DETECTION OF DELETERIOUS ON-TARGET EFFECTS AFTER CRISPR-MEDIATED GENOME EDITING IN HUMAN INDUCED PLURIPOTENT STEM CELLS

Isabel Weisheit

Graduate School of Systemic Neurosciences

LMU Munich

Dissertation der Graduate School of Systemic Neurosciences
der Ludwig-Maximilians-Universität München

2nd of February, 2021

Supervisor and first reviewer:
Prof. Dr. Dominik Paquet (Institute for Stroke and Dementia Research)

Second reviewer:
Prof. Dr. Martin Dichgans (Institute for Stroke and Dementia Research)

Date of Submission: 02.02.2021
Date of Defense: 02.07.2021

Table of Contents

Abstract

The CRISPR/Cas9 system is an exceedingly powerful technology for precise genome editing. Its ease of use, high editing efficiency and an ever-growing CRISPR-based toolbox has provided researchers with novel possibilities to unravel the molecular and systemic consequences of changes in the genetic code. For this reason, CRISPR is now applied for editing in a wide range of different cell lines and organisms for basic and translational research. Here, accurate and precise editing is an indispensable prerequisite to generate reliable research models. However, a lot remains to be understood about the molecular mechanism of double-strand-breaks (DSBs) in the DNA as introduced by the Cas9 nuclease during editing. In fact, CRISPR editing can be accompanied by inadvertent genomic changes at the targeted locus (on-target) as well as other genomic sites (off-target). These can have drastic consequences on gene activity or expression and therefore need to be carefully investigated.

Characterizing and avoiding unwanted off-target effects (OffTE) has been the focus of several studies and reliable tools for their detection have been developed. This is, however, not the case for on-target effects (OnTE) that have only been reported very recently. These can be large deletions, large insertions, complex rearrangements, or regions of copy-neutral loss of heterozygosity (LOH) around the target site. Several studies have described frequent occurrence of OnTEs in mice, but it has not been investigated if clinically relevant human cells, such as induced pluripotent stem cells (iPSCs) are also affected. The main problem with OnTEs is that they often remain unnoticed in standard quality controls like Sanger genotyping of the target locus, and additional checks are lacking in most CRISPR-based studies. This is also because there are no simple detection tools available.

Therefore, in this study, we developed simple and reliable tools for OnTE detection after CRISPR genome editing: Structural alterations like large deletions, large insertions or complex rearrangements can be identified by quantifying the number of intact alleles at the edited locus using our new method called quantitative genotyping PCR (qgPCR). In addition, we validated genotyping of neighboring single nucleotide polymorphisms (SNPs) either by Sanger sequencing or SNP microarrays to reveal editing-induced regions of LOH. The entire workflow is broadly applicable to different cell lines and organisms after editing by the NHEJ or HDR pathway.

We have applied our newly established detection technology to human iPSCs after HDR-mediated editing and demonstrate universal occurrence of OnTEs at multiple loci in up to 40% of edited single-cell clones. Furthermore, using an *in vitro* model of Alzheimer's disease, we illustrate deleterious consequences of OnTEs on expression of the edited gene that may reduce pathogenic effects and therefore interfere with experimental findings.

Overall, the threat of undetected OnTEs undermining the reliability of CRISPR-based studies has not received sufficient attention in the field so far. With this thesis, we hope to raise further awareness and propose that our simple and reliable on-target quality control workflow should be an essential part of all relevant genome editing experiments.

List of Abbreviations

APP^{Swe}	APPSwedish
bp	base-pair
Cas	CRISPR-associated
CGH	comparative genomic hybridization
CRISPR	clustered regularly interspaced short palindromic repeat
crRNA	CRISPR-RNA
C_t	threshold cycle
DSB	double-strand break
ESC	embryonic stem cell
GWAS	genome-wide association study
HDR	homology-directed repair
HiFi	high-fidelity
InDel	insertion/deletion
iPSC	induced pluripotent stem cell
LOH	loss of heterozygosity
MMEJ	microhomology-mediated end-joining
NHEJ	non-homologous end joining
nCas9	nickase Cas9
nt	nucleotide
OffTE	off-target effect
OnTE	on-target effect
PAM	protospacer adjacent motif
pegRNA	prime editing guide RNA
qgPCR	quantitative genotyping PCR
qPCR	quantitative PCR
RNP	ribonucleoprotein
sgRNA	single guide RNA
SNP	single nucleotide polymorphism
SpCas9	*Streptococcus pyogenes* Cas9
SSB	single-strand break
ssODN	single-stranded oligodeoxynucleotide
TALEN	transcription activator-like effector nucleases
tracrRNA	trans-activating crRNA
UGI	uracil DNA glycosylase inhibitor
ZFN	zinc finger nucleases

1. Introduction

1.1 The CRISPR/Cas9 system revolutionizes biomedical research

The idea of modifying the genomic sequence of living cells or organisms at will is a fascinating thought that has become reality with the recent discovery of the CRISPR/Cas9 system. It enables precise insertion or exchange of specific DNA sequences at virtually any genomic locus in a simple and rapid way. This drastically transforms the approach of many molecular biologists to unravel the connection between the genomic sequence and specific phenotypes, thereby opening endless new possibilities for basic research (reviewed in Hsu et al., 2014; Knott & Doudna, 2018; Komor et al., 2017; Wright et al., 2016). Furthermore, with several CRISPR-based clinical trials being currently conducted, its application for the treatment of genetically-encoded human diseases is in the immediate future (reviewed in Ernst et al., 2020; Mullard, 2020).

1.1.1 A bacterial immune system is transformed into a genome editing tool

The CRISPR/Cas9 system was initially discovered to be part of the adaptive immune system of bacteria and archaea as defense mechanism against foreign nucleic acids of invading bacteriophages (Barrangou et al., 2007; Brouns et al., 2008). During an invasion, foreign DNA is cleaved and incorporated as spacer sequences into the clustered regularly interspaced short palindromic repeats (CRISPR) array that contains multiple spacers from previous attacks. These spacers are then continuously expressed as CRISPR-RNA (crRNA) that bind to a trans-activating crRNA (tracrRNA) as well as the CRISPR-associated (Cas9) nuclease. This complex can then recognize a repeated attack by the same virus through the spacer sequence, which leads to targeted cuts in the bacteriophage DNA and its subsequent degradation (Deltcheva et al., 2011; Jinek et al., 2012) (Figure 1). Realizing the potential of this RNA-guided nuclease system for genome editing, it was shown that purified Cas9 can be programmed to cut and thereby introduce targeted changes into DNA *in vitro* (Gasiunas et al., 2012; Jinek et al., 2012) as well as *in vivo* into the genomic DNA of cultured human cancer cell lines and induced pluripotent stem cell lines (Cho et al., 2013; Cong et al., 2013; Jinek et al., 2013; Mali et al., 2013).

Starting with these first studies reporting successful genome editing using the CRISPR/Cas9 system, many researchers were fascinated by its ease of use and wide applicability. In the next years, a lot of effort was made to characterize the underlying molecular mechanism and further advance efficiency and versatility of the method.

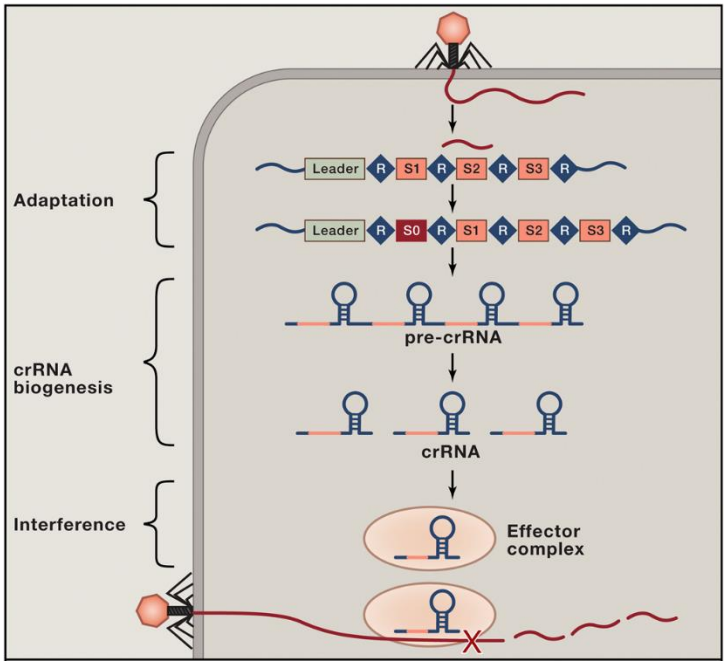


Figure 1: Mechanism of CRISPR-based adaptive immunity. Adaptation (top): Foreign nucleic acids from invading bacteriophages are processed and inserted as new spacer sequence into the CRISPR array. The CRISPR array contains spacer sequences from previous attacks segregated by repeat (R) elements. crRNA biogenesis (middle): The CRISPR array is transcribed and processed into individual crRNAs that assemble (e.g., with Cas9 and a tracrRNA) into an effector complex. Interference (bottom): The effector complex recognizes a repeated attack by the bacteriophage and initiates degradation of the foreign DNA. Image modified from Wright et al., 2016.

1.1.2 Introducing targeted DNA DSBs using the CRISPR/Cas9 system

The power of the CRISPR/Cas9 system is that targeted DSBs can be introduced into the DNA at a desired locus. This requires a complex of only two components: a single guide RNA (sgRNA) (which is an artificial fusion of crRNA and tracrRNA (Jinek et al., 2012)) and the Cas9 endonuclease. The sgRNA contains a scaffold sequence for interaction with the Cas9 nuclease and a 20 nucleotide (nt) spacer sequence at its 5' end that is complementary to the target locus. This spacer sequence guides the Cas9 nuclease to its DNA target (protospacer) by complementary base-pairing and thereby determines the specificity of the complex. To utilize this system for genome editing, the spacer sequence of the sgRNA can be exchanged freely. The only requirement is the presence of a 5'-NGG' protospacer adjacent motif (PAM) downstream of the protospacer sequence in the DNA (Figure 2A). These PAM motifs appear on average every 8 base-pairs (bp) in the human genome (Cong et al., 2013), which usually offers plenty of options to target a specific locus. The DSB will then be introduced 3 bp upstream of the PAM sequence.

Earlier site-specific nucleases that were used for genome editing, including zinc finger nucleases (ZFNs), transcription activator-like effector nucleases (TALENs) and meganucleases, required complex protein engineering to customize binding properties of the nuclease to the target locus (reviewed in Hsu et al., 2014). In contrast, the specificity of the

CRISPR system is solely based on complementary base pairing between nucleic acids that can easily be modified. This simplicity has made the CRISPR system rapidly into the leading tool for genome editing.

1.1.3 Exploiting cellular DNA repair mechanisms for genome editing

DNA DSBs are repaired by two main pathways: non-homologous end joining (NHEJ) or to a lesser extent homology-directed repair (HDR) (Rouet Philippe et al., 1994). The NHEJ pathway directly fuses two DNA ends together and typically results in different types of insertions or deletions (InDel) of usually 1-10 bp around the cut site (Mali et al., 2013; Paquet et al., 2016). For genome editing purposes, non-specific InDels that cause frameshift mutations can be used for gene knock-outs through premature stop codons on both alleles. The HDR pathway typically uses the homologous sister chromatid as template for errorless repair, but the pathway can also be exploited to introduce specific changes by providing an exogenous repair template, such as a single-stranded oligodeoxynucleotide (ssODN), that carries the desired edit with homology arms around the break site (Figure 2B) (Mali et al., 2013).

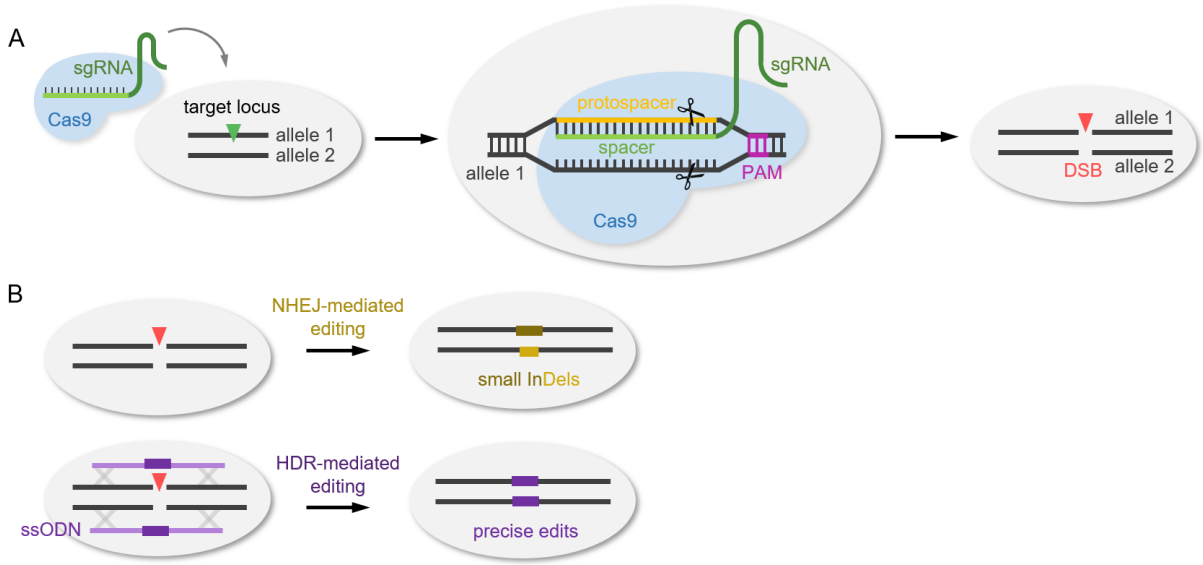


Figure 2: Principle of genome editing using the CRISPR/Cas9 system. (A) The Cas9 nuclease/sgRNA complex binds to a specific locus in the DNA through complementary base-pairing between the spacer sequence (light green) of the sgRNA and the protospacer sequence (yellow) in the DNA, with the PAM sequence (purple) 'NGG' directly downstream of the target sequence. A DSB (scissors and red triangle) is then introduced by the Cas9 nuclease 3-4 nucleotides upstream of the PAM sequence. (B) CRISPR-induced DSBs are mainly repaired in one of two ways: In the NHEJ pathway (top), both loose ends of the DSB are reconnected, which can result in non-specific small InDels at the cut site. Alternatively, to insert specific bp changes, one can exploit the cells intrinsic HDR pathway (bottom). Here, a repair template, e.g., a ssODN, carrying the desired bp changes can be supplied, which are then integrated into the DNA.

1.2 Using induced pluripotent stem cells to generate *in vitro* models of human disease

Another groundbreaking discovery that has revolutionized biomedical research was made a few years before the adaptation of the CRISPR system for genome editing: in 2008, Takahashi and Yamanaka described the reprogramming of mouse fibroblasts into cells that closely resemble embryonic stem cells (ESCs) by introduction of four specific exogenous factors: Oct3/4, Sox2, c-Myc, and Klf4. Hence, they called these cells induced pluripotent stem cells (iPSCs) (Takahashi & Yamanaka, 2006). A follow-up study one year later then reported that the same factors can also be used to reprogram human fibroblasts (Takahashi et al., 2007). In defined culture conditions, these iPSCs maintain their pluripotent state and can divide indefinitely, but the development of *in vitro* differentiation protocols also enables the generation of various somatic cell types from the same iPSC source. This is a great opportunity for disease-oriented research, as it allows studying phenotypes in different human cell types that are directly affected in patients. Especially research on diseases where the supply of affected cells is highly limited, such as neurodegenerative diseases, have greatly benefited from the iPSC technology: versatile protocols for the differentiation of brain cell types have been developed (reviewed in McComish & Caldwell, 2018) and are already widely applied to investigate different disease mechanisms (reviewed in TCW, 2019).

The generation of iPSCs eliminates ethical concerns about extraction of ESCs from fertilized embryos and additionally provides a sheer endless supply of cells that harbor different genetic backgrounds. This is highly useful to study diseases, as patient-derived iPSCs can be compared to iPSC lines from heathy controls. One caveat of this approach is, however, the high genetic variation between donors. These can either mask phenotypic effects of the mutation of interest or lead to non-specific phenotypes. For this reason, the CRISPR system provides an ideal addition to iPSC research: genetic modification to correct or introduce disease-associated mutations enables the generation of matching isogenic control lines, which allows studying the direct effects of genotypic changes in a homogenous background (Figure 3).

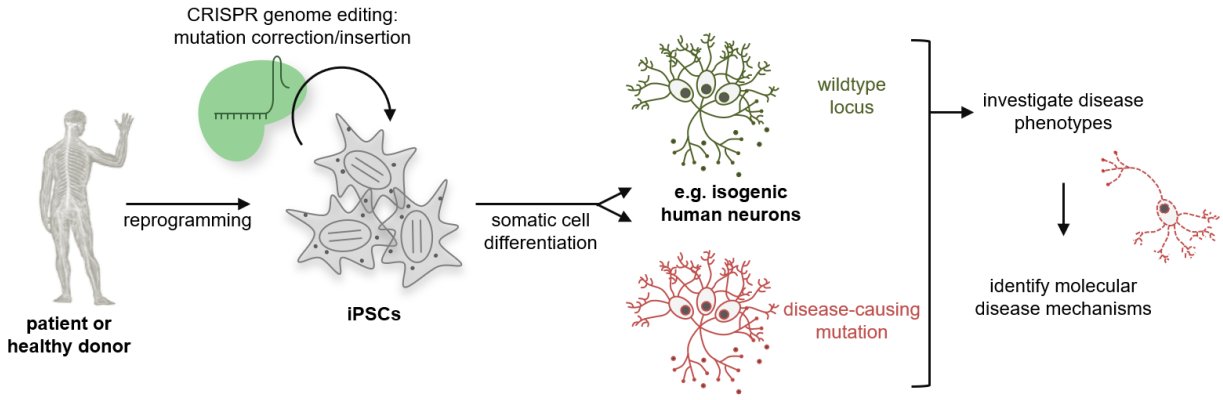


Figure 3: Overview of generating human *in vitro* disease models using CRISPR-edited iPSCs.

The approach of using CRISPR-edited iPSCs for the generation of *in vitro* disease models is now widely used with the goal to unravel molecular disease mechanisms and to identify novel drug targets (Hockemeyer & Jaenisch, 2016). Genome editing of stem cells, however, is very challenging and hampered by low efficiencies, which requires the implementation of protocols for human iPSCs (Kwart et al., 2017) and also the development of new strategies to increase efficiency and versatility of the CRISPR/Cas9 system.

1.3 Enhancing the efficiency of precise genome editing by the CRISPR/Cas9 system

As described above, CRISPR-induced DSBs in the genomic DNA are typically repaired by the NHEJ pathway, which leads to non-specific InDels at the cut site. Correction or insertion of specific disease-causing mutations, however, requires precise editing using the HDR pathway, which is usually much less active. This research field therefore greatly benefits from new ideas that are contributed by inventive scientists to further advance the development of genome editing (reviewed in Pickar-Oliver & Gersbach, 2019). Some of these strategies include developing general guidelines for accurate gene editing, increasing the efficiency of precise gene editing in particular, identifying and engineering new CRISPR variants as well as developing new gene editing technologies. These points will be outlined in more detail in the following.

1.3.1 Identifying fundamental principles for precise CRISPR editing

A study from Paquet et al. established fundamental guidelines to enable efficient and precise knock-in of homozygous and heterozygous mutations using the CRISPR/Cas9 system (Paquet et al., 2016): first, they showed that it is necessary to inhibit Cas9 from re-cutting and destroying the locus after a successful edit has been made, which can be done by introducing silent blocking mutations that alter the PAM sequence or gRNA recognition site. If introducing silent mutations is not possible or desired, a two-step editing process termed ‘CORRECT’ may be applied for scarless genome editing. Furthermore, the authors characterized an inverse relationship between distance of Cas9 cleavage site and mutation site with the efficiency of genome editing. This distance effect can be used to guide zygosity of the introduced edit.

1.3.2 Increasing the prevalence of DSB repair by the HDR pathway

One of the main challenges that researchers face to perform precision editing is that the NHEJ pathway is usually more efficient for DSB repair than the HDR pathway. However, only the latter one can be used to introduce specific changes. Therefore, many different strategies were developed to increase the rate of repair by HDR (reviewed in Danner et al., 2017). Some of these strategies include suppression of key NHEJ factors by small molecules (Chu et al., 2015; Maruyama et al., 2015) or synchronization of the cell cycle together with delivery of pre-assembled Cas9 ribonucleoprotein (RNP) complexes to enhance nuclease activity in stages

when the HDR pathway is most active (Lin et al., 2014). However, these approaches may interfere with the general cellular DNA damage response and could therefore lead to unwanted changes at loci other than the targeted site. Less invasive strategies were therefore developed that are restricted to the gene editing components itself, e.g., rational design of repair templates (Richardson et al., 2016), or cell cycle-dependent posttranslational regulation of Cas9 (Gutschner et al., 2016).

1.3.3 Identification of new CRISPR systems

CRISPR systems are adaptive defense mechanisms that are widespread in bacteria and archaea (Wiedenheft et al., 2012). Given the fact that these organisms are highly diverse and dynamic, also different CRISPR systems have evolved over time that are now classified into different classes, types and subtypes (Makarova et al., 2020). The Cas9 nuclease that was first used for genome editing purposes - and is still the most widely used variant today - was originally isolated from the pathogen *Streptococcus pyogenes* (SpCas9) and belongs to the class II CRISPR systems (Jinek et al., 2012). This class is characterized by the presence of one single-component effector protein, which make these systems very useful for genome editing purposes. In addition to SpCas9, more Cas proteins from different organisms have been identified, e.g., Cas12a, CasX and Casφ, which all have unique features in comparison to Cas9 in regard to PAM requirement, size, DNA cleavage mechanisms and/or targeting specificity, which make them promising candidates to help broaden the scope of genome editing applications (Liu et al., 2019; Pausch et al., 2020; Zetsche et al., 2015). Another improvement was the development of engineered Cas9 protein versions that have altered PAM requirements (Kleinstiver, Prew, Tsai, Nguyen, et al., 2015; Kleinstiver, Prew, Tsai, Topkar, et al., 2015). These greatly help to further expand the range of targetable sequences.

1.3.4 Development of new CRISPR-based editing technologies

Researchers aiming to do precise gene editing are often faced with inadvertent side-effects caused by DSB induction (see below) as well as high rates of unwanted InDel mutations at the target locus. Moreover, given that HDR editing is only active in the S and G2 phase of the cell cycle, this approach is restricted to actively dividing cells. To overcome these limitations of the CRISPR/Cas9 system to some extent, the lab of David Liu has developed new technologies for gene editing in recent years, called base editing and prime editing (reviewed in Anzalone et al., 2020). Both systems use a Cas9 version that is mutated at one nuclease domain yielding ‘nickase’ Cas9, which introduces only single-strand breaks (SSB) into genomic DNA. For base editors, a deaminase is fused to Cas9 that can initiate the direct conversion of one base to another (C•G to T•A base pair substitution for cytidine deaminases (Komor et al., 2016) or A•T to G•C for adenosine deaminases (Gaudelli et al., 2017)). For prime editing, a modified prime

editing guide RNA (pegRNA) both guides to a specific locus and also carries the desired edit, which is copied and inserted into the locus by a reverse transcriptase that is fused to Cas9 (Anzalone et al., 2019) (Figure 4). This technique - unlike base editing - does not only enable base substitutions, but also insertions and deletions at the target locus. In short, base editing and prime editing enable DSB-free and template-free editing with much lower rates of unwanted side-products.

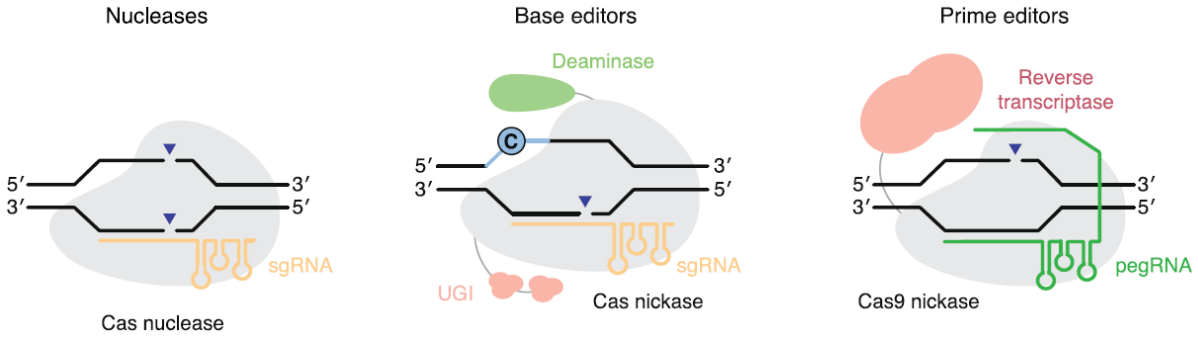


Figure 4: Overview of CRISPR-based technologies. In the original CRISPR/Cas9 approach, a sgRNA guides the Cas9 nuclease to a specific locus in the genomic DNA to introduce a DSB (left). For base editors, a nickase version of Cas9 (nCas9) introduces SSBs into the DNA, and a deaminase fused to Cas9 deaminates cytosines upstream of the cut site, which initiates their conversion to thymines (or adenine to guanine substitution in case of adenine base editors; not shown). Additional uracil DNA glycosylase inhibitors (UGI) prevent base excision repair at this site to increase the efficiency of base editing (middle). For prime editors, a reverse transcriptase is fused to nCas9 that copies the desired edit, which is encoded in the pegRNA, into the locus (right). Image modified from Anzalone et al., 2020.

With all the optimizations and developments described above, the CRISPR/Cas9 system has become an incredibly powerful tool for precise, efficient, and versatile gene editing. But it must be noted that gene editing applications do not only require the methods to be efficient, but they also need to be safe and accurate in order to generate reliable research models.

1.4 CRISPR gene editing can induce collateral damage

While genome editing using the CRISPR/Cas9 systems offers endless possibilities for various research fields, it is important to remember that this always requires the release of highly active molecular scissors into living cells to deliberately disturb the integrity of the genetic information. Optimally, this will only lead to introduction of the desired change at the target locus. However, genome editing can also be accompanied by a wide range of additional unwanted alterations, both at the target locus (on-target) and also at other genomic loci (off-target) that are often unpredictable (reviewed in Burgio & Teboul, 2020; Thomas et al., 2019). This observation is very unsettling, as specificity is of utmost importance in all genome editing applications: In experimental studies, edited cell lines or organisms need to be fully characterized to ascribe observed phenotypic changes to a specific change in the genomic sequence. For gene therapy, unnoticed genetic consequences could possibly lead to proto-oncogene activation

and hence threaten patient safety. These inadvertent genomic changes therefore need to be carefully investigated, which requires specific and dedicated detection tools.

1.4.1 Off-target effects

Soon after the discovery of the CRISPR/Cas9 system for genome editing, first reports about unwanted off-target effects were made (Fu et al., 2013; Hsu et al., 2013; Pattanayak et al., 2013). These are changes in the genomic sequence that arise from Cas9 nuclease activity at sites other than the target locus, which usually share high sequence similarity with the gRNA sequence but can contain up to 5 mismatches. Although a perfect prediction of off-target cleavage is not possible, several underlying principles were identified to predict affected sites: Off-target cleavage for one requires the presence of a PAM sequence, second, its likelihood decreases with number of mismatches and third, mismatches are less tolerated if they are more proximal to the PAM site. Furthermore, many researchers started to focus on developing new experimental methods to either minimize occurrence of off-target effects from the beginning or to develop sensitive methods for their detection after editing (reviewed in Naeem et al., 2020): first, it has become common practice to select sgRNAs for editing that have high specificity scores, i.e., with a low number and low probability of off-target cleavage sites as predicted by specific algorithms (Doench et al., 2016; Hsu et al., 2013). To facilitate editing design, such scores are now integrated into most gRNA selection tools, e.g., into CRISPOR (Concordet & Haeussler, 2018). Further strategies to improve editing specificity range from combining two Cas9 nickase enzymes that closely cut on opposite DNA strands to generate staggered DSBs (Ran et al., 2013), using truncated sgRNAs (Fu et al., 2014), or employing new engineered protein versions of Cas9 with higher fidelity, e.g., eSpCas9 (Slaymaker et al., 2016), SpCas9-HF1 (Kleinstiver et al., 2016), HiFi Cas9 (Vakulskas et al., 2018). In addition, delivering Cas9 and sgRNA as purified and pre-assembled ribonucleoprotein (RNP) complex instead of both components being encoded on plasmids, was shown to decrease the time of Cas9 within the cell, thereby leading to lower off-target activity (Kim et al., 2014; Liang et al., 2015). All of these methods, however, do not completely abolish off-target activity and therefore it is necessary to complement genome editing experiments with suitable OffTE detection methods afterwards: Sanger sequencing to genotype the top 5 to 10 off-target sites as predicted by above-mentioned algorithms can help to at least exclude the sites that are affected most likely. Alternatively, high-throughput whole-genome sequencing may also be performed. In addition, several methods have been developed to identify off-target sites experimentally in an unbiased way, with high specificity, and genome-wide. These include methods for *in vivo* detection, such as GUIDE-seq (Tsai et al., 2015) or DISCOVER-Seq (Wienert et al., 2019), as well as *in vitro*-based methods like CIRCLE-seq (Tsai et al., 2017) or CHANGE-seq (Lazzarotto et al., 2020).

Taken together, OffTEs have been at the focus of attention right after the discovery of CRISPR/Cas9 for genome editing. And although newly developed strategies have not yet led to perfect specificity, significant progress to tackle this problem has been made (Carroll, 2019; Cheng & Tsai, 2018). This situation, however, is much different for OnTEs that were described only very recently.

1.4.2 On-target effects

DNA DSBs, as induced by the Cas9 nuclease during genome editing experiments, are severe threats to the integrity of the genomic information. Cells have therefore developed diverse and robust mechanisms to repair this damage. On the one hand this is an advantage for genome editing purposes, as this leads to a wide range of different repair outcomes. On the other hand, these complex repair mechanisms can also lead to sequence changes that are usually not desired for genome editing, which are called OnTEs. These inadvertent changes can be large deletions, large insertions, complex rearrangements or also regions of copy-neutral LOH around the edited locus (Figure 5).

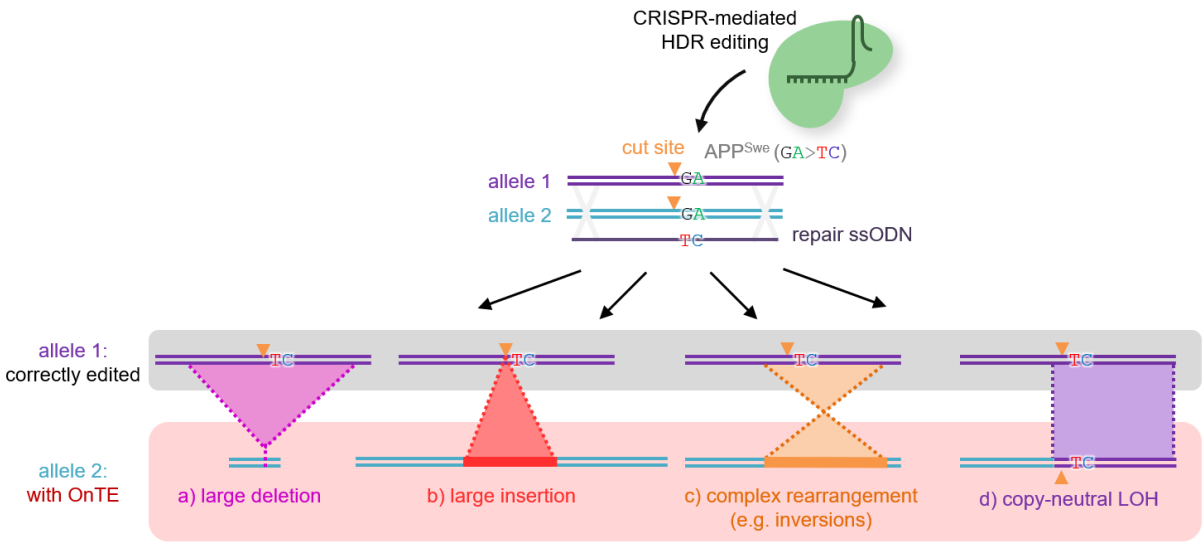


Figure 5: CRISPR-induced OnTEs. Genome editing using the CRISPR/Cas9 system can induce unwanted alterations at the target locus, like (a) large deletions, (b) large insertions, (c) complex rearrangements or (d) regions of copy-neutral LOH. Example of HDR-mediated insertion of the APPSwedish (APP^{Swe}) mutation, which leads to early-onset Alzheimer's disease in patients (Haass et al., 1995).

1.4.2.1 OnTEs are invisible in standard Sanger genotyping assays

As described above, DSBs within the genomic DNA are usually repaired by high rates of NHEJ that leads to non-specific InDels at the edited site affecting only a few bp. But soon after the discovery of the CRISPR/Cas9 system, there have also been few isolated studies reporting InDels of over several hundred bp (Cradick et al., 2013; Parikh et al., 2015; Zuckermann et al.,

2015). This phenomenon, however, was not directly investigated further, most likely because potential negative consequences were severely underestimated.

Confirming successful genome editing usually involves genotyping the locus by PCR amplification of a few hundred bp around the target site followed by Sanger sequencing. This enables detection of small changes around the target site, but large insertions or deletions on one allele might be invisible if they are not amplified in the PCR reaction (Figure 6). Edited single cell clones or organisms might therefore easily be falsely classified as homozygously edited, while they are actually hemizygous at the target locus (i.e., only one allele copy is present).

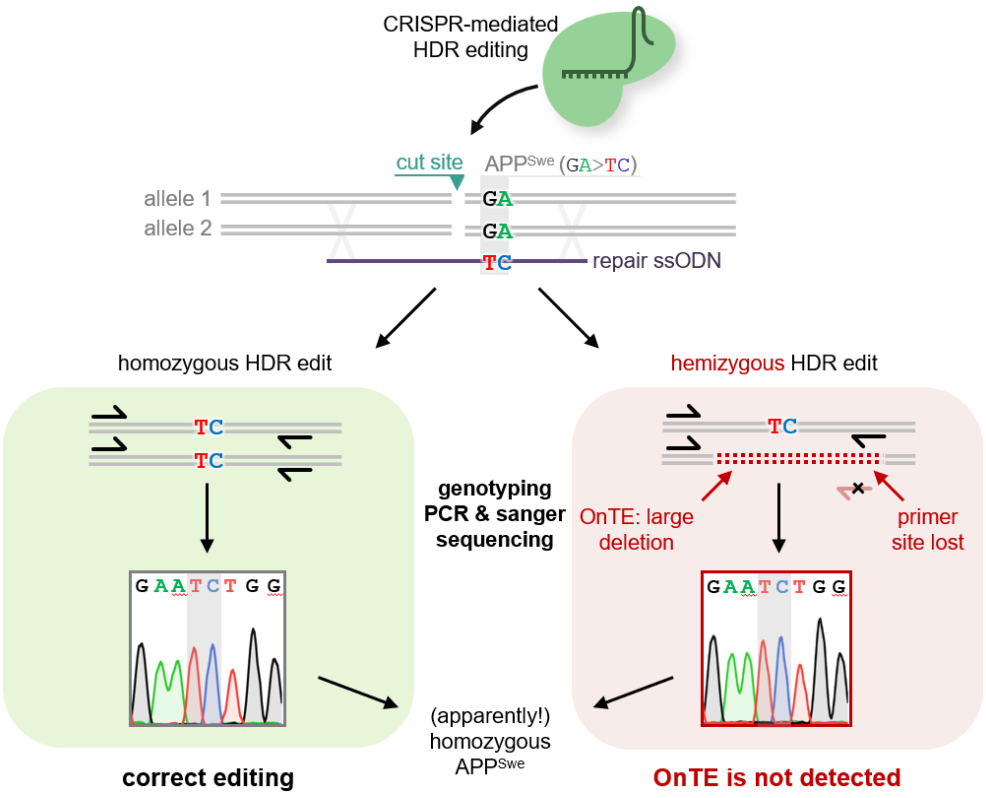


Figure 6: OnTEs are easily overlooked in conventional PCR genotyping. CRISPR genome editing can induce mono-allelic OnTEs such as large deletions, large insertions or complex rearrangements around the target site. If these overlap with one or both genotyping primer binding sites, the affected allele will not be amplified in the PCR reaction, hence, the Sanger sequencing trace only depicts the intact allele. Without the possibility of OnTEs in mind, this can easily lead to false interpretation of single cell clones as homozygously edited, although the locus is hemizygous. Example shows HDR-mediated insertion of the Alzheimer's disease associated mutation APP^{Swe}.

