - In frame	RKO_15
ZNRF3	No	p.V249fs*117	c.739_740delGT	0,38	TGGTCTCCACACACCGCGCT	Deletion - Frameshift	RKO_16
KIAA0430	No	p.G122_S123insG	c.362_363insTGG	0,64	TACTGGTACCTCCGCTACCG	Insertion - In frame	RKO_17
PABPC1	No	p.F145fs*1	c.434_435delTT	0,53	GGATTTGTACACTGAGACGC	Deletion - Frameshift	RKO_18
ZBTB48	No	p.H8fs*30	c.22_23delCA	0,43	CTGCAGAACCCTCACACTGC	Deletion - Frameshift	RKO_19
HSPA14_ENST00000309584	No	p.E84_S86delETS	c.250_258delGAAACCAGT	0,44	TACGAGACCTTGAGGACTCC	Deletion - In frame	RKO_20
EML2_ENST00000536630	No	p.M180delM	c.540_542delGAT	0,49	GCCCATGATCCCAGACGAGC	Deletion - In frame	RKO_21
POGZ_ENST00000491586	No	p.S545_Y546insS	c.1630_1631insCCT	0,32	ACATAGGAGGCCTGAAAATA	Insertion - In frame	RKO_22
TRAF7	Yes	p.K514delK	c.1536_1538delGAA	0,57	CTGAAGTTGAAGGAGCTCAC	Deletion - In frame	RKO_23
ZNF697_ENST00000421812	No	p.E4delE	c.12_14delAGA	0,48	CTGATTATCTTGTTTCATCC	Deletion - In frame	RKO_24
S100A3	No	p.C99fs*>4	c.294_295insCC	0,59	AGGCTACTGGGAGCACCGGG	Insertion - Frameshift	RKO_25
PSMC3IP	No	p.K69delK	c.206_208delAGA	0,53	CGGCAAGCAGATCTATTTTG	Deletion - In frame	RKO_26
POGZ_ENST00000491586	No	p.S545_Y546insS	c.1630_1631insCCT	0,52	TCAGGCCTCCTATGTGACTC	Insertion - In frame	RKO_27
ITGA2	No	p.V751delV	c.2247_2249delTGT	0,29	GCCCTCTGATGTCAACTCTT	Deletion - In frame	RKO_28
THAP6	No	p.S12delS	c.32_34delCTT	0,55	GCAGCGAGCACATCCAATGG	Deletion - In frame	RKO_29
ARL6IP4	No	p.H337fs*>24	c.1008_1009delAC	0,29	TTGATCTCTCTGTCGTTCTT	Deletion - Frameshift	RKO_30
ITGA2	No	p.V751delV	c.2247_2249delTGT	0,55	TCCAAAGAGTTGACATCAGA	Deletion - In frame	RKO_31
S100A3	No	p.C99fs*>4	c.294_295insCC	0,36	GGAGCACCGGGGGGCTCTGA	Insertion - Frameshift	RKO_32
DLG1	No	p.E732delE	c.2195_2197delAAG	0,36	CACAAACCTTGTTGATTCAC	Deletion - In frame	RKO_33
HMGXB3	No	p.V690delV	c.2065_2067delGTC	0,52	GGCTGCACACCACTTTTGAC	Deletion - In frame	RKO_34
ZNF181	No	p.R473fs*3	c.1415_1416delAA	0,42	CTCTCCATTATGTACTCTGA	Deletion - Frameshift	RKO_35
ENSG00000187522	No	p.E62_S64delETS	c.184_192delGAAACCAGT	0,49	AGTCCTCAAGGTCTCGTATC	Deletion - In frame	RKO_36
HMGXB3	No	p.V690delV	c.2065_2067delGTC	0,58	ACTAAGACGCCAGTCAAAAG	Deletion - In frame	RKO_37
PPL	No	p.W893_K894insW	c.2675_2676insGTG	0,68	TTCCTTCCTGATCTTCCACG	Insertion - In frame	RKO_38
PRAMEF10	No	p.E236fs*3	c.706_709delGAAC	0,38	AAGAACGAAGATTGCTCATC	Deletion - Frameshift	RKO_39
HSPA8	No	p.S276fs*7	c.825_826delTT	0,64	CGTACCCTCTCCCAGCACCC	Deletion - Frameshift	RKO_40
DVL3	No	p.F46delF	c.130_132delTTC	0,34	CTATAAGTTCTTCAAGTCTA	Deletion - In frame	RKO_41
EML2	No	p.M33delM	c.99_101delGAT	0,47	TCGTCTGGGATCATGGGCAC	Deletion - In frame	RKO_42
IPO5	No	p.I886fs*24	c.2652_2655delATTA	0,61	ACTTACAATGAGGTTGACAA	Deletion - Frameshift	RKO_43
PKDCC	No	p.F205delF	c.610_612delTTC	0,64	TGTGAAAAACCTGTGGCAGA	Deletion - In frame	RKO_44
TTN_ENST00000342175	No	p.E9441delE	c.28322_28324delAAG	0,52	TGCAAGAAGGTACTACTGAC	Deletion - In frame	RKO_45
ENSG00000212768	No	p.S40fs*3	c.119_120delCA	0,56	AACAGAACTTTCTTTTGACA	Deletion - Frameshift	RKO_46
PPL	No	p.W893_K894insW	c.2675_2676insGTG	0,46	AGGAGGCCACGTGGAAGATC	Insertion - In frame	RKO_47
JAKMIP1_ENST00000409021	No	p.R606fs*6	c.1816_1817delAG	0,59	GCGACCTCCTCTCTCTCCTG	Deletion - Frameshift	RKO_48
EML2	No	p.M33delM	c.99_101delGAT	0,56	CGTCTGGGATCATGGGCACA	Deletion - In frame	RKO_49
S100A3	No	p.C99fs*>4	c.294_295insCC	0,38	CAGAGGCTACTGGGAGCACC	Insertion - Frameshift	RKO_50
THAP6	No	p.S12delS	c.32_34delCTT	0,52	CAAGCAGCGAGCACATCCAA	Deletion - In frame	RKO_51
JAKMIP1_ENST00000409021	No	p.R606fs*6	c.1816_1817delAG	0,32	TGGCGACCTCCTCTCTCTCC	Deletion - Frameshift	RKO_52
SgK493	No	p.F348delF	c.1039_1041delTTC	0,70	GGAGGTATGTGAAAAACCTG	Deletion - In frame	RKO_53
ARFGEF1	No	p.I1106fs*19	c.3316_3319delATAG	0,40	ACTGGAAACAGCAAGTATTC	Deletion - Frameshift	RKO_54
KIAA0430	No	p.G122_S123insG	c.362_363insTGG	0,42	GGTGGTGGTGGCCACGGTAG	Insertion - In frame	RKO_55
RP1L1	No	p.K1578_L1579insK	c.4729_4730insAGA	0,41	TCCTCTAGAAGCTCCAGGGC	Insertion - In frame	RKO_56
PRAMEF10	No	p.E236fs*3	c.706_709delGAAC	0,37	TCTTCGTTCTTTTTAGCCTT	Deletion - Frameshift	RKO_57
ENSG00000212768	No	p.S40fs*3	c.119_120delCA	0,39	AAACAGAACTTTCTTTTGAC	Deletion - Frameshift	RKO_58
S100A3	No	p.C99fs*>4	c.294_295insCC	0,65	AGAGGCTACTGGGAGCACCG	Insertion - Frameshift	RKO_59
ZNRF3	No	p.V249fs*117	c.739_740delGT	0,42	GGTCTCCACACACCGCGCTT	Deletion - Frameshift	RKO_60
UTP14A	No	p.S99delS	c.292_294delTCT	0,34	GCCAAAGATGAAGTTTTAAC	Deletion - In frame	RKO_61

Tabel 1 RKO screen sgRNAs library

Gene	Census Gene	AA Mutation	CDS Mutation	Score (Doench 2016)	Selected gRNA	Type	sgRNA#
PRDM2	No	p.K1489fs*6	c.4459_4460delAA	0,09	GATGAATGAGAAACTTTTTT	Deletion - Frameshift	RKO_62
KIAA0430	No	p.G122_S123insG	c.362_363insTGG	0,56	GGTGGTGGCCACGGTAGCGG	Insertion - In frame	RKO_63
RP1L1	No	p.K1578_L1579insK	c.4729_4730insAGA	0,52	CCCGGCCCTGGAGCTTCTAG	Insertion - In frame	RKO_64
S100A3	No	p.C99fs*>4	c.294_295insCC	0,44	GCAGAGGCTACTGGGAGCAC	Insertion - Frameshift	RKO_65
MYH9	Yes	p.K1674delK	c.5020_5022delAAG	0,26	CTTCAGCTTCTCGTTCTCTT	Deletion - In frame	RKO_66
MSH6	Yes	p.Y1066fs*1	c.3196_3197delTA	0,52	CTAACTAGTCGAGGGGGTGA	Deletion - Frameshift	RKO_67
JAKMIP1_ENST00000409021	No	p.R606fs*6	c.1816_1817delAG	0,38	GGCGACCTCCTCTCTCTCCT	Deletion - Frameshift	RKO_68
CDH18	No	p.L382delL	c.1144_1146delCTA	0,68	TAGGAAGGCATGGAAAATGG	Deletion - In frame	RKO_69
ITGA2	No	p.V751delV	c.2247_2249delTGT	0,63	ATCCAAAGAGTTGACATCAG	Deletion - In frame	RKO_70
PPL	No	p.W893_K894insW	c.2675_2676insGTG	0,72	GGCCACGTGGAAGATCAGGA	Insertion - In frame	RKO_71
ZNF697_ENST00000421812	No	p.E4delE	c.12_14delAGA	0,33	CTGGATGAAACAAGATAATC	Deletion - In frame	RKO_72
CLEC16A	No	p.E400delE	c.1192_1194delGAA	0,62	AAACGTTGGGGAAGAAGATG	Deletion - In frame	RKO_73
PKDCC	No	p.F205delF	c.610_612delTTC	0,54	ATGTGAAAAACCTGTGGCAG	Deletion - In frame	RKO_74
CREB3	No	p.E125delE	c.373_375delGAG	0,45	GACAGATGAGAAGAGTCTAT	Deletion - In frame	RKO_75
S100A3	No	p.C99fs*>4	c.294_295insCC	0,33	GGGAGCACCGGGGGGCTCTG	Insertion - Frameshift	RKO_76
IGSF8	No	p.H197fs*39	c.590_591delAC	0,45	AGCACACAGAAGCACACACC	Deletion - Frameshift	RKO_77
EML2	No	p.M33delM	c.99_101delGAT	0,65	CTGGGATCATGGGCACAGGG	Deletion - In frame	RKO_78
ZBTB48	No	p.H8fs*30	c.22_23delCA	0,47	CAGCAGTGTGAGGGTTCTGC	Deletion - Frameshift	RKO_79
JAKMIP1_ENST00000409021	No	p.R606fs*6	c.1816_1817delAG	0,66	TTTTCCCCAGGAGAGAGAGG	Deletion - Frameshift	RKO_80
POGZ_ENST00000491586	No	p.S545_Y546insS	c.1630_1631insCCT	0,25	TCATTCTTAGCCTTATTTTC	Insertion - In frame	RKO_81
PSEN1	No	p.E69delE	c.202_204delGAA	0,58	GGAGCAAGATGAGGAAGATG	Deletion - In frame	RKO_82
KMT2A	Yes	p.D877fs*8	c.2620_2621delAG	0,44	GGACAAGAGTAGAGAGAGAC	Deletion - Frameshift	RKO_83
MYH9	Yes	p.K1674delK	c.5020_5022delAAG	0,53	GAACGAGAAGCTGAAGAGCA	Deletion - In frame	RKO_84
ZBTB48	No	p.H8fs*30	c.22_23delCA	0,51	CTCCTTCGTCCAGCAGTGTG	Deletion - Frameshift	RKO_85
ZNF181	No	p.R473fs*3	c.1415_1416delAA	0,47	CTGCCCATCAGAGTACATAA	Deletion - Frameshift	RKO_86
ZBTB48	No	p.H8fs*30	c.22_23delCA	0,53	TCCTTCGTCCAGCAGTGTGA	Deletion - Frameshift	RKO_87
S100A3	No	p.C99fs*>4	c.294_295insCC	0,43	GAGCACCGGGGGGCTCTGAG	Insertion - Frameshift	RKO_88
ARL6IP4	No	p.H337fs*>24	c.1008_1009delAC	0,79	GAACGACAGAGAGATCAACA	Deletion - Frameshift	RKO_89
TTK	No	p.R854fs*>4	c.2552_2553delAA	0,12	CTCCAAGACTTTTGAAAAAA	Deletion - Frameshift	RKO_90
HSPA8	No	p.S276fs*7	c.825_826delTT	0,43	ACTGGCCTGGGTGCTGGGAG	Deletion - Frameshift	RKO_91
PPL	No	p.W893_K894insW	c.2675_2676insGTG	0,67	CTGGAGTGGAGGAGGCCACG	Insertion - In frame	RKO_92
RP1L1	No	p.K1578_L1579insK	c.4729_4730insAGA	0,66	GAGCTCCTCTAGAAGCTCCA	Insertion - In frame	RKO_93
MSH6	Yes	p.Y1066fs*1	c.3196_3197delTA	0,68	GCCTGGCTAACTAGTCGAGG	Deletion - Frameshift	RKO_94
PRDM2	No	p.K1489fs*6	c.4459_4460delAA	0,45	AACTTTTTTTGGGAGGAGAA	Deletion - Frameshift	RKO_95
PDE3B	No	p.F131delF	c.384_386delCTT	0,52	GCAGGTGAGGAAGAAGGCAC	Deletion - In frame	RKO_96
RP1L1	No	p.K1578_L1579insK	c.4729_4730insAGA	0,47	CCTCTAGAAGCTCCAGGGCC	Insertion - In frame	RKO_97
EML2	No	p.M33delM	c.99_101delGAT	0,28	TGCCAGCTCGTCTGGGATCA	Deletion - In frame	RKO_98
TTK	No	p.R854fs*>4	c.2552_2553delAA	0,56	CCAAGACTTTTGAAAAAAAG	Deletion - Frameshift	RKO_99
STAT3	Yes	p.E616delE	c.1847_1849delAAG	0,54	TTCAGTGAAAGCAGCAAAGG	Deletion - In frame	RKO_100
	Positive & negative Control sgRNAs						
PolR1C-1					GAAGAATCTCATCCTGAACAA		Pos_ctrl1
RPL4-1					GCTAATTCACTGACAGCATA		Pos_ctrl2
U2AF1-1					GTGATTGACTTGAATAACCGT		Pos_ctrl3
MYC					GCGCCGTCGTTGTCTCCCCGA		Pos_ctrl4
RPL10					GCTACATACCTTCTGGCGGCC		Pos_ctrl5
MRPL10					GATAGCCGTCTGCCAGAATG		Pos_ctrl6
LACZ-1					GCTAACGCCTGGGTCGAACGC		Non-targeting1
LACZ-2					GCGGGTGAACTGATCGCGCAG		Non-targeting2
LUC-1					GAATTGTCTTGTCCCTATCGA		Non-targeting3
LUC-2					GATCAAGGCGTTGGTCGCTTC		Non-targeting4

Tabel 2 cont'd RKO screen sgRNAs library.

