

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 Malik¹, Benedikt Wefers^{3,4}, Peter Lichtner⁵, Wolfgang Wurst^{3,4,6,7}, Martin Dichgans^{1,3,7}, Dominik Paquet^{1,3,7*}

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 German Center for Neurodegenerative Diseases (DZNE) Munich, 81377 Munich, Germany

⁴ Institute of Developmental Genetics (IDG), HelmholtzZentrum München, 85764 Neuherberg, Germany

⁵ Core Facility NGS, HelmholtzZentrum München, 85764 Neuherberg, Germany

⁶ Technische Universität München-Weihenstephan, 85764 Neuherberg, Germany

⁷ Munich Cluster for Systems Neurology (SyNergy), 81377 Munich, Germany

* Correspondence: dominik.paquet@med.uni-muenchen.de

Abstract

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 (OnTE) 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 neighboring genes, OnTE 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 OnTE in genome-edited cells and organisms after non-homologous end joining (NHEJ-) and homology-directed repair (HDR-)mediated editing. Our protocol enables identification of OnTE 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 allows excluding clones with mono-allelic OnTE 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 neighboring 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 days.

Introduction

The 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 patients¹⁻⁴. 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, e.g. in 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). While 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 (OnTE) are still lacking. As we and others have recently shown, OnTE can occur at high frequency

in mouse and human cells, with up to 40% of edited clones affected in human iPSCs⁹⁻¹³. Here, we describe a procedure to efficiently detect OnTE and ensure locus integrity after CRISPR editing.

Development of the protocol

In our recent study¹³ we revealed frequent occurrence of OnTE, such as large deletions, large insertions, complex rearrangements and loss-of-heterozygosity caused by CRISPR genome editing in human iPSCs. These OnTE 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 knock-outs via the non-homologous end joining (NHEJ) pathway. Earlier work also demonstrated presence of OnTE in CRISPR-edited mice and cell lines⁹⁻¹², so it is conceivable that OnTE are a widespread problem relevant for CRISPR editing in many organisms