Another inadvertent alteration that can be introduced by CRISPR editing has been described only very recently by Ikeda et al., which is copy-neutral LOH (Ikeda et al., 2018). In these cases, DSB repair leads to replacement of areas of one homologous chromosome by the other that can potentially extend from the cut site to the chromosome end (Figure 7) but might also affect smaller areas. These changes will in most cases also be missed in Sanger genotyping,

as PCR amplicons usually do not include heterozygous variants around the target locus that could reveal the LOH.

All of these OnTEs might drastically alter gene expression of targeted or neighboring genes, which can easily lead to data misinterpretation if the unintended changes remain unnoticed.

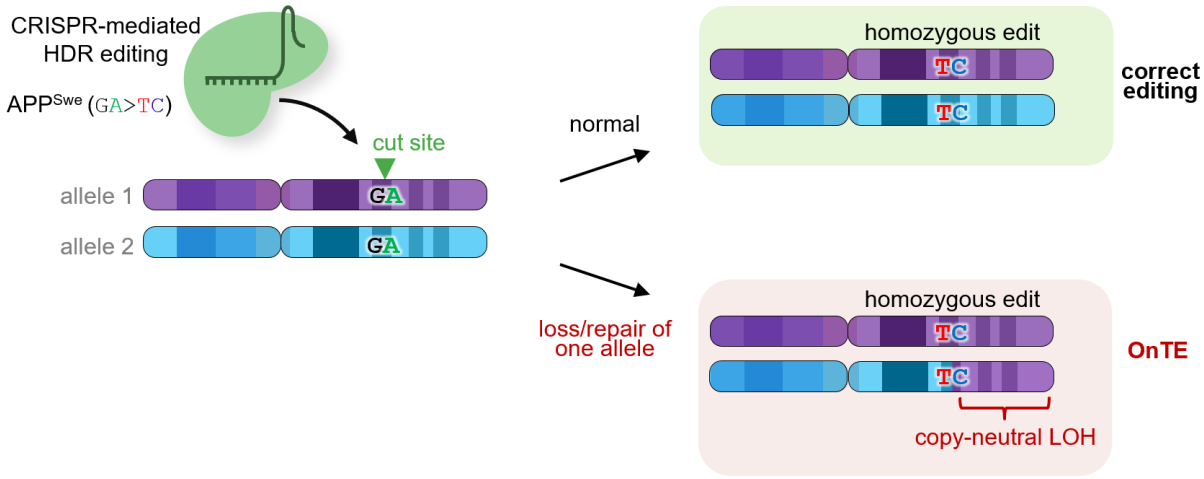


Figure 7: CRISPR genome editing may induce copy-neutral LOH. The repair of CRISPR-induced DSBs may lead to loss of chromosomal regions around the cut site that are replaced by the other homologous chromosome. This so-called copy-neutral LOH can affect small regions around the cut site (not shown) or entire chromosome arms. Example shows HDR-mediated insertion of the Alzheimer's disease associated mutation APP^{Swe}.

1.4.2.2 Open questions surrounding the occurrence of OnTEs

The problem of inadvertent alterations at the target locus after CRISPR-editing received more attention with a study from Shin et al. that reported the occurrence of large insertions and deletions at 17 sites in mouse zygotes (Shin et al., 2017). Next, a study by Kosicki et al. investigated if different cell lines are similarly affected and identified large deletions, insertions and complex rearrangements in mouse embryonic stem cells, mouse hematopoietic progenitors and an immortalized human cancer cell line (Kosicki et al., 2018). These and further studies clearly showed that OnTEs are a widespread problem after CRISPR editing in mice (Adikusuma et al., 2018; Owens et al., 2019), but the molecular mechanism leading to their occurrence remains largely unexplored (see discussion). Furthermore, it is not known if OnTEs also occur in clinically relevant human cells, such as iPSCs. Kosicki et al. identified OnTEs in human cells, but they only examined one locus and the cell line was immortalized, which is often accompanied by changes in DNA repair mechanisms. Previous studies also investigated OnTEs only after NHEJ-, but not HDR-mediated editing that has high relevance for disease-oriented research. All of these open questions highlight the need for more thorough analysis of OnTE occurrence after CRISPR editing.

1.4.2.3 Reliable OnTE detection methods are missing in the CRISPR field

To prevent undetected OnTEs from leading to data misinterpretation, thorough on-target quality control analysis after CRISPR editing is needed. However, such measures are lacking in most genome-editing studies. This is in some part due to the low level of awareness of this issue in the CRISPR field, but also due to the lack of suitable detection methods: Previous studies investigating OnTEs have used primer-walk PCRs (Adikusuma et al., 2018; Kosicki et al., 2018; Owens et al., 2019; Shin et al., 2017), deep-sequencing-based methods (Adikusuma et al., 2018; Kosicki et al., 2018; Owens et al., 2019) or droplet digital PCR (Owens et al., 2019), but these methods are laborious, expensive and/or require specialized expertise and equipment. Up to date, no specialized OnTE detection tools have been developed that allow low-cost and simple analysis of edited single cell clones or organisms after CRISPR editing.

Taken together, although undetected OnTEs can severely affect the reliability of CRISPR-based studies, this problem has been widely overlooked by the genome-editing community in recent years. Much more research is needed to address urgent questions surrounding this issue.

2. Aim of this study

The CRISPR/Cas9 system is an exceedingly powerful tool for genome editing, but its accuracy can be severely hampered by undetected OnTEs. These can be large deletions, insertions, complex rearrangements, or regions of LOH around the target site that are easily overlooked in standard quality control experiments like Sanger genotyping. OnTEs can strongly interfere with phenotype formation and thus, if being left unnoticed, affect the reliability of entire studies. Previous reports about frequent occurrence of OnTEs have created a lot of attention in the field, but they have mostly focused on mice edited by NHEJ. It is therefore still unclear if this problem also occurs in human cells subjected to HDR-mediated editing, which can be used to introduce precise changes, e.g., to study the mechanisms of disease-associated mutations. For this reason, one goal of this thesis was to investigate if also clinically relevant human iPSCs are affected. To determine how universally OnTEs occur, these iPSCs were edited at different disease-relevant loci by both, the NHEJ as well as the HDR pathway.

Identifying and excluding edited cell lines that carry OnTEs should be a crucial step after genome editing. However, on-target quality control analysis is currently not typically performed by CRISPR users, mainly due the lack of simple detection tools. Another goal of this thesis was therefore to address this problem by developing a new technology for reliable detection of OnTEs. One central focus was hereby to circumvent the necessity for specialized knowledge and equipment in the entire workflow. Furthermore, the methods should be broadly applicable to different types of edits and edited cell lines or organisms likewise. We hope that this will facilitate widespread implementation of our quality control technology in the field and improve reliability of CRISPR-based studies.

3. Research articles

3.1 Detection of deleterious on-target effects after HDR-mediated CRISPR editing

3.1.1 Summary

In this study, we address occurrence of OnTEs in human iPSCs after HDR-mediated editing. First, we compared different tools that were used in previous studies for OnTE detection like primer-walk PCR, nearby SNP genotyping and standard quantitative PCR (qPCR) but noticed that these methods did not produce reliable results, can be very laborious and/or are not feasible at any given locus. We therefore saw a great need for improved detection tools in the CRISPR field and hence developed a new method that we called quantitative genotyping PCR (qgPCR). This method quantifies a 300-450 bp genotyping PCR reaction around the locus using a fluorescently labelled probe that matches the genotyping primers. This probe binds the PCR amplicon during the reaction, thereby enabling the quantification of target locus amplification in each PCR cycle. This allows determining the number of intact alleles at the target locus to identify single cell clones with large deletions, large insertions or complex rearrangements. Next, during our analysis we noticed that there are single cell clones that are affected by another even more unknown CRISPR-induced OnTE, which is copy-neutral LOH that can affect small regions around the cut site, but also entire chromosome arms. Here, we validated nearby SNP genotyping and SNP microarrays as reliable detection methods: In nearby SNP genotyping, zygosity of potential SNPs around the cut site are determined individually by Sanger sequencing. In contrast, SNP microarrays determine zygosity and copy number of SNPs on a genome-wide scale.

Using our complete OnTE quality control workflow, we analyzed human iPSCs edited using the HDR pathway at different disease-relevant loci: We analyzed single cell clones after knock-in of the Alzheimer's disease-related mutations into the APP gene called 'Swedish' or 'Iberian'. Furthermore, we analyzed single-cell clones after editing in a non-coding region close to the HDAC9 gene at the position of a lead SNP (rs2107595) associated with stroke and coronary artery disease, which was identified in a recent genome-wide association study (GWAS) (Malik et al., 2018). Strikingly, we identified frequent occurrence of different types of OnTEs in up to 40% of edited clones independently of gRNA, locus, cell line, chromosome and coding region. Moreover, we illustrate deleterious consequences of OnTEs in an *in vitro* model of Alzheimer's disease: When we differentiated human iPSCs carrying the APP 'Swedish' mutation into neurons, we saw that cells with OnTEs produced significantly lower levels of the APP protein and Aβ peptide. This may affect pathogenic phenotypes and thus prevent reliable disease modeling.

Our results clearly indicate the widespread problem of OnTEs in human iPSCs after genome editing. To ensure locus integrity and prevent misleading results in CRISPR-based studies, we

developed a simple and reliable technology for OnTE detection that we recommend to all CRISPR users.

3.1.2 Reference

The paper was published in the Journal ‘Cell Reports’ under the following reference:

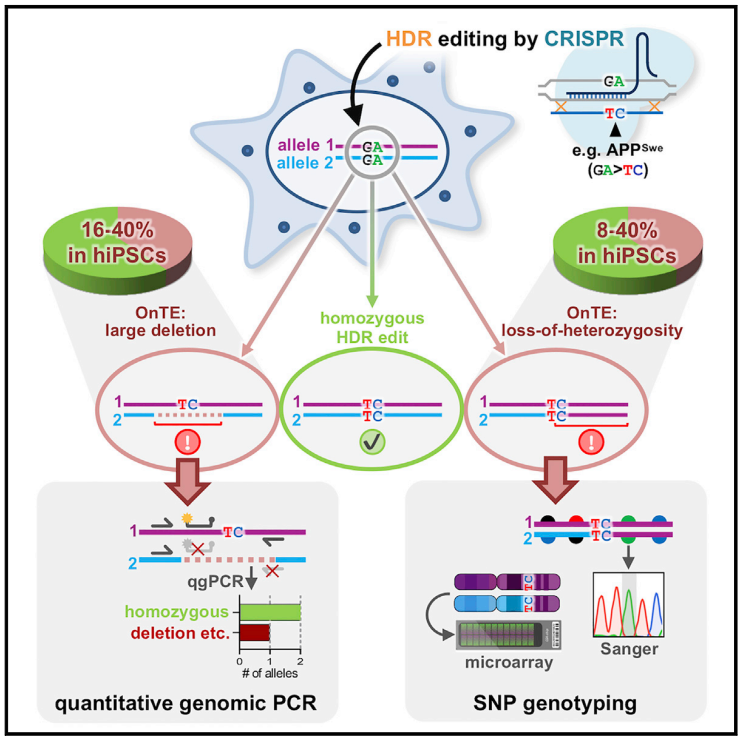
Weisheit, I., Kroeger, J. A., Malik, R., Klimmt, J., Crusius, D., Dannert, A., Dichgans, M., & Paquet, D. (2020). Detection of Deleterious On-Target Effects after HDR-Mediated CRISPR Editing. ***Cell Reports***, 31(8). https://doi.org/10.1016/j.celrep.2020.107689

Cell Reports

Report

Detection of Deleterious On-Target Effects after HDR-Mediated CRISPR Editing

Graphical Abstract



Highlights

- On-target effects (OnTEs) are present in up to 40% of human CRISPR-edited iPSC clones
- OnTEs are frequently missed by standard quality controls, such as locus sequencing
- Unnoticed OnTEs strongly affect phenotype formation in an iPSC Alzheimer model
- Simple and broadly applicable qgPCR and SNP genotyping-based tools reliably detect OnTEs

Authors

Isabel Weisheit, Joseph A. Kroeger, Rainer Malik, ..., Angelika Dannert, Martin Dichgans, Dominik Paquet

Correspondence

dominik.paquet@med.uni-muenchen.de

In Brief

Weisheit et al. show that deleterious on-target effects, such as large, mono-allelic deletions or loss of heterozygosity, are widespread in human iPSCs after CRISPR/Cas9 editing both via HDR and NHEJ. They describe simple, low-cost, and broadly applicable quantitative genomic PCR (qgPCR) and SNP genotyping-based tools to reliably identify such on-target effects.

Weisheit et al., 2020, Cell Reports *31*, 107689
May 26, 2020
https://doi.org/10.1016/j.celrep.2020.107689

CellPress

Cell Reports

CellPress
OPEN ACCESS

Report

Detection of Deleterious On-Target Effects after HDR-Mediated CRISPR Editing

Isabel Weisheit,[1,2] Joseph A. Kroeger,[1,2] Rainer Malik,[1] Julien Klimmt,[1,2] Dennis Crusius,[1] Angelika Dannert,[1,2] Martin Dichgans,[1,2,3] and Dominik Paquet[1,2,3,4,*]

[1]Institute for Stroke and Dementia Research (ISD), University Hospital, LMU Munich, 81377 Munich, Germany
[2]Graduate School of Systemic Neurosciences, LMU Munich, 82152 Planegg-Martinsried, Germany
[3]Munich Cluster for Systems Neurology (SyNergy), 81377 Munich, Germany
[4]Lead Contact
*Correspondence: dominik.paquet@med.uni-muenchen.de
https://doi.org/10.1016/j.celrep.2020.107689

SUMMARY

CRISPR genome editing is a promising tool for translational research but can cause undesired editing outcomes, both on target at the edited locus and off target at other genomic loci. Here, we investigate the occurrence of deleterious on-target effects (OnTEs) in human stem cells after insertion of disease-related mutations by homology-directed repair (HDR) and gene editing using non-homologous end joining (NHEJ). We identify large, mono-allelic genomic deletions and loss-of-heterozygosity escaping standard quality controls in up to 40% of edited clones. To reliably detect such events, we describe simple, low-cost, and broadly applicable quantitative genotyping PCR (qgPCR) and single-nucleotide polymorphism (SNP) genotyping-based tools and suggest their usage as additional quality controls after editing. This will help to ensure the integrity of edited loci and increase the reliability of CRISPR editing.

INTRODUCTION

CRISPR genome editing holds great promise for biomedical research because it allows precise and efficient genomic modifications for investigations of disease-associated variants, e.g., in disease-relevant human cell types derived from induced pluripotent stem cells (iPSCs) (Hockemeyer and Jaenisch, 2016; Paquet et al., 2016). However, application of CRISPR can be hampered by unwanted off- and on-target effects (Cheng and Tsai, 2018; Thomas et al., 2019). Recent studies in mice have described frequent occurrences of large deletions and complex rearrangements at CRISPR-edited loci after repair by non-homologous end joining (NHEJ) (Adikusuma et al., 2018; Kosicki et al., 2018; Owens et al., 2019; Shin et al., 2017). It is currently unclear whether such alterations also affect clinically relevant human cells, such as iPSCs, because repair pathways involved in CRISPR editing are differentially regulated (MacRae et al., 2015), as indicated, for example, by shorter human gene conversion tracts (Paquet et al., 2016). One report identified on-target effects (OnTEs) at a single locus in an immortalized human cell line edited using stable overexpression of Cas9 and a guide RNA (gRNA) (Kosicki et al., 2018) but did not address effects of transient expression of CRISPR machinery currently used in most editing protocols. Importantly, to our knowledge, it has not been investigated whether deleterious OnTEs also occur in cells edited by homology-directed repair (HDR) to introduce specific base changes, which has high relevance in disease research and gene and cell-replacement therapies. HDR- and NHEJ-edited clones are usually identified by PCR amplification of a few hundred bases around the edited locus, followed by Sanger sequencing (Kwart et al., 2017), but such genotyping will fail to identify clones with large, mono-allelic insertions or deletions overlapping with genotyping primer-binding sites. Instead, because the alterations prevent amplification of the affected allele, such hemizygous clones will appear to be homozygously edited (Figure 1A). Even though false identification of homozygously edited clones can corrupt the reliability of entire studies, tests for such deleterious OnTEs are still lacking in most genome-editing studies. Some reports have applied primer-walk PCR (Adikusuma et al., 2018; Kosicki et al., 2018; Owens et al., 2019; Shin et al., 2017), PacBio or another deep-sequencing method (Adikusuma et al., 2018; Kosicki et al., 2018; Owens et al., 2019), or droplet digital PCR (Owens et al., 2019) to detect large on-target alterations, but these methods are expensive, laborious, or require specific expertise and equipment. Here, we investigated whether large, mono-allelic deletions or insertions occur in human iPSCs after HDR-mediated CRISPR genome editing and developed quantitative genotyping PCR (qgPCR) as a simple and broadly applicable tool for their reliable detection. Strikingly, we identify these OnTEs in up to 40% of iPSC clones edited via HDR with CRISPR/Cas9 at different loci and demonstrate deleterious effects on phenotype formation in an Alzheimer's disease iPSC line. Extending on an earlier study (Ikeda et al., 2018), we also describe large regions of copy-neutral loss-of-heterozygosity (LOH) upon HDR-mediated editing in 8%–40% of clones and validate Sanger sequencing and microarray-based tools for LOH detection. Lastly, we investigated occurrence of large on-target deletions after NHEJ-mediated CRISPR editing using qgPCR and found the loss of one allele in 50% of apparently homozygous clones.

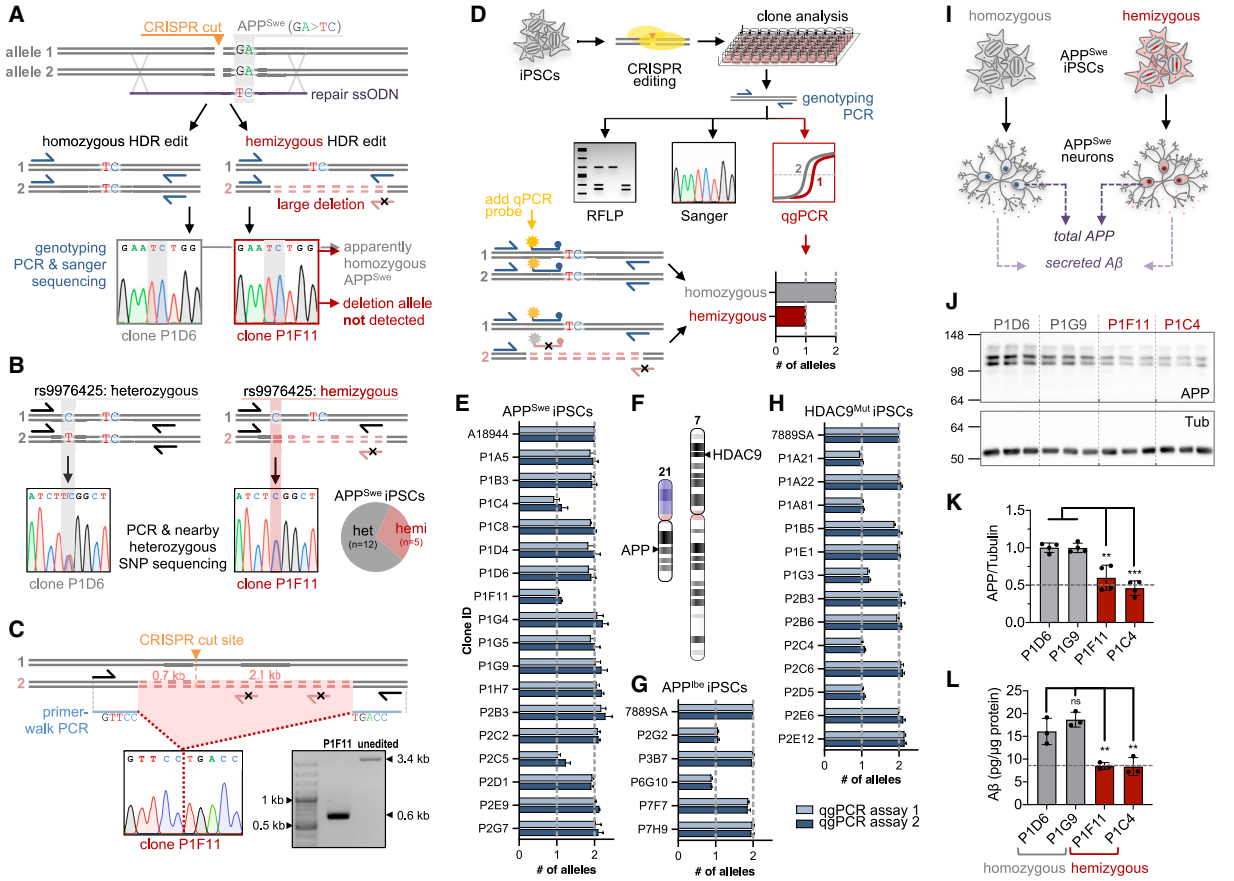


Figure 1. Deleterious OnTEs after HDR-Mediated Genome Editing in Human iPSCs

(A) Sanger genotyping fails to identify mono-allelic deletions in APP^{Swe} knockin clones.

(B) Hemizygous APP^{Swe} clones can be detected by extending genotyping PCRs to nearby heterozygous SNP rs9976425.

(C) Primer-walk PCR identified a 2.8 kb deletion in APP^{Swe} clone P1F11.

(D) Adding a qPCR probe to an existing genotyping PCR allows detection of reduced allele copy numbers by qgPCR.

(E) Allele copy numbers for two independent qgPCR assays reveal hemizygous clones with the loss of one allele after HDR knockin of APP^{Swe}. Values were normalized to the unedited parent cell line (A18944, n = 3). Data are represented as means ± SEM.

(F) Editing positions on chromosomes 21 and 7 at APP and HDAC9 loci.

(G) Identification of hemizygous clones edited at the APP^{Ibe} locus. Values were normalized to the unedited parent cell line (7889SA, n = 3). Data are represented as means ± SEM.

(H) Identification of hemizygous clones edited at the HDAC9^{Mut} locus. Values were normalized to the unedited parent cell line (7889SA, n = 3). Data are represented as means ± SEM.

(I) Two homozygous or hemizygous APP^{Swe} clones were differentiated into cortical neurons, and the levels of total APP and secreted Aβ were measured.

(J) Western blot of APP and tubulin indicates reduced APP expression in hemizygous clones.

(K) Quantification of (J) and biological replicates in Figure S2 (APP normalized to tubulin and means of homozygous clones on same gel, n = 4). Data are represented as means ± SEM. **p < 0.01, ***p < 0.001, one-way ANOVA. Dotted line indicates 50% of the means of homozygous clones.

(L) Aβ secretion (normalized to total protein amount, n = 3) is also reduced in hemizygous clones. Data are represented as means ± SEM. **p < 0.01, one-way ANOVA. Dotted line indicates 50% of the means of homozygous clones.

RESULTS

Analysis of OnTEs in HDR-Edited iPSC Clones by SNP Genotyping and PCR Primer-Walking Yields Inconsistent Results

To explore the incidence of deleterious OnTEs in CRISPR-edited iPSCs, we analyzed 17 clones with an apparently homozygous knock-in of the APP Swedish (APP^{Swe}) mutation generated using plasmid-based editing (Paquet et al., 2016; Figure 1A). APP^{Swe} causes early-onset Alzheimer's disease in patients and is used in many disease models. We reasoned that large, mono-allelic alterations could be identified by genotyping single-nucleotide polymorphisms (SNPs) near the target site that we identified to be heterozygous before editing. Large deletions or insertions in this region would prevent amplification of the aberrant allele in a PCR covering both the target and SNP site, leading to

homozygosity of the SNP in Sanger sequencing. Indeed, SNP rs9976425 appeared homozygous in five of 17 clones after editing, suggesting previously undetected mono-allelic changes (Figure 1B; Table S1). To identify possible deletions, we performed primer-walk PCRs up to 8 kb around the APP^{Swe} locus and increased the PCR extension times to detect insertions. We identified additional products in two of the five clones identified by SNP genotyping, revealing a deletion of 2.8 kb in clone P1F11 (Figure 1C) and an insertion of 4.1 kb in clone P1C4 (Figure S1; Table S1). Primer-walk PCRs were, however, not able to resolve alterations in the remaining three clones identified in the SNP assay, potentially because of PCR size limitations, illustrating the requirement for more reliable readouts. In addition, both SNP genotyping and primer-walk PCRs are not universally applicable because other loci may lack nearby heterozygous SNPs or contain regions difficult to amplify by PCR.

qgPCR Reliably Detects Widespread Occurrence of OnTEs in HDR-Edited iPSCs

An optimal assay should not only reliably identify deleterious OnTEs but also work on every edited locus, integrate well into existing gene-editing workflows, and be broadly applicable with low requirements for special knowledge and equipment. As genome-editing workflows usually contain a PCR for genotyping by RFLP and Sanger sequencing (Kwart et al., 2017), we reasoned that the simplest way of testing for mono-allelic alterations would be to determine allele copy number using the already established genotyping PCR. We addressed this by adding a labeled probe to the existing genotyping primers for quantitative genotyping PCR (qgPCR). Edited single-cell clones with large deletions or insertions will have higher cycle threshold (Ct) values, corresponding to a reduced allele copy number at the target site (Figure 1D; see design parameters in Figure S3). To test this approach, we analyzed all 17 APP^{Swe} clones by qgPCR and confirmed the results with a second, independent qgPCR assay. Compared with unedited parent cells, three clones showed copy numbers corresponding to only one allele, which all had been previously identified by SNP genotyping (Figure 1E; Table S1). Interestingly, two other clones with SNP homozygosity had normal allele numbers in both qgPCR assays, suggesting a different OnTE, such as LOH (confirmed in further analysis below). To investigate whether OnTEs occur independently of gRNA, locus, chromosome, coding region, and cell line, we repeated the analysis in a different iPSC line (7889SA; Paquet et al., 2016), edited with a different gRNA for the APP Iberian mutation (APP^{Ibe}). We also analyzed a line edited in a non-coding region near HDAC9 at rs2107595 (Figure 1F), a lead SNP identified in a recent genome-wide association study (GWAS) for stroke and coronary artery disease (Malik et al., 2018). qgPCR analysis revealed frequent loss of alleles at both loci and in both cell lines affecting two of five APP^{Ibe} and five of 13 HDAC9^{Mut} clones (Figures 1G and 1H). Again, primer-walk PCRs failed to identify all affected clones. Similar to the APP^{Swe} results, SNP genotyping revealed additional clones with SNP homozygosity but normal copy number, suggesting LOH (Table S1, see further analysis below). In agreement with previous studies (Owens et al., 2019; Shin et al., 2017), large deletions were preferentially located at sites with microhomologies, suggesting involvement of the microhomology-mediated end-joining (MMEJ) pathway (Table S2). Taken together, our data show that deleterious OnTEs, such as large deletions or insertions, occur in 18%–40% of CRISPR-edited human iPSCs and that these undesired editing events can be reliably identified by simple and universal qgPCR-based assays using already optimized genotyping PCRs.

"Standard Size" qPCR Assays Fail to Reliably Detect All OnTEs

Because our qgPCR assays had amplicon sizes of around 350 bp, we also tested assays with amplicon sizes of around 150 bp, which is set as "standard" in most qPCR primer design tools. However, these were not reliable because at least one assay for each analyzed locus failed to identify all abnormal clones (P1C4 for APP^{Swe}, P2G2 for APP^{Ibe}, and P1A21 for HDAC9; see Table S1 and Figure S3 for further details). In all these cases, the edited loci appeared to have two normal alleles, even though there were insertions or deletions present. These insertions or deletions (indels) were missed because they did not directly overlap with the cut sites, and, therefore, primers for short PCRs were still able to bind and support locus amplification. Hence, locus integrity cannot be reliably tested by "standard size" qPCRs but requires our longer qgPCR design.

OnTEs Affect Phenotype Formation in an iPSC-Based Model of Alzheimer's Disease

Most of the OnTEs we found in our HDR-edited lines caused large changes on the genomic loci, which, in many cases, could result in major changes in gene expression, unless the allelic damage is compensated by the other allele. As most HDR-mediated CRISPR editing is performed to insert or correct disease-associated mutations, defective alleles may also have unintended effects on disease modeling. To investigate potential consequences of undesired OnTEs on protein expression in a disease model, we differentiated APP^{Swe} iPSCs with and without mono-allelic alterations into cortical neurons and measured total APP levels, as well as secretion of the APP cleavage product Aβ (Figure 1I). Hemizygous APP lines displayed a reduction in APP expression and Aβ secretion by about 50% (Figures 1J–1L and S2). Such a reduction in Aβ levels may reduce pathogenic effects or even prevent formation of Alzheimer's disease phenotypes in an affected iPSC-based disease model, thus illustrating potential negative effects of undetected OnTEs on the reliability of studies using CRISPR/Cas9 editing for disease modeling.

CRISPR/Cas9 Editing in iPSCs Can Cause LOH of Entire Chromosome Arms

Our combined SNP genotyping and qgPCR analysis revealed clones with normal allelic copy number but homozygosity at nearby SNPs at all edited loci (APP^{Swe}: P1G9 and P2E9; APP^{Ibe}: P7H9; and HDAC9: P1E1; see Figures 1E, 1G, and 1H and Table S1). We reasoned that this may result from repair of a large, mono-allelic deletion by the homologous chromosome (Figure 2A). One previous report already indicated that copy-neutral LOH can occur after HDR-mediated CRISPR editing (Ikeda et al., 2018), but it is still unclear whether that is a general phenomenon or restricted to the cell line or transgene-based editing approach

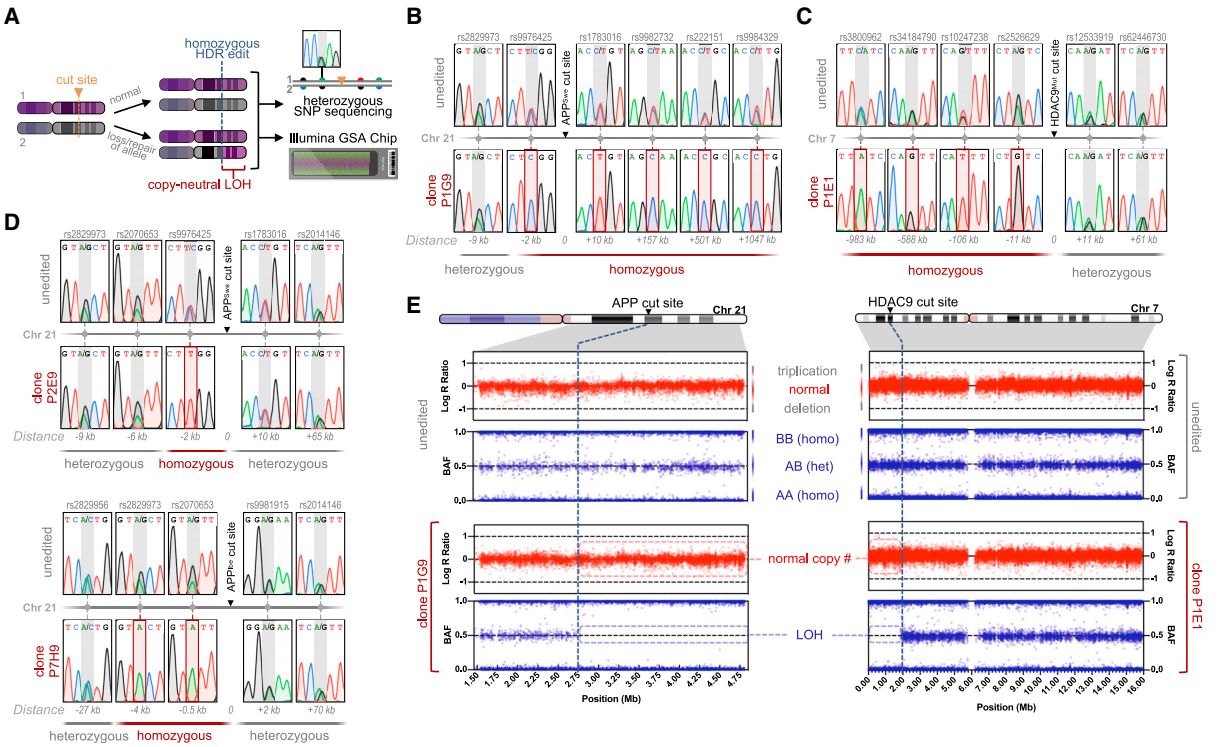


Figure 2. Detection of Copy-Neutral LOH after HDR-Mediated Genome Editing in Human iPSCs

(A) HDR editing may cause LOH, which can be detected via nearby SNP genotyping or SNP microarrays.

(B) Sanger sequencing traces of SNPs in control and edited clone P1G9 up to 1 Mb around the APP^{Swe} cut site.

(C) Sanger sequencing traces of SNPs in control and edited clone P1E1 up to 1 Mb around the HDAC9 cut site.

(D) Sanger sequencing traces of SNPs in control and edited clones around the APP^{Swe} (clone P2E9, top) or APP^{Ibe} (clone P7H9, bottom) cut sites.

(E) Log R ratio and BAF in control and edited clones for chromosome 21 (P1G9 edited for APP^{Swe}) (left) and 7 (P1E1 edited at HDAC9) (right).

described in that study. To investigate the extent of LOH in our edited iPSC lines, we identified SNPs that were heterozygous in the unedited lines on both sides of the target locus up to 1 Mb away from the cut site and analyzed their zygosity after editing. In one clone, edited at APP on chromosome 21 (P1G9), and another, edited at HDAC9 on chromosome 7 (P1E1), all tested SNPs in the direction to the end of the chromosome were homozygous (Figures 2B and 2C). Shorter regions were affected in the remaining clones (Figure 2D). To determine whether the LOH affected the entire chromosome arm in P1G9 and P1E1, we performed whole-genome SNP genotyping using the Illumina global screening array (GSA). Log R ratios showed normal copy number, but all heterozygous AB signals in B-allele frequency (BAF) were lost in the affected areas, indicating copy-neutral LOH from the cut site to the end of the targeted chromosome (Figure 2E). Taken together, our data indicate that LOH can occur after CRISPR/Cas9 editing, independent of chromosome, locus, cell line, or editing method.

OnTEs Are Also Widespread in iPSCs Edited via the NHEJ Pathway

Earlier work in mice and human cell lines indicated widespread occurrence of OnTEs after CRISPR editing via the NHEJ pathway (Adikusuma et al., 2018; Kosicki et al., 2018; Owens et al., 2019; Shin et al., 2017), but it is currently unclear whether OnTEs are also found in iPSCs, which differ in the regulation of repair pathways (MacRae et al., 2015). We, therefore, analyzed APP^{Swe} clones edited via NHEJ using the same plasmid-based CRISPR pipeline we also applied for HDR editing. We isolated clones for which loss of a restriction site overlapping the cut site indicated presence of indels and further analyzed the 28 clones in which presence of two alleles could not be shown by detection of two distinct bands in gel electrophoresis after locus PCR: 12 of these clones had differently edited alleles (i.e., double peaks in Sanger sequencing), indicating presence of two alleles, and this was confirmed by qgPCR in all cases (data not shown). Strikingly, out of the remaining 16 clones with apparently homozygous NHEJ editing (i.e., clean, single peaks in Sanger sequencing), eight had an allele copy number of only "one" in two independent qgPCR assays (Figure 3). These results were consistent with results from our nearby SNP genotyping assay: hemizygous clones identified by qgPCR were now homozygous at SNP rs9976425 (data not shown). Thus, if researchers preferentially select NHEJ clones with an apparently identical "clean" knockout on both alleles, they might have a 50% risk of using a clone with an OnTE.