	Oligo	Sequence	Oligo	Sequence
	***KRAS* Forward primer**	TGGACCCTGACATACTCCCA		
	***KRAS* Reverse primer**	AGCGTCGATGGAGGAGTTTG		
Base Editing sgRNAs	**sgG13D-TOP1**	CACCGCTGGTGACGTAGGCAAGAG	**sgG13D BOTTOM1**	AAACCTCTTGCCTACGTCACCAGC
	sgG13D-TOP2	CACCGGTGACGTAGGCAAGAG	**sgG13D BOTTOM2**	AAACCTCTTGCCTACGTCACC
	sgG13D-TOP3	CACCGAGCTGGTGACGTAGGCA	**sgG13D BOTTOM3**	AAACTGCCTACGTCACCAGCTC
	sgG13D-TOP4	CACCGTCACCAGCTCCAACTACCACA	**sgG13D BOTTOM4**	AAACTGTGGTAGTTGGAGCTGGTGAC
	sgG13D-TOP5	CACCGTGACGTAGGCAAGAGTGCCT	**sgG13D BOTTOM5**	AAACAGGCACTCTTGCCTACGTCAC
	sgG13D-TOP6	CACCGACGTAGGCAAGAGTGCCTTGA	**sgG13D BOTTOM6**	AAACTCAAGGCACTCTTGCCTACGTC
	sgG12D-TOP1	CACCGCTGATGGCGTAGGCAAGAG	**sgG12D BOTTOM1**	AAACCTCTTGCCTACGCCATCAGC
	sgG12D-TOP2	CACCGCCATCAGCTCCAACTACCACA	**sgG12D BOTTOM2**	AAACTGTGGTAGTTGGAGCTGATGGC
	sgG12D-TOP3	CACCGTGGAGCTGATGGCGTAGGCA	**sgG12D BOTTOM3**	AAACTGCCTACGCCATCAGCTCCAC
	sgG12D-TOP4	CACCGAGCTGATGGCGTAGGCA	**sgG12D BOTTOM4**	AAACTGCCTACGCCATCAGCTC
	sgG12D-TOP5	CACCGATGGCGTAGGCAAGAGTGCCT	**sgG12D BOTTOM5**	AAACAGGCACTCTTGCCTACGCCATC
	sgG12D-TOP6	CACCGAGCTGATGGCGTAGGCAAG	**sgG12D BOTTOM6**	AAACCTTGCCTACGCCATCAGCTC
	sgG12S-TOP1	CACCGCTAGTGGCGTAGGCAAGAG	**sgG12S BOTTOM1**	*AAACCTCTTGCCTACGCCACTAGC*
	sgG12S-TOP2	CACCGTGGAGCTAGTGGCGTAGGCA	**sgG12S BOTTOM2**	AAACTGCCTACGCCACTAGCTCCAC
	sgG12S-TOP3	CACCGAGCTAGTGGCGTAGGCA	**sgG12S BOTTOM3**	AAACTGCCTACGCCACTAGCTC
	sgG12S-TOP4	CACCGAGCTAGTGGCGTAGGCAAG	**sgG12S BOTTOM4**	AAACCTTGCCTACGCCACTAGCTC
Cas9 nuclease sgRNAs	**sgG13D-TOP**	CACCGTAGTTGGAGCTGGTGACGT	**sgG13D-BOTTOM**	AAACACGTCACCAGCTCCAACTAC
	sgG12D-TOP	CACCGTAGTTGGAGCTGATGGCGT	**sgG12D-BOTTOM**	AAACACGCCATCAGCTCCAACTAC
	sgG12S-TOP	CACCGCTTGTGGTAGTTGGAGCTAG	**sgG12S-BOTTOM**	AAACCTAGCTCCAACTACCACAAG
UTP14A Recombineering Oligos	**Fwd_Cassette - RecombineeringUTP14A**	ACCTTTCCACAGGATCAGGAGAAAAGCTGGTCCTTGCAGATCTGCTTGAGGAAGGCACGAACCCAGTTGAC		
	Rev_Cassette - RecombineeringUTP14A	TCCACTGTCTTCTTTGATTTGACTCTACTCAGTTGCTTTTTCACAGTGGCTTTGTTCAAAAAAAAGCCCGC		
	UTP14A- Repair OligoFwd	AGGATCAGGAGAAAAGCTGGTCCTTGCAGATCTGCTTGAGCCCGTCAAGACCAGTAGCCTTGCCACTGTGAAAAAGCAACTGAGTAGAGTCAAATCAAAGA		
	UTP14A- Repair OligoRev	TCTTTGATTTGACTCTACTCAGTTGCTTTTTCACAGTGGCAAGGCTACTGGTCTTGACGGGCTCAAGCAGATCTGCAAGGACCAGCTTTTCTCCTGATCCT		
ABE8e cloning in lentiviral vector	**NG-ABE8e T7-1**	gggcccggtaccTAATACGACTCACTATAGGG		
	NG-ABE8e CMV-2	gggcccggtaccCGCAAATGGGCGGTAGGCGTG		
	NG-ABE8e rev	gggggggtatacGACTTTCCTCTTCTTCTTGGG		

Tabel 3 Oligos used in this study.

	Oligo	Sequence
UTP14A oligos (protospacers, T7E1 and screen primers)	UTP14A-Fwd1	AGGGTCATGACATGCTGTGG
	UTP14A-Rev1	ATCTGAAGCAGGCAAGGCAT
	UTP14A-Fwd2	AATGGGGAGTCACTGGAGGA
	UTP14A-Rev2	GCTGGCTCCTCTTTCTCCAG
	UTP14A-exon1-TOP	CACCGTGCGAACCGGCTTGCAGAG
	UTP14A-exon1-Bottom	AAACCTCTGCAAGCCGGTTCGCAC
	UTP14A-exon1-FP	CTTCCGTTCTTGGTCCATGT
	UTP14A-exon1-RP	CTATACAGCTGAGGAGCGGAGT
	UTP14A-exon2-TOP	CACCGATCCGCTAGTTCTTCCTGT
	UTP14A-ex2-Bottom	AAACACAGGAAGAACTAGCGGATC
	UTP14A-exon2-FP	TCTGAGGGATAGGGCTATTCTG
	UTP14A-exon2-RP	AGTTTAGCAATTGTCACCCACC
	DeepS-PCR1-S1-L	GTAATAATTTCTTGGGTAGTTTGCA
	DeepS-PCR1-S1-R	ATTGTGGATGAATACTGCCATTTG
	DeepS-PCR2-L	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGGCTTTATATATCTTGTGGAAAGG
	DeepS-PCR2-R	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTCAAGTTGATAACGGACTAGCC
Cloning of xCas9.ABE7.10 in lentiviral vectors	MCS Fwd	gatccGGTACCcagctgcgaggctgtatacatgca
	MCS Rev	tgtatacagcctcgcagctgGGTACCg
	ABE Fwd	gggcccggtaccGCCACCATGTCCGAAGTCG
	ABERev	ggggggtatacGACTTTCCTCTTCTTCTTGGGGC
mRNA primers for Adenine base editors (xCas9.ABE7.10, NG-ABEmax and NG-ABE8e)	xCas9-ABE 5'	gctaatacgactcactatagggagagATGTCCGAAGTCGAGTTTTCCC
	xCas9-ABE 3'	ttttttttttttttttggtttattcTTAGACTTTCCTCTTCTTCTTGGGG
	xCas9-ABE-GFP 3'	ttttttttttttttttggtttattcTTACTTGTACAGCTCGTCCATGC
	ABE s	atgtccgaagtcgagttttccc
	ABE as	tcagcaggctgaagttggtggcgactttcctcttcttcttggggc
	P2A-GFP s	gccccaagaagaagaggaaagtcgccaccaacttcagcctgctga
	P2A-GFP as	TTACTTGTACAGCTCGTCCATGC

Tabel 4 cont'd oligos used in this study.

6 References

Adli, M. (2018). The CRISPR tool kit for genome editing and beyond. Nature Communications, 9(1), 1911. https://doi.org/10.1038/s41467-018-04252-2

Akcakaya, P., Bobbin, M. L., Guo, J. A., Malagon-Lopez, J., Clement, K., Garcia, S. P., Fellows, M. D., Porritt, M. J., Firth, M. A., Carreras, A., Baccega, T., Seeliger, F., Bjursell, M., Tsai, S. Q., Nguyen, N. T., Nitsch, R., Mayr, L. M., Pinello, L., Bohlooly-Y, M., … Joung, J. K. (2018). In vivo CRISPR editing with no detectable genome-wide off-target mutations. Nature, 561(7723), 416–419. https://doi.org/10.1038/s41586-018-0500-9

Allen, F., Crepaldi, L., Alsinet, C., Strong, A. J., Kleshchevnikov, V., Angeli, P. D., Páleníková, P., Khodak, A., Kiselev, V., Kosicki, M., Bassett, A. R., Harding, H., Galanty, Y., Muñoz-Martínez, F., Metzakopian, E., Jackson, S. P. & Parts, L. (2019). Predicting the mutations generated by repair of Cas9-induced double-strand breaks. Nature Biotechnology, 37(1), 64–72. https://doi.org/10.1038/nbt.4317

Anandakrishnan, R., Varghese, R. T., Kinney, N. A. & Garner, H. R. (2019). Estimating the number of genetic mutations (hits) required for carcinogenesis based on the distribution of somatic mutations. PLOS Computational Biology, 15(3), e1006881. https://doi.org/10.1371/journal.pcbi.1006881

Anderson, E. M., Haupt, A., Schiel, J. A., Chou, E., Machado, H. B., Strezoska, Ž., Lenger, S., McClelland, S., Birmingham, A., Vermeulen, A. & Smith, A. van B. (2015). Systematic analysis of CRISPR–Cas9 mismatch tolerance reveals low levels of off-target activity. Journal of Biotechnology, 211, 56–65. https://doi.org/10.1016/j.jbiotec.2015.06.427

Anzalone, A. V., Koblan, L. W. & Liu, D. R. (2020). Genome editing with CRISPR–Cas nucleases, base editors, transposases and prime editors. Nature Biotechnology, 38(7), 824–844. https://doi.org/10.1038/s41587-020-0561-9

Anzalone, A. V., Randolph, P. B., Davis, J. R., Sousa, A. A., Koblan, L. W., Levy, J. M., Chen, P. J., Wilson, C., Newby, G. A., Raguram, A. & Liu, D. R. (2019). Search-and-replace genome editing without double-strand breaks or donor DNA. Nature, 576(7785), 149–157. https://doi.org/10.1038/s41586-019-1711-4

Arbab, M., Shen, M. W., Mok, B., Wilson, C., Matuszek, Ż., Cassa, C. A. & Liu, D. R. (2020). Determinants of Base Editing Outcomes from Target Library Analysis and Machine Learning. Cell, 182(2), 463-480.e30. https://doi.org/10.1016/j.cell.2020.05.037

Armitage, P. & Doll, R. (1954). The Age Distribution of Cancer and a Multi-stage Theory of Carcinogenesis. British Journal of Cancer, 8(1), 1–12. https://doi.org/10.1038/bjc.1954.1

Arrington, A. K., Heinrich, E. L., Lee, W., Duldulao, M., Patel, S., Sanchez, J., Garcia-Aguilar, J. & Kim, J. (2012). Prognostic and Predictive Roles of KRAS Mutation in Colorectal Cancer. International Journal of Molecular Sciences, 13(10), 12153–12168. https://doi.org/10.3390/ijms131012153

Auton, A., Abecasis, G. R., Altshuler, D. M., Durbin, R. M., Abecasis, G. R., Bentley, D. R., Chakravarti, A., Clark, A. G., Donnelly, P., Eichler, E. E., Flicek, P., Gabriel, S. B., Gibbs, R. A., Green, E. D., Hurles, M. E., Knoppers, B. M., Korbel, J. O., Lander, E. S., Lee, C., … Abecasis, G. R. (2015). A global reference for human genetic variation. Nature, 526(7571), 68–74. https://doi.org/10.1038/nature15393

Avery, O. T., Macleod, C. M. & McCarty, M. (1944). Studies on the chemical nature of the substance inducing transformation of pneumococcal types: induction of transformation by a desoxyribonucleic acid fraction isolated from pneumococcus type III. The Journal of Experimental Medicine, 79(2), 137–158.

Babaei, S., Hulsman, M., Reinders, M. & Ridder, J. de. (2013). Detecting recurrent gene mutation in interaction network context using multi-scale graph diffusion. BMC Bioinformatics, 14(1), 29. https://doi.org/10.1186/1471-2105-14-29

Bartfeld, S., Bayram, T., Wetering, M. van de, Huch, M., Begthel, H., Kujala, P., Vries, R., Peters, P. J. & Clevers, H. (2015). In Vitro Expansion of Human Gastric Epithelial Stem Cells and Their Responses to Bacterial Infection. Gastroenterology, 148(1), 126-136.e6. https://doi.org/10.1053/j.gastro.2014.09.042

Berndt, N., Hamilton, A. D. & Sebti, S. M. (2011). Targeting protein prenylation for cancer therapy. Nature Reviews Cancer, 11(11), 775–791. https://doi.org/10.1038/nrc3151

Billon, P., Bryant, E. E., Joseph, S. A., Nambiar, T. S., Hayward, S. B., Rothstein, R. & Ciccia, A. (2017). CRISPR-Mediated Base Editing Enables Efficient Disruption of Eukaryotic Genes through Induction of STOP Codons. Molecular Cell, 67(6), 1068-1079.e4. https://doi.org/10.1016/j.molcel.2017.08.008

Boj, S. F., Hwang, C.-I., Baker, L. A., Chio, I. I. C., Engle, D. D., Corbo, V., Jager, M., Ponz-Sarvise, M., Tiriac, H., Spector, M. S., Gracanin, A., Oni, T., Yu, K. H., van Boxtel, R., Huch, M., Rivera, K. D., Wilson, J. P., Feigin, M. E., Öhlund, D., … Tuveson, D. A. (2015). Organoid Models of Human and Mouse Ductal Pancreatic Cancer. Cell, 160(1–2), 324–338. https://doi.org/10.1016/j.cell.2014.12.021

Bolukbasi, M. F., Gupta, A. & Wolfe, S. A. (2016). Creating and evaluating accurate CRISPR-Cas9 scalpels for genomic surgery. Nature Methods, 13(1), 41–50. https://doi.org/10.1038/nmeth.3684

Boveri, T. (1914). Zur frage der entstehung maligner tumoren.

Brinkman, E. K., Chen, T., Haas, M. de, Holland, H. A., Akhtar, W. & Steensel, B. van. (2018). Kinetics and Fidelity of the Repair of Cas9-Induced Double-Strand DNA Breaks. Molecular Cell, 70(5), 801-813.e6. https://doi.org/10.1016/j.molcel.2018.04.016

Brown, A.-L., Li, M., Goncearenco, A. & Panchenko, A. R. (2019). Finding driver mutations in cancer: Elucidating the role of background mutational processes. PLOS Computational Biology, 15(4), e1006981. https://doi.org/10.1371/journal.pcbi.1006981

Buljan, M., Blattmann, P., Aebersold, R. & Boutros, M. (2018). Systematic characterization of pan-cancer mutation clusters. Molecular Systems Biology, 14(3), e7974. https://doi.org/10.15252/msb.20177974

Bunn, A. G. (2010). Statistical and visual crossdating in R using the dplR library. Dendrochronologia, 28(4), 251–258.

Bycroft, C., Freeman, C., Petkova, D., Band, G., Elliott, L. T., Sharp, K., Motyer, A., Vukcevic, D., Delaneau, O., O'Connell, J., Cortes, A., Welsh, S., Young, A., Effingham, M., McVean, G., Leslie, S., Allen, N., Donnelly, P. & Marchini, J. (2018). The UK Biobank resource with deep phenotyping and genomic data. Nature, 562(7726), 203–209. https://doi.org/10.1038/s41586-018-0579-z

Carpi, S., Polini, B., Poli, G., Barata, G. A., Fogli, S., Romanini, A., Tuccinardi, T., Guella, G., Frontini, F. P., Nieri, P. & Giuseppe, G. D. (2018). Anticancer Activity of Euplotin C, Isolated from the Marine Ciliate Euplotes crassus, Against Human Melanoma Cells. Marine Drugs, 16(5), 166. https://doi.org/10.3390/md16050166

Chadwick, A. C., Wang, X. & Musunuru, K. (2018). In Vivo Base Editing of PCSK9 (Proprotein Convertase Subtilisin/Kexin Type 9) as a Therapeutic Alternative to Genome Editing. Arteriosclerosis, Thrombosis, and Vascular Biology, 37(9), 1741–1747. https://doi.org/10.1161/atvbaha.117.309881

Chapman, J. R., Taylor, M. R. G. & Boulton, S. J. (2012). Playing the End Game: DNA Double-Strand Break Repair Pathway Choice. Molecular Cell, 47(4), 497–510. https://doi.org/10.1016/j.molcel.2012.07.029

Chatterjee, P., Jakimo, N. & Jacobson, J. M. (2018). Minimal PAM specificity of a highly similar SpCas9 ortholog. Science Advances, 4(10), eaau0766. https://doi.org/10.1126/sciadv.aau0766

Chen, B., Gilbert, L. A., Cimini, B. A., Schnitzbauer, J., Zhang, W., Li, G.-W., Park, J., Blackburn, E. H., Weissman, J. S., Qi, L. S. & Huang, B. (2013). Dynamic Imaging of Genomic Loci in Living Human Cells by an Optimized CRISPR/Cas System. Cell, 155(7), 1479–1491. https://doi.org/10.1016/j.cell.2013.12.001

Chen, S., Sanjana, N. E., Zheng, K., Shalem, O., Lee, K., Shi, X., Scott, D. A., Song, J., Pan, J. Q., Weissleder, R., Lee, H., Zhang, F. & Sharp, P. A. (2015). Genome-wide CRISPR Screen in a Mouse Model of Tumor Growth and Metastasis. Cell, 160(6), 1246–1260. https://doi.org/10.1016/j.cell.2015.02.038

Chen, W., McKenna, A., Schreiber, J., Haeussler, M., Yin, Y., Agarwal, V., Noble, W. S. & Shendure, J. (2019). Massively parallel profiling and predictive modeling of the outcomes of CRISPR/Cas9-mediated double-strand break repair. Nucleic Acids Research, 47(15), 7989–8003. https://doi.org/10.1093/nar/gkz487

Cheng, D. & Tuveson, D. (2018). Kras in Organoids. Cold Spring Harbor Perspectives in Medicine, 8(10), a031575. https://doi.org/10.1101/cshperspect.a031575

Cheung, A. H.-K., Chow, C., Zhang, J., Zhou, Y., Huang, T., Ng, K. C.-K., Or, T. C.-T., Yao, Y. Y., Dong, Y., Fung, J. M.-W., Xiong, L., Chan, A. K.-Y., Lung, W.-M. R., Kang, W. & To, K.-F. (2018). Specific targeting of point mutations in EGFR L858R-positive lung cancer by CRISPR/Cas9. Laboratory Investigation, 98(7), 968–976. https://doi.org/10.1038/s41374-018-0056-1

Chin, L., Tam, A., Pomerantz, J., Wong, M., Holash, J., Bardeesy, N., Shen, Q., O'Hagan, R., Pantginis, J., Zhou, H., Horner, J. W., Cordon-Cardo, C., Yancopoulos, G. D. & DePinho#, R. A. (1999). Essential role for oncogenic Ras in tumour maintenance. Nature, 400(6743), 468–472. https://doi.org/10.1038/22788

Cho, S. W., Kim, S., Kim, J. M. & Kim, J.-S. (2019). Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease. Nature Biotechnology, 1–3.

Christensen, J. G., Olson, P., Briere, T., Wiel, C. & Bergo, M. O. (2020). Targeting Krasg12c-mutant cancer with a mutation-specific inhibitor. Journal of Internal Medicine, 288(2), 183–191. https://doi.org/10.1111/joim.13057

Ciccia, A. & Elledge, S. J. (2010). The DNA Damage Response: Making It Safe to Play with Knives. Molecular Cell, 40(2), 179–204. https://doi.org/10.1016/j.molcel.2010.09.019

Collins, M. A., Bednar, F., Zhang, Y., Brisset, J.-C., Galbán, S., Galbán, C. J., Rakshit, S., Flannagan, K. S., Adsay, N. V. & Magliano, M. P. di. (2012). Oncogenic Kras is required for both the initiation and maintenance of pancreatic cancer in mice. Journal of Clinical Investigation, 122(2), 639–653. https://doi.org/10.1172/jci59227

Cong, L., Ran, F. A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P. D., Wu, X., Jiang, W., Marraffini, L. A. & Zhang, F. (2013). Multiplex Genome Engineering Using CRISPR/Cas Systems. Science, 339(6121), 819–823. https://doi.org/10.1126/science.1231143

Cooper, G. M. & Shendure, J. (2011). Needles in stacks of needles: finding disease-causal variants in a wealth of genomic data. Nature Reviews Genetics, 12(9), 628–640. https://doi.org/10.1038/nrg3046

Cox, A. D. & Der, C. J. (2014). Ras history. Small GTPases, 1(1), 2–27.

Cox, D. B. T., Platt, R. J. & Zhang, F. (2015). Therapeutic genome editing: prospects and challenges. Nature Medicine, 21(2), 121–131.

D'Antonio, M. & Ciccarelli, F. D. (2013). Integrated analysis of recurrent properties of cancer genes to identify novel drivers. Genome Biology, 14(5), R52.

Darbyshire, M., Toit, Z. du, Rogers, M. F., Gaunt, T. R. & Campbell, C. (2019). Estimating the Frequency of Single Point Driver Mutations across Common Solid Tumours. Scientific Reports, 9(1), 13452. https://doi.org/10.1038/s41598-019-48765-2

Daver, N., Schlenk, R. F., Russell, N. H. & Levis, M. J. (2019). Targeting FLT3 mutations in AML: review of current knowledge and evidence. Leukemia, 33(2), 299–312. https://doi.org/10.1038/s41375-018-0357-9

Dees, N. D., Zhang, Q., Kandoth, C., Wendl, M. C., Schierding, W., Koboldt, D. C., Mooney, T. B., Callaway, M. B., Dooling, D., Mardis, E. R., Wilson, R. K. & Ding, L. (2012). MuSiC: Identifying mutational significance in cancer genomes. Genome Research, 22(8), 1589–1598. https://doi.org/10.1101/gr.134635.111

DeWitt, M. A., Magis, W., Bray, N. L., Wang, T., Berman, J. R., Urbinati, F., Heo, S.-J., Mitros, T., Muñoz, D. P., Boffelli, D., Kohn, D. B., Walters, M. C., Carroll, D., Martin, D. I. K. & Corn, J. E. (2016). Selection-free genome editing of the sickle mutation in human adult hematopoietic stem/progenitor cells. Science Translational Medicine, 8(360), 360ra134-360ra134. https://doi.org/10.1126/scitranslmed.aaf9336

Diederichs, S., Bartsch, L., Berkmann, J. C., Fröse, K., Heitmann, J., Hoppe, C., Iggena, D., Jazmati, D., Karschnia, P., Linsenmeier, M., Maulhardt, T., Möhrmann, L., Morstein, J., Paffenholz, S. V., Röpenack, P., Rückert, T., Sandig, L., Schell, M., Steinmann, A., … Wullenkord, R. (2016). The dark matter of the cancer genome: aberrations in regulatory elements, untranslated regions, splice sites, non-coding RNA and synonymous mutations. EMBO Molecular Medicine, 8(5), 442–457. https://doi.org/10.15252/emmm.201506055

Dietlein, F., Weghorn, D., Taylor-Weiner, A., Richters, A., Reardon, B., Liu, D., Lander, E. S., Allen, E. M. V. & Sunyaev, S. R. (2020). Identification of cancer driver genes based on nucleotide context. Nature Genetics, 52(2), 208–218. https://doi.org/10.1038/s41588-019-0572-y

Doench, J. G., Fusi, N., Sullender, M., Hegde, M., Vaimberg, E. W., Donovan, K. F., Smith, I., Tothova, Z., Wilen, C., Orchard, R., Virgin, H. W., Listgarten, J. & Root, D. E. (2016). Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. Nature Biotechnology, 34(2), 184–191. https://doi.org/10.1038/nbt.3437

Doench, J. G., Hartenian, E., Graham, D. B., Tothova, Z., Hegde, M., Smith, I., Sullender, M., Ebert, B. L., Xavier, R. J. & Root, D. E. (2014). Rational design of highly active sgRNAs for CRISPR-Cas9–mediated gene inactivation. Nature Biotechnology, 32(12), 1262–1267. https://doi.org/10.1038/nbt.3026

Domingo, E., Ramamoorthy, R., Oukrif, D., Rosmarin, D., Presz, M., Wang, H., Pulker, H., Lockstone, H., Hveem, T., Cranston, T., Danielsen, H., Novelli, M., Davidson, B., Xu, Z., Molloy, P., Johnstone, E., Holmes, C., Midgley, R., Kerr, D., … Tomlinson, I. (2013). Use of multivariate analysis to suggest a new molecular classification of colorectal cancer. The Journal of Pathology, 229(3), 441–448. https://doi.org/10.1002/path.4139

Dominguez, A. A., Lim, W. A. & Qi, L. S. (2016). Beyond editing: repurposing CRISPR–Cas9 for precision genome regulation and interrogation. Nature Reviews Molecular Cell Biology, 17(1), 5–15. https://doi.org/10.1038/nrm.2015.2

Driehuis, E. & Clevers, H. (2017). CRISPR/Cas 9 genome editing and its applications in organoids. American Journal of Physiology-Gastrointestinal and Liver Physiology, 312(3), G257–G265. https://doi.org/10.1152/ajpgi.00410.2016

Drost, J. & Clevers, H. (2018). Organoids in cancer research. Nature Publishing Group, 1–12.

Du, D., Roguev, A., Gordon, D. E., Chen, M., Chen, S.-H., Shales, M., Shen, J. P., Ideker, T., Mali, P., Qi, L. S. & Krogan, N. J. (2017). Genetic interaction mapping in mammalian cells using CRISPR interference. Nature Methods, 14(6), 577–580. https://doi.org/10.1038/nmeth.4286

Dwane, L., Behan, F. M., Gonçalves, E., Lightfoot, H., Yang, W., van der Meer, D., Shepherd, R., Pignatelli, M., Iorio, F. & Garnett, M. J. (2020). Project Score database: a resource for investigating cancer cell dependencies and prioritizing therapeutic targets. Nucleic Acids Research, 49(D1), gkaa882-. https://doi.org/10.1093/nar/gkaa882

Edraki, A., Mir, A., Ibraheim, R., Gainetdinov, I., Yoon, Y., Song, C.-Q., Cao, Y., Gallant, J., Xue, W., Rivera-Pérez, J. A. & Sontheimer, E. J. (2019). A Compact, High-Accuracy Cas9 with a Dinucleotide PAM for In Vivo Genome Editing. Molecular Cell, 73(4), 714-726.e4. https://doi.org/10.1016/j.molcel.2018.12.003

Esvelt, K. M., Mali, P., Braff, J. L., Moosburner, M., Yaung, S. J. & Church, G. M. (2013). Orthogonal Cas9 proteins for RNA-guided gene regulation and editing. Nature Methods, 10(11), 1116–1121. https://doi.org/10.1038/nmeth.2681

Ferrer, I., Zugazagoitia, J., Herbertz, S., John, W., Paz-Ares, L. & Schmid-Bindert, G. (2018). KRAS-Mutant non-small cell lung cancer: From biology to therapy. Lung Cancer, 124, 53–64. https://doi.org/10.1016/j.lungcan.2018.07.013

Fiala, O., Buchler, T., Mohelnikova-Duchonova, B., Melichar, B., Matejka, V. M., Holubec, L., Kulhankova, J., Bortlicek, Z., Bartouskova, M., Liska, V., Topolcan, O., Sedivcova, M. & Finek, J. (2016). G12V and G12A KRAS mutations are associated with poor outcome in patients with metastatic colorectal cancer treated with bevacizumab. Tumor Biology, 37(5), 6823–6830. https://doi.org/10.1007/s13277-015-4523-7

Fisher, G. H., Wellen, S. L., Klimstra, D., Lenczowski, J. M., Tichelaar, J. W., Lizak, M. J., Whitsett, J. A., Koretsky, A. & Varmus, H. E. (2001). Induction and apoptotic regression of lung adenocarcinomas by regulation of a K-Ras transgene in the presence and absence of tumor suppressor genes. Genes & Development, 15(24), 3249–3262. https://doi.org/10.1101/gad.947701

Florian, G., Albion, B., Kiran, M., Gerard, M. & Zemin, Z. (2013). Assessment of computational methods for predicting the effects of missense mutations in human cancers. BMC Genomics, 14(3), S7. https://doi.org/10.1186/1471-2164-14-s3-s7

Forbes, S. A., Tang, G., Bindal, N., Bamford, S., Dawson, E., Cole, C., Kok, C. Y., Jia, M., Ewing, R., Menzies, A., Teague, J. W., Stratton, M. R. & Futreal, P. A. (2010). COSMIC (the Catalogue of Somatic Mutations in Cancer): a resource to investigate acquired mutations in human cancer. Nucleic Acids Research, 38(suppl_1), D652–D657. https://doi.org/10.1093/nar/gkp995

Fu, J., Teucher, M., Anastassiadis, K., Skarnes, W. & Stewart, A. F. (2010). Chapter Eight A Recombineering Pipeline to Make Conditional Targeting Constructs. Methods in Enzymology, 477, 125–144. https://doi.org/10.1016/s0076-6879(10)77008-7

Garay, J. P. & Gray, J. W. (2012). Omics and therapy – A basis for precision medicine. Molecular Oncology, 6(2), 128–139. https://doi.org/10.1016/j.molonc.2012.02.009

Garrido, J. J., Giraud, P., Carlier, E., Fernandes, F., Moussif, A., Fache, M.-P., Debanne, D. & Dargent, B. (2003). A Targeting Motif Involved in Sodium Channel Clustering at the Axonal Initial Segment. Science, 300(5628), 2091–2094. https://doi.org/10.1126/science.1085167

Gaudelli, N. M., Komor, A. C., Rees, H. A., Packer, M. S., Badran, A. H., Bryson, D. I. & Liu, D. R. (2017). Programmable base editing of A•T to G•C in genomic DNA without DNA cleavage. Nature, 551(7681), 464–471. https://doi.org/10.1038/nature24644

Gaudelli, N. M., Lam, D. K., Rees, H. A., Solá-Esteves, N. M., Barrera, L. A., Born, D. A., Edwards, A., Gehrke, J. M., Lee, S.-J., Liquori, A. J., Murray, R., Packer, M. S., Rinaldi, C., Slaymaker, I. M., Yen, J., Young, L. E. & Ciaramella, G. (2020). Directed evolution of adenine base editors with increased activity and therapeutic application. Nature Biotechnology, 38(7), 892–900. https://doi.org/10.1038/s41587-020-0491-6

Gebler, C., Lohoff, T., Paszkowski-Rogacz, M., Mircetic, J., Chakraborty, D., Camgoz, A., Hamann, M. V., Theis, M., Thiede, C. & Buchholz, F. (2016). Inactivation of Cancer Mutations Utilizing CRISPR/Cas9. JNCI: Journal of the National Cancer Institute, 109(1). https://doi.org/10.1093/jnci/djw183

Gerstung, M., Papaemmanuil, E., Martincorena, I., Bullinger, L., Gaidzik, V. I., Paschka, P., Heuser, M., Thol, F., Bolli, N., Ganly, P., Ganser, A., McDermott, U., Döhner, K., Schlenk, R. F., Döhner, H. & Campbell, P. J. (2017). Precision oncology for acute myeloid leukemia using a knowledge bank approach. Nature Genetics, 49(3), 332–340. https://doi.org/10.1038/ng.3756

Getz, G., Hofling, H., Mesirov, J. P., Golub, T. R., Meyerson, M., Tibshirani, R. & Lander, E. S. (2007). Comment on “The Consensus Coding Sequences of Human Breast and Colorectal Cancers.” Science, 317(5844), 1500–1500. https://doi.org/10.1126/science.1138764

Gonzalez-Perez, A. & Lopez-Bigas, N. (2012). Functional impact bias reveals cancer drivers. Nucleic Acids Research, 40(21), e169–e169. https://doi.org/10.1093/nar/gks743

Gonzalez-Perez, A., Mustonen, V., Reva, B., Ritchie, G. R. S., Creixell, P., Karchin, R., Vazquez, M., Fink, J. L., Kassahn, K. S., Pearson, J. V., Bader, G. D., Boutros, P. C., Muthuswamy, L., Ouellette, B. F. F., Reimand, J., Linding, R., Shibata, T., Valencia, A., Butler, A., … Group, I. C. G. C. M. P. and C. S. of the B. A. W. (2013). Computational approaches to identify functional genetic variants in cancer genomes. Nature Methods, 10(8), 723–729. https://doi.org/10.1038/nmeth.2562

Greenman, C., Wooster, R., Futreal, P. A., Stratton, M. R. & Easton, D. F. (2006). Statistical Analysis of Pathogenicity of Somatic Mutations in Cancer. Genetics, 173(4), 2187–2198. https://doi.org/10.1534/genetics.105.044677

Haigis, K. M. (2017). KRAS Alleles: The Devil Is in the Detail. TRENDS in CANCER, 1–12.