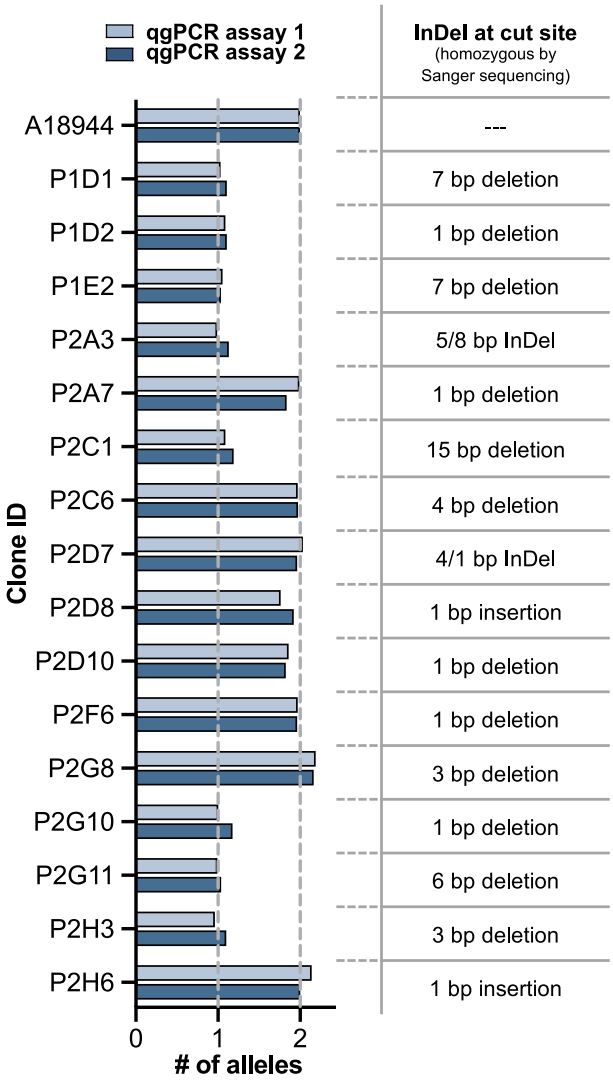


Figure 3. Widespread Formation of OnTEs after NHEJ-Mediated Genome Editing in Human iPSCs

Allele copy numbers for two independent qgPCR assays reveal hemizygous clones with the loss of one allele at the APP^{Swe} locus after NHEJ editing (left); 50% of clones with apparently homozygous editing are affected. Indel sizes as determined by Sanger sequencing (right).

DISCUSSION

The recent CRISPR revolution has provided researchers with powerful genome-editing tools that are widely applied in basic and translational research and currently also cross barriers into therapeutic applications of CRISPR-edited cells and editing directly in patients (Fellmann et al., 2017). However, CRISPR editing can cause unintended effects at the edited site and elsewhere in the genome. Although off-target effects can be efficiently detected with a variety of tools, the occurrence of OnTEs has only been described recently in mice and an immortalized human cell line. In these studies, OnTEs occurred frequently upon genome editing via the NHEJ pathway, independent of the applied CRISPR system (plasmid, RNP, and mRNA) (Adikusuma et al., 2018; Kosicki et al., 2018; Owens et al., 2019; Shin et al., 2017). However, it has been unclear whether OnTEs also occur in clinically relevant human stem cells or after editing by HDR, which is used to introduce specific base changes. We show that large, mono-allelic deletions and insertions occurred in 18%–40% of human iPSC clones after HDR-mediated CRISPR editing. These deleterious OnTEs appeared independent of the targeted locus, gRNA, coding regions, or edited cell line, suggesting widespread prevalence of on-target issues in iPSCs and also in other organisms and systems. By differentiating edited iPSCs with and without such unintended alterations into cortical neurons and comparing levels of Alzheimer's-disease-relevant Aβ secretions, we demonstrate the drastic effects unnoticed genomic alterations can have on studies using CRISPR-edited cells. Confirming and extending on earlier work in other systems, we also demonstrate the presence of OnTEs in up to 50% of iPSCs edited via the NHEJ pathway.

Furthermore, we also observed the occurrence of copy-neutral LOH after CRISPR editing, affecting entire chromosome arms. Similar LOH has also been described in human pre-implantation embryos edited by CRISPR to correct heterozygous mutations by interhomolog recombination (Ma et al., 2017). However, a major difference to our study is that the LOH allele did not simply acquire the sequence of the other allele but, in addition, contained the mutation introduced by the repair template used for HDR. This difference indicates a more complex repair scenario, in which it is not obvious that one allele acquired the sequence of the other. Such loss of SNP heterozygosity may potentially alter gene expression or expose effects of recessive mutations, which could be detrimental, especially in edited human embryos and clinical applications of iPSCs. Our findings highlight the need for technologies that reliably detect all unwanted OnTEs. Standard quality controls broadly performed in the field, such as genotyping or karyotyping, will only detect small events restricted to genotyping amplicons or very large chromosomal aberrations, such as megabase-sized deletions, translocations, and inversions, but miss the CRISPR-induced OnTEs that we and others have revealed (Adikusuma et al., 2018; Ikeda et al., 2018; O'Keefe et al., 2010; Owens et al., 2019; Shin et al., 2017).

Moreover, high-density SNP arrays sometimes used for quality controls faithfully detect only larger deletions, inversions, and LOH because their reliability increases with the number of affected SNPs. LOH affecting single SNPs may be visible, but the reliability of chip data for single SNPs is less than it is for Sanger sequencing and often depends on the detection probe, genomic location, etc. Copy-neutral inversions are usually invisible in chip assays. Many of the OnTEs that we found were small, affecting only a couple hundred to a few thousand base pairs. Accordingly, these events overlapped with no, a single, or only a few SNPs. Although such small events could be reliably detected by qgPCR (deletions, insertions, and inversions) or our Sanger sequencing-based assay (LOH), they could not be faithfully detected by the standard GWAS chip technology we used. Using higher-density chips would not solve that problem

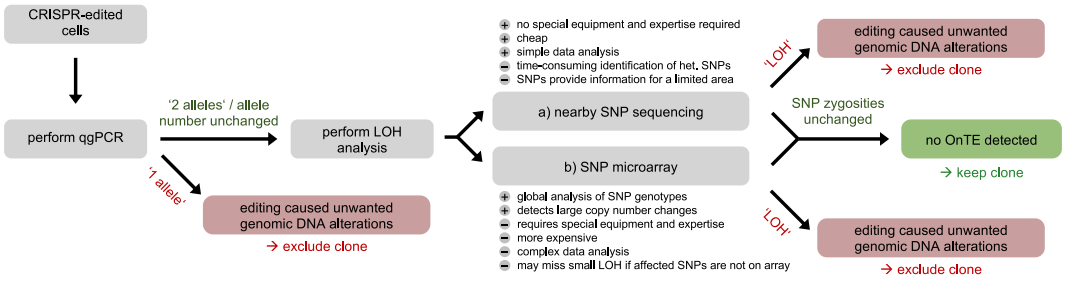


Figure 4. Workflow of Suggested Quality Control Experiments to Determine OnTEs after CRISPR Editing

Single-cell clones edited by CRISPR/Cas9 are first subjected to analysis by qgPCR to confirm unchanged allele numbers in edited clones and to exclude clones with altered allelic copy number. To check clones for loss of heterozygosity (LOH), there are two possibilities: nearby SNP sequencing and SNP microarrays. Both methods have their individual advantages, and the selection needs to be made according to the researchers' needs: Nearby SNP sequencing is cheap and does not require special equipment or expertise for analysis, whereas SNP microarrays are more expensive and involve complex data analysis. Local SNP sequencing is more sensitive toward small regions of LOH that overlap with only a few SNPs, but identifying those heterozygous SNPs on both sides of the target site can be laborious in contrast to a fast analysis by microarrays. Furthermore, SNP microarrays analyze SNP genotypes genome-wide and enable characterizing the dimension of large regions of LOH, whereas nearby SNP genotyping is restricted to a few loci around the edited site. Taken together, a combination of qgPCR analysis, and nearby SNP genotyping, and/or clone analysis by SNP microarrays should be conducted after CRISPR/Cas9 genome editing to ensure the integrity of the edited loci.

because the detection is not limited by the overall number of measured SNPs on the chip but by the number of measurable affected SNPs around the edited locus. We, therefore, developed and validated assays based on qgPCR, Sanger sequencing, and microarrays, which in combination allow reliable detection of OnTEs in iPSCs and other systems. We selected these techniques because of their simplicity, low cost, easy integration into existing workflows, universal applicability for HDR- and NHEJ-mediated CRISPR editing in various systems, and feasibility for non-specialist laboratories to allow broad dissemination and acceptance in the field. We suggest using both qgPCR and nearby or global SNP genotyping as additional quality-control measures to increase the reliability of CRISPR editing (see detailed workflow in Figure 4) in iPSCs and other systems.

In this study, we focused on developing reliable assays for OnTE detection to meet the urgent need of the CRISPR field for thorough quality-control measures of edited cells and animals. However, future work should be aimed at not only detecting these OnTEs but also understanding their biological roots and reasons for occurrence, leading to strategies to avoid their formation in the first place. This could be addressed by studying (1) the locus-dependent influences, such as chromatin structure; (2) the effects of genome-editing reagents, e.g., by using Cas9 nickase or another nuclease; (3) the effects of repair templates by modulating the single-stranded oligodeoxynucleotides (ssODN) design and orientation; and (4) the influences of other repair pathways, e.g., by modulating the NHEJ or MMEJ pathways using knockdowns or specific inhibitors.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCE TABLE
- RESOURCE AVAILABILITY
  - Lead Contact
  - Materials Availability
  - Data and Code Availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - iPSC lines
  - CRISPR/Cas9 genome editing
  - iPSC culture, electroporation and cortical differentiation
- METHOD DETAILS
  - Genotyping assay design and copy number analysis by quantitative genotyping PCR (qgPCR)
  - GSA Illumina Chip
  - Genomic variant identification
  - Primer-walk PCR
  - Measurements of total APP and Amyloid-β
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.celrep.2020.107689.

ACKNOWLEDGMENTS

This work was supported by grants from the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy within the framework of the Munich Cluster for Systems Neurology (EXC 2145 SyNergy; grant no. 390857198), Vascular Dementia Research Foundation, VERUM Foundation, Wilhelm-Vaillant-Foundation, and the donors of the ADR AD2019604S, a program of the BrightFocus Foundation (to D.P.). We thank Peter Lichtner for GSA chip analysis, Johannes Trambauer and Harald Steiner for help with the APP western blots, Brigitte Nuscher for help with the Aβ MSD, and Benedikt Wefers for helpful comments.

AUTHOR CONTRIBUTIONS

Conceptualization, I.W. and D.P.; Methodology, I.W., J.A.K., R.M., and D.P.; Investigation, I.W., J.A.K., R.M., J.K., D.C., and A.D.; Writing – Original Draft,

I.W., J.A.K., R.M., and D.P.; Writing – Review & Editing, I.W., J.A.K., R.M., J.K., A.D., M.D., and D.P.; Funding Acquisition, M.D. and D.P.; Supervision, M.D. and D.P.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: November 27, 2019
Revised: April 22, 2020
Accepted: May 4, 2020
Published: May 26, 2020

REFERENCES

Adikusuma, F., Piltz, S., Corbett, M.A., Turvey, M., McColl, S.R., Helbig, K.J., Beard, M.R., Hughes, J., Pomerantz, R.T., and Thomas, P.Q. (2018). Large deletions induced by Cas9 cleavage. Nature *560*, E8–E9.

Cheng, Y., and Tsai, S.Q. (2018). Illuminating the genome-wide activity of genome editors for safe and effective therapeutics. Genome Biol. *19*, 226.

Fellmann, C., Gowen, B.G., Lin, P.-C., Doudna, J.A., and Corn, J.E. (2017). Cornerstones of CRISPR-Cas in drug discovery and therapy. Nat. Rev. Drug Discov. *16*, 89–100.

Guo, Y., He, J., Zhao, S., Wu, H., Zhong, X., Sheng, Q., Samuels, D.C., Shyr, Y., and Long, J. (2014). Illumina human exome genotyping array clustering and quality control. Nat. Protoc. *9*, 2643–2662.

Hockemeyer, D., and Jaenisch, R. (2016). Induced pluripotent stem cells meet genome editing. Cell Stem Cell *18*, 573–586.

Ikeda, K., Uchida, N., Nishimura, T., White, J., Martin, R.M., Nakauchi, H., Sebastiano, V., Weinberg, K.I., and Porteus, M.H. (2018). Efficient scarless genome editing in human pluripotent stem cells. Nat. Methods *15*, 1045–1047.

Karlsson, J.O., Ostwald, K., Kåbjörn, C., and Andersson, M. (1994). A method for protein assay in Laemmli buffer. Anal. Biochem. *219*, 144–146.

Kosicki, M., Tomberg, K., and Bradley, A. (2018). Repair of double-strand breaks induced by CRISPR-Cas9 leads to large deletions and complex rearrangements. Nat. Biotechnol. *36*, 765–771.

Kwart, D., Paquet, D., Teo, S., and Tessier-Lavigne, M. (2017). Precise and efficient scarless genome editing in stem cells using CORRECT. Nat. Protoc. *12*, 329–354.

Ma, H., Marti-Gutierrez, N., Park, S.-W., Wu, J., Lee, Y., Suzuki, K., Koski, A., Ji, D., Hayama, T., Ahmed, R., et al. (2017). Correction of a pathogenic gene mutation in human embryos. Nature *548*, 413–419.

MacRae, S.L., Croken, M.M., Calder, R.B., Aliper, A., Milholland, B., White, R.R., Zhavoronkov, A., Gladyshev, V.N., Seluanov, A., Gorbunova, V., et al. (2015). DNA repair in species with extreme lifespan differences. Aging (Albany NY) *7*, 1171–1184.

Malik, R., Chauhan, G., Traylor, M., Sargurupremraj, M., Okada, Y., Mishra, A., Rutten-Jacobs, L., Giese, A.-K., van der Laan, S.W., Gretarsdottir, S., et al.; AFGen Consortium; Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) Consortium; International Genomics of Blood Pressure (iGEN-BP) Consortium; INVENT Consortium; STARNET; BioBank Japan Cooperative Hospital Group; COMPASS Consortium; EPIC-CVD Consortium; EPIC-InterAct Consortium; International Stroke Genetics Consortium (ISGC); METASTROKE Consortium; Neurology Working Group of the CHARGE Consortium; NINDS Stroke Genetics Network (SiGN); UK Young Lacunar DNA Study; MEGASTROKE Consortium (2018). Multiancestry genome-wide association study of 520,000 subjects identifies 32 loci associated with stroke and stroke subtypes. Nat. Genet. *50*, 524–537.

O'Keefe, C., McDevitt, M.A., and Maciejewski, J.P. (2010). Copy neutral loss of heterozygosity: a novel chromosomal lesion in myeloid malignancies. Blood *115*, 2731–2739.

Owens, D.D.G., Caulder, A., Frontera, V., Harman, J.R., Allan, A.J., Bucakci, A., Greder, L., Codner, G.F., Hublitz, P., McHugh, P.J., et al. (2019). Microhomologies are prevalent at Cas9-induced larger deletions. Nucleic Acids Res. *47*, 7402–7417.

Paquet, D., Kwart, D., Chen, A., Sproul, A., Jacob, S., Teo, S., Olsen, K.M., Gregg, A., Noggle, S., and Tessier-Lavigne, M. (2016). Efficient introduction of specific homozygous and heterozygous mutations using CRISPR/Cas9. Nature *533*, 125–129.

Richardson, C.D., Ray, G.J., DeWitt, M.A., Curie, G.L., and Corn, J.E. (2016). Enhancing homology-directed genome editing by catalytically active and inactive CRISPR-Cas9 using asymmetric donor DNA. Nat. Biotechnol. *34*, 339–344.

Shin, H.Y., Wang, C., Lee, H.K., Yoo, K.H., Zeng, X., Kuhns, T., Yang, C.M., Mohr, T., Liu, C., and Hennighausen, L. (2017). CRISPR/Cas9 targeting events cause complex deletions and insertions at 17 sites in the mouse genome. Nat. Commun. *8*, 15464.

Steyer, B., Bu, Q., Cory, E., Jiang, K., Duong, S., Sinha, D., Steltzer, S., Gamm, D., Chang, Q., and Saha, K. (2018). Scarless genome editing of human pluripotent stem cells via transient puromycin selection. Stem Cell Reports *10*, 642–654.

Thomas, M., Burgio, G., Adams, D.J., and Iyer, V. (2019). Collateral damage and CRISPR genome editing. PLoS Genet. *15*, e1007994–e1007998.

STAR★METHODS

KEY RESOURCE TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
APP-Y188	Abcam	ab32136; RRID:AB_2289606
Tubulin	Sigma	T5168; RRID:AB_477579
Anti-Rabbit IgG (H+L), HRP Conjugate	Promega	W4011, RRID:AB_430833
Anti-Mouse IgG (H+L), HRP Conjugate	Promega	W4021, RRID:AB_430834
Chemicals, Peptides, and Recombinant Proteins		
TfiI	NEB	R0546S
DdeI	NEB	R0175S
XmnI	NEB	R0194S
2x PrimeTime Gene Expression Master Mix	IDT	1055772
20x human TERT TaqMan Copy Number Reference Assay	ThermoFisher	4403316
OneTaq 2x Master Mix	NEB	M0486L
GeneRuler 100 bp Plus DNA ladder	ThermoFisher	SM0321
Critical Commercial Assays		
NucleoSpin Tissue Kit	Macherey-Nagel	740952
NucleoSpin Gel and PCR Clean-Up Kit	Macherey-Nagel	740609
TOPO TA Cloning Kit for Sequencing	ThermoFisher	450030
NucleoSpin Plasmid Kit	Macherey Nagel	740588
NucleoSpin RNA/Protein Kit	Macherey-Nagel	740933
MSD Human (6E10) Aβ V-PLEX Kit	Meso Scale Discovery	K15200E
Experimental Models: Cell Lines		
7889SA	Paquet et al., 2016, NYSCF	7889SA
A18944	ThermoFisher	A18945
Oligonucleotides		
sgRNAs	This paper	Table S3
ssODNs for HDR-mediated editing	This paper	Table S3
Primers	This paper	Table S3
Recombinant DNA		
MLM3636	a gift from K. Joung, Addgene	43860
pSpCas9(BB)-2A-Puro (PX459) V2.0	a gift from F. Zhang, Addgene	62988
pCas9_GFP	a gift from K. Musunuru, Addgene	44719
Software and Algorithms		
CRISPOR design tool	Tefor	http://crispor.tefor.net/
PrimerQuest design tool	IDT	N/A
PLINK	N/A	https://www.cog-genomics.org/plink/2.0/
Genome Studio 2.0	Illumina	https://emea.illumina.com/techniques/microarrays/array-data-analysis-experimental-design/genomestudio.html
Ensembl Biomart tool	Ensembl	https://www.ensembl.org/info/data/biomart/index.html
Primer3Plus	Primer3Plus	https://primer3plus.com
GraphPad Prism 8	GraphPad	N/A
Other		
Illumina Global Screening Array v2 genotyping chip	Illumina	20030770

RESOURCE AVAILABILITY

Lead Contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Dominik Paquet (dominik.paquet@med.uni-muenchen.de).

Materials Availability

All unique/stable reagents generated in this study are available from the Lead Contact with a completed Materials Transfer Agreement.

Data and Code Availability

All uncropped gels, raw qPCR data, APP and Abeta quantifications, Sanger sequencing reads and Illumina GSA chip data are available in Mendeley Data (https://dx.doi.org/10.17632/87kh5vj429.2).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

iPSC lines

iPSC experiments were performed in accordance with all relevant guidelines and regulations. Work with male line 7889SA (Paquet et al., 2016) (NYSCF) was approved by the Rockefeller University Institutional Review Board after informed consent was obtained from subjects by Coriell Institute. Female iPSC line A18944 was purchased from ThermoFisher (A18945).

CRISPR/Cas9 genome editing

Single guide RNAs (sgRNAs) were designed using the CRISPR design tool (http://crispor.tefor.net/). sgRNA sequences were cloned into the BsmBI restriction site of plasmid MLM3636 (a gift from K. Joung, Addgene 43860). CRISPR editing was performed as described previously (Paquet et al., 2016) using Cas9 plasmids pSpCas9(BB)-2A-Puro (PX459) V2.0 (a gift from F. Zhang, Addgene 62988) or pCas9_GFP (a gift from K. Musunuru, Addgene 44719). Repair oligos were either symmetric 100 bp ssODNs with the same orientation as the gRNA sequence (APP^{Swe}, (Paquet et al., 2016)) or asymmetric 107 bp ssODNs (71 and 36 bp, long arm on the PAM-proximal side (Richardson et al., 2016)) with sequence complementary to the gRNA (APP^{Ibe} and HDAC9), and ordered as Ultramers from IDT.

iPSC culture, electroporation and cortical differentiation

iPSCs were maintained on Vitronectin-coated (ThermoFisher A14700) cell culture plates and grown in Essential 8 Flex Medium (ThermoFisher A2858501) at 37°C with 5% CO2. Prior to transfection, iPSCs were transferred to Geltrex-coated (ThermoFisher A1413302) cell culture plates and grown in StemFlex Medium (ThermoFisher A3349401) containing 10 μM ROCK inhibitor (Selleckchem S1049) for two days. iPS cells were transfected by electroporation as described (Kwart et al., 2017). Briefly, two million cells were resuspended in 100 μL cold BTXpress electroporation solution (VWR 732-1285) with 20 μg Cas9, 5 μg sgRNA plasmid, and 30 μg ssODN. Cells were electroporated with 2 pulses at 65 mV for 20 ms in a 1 mm cuvette (Fisher Scientific 15437270). After electroporation, cells were transferred to Geltrex-coated 10 cm plates and grown in StemFlex Medium containing 10 μM ROCK inhibitor. Cells expressing Cas9 were selected either by sorting for GFP (Kwart et al., 2017) or selection with 350 ng/ml Puromycin dihydrochloride (VWR J593) for three consecutive days starting one day after electroporation (Steyer et al., 2018). Single-cell clone colonies were picked and analyzed by RFLP assay, using NEB enzymes TfiI for APP^{Swe}, DdeI for APP^{Ibe}, XmnI for HDAC9, and Sanger sequencing as previously described (Kwart et al., 2017). Cortical neuron differentiation was performed using a dual-SMAD inhibition-based protocol as described (Paquet et al., 2016).

METHOD DETAILS

Genotyping assay design and copy number analysis by quantitative genotyping PCR (qgPCR)

Assays for qgPCR analysis of edited single-cell clones were designed using the IDT PrimerQuest design tool. Briefly, a 400-550 bp region surrounding the edited locus was entered and the amplicon size range set to 300-450 bp. The edited site was selected as excluded region for the probe to prevent overlap. If genotyping primers were available, the primer sequences were entered under partial design input. Assays in which the probe was close to the edited site were favored. For copy number analysis, genomic DNA (gDNA) was isolated with a NucleoSpin Tissue Kit (Macherey-Nagel 740952) according to manufacturer's instructions and 60 ng were used for analysis. As we occasionally observed variation in gDNA integrities from stored gDNA samples we recommend using fresh gDNA isolated at the same time from control and assayed clones. Freshly isolated gDNA was mixed with 2x PrimeTime Gene Expression Master Mix (IDT 1055772), 20x human TERT TaqMan Copy Number Reference Assay (ThermoFisher 4403316) as internal reference control, genotyping primers (0.5 pmol/μl) and the designed PrimeTime Eco Probe 5′ 6-FAM/ZEN/3′ IBFQ (0.25 pmol/μl, HPLC-purified, IDT). The qgPCR reaction was run for 2 min at 50°C, 10 min at 95°C, followed by 40 cycles of 15 s at 95°C and 1 min at 60°C. Allele copy numbers were determined by ddCt calculation relative to internal TERT reference and unedited

control; values were multiplied by two to get total number of alleles. qgPCR experiments were performed in three independent technical replicates.

GSA Illumina Chip

gDNA from all iPSC lines to be analyzed was isolated with a NucleoSpin Tissue Kit and diluted to a concentration of 75 ng/μl. Whole-genome genotyping was performed at the Helmholtz Zentrum München Genome Analysis Center (Neuherberg, Germany) using the Illumina Global Screening Array v2 genotyping chip (Illumina, San Diego, California, USA). Single nucleotide polymorphisms (SNPs) were called using the GenCall algorithm. All samples analyzed showed a sample call rate > 0.99. Gender checks were performed as an additional quality control step using PLINK2 (https://www.cog-genomics.org/plink/2.0/). SNPs with a call rate < 0.9 were discarded. All SNPs were filtered using a Hardy-Weinberg equilibrium p-value cutoff of 1E-4 and a GenTrain score cutoff of 0.7 to ensure correct clustering (Guo et al., 2014). Log R Ratio and B Allele Frequency were extracted using Genome Studio 2.0 (Illumina, San Diego, California, USA).

Genomic variant identification

Potential genomic variants within 5 kb around the edited loci were identified using the Ensembl Biomart tool (http://www.ensembl.org//useast.ensembl.org/info/data/biomart/index.html?redirectsrc=//www.ensembl.org%2Finfo%2Fdata%2Fbiomart%2Findex.html) with the following settings and filters: Ensembl variation 98 database, human Short Variants (SNPs and InDels excluding flagged variants), respective chromosome with a region of around 5kb around the edited site, global minor allele frequency > = 0.2. The flanking sequence around the retrieved variants was downloaded from Ensembl and used in Primer3Plus (https://primer3plus.com/) to design primers for SNP genotyping. Prior to Sanger sequencing, the amplicons were analyzed for size differences by agarose gel electrophoresis to check for length polymorphisms. Heterozygosity of SNPs was confirmed by identification of double peaks in Sanger sequencing in unedited versus edited iPSC. Heterozygous SNPs in a 1 Mb region around the edited loci were identified by parsing data from a previous molecular karyotyping experiment performed in unedited parent lines using the Illumina bead array HumanOmni2.5Exome-8 BeadChip v1.3 (Life & Brain GmbH, Bonn) (data not shown).

Primer-walk PCR

Primer-walk PCRs were performed with edited single-cell clones to identify aberrant PCR products with OneTaq 2x Master Mix (NEB M0486L) following manufacturer's instructions. Primers with increasing distance to the cut site in steps of around 500 bp were tested. PCR products were analyzed by agarose gel electrophoresis with a GeneRuler 100 bp Plus DNA ladder (ThermoFisher SM0321). If additional bands, not present in the unedited control cell line, were detected, PCR products were gel-purified using a NucleoSpin Gel and PCR Clean-Up Kit (Macherey-Nagel 740609) followed by Sanger sequencing. If sequencing was not successful, PCR products were TOPO cloned following manufacturer's instructions (TOPO TA Cloning Kit for Sequencing, ThermoFisher 450030). Plasmids with TOPO-cloned inserts were isolated using the NucleoSpin Plasmid kit (Macherey Nagel 740588) and Sanger sequenced.

Measurements of total APP and Amyloid-β

Total protein was extracted from differentiated neurons at DIV 35 with the NucleoSpin RNA/Protein Kit (Macherey-Nagel 740933) according to manufacturer's instructions, separated on 8% TRIS-Glycine hand-casted gels, transferred to nitrocellulose membranes (Amersham Protran 0.45 NC, GE Healthcare), boiled for 5 min in PBS, and blocked for 1 h using 0,2% I-Block (ThermoFisher T2015) with 0,1% Tween20 (Merck) in PBS. Primary antibodies (APP-Y188, Abcam ab32136, 1:4,000; Tubulin, Sigma T5168, 1:4000) were diluted in blocking solution and incubated with the membrane overnight at 4°C. After three washes in PBS + 1% Tween20, HRP-labeled secondary antibodies (Anti-Rabbit IgG (H+L), HRP Conjugate, Promega, W4011; Anti-Mouse IgG (H+L), HRP Conjugate, Promega, W4021) were added for 1h and protein signals were detected using Pierce ECL Western Blotting Substrate kit (ThermoFisher 32109), using a Fujifilm LAS4000 luminescence imager and band intensities quantified using ImageJ. For Aβ measurements, cell supernatant was conditioned for 5 days and experiments were performed in 3 biological replicates. Supernatants from experiments collected at different time points were flash-frozen in liquid nitrogen and stored at −80°C. Secreted Aβ1−38, Aβ1−40 and Aβ1−42 were measured with MSD Human (6E10) Aβ V-PLEX kits (Meso Scale Discovery) according to the manufacturer's directions. Aβ values were combined to obtain total Aβ and normalized to total protein levels from cell lysate determined by the Karlsson et al. (1994) method, as described in the NucleoSpin RNA/Protein Kit.

QUANTIFICATION AND STATISTICAL ANALYSIS

No statistical methods were used to predetermine sample size and the experiments were not randomized. Experimental data was analyzed for significance using GraphPad Prism 8. Multiplicity-adjusted $p < 0.05$ was considered statistically significant. Significance was analyzed by one-way ANOVA comparing the mean of each column with the mean of the control followed by multiple-comparison post-testing with Dunnett's method. The analysis approaches have been justified as appropriate by previous biological studies, and all data met the criteria of the tests. The investigators were not blinded to allocation during experiments and outcome assessment.

Cell Reports, Volume *31*

Supplemental Information

Detection of Deleterious On-Target Effects after HDR-Mediated CRISPR Editing

Isabel Weisheit, Joseph A. Kroeger, Rainer Malik, Julien Klimmt, Dennis Crusius, Angelika Dannert, Martin Dichgans, and Dominik Paquet

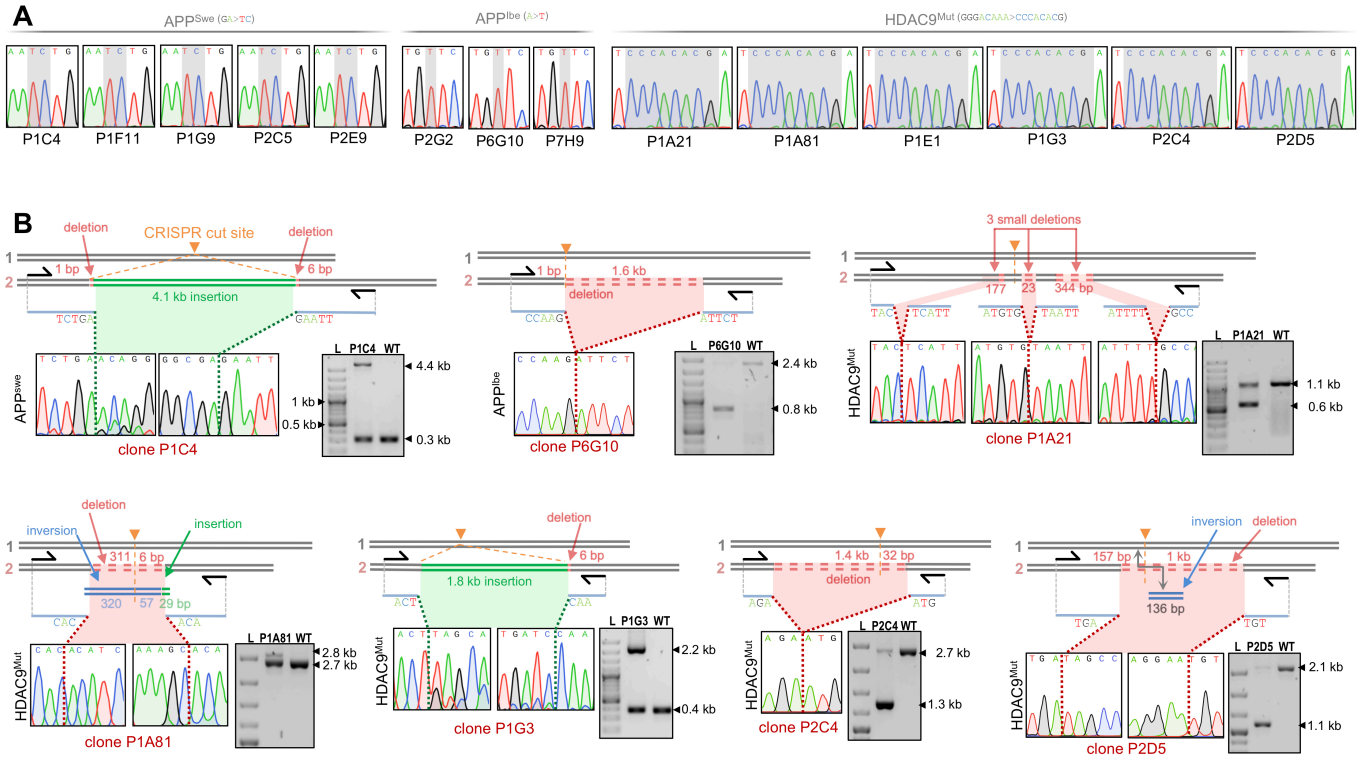


Figure S1. Genotyping results and primer-walk PCRs in CRISPR-edited clones, related to Figure 1.
(**A**) Sanger genotyping suggests homozygous HDR editing of APP^{Swe}, APP^{Ibe} and HDAC9^{Mut} in single-cell clones later identified to have mono-allelic deletions, insertions or LOH (see also Supplementary Table 1).
(**B**) Primer-walk PCR identified alleles with insertions, deletions or inversions for APP^{Swe} (top left), APP^{Ibe} (top middle), and HDAC9^{Mut} clones (remaining 5 clones). L: DNA ladder, WT: wildtype. The wildtype allele is hardly visible in some clones, potentially due to preferential amplification of the shorter PCR product.

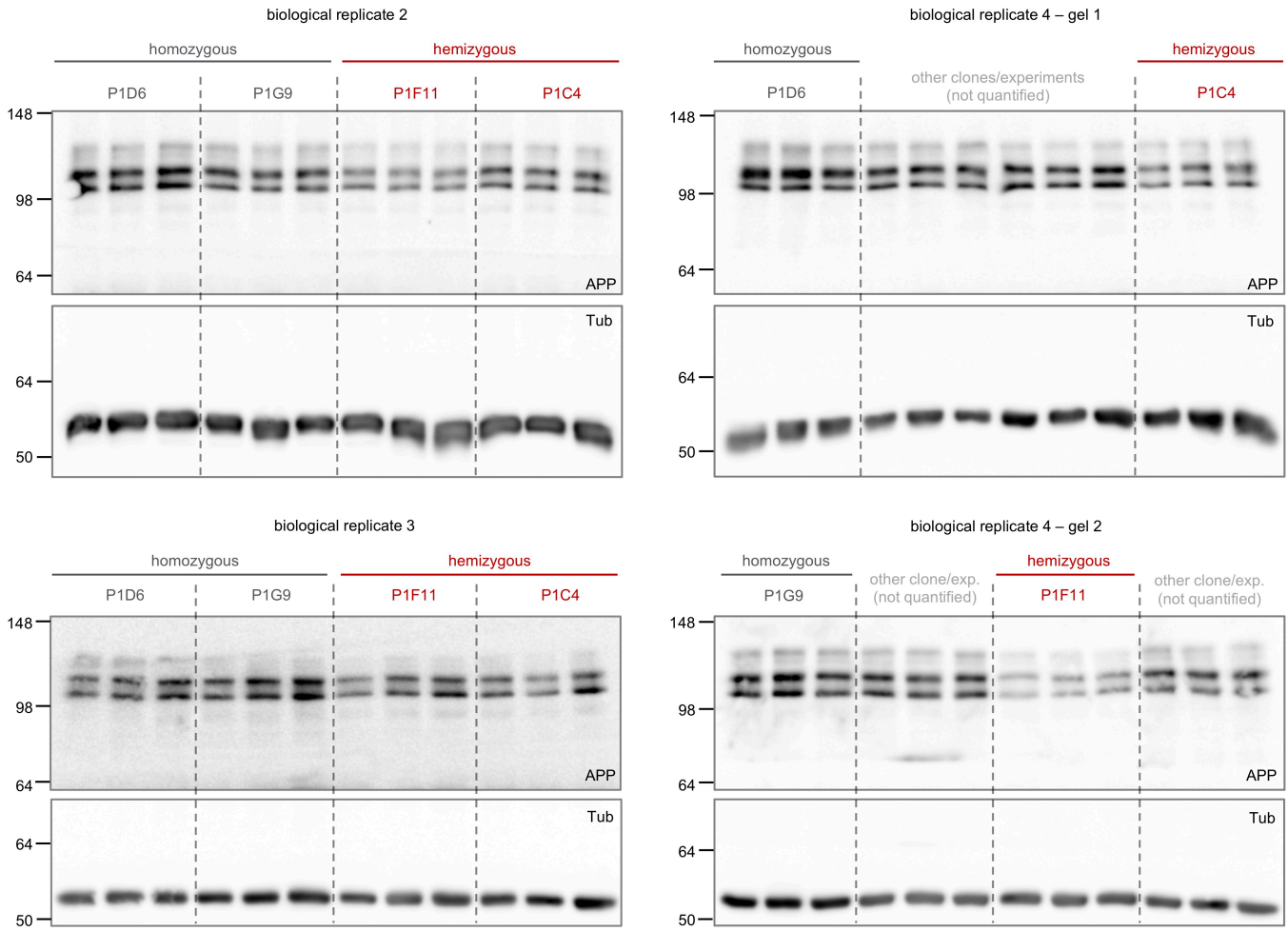


Figure S2. Quantification of APP expression in CRISPR-edited clones, related to Figure 1.
Western blot images of biological replicates 2-4 with APP and β-Tubulin expression used for quantification shown in Figure 1K. Loading controls (β-Tubulin) were run on the same blot as APP and quantitative comparisons were only performed between samples on the same blot.

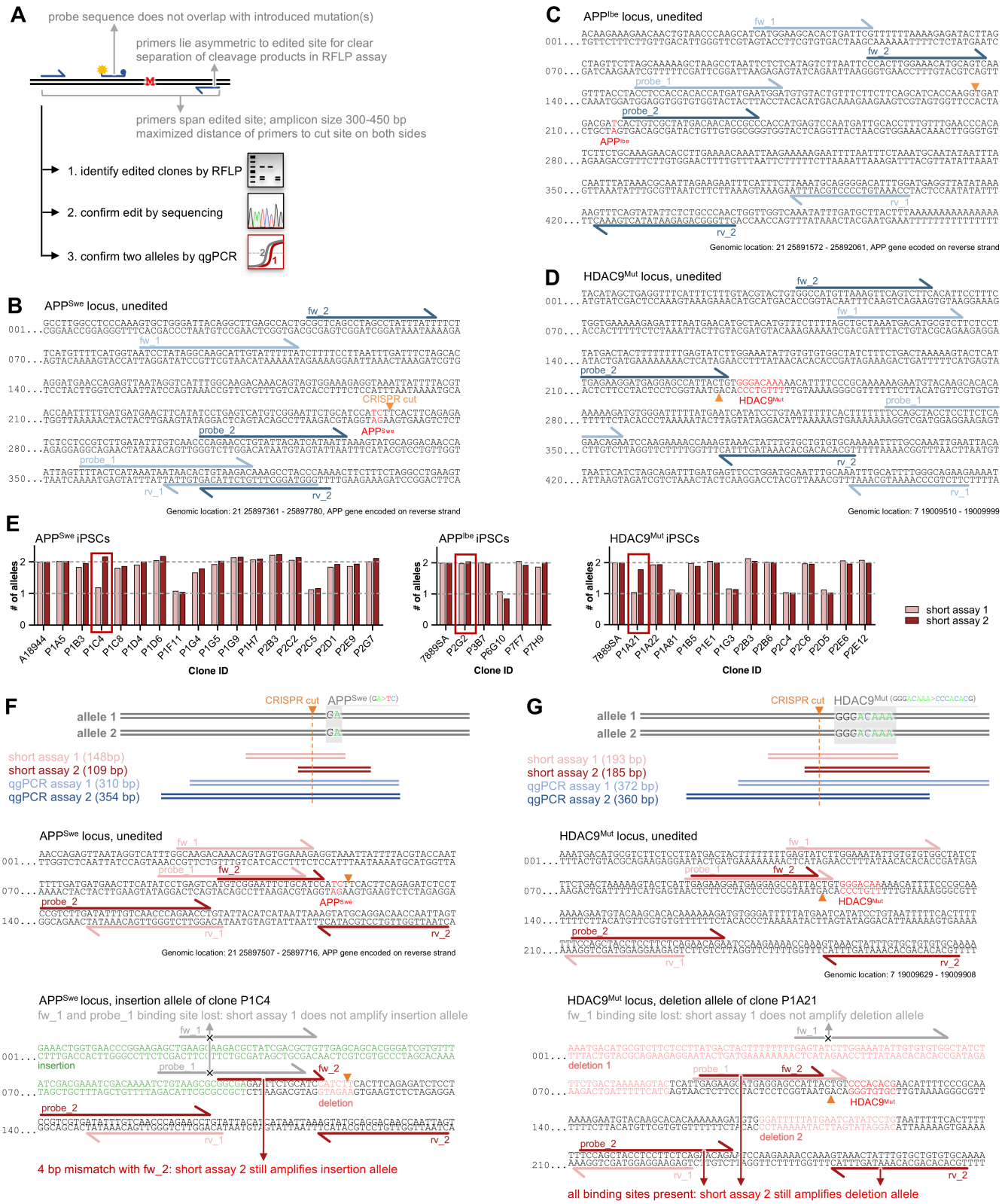

A
probe sequence does not overlap with introduced mutation(s)
primers lie asymmetric to edited site for clear separation of cleavage products in RFLP assay
M
primers span edited site; amplicon size 300-450 bp
maximized distance of primers to cut site on both sides
1. identify edited clones by RFLP
2. confirm edit by sequencing
3. confirm two alleles by qgPCR
B
APPSwe locus, unedited
fw_2
fw_1
CRISPR cut
APPSwe
probe_2
probe_1
rv_1
rv_2
Genomic location: 21 25897361 - 25897780, APP gene encoded on reverse strand
C
APPIbe locus, unedited
fw_1
fw_2
probe_1
probe_2
APPIbe
rv_1
rv_2
Genomic location: 21 25891572 - 25892061, APP gene ecoded on reverse strand
D
HDAC9Mut locus, unedited
fw_2
fw_1
probe_2
HDAC9Mut
probe_1
rv_2
rv_1
Genomic location: 7 19009510 - 19009999
E
APPSwe iPSCs
APPIbe iPSCs
HDAC9Mut iPSCs
# of alleles
A18944 P1A5 P1B3 P1C4 P1C8 P1D4 P1D6 P1F11 P1G4 P1G5 P1G9 P1H7 P2B3 P2C2 P2C5 P2D1 P2E9 P2G7
7889SA P2G2 P3B7 P6G10 P7F7 P7H9
7889SA P1A21 P1A22 P1A81 P1B5 P1E1 P1G3 P2B3 P2B6 P2C4 P2C6 P2D5 P2E6 P2E12
Clone ID
short assay 1
short assay 2
F
CRISPR cut
APPSwe (GA>TC)
allele 1
allele 2
short assay 1 (148bp)
short assay 2 (109 bp)
qgPCR assay 1 (310 bp)
qgPCR assay 2 (354 bp)
APPSwe locus, unedited
fw_1
fw_2
probe_1
probe_2
APPSwe
rv_1
rv_2
Genomic location: 21 25897507 - 25897716, APP gene encoded on reverse strand
APPSwe locus, insertion allele of clone P1C4
fw_1 and probe_1 binding site lost: short assay 1 does not amplify insertion allele
insertion
deletion
4 bp mismatch with fw_2: short assay 2 still amplifies insertion allele
G
CRISPR cut
HDAC9Mut (GGGACAAA>CCCACACG)
allele 1
allele 2
short assay 1 (193 bp)
short assay 2 (185 bp)
qgPCR assay 1 (372 bp)
qgPCR assay 2 (360 bp)
HDAC9Mut locus, unedited
Genomic location: 7 19009629 - 19009908
HDAC9Mut locus, deletion allele of clone P1A21
fw_1 binding site lost: short assay 1 does not amplify deletion allele
deletion 1
deletion 2
all binding sites present: short assay 2 still amplifies deletion allele

Figure S3. Design of quantitative PCR assays for detection of allele copy numbers in CRISPR-edited iPSCs, related to Figures 1 and 3. (**A**) Design parameters and guidelines for qgPCR assays around inserted mutation(s) 'M'. By using the same "base PCR" as for RFLP and Sanger sequencing, qgPCR assays on edited single-cell clones easily integrate into existing genome editing workflows.
(**B-D**) Positions of two independent qgPCR assays around APP^{Swe} (**B**), APP^{Ibe} (**C**) and $HDAC9^{Mut}$ (**D**) loci shown in Figure 1, with forward primer (fw), reverse primer (rv) and qPCR probe (Full primer names, as listed in Methods: APP_Swe_Gt…, APP_Ibe_Gt…, HDAC9_Gt_...).
(**E**) 'Standard' short amplicon qPCR assays fail to detect aberrant clones: Allele copy numbers for two independent short amplicon qPCR assays reveal several hemizygous clones, but at each locus one aberrant clone is not detected by either one or both short assays (red box). All values normalized to unedited parent cell line (A18944 or 7889SA).
(**F**) Overview of qPCR probe designs (top), position of primers (fw, rv) and probes for short qPCR assays 1+2 at APP^{Swe} (full primer names, as listed in methods: APP_Swe_short…) (middle), and explanation of failed detection of insertion in clone P1C4 at APP^{Swe} locus (bottom).
(**G**) Overview of qPCR probe designs (top), position of primers (fw, rw) and probes for short qPCR assays 1+2 at HDAC9 (full primer names, as listed in methods: HDAC9_short…) (middle), and explanation of failed detection of insertion in clone P1A21 at $HDAC9^{Mut}$ locus (bottom).

		Copy number analysis by qPCR		LOH analysis			
Genotype by Sanger sequencing (homo-/hemizygous)	**Clone ID**	**qgPCR assay 1+2** (Figure 1E)	**Short amplicon assay 1+2** (Figure S3E)	**Zygosity at rs9976425*** (2.4 kb upstream)	**Zygosity at rs1783016**** (10 kb downstream)	**Global SNP genotyping**	**Aberrant allele**
APP^{Swe}	P1A5	2	2	heterozygous	heterozygous	not tested	-
APP^{Swe}	P1B3	2	2	heterozygous	heterozygous	not tested	-
APP^{Swe} (Figure S1A)	P1C4	**1**	**1/2 (Figure S3F)**	**homozygous**	heterozygous	no	**4.1 kb insertion, small deletions (Figure S1B)**
APP^{Swe}	P1C8	2	2	heterozygous	heterozygous	not tested	-
APP^{Swe}	P1D4	2	2	heterozygous	heterozygous	not tested	-
APP^{Swe}	P1D6	2	2	heterozygous	heterozygous	no	-
APP^{Swe} (Figure S1A)	P1F11	**1**	**1**	**homozygous (Figure 1B)**	heterozygous	no	**2.8 kb deletion (Figure 1C)**
APP^{Swe}	P1G4	2	2	heterozygous	heterozygous	not tested	-
APP^{Swe}	P1G5	2	2	heterozygous	heterozygous	not tested	-
APP^{Swe} (Figure S1A)	P1G9	2	2	**homozygous ~1 Mb LOH by further SNP genotyping analysis (Figure 2B)**	**homozygous**	**LOH until end of chromosome (Figure 2E)**	**LOH**
APP^{Swe}	P1H7	2	2	heterozygous	heterozygous	not tested	-
APP^{Swe}	P2B3	2	2	heterozygous	heterozygous	not tested	-
APP^{Swe}	P2C2	2	2	heterozygous	heterozygous	no	-
APP^{Swe} (Figure S1A)	P2C5	**1**	**1**	**homozygous**	heterozygous	no	**not resolved, potentially very large deletion**
APP^{Swe}	P2D1	2	2	heterozygous	heterozygous	not tested	-
APP^{Swe} (Figure S1A)	P2E9	2	2	**homozygous ~2 kb LOH by further SNP genotyping analysis (Figure 2D)**	**heterozygous**	no	**LOH**
APP^{Swe}	P2G7	2	2	heterozygous	heterozygous	not tested	-
Genotype by Sanger sequencing (homo-/hemizygous)	**Clone ID**	**qgPCR assay 1+2** (Figure 1G)	**Short amplicon assay 1+2** (Figure S3E)	**Zygosity at rs2070653*** (0.5 kb upstream)	**Zygosity at rs5843179*** (0.7 kb downstream)	**Global SNP genotyping**	**Aberrant allele**
APP^{Ibe} (Figure S1A)	P2G2	**1**	**2**	**homozygous**	**homozygous**	**~ 79 kb LOH**	**not resolved, potentially very large deletion and LOH**
APP^{Ibe}	P3B7	2	2	heterozygous	heterozygous	no	-
APP^{Ibe} (Figure S1A)	P6G10	**1**	**1**	**homozygous**	**homozygous**	not tested	**1.6 kb deletion (Figure S1B)**
APP^{Ibe}	P7F7	2	2	heterozygous	heterozygous	no	-
APP^{Ibe} (Figure S1A)	P7H9	2	2	**homozygous ~3 kb LOH by further SNP genotyping analysis (Figure 2D)**	**homozygous**	**~ 10 kb LOH**	**LOH**
Genotype by Sanger sequencing (homo-/hemizygous)	**Clone ID**	**qgPCR assay 1+2** (Figure 1H)	**Short amplicon assay 1+2** (Figure S3E)	**Zygosity at rs2717369*** (2 kb upstream)	**Zygosity at rs2717368*** (2 kb downstream)	**Global SNP genotyping**	**Aberrant allele**
$HDAC9^{Mut}$ (Figure S1A)	P1A21	**1**	**1/2 (Figure S3G)**	**homozygous**	**homozygous**	not tested	**3 small deletions (Figure S1B)**
$HDAC9^{Mut}$	P1A22	2	2	heterozygous	heterozygous	not tested	-
$HDAC9^{Mut}$ (Figure S1A)	P1A81	**1**	**1**	**2 PCR products of different length**	**homozygous**	no	**inversion, deletion, insertion (Figure S1B)**
$HDAC9^{Mut}$	P1B5	2	2	heterozygous	heterozygous	no	-
$HDAC9^{Mut}$ (Figure S1A)	P1E1	2	2	**homozygous ~1 Mb LOH by further SNP genotyping analysis (Figure 2C)**	**heterozygous**	**LOH until end of chromosome (Figure 2E)**	**LOH**
$HDAC9^{Mut}$ (Figure S1A)	P1G3	**1**	**1**	**homozygous**	**homozygous**	no	**1.8 kb insertion, small deletion (Figure S1B)**
$HDAC9^{Mut}$	P2B3	2	2	heterozygous	heterozygous	not tested	-
$HDAC9^{Mut}$	P2B6	2	2	heterozygous	heterozygous	not tested	-
$HDAC9^{Mut}$ (Figure S1A)	P2C4	**1**	**1**	**2 PCR products of different length**	**homozygous**	not tested	**1.4 kb deletion (Figure S1B)**
$HDAC9^{Mut}$	P2C6	2	2	heterozygous	heterozygous	not tested	-
$HDAC9^{Mut}$ (Figure S1A)	P2D5	**1**	**1**	**homozygous**	**homozygous**	not tested	**1.2 kb deletion, inversion (Figure S1B)**
$HDAC9^{Mut}$	P2E6	2	2	heterozygous	heterozygous	not tested	-
$HDAC9^{Mut}$	P2E12	2	2	heterozygous	heterozygous	no	-

* PCR for SNP genotyping was spanning the cut site
** PCR for SNP genotyping was **not** spanning the cut site

Table S1. Overview of CRISPR-edited iPSC clones at APP^{Swe}, APP^{Ibe}, $HDAC9^{Mut}$, related to Figures 1 and 2. Data of altered clones shown in indicated figures, other data not shown.

Locus	Clone ID	Retained I Deleted I Retained	Homologous bases
APP^{Swe}	P1C4	Insertion	N/A
APP^{Swe}	P1F11	AGACAGT**TCC**|GGATGTGAAT.....TCCCAAA**TCC**|TGACCTATAA	3
APP^{Ibe}	P6G10	CATCA**CCAAG**|GTGATGACGA.....GAAAG**CCAAG**|ATTCTTGTGC	5
$HDAC9^{Mut}$	P1A21 (1st Deletion)	TTCTTT**GTAC**|GTACTGTGGC.....TAAAAA**GTAC**|TCATTGAGAA	4
$HDAC9^{Mut}$	P1A21 (2nd Deletion)	AAAAGATG**TG**|GGATTTTTAT.....TCATATCC**TG**|TAATTTTTCA	2
$HDAC9^{Mut}$	P1A21 (3rd Deletion)	CAAAAATTTT|**GCCAA**ATTGA.....TAAATATTTG|**GCCAA**CTTTT	5
$HDAC9^{Mut}$	P1A81	Complex changes with duplication and inversion	N/A
$HDAC9^{Mut}$	P1G3	Insertion	N/A
$HDAC9^{Mut}$	P2C4	GGATTG**AAGA**|CATATCCCTC.....GCAAAA**AAGA**|ATGTACAAGC	4
$HDAC9^{Mut}$	P2D5	Complex changes with duplication and inversion	N/A

Table S2. Microhomologies are prevalent at large deletion sites in CRISPR-edited iPSCs, related to Figure 1. Sequences around deletion sites with microhomologies (indicated with red letters) suggesting involvement of the MMEJ pathway. Black bars indicate sites of fusion between flanking regions, intervening part is deleted.

3.2 Simple and reliable detection of CRISPR-induced on-target effects by qgPCR and SNP genotyping

3.2.1 Summary

In this protocol, we provide a detailed description for reliable OnTE detection after CRISPR-editing, based on the tools that we have established in our first study described above. All steps are described with substantial detail as well as suggestions for troubleshooting to also enable conduction of these methods by less experienced users. Furthermore, we extend protocol and discussion to genome-edited organisms that might also be affected by OnTE occurrence.

The protocol covers six main parts: 1) extraction of genomic DNA from edited samples for further analysis, 2) optimization of a genotyping PCR, 3) design of a qgPCR assay, 4) analysis of edited cells and organisms by qgPCR, and LOH analysis by 5) nearby SNP sequencing or the alternative option 6) SNP microarrays.

For the first step, we provide two different options: harvesting of cells from edited clones or collection of tissues from edited organisms followed by extraction of genomic DNA that will be used in the subsequent analysis. Next, we describe optimization of a 300-450 bp genotyping PCR reaction around the cut site by comparing the efficiency of different primer combinations. The third step involves designing an internal probe that is compatible with the genotyping PCR primers, which should then be ordered as fluorescently labelled oligo to enable quantification in the following qgPCR analysis. In the fourth step, we describe analysis of edited cells or organisms by qgPCR, including preparation of the reaction, running it on a standard qPCR machine and finally correlating the resulting threshold cycle (C_t) values with the number of intact alleles at the target locus to exclude samples with large deletions, insertions, or complex rearrangements. Nearby SNP genotyping in the fifth step starts with retrieving a list of 10-20 SNPs around the edited locus from an online database. Genotyping of these SNPs in the unedited parent cell line or organism should be performed until at least one heterozygous SNP is identified on both sides of the target locus, which is followed by validation of zygosities after editing. For the alternative option for LOH analysis by SNP microarrays, we start with a description how to identify suitable microarray chips, which should have a sufficient SNP coverage around the target site. Subsequent steps are based on specific equipment like a hybridization oven, pipetting robots and an iScan scanner. Data analysis involves extracting B allele frequencies for SNP zygosities and Log R ratios for copy numbers for each SNP to identify regions of LOH. Edited single cell clones or organisms that show any abnormalities in the assays, should be excluded from further experiments due to OnTEs.

3.2.2 Reference

The paper was published in the Journal ‘Nature Protocols’ under the following reference:

Weisheit, I., Kroeger, J. A., Malik, R., Benedikt Wefers, B., Lichtner, P., Wurst, W., Dichgans, M., & Paquet, D. (2021). Simple and reliable detection of CRISPR-induced on-target effects by qgPCR and SNP genotyping. ***Nature Protocols***, 16. https://doi.org/10.1038/s41596-020-00481-2

nature protocols

PROTOCOL

https://doi.org/10.1038/s41596-020-00481-2

Check for updates

Simple and reliable detection of CRISPR-induced on-target effects by qgPCR and SNP genotyping

Isabel Weisheit[1,2], Joseph A. Kroeger[1,2], Rainer Malik[1], Benedikt Wefers[3,4], Peter Lichtner[5], Wolfgang Wurst[3,4,6,7], Martin Dichgans[1,3,7] and Dominik Paquet[1,2,7]✉

The recent CRISPR revolution has provided researchers with powerful tools to perform genome editing in a variety of organisms. However, recent reports indicate widespread occurrence of unintended CRISPR-induced on-target effects (OnTEs) at the edited site in mice and human induced pluripotent stem cells (iPSCs) that escape standard quality controls. By altering gene expression of targeted or neighbouring genes, OnTEs can severely affect phenotypes of CRISPR-edited cells and organisms and thus lead to data misinterpretation, which can undermine the reliability of CRISPR-based studies. Here we describe a broadly applicable framework for detecting OnTEs in genome-edited cells and organisms after non-homologous end joining-mediated and homology-directed repair-mediated editing. Our protocol enables identification of OnTEs such as large deletions, large insertions, rearrangements or loss of heterozygosity (LOH). This is achieved by subjecting genomic DNA first to quantitative genotyping PCR (qgPCR), which determines the number of intact alleles at the target site using the same PCR amplicon that has been optimized for genotyping. This combination of genotyping and quantitation makes it possible to exclude clones with monoallelic OnTEs and hemizygous editing, which are often mischaracterized as correctly edited in standard Sanger sequencing. Second, occurrence of LOH around the edited locus is detected by genotyping neighbouring single-nucleotide polymorphisms (SNPs), using either a Sanger sequencing-based method or SNP microarrays. All steps are optimized to maximize simplicity and minimize cost to promote wide dissemination and applicability across the field. The entire protocol from genomic DNA extraction to OnTE exclusion can be performed in 6–9 d.

Introduction

CRISPR–Cas9 technology is revolutionizing biomedical research because it allows researchers to directly and specifically modify genes in many different organisms, including human cells and patients[1–4]. New protocols for genome editing are constantly generated and improved to further increase efficiency and applicability[5,6]. In the course of this advancement, the technology has been widely used to generate models of human disease in, e.g., mice or clinically relevant induced pluripotent stem cells (iPSCs). However, CRISPR editing is not always precise and can lead to inadvertent alterations in the edited genome, both at the edited site (on-target) and at other genomic loci (off-target). Although the problem of off-target effects has been well recognized and addressed with reliable detection technologies in the field[5,7,8], efficient and broadly applicable tools for detection of on-target effects (OnTEs) are still lacking. As we and others have recently shown, OnTEs can occur at high frequency in mouse and human cells, with up to 40% of edited clones affected in human iPSCs[9–13]. In this article we describe a procedure to efficiently detect OnTEs and ensure locus integrity after CRISPR editing.

Development of the protocol

In our recent study[13] we revealed frequent occurrence of OnTEs, such as large deletions, large insertions, complex rearrangements and loss of heterozygosity (LOH) caused by CRISPR genome editing in human iPSCs. These OnTEs occurred in up to 40% of clones edited to introduce targeted mutations via the homology-directed repair (HDR) pathway and to a similar degree in clones with apparently homozygous frameshift edits to generate knockouts via the non-homologous end joining

[1]Institute for Stroke and Dementia Research (ISD), University Hospital, LMU Munich, Munich, Germany. [2]Graduate School of Systemic Neurosciences, LMU Munich, Planegg-Martinsried, Germany. [3]German Center for Neurodegenerative Diseases (DZNE) Munich, Munich, Germany. [4]Institute of Developmental Genetics (IDG), HelmholtzZentrum München, Neuherberg, Germany. [5]Core Facility NGS, HelmholtzZentrum München, Neuherberg, Germany. [6]Technische Universität München-Weihenstephan, Neuherberg, Germany. [7]Munich Cluster for Systems Neurology (SyNergy), Munich, Germany. ✉e-mail: dominik.paquet@med.uni-muenchen.de

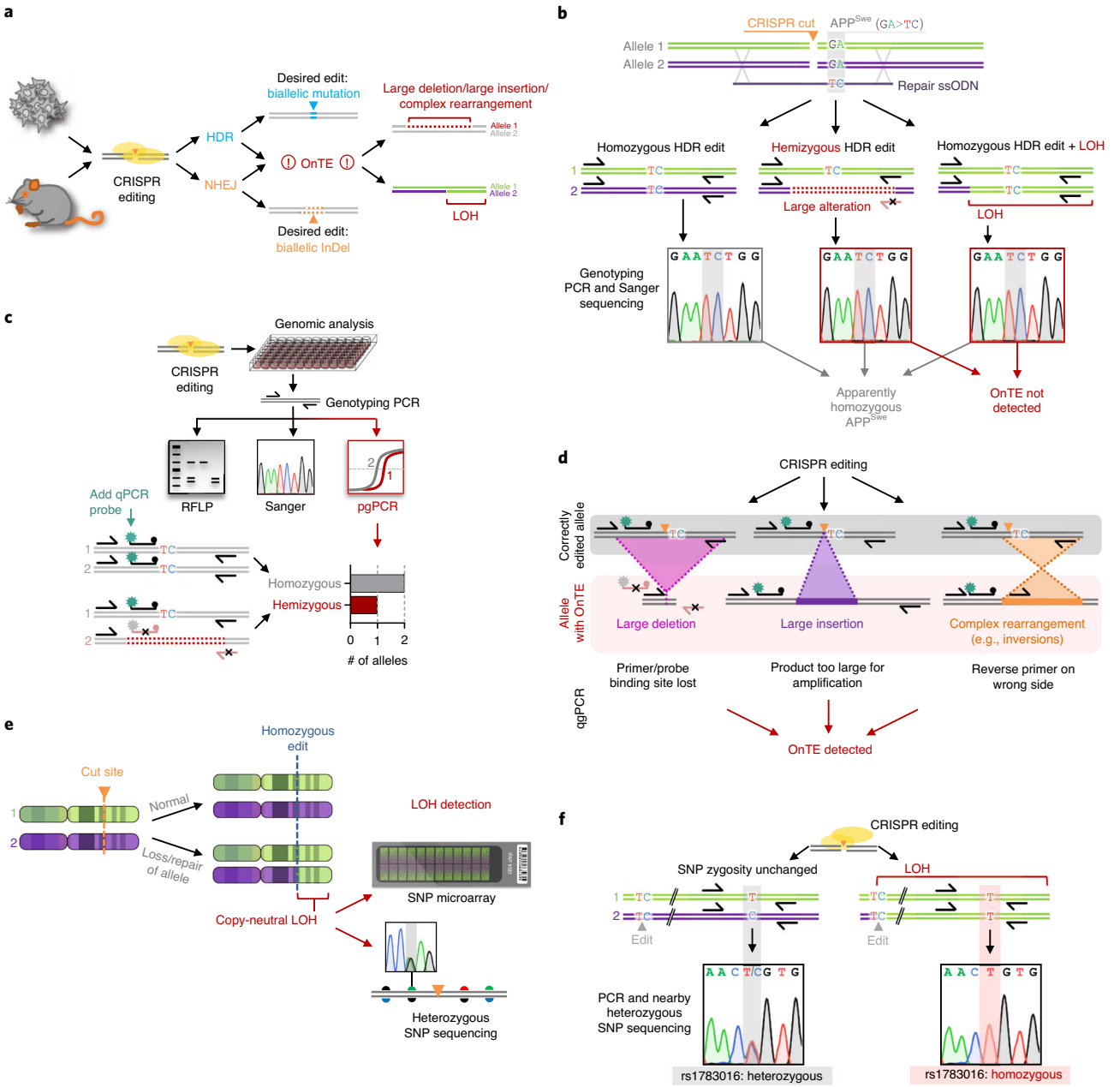


Fig. 1 | CRISPR–Cas9 genome editing induces deleterious OnTEs that can be detected by qgPCR and SNP genotyping in a simple and reliable manner. a, CRISPR editing can induce unintended OnTEs such as large deletions, large insertions, complex rearrangements or regions of copy-neutral loss of heterozygosity (LOH) around the target locus in human cells, mice and other organisms during HDR- or NHEJ-mediated editing. **b**, OnTEs often escape standard quality controls such as Sanger genotyping: large alterations might prevent primer binding at the affected allele, which leads to seemingly homozygous sequencing traces (middle panel). Regions of LOH are also routinely missed as there are usually no heterozygous SNPs present within the genotyping PCR (right panel). ssODN, single-stranded oligodeoxynucleotide. **c**, qgPCR analysis can be easily integrated into existing genome-editing workflows by adding a fluorescently labeled probe to the established genotyping PCR and measuring copy numbers by qPCR. Hemizygous clones with only one intact allele will be revealed by increased C_t values. RFLP, restriction fragment length polymorphism. **d**, qgPCR allows detection of monoallelic large deletions (left), large insertions (middle) and complex rearrangements (right) that prevent amplification of the affected allele. **e**, LOH is detected by Sanger sequencing of nearby heterozygous SNPs or SNP microarrays that investigate SNP zygosities and copy numbers genome-wide. **f**, Nearby SNP sequencing involves identifying heterozygous variants around the target locus in unedited cell lines and genotyping these in edited cell lines. **c** and **e** are modified from ref. [13].

(NHEJ) pathway. Earlier work also demonstrated presence of OnTEs in CRISPR-edited mice and cell lines[9–12], and it is therefore conceivable that OnTEs are a widespread problem relevant for CRISPR editing in many organisms (Fig. 1a). Highlighting the importance of their detection, we showed that OnTEs can have deleterious consequences on the formation of relevant phenotypes in disease

models[13]. We therefore developed and validated assays for reliable detection of OnTEs based on quantitative PCR (qPCR), Sanger sequencing and single-nucleotide polymorphism (SNP) microarrays. Rather than developing a complex technology accessible only by specialist laboratories, we reasoned that the best way to perform OnTE detection is with a simple and affordable technology that is broadly applicable to established editing workflows. We therefore set out to develop a protocol that is easy to implement by every laboratory performing CRISPR genome editing with minimal requirements for specialized machinery. The entire workflow of OnTE detection, including analysis of LOH, can be performed using standard molecular biology equipment such as PCR and qPCR machines and a gel electrophoresis setup. Alternative LOH detection by SNP microarrays can be performed if relevant equipment or access to a core facility is available. Analysis of all steps requires only freely available software. We hope this will help to ensure wide dissemination of OnTE quality control in the genome editing field.

OnTEs can be large deletions, large insertions or other complex rearrangements at edited genomic loci. Standard genotyping by PCR and Sanger sequencing will only detect small changes within the PCR amplicon; however, larger events are missed if they only occur on one allele. While the unaffected allele will display the desired edit, the aberrant allele will be invisible when a loss of primer binding sites prevents amplification and sequencing. In addition, it has been described very recently that CRISPR editing can also cause LOH, i.e., the replacement of potentially large areas of one homologous chromosome by the other[13,14]. This type of OnTE is also invisible in standard genotyping assays (Fig. 1b), except in rare cases where heterozygous SNPs are present in the PCR amplicon on both sides of the edit. To exclude LOH or the complete loss of an edited allele, we developed an OnTE detection framework based on a modified application of two basic techniques that are available in most molecular biology laboratories: qPCR and genotyping of SNPs.

We first describe quantitative genomic PCR (qgPCR), a new method in which a 300–450 bp genotyping PCR of the region around the edited site is used to confirm presence of two correctly edited alleles. This step involves designing a fluorescently labeled probe that binds to the PCR amplicon according to specific design principles described in our protocol, extracting genomic DNA (gDNA) from a clonal population of edited cells or gene-edited organisms, and performing and analyzing qgPCR (Fig. 1c). In the qgPCR reaction, the probe allows real-time monitoring of target locus amplification rates during each PCR cycle. Calculation of gene copy numbers can then be performed by determining the threshold cycle (C_t), i.e., the cycle when the amplification exceeds a certain threshold above background fluorescence[15,16]. This allows detection of cells or organisms with OnTEs caused by large deletions, large insertions or complex rearrangements (Fig. 1d). Performing reliable qgPCR requires avoiding several pitfalls of 'standard' qPCR assays, which we explain in our protocol. Following qgPCR, absence of LOH is confirmed by genotyping heterozygous SNPs around the edited site. This step verifies that genomic regions that were heterozygous in the original unedited cell line or organism are still heterozygous after editing. We provide two different options for this analysis: nearby SNP genotyping and genome-wide SNP microarray analysis (Fig. 1e). Nearby SNP genotyping determines zygosity of neighbouring SNPs by Sanger sequencing (Fig. 1f). We describe how to identify, select and analyze suitable SNPs. For SNP microarray analysis we describe a detailed workflow including selection of suitable chips, sample preparation and detailed instructions for data analysis. We discuss advantages and disadvantages of both approaches and provide recommendations and detailed descriptions for use depending on the edited locus and technological availabilities in laboratories (see Box 1 for details).

Applications of the method

OnTEs are widespread in edited human iPSCs, other human cells and mice[9–13], but probably also occur in other organisms subjected to CRISPR editing. We have established our protocol using CRISPR-edited human iPSCs. However, since complete analysis of OnTE is achievable using only purified gDNA from the edited sample, this protocol can also be applied for detection in other edited cell lines or animals (for discussion regarding application and limitations of the method in animals, see below). Furthermore, the OnTE detection framework described in this protocol is applicable to different types of introduced genomic changes: the main requirement for performing qgPCR is that successful editing can be confirmed by a single genotyping PCR that spans across the edited site, where we recommend a size not larger than 450 bp for efficient amplification of the PCR product. Thus, OnTE evaluation by qgPCR is ideally suited for all cases in which either specific homozygous single-base-pair changes were introduced by HDR-mediated editing, or small insertions–deletions

Box 1 | Choosing nearby SNP genotyping or SNP microarrays for LOH analysis

Our protocol describes two options for detecting LOH after editing: nearby SNP genotyping and SNP microarrays. Both assays have distinct advantages and disadvantages that every laboratory needs to consider. A decision as to which method should be used will be based on available expertise, equipment and the desired information. Users requiring a maximal level of certainty may also consider performing both assays in parallel.

Nearby SNP genotyping investigates editing-induced zygosity changes of nearby heterozygous SNPs by Sanger genotyping. This is a cheap and simple assay that only requires identifying and genotyping heterozygous variants surrounding the target locus, and can therefore be easily implemented with standard technologies available in most laboratories, such as Sanger sequencing. It is also more flexible than predefined SNP microarrays because the genotyped SNPs can be freely chosen according to experimental needs. This custom design can improve detection of small regions of LOH that may be missed by microarrays that might not assay affected SNPs close enough to the edited site. However, nearby SNP sequencing only provides genotyping information for a restricted area and therefore further characterizing the extent of LOH events affecting large parts of a chromosome is challenging. Another limitation of this method is that the zygosity of each SNP needs to be determined experimentally, which can be laborious if editing was done in a region of low variation and no heterozygous SNPs are present directly near the edited site. If a model cell line is used repeatedly in a laboratory for genome editing, it can be beneficial to obtain a whole-genome SNP profile using a high-density SNP microarray once (see below), which can then guide selection of heterozygous SNPs for Sanger genotyping in subsequent experiments. Lastly, SNP sequencing cannot be easily upscaled or automated, preventing efficient analysis of experiments with multiple edited loci or multiple differentially edited lines.