Haigis, K. M., Kendall, K. R., Wang, Y., Cheung, A., Haigis, M. C., Glickman, J. N., Niwa-Kawakita, M., Sweet-Cordero, A., Sebolt-Leopold, J., Shannon, K. M., Settleman, J., Giovannini, M. & Jacks, T. (2008). Differential effects of oncogenic K-Ras and N-Ras on proliferation, differentiation and tumor progression in the colon. Nature Genetics, 40(5), 600–608. https://doi.org/10.1038/ng.115

Hansemann, D. (1890). Ueber asymmetrische Zelltheilung in Epithelkrebsen und deren biologische Bedeutung. Archiv Für Pathologische Anatomie Und Physiologie Und F̈ur Klinische Medicin, 119(2), 299–326.

Harrington, L. B., Paez-Espino, D., Staahl, B. T., Chen, J. S., Ma, E., Kyrpides, N. C. & Doudna, J. A. (2017). A thermostable Cas9 with increased lifetime in human plasma. Nature Communications, 8(1), 1424. https://doi.org/10.1038/s41467-017-01408-4

Hart, T., Chandrashekhar, M., Aregger, M., Steinhart, Z., Brown, K. R., MacLeod, G., Mis, M., Zimmermann, M., Fradet-Turcotte, A., Sun, S., Mero, P., Dirks, P., Sidhu, S., Roth, F. P., Rissland, O. S., Durocher, D., Angers, S. & Moffat, J. (2015). High-Resolution CRISPR Screens Reveal Fitness Genes and Genotype-Specific Cancer Liabilities. Cell, 163(6), 1515–1526. https://doi.org/10.1016/j.cell.2015.11.015

Heckl, D., Kowalczyk, M. S., Yudovich, D., Belizaire, R., Puram, R. V., McConkey, M. E., Thielke, A., Aster, J. C., Regev, A. & Ebert, B. L. (2014). Generation of mouse models of myeloid malignancy with combinatorial genetic lesions using CRISPR-Cas9 genome editing. Nature Biotechnology, 32(9), 941–946. https://doi.org/10.1038/nbt.2951

Hennig, A., Wolf, L., Jahnke, B., Polster, H., Seidlitz, T., Werner, K., Aust, D. E., Hampe, J., Distler, M., Weitz, J., Stange, D. E. & Welsch, T. (2019). CFTR Expression Analysis for Subtyping of Human Pancreatic Cancer Organoids. Stem Cells International, 2019, 1–8. https://doi.org/10.1155/2019/1024614

Higasa, K., Miyake, N., Yoshimura, J., Okamura, K., Niihori, T., Saitsu, H., Doi, K., Shimizu, M., Nakabayashi, K., Aoki, Y., Tsurusaki, Y., Morishita, S., Kawaguchi, T., Migita, O., Nakayama, K., Nakashima, M., Mitsui, J., Narahara, M., Hayashi, K., … Matsuda, F. (2016). Human genetic variation database, a reference database of genetic variations in the Japanese population. Journal of Human Genetics, 61(6), 547–553. https://doi.org/10.1038/jhg.2016.12

Hirano, H., Gootenberg, J. S., Horii, T., Abudayyeh, O. O., Kimura, M., Hsu, P. D., Nakane, T., Ishitani, R., Hatada, I., Zhang, F., Nishimasu, H. & Nureki, O. (2016). Structure and Engineering of Francisella novicida Cas9. CELL, 164(5), 950–961.

Hobbs, G. A., Der, C. J. & Rossman, K. L. (2016). RAS isoforms and mutations in cancer at a glance. Journal of Cell Science, 129(7), 1287–1292. https://doi.org/10.1242/jcs.182873

Hou, Z., Zhang, Y., Propson, N. E., Howden, S. E., Chu, L.-F., Sontheimer, E. J. & Thomson, J. A. (2013). Efficient genome engineering in human pluripotent stem cells using Cas9 from Neisseria meningitidis. Proceedings of the National Academy of Sciences, 110(39), 15644–15649. https://doi.org/10.1073/pnas.1313587110

Hsu, P. D., Lander, E. S. & Zhang, F. (2014). Development and Applications of CRISPR-Cas9 for Genome Engineering. Cell, 157(6), 1262–1278. https://doi.org/10.1016/j.cell.2014.05.010

Hsu, P. D., Scott, D. A., Weinstein, J. A., Ran, F. A., Konermann, S., Agarwala, V., Li, Y., Fine, E. J., Wu, X., Shalem, O., Cradick, T. J., Marraffini, L. A., Bao, G. & Zhang, F. (2013). DNA targeting specificity of RNA-guided Cas9 nucleases. Nature Biotechnology, 31(9), 827–832. https://doi.org/10.1038/nbt.2647

Hu, J. H., Miller, S. M., Geurts, M. H., Tang, W., Chen, L., Sun, N., Zeina, C. M., Gao, X., Rees, H. A., Lin, Z. & Liu, D. R. (2018). Evolved Cas9 variants with broad PAM compatibility and high DNA specificity. Nature, 556(7699), 57–63. https://doi.org/10.1038/nature26155

Hu, L., Wang, J., Liu, Y., Zhang, Y., Zhang, L., Kong, R., Zheng, Z., Du, X. & Ke, Y. (2011). A Small Ribosomal Subunit (SSU) Processome Component, the Human U3 Protein 14A (hUTP14A) Binds p53 and Promotes p53 Degradation*. Journal of Biological Chemistry, 286(4), 3119–3128. https://doi.org/10.1074/jbc.m110.157842

Huang, L., Holtzinger, A., Jagan, I., BeGora, M., Lohse, I., Ngai, N., Nostro, C., Wang, R., Muthuswamy, L. B., Crawford, H. C., Arrowsmith, C., Kalloger, S. E., Renouf, D. J., Connor, A. A., Cleary, S., Schaeffer, D. F., Roehrl, M., Tsao, M.-S., Gallinger, S., … Muthuswamy, S. K. (2015). Ductal pancreatic cancer modeling and drug screening using human pluripotent stem cell– and patient-derived tumor organoids. Nature Medicine, 21(11), 1364–1371. https://doi.org/10.1038/nm.3973

Huang, T. P., Zhao, K. T., Miller, S. M., Gaudelli, N. M., Oakes, B. L., Fellmann, C., Savage, D. F. & Liu, D. R. (2019). Circularly permuted and PAM-modified Cas9 variants broaden the targeting scope of base editors. Nature Biotechnology, 37(6), 626–631. https://doi.org/10.1038/s41587-019-0134-y

Hubner, N. C., Bird, A. W., Cox, J., Splettstoesser, B., Bandilla, P., Poser, I., Hyman, A. & Mann, M. (2010). Quantitative proteomics combined with BAC TransgeneOmics reveals in vivo protein interactions. Journal of Cell Biology, 189(4), 739–754. https://doi.org/10.1083/jcb.200911091

Imamura, Y., Morikawa, T., Liao, X., Lochhead, P., Kuchiba, A., Yamauchi, M., Qian, Z. R., Nishihara, R., Meyerhardt, J. A., Haigis, K. M., Fuchs, C. S. & Ogino, S. (2012). Specific Mutations in KRAS Codons 12 and 13, and Patient Prognosis in 1075 BRAF Wild-Type Colorectal Cancers. Clinical Cancer Research, 18(17), 4753–4763. https://doi.org/10.1158/1078-0432.ccr-11-3210

Iqbal, N. & Iqbal, N. (2014). Imatinib: A Breakthrough of Targeted Therapy in Cancer. Chemotherapy Research and Practice, 2014, 1–9. https://doi.org/10.1155/2014/357027

Iranzo, J., Martincorena, I. & Koonin, E. V. (2018). Cancer-mutation network and the number and specificity of driver mutations. Proceedings of the National Academy of Sciences, 115(26), 201803155. https://doi.org/10.1073/pnas.1803155115

Iyer, S., Suresh, S., Guo, D., Daman, K., Chen, J. C. J., Liu, P., Zieger, M., Luk, K., Roscoe, B. P., Mueller, C., King, O. D., Emerson, C. P. & Wolfe, S. A. (2019). Precise therapeutic gene correction by a simple nuclease-induced double-stranded break. Nature, 568(7753), 561–565. https://doi.org/10.1038/s41586-019-1076-8

Jiang, F., Taylor, D. W., Chen, J. S., Kornfeld, J. E., Zhou, K., Thompson, A. J., Nogales, E. & Doudna, J. A. (2016). Structures of a CRISPR-Cas9 R-loop complex primed for DNA cleavage. Science, 351(6275), 867–871. https://doi.org/10.1126/science.aad8282

Jiang, L., Zheng, J., Kwan, J. S. H., Dai, S., Li, C., Li, M. J., Yu, B., TO, K. F., Sham, P. C., Zhu, Y. & Li, M. (2019). WITER: a powerful method for estimation of cancer-driver genes using a weighted iterative regression modelling background mutation counts. Nucleic Acids Research, 47(16), e96–e96. https://doi.org/10.1093/nar/gkz566

Jinek, M., Jiang, F., Taylor, D. W., Sternberg, S. H., Kaya, E., Ma, E., Anders, C., Hauer, M., Zhou, K., Lin, S., Kaplan, M., Iavarone, A. T., Charpentier, E., Nogales, E. & Doudna, J. A. (2014). Structures of Cas9 Endonucleases Reveal RNA-Mediated Conformational Activation. Science, 343(6176), 1247997–1247997.

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A. & Charpentier, E. (2012). A Programmable Dual-RNA–Guided DNA Endonuclease in Adaptive Bacterial Immunity. Science, 337(6096), 816–821. https://doi.org/10.1126/science.1225829

Jinek, M., East, A., Cheng, A., Lin, S., Ma, E. & Doudna, J. (2013). RNA-programmed genome editing in human cells. ELife, 2, e00471. https://doi.org/10.7554/elife.00471

Jonsson, P. F. & Bates, P. A. (2006). Global topological features of cancer proteins in the human interactome. Bioinformatics, 22(18), 2291–2297. https://doi.org/10.1093/bioinformatics/btl390

Jun, S., Lim, H., Chun, H., Lee, J. H. & Bang, D. (2020). Single-cell analysis of a mutant library generated using CRISPR-guided deaminase in human melanoma cells. Communications Biology, 3(1), 154. https://doi.org/10.1038/s42003-020-0888-2

Kastenhuber, E. R. & Lowe, S. W. (2017). Putting p53 in Context. Cell, 170(6), 1062–1078. https://doi.org/10.1016/j.cell.2017.08.028

Kaufman, H. L., Kohlhapp, F. J. & Zloza, A. (2015). Oncolytic viruses: a new class of immunotherapy drugs. Nature Reviews Drug Discovery, 14(9), 642–662. https://doi.org/10.1038/nrd4663

Kennedy, S. A., Jarboui, M.-A., Srihari, S., Raso, C., Bryan, K., Dernayka, L., Charitou, T., Bernal-Llinares, M., Herrera-Montavez, C., Krstic, A., Matallanas, D., Kotlyar, M., Jurisica, I., Curak, J., Wong, V., Stagljar, I., LeBihan, T., Imrie, L., Pillai, P., … Kolch, W. (2020). Extensive rewiring of the EGFR network in colorectal cancer cells expressing transforming levels of KRASG13D. Nature Communications, 11(1), 499. https://doi.org/10.1038/s41467-019-14224-9

Kerr, E. M., Gaude, E., Turrell, F. K., Frezza, C. & Martins, C. P. (2016). Mutant Kras copy number defines metabolic reprogramming and therapeutic susceptibilities. Nature, 531(7592), 110–113. https://doi.org/10.1038/nature16967

Kiessling, M. K., Schuierer, S., Stertz, S., Beibel, M., Bergling, S., Knehr, J., Carbone, W., Vallière, C. de, Tchinda, J., Bouwmeester, T., Seuwen, K., Rogler, G. & Roma, G. (2016). Identification of oncogenic driver mutations by genome-wide CRISPR-Cas9 dropout screening. BMC Genomics, 17(1), 723. https://doi.org/10.1186/s12864-016-3042-2

Kim, A. & Cohen, M. S. (2016). The discovery of vemurafenib for the treatment of BRAF-mutated metastatic melanoma. Expert Opinion on Drug Discovery, 11(9), 907–916. https://doi.org/10.1080/17460441.2016.1201057

Kim, D., Kim, D., Lee, G., Cho, S.-I. & Kim, J.-S. (2019). Genome-wide target specificity of CRISPR RNA-guided adenine base editors. Nature Biotechnology, 37(4), 430–435. https://doi.org/10.1038/s41587-019-0050-1

Kim, E., Koo, T., Park, S. W., Kim, D., Kim, K., Cho, H.-Y., Song, D. W., Lee, K. J., Jung, M. H., Kim, S., Kim, J. H., Kim, J. H. & Kim, J.-S. (2019). In vivo genome editing with a small Cas9 orthologue derived from Campylobacter jejuni. Nature Communications, 1–12.