SNP microarrays will investigate SNP zygosities and corresponding copy numbers on a genome-wide scale. A major advantage is that multiple cell lines, in particular cell lines with different genome edits, can be analyzed in parallel in a timely manner. SNP microarrays are especially useful to identify larger regions of LOH and to determine their full dimension. The reliability of SNP microarray data increases with the number of detected SNPs in the affected area, and therefore small on-target zygosity changes of only one or a few SNPs can be challenging to resolve, depending on the SNP coverage of the chip. This can be partially improved by choosing the most appropriate chip after investigating the SNP coverage of different SNP microarrays at the target locus, and/or custom tailoring assayed SNPs, but even high-density chips will only cover a subset of SNPs present in a cell line or organism. Lastly, LOH analysis by SNP microarrays is more expensive than nearby SNP genotyping, requires access to suitable equipment and involves more complex data analysis.

It should be added that besides OnTE detection, microarray analysis can also be applied to investigate chromosomal aberrations genome wide by 'molecular karyotyping'. This way, deletions or duplications larger than ~600 kb–1 Mb can be detected; however, smaller aberrations, and copy-neutral inversions and balanced translocations, will be missed by microarray analysis.

(indels), as typically introduced by NHEJ-mediated editing to generate gene knockouts. If heterozygous edits were introduced, the presence of both alleles can already be confirmed by Sanger sequencing, making further analysis by qgPCR unnecessary. LOH analysis, however, should still be performed (see below). Furthermore, larger genomic changes up to ~350 bp, such as targeted large deletions/inversions or large knockins (e.g., for protein tagging) could also be checked for OnTEs as long as the PCR spans the edit, although we have not tested this application. However, different amplification efficiencies of the two differentially sized amplicons may complicate the analysis, and therefore comparable PCR efficiencies should be ensured. Lastly, qgPCR may also be applied to edits that are larger than the recommended 350 bp, such as fusion of GFP-tags to endogenous proteins. However, this would require an amplicon specific to the edit, thus preventing normalization of the qgPCR to the unedited 'parent' cell line or animal. In such a case, normalization to the reference assay (e.g., at the telomerase reverse transcriptase (TERT) gene locus for human or transferrin receptor (Tfrc) for mouse samples) could be performed, which also requires comparable qgPCR efficiencies to determine defined copy numbers.

The LOH analysis is even more widely applicable and can be used to quality control all genome-editing procedures without further adjustments, including those where large knockins were performed. Furthermore, our methods for OnTE detection are in principle also applicable to non-CRISPR genome editing tools, including ZFNs and TALEN. However, to our knowledge, it has not been investigated yet whether cells or organisms edited with these tools are also affected by OnTEs.

Comparison with other methods

Several other methods have previously been used for OnTE detection.

Primer-walking is a PCR-based method that uses a series of primers with increasing distance to the edited site to overcome the problem that closer primer binding sites are deleted by the OnTEs. Although primer-walk PCRs in principle allow detection of large deletions at low cost, we have shown recently that they are not always meaningful or reliable for detecting all OnTEs, especially in the presence of large insertions or complex alterations that exceed the amplification limits of traditional PCR[13]. Moreover, primer-walking PCR does not detect LOH.

Next-generation sequencing technologies allow the identification of OnTEs after editing at high sensitivity and specificity[10–12,17] by massively parallel sequencing of the area around the edited region and analysis of read coverage and sequence. However, deep-sequencing-based methods are expensive,

require specialized equipment and expertise for implementation, are only available to specialized laboratories and require complex data analysis, especially for short-read-based sequencing methods (producing reads of a few hundred base pairs): multiple ensemble algorithms for different types of OnTEs need to be applied to map discordances between sample and reference genomes. However, since OnTEs can include very diverse structural variations of the genome, the method may not detect all OnTEs[18,19]. To overcome this problem, targeted long-read-sequencing approaches, e.g., from Pacific Biosciences or Oxford Nanopore Technologies, can be applied. These will generate continuous reads of several kilobases around the edited locus, which eliminates the need for complex read assembly, but always restricts OnTE detection to the read length[17,18].

Another alternative for OnTE detection could be droplet digital PCR, which is based on the principle of a qPCR reaction, but additionally partitions the sample into thousands of individual droplets (i.e., 'reaction chambers'), allowing absolute quantification of target sequences[12]. This increases sensitivity and specificity, but requires special equipment and more complex data analysis, without—in our view—providing a major benefit for the purpose of OnTE detection.

Lastly, array-based technologies, such as comparative genomic hybridization (array-CGH) or SNP microarrays, may be applied. Array-CGH is based on hybridization of sample and control gDNA and is routinely used to perform molecular karyotyping with a resolution of up to 60 kb, which largely surpasses standard G banding[20]. However, this resolution is still insufficient to detect many of the OnTEs we detected in our iPSCs, which were in many cases well <10 kb (ref. [13]). In addition, array-CGH cannot reliably detect LOH. SNP microarrays are based on genome-wide SNP analysis by hybridization of gDNA to oligo probes on a chip and provide information on SNP zygosity as well as copy number. Although this method offers relatively high resolutions to also detect small LOH, it is insufficient to also detect small deletions, small insertions or rearrangements affecting none, only one or a few SNPs that we primarily found in our studies (see Box 1 for more details). Inversions or balanced translocations at the edited locus will also be missed by SNP arrays.

Standard qPCR assays, without the design optimizations we implemented for qgPCR, are also inadequate to detect OnTEs because they failed to reliably detect all affected clones (see qgPCR design section for more details)[13].

Experimental design

Detection of OnTEs after CRISPR editing consists of two main phases: (i) performing qgPCR to confirm locus integrity and exclude the possibility of large deletions, large insertions or complex rearrangements at the target locus; and (ii) investigate occurrence of LOH around the target site by SNP genotyping, which can be performed either by nearby SNP genotyping and/or SNP microarrays (see workflow in Fig. 2).

Timing of OnTE detection

Due to the frequent occurrence of OnTEs (up to 40% of HDR-edited human iPSCs[13]), we recommend prioritizing their detection over other quality-control measures after CRISPR editing. In our experiments exclusion of clones was more frequently due to OnTEs than other aberrations, such as off-target effects, issues with pluripotency, integration of Cas9 plasmids, clonality or karyotype (Box 2 lists our general recommendations for quality controls after CRISPR editing in human iPSCs, other cultured cells or animals). We therefore typically isolate gDNA of single-cell clones immediately after editing and confirm successful editing by Sanger genotyping, directly followed by OnTE analysis.

For OnTE analysis, we recommend starting with qgPCR because OnTEs detected by this method, such as large deletions, large insertions or complex rearrangements, seemed to occur more frequently than LOH in human iPSCs. Furthermore, once optimized, qgPCR can easily be performed for multiple samples in parallel in a few hours in a single qPCR reaction. Single-cell clones that do not exhibit any abnormalities in the qgPCR assay are then further subjected to SNP genotyping, either by nearby SNP sequencing or SNP microarray analysis, which takes ~2–3 d.

Controls

OnTEs are always detected by comparing gDNA from edited cells to that from unedited 'parent' cells or organisms. As we occasionally observed variations in gDNA integrity from stored samples, we recommend that control gDNA should be freshly isolated together with all other samples using a method that provides high-purity gDNA. For qgPCR, the control sample serves as a normalization factor to exclude clones that have acquired copy number changes by genome editing. For LOH

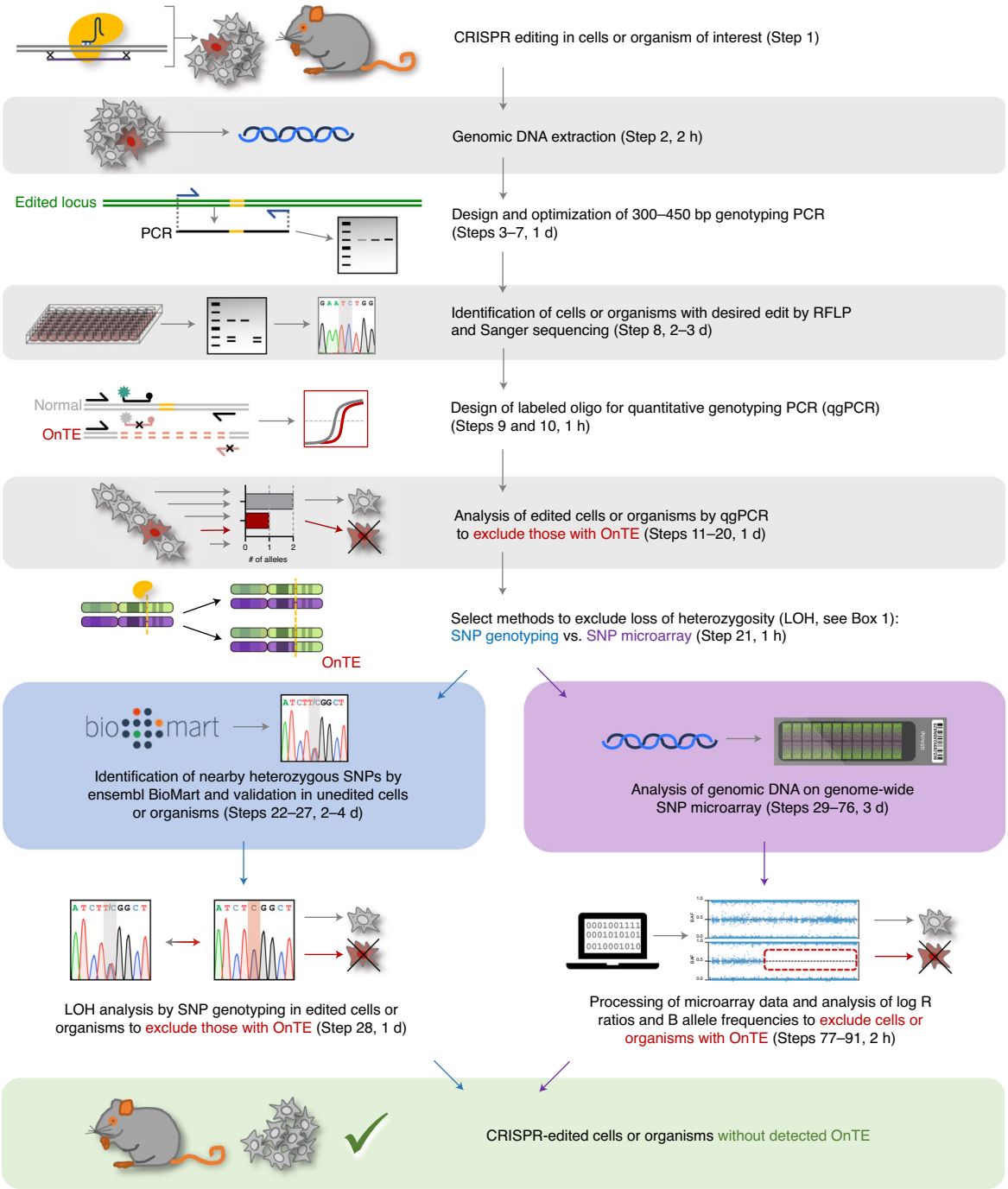


Fig. 2 | Overview of the procedure for simple and reliable OnTE detection after CRISPR genome editing. First, genome editing is performed in a cell line or organism of interest (Step 1), followed by extraction of genomic DNA (Step 2). After design and optimization of a 300–450 bp genotyping PCR (Steps 3–7), this PCR is used to identify cells or organisms with the desired edit by restriction fragment length polymorphism (RFLP) and Sanger sequencing (Step 8). A labeled oligo for the quantitative genotyping PCR (qgPCR) reaction is then designed that matches the genotyping primers (Steps 9 and 10). Next, edited cells or organisms are analyzed by qgPCR to exclude those with OnTEs like large deletions, large insertions or other complex rearrangements (Steps 11–20). For the following LOH analysis, one of the two options should be selected: nearby SNP genotyping or SNP microarray analysis (Step 21). For the first option, SNPs are identified by ensemble BioMart and validated in unedited cells or organisms (Steps 22–27). The zygosity of heterozygous SNPs is then confirmed in edited cells or organisms to exclude those with OnTEs (Step 28). For the second option, genomic DNA is first analyzed on a genome-wide SNP microarray (Steps 29–76), followed by processing of log R ratios and B allele frequencies to exclude cells or organisms with OnTEs (Steps 77–91).

analysis, the zygosity of different SNPs is first identified in the control sample. Then, using gDNA from edited samples, the same SNPs are genotyped to ensure that heterozygous SNPs have maintained their zygosities after editing.

Box 2 | Overview of quality controls after CRISPR editing in iPSCs, other cultured cells or animals

CRISPR editing can lead to a variety of non-intended alterations in the genome. These can be induced by the CRISPR-associated nuclease around the edited site (OnTEs) or in other regions of the genome (off-target effects). Additionally, the process of editing and cultivation of single cells subjects cells to stress and can induce or enrich chromosomal aberrations. To allow meaningful comparisons between edited cells or organisms and their non-edited isogenic controls, non-intended alterations need to be excluded as much as this is technically possible. We apply a quality control workflow for each edited iPSC line[35] and founder animal, which is prioritized according to frequency of occurrence as well as the cost and complexity of the assay:

1. **OnTEs (cells and animals)**: as described in this protocol, every line or founder is screened for OnTEs by qgPCR and SNP genotyping (see discussion above about applicability in animals).
2. **Chromosomal aberrations (cells and animals)**: both cutting a chromosome by CRISPR and single-cell cloning can induce chromosomal aberrations, such as translocations, trisomies, deletions or inversions. Such alterations can be detected by karyotyping, using traditional G banding, or SNP microarray-based molecular karyotyping (see Box 1). Molecular karyotyping is more affordable and provides higher-resolution data but cannot detect balanced translocations and inversions. As both assays are still relatively expensive, we recommend prescreening for the most typical aberration by qPCR[36]. In our human iPSCs this is a local trisomy of Chr 20q11.21 (BCL2L1), but qPCR detection of other aneuploidies can be added, depending on expected occurrence in the edited cell line or organism.
3. **Off-target effects (cells and animals)**: guide RNAs with low predicted off-target rates are chosen using CRISPOR[37]. After editing, potential off-target effects are determined by Sanger sequencing the top five off-target hits from both MIT and CFD scoring algorithms[38–40]. Off-target detection can also be performed by more unbiased methods such as DISCOVER-seq, Guide-Seq or whole-genome sequencing[5,7,8]. In animals, backcrossing is usually applied to remove non-linked off-target effects. Another possibility to exclude negative consequences of off-target effects is to work with two lines edited by different gRNAs with non-overlapping off-target profiles.
4. **Identity of cell line or animals**: fingerprinting of cell lines can be applied to exclude mix-up during editing or contaminations with different cell lines. Lines can be fingerprinted thoroughly by STR profiling[35], but if a laboratory only uses a limited number of lines, a simple PCR assaying one variable genomic locus can be sufficient to distinguish lines from several iPSC donors. We assay a microvariation at the human D1S80 locus on Chromosome 1 (ref. [41]). Variability of alleles between individuals can be visualized by analyzing PCR products on a 2% (wt/vol) agarose gel (PCR described in step 4: annealing temperature, 68 °C; extension time, 1 min; primers, GTCTTGTTGGAGATGCACGTGCCCCTTGC, GAAACTGGCCTCCAAACACTGCCCGCCG). In animals, genetic background can be determined. If untreated or littermate controls are unavailable, and the exact genetic background is unknown, identification of the correct strain can be easily achieved using genetic strain panels from commercial providers (e.g., Genome Scanning Service, JAX). Alternatively, known variations between different (sub)strains (such as the Nnt deletion or Crb1rd8 mutation in mice[42–44]) can be analyzed by PCR-based assays to identify the actual genetic background.
5. **Mixed cell population (cells only)**: we check clonality of edited iPSC lines by plating edited clones as single cells on a 10 cm dish, followed by picking and genotyping 30–50 clones as described in ref. [27].
6. **Genomic integration of editing components (cells and animals)**: if plasmids were used for editing, their genomic integration needs to be excluded, e.g., by PCR and/or absence of antibiotic resistance mediated by the plasmid. If exclusion of HDR repair template off-target integration is desired (integration is less likely when using ssODNs, instead of double-stranded donors), qgPCR could be redesigned to determine repair template copy numbers within the genome. For this, both primers and probe would have to lie within the repair template.
7. **Pluripotency (stem cells only)**: as CRISPR editing of an iPSC line usually does not affect its pluripotency, it is sufficient to confirm typical stem cell morphology by live brightfield microscopy and the presence of pluripotency factors (e.g., Oct4, Tra1-60, SSEA4, Nanog)[28] by immunofluorescence staining. To further confirm pluripotency, scorecard assays or in vitro undirected embryoid-body-based differentiation followed by identification of expression markers for all three germ layers can also be performed[45].

qgPCR design

Performing qgPCR requires applying a qPCR assay amplifying a region of ~300–450 bp spanning across the edited genomic locus. The qgPCR assay consists of two primers amplifying the target region and a fluorescently labeled probe that allows quantification of the PCR reaction. We recommend reusing primers that were previously designed for genotyping the locus by Sanger sequencing and combining them with a matching probe using a qPCR design tool that allows probe design for existing primer sequences. The probe is then ordered as a custom oligo using dye and quencher fitting to local qPCR infrastructures. In the qPCR reaction, a copy number reference assay should be run together with the qgPCR assay for the edited gene. Amplifying a reference gene known to be present in two copies in the genome enables the gene of interest to be normalized to an endogenous control. Thus, C_t value differences caused by different amounts of gDNA in the reaction tube can be normalized.

When we developed the method in our original study, we investigated the possibility of assay artefacts by testing two different qgPCR assays with primers being shifted in different directions relative to the cut site as well as probes positioned on different sides of the cut site at three different loci in human iPSCs[13]. In these experiments, we did not observe conflicting results between different assay sets and therefore conclude that one qgPCR is sufficient for reliable OnTE detection. However, users requiring a maximal level of certainty may still consider using two or more independent assays based on our design principles.

Commonly used design tools for quantitative PCR (qPCR) typically recommend amplicon sizes of no more than 150 bp for qPCR assays. But, as we have shown recently[13], this size is not sufficient for reliable on-target analysis after genome editing. In several cases the assays failed to reliably flag clones affected by OnTEs where insertions or deletions did not directly overlap with the cut site. This allowed primers from the short amplicon assay to still bind and amplify the aberrant allele, leading to false-negative signals in the qgPCR. Longer amplicons with the recommended length of 300–450 bp were sufficient to reveal OnTEs non-contiguous with the cut site. Therefore, this size is a good

compromise between efficiency of qPCR (which decreases with increasing amplicon length) and likelihood of OnTEs (which decreases with increasing distance from the cut site).

LOH analysis by SNP genotyping

We applied and validated two methods for LOH analysis: nearby SNP sequencing and SNP microarrays. Both assays provide genotype information that can be applied to exclude the occurrence of LOH around the edited locus. It is up to the user to decide which assay is more appropriate for individual experiments and laboratory requirements (see detailed discussion of benefits and challenges in Box 1).

Nearby SNP sequencing. Performing nearby SNP sequencing for LOH analysis requires first identifying potentially heterozygous SNPs on both sides of the edited locus in the 'parent' cell line or organism using Ensembl BioMart[21], followed by validation of heterozygosity using Sanger sequencing. BioMart accesses the Ensembl databases with genomic variants that can be filtered by, e.g., genomic region, type of variants and allele frequency. We recommend first concentrating on a region of ~10 kb around the edited site and human short variants with a global minor allele frequency (MAF) ≥0.3. If these settings do not allow retrieval of ~5–10 SNPs on each side, the genomic region can be increased or the cutoff for MAF decreased. The flanking sequence around the variants is downloaded from Ensembl and used as a basis for genotyping SNPs by PCR and Sanger sequencing. To exclude the occurrence of LOH, it is necessary to identify one heterozygous SNP on each side of the targeted locus in the unedited cell line or organism to allow thorough analysis after editing. Assuming independent inheritance of the SNPs (no linkage disequilibrium), a maximum of nine SNPs with an MAF ≥0.3 have to be checked to reach a 99% chance of finding a heterozygous one (allele frequency according to Hardy–Weinberg $= 2 \times 0.7 \times 0.3 = 42\%$, minimum probability $n = \ln(1 - 0.99)/\ln(1 - 0.42) = 8.5$). If desired, nearby SNP genotyping can be extended to loci further away from the cut site to determine the dimension of large regions of LOH.

SNP microarray. For analysis by SNP microarrays it is important to select chips with sufficient SNP coverage in the edited region. Several biallelic variants should be assayed near the targeted site (preferably within 10–50 kb) on each side of the targeted locus. Only the SNPs that are heterozygous in the unedited 'parent' cell line/organism will be informative for LOH analysis, and their availability will vary depending on genomic region and the cell line/organism to be assayed. Chips assaying more SNPs around the locus will yield more reliable data, and assaying closer to the target site will allow detection of smaller regions of LOH. SNP microarrays can also be customized to a certain degree, which allows more SNPs in a region of interest to be assayed. We applied the Illumina Global Screening Array v.2 genotyping chip, which is an affordable option commonly used for genome-wide association studies[22] and has a reasonably dense and well-balanced genome-wide coverage (one SNP every 4.4 kb). Other higher-density chips by Illumina or other manufacturers, such as Infinium Omni2.5-8 or UK Biobank Axiom Array (Thermo Fisher Scientific), and so on, can likewise be used to further increase SNP coverage.

Most SNP arrays are designed for high-throughput genotyping—for instance, the Illumina Global Screening array processes 24 samples on one chip—and therefore it is most cost effective to analyze multiple samples in parallel. In addition, specific equipment is needed for chip analysis which makes LOH analysis by SNP microarray more expensive than nearby SNP sequencing. For these reasons, SNP chip genotyping is often outsourced to companies or core facilities. Analysis of raw chip data is performed using GenomeStudio software from Illumina, which is downloaded together with chip-specific manifest and cluster files from the manufacturer's website.

Expertise needed to implement the protocol

All steps of our protocol are optimized wherever possible for maximal accessibility and modest requirements for specialized infrastructure, and can be performed with the expertise and equipment typically available in a laboratory performing CRISPR genome editing. This includes performing and optimizing PCR, qPCR and Sanger sequencing, the latter typically being outsourced to a company or core facility. The same technologies are required to perform SNP genotyping. If LOH analysis is performed using SNP microarrays, access to specific equipment, such as pipetting robots, hybridization ovens and an iScan scanner, is required, but this can also be outsourced to a core facility or commercial provider.

Limitations

A limitation of our technology is that qgPCR relies on PCR amplification of the edited locus, which may not work for some loci due to local sequence composition, e.g., very high GC content or

repetitive sequences. However, if the edited locus is not compatible with PCR, verification of successful editing by Sanger sequencing would also be prevented, thus causing a general obstacle to validating CRISPR editing. Furthermore, qgPCR is not an optimal OnTE detection method if very large edits are desired, such as kilobase-sized insertions, because the PCR could not span the entire edit with primers in the flanking genomic regions. As discussed above, careful redesign and validation for PCR efficiencies may still allow adapting qgPCR to such a scenario.

Another potential limit of our assays is that they will not identify chromosomal alterations that occur not directly at the target site and therefore do not affect the amplicon of the qgPCR but are also too small to be detected by SNP sequencing or microarrays. For example, translocation of a small region containing the entire qgPCR amplicon to another genomic position or large inversions would not affect copy number values of qgPCR and zygosities in SNP microarrays. However, as the genomic cut by Cas9 is the initial trigger for the formation of OnTEs, we expect that the large majority of OnTEs will affect the immediate region around the target site, and therefore events escaping detection would be extremely rare.

qgPCR and SNP genotyping is a crucial quality-control measure to increase reliability of CRISPR-based studies by excluding edited cells and organisms with OnTEs. However, it is important to keep in mind that although our tools are useful to exclude OnTEs, other editing-induced issues such as off-target effects can still be present in the edited cell or organism and therefore should also be investigated carefully (see Box 2).

Application and limitations of the method in animals

Although we have tested our OnTE detection technology only in human iPSCs, its general principles also allow OnTE detection to be performed in edited animals. However, its full applicability and usefulness depends on several factors that differ between edited cells and animals, necessitating careful consideration.

First, our protocol requires clonality of analyzed organisms, which is often not achieved in edited animals. The F0 generation of many model organisms may be mosaic after genome editing due to editing in multicellular stages and/or multiple parallel editing events. Assaying mosaic samples with a mixture of different genomes is complex and requires more sophisticated methods than qgPCR/bulk SNP genotyping, such as targeted deep sequencing. For this reason, prevention of unwanted OnTEs is currently mainly performed by backcrossing of edited animals, which separates alleles with OnTEs from correctly edited alleles. However, genome-editing protocols are constantly becoming more versatile, and there is increasing interest in using edited animals already in the F0 generation (i.e., the animal edited as an embryo/zygote is directly assayed) to generate more rapid experimental results. It has been shown that mosaicism can be reduced by performing editing in one-cell-stage embryos with CRISPR ribonucleoproteins (RNPs) instead of plasmids or RNA[23] or by accelerated proteasomal degradation of Cas9 (ref. [24]). Li et al.[25] recently reported a dramatic decrease in genetic mosaicism by spatiotemporal control of Cas9 activity in mouse embryos, leading to the generation of identical biallelic F0 mutants. As a proof of concept, especially for species with long generation lengths and/or costly husbandry, they achieved high frequency of identical biallelic editing with little mosaicism. It is therefore conceivable that further improvements in editing and preimplantation technologies will lead to wider use of edited animals directly in the F0 generation, making our OnTE detection protocol a useful addition to quality-control measures.

Second, testing for LOH is only relevant in cell lines and animals with a variable genetic background. Whereas most human cell lines bear the genetic variability of their donors, inbred animals and cell lines derived from them, which are often used as model organisms in laboratories, are considered to be genetically identical and homozygous at every SNP (apart from spontaneous mutations). Therefore, LOH cannot occur, which makes testing for it unnecessary. However, if genetically variable outbred animals (such as many common rat models or diversity outbred mice[26], which mimic human genetic diversity) are used, our technology can be applied to exclude LOH after genome editing, regardless of the species.

Materials

Biological materials

! CAUTION For work with iPSCs informed consent must be obtained from the respective subjects. Studies with iPSCs or animals must conform to all relevant institutional and governmental regulations.

! CAUTION Cell lines should be routinely checked for authenticity and contamination with mycoplasma.

- Female iPSC line A18944 (ThermoFisher Scientific, cat. no. A18945; RRID: CVCL_RM92)

Reagents

Genomic DNA extraction

- PBS (Sigma, cat. no. D8537)
- Ethanol absolute (VWR, cat. no. 20821.310)
- NucleoSpin Tissue Kit (Machery-Nagel, cat. no. 740952.50)

PCR and gel electrophoresis

- Custom primers designed to amplify a 300–450 bp region around the edited locus (IDT, see Supplementary Table 1 for primer sequences used in our experiments)
- OneTaq Quick-Load 2× Master Mix with standard Buffer (NEB, cat. no. M0486S)
- UltraPure, DNase/RNase-Free Distilled Water (Invitrogen, cat. no. 10977-035)
- Agarose SERVA Wide Range (SERVA, cat. no. 11406.01)
- TAE buffer, 50× (Thermo Fisher Scientific, cat. no. B49)
- SYBR Safe DNA Gel Stain (Invitrogen, cat. no. S33102) **! CAUTION** SYBR Safe has possible mutagenic affects. Wear full personal protective equipment (PPE) and handle with care in a designated separate working area.
- Gene ruler 100 bp plus DNA ladder (Thermo Fisher Scientific, cat. no. SM0323)

qgPCR

- PrimeTime Gene Expression Master Mix (IDT, cat. no. 1055770)
- Human TaqMan Copy Number Reference Assays (human TERT on chromosome 5, Thermo Fisher Scientific, cat. no. 4403316, or human RNase P on chromosome 14, Thermo Fisher Scientific, cat. no. 4403326) **▲ CRITICAL** Reference gene assays should target a locus on a different chromosome than the edited locus to exclude problems due to large chromosomal aberrations on the edited chromosome.
- Mouse TaqMan Copy Number Reference Assays (mouse Tfrc on chromosome 16, Thermo Fisher Scientific, cat. no. 4458366, or mouse Tert on chromosome 13, Thermo Fisher Scientific, cat. no. 4458368) **▲ CRITICAL** Reference gene assays should target a locus on a different chromosome than the edited locus to exclude problems due to large chromosomal aberrations on the edited chromosome.
- Fluorescently labeled probe for target-specific qgPCR (see Supplementary Table 1 for probe sequences used in our experiments), e.g., designed using the IDT PrimerQuest tool and ordered from IDT as PrimeTime Eco Probes 5′ 6-FAM/ZEN 3′ IBFQ (FAM label on the 5′ end and a double-quencher with ZEN/Iowa Black FQ at the 3′ end). The 'Eco' scale of 2.5 nmol is sufficient for a maximum of 668 single qgPCR reactions as described in step 12 **▲ CRITICAL** The design of the probe must not overlap with the cut site or intended mutation; see Procedure for details.

Nearby SNP sequencing

- Custom PCR primers amplifying a 300–500 bp region flanking SNPs derived from BioMart (IDT, see Supplementary Table 1 for primer sequences used in our experiments)

SNP microarray

- Infinium Global Screening Array-24+ v.2 Kit (48 samples, Illumina, cat. no. 20030773)
- NaOH, 0.1 M (Merck, cat. no. 43617) **! CAUTION** NaOH is highly corrosive when concentrated. Wear appropriate PPE and work in a chemical fume hood.
- 2-Propanol (Merck, cat. no. I9516)
- Formamide (Carl Roth, cat. no. 6749.1) **! CAUTION** Formamide is toxic; use under a fume hood and wear full PPE.
- EDTA, 500 mM (Merck, cat. no. 20-158)

Equipment

General laboratory consumables

- PCR plate, 96 well, semi-skirted (BRAND, cat. no. 781374) **▲ CRITICAL** For qgPCR, use plates that fit the qPCR machine (as recommended by the manufacturer).
- LightCycler 480 Sealing Foil (Roche, cat. no. 04729757001)
- NucleoSpin Gel and PCR clean-up kit (Machery-Nagel, cat. no. 740609.50)
- 8-Strip PCR tubes with domed lids (Biozym, cat. no. 711047)
- Standard 1.5 ml centrifuge tubes (Eppendorf, cat. no. 0030 120.086)
- Standard 2.0 ml centrifuge tubes (Eppendorf, cat. no. 0030 120.094)

- Sterile pipette tips with filters, 2.5 µl (Sarstedt, cat. no. 70.1130.212)
- Sterile pipette tips with filters, 200 µl (Sarstedt, cat. no. 70.760.211)
- Sterile pipette tips with filters, 1,250 µl, long (Sarstedt, cat. no. 70.1186.210)
- Storage plate, 0.8 ml (MIDI plate), conical well-bottom (Abgene, cat. no. AB-0765)
- 96-Well cap mats (Abgene, cat. no. AB-0566) ▲ **CRITICAL** Ensure that plates and cap mats fit heating blocks and pipetting robots.
- Heat sealing foil (Abgene, cat. no. 0559)
- Tape Pads adhesive foil (Qiagen, cat. no. 19570)
- Cell lifter (Corning, cat. no. CLS3008)
- Qubit Assay Tubes (Thermo Fisher Scientific, cat. no. Q32856)
- Qubit dsDNA BR Assay Kit (Thermo Fisher Scientific, cat. no. Q32850)

General laboratory equipment

- ProFlex 96-well PCR System (Applied Biosystems, cat. no. 4484075)
- Microcentrifuge (Eppendorf, cat. no. 5417R)
- PeqPOWER 300 V power supply (Peqlab, cat. no. 55-E300-230V)
- Horizontal gel electrophoresis chamber (Peqlab, cat. no. 40-1214)
- ThermoMixer C (Eppendorf, cat. no. 5382000023) and Thermoblock (Eppendorf, cat. no. 5361000031)
- Gel imaging system (Bio-Rad, cat. no. 1708265)
- Pipette, 0.1–2.5 µl (Eppendorf, cat. no. 3120000011)
- Pipette, 2–20 µl (Eppendorf, cat. no. 3120000038)
- Pipette, 20–200 µl (Eppendorf, cat. no. 3121000082)
- Pipette, 100–1,000 µl (Eppendorf, cat. no. 3121000120)
- StepOnePlus Real-Time PCR System (Thermo Fisher Scientific, cat. no. 4376600)
- NanoDrop 2000c Spectrophotometer (Thermo Fisher Scientific, cat. no. ND-2000C)
- Qubit 4 Fluorometer (Thermo Fisher Scientific, cat. no. Q33238)

Processing of Illumina chips

- Tecan eight-tip robot (Illumina, cat. no. SC-30-402)
- Infinium HD Starter Kit (24 beadchip) 220V (Illumina, cat. no. 20028879)
- iScan scanner with software (Illumina, cat. no. SY-101-1001, https://emea.illumina.com/systems/array-scanners/iscan.html, user guide: https://support.illumina.com/content/dam/illumina-support/documents/documentation/system_documentation/iscan/iscan-system-guide-11313539-01.pdf)
- Cooling microplate centrifuge with adapters for 0.8 ml MIDI plates and tubes (Sigma, cat. no. 4K15C)

Software

- SnapGene Viewer v.4.3.10 (SnapGene, https://www.snapgene.com/snapgene-viewer/)
- Illumina Decode File Client (Illumina, https://emea.support.illumina.com/array/array_software/decode_file_client/downloads.html)
- GenomeStudio v.2.0.5 with genotyping module v.2.0.5. (Illumina, https://emea.illumina.com/techniques/microarrays/array-data-analysis-experimental-design/genomestudio.html)
- iScan Control software v.4.0.0 (Illumina, https://emea.support.illumina.com/downloads/iscan-control-software-release-notes.html)

Reagent setup

Preparation of qgPCR probe

Briefly centrifuge the tube at 500*g* for 10 s at room temperature and resuspend in TE buffer pH 8.0 to a final concentration of 15 µM. Probe is light sensitive and should always be stored in dark conditions at −20 °C, where it is stable for up to 2 years.

Preparation of qgPCR assay

Mix assay components to a final concentration of 10 µM (each primer) and 5 µM (probe). Prepare 30 µl aliquots and store reconstituted assay in dark conditions at −20 °C, where it is stable for up to 2 years.

TAE buffer solution for electrophoresis

Dilute 50× TAE buffer to a 1× working solution using deionized water.

Agarose gels for electrophoresis

Prepare 2% (wt/vol) agarose solution in 1× TAE buffer. Microwave solution in a shatterproof glass container without the lid until agarose has dissolved. Add 10,000× SybrSafe and pour the liquid into a gel chamber with the desired combs. Gels can be used when solid or stored at 4 °C for 1 week.

Procedure

CRISPR editing in cells or organism of interest ● Timing 1–3 months, depending on cell type and organism

1 Perform genome editing in cells or organisms of interest according to published efficient CRISPR protocols (e.g., refs. [27–31]). Briefly, this includes designing a suitable guide RNA (gRNA) for the targeted locus and expressing the guide RNA together with WT Cas9, a Cas9 variant or a related other Cas nuclease in the cell line or organism to be edited, either by plasmid, mRNA or as RNP.

gDNA extraction ● Timing 2 h

▲ **CRITICAL** Subsequent steps require high-quality gDNA. Crude DNA extracts from rapid protocols lacking proteinase K digest will not work reliably.

▲ **CRITICAL** Stored gDNA samples often show varying integrities, which is problematic for the following qgPCR. We therefore recommend simultaneous harvesting and gDNA extraction of all samples, including controls (see step 11).

2 *Harvest cells for gDNA extraction*. For edited single-cell clones, follow option A; for tissues of edited mice or other organisms, follow option B.

(A) **Harvest and gDNA extraction of edited single-cell clones**

(i) After genome editing, expand single-cell clones in tissue culture to obtain 500,000–1,000,000 cells for gDNA extraction. Wash attached cells once with PBS without dislodging them from the plate, aspirate and add 1 ml of fresh PBS for harvesting. Simultaneously harvest unedited control sample for gDNA extraction.

(ii) Scrape cells off the plate using a cell lifter and triturate to remove clumps; transfer the cells to a 1.5 ml Eppendorf tube.

(iii) Pellet cells by centrifugation at 1,000*g* for 2 min at room temperature (20–25 °C). Aspirate PBS and place cell pellets on ice.

■ **PAUSE POINT** Cell pellets can also be stored at −80 °C for up to 1 year.