Kim, W., Lee, S., Kim, H. S., Song, M., Cha, Y. H., Kim, Y.-H., Shin, J., Lee, E.-S., Joo, Y., Song, J. J., Choi, E. J., Choi, J. W., Lee, J., Kang, M., Yook, J. I., Lee, M. G., Kim, Y.-S., Paik, S. & Kim, H. (Henry). (2018). Targeting mutant KRAS with CRISPR-Cas9 controls tumor growth. Genome Research, 28(3), 374–382. https://doi.org/10.1101/gr.223891.117

Kim, Y. B., Komor, A. C., Levy, J. M., Packer, M. S., Zhao, K. T. & Liu, D. R. (2017). Increasing the genome-targeting scope and precision of base editing with engineered Cas9-cytidine deaminase fusions. Nature Biotechnology, 35(4), 371–376. https://doi.org/10.1038/nbt.3803

Kleinstiver, B. P., Pattanayak, V., Prew, M. S., Tsai, S. Q., Nguyen, N. T., Zheng, Z. & Joung, J. K. (2016). High-fidelity CRISPR–Cas9 nucleases with no detectable genome-wide off-target effects. Nature, 529(7587), 490–495. https://doi.org/10.1038/nature16526

Kleinstiver, B. P., Prew, M. S., Tsai, S. Q., Nguyen, N. T., Topkar, V. V., Zheng, Z. & Joung, J. K. (2015a). Broadening the targeting range of Staphylococcus aureus CRISPR-Cas9 by modifying PAM recognition. Nature Biotechnology, 33(12), 1293–1298. https://doi.org/10.1038/nbt.3404

Kleinstiver, B. P., Prew, M. S., Tsai, S. Q., Topkar, V. V., Nguyen, N. T., Zheng, Z., Gonzales, A. P. W., Li, Z., Peterson, R. T., Yeh, J.-R. J., Aryee, M. J. & Joung, J. K. (2015b). Engineered CRISPR-Cas9 nucleases with altered PAM specificities. Nature, 523(7561), 481–485. https://doi.org/10.1038/nature14592

Kleinstiver, B. P., Sousa, A. A., Walton, R. T., Tak, Y. E., Hsu, J. Y., Clement, K., Welch, M. M., Horng, J. E., Malagon-Lopez, J., Scarfò, I., Maus, M. V., Pinello, L., Aryee, M. J. & Joung, J. K. (2019). Engineered CRISPR–Cas12a variants with increased activities and improved targeting ranges for gene, epigenetic and base editing. Nature Biotechnology, 37(3), 276–282. https://doi.org/10.1038/s41587-018-0011-0

Kluesner, M. G., Nedveck, D. A., Lahr, W. S., Garbe, J. R., Abrahante, J. E., Webber, B. R. & Moriarity, B. S. (2018). EditR: A Method to Quantify Base Editing from Sanger

Sequencing. The CRISPR Journal, 1(3), 239–250. https://doi.org/10.1089/crispr.2018.0014

Koblan, L. W., Doman, J. L., Wilson, C., Levy, J. M., Tay, T., Newby, G. A., Maianti, J. P., Raguram, A. & Liu, D. R. (2018). Improving cytidine and adenine base editors by expression optimization and ancestral reconstruction. Nature Biotechnology, 36(9), 843–846. https://doi.org/10.1038/nbt.4172

Koike-Yusa, H., Li, Y., Tan, E.-P., Velasco-Herrera, M. D. C. & Yusa, K. (2014). Genome-wide recessive genetic screening in mammalian cells with a lentiviral CRISPR-guide RNA library. Nature Biotechnology, 32(3), 267–273. https://doi.org/10.1038/nbt.2800

Komor, A. C., Badran, A. H. & Liu, D. R. (2017). CRISPR-Based Technologies for the Manipulation of Eukaryotic Genomes. Cell, 168(1–2), 20–36. https://doi.org/10.1016/j.cell.2016.10.044

Komor, A. C., Kim, Y. B., Packer, M. S., Zuris, J. A. & Liu, D. R. (2016). Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. Nature, 533(7603), 420–424. https://doi.org/10.1038/nature17946

Kong, X., Chen, J., Xie, W., Brown, S. M., Cai, Y., Wu, K., Fan, D., Nie, Y., Yegnasubramanian, S., Tiedemann, R. L., Tao, Y., Yen, R.-W. C., Topper, M. J., Zahnow, C. A., Easwaran, H., Rothbart, S. B., Xia, L. & Baylin, S. B. (2019). Defining UHRF1 Domains that Support Maintenance of Human Colon Cancer DNA Methylation and Oncogenic Properties. Cancer Cell, 35(4), 633-648.e7.

Koo, T., Yoon, A.-R., Cho, H.-Y., Bae, S., Yun, C.-O. & Kim, J.-S. (2017). Selective disruption of an oncogenic mutant allele by CRISPR/Cas9 induces efficient tumor regression. Nucleic Acids Research, 45(13), gkx490-. https://doi.org/10.1093/nar/gkx490

Krasinskas, A. M., Moser, A. J., Saka, B., Adsay, N. V. & Chiosea, S. I. (2013). KRAS mutant allele-specific imbalance is associated with worse prognosis in pancreatic cancer and progression to undifferentiated carcinoma of the pancreas. Modern Pathology, 26(10), 1346–1354. https://doi.org/10.1038/modpathol.2013.71

Krontiris, T. G. & Cooper, G. M. (1981). Transforming activity of human tumor DNAs. Proceedings of the National Academy of Sciences, 78(2), 1181–1184. https://doi.org/10.1073/pnas.78.2.1181

Kunapuli, P., Somerville, R., Still, I. H. & Cowell, J. K. (2003). ZNF198 protein, involved in rearrangement in myeloproliferative disease, forms complexes with the DNA repair-associated HHR6A/6B and RAD18 proteins. Oncogene, 22(22), 3417–3423. https://doi.org/10.1038/sj.onc.1206408

Kuscu, C., Arslan, S., Singh, R., Thorpe, J. & Adli, M. (2014). Genome-wide analysis reveals characteristics of off-target sites bound by the Cas9 endonuclease. Nature Biotechnology, 32(7), 677–683. https://doi.org/10.1038/nbt.2916

Kuscu, C., Parlak, M., Tufan, T., Yang, J., Szlachta, K., Wei, X., Mammadov, R. & Adli, M. (2017). CRISPR-STOP: gene silencing through base-editing-induced nonsense mutations. Nature Methods, 14(7), 710–712. https://doi.org/10.1038/nmeth.4327

Landrum, M. J., Lee, J. M., Benson, M., Brown, G., Chao, C., Chitipiralla, S., Gu, B., Hart, J., Hoffman, D., Hoover, J., Jang, W., Katz, K., Ovetsky, M., Riley, G., Sethi, A., Tully, R., Villamarin-Salomon, R., Rubinstein, W. & Maglott, D. R. (2016). ClinVar: public archive

of interpretations of clinically relevant variants. Nucleic Acids Research, 44(D1), D862–D868. https://doi.org/10.1093/nar/gkv1222

Lapinaite, A., Knott, G. J., Palumbo, C. M., Lin-Shiao, E., Richter, M. F., Zhao, K. T., Beal, P. A., Liu, D. R. & Doudna, J. A. (2020). DNA capture by a CRISPR-Cas9–guided adenine base editor. Science, 369(6503), 566–571. https://doi.org/10.1126/science.abb1390

Lawrence, M. S., Stojanov, P., Polak, P., Kryukov, G. V., Cibulskis, K., Sivachenko, A., Carter, S. L., Stewart, C., Mermel, C. H., Roberts, S. A., Kiezun, A., Hammerman, P. S., McKenna, A., Drier, Y., Zou, L., Ramos, A. H., Pugh, T. J., Stransky, N., Helman, E., … Getz, G. (2013). Mutational heterogeneity in cancer and the search for new cancer-associated genes. Nature, 499(7457), 214–218. https://doi.org/10.1038/nature12213

Lee, C., Jo, D. H., Hwang, G.-H., Yu, J., Kim, J. H., Park, S., Kim, J.-S., Kim, J. H. & Bae, S. (2019). CRISPR-Pass: Gene Rescue of Nonsense Mutations Using Adenine Base Editors. Molecular Therapy, 27(8), 1364–1371. https://doi.org/10.1016/j.ymthe.2019.05.013

Lee, J. K., Jeong, E., Lee, J., Jung, M., Shin, E., Kim, Y., Lee, K., Jung, I., Kim, D., Kim, S. & Kim, J.-S. (2018). Directed evolution of CRISPR-Cas9 to increase its specificity. Nature Communications, 9(1), 3048. https://doi.org/10.1038/s41467-018-05477-x

Lee, William, Jiang, Z., Liu, J., Haverty, P. M., Guan, Y., Stinson, J., Yue, P., Zhang, Y., Pant, K. P., Bhatt, D., Ha, C., Johnson, S., Kennemer, M. I., Mohan, S., Nazarenko, I., Watanabe, C., Sparks, A. B., Shames, D. S., Gentleman, R., … Zhang, Z. (2010). The mutation spectrum revealed by paired genome sequences from a lung cancer patient. Nature, 465(7297), 473–477. https://doi.org/10.1038/nature09004

Lee, Wookjae, Lee, J. H., Jun, S., Lee, J. H. & Bang, D. (2018). Selective targeting of KRAS oncogenic alleles by CRISPR/Cas9 inhibits proliferation of cancer cells. Scientific Reports, 8(1), 11879. https://doi.org/10.1038/s41598-018-30205-2

Leenay, R. T. & Beisel, C. L. (2017). Deciphering, Communicating, and Engineering the CRISPR PAM. Journal of Molecular Biology, 429(2), 177–191. https://doi.org/10.1016/j.jmb.2016.11.024

Levy, J. M., Yeh, W.-H., Pendse, N., Davis, J. R., Hennessey, E., Butcher, R., Koblan, L. W., Comander, J., Liu, Q. & Liu, D. R. (2020). Cytosine and adenine base editing of the brain, liver, retina, heart and skeletal muscle of mice via adeno-associated viruses. Nature Biomedical Engineering, 4(1), 97–110. https://doi.org/10.1038/s41551-019-0501-5

Li, B.-Q., You, J., Chen, L., Zhang, J., Zhang, N., Li, H.-P., Huang, T., Kong, X.-Y. & Cai, Y.-D. (2013). Identification of Lung-Cancer-Related Genes with the Shortest Path Approach in a Protein-Protein Interaction Network. BioMed Research International, 2013, 1–8. https://doi.org/10.1155/2013/267375

Li, H. (2018). Minimap2: pairwise alignment for nucleotide sequences. Bioinformatics, 34(18), 3094–3100. https://doi.org/10.1093/bioinformatics/bty191

Li, K., Mao, C., Zhang, J., Ma, Q., Wang, Y., Liu, X., Bao, T. & Guo, W. (2019). Overexpression of U three protein 14a (UTP14a) is associated with poor prognosis of esophageal squamous cell carcinoma. Thoracic Cancer, 10(11), 2071–2080. https://doi.org/10.1111/1759-7714.13176

Liang, P., Xie, X., Zhi, S., Sun, H., Zhang, X., Chen, Y., Chen, Y., Xiong, Y., Ma, W., Liu, D., Huang, J. & Songyang, Z. (2019). Genome-wide profiling of adenine base editor specificity by EndoV-seq. Nature Communications, 10(1), 67. https://doi.org/10.1038/s41467-018-07988-z

Lisenbee, C. S., Karnik, S. K. & Trelease, R. N. (2003). Overexpression and mislocalization of a tail-anchored GFP redefines the identity of peroxisomal ER. Traffic (Copenhagen, Denmark), 4(7), 491–501.

Liu, J.-J., Orlova, N., Oakes, B. L., Ma, E., Spinner, H. B., Baney, K. L. M., Chuck, J., Tan, D., Knott, G. J., Harrington, L. B., Al-Shayeb, B., Wagner, A., Brötzmann, J., Staahl, B. T., Taylor, K. L., Desmarais, J., Nogales, E. & Doudna, J. A. (2019). CasX enzymes comprise a distinct family of RNA-guided genome editors. Nature, 566(7743), 218–223. https://doi.org/10.1038/s41586-019-0908-x

Liu, Y., Hu, X., Han, C., Wang, L., Zhang, X., He, X. & Lu, X. (2015). Targeting tumor suppressor genes for cancer therapy. BioEssays, 37(12), 1277–1286. https://doi.org/10.1002/bies.201500093

Liu, Z., Lu, Z., Yang, G., Huang, S., Li, G., Feng, S., Liu, Y., Li, J., Yu, W., Zhang, Y., Chen, J., Sun, Q. & Huang, X. (2018). Efficient generation of mouse models of human diseases via ABE- and BE-mediated base editing. Nature Communications, 9(1), 2338. https://doi.org/10.1038/s41467-018-04768-7

Lo, Y.-H., Karlsson, K. & Kuo, C. J. (2020). Applications of organoids for cancer biology and precision medicine. Nature Cancer, 1(8), 761–773. https://doi.org/10.1038/s43018-020-0102-y

Loeb, L. A. & Harris, C. C. (2008). Advances in Chemical Carcinogenesis: A Historical Review and Prospective. Cancer Research, 68(17), 6863–6872. https://doi.org/10.1158/0008-5472.can-08-2852

Lu, H., Chen, I., Shimoda, L. A., Park, Y., Zhang, C., Tran, L., Zhang, H. & Semenza, G. L. (2017). Chemotherapy-Induced Ca2+ Release Stimulates Breast Cancer Stem Cell Enrichment. Cell Reports, 18(8), 1946–1957. https://doi.org/10.1016/j.celrep.2017.02.001

Lukinavičius, G., Blaukopf, C., Pershagen, E., Schena, A., Reymond, L., Derivery, E., Gonzalez-Gaitan, M., D'Este, E., Hell, S. W., Gerlich, D. W. & Johnsson, K. (2015). SiR–Hoechst is a far-red DNA stain for live-cell nanoscopy. Nature Communications, 6(1), 8497. https://doi.org/10.1038/ncomms9497

Mali, P., Yang, L., Esvelt, K. M., Aach, J., Guell, M., DiCarlo, J. E., Norville, J. E. & Church, G. M. (2013). RNA-Guided Human Genome Engineering via Cas9. Science, 339(6121), 823–826. https://doi.org/10.1126/science.1232033

Mariot, P., Prevarskaya, N., Roudbaraki, M. M., Bourhis, X. L., Coppenolle, F. V., Vanoverberghe, K. & Skryma, R. (2000). Evidence of functional ryanodine receptor involved in apoptosis of prostate cancer (LNCaP) cells. The Prostate, 43(3), 205–214. https://doi.org/10.1002/(sici)1097-0045(20000515)43:3<205::aid-pros6>3.0.co;2-m

Martincorena, I. & Campbell, P. J. (2015). Somatic mutation in cancer and normal cells. Science, 349(6255), 1483–1489. https://doi.org/10.1126/science.aab4082

Martincorena, I., Raine, K. M., Gerstung, M., Dawson, K. J., Haase, K., Loo, P. V., Davies, H., Stratton, M. R. & Campbell, P. J. (2017). Universal Patterns of Selection in Cancer and Somatic Tissues. Cell, 171(5), 1029-1041.e21. https://doi.org/10.1016/j.cell.2017.09.042

McCarthy, T. V., Datar, S. & Mackrill, J. J. (2003). Activation of ryanodine receptor/Ca2+ release channels downregulates CD38 in the Namalwa B lymphoma. FEBS Letters, 554(1–2), 133–137. https://doi.org/10.1016/s0014-5793(03)01122-0

McCormick, F. (2015). KRAS as a Therapeutic Target. Clinical Cancer Research, 21(8), 1797–1801. https://doi.org/10.1158/1078-0432.ccr-14-2662

McDonald, E. R., Weck, A. de, Schlabach, M. R., Billy, E., Mavrakis, K. J., Hoffman, G. R., Belur, D., Castelletti, D., Frias, E., Gampa, K., Golji, J., Kao, I., Li, L., Megel, P., Perkins, T. A., Ramadan, N., Ruddy, D. A., Silver, S. J., Sovath, S., … Sellers, W. R. (2017). Project DRIVE: A Compendium of Cancer Dependencies and Synthetic Lethal Relationships Uncovered by Large-Scale, Deep RNAi Screening. Cell, 170(3), 577-592.e10. https://doi.org/10.1016/j.cell.2017.07.005

Miles, L. A., Garippa, R. J. & Poirier, J. T. (2016). Design, execution, and analysis of pooled in vitro CRISPR/Cas9 screens. The FEBS Journal, 283(17), 3170–3180. https://doi.org/10.1111/febs.13770

Milne, I., Stephen, G., Bayer, M., Cock, P. J. A., Pritchard, L., Cardle, L., Shaw, P. D. & Marshall, D. (2013). Using Tablet for visual exploration of second-generation sequencing data. Briefings in Bioinformatics, 14(2), 193–202. https://doi.org/10.1093/bib/bbs012

Mojica, F. J. M., Díez-Villaseñor, C., García-Martínez, J. & Soria, E. (2005). Intervening Sequences of Regularly Spaced Prokaryotic Repeats Derive from Foreign Genetic Elements. Journal of Molecular Evolution, 60(2), 174–182. https://doi.org/10.1007/s00239-004-0046-3

Molla, K. A. & Yang, Y. (2019). CRISPR/Cas-Mediated Base Editing: Technical Considerations and Practical Applications. Trends in Biotechnology, 37(10), 1121–1142. https://doi.org/10.1016/j.tibtech.2019.03.008

Montague, T. G., Cruz, J. M., Gagnon, J. A., Church, G. M. & Valen, E. (2014). CHOPCHOP: a CRISPR/Cas9 and TALEN web tool for genome editing. Nucleic Acids Research, 42(W1), W401–W407. https://doi.org/10.1093/nar/gku410

Mouneimne, G., Hansen, S. D., Selfors, L. M., Petrak, L., Hickey, M. M., Gallegos, L. L., Simpson, K. J., Lim, J., Gertler, F. B., Hartwig, J. H., Mullins, R. D. & Brugge, J. S. (2012). Differential Remodeling of Actin Cytoskeleton Architecture by Profilin Isoforms Leads to Distinct Effects on Cell Migration and Invasion. Cancer Cell, 22(5), 615–630. https://doi.org/10.1016/j.ccr.2012.09.027

Mukaihara, K., Suehara, Y., Kohsaka, S., Kubota, D., Toda-Ishii, M., Akaike, K., Fujimura, T., Kobayashi, E., Yao, T., Ladanyi, M., Kaneko, K. & Saito, T. (2016). Expression of F-actin-capping protein subunit beta, CAPZB, is associated with cell growth and motility in epithelioid sarcoma. BMC Cancer, 16(1), 206. https://doi.org/10.1186/s12885-016-2235-z

Mularoni, L., Sabarinathan, R., Deu-Pons, J., Gonzalez-Perez, A. & López-Bigas, N. (2016). OncodriveFML: a general framework to identify coding and non-coding regions with

cancer driver mutations. Genome Biology, 17(1), 128. https://doi.org/10.1186/s13059-016-0994-0

Mullard, A. (2020). Gene-editing pipeline takes off. Nature Reviews Drug Discovery, 1–6. https://doi.org/10.1038/d41573-020-00096-y

Müller, M., Lee, C. M., Gasiunas, G., Davis, T. H., Cradick, T. J., Siksnys, V., Bao, G., Cathomen, T. & Mussolino, C. (2016). Streptococcus thermophilus CRISPR-Cas9 Systems Enable Specific Editing of the Human Genome. Molecular Therapy, 24(3), 636–644. https://doi.org/10.1038/mt.2015.218

Muzumdar, M. D., Chen, P.-Y., Dorans, K. J., Chung, K. M., Bhutkar, A., Hong, E., Noll, E. M., Sprick, M. R., Trumpp, A. & Jacks, T. (2017). Survival of pancreatic cancer cells lacking KRAS function. Nature Communications, 1–18.

Neal, J. T. & Kuo, C. J. (2015). Organoids as Models for Neoplastic Transformation. Annual Review of Pathology: Mechanisms of Disease, 11(1), 1–22. https://doi.org/10.1146/annurev-pathol-012615-044249

Nishimasu, H., Ran, F. A., Hsu, P. D., Konermann, S., Shehata, S. I., Dohmae, N., Ishitani, R., Zhang, F. & Nureki, O. (2014). Crystal Structure of Cas9 in Complex with Guide RNA and Target DNA. Cell, 156(5), 935–949. https://doi.org/10.1016/j.cell.2014.02.001

Nishimasu, H., Shi, X., Ishiguro, S., Gao, L., Hirano, S., Okazaki, S., Noda, T., Abudayyeh, O. O., Gootenberg, J. S., Mori, H., Oura, S., Holmes, B., Tanaka, M., Seki, M., Hirano, H., Aburatani, H., Ishitani, R., Ikawa, M., Yachie, N., … Nureki, O. (2018). Engineered CRISPR-Cas9 nuclease with expanded targeting space. Science, 361(6408), 1259–1262.

Nordling, C. O. (1953). A New Theory on the Cancer-inducing Mechanism. British Journal of Cancer, 7(1), 68–72. https://doi.org/10.1038/bjc.1953.8

Nowak, C. M., Lawson, S., Zerez, M. & Bleris, L. (2016). Guide RNA engineering for versatile Cas9 functionality. Nucleic Acids Research, 44(20), 9555–9564. https://doi.org/10.1093/nar/gkw908

Orth, M., Metzger, P., Gerum, S., Mayerle, J., Schneider, G., Belka, C., Schnurr, M. & Lauber, K. (2019). Pancreatic ductal adenocarcinoma: biological hallmarks, current status, and future perspectives of combined modality treatment approaches. Radiation Oncology, 14(1), 141. https://doi.org/10.1186/s13014-019-1345-6

Ostrem, J. M. L. & Shokat, K. M. (2016). Direct small-molecule inhibitors of KRAS: from structural insights to mechanism-based design. Nature Reviews Drug Discovery, 15(11), 771–785. https://doi.org/10.1038/nrd.2016.139

Ostrem, J. M., Peters, U., Sos, M. L., Wells, J. A. & Shokat, K. M. (2013). K-Ras(G12C) inhibitors allosterically control GTP affinity and effector interactions. Nature, 503(7477), 548–551. https://doi.org/10.1038/nature12796

Padilla-Rodriguez, M., Parker, S. S., Adams, D. G., Westerling, T., Puleo, J. I., Watson, A. W., Hill, S. M., Noon, M., Gaudin, R., Aaron, J., Tong, D., Roe, D. J., Knudsen, B. & Mouneimne, G. (2018). The actin cytoskeletal architecture of estrogen receptor positive breast cancer cells suppresses invasion. Nature Communications, 9(1), 2980. https://doi.org/10.1038/s41467-018-05367-2

Papke, B. & Der, C. J. (2017). Drugging RAS: Know the enemy. Science, 355(6330), 1158–1163. https://doi.org/10.1126/science.aam7622

Parnas, O., Jovanovic, M., Eisenhaure, T. M., Herbst, R. H., Dixit, A., Ye, C. J., Przybylski, D., Platt, R. J., Tirosh, I., Sanjana, N. E., Shalem, O., Satija, R., Raychowdhury, R., Mertins, P., Carr, S. A., Zhang, F., Hacohen, N. & Regev, A. (2015). A Genome-wide CRISPR Screen in Primary Immune Cells to Dissect Regulatory Networks. Cell, 162(3), 675–686. https://doi.org/10.1016/j.cell.2015.06.059

Parsons, D. W., Jones, S., Zhang, X., Lin, J. C.-H., Leary, R. J., Angenendt, P., Mankoo, P., Carter, H., Siu, I.-M., Gallia, G. L., Olivi, A., McLendon, R., Rasheed, B. A., Keir, S., Nikolskaya, T., Nikolsky, Y., Busam, D. A., Tekleab, H., Diaz, L. A., … Kinzler, K. W. (2008). An Integrated Genomic Analysis of Human Glioblastoma Multiforme. Science, 321(5897), 1807–1812. https://doi.org/10.1126/science.1164382

Patricelli, M. P., Janes, M. R., Li, L.-S., Hansen, R., Peters, U., Kessler, L. V., Chen, Y., Kucharski, J. M., Feng, J., Ely, T., Chen, J. H., Firdaus, S. J., Babbar, A., Ren, P. & Liu, Y. (2016). Selective Inhibition of Oncogenic KRAS Output with Small Molecules Targeting the Inactive State. Cancer Discovery, 6(3), 316–329. https://doi.org/10.1158/2159-8290.cd-15-1105

Peng, J., Zhou, Y., Zhu, S. & Wei, W. (2015). High-throughput screens in mammalian cells using the CRISPR-Cas9 system. FEBS Journal, 282(11), 2089–2096. https://doi.org/10.1111/febs.13251

Pestana, A., Vinagre, J., Sobrinho-Simões, M. & Soares, P. (2017). TERT biology and function in cancer: beyond immortalisation. Journal of Molecular Endocrinology, 58(2), R129–R146. https://doi.org/10.1530/jme-16-0195

Pickar-Oliver, A. & Gersbach, C. A. (2019). The next generation of CRISPR–Cas technologies and applications. Nature Reviews Molecular Cell Biology, 20(8), 490–507. https://doi.org/10.1038/s41580-019-0131-5

Poser, I., Sarov, M., Hutchins, J. R. A., Hériché, J.-K., Toyoda, Y., Pozniakovsky, A., Weigl, D., Nitzsche, A., Hegemann, B., Bird, A. W., Pelletier, L., Kittler, R., Hua, S., Naumann, R., Augsburg, M., Sykora, M. M., Hofemeister, H., Zhang, Y., Nasmyth, K., … Hyman, A. A. (2008). BAC TransgeneOmics: a high-throughput method for exploration of protein function in mammals. Nature Methods, 5(5), 409–415. https://doi.org/10.1038/nmeth.1199

Prior, I. A., Lewis, P. D. & Mattos, C. (2012). A Comprehensive Survey of Ras Mutations in Cancer. Cancer Research, 72(10), 2457–2467. https://doi.org/10.1158/0008-5472.can-11-2612

Quante, A. S., Ming, C., Rottmann, M., Engel, J., Boeck, S., Heinemann, V., Westphalen, C. B. & Strauch, K. (2016). Projections of cancer incidence and cancer-related deaths in Germany by 2020 and 2030. Cancer Medicine, 5(9), 2649–2656. https://doi.org/10.1002/cam4.767

Rambaldi, D., Giorgi, F. M., Capuani, F., Ciliberto, A. & Ciccarelli, F. D. (2008). Low duplicability and network fragility of cancer genes. Trends in Genetics, 24(9), 427–430. https://doi.org/10.1016/j.tig.2008.06.003

Ran, F. A., Cong, L., Yan, W. X., Scott, D. A., Gootenberg, J. S., Kriz, A. J., Zetsche, B., Shalem, O., Wu, X., Makarova, K. S., Koonin, E. V., Sharp, P. A. & Zhang, F. (2015). In

vivo genome editing using Staphylococcus aureus Cas9. Nature, 520(7546), 186–191. https://doi.org/10.1038/nature14299

Ran, F. A., Hsu, P. D., Lin, C.-Y., Gootenberg, J. S., Konermann, S., Trevino, A. E., Scott, D. A., Inoue, A., Matoba, S., Zhang, Y. & Zhang, F. (2013a). Double Nicking by RNA-Guided CRISPR Cas9 for Enhanced Genome Editing Specificity. Cell, 154(6), 1380–1389. https://doi.org/10.1016/j.cell.2013.08.021

Ran, F. A., Hsu, P. D., Wright, J., Agarwala, V., Scott, D. A. & Zhang, F. (2013b). Genome engineering using the CRISPR-Cas9 system. Nature Protocols, 8(11), 2281–2308. https://doi.org/10.1038/nprot.2013.143

Rauscher, B., Heigwer, F., Henkel, L., Hielscher, T., Voloshanenko, O. & Boutros, M. (2018). Toward an integrated map of genetic interactions in cancer cells. Molecular Systems Biology, 14(2), e7656. https://doi.org/10.15252/msb.20177656

Reddy, E. P., Reynolds, R. K., Santos, E. & Barbacid, M. (1982). A point mutation is responsible for the acquisition of transforming properties by the T24 human bladder carcinoma oncogene. Nature, 300(5888), 149–152.

Rees, H. A. & Liu, D. R. (2018). Base editing: precision chemistry on the genome and transcriptome of living cells. Nature Reviews Genetics, 19(12), 770–788. https://doi.org/10.1038/s41576-018-0059-1

Ren, B., Liu, L., Li, S., Kuang, Y., Wang, J., Zhang, D., Zhou, X., Lin, H. & Zhou, H. (2019). Cas9-NG Greatly Expands the Targeting Scope of the Genome-Editing Toolkit by Recognizing NG and Other Atypical PAMs in Rice. Molecular Plant, 12(7), 1015–1026. https://doi.org/10.1016/j.molp.2019.03.010

Ren, P., Sun, X., Zhang, C., Wang, L., Xing, B. & Du, X. (2019). Human UTP14a promotes angiogenesis through upregulating PDGFA expression in colorectal cancer. Biochemical and Biophysical Research Communications, 512(4), 871–876. https://doi.org/10.1016/j.bbrc.2019.03.142

Richter, M. F., Zhao, K. T., Eton, E., Lapinaite, A., Newby, G. A., Thuronyi, B. W., Wilson, C., Koblan, L. W., Zeng, J., Bauer, D. E., Doudna, J. A. & Liu, D. R. (2020). Phage-assisted evolution of an adenine base editor with improved Cas domain compatibility and activity. Nature Biotechnology, 38(7), 883–891. https://doi.org/10.1038/s41587-020-0453-z

Rossidis, A. C., Stratigis, J. D., Chadwick, A. C., Hartman, H. A., Ahn, N. J., Li, H., Singh, K., Coons, B. E., Li, L., Lv, W., Zoltick, P. W., Alapati, D., Zacharias, W., Jain, R., Morrisey, E. E., Musunuru, K. & Peranteau, W. H. (2018). In utero CRISPR-mediated therapeutic editing of metabolic genes. Nature Medicine, 24(10), 1513–1518. https://doi.org/10.1038/s41591-018-0184-6

Rowley, J. D. (1973). A New Consistent Chromosomal Abnormality in Chronic Myelogenous Leukaemia identified by Quinacrine Fluorescence and Giemsa Staining. Nature, 243(5405), 290–293. https://doi.org/10.1038/243290a0

Sander, J. D. & Joung, J. K. (2014). CRISPR-Cas systems for editing, regulating and targeting genomes. Nature Biotechnology, 32(4), 347–355. https://doi.org/10.1038/nbt.2842

Sanjana, N. E. (2017). Genome-scale CRISPR pooled screens. Analytical Biochemistry, 532, 95–99. https://doi.org/10.1016/j.ab.2016.05.014

Sayed, S., Paszkowski-Rogacz, M., Schmitt, L. T. & Buchholz, F. (2019). CRISPR/Cas9 as a tool to dissect cancer mutations. Methods, 164, 36–48. https://doi.org/10.1016/j.ymeth.2019.05.007

Scolnick, E. M., Rands, E., Williams, D. & Parks, W. P. (1973). Studies on the nucleic acid sequences of Kirsten sarcoma virus: a model for formation of a mammalian RNA-containing sarcoma virus. Journal of Virology, 12(3), 458–463.