(iv) Extract gDNA from each sample, including controls, using a gDNA extraction kit, such as NucleoSpin Tissue (Macherey-Nagel), according to the manufacturer's instructions.

(B) **Harvest and gDNA extraction of edited organisms**

(i) Collect animal tissue (e.g., 5 mm of mouse tail, 10–25 mg of organ or ear/tail clip biopsy).

▲ **CRITICAL STEP** Fresh tissue should be processed as soon as possible to prevent degradation of gDNA. Until then, store tissue at −20 or −80 °C.

(ii) Extract gDNA from each sample, including controls, using a gDNA extraction kit, such as NucleoSpin Tissue (Macherey-Nagel), according to manufacturer's instructions.

Design and optimization of 300–450 bp genotyping PCR ● Timing 1 d

▲ **CRITICAL** All genotyping and qgPCR reactions require an optimal PCR product without additional bands or low product yield. This is achieved by selecting an optimal primer pair from several primer candidates as well as optimizing yield and annealing temperature using gradient PCR.

3 Locate edited locus, e.g., by BLASTing guide RNA sequences, and export ~300 bp on each side of the target locus from Ensembl (http://www.ensembl.org) or another genome browser. Paste the sequence into a primer design tool, such as Primer3Plus (https://primer3plus.com/), and use default settings to select primers. The intended PCR product should amplify a 300–450 bp region with the target locus placed near the middle (distance between primer and target site should be at least 100 bp). Order three different forward and reverse primers.

▲ **CRITICAL STEP** In case of HDR-mediated editing, ensure that at least one primer lies outside of the repair template to avoid misleading results in the following qgPCR reaction due to residual template still being present in the cells after editing.

4 Prepare the following PCR reaction using the primers to amplify the target locus from gDNA of the unedited 'parent' cells or organism. We recommend testing all nine possible primer combinations

at different annealing temperatures (see next step). For this, prepare a master mix for each combination with individual 10 µl PCR reactions.

Components	Amount	Final concentration
Genomic DNA	10 ng	1 ng/µl
2× OneTaq Quick-Load Master Mix	5 µl	1×
Forward primer (10 µM)	0.2 µl	0.2 µM
Reverse primer (10 µM)	0.2 µl	0.2 µM
DNase-free H_2O	To 10 µl	—

5 Run a gradient PCR using the following thermocycler program with annealing temperatures varying between 50 and 70 °C in 4 °C steps (six temperatures for each combination).
▲ **CRITICAL STEP** NEB recommends an extension temperature of 68 °C for OneTaq Master mix. If a different polymerase enzyme or master mix is used for the PCR reaction, thermocycling conditions might need to be adjusted according to manufacturer's instructions.

Step	Temperature (°C)	Time	Number of cycles
1. Initial denaturation	94	5 min	1×
2. Denaturation	94	30 s	35×
3. Annealing	50–70	30 s	
4. Extension	68	30 s	
5. Final extension	68	5 min	1×
6. Hold	10	Until ready to process	1×

6 *Analyze PCR products by agarose gel electrophoresis*. Load 5–10 µl of each sample together with DNA ladder on a 2% (wt/vol) agarose gel with SYBR Safe (1×). Run gel at 150 V to separate DNA bands until the loading dye front migrates through ~75% of the gel. Visualize bands using an appropriate gel imager. Determine the best temperature and primer combination that yields a single, strong PCR product.
? TROUBLESHOOTING

7 Scale up all reagents to a total volume of 50 µl and repeat the reaction for the best PCR product. Purify the reaction using a PCR clean-up kit, such as NucleoSpin Gel and PCR Clean-Up (Macherey-Nagel). Sanger sequence the product to confirm that it matches the sequence obtained from the genome browser.
▲ **CRITICAL STEP** Genomic sequence of target cells may differ from genome databases because of SNPs or database errors. Always confirm that the correct sequence is used for designing all genome-editing and quality-control reagents.
? TROUBLESHOOTING

Identification of cells or organisms with desired edit by restriction fragment length polymorphism and Sanger sequencing ● Timing 2–3 d

8 Identify clones or organisms with desired edit by amplifying the edited region with the above-optimized PCR, analyzing PCR products by gel electrophoresis and Sanger sequencing. If a large number of clones or founder animals needs to be screened and the edit introduces a restriction enzyme site, a restriction fragment length polymorphism assay (as described, e.g., in refs. [27,28]) can also be used for screening of positive samples.
▲ **CRITICAL STEP** It is important to analyze genotyping PCR products by agarose gel electrophoresis before Sanger sequencing. This will help to exclude clones with small deletions or insertions that occurred within the genotyping amplicon, which will lead to double bands or size-shifted bands on a gel.

Design of labeled oligo for qgPCR ● Timing 1 h

9 Enter the genomic sequence surrounding the target locus into a qPCR design tool, such as IDT PrimerQuest. Enter the available primer sequences from the optimized genotyping PCR (Step 6) under 'partial design input' and leave default settings for other parameters.

▲ **CRITICAL STEP** To avoid overlap with edited bases that would affect the assay, enter the edited site as an excluded region for the probe. If NHEJ-edited cells or organisms need to be analyzed, exclude 5–20 bp on each side of the cut site, depending on the maximal size of indels that are desired.

▲ **CRITICAL STEP** Default design parameters and outcomes for qPCR probes and corresponding primers can vary for other design tools, such as Primer3Plus or Primer-BLAST. We therefore recommend checking amplification curves of the qPCR reactions (Step 16) and single bands by agarose gel electrophoresis (Step 19).

10 Order fluorescently labeled probes for target-specific qgPCR with modifications fitting the specifications of your qPCR machine.

▲ **CRITICAL STEP** The copy number reference assays we suggest (e.g., TERT for humans and Tfrc for mice) are labeled with VIC dyes. To enable duplex qPCR reactions (i.e., two targets are amplified in a single well) to be performed, it is necessary that probes for the gene of interest and reference gene are labeled with two different colors. Compatibility of colors with the qPCR machine needs to be checked individually, but typically, a FAM-dye-labeled probe is used together with a VIC-dye-labeled probe. In both cases, a quencher molecule on the probe prevents the dye from emitting a signal until it is cleaved off and hence activated by the polymerase's exonuclease activity.

Analysis of edited cells or organisms by qgPCR to exclude those with OnTEs ● Timing 1 d

▲ **CRITICAL** Avoid contamination of your samples and reagents with foreign nucleic acids by using filter tips for pipetting throughout the qPCR experiment. We also recommend using a dedicated space and pipettes and, if available, a ultraviolet-treated clean hood for pipetting the qgPCR.

▲ **CRITICAL** Include one unedited control gDNA sample and one no-template (only water) control. If edited cell lines/organisms from previous editing rounds have already been characterized by qgPCR and yielded samples with only one allele at the same targeted locus, fresh gDNA from these samples can be added as an additional positive control.

11 Determine DNA concentration of gDNA samples using a spectrophotometer. Prepare 9.5 µl of a 10 ng/µl dilution of gDNA with DNase-free H_2O for triplicate reactions (see step 12) in a new microcentrifuge tube (or 6.5 µl for duplicate reactions).

▲ **CRITICAL STEP** Accurate determination of sample concentration is important for the subsequent qPCR reaction. A spectrophotometer is usually sufficient, but if desired a more precise measurement using a Qubit fluorometer can be performed.

12 *Prepare the following qgPCR reaction master mix.* Check the PrimeTime Gene Expression Master Mix instructions or the manual of your qPCR machine if a reference dye needs to be added. Thaw all components on ice and prepare a master mix for all samples without gDNA. Briefly vortex to mix all reagents thoroughly. We recommend running each sample in triplicate. Volumes provided are for one reaction in one well; scale up all components as necessary and add ~10% extra volume to account for small pipetting errors.

Components	Amount (µl)	Final concentration
2× PrimeTime Gene Expression Master Mix	7.5	1×
20× qgPCR assay (see 'Reagent setup' section)	0.75	1×
20× TaqMan Copy Number Reference Assay	0.75	1×
DNase-free H_2O	3	—

▲ **CRITICAL STEP** Analyzing samples in three technical replicates is helpful to identify outliers. For more experienced users of qPCR experiments, it might be sufficient to analyze samples in duplicate to save material cost.

13 For triplicate reactions, add 38 µl of qgPCR reaction master mix (or 26 µl for duplicate reactions) from step 12 to each sample from step 11.

14 Pipette 15 µl of qgPCR reaction mix from the previous step into each well. Seal the plate with optically transparent film suitable for qPCR reactions. Avoid fingerprints or other marks on the film. Centrifuge the plate at 1,000*g* for 1 min at 4 °C.
▲ **CRITICAL STEP** To decrease technical variability in your qPCR reaction resulting from different amounts of reagents in each well, it is extremely important to focus on a consistent pipetting technique.
■ **PAUSE POINT** The prepared qPCR plate can be stored in dark conditions at 4 °C for a few hours. Briefly vortex and centrifuge before continuing with the next step.

15 Run qgPCR reaction on a qPCR thermocycler using the following program.

Step	Temperature (°C)	Time	Number of cycles
1. Polymerase activation	95	5 min	1×
2. Denaturation	95	15 s	40×
3. Annealing/extension	60	1 min	
4. Hold	10	Until ready to process	1×

16 Examine the amplification plot of qPCR reactions. It should show a baseline at the beginning, a logarithmic amplification phase (this phase is usually within qgPCR reaction cycles 20 to 30 and includes the Ct value; see Fig. 3c,d) and a plateau toward the end of the run. Discard samples with abnormal amplification plots from the analysis. Detailed discussion on baselines and amplification plots can be found elsewhere[16].

17 *Analyze C_t values of samples*. Obtain C_t values from qPCR machine software using auto baseline settings and export values to an Excel file. Calculate mean C_t values from technical replicates and standard deviation. The 'no template control' should not show any target amplification and the standard deviation should not be >0.2.

From our experience, C_t values of qgPCR assays for target genes may vary but are usually ~25–26. C_t values for TERT are usually between 27 and 28. Variations can be caused by different amounts of gDNA between wells or different efficiencies of PCR reactions. Low amplicon levels will result in high C_t values with a greater chance of high variation between samples. Therefore, a C_t value threshold of 30 may be used to prevent unreliable results[32].
? TROUBLESHOOTING

18 For relative quantification calculate $\Delta\Delta C_t$ by normalizing target gene values to the internal reference gene and unedited control sample using the following equation (with mean C_t values for each):

$$\Delta\Delta C_t = \left(C_t\,\text{mean}_{\text{target}} - C_t\,\text{mean}_{\text{reference}}\right)_{\text{edited sample}} - \left(C_t\,\text{mean}_{\text{target}} - C_t\,\text{mean}_{\text{reference}}\right)_{\text{unedited control}}$$

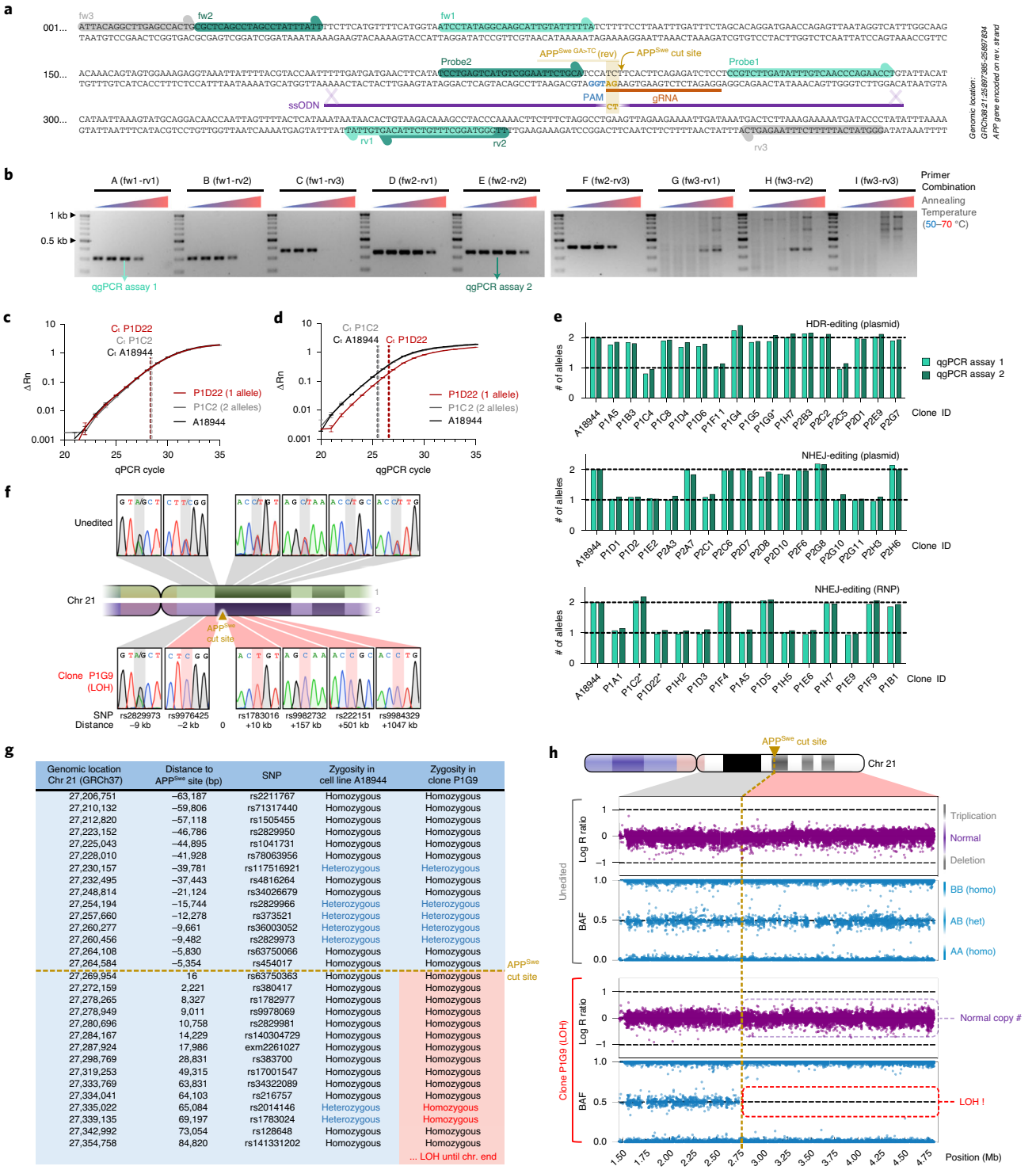
To get the fold difference, calculate $2^{-\Delta\Delta Ct}$. Multiply values by 2 to get the total number of alleles. Edited samples should either have two alleles (i.e., no OnTEs) or one allele (i.e., OnTEs detected).
? TROUBLESHOOTING

19 Analyze qgPCR products by agarose gel electrophoresis. Load one entire 15 µl reaction together with 2 µl DNA ladder on a 2% (wt/vol) agarose gel with SYBR Safe (10,000×). Run gel at 150 V to separate DNA bands until the loading dye front migrates through ~75% of the gel. Visualize bands using an appropriate gel imager.
▲ **CRITICAL STEP** All analyzed samples should have one single band from target locus amplification and one band from the reference assay (e.g., human TERT at 88 bp or mouse Tfrc at 91 bp).
? TROUBLESHOOTING

20 Edited single-cell clones/animals with two alleles in the qgPCR reaction and one single PCR product on the agarose gel do not contain OnTEs such as large deletions, large insertions or other complex rearrangements. Exclude all other clones or animals.

Select methods to exclude LOH ● Timing 1 h

21 After qgPCR, edited single-cell clones or organisms should be subjected to LOH analysis. For this purpose, we describe two distinct options in the following protocol: nearby SNP genotyping and SNP microarrays. If unsure which one to select, consider our discussion of the advantages and disadvantages in Box 1 or perform both assays in parallel.



Genomic location Chr 21 (GRCh37)	Distance to APPSwe site (bp)	SNP	Zygosity in cell line A18944	Zygosity in clone P1G9
27,206,751	−63,187	rs2211767	Homozygous	Homozygous
27,210,132	−59,806	rs71317440	Homozygous	Homozygous
27,212,820	−57,118	rs1505455	Homozygous	Homozygous
27,223,152	−46,786	rs2829950	Homozygous	Homozygous
27,225,043	−44,895	rs1041731	Homozygous	Homozygous
27,228,010	−41,928	rs78063956	Homozygous	Homozygous
27,230,157	−39,781	rs117516921	Heterozygous	Heterozygous
27,232,495	−37,443	rs4816264	Homozygous	Homozygous
27,248,814	−21,124	rs34026679	Homozygous	Homozygous
27,254,194	−15,744	rs2829966	Heterozygous	Heterozygous
27,257,660	−12,278	rs373521	Heterozygous	Heterozygous
27,260,277	−9,661	rs36003052	Heterozygous	Heterozygous
27,260,456	−9,482	rs2829973	Heterozygous	Heterozygous
27,264,108	−5,830	rs63750066	Homozygous	Homozygous
27,264,584	−5,354	rs454017	Homozygous	Homozygous
27,269,954	16	rs63750363	Homozygous	Homozygous
27,272,159	2,221	rs380417	Homozygous	Homozygous
27,278,265	8,327	rs1782977	Homozygous	Homozygous
27,278,949	9,011	rs9978069	Homozygous	Homozygous
27,280,696	10,758	rs2829981	Homozygous	Homozygous
27,284,167	14,229	rs140304729	Homozygous	Homozygous
27,287,924	17,986	exm2261027	Homozygous	Homozygous
27,298,769	28,831	rs383700	Homozygous	Homozygous
27,319,253	49,315	rs17001547	Homozygous	Homozygous
27,333,769	63,831	rs34322089	Homozygous	Homozygous
27,334,041	64,103	rs216757	Homozygous	Homozygous
27,335,022	65,084	rs2014146	Heterozygous	Homozygous
27,339,135	69,197	rs1783024	Heterozygous	Homozygous
27,342,992	73,054	rs128648	Homozygous	Homozygous
27,354,758	84,820	rs141331202	Homozygous	Homozygous
				... LOH until chr. end

APPSwe cut site

Identification of nearby heterozygous SNPs by Ensembl BioMart and validation in unedited cells or organism ● Timing 2–4 d

22 Open Ensembl BioMart (https://www.ensembl.org/info/data/biomart/index.html) and select the Ensembl Variation 102 database with the respective dataset for your species, e.g., 'Human Short Variants (SNPs and indels excluding flagged variants) (GRCh38.p13)'. Specify filters to define a region of 10 kb around the cut site and a global MAF of ≥0.3 under general variant filters.

◀ **Fig. 3 | Anticipated results for OnTE detection—example of CRISPR editing at the APP^{Swe} locus in human iPSCs. a**, Overview of CRISPR editing and positions of primers and probes for two independent qgPCR assays at the APP^{Swe} locus. fw, forward; rv, reverse. **b**, PCR optimization at the APP^{Swe} locus with combinations of three forward and three reverse primers at increasing annealing temperatures (50–70 °C in 4 °C steps), analyzed on a 2% (wt/vol) agarose gel. Two primer combinations (A and E) were selected for the qgPCR assays 1 and 2 shown in **e**. **c**, qPCR amplification plot of human TERT reference gene for two NHEJ-/RNP-edited clones and the unedited 'parent' line A18944. (ΔRn = normalized reporter fluorescence signal – baseline). Data are represented as means ± s.e.m. ($n = 2$). **d**, qPCR amplification plot of APP^{Swe} qgPCR assay 1 for two NHEJ-/RNP-edited clones and the unedited A18944 line. Note shift in ΔRn for clone P1D22 indicating an OnTE. Data are represented as means ± s.e.m. ($n = 2$). **e**, qgPCR analysis with two independent assays reveals clones with decreased allele copy number for HDR-clones edited by plasmid delivery of editing components (top), NHEJ clones with plasmid delivery (middle) and NHEJ clones with RNP delivery (bottom). Edited clones were genotyped beforehand to confirm insertion of the APP^{Swe} mutation for HDR editing or generation of putative homozygous indels for NHEJ editing. Values were normalized to the unedited 'parent' cell line A18944. Highlighted clones (*) also shown in **c** and **d** (P1C2 + P1D22) or **f**–**h** (P1G9). **f**, Sanger sequencing traces from nearby SNP sequencing for unedited line and APP^{Swe} LOH clone P1G9. **g**, Overview of ±15 SNPs analyzed around the APP^{Swe} locus by the Illumina Global Screening Array v.2 genotyping chip and zygosities in unedited cell line A18944 and LOH clone P1G9 (HDR-plasmid editing). **h**, SNP microarray analysis determining log R ratios and B-allele frequencies in control and APP^{Swe} clone P1G9 (HDR-plasmid editing) reveals LOH from the cut site to the end of the long arm of chromosome 21. iPSC line A18944 (RRID: CVCL_RM92) has been validated by fingerprinting, pluripotency assays and karyotyping, and for absence of mycoplasma. **e** (upper two graphs), **f** and **h** are modified from ref. [13].

▲ **CRITICAL STEP** The frequency of genomic variation strongly varies between different loci. We recommend performing the analysis with ~5–10 different SNPs on each side of the cut site to have a sufficient chance of identifying heterozygous SNPs in the 'parent' cell line or organism. If the above-mentioned filters do not yield enough variants, double the size of the analyzed genomic region and lower the cutoff for the global MAF by 0.1; repeat these adjustments if necessary.

23 Once the potential heterozygous SNPs are identified, download the flanking sequence around the SNP from Ensembl.

▲ **CRITICAL STEP** Ensure that the genomic coordinates from the database are identical to those used in BioMart by using the same version of the edited genome (GRCh38.p13 for human cells).

24 Using the downloaded sequence, design primers for a 300–500 bp PCR around the SNP as described in the PCR optimization above (Step 3) and assemble a 50 µl PCR reaction using gDNA from the unedited 'parent' cell line as described in the table below. We recommend testing one primer pair at each locus first and designing more primers only if necessary.

Components	Amount	Final concentration
Genomic DNA	50 ng	1 ng/µl
2× OneTaq Quick-Load Master Mix	25 µl	1×
Forward primer (10 µM)	1 µl	0.2 µM
Reverse primer (10 µM)	1 µl	0.2 µM
DNase-free H_2O	To 50 µl	

25 Run the following program on a thermocycler:

Step	Temperature (ºC)	Time	No. of cycles
1. Initial denaturation	94	2 min	1×
2. Denaturation	94	30 s	35×
3. Annealing	60 (50–70)	30 s	
4. Extension	68	1 min/1,000 bp	
5. Final extension	68	5 min	1×
6. Hold	10	Until ready to process	1×

26 Perform electrophoresis by analyzing 3 µl of the reaction mix as described above (step 6) and inspect PCR bands. If a single strong band is present at the expected size, purify the remaining 47 µl of PCR reaction using a PCR clean-up kit, such as NucleoSpin Gel and PCR Clean-Up (Macherey-Nagel), according to the manufacturer's instructions. If the band size is incorrect or more than one band is present, perform a gradient PCR with annealing temperatures between 50–70 °C. If this does not improve the PCR result, design and order new primer pairs (step 3).

? TROUBLESHOOTING

27 Sanger sequence the purified PCR product and determine the zygosity of the SNP in the unedited 'parent' line. Continue until you have identified at least one heterozygous SNP on each side of the cut site.
? TROUBLESHOOTING

LOH analysis by SNP genotyping in edited cells or organisms to exclude those with OnTEs ● Timing 1 d

28 Using the same assay, analyze edited cell lines to check if SNP zygosities have changed from heterozygous to homozygous. Loss of heterozygous SNPs indicates occurrence of LOH; exclude all clones or animals with LOH from further analysis.

Analysis of gDNA on genome-wide SNP microarray ● Timing 3 d

▲ **CRITICAL** The SNP microarray chip needs to match the edited organism and provide sufficient coverage in the edited region. Targeted content chips (e.g., exome chips, panel chips) are not suitable. We used the Illumina Global Screening Array v.2 genotyping chip for analysis in human cells, which provides reasonably dense and well-balanced genome-wide coverage at a relatively low price (about US $50). A chip suitable for mouse could be the GGP GIGA-MUGA-24, which provides more than 143,000 SNPs. In general, the number of informative SNPs depends very much on the animal strain and subspecies used. Therefore, applicability needs to be determined for each individual experiment.

29 *Check suitability of chip*. To confirm that the selected chip has sufficient coverage in the edited genomic region, download the chip manifest file detailing available variants and chromosomal positions as a CSV file from the manufacturer's webpage (e.g., ftp://webdata2:webdata2@ussd-ftp.illumina.com/downloads/productfiles/global-screening-array/v2-0/gsa-24-v2-0-A1-manifest-file-csv.zip).

30 Open the file with Excel and sort for chromosome and position to locate your editing site.

31 Check for available variants around the targeted site. If there are no biallelic variants available on the chip within ±10–50 kb of the editing site, we suggest finding a chip with better SNP coverage in the respective area.

32 *Whole-genome amplification*. Measure DNA concentration by Qubit and check integrity with 1 µl DNA on a 0.8% (wt/vol) agarose gel. Add 4 µl gDNA (30–100 ng/µl) to a 96 well 0.8 ml microplate (MIDI plate).
▲ **CRITICAL STEP** The microarray genotyping workflows of the leading chip manufacturers Illumina and Thermo Fisher Scientific require high-molecular-weight DNA as input because all protocols start with a whole-genome amplification. Accurate quantification is crucial and therefore ultraviolet measurement methods are not recommended because RNA or extraction contaminants may lead to an overestimation of the quantity. Instead, methods using fluorophores specific for double-stranded DNA, e.g., Qubit assays, give more reliable concentration data.
▲ **CRITICAL STEP** In the first SNP microarray experiment, include one unedited control gDNA sample for analysis to be able to detect changes in edited cell lines or organisms. In subsequent experiments with the same original cell line, no further controls are required.

33 Thaw MA1, MA2 and MSM buffers from the Illumina kit and mix by inverting. Preheat the Illumina hybridization oven to 37 °C.

34 Dispense 20 µl MA1 and 4 µl 0.1 N NaOH into each well using a multichannel pipette, seal the plate with adhesive foil and vortex at 1,500 r.p.m. for 1 min.

35 Incubate for 10 min at room temperature.
▲ **CRITICAL STEP** Incubation time must not exceed 12 min.

36 Remove adhesive foil, add 34 µl MA2 and 38 µl MSM into each well using a multichannel pipette with filter tips, seal the plate with a cap mat and mix by inverting. Centrifuge at 280*g* for 1 min at room temperature and incubate the plate in the Illumina hybridization oven at 37 °C for 20–24 h.

37 *Fragmentation*. Thaw the FMS tube and mix by inverting. Preheat a heat block with a MIDI plate insert to 37 °C.

38 Remove the MIDI plate from the hybridization oven and centrifuge at 280*g* for 1 min at room temperature. Remove the cap mat, add 25 µl FMS using a multichannel pipette with filter tips, seal the plate again with the same cap mat, vortex for 1 min at 1,500 r.p.m. and centrifuge at 280*g* for 1 min at room temperature.

39 Place plate on the 37 °C heat block for 1 h.
▲ **CRITICAL STEP** Incubation time must not exceed 2 h.

40 *Precipitation*. Thaw PM1 at room temperature, remove the cap mat from the MIDI plate and add 50 µl PM1 to each well with a multichannel pipette. Seal the plate with the same cap mat, vortex at 1,500 r.p.m. for 1 min and centrifuge at 280*g* for 1 min at room temperature.

41 Incubate the plate in the 37 °C heat block for 5 min followed by centrifugation at 280*g* for 1 min at room temperature. Remove the cap mat and add 155 µl 2-propanol to each well using a multichannel pipette. Seal with a new cap mat and mix by carefully inverting 10 times.

42 Incubate the plate in a refrigerator at 4 °C for 30 min.

43 Precool centrifuge to 4 °C and centrifuge the plate at 3,000*g* for 20 min at 4 °C.
▲ **CRITICAL STEP** Check if blue pellets are visible at the bottom of each well at the end of the centrifugation. If this is not the case, repeat this step.

44 Remove the cap mat and drain the liquid by inverting the whole plate. Put the plate on an absorbent pad and smack down several times to remove liquid drops. Leave the inverted plate on a tube rack for 1–2 h to air dry the pellets.
▲ **CRITICAL STEP** Keep the plate inverted all the time to avoid contamination until the pellets are dry. Do not dry pellets longer than 2 h otherwise it may be difficult to resuspend them.

45 *Resuspension*. Preheat the Illumina hybridization oven to 48 °C. Turn on the heat sealer and thaw RA1 in a 25 °C water bath. Pipette 7 ml RA1 into a reservoir.

46 From the reservoir add 23 µl RA1 to each well using a multichannel pipette, seal the plate with a heat seal foil on a heat sealer at 165 °C for 5 s and incubate the plate in the Illumina hybridization oven at 48 °C for 1 h.

47 Remove the plate from the oven, vortex at 1,700 r.p.m. for 1 min and centrifuge at 280*g* for 1 min at room temperature.
■ **PAUSE POINT** The plate can be stored at −20 °C for up to 24 h or at −80 °C for up to 3 months.

48 *Hybridization to array*. Preheat a heat block with a MIDI plate insert to 95 °C and the Illumina hybridization oven to 48 °C. Set rocker speed to 5.

49 Assemble the Illumina hybridization chambers according to the Infinium HTS Assay Reference Guide (https://support.illumina.com/content/dam/illumina-support/documents/documentation/chemistry_documentation/infinium_assays/infinium-hts/infinium-hts-assay-reference-guide-15045738-04.pdf). You need one chamber for each 96-well plate, i.e., four bead chips. Pipette 400 µl PB2 into each of the eight chamber reservoirs, close the chamber with the lid and leave it until chips are loaded with DNA.

50 Place the plate from step 49 on the 95 °C heat block for 20 min, let it cool down on the bench for 30 min and centrifuge at 280*g* for 1 min at room temperature. Remove the heat seal.

51 Place four Illumina GSA bead chips on two alignment fixtures and cover the fixtures with the Robot Tip Alignment Guide-G.

52 Place the MIDI plate and the prepared alignment fixtures on the pipetting robot.

53 Start the program Illumina Automation Control and select 'MSA3-Hyb Multi BC2'. Pipetting one 96- well plate onto four bead chips takes ~25 min.

54 Place each bead chip in the prepared Illumina hybridization chamber. Cover the chamber with the lid by closing all four clamps. You will need four bead chips for 96 samples.

55 Place the hybridization chamber in the prepared Illumina hybridization oven and incubate at 48 °C for 16–24 h.

56 Prepare the XC4 reagent by adding 330 ml absolute ethanol. Shake the bottle vigorously for 15 min and leave it on the laboratory bench at room temperature overnight.

57 *Washing, extension, staining*. Thaw LX1, X2, SML, ATM and EML tubes (one tube each for four bead chips). Prepare the 95% (vol/vol) formamide/1 mM EDTA solution by mixing 23.75 ml formamide, 1.2 ml water and 50 µl 0.5 M EDTA.
! **CAUTION** Formamide is toxic! Wear full PPE.

58 For four bead chips fill one Multi-sample BeadChip alignment fixture with 150 ml PB1 and place four black frames into it. Fill two wash dishes with 200 ml PB1 each.

59 Remove the hybridization chamber from the oven and let it cool on the laboratory bench for 30 min before opening.

60 Place LX1, X2, SML, ATM and EML tubes in the pipetting robot and remove all caps. Place three reservoirs with 15 ml 95% (vol/vol) formamide/1 mM EDTA, 10 ml RA1 and 50 ml XC3 in the pipetting robot.

61 Turn on the water circulator and set the temperature to 44 °C.

62 Submerge the wash rack in one wash dish with 200 ml PB1. Open the hybridization chamber, remove one chip at a time, remove the seal on the chip starting with a corner on the barcode end and immediately place the chip into the wash rack in PB1.

63 After all chips are in the wash rack, move the rack up and down for 1 min. Move the wash rack to the second wash dish with 200 ml PB1 and repeat.

64 Place each chip into one black frame from the Multi-Sample BeadChip Alignment Fixture from step 58. Place a clear LCG spacer onto the top of each chip and place the alignment bar onto the fixture. Place a glass back-plate on each chip and attach the metal clamps to the flow-through chambers.

65 Remove the flow-through chambers from the fixture and trim the ends of the spacers with scissors.

▲ **CRITICAL STEP** Do not let the bead chips dry. Bead chips should always be in PB1.

66 Start the program Illumina Automation Control and select 'XStain Tasks | XStain LCG BeadChip'. Enter the number of bead chips and make sure that you placed all items properly on the robot bed.

67 Click 'Run'. Enter the stain temperature indicated on the SML tube and when the temperature probe registers 44 °C, click 'OK'.

68 When prompted, place each assembled flow-through chamber in the chamber rack and click 'OK'. The entire program takes 2–3 h.

69 Fill 310 ml PB1 in a wash dish and submerge the staining rack. Remove the assembled flow-through chambers from the robot, disassemble each flow-through chamber, remove the spacer and place the bead chip in the staining rack. Move the rack slowly up and down 10 times and leave it in PB1 for another 5–30 min.

70 Fill 310 ml XC4 from step 56 in a wash dish. Immediately remove the staining rack from the PB1 dish, place it directly into the XC4 dish, move the rack slowly up and down ten times and leave it in XC4 for additional 5 min.

71 Remove the staining rack from the XC4 dish and place it on a tube rack. Remove the bead chips from the staining rack and put them on another tube rack.

72 Put all bead chips into a vacuum desiccator, apply the vacuum and dry for 60–90 min. Remove each bead chip and clean the underside with a wipe soaked with ethanol.

▲ **CRITICAL STEP** The bead chips are now ready for scanning.

■ **PAUSE POINT** The bead chips can be stored in the dark at room temperature for several weeks before scanning.

73 *Scanning*. Download dmap files from Illumina with the Decode File Client. There is one dmap file folder per bead chip.

74 Turn on the iScan scanner. Warm up for 30 min and place up to four bead chips onto one carrier.

75 Start iScan Control software and click 'Start'. The iScan tray opens. Place the carrier onto the tray and select 'Scanner | CloseTray'. Click 'Next'. Use the Infinium LCG scan setting. Scanning time is ~30 min for one bead chip.

76 Data are stored in the ScanData folder. There is one folder per bead chip. For data analysis you will need the .idat files. There are 48 .idat files per bead chip.

Processing of microarray data and analysis of log R ratios and B allele frequencies to exclude cells or organisms with OnTEs ● Timing 2 h

▲ **CRITICAL** The following steps are valid for all Illumina genotyping chips. Procedures for Affymetrix chips might differ.

▲ **CRITICAL** Ensure that in the regional settings of your operating system, the decimal symbol is set to '.'.

77 Before starting the chip analysis, acquire the following chip-specific files from the Illumina support center webpage (e.g., https://emea.support.illumina.com/downloads/infinium-global-screening-array-v2-0-product-files.html for GSAv2.0):

- Bead pool manifest files (.bpm) that contain information about SNP/probe content on the BeadChip.
- Cluster files (.egt or .egtp), which have reference cluster information for each interrogated locus.

▲ **CRITICAL STEP** Please verify that you have downloaded the correct files for the chip, as there will be differences, even between different versions of the same chip.

78 After genotyping, acquire the following files from the chip analysis:

- Raw intensity data files (.idat files).
- Sample sheets (.csv files that contain sample information, such as plate ID, cell ID, gender, and so on; open with a text editor (not Excel) and confirm that file paths fit to the location of the files on your computer).

79 Open GenomeStudio and start a new Genotyping project. Genome studio will ask for a Project Repository folder in which to save analysis files and a project name; the sample sheet and raw data files should be provided in a Data Repository folder, and the bead pool manifest file in a Manifest Repository folder. In the next dialog box, give the location of the cluster file from Step 77, and 'import cluster positions'.

80 Choose to 'Calculate Sample and SNP Statistics' under 'Project Creation Actions' and click 'Finish'. Automatic clustering will be performed by GenomeStudio during the manual data loading step.
? TROUBLESHOOTING

81 Go to 'samples table' (usually lower left in GenomeStudio) and confirm the gender of your samples by right-click and 'Estimate Gender'.
▲ CRITICAL STEP If there is a gender mismatch, review the sample sheet and confirm that the samples were matched correctly. If the error persists, samples may have been mixed up.