Seidlitz, T., Merker, S. R., Rothe, A., Zakrzewski, F., Neubeck, C. von, Grützmann, K., Sommer, U., Schweitzer, C., Schölch, S., Uhlemann, H., Gaebler, A.-M., Werner, K., Krause, M., Baretton, G. B., Welsch, T., Koo, B.-K., Aust, D. E., Klink, B., Weitz, J. & Stange, D. E. (2019). Human gastric cancer modelling using organoids. Gut, 68(2), 207. https://doi.org/10.1136/gutjnl-2017-314549

Seino, T., Kawasaki, S., Shimokawa, M., Tamagawa, H., Toshimitsu, K., Fujii, M., Ohta, Y., Matano, M., Nanki, K., Kawasaki, K., Takahashi, S., Sugimoto, S., Iwasaki, E., Takagi, J., Itoi, T., Kitago, M., Kitagawa, Y., Kanai, T. & Sato, T. (2018). Human Pancreatic Tumor Organoids Reveal Loss of Stem Cell Niche Factor Dependence during Disease Progression. Cell Stem Cell, 22(3), 454-467.e6. https://doi.org/10.1016/j.stem.2017.12.009

Serebriiskii, I. G., Connelly, C., Frampton, G., Newberg, J., Cooke, M., Miller, V., Ali, S., Ross, J. S., Handorf, E., Arora, S., Lieu, C., Golemis, E. A. & Meyer, J. E. (2019). Comprehensive characterization of RAS mutations in colon and rectal cancers in old and young patients. Nature Communications, 10(1), 3722. https://doi.org/10.1038/s41467-019-11530-0

Shah, S. P., Roth, A., Goya, R., Oloumi, A., Ha, G., Zhao, Y., Turashvili, G., Ding, J., Tse, K., Haffari, G., Bashashati, A., Prentice, L. M., Khattra, J., Burleigh, A., Yap, D., Bernard, V., McPherson, A., Shumansky, K., Crisan, A., … Aparicio, S. (2012). The clonal and mutational evolution spectrum of primary triple-negative breast cancers. Nature, 486(7403), 395–399. https://doi.org/10.1038/nature10933

Shalem, O., Sanjana, N. E., Hartenian, E., Shi, X., Scott, D. A., Mikkelsen, T. S., Heckl, D., Ebert, B. L., Root, D. E., Doench, J. G. & Zhang, F. (2014). Genome-Scale CRISPR-Cas9 Knockout Screening in Human Cells. Science, 343(6166), 84–87. https://doi.org/10.1126/science.1247005

Shalem, O., Sanjana, N. E. & Zhang, F. (2015). High-throughput functional genomics using CRISPR–Cas9. Nature Reviews Genetics, 16(5), 299–311. https://doi.org/10.1038/nrg3899

Sharma, S. V., Gajowniczek, P., Way, I. P., Lee, D. Y., Jiang, J., Yuza, Y., Classon, M., Haber, D. A. & Settleman, J. (2006). A common signaling cascade may underlie “addiction” to the Src, BCR-ABL, and EGF receptor oncogenes. Cancer Cell, 10(5), 425–435. https://doi.org/10.1016/j.ccr.2006.09.014

Shen, M. W., Arbab, M., Hsu, J. Y., Worstell, D., Culbertson, S. J., Krabbe, O., Cassa, C. A., Liu, D. R., Gifford, D. K. & Sherwood, R. I. (2018). Predictable and precise template-free CRISPR editing of pathogenic variants. Nature, 563(7733), 646–651. https://doi.org/10.1038/s41586-018-0686-x

Shi, J., Wang, E., Milazzo, J. P., Wang, Z., Kinney, J. B. & Vakoc, C. R. (2015). Discovery of cancer drug targets by CRISPR-Cas9 screening of protein domains. Nature Biotechnology, 33(6), 661–667. https://doi.org/10.1038/nbt.3235

Shi, W., Dong, F., Jiang, Y., Lu, L., Wang, C., Tan, J., Yang, W., Guo, H., Ming, J. & Huang, T. (2019). Construction of prognostic microRNA signature for human invasive breast cancer by integrated analysis . OncoTargets and Therapy, Volume 12, 1979–2010.

Shih, C., Padhy, L. C., Murray, M. & Weinberg, R. A. (1981). Transforming genes of carcinomas and neuroblastomas introduced into mouse fibroblasts. Nature, 290(5803), 261–264. https://doi.org/10.1038/290261a0

Shimatani, Z., Kashojiya, S., Takayama, M., Terada, R., Arazoe, T., Ishii, H., Teramura, H., Yamamoto, T., Komatsu, H., Miura, K., Ezura, H., Nishida, K., Ariizumi, T. & Kondo, A. (2017). Targeted base editing in rice and tomato using a CRISPR-Cas9 cytidine deaminase fusion. Nature Biotechnology, 35(5), 441–443. https://doi.org/10.1038/nbt.3833

Shirasawa, S., Furuse, M., Yokoyama, N. & Sasazuki, T. (1993). Altered growth of human colon cancer cell lines disrupted at activated Ki-ras. Science, 260(5104), 85–88. https://doi.org/10.1126/science.8465203

Shlien, A., Campbell, B. B., Borja, R. de, Alexandrov, L. B., Merico, D., Wedge, D., Loo, P. V., Tarpey, P. S., Coupland, P., Behjati, S., Pollett, A., Lipman, T., Heidari, A., Deshmukh, S., Avitzur, N., Meier, B., Gerstung, M., Hong, Y., Merino, D. M., … Consortium, B. M. R. D. (2015). Combined hereditary and somatic mutations of replication error repair genes result in rapid onset of ultra-hypermutated cancers. Nature Genetics, 47(3), 257–262. https://doi.org/10.1038/ng.3202

Shmakov, S., Abudayyeh, O. O., Makarova, K. S., Wolf, Y. I., Gootenberg, J. S., Semenova, E., Minakhin, L., Joung, J., Konermann, S., Severinov, K., Zhang, F. & Koonin, E. V. (2015). Discovery and Functional Characterization of Diverse Class 2 CRISPR-Cas Systems. Molecular Cell, 60(3), 385–397. https://doi.org/10.1016/j.molcel.2015.10.008

Shrubsole, M. J., Gao, Y.-T., Cai, Q., Shu, X. O., Dai, Q., Jin, F. & Zheng, W. (2006). MTR and MTRR Polymorphisms, Dietary Intake, and Breast Cancer Risk. Cancer Epidemiology and Prevention Biomarkers, 15(3), 586–588. https://doi.org/10.1158/1055-9965.epi-05-0576

Simanshu, D. K., Nissley, D. V. & McCormick, F. (2017). RAS Proteins and Their Regulators in Human Disease. Cell, 170(1), 17–33. https://doi.org/10.1016/j.cell.2017.06.009

Sondka, Z., Bamford, S., Cole, C. G., Ward, S. A., Dunham, I. & Forbes, S. A. (2018). The COSMIC Cancer Gene Census: describing genetic dysfunction across all human cancers. Nature Reviews Cancer, 18(11), 696–705. https://doi.org/10.1038/s41568-018-0060-1

Song, C.-Q., Jiang, T., Richter, M., Rhym, L. H., Koblan, L. W., Zafra, M. P., Schatoff, E. M., Doman, J. L., Cao, Y., Dow, L. E., Zhu, L. J., Anderson, D. G., Liu, D. R., Yin, H. & Xue, W. (2020). Adenine base editing in an adult mouse model of tyrosinaemia. Nature Biomedical Engineering, 4(1), 125–130. https://doi.org/10.1038/s41551-019-0357-8

Song, M., Kim, H. K., Lee, S., Kim, Y., Seo, S.-Y., Park, J., Choi, J. W., Jang, H., Shin, J. H., Min, S., Quan, Z., Kim, J. H., Kang, H. C., Yoon, S. & Kim, H. H. (2020). Sequence-specific prediction of the efficiencies of adenine and cytosine base editors. Nature Biotechnology, 38(9), 1037–1043. https://doi.org/10.1038/s41587-020-0573-5

Steckel, M., Molina-Arcas, M., Weigelt, B., Marani, M., Warne, P. H., Kuznetsov, H., Kelly, G., Saunders, B., Howell, M., Downward, J. & Hancock, D. C. (2012). Determination of

synthetic lethal interactions in KRAS oncogene-dependent cancer cells reveals novel therapeutic targeting strategies. Cell Research, 22(8), 1227–1245. https://doi.org/10.1038/cr.2012.82

Stemmer, M., Thumberger, T., Keyer, M. del S., Wittbrodt, J. & Mateo, J. L. (2015). CCTop: An Intuitive, Flexible and Reliable CRISPR/Cas9 Target Prediction Tool. PLOS ONE, 10(4), e0124633. https://doi.org/10.1371/journal.pone.0124633

Sternberg, S. H., LaFrance, B., Kaplan, M. & Doudna, J. A. (2015). Conformational control of DNA target cleavage by CRISPR–Cas9. Nature, 527(7576), 110–113. https://doi.org/10.1038/nature15544

Sternberg, S. H., LaFrance, B., Kaplan, M. & Doudna, J. A. (2019). Conformational control of DNA target cleavage by CRISPR–Cas9. Nature, 1–14.

Sternberg, S. H., Redding, S., Jinek, M., Greene, E. C. & Doudna, J. A. (2014). DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. Nature, 507(7490), 62–67. https://doi.org/10.1038/nature13011

Stirewalt, D. L. & Radich, J. P. (2003). The role of FLT3 in haematopoietic malignancies. Nature Reviews Cancer, 3(9), 650–665. https://doi.org/10.1038/nrc1169

Stratton, M. R., Campbell, P. J. & Futreal, P. A. (2009). The cancer genome. Nature, 458(7239), 719–724. https://doi.org/10.1038/nature07943

Suh, S., Choi, E. H., Leinonen, H., Foik, A. T., Newby, G. A., Yeh, W.-H., Dong, Z., Kiser, P. D., Lyon, D. C., Liu, D. R. & Palczewski, K. (2021). Restoration of visual function in adult mice with an inherited retinal disease via adenine base editing. Nature Biomedical Engineering, 5(2), 169–178. https://doi.org/10.1038/s41551-020-00632-6

Sun, L.-C. & Qian, H.-X. (2018). Screening for implicated genes in colorectal cancer using whole-genome gene expression profiling. Molecular Medicine Reports, 17(6), 8260–8268. https://doi.org/10.3892/mmr.2018.8862

Sürün, D., Schneider, A., Mircetic, J., Neumann, K., Lansing, F., Paszkowski-Rogacz, M., Hänchen, V., Lee-Kirsch, M. A. & Buchholz, F. (2020). Efficient Generation and Correction of Mutations in Human iPS Cells Utilizing mRNAs of CRISPR Base Editors and Prime Editors. Genes, 11(5), 511. https://doi.org/10.3390/genes11050511

Szczelkun, M. D., Tikhomirova, M. S., Sinkunas, T., Gasiunas, G., Karvelis, T., Pschera, P., Siksnys, V. & Seidel, R. (2014). Direct observation of R-loop formation by single RNA-guided Cas9 and Cascade effector complexes. Proceedings of the National Academy of Sciences, 111(27), 9798–9803. https://doi.org/10.1073/pnas.1402597111

Tabin, C. J., Bradley, S. M., Bargmann, C. I., Weinberg, R. A., Papageorge, A. G., Scolnick, E. M., Dhar, R., Lowy, D. R. & Chang, E. H. (1982). Mechanism of activation of a human oncogene. Nature, 300(5888), 143–149.

Tamborero, D., Gonzalez-Perez, A., Perez-Llamas, C., Deu-Pons, J., Kandoth, C., Reimand, J., Lawrence, M. S., Getz, G., Bader, G. D., Ding, L. & Lopez-Bigas, N. (2013). Comprehensive identification of mutational cancer driver genes across 12 tumor types. Scientific Reports, 3(1), 2650. https://doi.org/10.1038/srep02650

Tan, J., Zhang, F., Karcher, D. & Bock, R. (2019). Engineering of high-precision base editors for site-specific single nucleotide replacement. Nature Communications, 10(1), 439. https://doi.org/10.1038/s41467-018-08034-8

Tang, W. & Liu., D. R. (2018). Rewritable multi-event analog recording in bacterial and mammalian cells. Science, 360(6385), eaap8992. https://doi.org/10.1126/science.aap8992

Tate, J. G., Bamford, S., Jubb, H. C., Sondka, Z., Beare, D. M., Bindal, N., Boutselakis, H., Cole, C. G., Creatore, C., Dawson, E., Fish, P., Harsha, B., Hathaway, C., Jupe, S. C., Kok, C. Y., Noble, K., Ponting, L., Ramshaw, C. C., Rye, C. E., … Forbes, S. A. (2018). COSMIC: the Catalogue Of Somatic Mutations In Cancer. Nucleic Acids Research, 47(D1), gky1015-. https://doi.org/10.1093/nar/gky1015

TCGA Network. (2012). Comprehensive molecular characterization of human colon and rectal cancer. Nature, 487(7407), 330–337. https://doi.org/10.1038/nature11252

Thakore, P. I., Black, J. B., Hilton, I. B. & Gersbach, C. A. (2016). Editing the epigenome: technologies for programmable transcription and epigenetic modulation. Nature Methods, 13(2), 127–137. https://doi.org/10.1038/nmeth.3733

Thuronyi, B. W., Koblan, L. W., Levy, J. M., Yeh, W.-H., Zheng, C., Newby, G. A., Wilson, C., Bhaumik, M., Shubina-Oleinik, O., Holt, J. R. & Liu, D. R. (2019). Continuous evolution of base editors with expanded target compatibility and improved activity. Nature Biotechnology, 37(9), 1070–1079. https://doi.org/10.1038/s41587-019-0193-0

Tiriac, H., Plenker, D., Baker, L. A. & Tuveson, D. A. (2019). Organoid models for translational pancreatic cancer research. Current Opinion in Genetics & Development, 54, 7–11. https://doi.org/10.1016/j.gde.2019.02.003

Tomasetti, C., Marchionni, L., Nowak, M. A., Parmigiani, G. & Vogelstein, B. (2015). Only three driver gene mutations are required for the development of lung and colorectal cancers. Proceedings of the National Academy of Sciences, 112(1), 118–123. https://doi.org/10.1073/pnas.1421839112

Torkamani, A. & Schork, N. J. (2008). Prediction of Cancer Driver Mutations in Protein Kinases. Cancer Research, 68(6), 1675–1682. https://doi.org/10.1158/0008-5472.can-07-5283

Tsai, S. Q., Nguyen, N. T., Malagon-Lopez, J., Topkar, V. V., Aryee, M. J. & Joung, J. K. (2017). CIRCLE-seq: a highly sensitive in vitro screen for genome-wide CRISPR–Cas9 nuclease off-targets. Nature Methods, 14(6), 607–614. https://doi.org/10.1038/nmeth.4278

Tsherniak, A., Vazquez, F., Montgomery, P. G., Weir, B. A., Kryukov, G., Cowley, G. S., Gill, S., Harrington, W. F., Pantel, S., Krill-Burger, J. M., Meyers, R. M., Ali, L., Goodale, A., Lee, Y., Jiang, G., Hsiao, J., Gerath, W. F. J., Howell, S., Merkel, E., … Hahn, W. C. (2017). Defining a Cancer Dependency Map. Cell, 170(3), 564-576.e16. https://doi.org/10.1016/j.cell.2017.06.010

Tzelepis, K., Koike-Yusa, H., De Braekeleer, E., Li, Y., Metzakopian, E., Dovey, O. M., Mupo, A., Grinkevich, V., Li, M., Mazan, M., Gozdecka, M., Ohnishi, S., Cooper, J., Patel, M., McKerrell, T., Chen, B., Domingues, A. F., Gallipoli, P., Teichmann, S., … Yusa, K. (2016). A CRISPR Dropout Screen Identifies Genetic Vulnerabilities and Therapeutic

Targets in Acute Myeloid Leukemia. Cell Reports, 17(4), 1193–1205. https://doi.org/10.1016/j.celrep.2016.09.079

van Overbeek, M., Capurso, D., Carter, M. M., Thompson, M. S., Frias, E., Russ, C., Reece-Hoyes, J. S., Nye, C., Gradia, S., Vidal, B., Zheng, J., Hoffman, G. R., Fuller, C. K. & May, A. P. (2016). DNA Repair Profiling Reveals Nonrandom Outcomes at Cas9-Mediated Breaks. Molecular Cell, 63(4), 633–646. https://doi.org/10.1016/j.molcel.2016.06.037

Vogelstein, B., Papadopoulos, N., Velculescu, V. E., Zhou, S., Diaz, L. A. & Kinzler, K. W. (2013). Cancer Genome Landscapes. Science, 339(6127), 1546–1558. https://doi.org/10.1126/science.1235122

Wallace, J., Hu, R., Mosbruger, T. L., Dahlem, T. J., Stephens, W. Z., Rao, D. S., Round, J. L. & O'Connell, R. M. (2016). Genome-Wide CRISPR-Cas9 Screen Identifies MicroRNAs That Regulate Myeloid Leukemia Cell Growth. PLOS ONE, 11(4), e0153689. https://doi.org/10.1371/journal.pone.0153689

Walton, R. T., Christie, K. A., Whittaker, M. N. & Kleinstiver, B. P. (2020). Unconstrained genome targeting with near-PAMless engineered CRISPR-Cas9 variants. Science, 368(6488), 290–296. https://doi.org/10.1126/science.aba8853

Wan, T., Chen, Y., Pan, Q., Xu, X., Kang, Y., Gao, X., Huang, F., Wu, C. & Ping, Y. (2020). Genome editing of mutant KRAS through supramolecular polymer-mediated delivery of Cas9 ribonucleoprotein for colorectal cancer therapy. Journal of Controlled Release, 322, 236–247.

Wang, H., Bian, X., Xia, L., Ding, X., Müller, R., Zhang, Y., Fu, J. & Stewart, A. F. (2014). Improved seamless mutagenesis by recombineering using ccdB for counterselection. Nucleic Acids Research, 42(5), e37–e37. https://doi.org/10.1093/nar/gkt1339

Wang, T., Wei, J. J., Sabatini, D. M. & Lander, E. S. (2014). Genetic Screens in Human Cells Using the CRISPR-Cas9 System. Science, 343(6166), 80–84. https://doi.org/10.1126/science.1246981

Wang, X., Huang, X., Fang, X., Zhang, Y. & Wang, W. (2016). CRISPR-Cas9 System as a Versatile Tool for Genome Engineering in Human Cells. Molecular Therapy - Nucleic Acids, 5(11), e388. https://doi.org/10.1038/mtna.2016.95

Watson, J. D. & Crick, F. H. C. (1953). Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid. Nature, 171(4356), 737–738.

Weghorn, D. & Sunyaev, S. (2017). Bayesian inference of negative and positive selection in human cancers. Nature Genetics, 49(12), 1785–1788. https://doi.org/10.1038/ng.3987

Welman, A., Burger M. M., Hagmann, J. (2000). Structure and function of the C-terminal hypervariable region of K-Ras4B in plasma membrane targeting and transformation. Oncogene, 19, 4582-4591. https://doi.org/10.1038/sj.onc.1203818

Wong, A. S. L., Choi, G. C. G., Cui, C. H., Pregernig, G., Milani, P., Adam, M., Perli, S. D., Kazer, S. W., Gaillard, A., Hermann, M., Shalek, A. K., Fraenkel, E. & Lu, T. K. (2016). Multiplexed barcoded CRISPR-Cas9 screening enabled by CombiGEM. Proceedings of the National Academy of Sciences, 113(9), 2544–2549. https://doi.org/10.1073/pnas.1517883113

Xue, J. Y., Zhao, Y., Aronowitz, J., Mai, T. T., Vides, A., Qeriqi, B., Kim, D., Li, C., Stanchina, E. de, Mazutis, L., Risso, D. & Lito, P. (2020). Rapid non-uniform adaptation to conformation-specific KRAS(G12C) inhibition. Nature, 577(7790), 421–425. https://doi.org/10.1038/s41586-019-1884-x

Yang, M., Wei, H., Wang, Y., Deng, J., Tang, Y., Zhou, L., Guo, G. & Tong, A. (2017). Targeted Disruption of V600E-Mutant BRAF Gene by CRISPR-Cpf1. Molecular Therapy - Nucleic Acids, 8, 450–458. https://doi.org/10.1016/j.omtn.2017.05.009

Yeh, W.-H., Chiang, H., Rees, H. A., Edge, A. S. B. & Liu, D. R. (2018). In vivo base editing of post-mitotic sensory cells. Nature Communications, 9(1), 2184. https://doi.org/10.1038/s41467-018-04580-3

Yu, H. A., Sima, C. S., Shen, R., Kass, S., Gainor, J., Shaw, A., Hames, M., Iams, W., Aston, J., Lovly, C. M., Horn, L., Lydon, C., Oxnard, G. R., Kris, M. G., Ladanyi, M. & Riely, G. J. (2015). Prognostic Impact of KRAS Mutation Subtypes in 677 Patients with Metastatic Lung Adenocarcinomas. Journal of Thoracic Oncology, 10(3), 431–437. https://doi.org/10.1097/jto.0000000000000432

Yuan, J., Ma, Y., Huang, T., Chen, Y., Peng, Y., Li, B., Li, J., Zhang, Y., Song, B., Sun, X., Ding, Q., Song, Y. & Chang, X. (2018). Genetic Modulation of RNA Splicing with a CRISPR-Guided Cytidine Deaminase. Molecular Cell, 72(2), 380-394.e7. https://doi.org/10.1016/j.molcel.2018.09.002

Yuan, L., Qin, X., Li, L., Zhou, J., Zhou, M., Li, X., Xu, Y., Wang, X. & Xing, H. (2019). The transcriptome profiles and methylation status revealed the potential cancer-related lncRNAs in patients with cervical cancer. Journal of Cellular Physiology, 234(6), 9756–9763. https://doi.org/10.1002/jcp.27661

Yuan, T. L., Fellmann, C., Lee, C.-S., Ritchie, C. D., Thapar, V., Lee, L. C., Hsu, D. J., Grace, D., Carver, J. O., Zuber, J., Luo, J., McCormick, F. & Lowe, S. W. (2014). Development of siRNA Payloads to Target KRAS-Mutant Cancer. Cancer Discovery, 4(10), 1182–1197. https://doi.org/10.1158/2159-8290.cd-13-0900

Yuan, X., Larsson, C. & Xu, D. (2019). Mechanisms underlying the activation of TERT transcription and telomerase activity in human cancer: old actors and new players. Oncogene, 38(34), 6172–6183. https://doi.org/10.1038/s41388-019-0872-9

Zafra, M. P., Parsons, M. J., Kim, J., Alonso-Curbelo, D., Goswami, S., Schatoff, E. M., Han, T., Katti, A., Fernandez, M. T. C., Wilkinson, J. E., Piskounova, E. & Dow, L. E. (2020). An In Vivo Kras Allelic Series Reveals Distinct Phenotypes of Common Oncogenic Variants. Cancer Discov, 10(11), 1654. E

Zafra, M. P., Schatoff, E. M., Katti, A., Foronda, M., Breinig, M., Schweitzer, A. Y., Simon, A., Han, T., Goswami, S., Montgomery, E., Thibado, J., Kastenhuber, E. R., Sánchez-Rivera, F. J., Shi, J., Vakoc, C. R., Lowe, S. W., Tschaharganeh, D. F. & Dow, L. E. (2018). Optimized base editors enable efficient editing in cells, organoids and mice. Nature Biotechnology, 36(9), 888–893. https://doi.org/10.1038/nbt.4194

Zeng, J., Wu, Y., Ren, C., Bonanno, J., Shen, A. H., Shea, D., Gehrke, J. M., Clement, K., Luk, K., Yao, Q., Kim, R., Wolfe, S. A., Manis, J. P., Pinello, L., Joung, J. K. & Bauer, D. E. (2020). Therapeutic base editing of human hematopoietic stem cells. Nature Medicine, 26(4), 535–541. https://doi.org/10.1038/s41591-020-0790-y

Zhang, J., Ren, P., Xu, D., Liu, X., Liu, Z., Zhang, C., Li, Y., Wang, L., Du, X. & Xing, B. (2019). Human UTP14a promotes colorectal cancer progression by forming a positive regulation loop with c-Myc. Cancer Letters, 440, 106–115. https://doi.org/10.1016/j.canlet.2018.10.010

Zhang, J.-Y., Xu, D., Liu, Z.-Z., Li, Y., Wang, L.-J. & Xing, B.-C. (2017). Human U Three Protein 14a Expression is Increased in Hepatocellular Carcinoma and Associated with Poor Prognosis. Chinese Medical Journal, 130(4), 470–476. https://doi.org/10.4103/0366-6999.199839

Zhang, R., Miner, J. J., Gorman, M. J., Rausch, K., Ramage, H., White, J. P., Zuiani, A., Zhang, P., Fernandez, E., Zhang, Q., Dowd, K. A., Pierson, T. C., Cherry, S. & Diamond, M. S. (2016). A CRISPR screen defines a signal peptide processing pathway required by flaviviruses. Nature, 535(7610), 164–168. https://doi.org/10.1038/nature18625

Zhang, W., Bojorquez-Gomez, A., Velez, D. O., Xu, G., Sanchez, K. S., Shen, J. P., Chen, K., Licon, K., Melton, C., Olson, K. M., Yu, M. K., Huang, J. K., Carter, H., Farley, E. K., Snyder, M., Fraley, S. I., Kreisberg, J. F. & Ideker, T. (2018). A global transcriptional network connecting noncoding mutations to changes in tumor gene expression. Nature Genetics, 50(4), 613–620. https://doi.org/10.1038/s41588-018-0091-2

Zheng, T., Hou, Y., Zhang, P., Zhang, Z., Xu, Y., Zhang, L., Niu, L., Yang, Y., Liang, D., Yi, F., Peng, W., Feng, W., Yang, Y., Chen, J., Zhu, Y. Y., Zhang, L.-H. & Du, Q. (2017). Profiling single-guide RNA specificity reveals a mismatch sensitive core sequence. Scientific Reports, 7(1), 40638. https://doi.org/10.1038/srep40638

Zhu, J., Liu, X., Anjos, M., Correll, C. C. & Johnson, A. W. (2016). Utp14 Recruits and Activates the RNA Helicase Dhr1 To Undock U3 snoRNA from the Preribosome. Molecular and Cellular Biology, 36(6), 965–978. https://doi.org/10.1128/mcb.00773-15

7 Acknowledgments

During my PhD journey, I got to meet many wonderful people who inspired me and, directly or indirectly, helped me to accomplish this work. I would like to say a big “Thank you!” to everyone I have encountered along the way, in particular I want to thank:

My supervisor, **Frank Buchholz** for the opportunity to work on this exciting project, and for believing in me to follow my own ideas while always being there for guidance and support. I owe him a lot of my scientific growth over the past four years, and to him I am forever grateful. Second note of gratitude goes to my PhD committee members, **Martin Bornhäuser** and **Dirk Lindemann**, for their support and valuable input in our annual meetings.

I would like to thank all the Buchholz group members, both past and present, for the nice and collaborative environment, in particular I want to thank: **Sandeep Sreevalsan** whom I had the privilege to sit next to in the office, for Western blot tips and help with image analysis. **Jovan Mirčetić** for introducing me to the organoids world, and I am truly grateful to him for his insightful comments and critical feedback on this dissertation. **Aylin Camgöz** for initially demonstrating various techniques including cell culture, virus productions and for always being available for discussions. **Olga Sidorova** for sharing this journey with me, for occasionally taking care of the cells in my absence and for SOS help with R scripts. **Martina Augsburg** for her excellent technical assistance and help with Time-lapse Microscopy. **Duran Sürün** for advice on cloning, occasional help with cell sorting and useful suggestions while writing this dissertation. My co-authors: **Lukas Schmitt,** for helping with Nanopore sequencing experiment and **Maciej Paszkowski-Rogacz** for designing sgRNAs for RKO screen, (and for maintaining our blood caffeine levels). **Catherine Cortés**, my rotation student, for help with the arrayed screen in HCT116 cells. **Felix Lansing**, for the interesting science and non-science lunch discussions, and generally for his genuine personality over the years. **The Schneiders: Aksana** and **Martin** for their sincere friendship and support, be it experimental or emotional. **Jenna Hoersten** for the running motivation and feedback on figures. A huge thank you to **Mandy Erlitz**, **Carolina Blanck** and **Sebastian Rose** for keeping things together at the lab in every aspect, and for the continuous administrative and logistical support. Finally, I have used **BioRender** for all the figures I made for this dissertation as well as for various talks throughout my PhD- massive thanks to the team behind for making my life a lot easier!

I am grateful to **Frank Groß** for sharing the bacterial strains necessary for the recombineering experiment. And to our collaborators **Daniel Stange** and **Alexander Hennig**, for the pancreatic organoids samples and advice with regards to organoids culture.

Outside the lab, there are several people whom I would like to thank for creating a wonderful social life in Dresden: My squash partner: **Avinash Patel** for the long rallies and tiebreakers, and generally the great time on and off the court. The Egyptian crew in Dresden, especially **Ezz Alfar** and **Marwa Magdy** for offering their place a safe harbor to dock and stock up on good food, and good laughs. My violin teacher **Josephine Bernewitz**, for the great mood she sets with a simple bow stroke on her violin and for teaching me all the cool notes. I am grateful to **David Kilian**, **Andres Sote**, and **Nicola Mitwasi** for their friendship over the past six years. Being part of the Student Representatives has been an invaluable experience, I want to thank the **StudReps** squad, both past and present, for a lively science community, for all the fun and organising activities. And I cannot forget the DIPP project “Science goes to school”. A shout-out to **Anja Glenk** for organizing this amazing project. Hope we can resume once again soon.

I grew up in Cairo, where like big cities life is hectic and fast-paced. From there I treasure having several life-long friends who have acted as solid reference points whenever needed. All my love to **Elhamy**, **Sharkawy**, **Nour**, **Luma**, **Moataz**, **Ezz**, **Hamido**, and **Mokhtar**. I am most grateful to **Shehab** for his friendship over the years, and for the unwavering supply of Humour/memes, which never failed to brighten the days up.

Before I end up, I’d like to dedicate a note of support and compassion to future PhD students, who might read this dissertation: I hope you remember to stay positive when experiments aren’t going as expected. I hope you remember to remain consistent and to be patient, it does pay off. Finally, and most importantly, enjoy every second of the entire process.

Finally, a note of appreciation to my big brother, **Ramy**, for always having my back since the very beginning, and to my little sister, **Fayrouz**, for being the sweetest sister anybody can ask for. As stated in the beginning of this book, this work is dedicated to **my parents**, I hope everyone has as supportive and open-minded family as mine. I would like to end this chapter (and thesis) by addressing them in a few words:

أبويا وأمي الغاليين
الجزء ده من الرسالة بيكون شكر للناس اللي ساهمت في نجاح هذا العمل.. والحقيقة مش هقدر أختم من غير ما أشكركم. أشكركم من كل قلبي على تعبكم ومساندتكم لي بإستمرار.. أشكركم على تحمل غيابي طوال السنين اللي فاتت. الفضل في كل شيء حققته يرجع لكم وإن شاءالله أحقق كل ماتتمنوه (سامعك يا غالية:)

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 ausschließlich die im Manuskript angegebenen Unterstützungsleistungen erhalten:
-nicht zutreffend-
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:

Sayed, Shady, Maciej Paszkowski-Rogacz, Lukas Theo Schmitt, and Frank Buchholz. 2019. "CRISPR/Cas9 as a Tool to Dissect Cancer Mutations." *Methods* 164-165 (August). Elsevier: 36–48.

6. Ich bestätige, dass es keine zurückliegenden erfolgslosen 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.

Dresden, den 21 Dezember 2020

Shady Sayed

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

  -nicht zutreffend-

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

  -nicht zutreffend-

- die Einhaltung des Gentechnikgesetzes

  Projektnummer: 55-8811.71/197 (S1)
  55-8811.72 / 89 (S2)

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

Dresden, den 21 Dezember 2020

Shady Sayed