82 Exclude samples with <95% call rate and select 'Analysis/Update SNP statistics'.

83 Select the 'SNP Table' tab and filter based on 'Call Freq' > 0.95.

84 In the 'Full Data Table' tab filter SNPs (filter rows) based on the chromosome of interest.
? TROUBLESHOOTING

85 Add Filter to display only SNPs with GenTrain score >0.7.
? TROUBLESHOOTING

86 Add Filter 'Name lacks cnv' to display only biallelic SNPs.

87 In 'Column chooser', for each sample, display 'B Allele Freq', 'Log R Ratio' and 'GType'.

88 Generate a scatter plot using the ScatterPlot function. Select 'Position' as the *x*-axis and, after selecting individual sample in '*Columns*', select 'B Allele Freq' as the *y*-axis in 'Sub columns'. The resulting image displays frequencies of A and B alleles along the selected chromosome. Regions of potential LOH are characterized by long stretches of homozygous SNPs.

89 Generate a scatter plot using the ScatterPlot function. Select 'Position' as the *x*-axis and 'Log R ratio' as the *y*-axis for each sample of interest. The resulting image displays copy number information about all assayed SNPs along the selected chromosome. If regions of LOH have normal copy numbers ~0 to ±0.25, the aberration is a LOH; if copy number is increased or decreased, the area is duplicated or deleted, respectively.

90 Select 'export displayed data to a file'. Select all visible rows.

91 The exported tab-separated file can be loaded into Excel or statistical software for further processing.

Troubleshooting

Troubleshooting advice can be found in Table 1.

Table 1 | Troubleshooting table

Step	Problem	Possible reason	Solution
6,26	PCR does not work efficiently with any tested primer combination and temperature	Target DNA concentration too low	Increase amounts of gDNA in the PCR reaction
		Poor DNA quality or integrity	Check for gDNA degradation by gel electrophoresis. gDNA can also be contaminated by components from the lysed cells, such as salts or nucleases, or components of the purification kits, including incompletely inactivated Proteinase K, SDS, EDTA or chaotropic salts, which can affect photometer readings. Carefully perform all purification steps according to the manufacturer's protocol and check the 260/280 ratio using a spectrophotometer. If required, further purify gDNA by ethanol precipitation
		Locus difficult to amplify by standard PCR	Check locus for overall GC content and hard-to-amplify regions, such as single-base stretches. These issues can often be solved by changing the polymerase (e.g., to NEB Q5, Agilent Herculase, KAPA Master mix), adding high-GC enhancer (comes with Q5), adding 1–5% (vol/vol) DMSO or moving primers to avoid difficult regions
			Perform a nested PCR, which can improve sensitivity and specificity

Table continued

Table 1 (continued)

Step	Problem	Possible reason	Solution
7,27	Sanger sequencing does not work	Poor DNA quality	Purify PCR product using recommended kit
		Inappropriate amounts of DNA	Check for presence of PCR product by gel electrophoresis after purification
			Accurately determine DNA concentration by spectrophotometer or fluorometer measurements and dilute to concentration recommended by Sanger sequencing provider
		Locus difficult to sequence	Use primer for sequencing lying inward of PCR product
			Perform TOPO cloning of PCR products into a plasmid vector. Note that abundance of alleles may vary in the TOPO clones. Heterozygous events will therefore be more difficult to detect
17	Standard deviation of replicates is >0.2	Unequal amounts of reaction mix in each well	Thoroughly mix samples before pipetting
			Ensure accurate pipetting for each sample, which can be achieved by not reusing pipette tips for the same sample or using low-retention pipette tips. If available, electronic pipettes can also help to improve precision of pipetting
		Unequal detection by qPCR machine	Run qPCR machine maintenance to ensure equal detection of all wells; if problem persists, avoid areas with unequal detection, which can occur especially toward the edges of the plate
	C_t values are >30	Too little gDNA within the reaction	Increase gDNA concentration
		Inhibitors of PCR reaction present	Purify gDNA using recommended kit
	No template control gives a signal	Contamination of reaction mix	Repeat qgPCR using all fresh reagents and filter tips for pipetting
			If available, using a dedicated and/or ultraviolet-treated area for qPCR reactions can further minimize risk of contamination
	No amplification detectable in all or some samples	Issues with primers, probe, template, buffers or qPCR machine	See detailed infomation on qPCR basics in *Real-time PCR Handbook* (https://www.thermofisher.com/content/dam/LifeTech/global/Forms/PDF/real-time-pcr-handbook.pdf), and troubleshooting in protocol by Weissenborn et al.[33] for detailed discussion and suggestions
18	Copy number values strongly vary between samples and are not close to values of '1' or '2' allele numbers	Varying integrities of gDNA samples	Repeat gDNA extraction of all samples together and make sure to treat all samples exactly the same
		Unequal amounts of gDNA in each sample	For precise determination of gDNA concentrations, use, for example, a Qubit fluorometer instead of a spectrophotometer. Take extreme care with accurate pipetting when diluting samples to the working concentration
		Restricted accessibility of qPCR assay to gDNA	Predigestion of gDNA with a restriction enzyme that does not cut within the qPCR amplicon can increase accessibility of primers to DNA and improve accuracy of results
	NHEJ clones have copy number values <1	Probe overlaps with indel	Redesign qgPCR assay to prevent overlap of probe with indel
19	More than two bands are visible on agarose gel	Unspecific amplification of qPCR assays	Unspecific amplification occurring in multiple samples could affect the efficiency of target amplification and therefore cause unreliable results. Adjusting qPCR conditions such as input gDNA amount or primer concentrations or eventually exchanging primers should be tested
		Sample has OnTEs at edited locus	Multiple bands in a single sample might be caused by OnTEs that occurred within the genotyping PCR amplicon. This clone should therefore be excluded
27	No tested SNPs are heterozygous in the unedited line	The edit was done in a region of low variation	Expand the region in which the zygosity of SNPs is investigated by doubling the size of the analyzed genomic region and lowering the cutoff for the global MAF by 0.1. In principle, there is no limit to how far away from the cut site LOH can be investigated; however, analyzing heterozygous SNPs close to the cut site can exclude occurrence of smaller regions of LOH
			Perform SNP microarray to obtain genome-wide SNP profile
		An inbred organism was used	Perform editing in a cell line or animal strain with higher genomic variation
			LOH analysis is not necessary for inbred organisms

Table continued

Table 1 (continued)

Step	Problem	Possible reason	Solution
80	GenomeStudio fails to locate intensity data	Incorrect directory is provided	Usually, the intensity files (.idat) are stored in folder(s) with names consisting of just numbers that refer to the beadchip ID. The directory provided to GenomeStudio needs to be the folder that contains all subfolders that store the actual .idat files. Also, using a text editor, check in the Samplesheet if the right directory is provided. For further info, check general documentation available at https://support.illumina.com/content/dam/illumina-support/documents/documentation/software_documentation/genomestudio/genomestudio-2-0/genomestudio-genotyping-module-v2-user-guide-11319113-01.pdf
	GenomeStudio fails to load certain samples	Incorrect Samplesheet	Using a text editor, check in the Samplesheet if the right directory is provided and if all samples are included in the Samplesheet
84,85	GenomeStudio does not perform correct filtering	Conflicting regional system settings	Open the Windows Start menu and click Control Panel. Open the Regional and Language Options dialog box. Click the Regional Options tab. Click 'Customize/Additional settings' (Windows 10). Type a period into the 'Decimal separator' box (.) Click 'OK' twice to confirm the change

Timing

Step 1, CRISPR editing in cells or organisms of interest: 1–3 months, depending on organism
Step 2, gDNA extraction: 2 h
Steps 3–7, design and optimization of 300–450 bp genotyping PCR: 1 d
Step 8, identification of cells or organisms with desired edit by RFLP assay and Sanger sequencing: 2–3 d
Steps 9 and 10, design of labeled oligo for qgPCR: 1 h
Steps 11–20, analysis of edited cells or organisms by qgPCR to exclude those with OnTEs: 1 d
Step 21, select methods to exclude LOH: 1 h
Steps 22–27, identification of nearby heterozygous SNPs by Ensembl BioMart and validation in unedited cells or organisms: 2–4 d
Step 28, LOH analysis by SNP genotyping in edited cells or organisms to exclude those with OnTEs: 1 d
Steps 29–76, analysis of gDNA on genome-wide SNP microarray: 3 d
Steps 77–91, processing of microarray data and analysis of log R ratios and B allele frequencies to exclude cells or organisms with OnTEs: 2 h

Anticipated results

Thorough quality control after genome editing is essential to ensure reliability of results based on edited cells or organisms (Box 2). The presented protocol describes detection of OnTEs in cells or animals after NHEJ- and HDR-mediated CRISPR–Cas9 genome editing. As an example, we demonstrate our OnTE quality-control assays in human iPSCs edited at the APP 'Swedish' locus (APP^{Swe}; Fig. 3a), which is mutated in patients with early-onset Alzheimer's disease[34]. Using our guidelines, the protocol can be adapted to other edited loci in iPSCs and also to other human and non-human cell lines and organisms.

Before performing OnTE detection, successful genome editing should be confirmed in single-cell clones or organisms using a 300–450 bp genotyping PCR. Optimal PCR conditions should be identified by testing combinations of different forward and reverse primers at different annealing temperatures, which usually yields at least one suitable combination (Fig. 3b). Then, a matching fluorescently labeled probe should be designed that can be used in the following qgPCR analysis. The underlying principle of the qgPCR is that a lower number of intact alleles at the edited locus leads to lower fluorescence values during target amplification and therefore an increase of the C_t value of the target-specific qPCR assay. As template amount and quality can vary between different samples, leading to varying PCR amplification of the target, a reference assay, such as TERT for human cells, is used in the same amplification reaction for (C_t) normalization. If C_t values of the reference assays do

not differ between samples (Fig. 3c), samples without OnTEs will have the same C_t value as the edited control. Samples with OnTEs will vary by one cycle (Fig. 3d). In case of differing C_t values for the reference assay (e.g., due to varying gDNA amounts in different samples), samples are normalized by determining the $\Delta\Delta C_t$ value, which then directly correlates with the number of intact alleles at the target locus (Fig. 3e). For our human iPSCs edited at the APP^{Swe} locus we observed clones with OnTEs at varying frequencies, ranging from 17 to 57% depending on the editing system (HDR and plasmid versus NHEJ and RNP, respectively). These clones need to be excluded from further analysis. In general, OnTE frequency may also vary strongly depending on the edited locus or organism.

After successful exclusion of OnTEs such as large deletions or complex rearrangements by qgPCR, one should continue with LOH analysis by confirming the presence of heterozygous SNPs on both sides of the target locus. This can be done either by nearby SNP sequencing or SNP microarrays (Box 1). Again, frequencies of LOH occurrence may strongly depend on the target locus or organism. Furthermore, the number of affected SNPs might also range from only one or a few SNPs, to entire chromosome arms. To exclude the occurrence of LOH at the target site, it is sufficient to genotype one SNP on each side of the cut site after editing, but SNP sequencing can also be expanded further around the target locus to also determine the dimension of larger regions of LOH (Fig. 3f). If LOH detection by SNP microarray analysis is desired, it should first be confirmed that the number and density of SNPs around the target locus detected by the used chip is sufficient (preferably multiple variants within ±10–50 kb). The Illumina Global Screening Array v.2 genotyping chip we used will analyze on average one SNP approximately every 4.4 kb. However, only a subset of analyzed SNPs will be heterozygous in the unedited cell line and only these are useful for detection of LOH in this area (Fig. 3g). In cases of copy-neutral LOH detected by SNP microarray analysis, the log R ratio would stay constant due to unchanged copy numbers, but all heterozygous signals in B-allele frequency would be lost (Fig. 3h provides an example of LOH from the cut site to the chromosome end). Since LOH can affect gene expression, clones with LOH around the target site need to be excluded from further analysis. LOH may also occur distal to the edited locus, but because such cases are probably not caused by Cas9 chromosome cleavage at the targeted site, they would not be classified as OnTEs. In principle, distal LOH could be caused by off-target activity of Cas9, but also by Cas9-independent chromosomal rearrangements that occur spontaneously.

Data availability

Data that support the findings of this study are available in Mendeley Data, https://doi.org/10.17632/v3xg37d77t.1

References

1. Hsu, P. D., Lander, E. S. & Zhang, F. Development and applications of CRISPR-Cas9 for genome engineering. *Cell* **157**, 1262–1278 (2014).
2. Wright, A. V., Nuñez, J. K. & Doudna, J. A. Biology and applications of CRISPR systems: harnessing nature's toolbox for genome engineering. *Cell* **164**, 29–44 (2016).
3. Komor, A. C., Badran, A. H. & Liu, D. R. CRISPR-based technologies for the manipulation of eukaryotic genomes. *Cell* **168**, 20–36 (2017).
4. Fellmann, C., Gowen, B. G., Lin, P.-C., Doudna, J. A. & Corn, J. E. Cornerstones of CRISPR-Cas in drug discovery and therapy. *Nat. Rev. Drug Discov.* **16**, 89–100 (2017).
5. Clement, K., Hsu, J. Y., Canver, M. C., Joung, J. K. & Pinello, L. Technologies and computational analysis strategies for CRISPR applications. *Mol Cell* **79**, 11–29 (2020).
6. Pickar-Oliver, A. & Gersbach, C. A. The next generation of CRISPR-Cas technologies and applications. *Nat. Rev. Mol. Cell Biol.* **20**, 1–18 (2019).
7. Wienert, B., Wyman, S. K., Yeh, C. D., Conklin, B. R. & Corn, J. E. CRISPR off-target detection with DISCOVER-seq. *Nat. Protoc.* **15**, 1–28 (2020).
8. Gkazi, S. A. Quantifying CRISPR off-target effects. *Emerg. Top. Life Sci.* **3**, 327–334 (2019).
9. Shin, H. Y. et al. CRISPR/Cas9 targeting events cause complex deletions and insertions at 17 sites in the mouse genome. *Nat. Commun.* **8**, 15464 (2017).
10. Adikusuma, F. et al. Large deletions induced by Cas9 cleavage. *Nature* **560**, E8–E9 (2018).
11. Kosicki, M., Tomberg, K. & Bradley, A. Repair of double-strand breaks induced by CRISPR-Cas9 leads to large deletions and complex rearrangements. *Nat. Biotechnol.* **36**, 765–771 (2018).
12. Owens, D. D. G. et al. Microhomologies are prevalent at Cas9-induced larger deletions. *Nucleic Acids Res.* **47**, 7402–7417 (2019).
13. Weisheit, I. et al. Detection of deleterious on-target effects after HDR-mediated CRISPR editing. *Cell Rep.* **31**, 107689 (2020).

14. Ikeda, K. et al. Efficient scarless genome editing in human pluripotent stem cells. *Nat. Methods* **15**, 1–7 (2018).
15. Bustin, S. A. & Nolan, T. Pitfalls of quantitative real-time reverse-transcription polymerase chain reaction. *J. Biomol. Tech.* **15**, 155–166 (2004).
16. Nolan, T., Hands, R. E. & Bustin, S. A. Quantification of mRNA using real-time RT-PCR. *Nat. Protoc.* **1**, 1559–1582 (2006).
17. Bi, C. et al. Long-read individual-molecule sequencing reveals CRISPR-induced genetic heterogeneity in human ESCs. *Genome Biol.* **21**, 213 (2020).
18. Ho, S. S., Urban, A. E. & Mills, R. E. Structural variation in the sequencing era. *Nat. Rev. Genet.* **21**, 171–189 (2019).
19. Kosugi, S. et al. Comprehensive evaluation of structural variation detection algorithms for whole genome sequencing. *Genome Biol.* **20**, 117 (2019).
20. Cullot, G. et al. CRISPR-Cas9 genome editing induces megabase-scale chromosomal truncations. *Nat. Commun.* **10**, 1–14 (2019).
21. Yates, A. D. et al. Ensembl 2020. *Nucleic Acids Res.* **41**, D21–D27 (2019).
22. Illumina. *Infinium Global Screening Array-24 v3.0 BeadChip* 1–6 (accessed 22 July 2020); https://www.illumina.com/products/by-type/microarray-kits/infinium-global-screening.html
23. Hashimoto, M., Yamashita, Y. & Takemoto, T. Electroporation of Cas9 protein/sgRNA into early pronuclear zygotes generates non-mosaic mutants in the mouse. *Dev. Biol.* **418**, 1–9 (2016).
24. Tu, Z. et al. Promoting Cas9 degradation reduces mosaic mutations in non-human primate embryos. *Sci. Rep.* **7**, 1–11 (2017).
25. Li, Y. et al. Precise allele-specific genome editing by spatiotemporal control of CRISPR-Cas9 via pronuclear transplantation. *Nat. Commun.* **11**, 4593 (2020).
26. Bogue, M. A., Churchill, G. A. & Chesler, E. J. Collaborative cross and diversity outbred data resources in the mouse phenome database. *Mamm. Genome* **26**, 511–520 (2015).
27. Kwart, D., Paquet, D., Teo, S. & Tessier-Lavigne, M. Precise and efficient scarless genome editing in stem cells using CORRECT. *Nat. Protoc.* **12**, 329–354 (2017).
28. Paquet, D. et al. Efficient introduction of specific homozygous and heterozygous mutations using CRISPR/Cas9. *Nature* **533**, 125–129 (2016).
29. Wefers, B., Bashir, S., Rossius, J., Wurst, W. & Kühn, R. Gene editing in mouse zygotes using the CRISPR/Cas9 system. *Methods* **121-122**, 55–67 (2017).
30. Modzelewski, A. J. et al. Efficient mouse genome engineering by CRISPR-EZ technology. *Nat. Protoc.* **13**, 1253–1274 (2018).
31. Hruscha, A. et al. Efficient CRISPR/Cas9 genome editing with low off-target effects in zebrafish. *Development* **140**, 4982–4987 (2013).
32. Taylor, S. C. et al. The ultimate qPCR experiment: producing publication quality, reproducible data the first time. *Trends Biotechnol.* **37**, 761–774 (2019).
33. Weissenborn, S. J., Wieland, U., Junk, M. & Pfister, H. Quantification of beta-human papillomavirus DNA by real-time PCR. *Nat. Protoc.* **5**, 1–13 (2010).
34. Haass, C. et al. The Swedish mutation causes early-onset Alzheimer's disease by beta-secretase cleavage within the secretory pathway. *Nat. Med.* **1**, 1291–1296 (1995).
35. Sullivan, S. et al. Quality control guidelines for clinical-grade human induced pluripotent stem cell lines. *Regen. Med.* **13**, 859–866 (2018).
36. Baker, D. et al. Detecting genetic mosaicism in cultures of human pluripotent stem cells. *Stem Cell Rep.* **7**, 998–1012 (2016).
37. Concordet, J.-P. & Haeussler, M. CRISPOR: intuitive guide selection for CRISPR/Cas9 genome editing experiments and screens. *Nucleic Acids Res.* **46**, W242–W245 (2018).
38. Hsu, P. D. et al. DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat. Biotechnol.* **31**, 827–832 (2013).
39. Doench, J. G. et al. Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. *Nat. Biotechnol.* **34**, 184–191 (2016).
40. Tycko, J. et al. Mitigation of off-target toxicity in CRISPR-Cas9 screens for essential non-coding elements. *Nat. Commun.* **10**, 4063 (2019).
41. Duncan, G. T., Balamurugan, K., Budowle, B., Smerick, J. & Tracey, M. L. Microvariation at the human D1S80 locus. *Int. J. Legal Med.* **110**, 150–154 (1997).
42. Zurita, E. et al. Genetic polymorphisms among C57BL/6 mouse inbred strains. *Transgenic Res.* **20**, 481–489 (2011).
43. Mattapallil, M. J. et al. The Rd8 mutation of the Crb1 gene is present in vendor lines of C57BL/6N mice and embryonic stem cells, and confounds ocular induced mutant phenotypes. *Invest. Ophthalmol. Vis. Sci.* **53**, 2921–2927 (2012).
44. Freeman, H. C., Hugill, A., Dear, N. T., Ashcroft, F. M. & Cox, R. D. Deletion of nicotinamide nucleotide transhydrogenase: a new quantitive trait locus accounting for glucose intolerance in C57BL/6J mice. *Diabetes* **55**, 2153–2156 (2006).
45. Bouma, M. J. et al. Differentiation-defective human induced pluripotent stem cells reveal strengths and limitations of the teratoma assay and in vitro pluripotency assays. *Stem Cell Rep.* **8**, 1340–1353 (2017).

Acknowledgements

This work was supported by grants from the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy within the framework of the Munich Cluster for Systems Neurology (EXC 2145 SyNergy—ID 390857198), Vascular Dementia Research Foundation, VERUM Foundation, Wilhelm-Vaillant-Foundation, and the donors of the ADR AD2019604S, a program of the BrightFocus Foundation (to D.P.), and Helmholtz Association 'ExNet-0041-Phase2-3 ('SyNergy-HMGU')', Else Kröner Fresenius Stiftung (ForTra-gGmbH; genome editing, to W.W.). We also thank ISAR Bioscience GmbH for technical help with performance of RNP editing experiments.

Author contributions

Conceptualization, I.W. and D.P.; methodology, I.W., J.A.K, R.M., B.W., P.L. and D.P.; writing—original draft, I.W., J.A.K., R.M., B.W., P.L. and D.P.; writing—review and editing, I.W., J.A.K., R.M., B.W., P.L., W.W., M.D. and D.P.; funding acquisition, W.W., M.D. and D.P.; supervision, D.P.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41596-020-00481-2.

Correspondence and requests for materials should be addressed to D.P.

Peer review information *Nature Protocols* thanks Marco Herold, Chengyu Liu and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 23 July 2020; Accepted: 9 December 2020;
Published online: 17 February 2021

Related links
Key reference using this protocol
Weisheit, I. et al. *Cell Rep.* **31**, 107689 (2020): https://doi.org/10.1016/j.celrep.2020.107689

4. Discussion

This thesis deals with a widely overlooked issue in the field of CRISPR genome editing: the occurrence of deleterious OnTEs that frequently escape standard quality controls. We show that this is a widespread problem in genome-edited human iPSCs and develop simple tools for their reliable detection (Figure 8). However, a lot remains to be done to entirely prevent OnTEs from compromising the reliability of CRISPR-based studies: First, we hope to achieve a better awareness of this problem by CRISPR users, which should then lead to widespread implementation of our detection methods. In a next step, it would be beneficial to identify factors that influence OnTE occurrence to unravel underlying molecular mechanisms. This could eventually help to establish guidelines for genome editing protocols that prevent OnTE occurrence in the first place. All steps will be discussed in more detail in the following paragraphs.

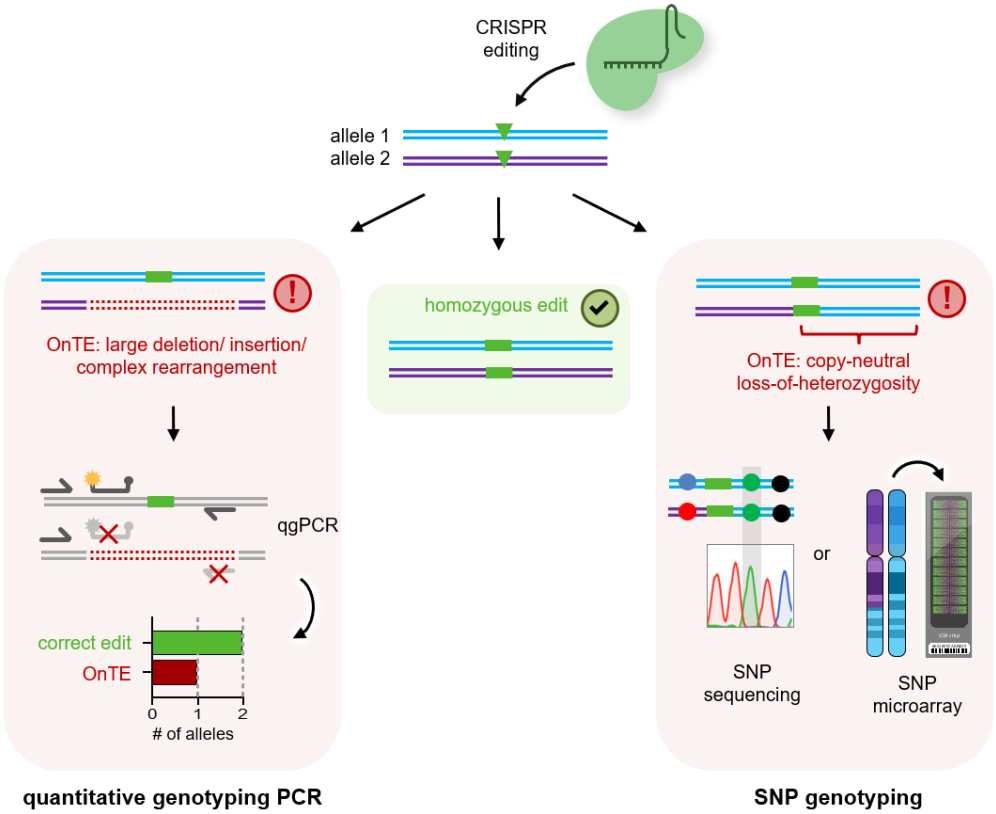


Figure 8: Overview of OnTE detection framework after CRISPR editing. Genome editing using the CRISPR/Cas9 system may induce large deletions, large insertions or complex rearrangements around the target site (left). These can be identified using qgPCR that determines the number of intact alleles. Additionally, the repair of DNA DSBs may lead to regions of LOH around the cut site that can be identified using SNP genotyping by Sanger sequencing or SNP microarrays (right).

4.1 The problem of CRISPR-induced OnTEs requires more attention

The first step towards solving the problem of undetected OnTEs is its recognition as a severe problem by the genome-editing community. A recent study (Kosicki et al., 2018) has created a

lot of attention and researchers were “calling for greater caution when using it in clinical contexts” (Baumann, 2018), as many realized that “Some of these changes could be the first hits that cause cells to eventually become neoplastic” (Leslie, 2018). However, the problem of OnTEs is still widely overlooked in the CRISPR field and measures for their detection are missing in over 90% of recently published papers on edited iPSCs[1]. This thesis clearly demonstrates the severity of this issue: we identified OnTEs in human iPSCs at multiple loci in up to 40% of clones. Thus, it is conceivable that affected clones are sometimes used in similar CRISPR-based studies, which might lead to data misinterpretation and false conclusions. We therefore tried to alert more researchers to this problem by showing an example of deleterious consequences of undetected OnTEs: Using an *in vitro* Alzheimer’s disease model, we show that monoallelic OnTEs can lead to a significant reduction of expression of the targeted gene. This reduction could hamper with pathogenic phenotypes and therefore affect the reliability of the entire experiment.

4.2 Preventing undetected OnTEs by using reliable detection methods

The low level of awareness about the issue of undetected OnTEs in the CRISPR community leads to a lack of on-target quality control measures in most studies. We reasoned that this is to a great extent due to the fact that there are no suitable methods available for this purpose. In this study, we therefore developed a new OnTE detection technology that is reliable, simple, and broadly applicable, which we hope will facilitate its widespread application after genome editing. The advantages and limitations of our methods are discussed in the following, along with a comparison with alternative methods that were previously used in the field.

4.2.1 Advantages of the presented OnTE detection technology

The first advantage of our OnTE detection technology is that the entire workflow can be performed using PCR amplification, gel electrophoresis, Sanger sequencing and qPCR, which are all standard methods in molecular biology labs that are typically available to all CRISPR-users. In addition, we provide an alternative for LOH detection by genome-wide SNP microarrays that requires access to specific equipment but enables more high-throughput analysis.

Second, the combination of only two methods, qgPCR analysis and SNP genotyping, allows reliable detection of the major types of OnTEs after CRISPR editing: qgPCR analysis combines quantification with genotyping and can therefore directly identify copy number changes at the target locus that can be caused by large deletions, insertions, or complex rearrangements. For detection of LOH after editing, each researcher can choose between two different methods

[1] determined by parsing all papers from 2019 on PubMed that had the search terms ‘iPS*’ and ‘CRISPR’ in title or abstract and that performed compatible editing experiments.

according to their advantages and disadvantages: Briefly, nearby SNP genotyping is solely based on PCR and Sanger sequencing and is therefore a cheap and simple assay. SNPs zygosities have to be determined experimentally, which can be laborious but also allows more flexibility with regard to the analyzed area. LOH detection by SNP microarrays is more expensive and requires special equipment but offers a fast option to test multiple samples in parallel, also with different edits. The genome-wide analysis can also serve as additional check for other large chromosomal aberrations.

A third advantage is that the entire analysis is performed using only purified genomic DNA from the edited sample and controls, which allows broad use of our protocol in various cell lines or organisms that all might be affected by OnTEs.

Lastly, our technology is widely applicable to different types of genomic changes after NHEJ- as well as HDR-mediated editing (see below).

4.2.2 Comparison with other methods

Previous studies have used different tools for OnTE detection after CRISPR editing, however, these tools have certain drawbacks that do not make them suitable for reliable and widespread application by all CRISPR users:

Primer-walk PCR, which we have also tested in this study, is very laborious and does not produce reliable results, as large or complex OnTEs might not be detectable by PCR reactions. Furthermore, we also tested standard qPCR assays with a length of maximal 150 bp, but these failed to detect clones with OnTEs that did not directly overlap with the cut site. Droplet digital PCR, similar to qPCR, quantifies a PCR reaction around the edited locus, but this method additionally requires special equipment and complex data analysis. Furthermore, all of the above-mentioned methods do not detect LOH. Next, array-based technologies such as comparative genomic hybridization (CGH) or SNP microarrays are frequently used to detect chromosomal aberrations genome-wide, but their resolution is too low for detection of smaller changes of few kb. Lastly, next-generation sequencing technologies enable reliable detection of OnTEs with high sensitivity, but they are not available to all CRISPR users due to their high cost and complex data analysis. Overall, our technology outcompetes previously used technologies due to its simplicity and broad applicability, but it certainly also has limitations.

4.2.3 Limitations of the presented OnTE detection technology

In this study, we have developed methods for quality control analysis of individual cell lines or organisms to ensure locus integrity after genome editing. Our methods, however, do not allow analysis of OnTE prevalence in genomic DNA from pooled samples. Such an application could prove very useful to further investigate the effect of different genome editing components or to test more loci, as it is described below.

Furthermore, in this study, we have analyzed single-cell clones edited by the NHEJ - as well as the HDR pathway, introducing small sequence changes up to 10 bp. In these cases, the qgPCR amplicon can easily span the edited locus and the unedited 'parent' cell line can be used as control. We have, however, not tested our methods after introduction of larger changes, e.g., after insertion of fluorescent protein tags. Here it is not possible to perform qgPCR analysis using our design principles, as the PCR-product cannot entirely span the edited site. But the same principle, quantification of PCR reactions, might also be applicable to these cases with further adaptations. In contrast, LOH analysis by SNP genotyping is not affected by the size of the edit.
Lastly, qgPCR analysis does not detect genomic alterations that do not affect copy numbers such as translocations or inversions of the entire genotyping PCR amplicon.

We hope that the OnTE detection tools we present in this thesis will provide a valuable resource for the genome editing field and will soon be used by many researchers. The next steps for research on OnTEs should then aim towards better understanding why OnTEs occur and how their occurrence can be avoided.

4.3 Occurrence of OnTEs is a widespread problem

Understanding the general extent of the problem of OnTEs is an important task for the genome editing field. This would help to develop guidelines in which cases it is necessary to perform suitable quality control experiments.
Previous studies have identified OnTEs in mice and one immortalized human cell line and our findings clearly illustrate the universality of this problem in human iPSCs. This suggests widespread conservation of involved molecular mechanisms, but it is important to confirm this hypothesis in further cell lines and model organisms that are frequently used for genome editing. It is likely that there will also be organism- and cell line-specific differences for OnTE prevalence due to varying activities of alternative DNA repair pathways. For example, it has been shown previously that mouse and human stem cells have higher rates of DSB repair by the microhomology-mediated end-joining (MMEJ) pathway (Allen et al., 2019), which accounts for large deletions or insertions at the edited site (see below). Furthermore, all studies, including ours, have only applied CRISPR-based tools for investigation of OnTEs after editing. It is therefore still unclear if other systems, e.g., based on TALEN or ZFN nucleases are affected likewise. However, it is likely that OnTEs are also a problem using these systems, as they also rely on DSB induction, which is most likely the dominant factor for OnTE occurrence, rather than the type of used nuclease.

As mentioned above, the OnTE detection methods developed in this thesis can be applied to different cell lines, organisms and editing systems and can therefore be very helpful to address questions surrounding OnTE occurrence.

4.4 Unraveling the molecular mechanisms underlying OnTE occurrence

When taking a closer look at breakpoint junctions of large deletions around the cut site, we and others (Owens et al., 2019) noticed that these were predominantly located at sites of microhomologies of 2-5 bp. This strongly indicates involvement of the MMEJ pathway, a less common alternative for DSB repair known to induce large deletions or insertions (Sfeir & Symington, 2015). In support of this hypothesis, experiments in mouse ESCs have shown a decreased frequency of large deletions upon deficiency of MMEJ repair genes (Kosicki et al., 2020). A pathway responsible for chromosomal rearrangements leading to copy-neutral LOH has - to our knowledge - not yet been identified.

Deciphering the mechanism that leads to the creation of OnTEs is crucial to develop strategies to avoid them. And in line with this, identifying additional factors in the genome editing process that influence their occurrence, could prove very valuable to increase the efficiency of precise editing. Some of the factors that might affect OnTEs are 1) the cell cycle stage, 2) induction of SSBs by Cas9 nickase 3) DNA sequence characteristics or chromatin state of the surrounding sequence, 4) the type of Cas9 nuclease and 5) the design of the ssODN repair template for HDR-mediated editing. These factors will be discussed in more detail in the following paragraphs.

4.4.1 Cell cycle stage

The cells choice of a specific pathway to repair DSBs highly depends on the cell cycle stage (reviewed in Chapman et al., 2012; Hustedt & Durocher, 2017): the HDR pathway is restricted to the late S and G2 phase when the sister chromatid is available as a repair template. In contrast, the NHEJ pathway is active throughout the cell cycle, but mostly in the G1 and S phase. With the intent to increase the efficiency of precise genome editing by HDR, researchers have therefore restricted Cas9 activity to a specific cell cycle stage by posttranslational regulation (Gutschner et al., 2016). It is conceivable that also the mechanisms leading to OnTE occurrence underlay cell cycle stage-dependent regulation and it would therefore be interesting to investigate if such an approach would affect OnTE prevalence.

4.4.2 Single-strand breaks by Cas9 nickase

Previous studies investigated occurrence of OnTEs after induction of DSBs either by using WT Cas9 (Adikusuma et al., 2018; Kosicki et al., 2018; Owens et al., 2019) or dual Cas9 nickase with paired sgRNAs (Owens et al., 2019) to generate staggered DSBs. Newly developed

CRISPR-based genome editing tools like base editors and prime editors, however, use Cas9 nickase that induces only SSBs into the genomic DNA (Anzalone et al., 2019; Komor et al., 2016) and so far it is unclear if OnTEs also occur in these cases. Since cells usually choose less invasive pathways for repair of SSBs that involve only the replacement of a few nucleotides of the broken strand (reviewed in Caldecott, 2008), it seems less likely that repair mechanisms leading to large deletions or rearrangements are also active at these sites. However, it has been reported that MMEJ, the predominant pathway producing large deletions, also contributes to SSB repair, albeit with lower efficiency (Wang et al., 2019). This might be due to the fact that SSBs can also turn into DSBs, which can then lead to OnTEs. It therefore remains to be investigated to what extent this occurs after genome editing with Cas9 nickase versions.

4.4.3 Surrounding DNA sequence and chromatin structure

Several previous studies have shown that local DNA sequence (Allen et al., 2019; Shen et al., 2018) as well as chromatin context (Schep et al., 2020; reviewed in Scully et al., 2019) surrounding the target locus greatly influence the choice of repair pathways after induction of DSBs by Cas9. However, experimental design often limited these studies to the investigation of smaller genomic changes well below 100 bp and could therefore not specifically determine the occurrence of large OnTEs. Nevertheless, it has been identified that DSB repair by the MMEJ pathway seems to be enhanced in heterochromatic regions (Schep et al., 2020), which could also suggest an increase of OnTEs caused by large deletions at these loci. Furthermore, large microhomologies around the cut site lead to an increase in large deletions, but in general there are strong differences between repair pathway choices at different loci with yet unknown underlying principles (Kosicki et al., 2020).

In the presented study, we have analyzed three different loci, which is not sufficient to extract more features influencing the prevalence of other OnTEs like large insertions or LOH. It would therefore be interesting to investigate this further by testing more loci. This could help to design algorithms that predict OnTE occurrence to potentially guide the choice of sgRNAs for editing.

4.4.4 Cas9 nuclease variants

To overcome the problem of OffTE after genome editing, Cas9 nuclease variants with higher fidelity have been developed. For example, high-fidelity (HiFi) Cas9 carrying a single point mutation has been identified to retain on-target activity, while significantly reducing off-target activity (Vakulskas et al., 2018). The authors speculate that the mutation leads to impaired conformational change of the Cas9 nuclease from an inactive to an active state when mismatches between sgRNA and the target locus are present. It would therefore be interesting to investigate if this or other Cas9 variants, as identified by unbiased screens or rational protein

engineering, could also affect OnTE occurrence by slight changes in the molecular mechanism of initial Cas9 cleavage. Furthermore, not only the exact type of the nuclease used, but also different Cas9 delivery methods might influence OnTE prevalence, e.g., due to varying levels of cutting activity at the target site, which has been shown previously to affect DSB repair outcomes (Kosicki et al., 2020). We and others that investigated OnTE occurrence after CRIPSR editing, have used different Cas9 delivery methods, such as mRNA (Adikusuma et al., 2018; Shin et al., 2017), plasmids (Kosicki et al., 2018; Owens et al., 2019, our study) or RNPs (Kosicki et al., 2018, our study), but differences between those have never been investigated systematically.

4.4.5 HDR repair template design

Previous studies have shown an influence of ssODN orientation, asymmetry and length on the efficiency of genome editing by HDR (Richardson et al., 2016). It is therefore conceivable that these factors also have an effect on repair pathways that lead to OnTE occurrence. If such a correlation is identified at some point, modifying the ssODN design is relatively simple, which is a great advantage of this approach.

Taken together, deciphering factors that influence OnTE occurrence and its underlying molecular mechanism could help preventing OnTEs from being generated in the first place.

4.5 Preventing occurrence of OnTEs after CRISPR-editing

Implementing above-mentioned strategies with potential influence on OnTE prevalence, could eventually lead to complete avoidance or at least a reduction of adverse events at the target locus. In this study, we identified OnTEs in up to 40% of CRISPR-edited clones, which we all had to exclude for subsequent experiments. Reducing the amount of faulty single-cell clones would have saved a lot of cost for media and consumables as well as time for clone picking and analysis. In addition, inhibiting OnTE occurrence would not only generally lower efforts for maintenance but could also automatically increase the rate of precision editing, which is still a problem in many genome editing experiments.

5. Outlook

With the simplicity of the CRISPR system, it is easy to forget that successful genome editing does not only involve creating the desired edit, but additionally requires thorough quality control analysis afterwards. Verifying accurate editing can sometimes take up more time than the editing process itself. This is due to the fact that CRISPR editing can be impaired by several inadvertent alterations in the genome (reviewed in Burgio & Teboul, 2020; Thomas et al., 2019). Briefly, these include OffTE, random integration of editing components, or other chromosomal abnormalities like trisomies or translocations induced by the stressful editing process. Furthermore, it is important to exclude mix up of cell lines or animals during editing as well confirming clonality of cell lines (details for quality control analysis experiments can also be found in our second study). In this thesis, we covered another problem after CRISPR editing, which is the occurrence of deleterious OnTEs, which we hope will soon gain widespread awareness in the field. This should lead - first and foremost - to increased use of OnTE detection methods.

Overall, we believe it will be beneficial for the genome editing field to develop general guidelines for quality control after CRISPR editing that are adapted to the type of genome edit and the edited system. This could help to strengthen the reliability of CRSPR-based studies.

6. References

Adikusuma, F., Piltz, S., Corbett, M. A., Turvey, M., McColl, S. R., Helbig, K. J., Beard, M. R., Hughes, J., Pomerantz, R. T., & Thomas, P. Q. (2018). Large deletions induced by Cas9 cleavage. *Nature*, *560*(7717). https://doi.org/10.1038/s41586-018-0380-z

Allen, F., Crepaldi, L., Alsinet, C., Strong, A. J., Kleshchevnikov, V., De Angeli, P., 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), 64–82. https://doi.org/10.1038/nbt.4317

Anzalone, A. V., Koblan, L. W., & Liu, D. R. (2020). Genome editing with CRISPR–Cas nucleases, base editors, transposases and prime editors. In *Nature Biotechnology* (Vol. 38, Issue 7, pp. 824–844). Nature Research. 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

Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., Romero, D. A., & Horvath, P. (2007). CRISPR Provides Acquired Resistance Against Viruses in Prokaryotes. *Science*, *315*(5819). https://doi.org/10.1126/science.1138140

Baumann, K. (2018). Not so CRISP(R). *Nature Reviews Molecular Cell Biology*.

Brouns, S. J. J., Jore, M. M., Lundgren, M., Westra, E. R., Slijkhuis, R. J. H., Snijders, A. P. L., Dickman, M. J., Makarova, K. S., Koonin, E. V., & Van Der Oost, J. (2008). Small CRISPR RNAs guide antiviral defense in prokaryotes. *Science*, *321*(5891), 960–964. https://doi.org/10.1126/science.1159689

Burgio, G., & Teboul, L. (2020). Anticipating and Identifying Collateral Damage in Genome Editing. In *Trends in Genetics* (Vol. 36, Issue 12, pp. 905–914). Elsevier Ltd. https://doi.org/10.1016/j.tig.2020.09.011

Caldecott, K. W. (2008). Single-strand break repair and genetic disease. In *Nature Reviews Genetics* (Vol. 9, Issue 8, pp. 619–631). https://doi.org/10.1038/nrg2380

Carroll, D. (2019). Collateral damage: Benchmarking off-target effects in genome editing. In *Genome Biology* (Vol. 20, Issue 1). BioMed Central Ltd. https://doi.org/10.1186/s13059-019-1725-0

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

Cheng, Y., & Tsai, S. Q. (2018). Illuminating the genome-wide activity of genome editors for safe and effective therapeutics. In *Genome Biology* (Vol. 19, Issue 1). BioMed Central Ltd. https://doi.org/10.1186/s13059-018-1610-2

Cho, S. W., Kim, S., Kim, J. M., & Kim, J. S. (2013). Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease. *Nature Biotechnology*, *31*(3), 230–232. https://doi.org/10.1038/nbt.2507

Chu, V. T., Weber, T., Wefers, B., Wurst, W., Sander, S., Rajewsky, K., & Kühn, R. (2015). Increasing the efficiency of homology-directed repair for CRISPR-Cas9-induced precise gene editing in mammalian cells. *Nature Biotechnology*, *33*(5), 543–548. https://doi.org/10.1038/nbt.3198

Concordet, J. P., & Haeussler, M. (2018). CRISPOR: Intuitive guide selection for CRISPR/Cas9 genome editing experiments and screens. *Nucleic Acids Research*, *46*(W1), W242–W245. https://doi.org/10.1093/nar/gky354

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

Cradick, T. J., Fine, E. J., Antico, C. J., & Bao, G. (2013). CRISPR/Cas9 systems targeting β-globin and CCR5 genes have substantial off-target activity. *Nucleic Acids Research*, *41*(20), 9584–9592. https://doi.org/10.1093/nar/gkt714

Danner, E., Bashir, S., Yumlu, S., Wurst, W., Wefers, B., & Kühn, R. (2017). Control of gene editing by manipulation of DNA repair mechanisms. *Mammalian Genome*, *28*(7–8), 262–274. https://doi.org/10.1007/s00335-017-9688-5

Deltcheva, E., Chylinski, K., Sharma, C. M., Gonzales, K., Chao, Y., Pirzada, Z. A., Eckert, M. R., Vogel, J., & Charpentier, E. (2011). CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature*, *471*(7340), 602–607. https://doi.org/10.1038/nature09886

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

Ernst, M. P. T., Broeders, M., Herrero-Hernandez, P., Oussoren, E., van der Ploeg, A. T., & Pijnappel, W. W. M. P. (2020). Ready for Repair? Gene Editing Enters the Clinic for the Treatment of Human Disease. In *Molecular Therapy - Methods and Clinical Development* (Vol. 18, pp. 532–557). Cell Press. https://doi.org/10.1016/j.omtm.2020.06.022

Fu, Y., Foden, J. A., Khayter, C., Maeder, M. L., Reyon, D., Joung, J. K., & Sander, J. D. (2013). High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nature Biotechnology*, *31*(9), 822–826. https://doi.org/10.1038/nbt.2623

Fu, Y., Sander, J. D., Reyon, D., Cascio, V. M., & Joung, J. K. (2014). Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. *Nature Biotechnology*, *32*(3), 279–284. https://doi.org/10.1038/nbt.2808

Gasiunas, G., Barrangou, R., Horvath, P., & Siksnys, V. (2012). Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(39). https://doi.org/10.1073/pnas.1208507109

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 T to G C in genomic DNA without DNA cleavage. *Nature*, *551*(7681), 464–471. https://doi.org/10.1038/nature24644

Gutschner, T., Haemmerle, M., Genovese, G., Draetta, G. F., & Chin, L. (2016). Post-translational Regulation of Cas9 during G1 Enhances Homology-Directed Repair. *Cell Reports*, *14*(6), 1555–1566. https://doi.org/10.1016/j.celrep.2016.01.019

Hockemeyer, D., & Jaenisch, R. (2016). Induced pluripotent stem cells meet genome editing. In *Cell Stem Cell* (Vol. 18, Issue 5, pp. 573–586). Cell Press. https://doi.org/10.1016/j.stem.2016.04.013

Hsu, P. D., Lander, E. S., & Zhang, F. (2014). Development and applications of CRISPR-Cas9 for genome engineering. In *Cell* (Vol. 157, Issue 6, pp. 1262–1278). Cell Press. 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

Hustedt, N., & Durocher, D. (2017). The control of DNA repair by the cell cycle. *Nature Cell Biology*, *19*(1), 1–9. https://doi.org/10.1038/ncb3452

Ikeda, K., Uchida, N., Nishimura, T., White, J., Martin, R. M., Nakauchi, H., Sebastiano, V., Weinberg, K. I., & Porteus, M. H. (2018). Efficient scarless genome editing in human pluripotent stem cells. *Nature Methods*, *15*(12), 1045–1047. https://doi.org/10.1038/s41592-018-0212-y

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. In *Science* (Vol. 337, Issue 6096). 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*, *2013*(2). https://doi.org/10.7554/eLife.00471

Kim, S., Kim, D., Cho, S. W., Kim, J., & Kim, J. S. (2014). Highly efficient RNA-guided genome editing in human cells via delivery of purified Cas9 ribonucleoproteins. *Genome Research*, *24*(6), 1012–1019. https://doi.org/10.1101/gr.171322.113

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. (2015). 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. (2015). Engineered CRISPR-Cas9 nucleases with altered PAM specificities. *Nature*, *523*(7561), 481–485. https://doi.org/10.1038/nature14592

Knott, G. J., & Doudna, J. A. (2018). CRISPR-Cas guides the future of genetic engineering. In *Science* (Vol. 361, Issue 6405, pp. 866–869). https://doi.org/10.1126/science.aat5011

Komor, A. C., Badran, A. H., & Liu, D. R. (2017). CRISPR-Based Technologies for the Manipulation of Eukaryotic Genomes. In *Cell* (Vol. 168, Issues 1–2, pp. 20–36). Cell Press. 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*, 420–424.

https://doi.org/10.1038/nature17946
Kosicki, M., Allen, F., & Bradley, A. (2020). Cas9-induced large deletions and small indels are controlled in a convergent fashion. *BioRxiv*, *1*(1), 2020.08.05.216739. https://doi.org/10.1101/2020.08.05.216739
Kosicki, M., Tomberg, K., & Bradley, A. (2018). Repair of double-strand breaks induced by CRISPR–Cas9 leads to large deletions and complex rearrangements. *Nature Biotechnology*, *36*(8). https://doi.org/10.1038/nbt.4192
Kwart, D., Paquet, D., Teo, S., & Tessier-Lavigne, M. (2017). Precise and efficient scarless genome editing in stem cells using CORRECT. *Nature Protocols*, *12*(2), 329–334. https://doi.org/10.1038/nprot.2016.171
Lazzarotto, C. R., Malinin, N. L., Li, Y., Zhang, R., Yang, Y., Lee, G. H., Cowley, E., He, Y., Lan, X., Jividen, K., Katta, V., Kolmakova, N. G., Petersen, C. T., Qi, Q., Strelcov, E., Maragh, S., Krenciute, G., Ma, J., Cheng, Y., & Tsai, S. Q. (2020). CHANGE-seq reveals genetic and epigenetic effects on CRISPR–Cas9 genome-wide activity. *Nature Biotechnology*, *38*(11), 1317–1327. https://doi.org/10.1038/s41587-020-0555-7
Leslie, M. (2018). CRISPR Causes Unexpected Genomic Damage. *Cancer Discovery*, *8*(9). https://doi.org/10.1158/2159-8290.CD-NB2018-106
Liang, X., Potter, J., Kumar, S., Zou, Y., Quintanilla, R., Sridharan, M., Carte, J., Chen, W., Roark, N., Ranganathan, S., Ravinder, N., & Chesnut, J. D. (2015). Rapid and highly efficient mammalian cell engineering via Cas9 protein transfection. *Journal of Biotechnology*, *208*, 44–53. https://doi.org/10.1016/j.jbiotec.2015.04.024
Lin, S., Staahl, B. T., Alla, R. K., & Doudna, J. A. (2014). Enhanced homology-directed human genome engineering by controlled timing of CRISPR/Cas9 delivery. *ELife*, *3*, e04766. https://doi.org/10.7554/eLife.04766
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
Makarova, K. S., Wolf, Y. I., Iranzo, J., Shmakov, S. A., Alkhnbashi, O. S., Brouns, S. J. J., Charpentier, E., Cheng, D., Haft, D. H., Horvath, P., Moineau, S., Mojica, F. J. M., Scott, D., Shah, S. A., Siksnys, V., Terns, M. P., Venclovas, Č., White, M. F., Yakunin, A. F., … Koonin, E. V. (2020). Evolutionary classification of CRISPR–Cas systems: a burst of class 2 and derived variants. In *Nature Reviews Microbiology* (Vol. 18, Issue 2, pp. 67–83). Nature Research. https://doi.org/10.1038/s41579-019-0299-x
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
Malik, R., Chauhan, G., Traylor, M., Sargurupremraj, M., Okada, Y., Mishra, A., Rutten-Jacobs, L., Giese, A. K., Van Der Laan, S. W., Gretarsdottir, S., Anderson, C. D., Chong, M., Adams, H. H. H., Ago, T., Almgren, P., Amouyel, P., Ay, H., Bartz, T. M., Benavente, O. R., … Yamaji, T. (2018). Multiancestry genome-wide association study of 520,000 subjects identifies 32 loci associated with stroke and stroke subtypes. *Nature Genetics*, *50*(4), 524–537. https://doi.org/10.1038/s41588-018-0058-3
Maruyama, T., Dougan, S. K., Truttmann, M. C., Bilate, A. M., Ingram, J. R., & Ploegh, H. L. (2015). Increasing the efficiency of precise genome editing with CRISPR-Cas9 by inhibition of nonhomologous end joining. *Nature Biotechnology*, *33*(5), 538–542. https://doi.org/10.1038/nbt.3190
McComish, S. F., & Caldwell, M. A. (2018). Generation of defined neural populations from pluripotent stem cells. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *373*(1750). https://doi.org/10.1098/rstb.2017.0214
Mullard, A. (2020). Gene-editing pipeline takes off. In *Nature reviews. Drug discovery* (Vol. 19, Issue 6, pp. 367–372). NLM (Medline). https://doi.org/10.1038/d41573-020-00096-y
Naeem, M., Majeed, S., Hoque, M. Z., & Ahmad, I. (2020). Latest Developed Strategies to Minimize the Off-Target Effects in CRISPR-Cas-Mediated Genome Editing. In *Cells* (Vol. 9, Issue 7). NLM (Medline). https://doi.org/10.3390/cells9071608
Owens, D. D. G., Caulder, A., Frontera, V., Harman, J. R., Allan, A. J., Bucakci, A., Greder, L., Codner, G. F., Hublitz, P., McHugh, P. J., Teboul, L., & de Bruijn, M. F. T. R. (2019). Microhomologies are prevalent at Cas9-induced larger deletions. *Nucleic Acids Research*, *47*(14), 7402–7417. https://doi.org/10.1093/nar/gkz459
Paquet, D., Kwart, D., Chen, A., Sproul, A., Jacob, S., Teo, S., Olsen, K. M., Gregg, A., Noggle, S., &

Tessier-Lavigne, M. (2016). Efficient introduction of specific homozygous and heterozygous mutations using CRISPR/Cas9. *Nature*, *533*(7601), 125–129. https://doi.org/10.1038/nature17664

Parikh, B. A., Beckman, D. L., Patel, S. J., White, J. M., & Yokoyama, W. M. (2015). Detailed phenotypic and molecular analyses of genetically modified mice generated by CRISPR-Cas9-mediated editing. *PLoS ONE*, *10*(1). https://doi.org/10.1371/journal.pone.0116484

Pattanayak, V., Lin, S., Guilinger, J. P., Ma, E., Doudna, J. A., & Liu, D. R. (2013). High-throughput profiling of off-target DNA cleavage reveals RNA-programmed Cas9 nuclease specificity. *Nature Biotechnology*, *31*(9), 839–843. https://doi.org/10.1038/nbt.2673

Pausch, P., Al-Shayeb, B., Bisom-Rapp, E., Tsuchida, C. A., Li, Z., Cress, B. F., Knott, G. J., Jacobsen, S. E., Banfield, J. F., & Doudna, J. A. (2020). CRISPR-CasF from huge phages is a hypercompact genome editor. *Science*. http://science.sciencemag.org/

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

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. (2013). 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

Richardson, C. D., Ray, G. J., DeWitt, M. A., Curie, G. L., & Corn, J. E. (2016). Enhancing homology-directed genome editing by catalytically active and inactive CRISPR-Cas9 using asymmetric donor DNA. *Nature Biotechnology*, *34*(3), 339–344. https://doi.org/10.1038/nbt.3481

Rouet Philippe, Smih Fatima, & Jasin Maria. (1994). Introduction of Double-Strand Breaks into the Genome of Mouse Cells by Expression of a Rare-Cutting Endonuclease. *Molecular and Cellular Biology*. http://mcb.asm.org/

Schep, R., Brinkman, E., Leemans, C., Vergara, X., Morris, B., van Schaik, T., Manzo, S., Peric-Hupkes, D., van den Berg, J., Beijersbergen, R., Medema, R., & van Steensel, B. (2020). Impact of chromatin context on Cas9-induced DNA double-strand break repair pathway balance. *BioRxiv*. https://doi.org/10.1101/2020.05.05.078436

Scully, R., Panday, A., Elango, R., & Willis, N. A. (2019). DNA double-strand break repair-pathway choice in somatic mammalian cells. *Nature Reviews Molecular Cell Biology*, *20*(11), 698–714. https://doi.org/10.1038/s41580-019-0152-0

Sfeir, A., & Symington, L. S. (2015). Microhomology-Mediated End Joining: A Back-up Survival Mechanism or Dedicated Pathway? *Trends in Biochemical Sciences*, *40*(11), 701–714. https://doi.org/10.1016/j.tibs.2015.08.006

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

Shin, H. Y., Wang, C., Lee, H. K., Yoo, K. H., Zeng, X., Kuhns, T., Yang, C. M., Mohr, T., Liu, C., & Hennighausen, L. (2017). CRISPR/Cas9 targeting events cause complex deletions and insertions at 17 sites in the mouse genome. *Nature Communications*, *8*. https://doi.org/10.1038/ncomms15464

Slaymaker, I. M., Gao, L., Zetsche, B., Scott, D. A., Yan, W. X., & Zhang, F. (2016). Rationally engineered Cas9 nucleases with improved specificity. *Science*, *351*(6268). https://doi.org/10.1126/science.aad5227

Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K., & Yamanaka, S. (2007). Induction of Pluripotent Stem Cells from Adult Human Fibroblasts by Defined Factors. *Cell*, *131*(5), 861–872. https://doi.org/10.1016/j.cell.2007.11.019

Takahashi, K., & Yamanaka, S. (2006). Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors. *Cell*, *126*(4), 663–676. https://doi.org/10.1016/j.cell.2006.07.024

TCW, J. (2019). Human iPSC application in Alzheimer's disease and Tau-related neurodegenerative diseases. *Neuroscience Letters*, *699*(January), 31–40. https://doi.org/10.1016/j.neulet.2019.01.043

Thomas, M., Burgio, G., Adams, D. J., & Iyer, V. (2019). Collateral damage and CRISPR genome editing. *PLoS Genetics*, *15*(3). https://doi.org/10.1371/journal.pgen.1007994

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

Tsai, S. Q., Zheng, Z., Nguyen, N. T., Liebers, M., Topkar, V. V., Thapar, V., Wyvekens, N., Khayter, C., Iafrate, A. J., Le, L. P., Aryee, M. J., & Joung, J. K. (2015). GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases. *Nature Biotechnology*, *33*(2), 187–198.

https://doi.org/10.1038/nbt.3117
Vakulskas, C. A., Dever, D. P., Rettig, G. R., Turk, R., Jacobi, A. M., Collingwood, M. A., Bode, N. M., McNeill, M. S., Yan, S., Camarena, J., Lee, C. M., Park, S. H., Wiebking, V., Bak, R. O., Gomez-Ospina, N., Pavel-Dinu, M., Sun, W., Bao, G., Porteus, M. H., & Behlke, M. A. (2018). A high-fidelity Cas9 mutant delivered as a ribonucleoprotein complex enables efficient gene editing in human hematopoietic stem and progenitor cells. *Nature Medicine*, *24*(8), 1216–1224. https://doi.org/10.1038/s41591-018-0137-0
Wang, Z., Song, Y., Li, S., Kurian, S., Xiang, R., Chiba, T., & Wu, X. (2019). DNA polymerase (POLQ) is important for repair of DNA double-strand breaks caused by fork collapse. *Journal of Biological Chemistry*, *294*(11), 3909–3919. https://doi.org/10.1074/jbc.RA118.005188
Wiedenheft, B., Sternberg, S. H., & Doudna, J. A. (2012). RNA-guided genetic silencing systems in bacteria and archaea. In *Nature* (Vol. 482, Issue 7385, pp. 331–338). https://doi.org/10.1038/nature10886
Wienert, B., Wyman, S. K., Richardson, C. D., Yeh, C. D., Akcakaya, P., Porritt, M. J., Morlock, M., Vu, J. T., Kazane, K. R., Watry, H. L., Judge, L. M., Conklin, B. R., Maresca, M., & Corn, J. E. (2019). Unbiased detection of CRISPR off-targets in vivo using DISCOVER-Seq. *Science*, *364*(6437), 286–289. https://doi.org/10.1126/science.aav9023
Wright, A. V., Nuñez, J. K., & Doudna, J. A. (2016). Biology and Applications of CRISPR Systems: Harnessing Nature's Toolbox for Genome Engineering. In *Cell* (Vol. 164, Issues 1–2, pp. 29–44). Cell Press. https://doi.org/10.1016/j.cell.2015.12.035
Zetsche, B., Gootenberg, J. S., Abudayyeh, O. O., Slaymaker, I. M., Makarova, K. S., Essletzbichler, P., Volz, S. E., Joung, J., Van Der Oost, J., Regev, A., Koonin, E. V., & Zhang, F. (2015). Cpf1 Is a Single RNA-Guided Endonuclease of a Class 2 CRISPR-Cas System. *Cell*, *163*(3), 759–771. https://doi.org/10.1016/j.cell.2015.09.038
Zuckermann, M., Hovestadt, V., Knobbe-Thomsen, C. B., Zapatka, M., Northcott, P. A., Schramm, K., Belic, J., Jones, D. T. W., Tschida, B., Moriarity, B., Largaespada, D., Roussel, M. F., Korshunov, A., Reifenberger, G., Pfister, S. M., Lichter, P., Kawauchi, D., & Gronych, J. (2015). Somatic CRISPR/Cas9-mediated tumour suppressor disruption enables versatile brain tumour modelling. *Nature Communications*, *6*. https://doi.org/10.1038/ncomms8391

7. Copyright information

Figure number	1
License number	4996000230237
License date	Jan 25, 2021
Licensed content publisher	Elsevier
Licensed content publication	Cell
Licensed content title	Biology and Applications of CRISPR Systems: Harnessing Nature's Toolbox for Genome Engineering
Licensed content author	Addison V. Wright, James K. Nuñez, Jennifer A. Doudna

Figure number	4
License number	4995990984718
License date	Jan 25, 2021
Licensed content publisher	Springer Nature
Licensed content publication	Nature Biotechnology
Licensed content title	Genome editing with CRISPR–Cas nucleases, base editors, transposases and prime editors
Licensed content author	Andrew V. Anzalone et al.

8. Curriculum Vitae

Isabel Weisheit

EDUCATION

Since 04/2017	**Institute for Stroke and Dementia Research, Munich** **Graduate School of Systemic Neurosciences** **Ludwig-Maximilians-University, Lab of Prof. Dominik Paquet** *PhD student*
10/2014–01/2017	**Ludwig-Maximilians-University, Munich** *Master of Science* in Biochemistry, grade: 1.1
10/2011–10/2014	**University of Ulm** *Bachelor of Science* in Biochemistry, grade: 1.2
2009–2011	**Jakob-Fugger-Gymnasium, Augsburg** *Abitur*, grade: 1.6
2008–2009	**Alamosa High School, Colorado, USA** *High School Diploma*, grade point average: 3.9 of 4.0
2004–2008	**Jakob-Fugger-Gymnasium, Augsburg**
2002–2004	**Hölderlin-Gymnasium, Cologne**

PROFESSIONAL & RESEARCH EXPERIENCE

10/2020–01/2021	**BIOTOPIA Lab Museum, Munich** *Research assistant* in science museum
02/2020–05/2020	**ISAR Bioscience GmbH, Munich** *Research intern* in translational research company
06/2016–01/2017	**Francis Crick Institute, London, Great Britain** Lab of Prof. Jernej Ule *Master's thesis:* modifying the RNA-binding properties of TDP-43 by CRISPR/Cas9 genome editing
02–03/2016	**Ludwig-Maximilians-University, Biomedical Center, Munich** Lab of Prof. Michael Kiebler *Student assistant:* analysis of protein expression in mouse brain lysates
09–10/2015	**Ludwig-Maximilians-University, Gene Center, Munich** Lab of Prof. Mario Halic *Research intern:* expression and purification of histones for recombinant nucleosome reconstitution
02–06/2015	**Helmoltz Center, Stem Cell Research Institute, Munich** Lab of Dr. Micha Drukker *Research intern and student assistant*: studies on the role of lncRNAs on the pluripotency state of embryonic stem cells
04–06/2014	**University of Ulm, Organic Chemistry Institute III** Lab of Prof. Tanja Weil *Bachelor's thesis*: protein-DNA conjugation for artificial fusion proteins

9. List of Publications

1. **Weisheit, I.**, Kroeger, J. A., Malik, R., Klimmt, J., Crusius, D., Dannert, A., Dichgans, M., & Paquet, D. (2020). Detection of Deleterious On-Target Effects after HDR-Mediated CRISPR Editing. ***Cell Reports***, 31(8). https://doi.org/10.1016/j.celrep.2020.107689

2. **Weisheit, I**., Kroeger, J. A., Malik, R., Benedikt Wefers, B., Lichtner, P., Wurst, W., Dichgans, M., & Paquet, D. (2021). Simple and reliable detection of CRISPR-induced on-target effects by qgPCR and SNP genotyping. ***Nature Protocols***, 16. https://doi.org/10.1038/s41596-020-00481-2

10. Affidavit

Eidesstattliche Versicherung/Affidavit

Hiermit versichere ich an Eides statt, dass ich die vorliegende Dissertation “Detection of Deleterious On-Target Effects after CRISPR-mediated Editing in Human Induced Pluripotent Stem Cells” selbstständig angefertigt habe, mich außer der angegebenen keiner weiteren Hilfsmittel bedient und alle Erkenntnisse, die aus dem Schrifttum ganz oder annähernd übernommen sind, als solche kenntlich gemacht und nach ihrer Herkunft unter Bezeichnung der Fundstelle einzeln nachgewiesen habe.

I hereby confirm that the dissertation “Detection of Deleterious On-Target Effects after CRISPR-mediated Editing in Human Induced Pluripotent Stem Cells” is the result of my own work and that I have only used sources or materials listed and specified in the dissertation.

München, den
Munich, 08.08.2021

_____

Isabel Weisheit

11. Declaration of Author Contributions

Detection of Deleterious On-Target Effects after HDR-Mediated CRISPR Editing.

Weisheit, I., Kroeger, J. A., Malik, R., Klimmt, J., Crusius, D., Dannert, A., Dichgans, M., Paquet, D.

Cell Reports, 31(8), 107689. doi: 10.1016/j.celrep.2020.107689.

Author contributions: Conceptualization, I.W. and D.P.; Methodology, I.W., J.A.K., R.M., and D.P.; Investigation, I.W., J.A.K., R.M., J.K., D.C., and A.D.; Writing – Original Draft, I.W., J.A.K., R.M., and D.P.; Writing – Review & Editing, I.W., J.A.K., R.M., J.K., A.D., M.D., and D.P.; Funding Acquisition, M.D. and D.P.; Supervision, M.D. and D.P.

Simple and reliable detection of CRISPR-induced on-target effects by qgPCR and SNP genotyping.

Weisheit, I., Kroeger, J. A., Malik, R., Wefers, B., Lichtner, P., Wurst, W., Dichgans, M., Paquet, D.

Nature Protocols, 16, 1714–1739. doi: 10.1038/s41596-020-00481-2

Author contributions: Conceptualization, I.W. and D.P.; Methodology, I.W., J.A.K, R.M., B.W., P.L., D.P.; Writing – Original Draft, I.W., J.A.K., R.M., B.W., P.L., D.P.; Writing – Review & Editing, I.W., J.A.K., R.M., B.W., P.L., W.W., M.D., D.P.; Funding Acquisition, W.W., M.D., D.P.; Supervision, D.P.

Herewith, I confirm the contributions of Isabel Weisheit to the articles.

München, den
Munich, 08.08.2021

Isabel Weisheit	Prof. Dominik Paquet (Supervisor)

12. Acknowledgements

Danksagung

Zuallererst möchte ich Prof. Dominik Paquet danken. Er ermöglichte es mir, meine zwei größten Interessen - Stammzellen und CRISPR - in einem spannenden Projekt zu vereinen. Ich danke ihm für die intensive Betreuung und die anregenden Diskussionen auf Augenhöhe. Seine zahlreichen Ideen und die strukturierte Arbeitsweise schufen die Grundlage einer inspirierenden, produktiven Umgebung für meine Doktorarbeit.

Mein Dank gilt auch meinem zweiten Betreuer, Prof. Martin Dichgans, für seine Unterstützung und das fortwährende Interesse an meinem Projekt. Es war eine besondere Erfahrung in einem Institut zu promovieren, das Grundlagenforschung und Patientenkontakt verbindet.

Dr. Dorothee Dormann danke ich für ihre hilfreichen Ratschläge in meinen TAC Meetings. Prof. Michael Kiebler und Dr. Stefan Stricker danke ich für die Prüfung meiner Doktorarbeit.

Ich danke auch von Herzen allen Kollegen in der Arbeitsgruppe. Allen voran danke ich hierbei Julien Klimmt, mit dem ich in produktiver Zusammenarbeit Vieles zu Beginn im Labor etablierte. Ein besonderer Dank gilt auch Dennis Crusius, der mit seinem Organisationstalent all unsere Projekte unterstützte. Ich danke Judit(h) González-Gallego für ihre fröhliche Art und unsere enge Freundschaft in den letzten Monaten. Angelika Dannert danke ich für die zahlreichen kleinen Tipps und die oft tiefgründigen Unterhaltungen an langen Labortagen. Ich danke Jennifer Yilmaz für die organisatorische Unterstützung im Hintergrund. Während meiner Doktorarbeit durfte ich außerdem zwei Masterstudenten betreuen: Ich danke Joseph Kroeger, der mit seinem Einsatz zum Erfolg meiner Doktorarbeit beitrug. Annika Wagener danke ich, dass sie dieses Projekt weiter voranbringt. Außerdem danke ich aus unserer Gruppe Liliana Pedro, Sophie Robinson, Carolina Cardoso Gonçalves, Einar Krogsæter, Marvin Reich und Bérénice Horlacher. Ich bin sehr froh, dass wir uns alle gegenseitig unterstützten und man Jeder/m seine Zellen am Wochenende zum Füttern anvertrauen konnte. Diese positive Atmosphäre trug dazu bei, dass ich jeden Tag gerne in die Arbeit ging.

Meinen Kollaborationspartnern, Dr. Rainer Malik, Dr. Benedikt Wefers, Dr. Peter Lichtner und Prof. Wolfgang Wurst danke ich für die produktive und erfolgreiche Zusammenarbeit. Dr. Matthias Prestel, Dr. Guangyao Yan, Dr. Yaw Asare, Barbara Lindner und Dr. Stefan Roth danke ich für die Unterstützung bei meinem HDAC9 Projekt. Bei ISAR Bioscience danke ich Dr. Sabrina Desbordes, sowie Dr. Miriam Vogt und Dr. Gernot Kleinberger.

Lena Bittl und Stefanie Bosse danke ich für die organisatorische Unterstützung in der GSN.

Einen besonderen Dank möchte ich auch Dr. Markus Grosch und Dr. Miha Modic widmen, die meine Neugier für Stammzellen und CRISPR weckten. Ich danke auch einigen weiteren

Personen, die mich während meines Studiums inspirierten und prägten: Dr. Micha Drukker, Dr. Silvija Bilokapic, Dr. Rico Schieweck, Dr. Martina Hallegger und Prof. Jernej Ule.

Außerdem danke ich meinen Kollegen aus anderen Arbeitsgruppen, die zu engen Freunden wurden und dazu beitrugen, dass die letzten vier Jahre eine unvergessliche Zeit wurden: Ich danke Dr. Rebecca Sadler, Dr. Andreas Zellner, Madita (Madiiitz) Reimer und Bernhard (Bernz) Groschup für die unterhaltsamen Kaffeepausen, tiefsinnigen Gespräche am Mikroskop, schönen Abende im Münchner Nachtleben, Ausflüge in die Berge und ereignisreichen Urlaube.

Des Weiteren möchte ich mich bei meinen Freunden außerhalb des Labors bedanken: Ich danke Achim für eine schöne Zeit, in der wir gemeinsam die Höhen und Tiefen des Studiums meisterten. Ich danke Kathi, mit der ich so gut abschalten kann und die mich davon abhält eine verrückte Wissenschaftlerin zu werden. Samira danke ich für die vielen schönen Stunden in den Bergen und beim Bouldern, die mir immer wieder neue Kraft gaben. Migjen danke ich für viel Unterstützung und inspirierende Gespräche, die mich meine Arbeit von weiter weg betrachten ließen.

Mein größter Dank gilt meiner Familie: Ich danke meiner Schwester Lara, die immer ein Vorbild für mich ist. Unzählige Male beriet sie mich bei kleinen und großen Entscheidungen und ihre aufmunternden Worte gaben mir neue Kraft. Meiner Mutter und meinem Vater danke ich, dass sie immer an mich glauben und so viel Interesse an meiner Arbeit zeigen. Ihr brachtet mir bei, für meine Träume zu kämpfen, kritisch zu denken und Allem im Leben mit Respekt zu begegnen.