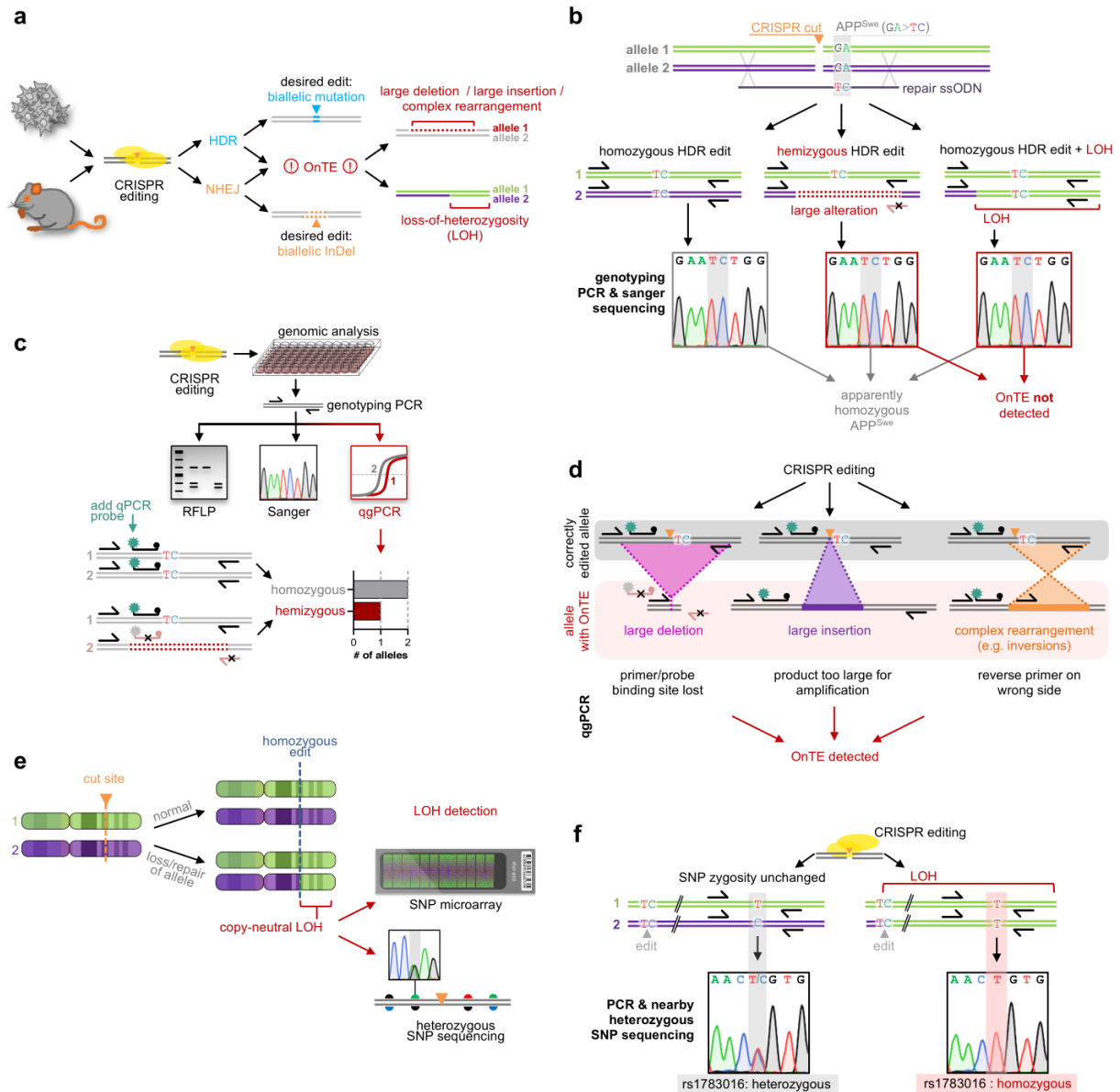


Fig. 1 | CRISPR/Cas9 genome editing induces deleterious OnTE that can be detected by qPCR and SNP genotyping in a simple and reliable manner. **a**, CRISPR editing can induce unintended OnTE such as large deletions, large insertions, complex rearrangements or regions of copy-neutral loss-of-heterozygosity (LOH) around the target locus in human cells, mice, and other organisms during HDR- as well as NHEJ-mediated editing. **b**, OnTE often escape standard quality controls like 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**, qPCR 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 Ct values. RFLP: restriction fragment length polymorphism. **d**, qPCR allows detecting mono-allelic large deletions (left), large insertions (middle), and complex rearrangements (e.g. inversions) (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 zygosity 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. Figure panels c and e are modified from Ref. 13.

(Fig. 1a). Highlighting the importance of their detection, we showed that OnTE can have deleterious consequences on formation of relevant phenotypes in disease models¹³. We therefore developed and validated assays for reliable detection of OnTE based on quantitative PCR, Sanger sequencing and SNP microarrays.

Rather than developing a complex technology accessible only by specialist labs, we reasoned 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 lab performing CRISPR

genome editing with minimal requirements for specialized machinery. The entire workflow of OnTE detection including analysis of loss-of-heterozygosity (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.

OnTE 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: quantitative PCR 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

Box 1 | Choosing nearby SNP genotyping or SNP microarrays for loss-of-heterozygosity (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 each lab 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 labs, such as Sanger sequencing. It is also more flexible than predefined SNP microarrays, as the genotyped SNPs can be freely chosen according to experimental needs. This custom design can improve detection of small regions of LOH that can 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 lab 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 zygosity 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 compared to 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 about 600 kb to 1 Mb can be detected, however smaller aberrations, as well as copy-neutral inversions and balanced translocations will be missed by microarray analysis.

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 OnTE 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 neighboring 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 labs (see Box 1 for details).

Applications of the method

OnTE are widespread in edited human iPSCs, other human cells and mice⁹⁻¹³, but likely also occur in other organisms subjected to CRISPR editing. We have established our protocol using CRISPR-edited human iPSCs. But 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-bp changes were introduced by HDR-mediated editing, or small InDels, as typically introduced by NHEJ-mediated editing to generate gene knock-outs. If heterozygous edits were introduced, 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 about 350 bp, such as targeted large deletions/inversions or large knock-ins (e.g. for protein tagging) could also be checked for OnTE 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 knock-ins 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 if cells or organisms edited with these tools are also affected by OnTEs.

Comparison with other methods

Several other methods have been used for OnTE detection previously:

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 OnTE. While 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 to detect all OnTE, especially in the presence of large insertions or

complex alterations that surpass the amplification limits of traditional PCR¹³. Moreover, primer-walking PCR does not detect LOH.

Next-generation sequencing technologies allow determining OnTE occurrence 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 labs, 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 OnTE need to be applied in order to map discordances between sample and reference genomes. But since OnTE can include very diverse structural variations of the genome, the method may not detect all OnTE^{18,19}. To overcome this problem, targeted long-read sequencing approaches, e.g. from Pacific Biosciences (PacBio) or Oxford Nanopore Technologies (ONT) 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 quantitative PCR reaction, but additionally partitions the sample into thousands of individual droplets (i.e. 'reaction chambers'), allowing absolute quantification of target sequences¹². 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 up to 60 kb, which largely surpasses standard G-banding²⁰. However, this resolution is still insufficient to detect many of the OnTE we detected in our iPSCs, which were in many cases well below 10 kb¹³. 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. While this method offers relatively high resolution to detect

also 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 OnTE, as they failed to reliably detect all affected clones (see qgPCR design section for more details)¹³.

Experimental design

Detection of OnTE 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 on-target effect detection

Due to the frequent occurrence of OnTE (up to 40% of HDR-edited human iPSCs¹³), we recommend prioritizing their detection over other quality control measures after CRISPR editing. In our experiments exclusion of clones was more frequently due to OnTE 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 genomic DNA of single cell clones right after editing and confirm successful editing by Sanger genotyping, directly followed by OnTE analysis. For OnTE analysis, we recommend starting with qgPCR because OnTE 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 approximately 2-3 days.

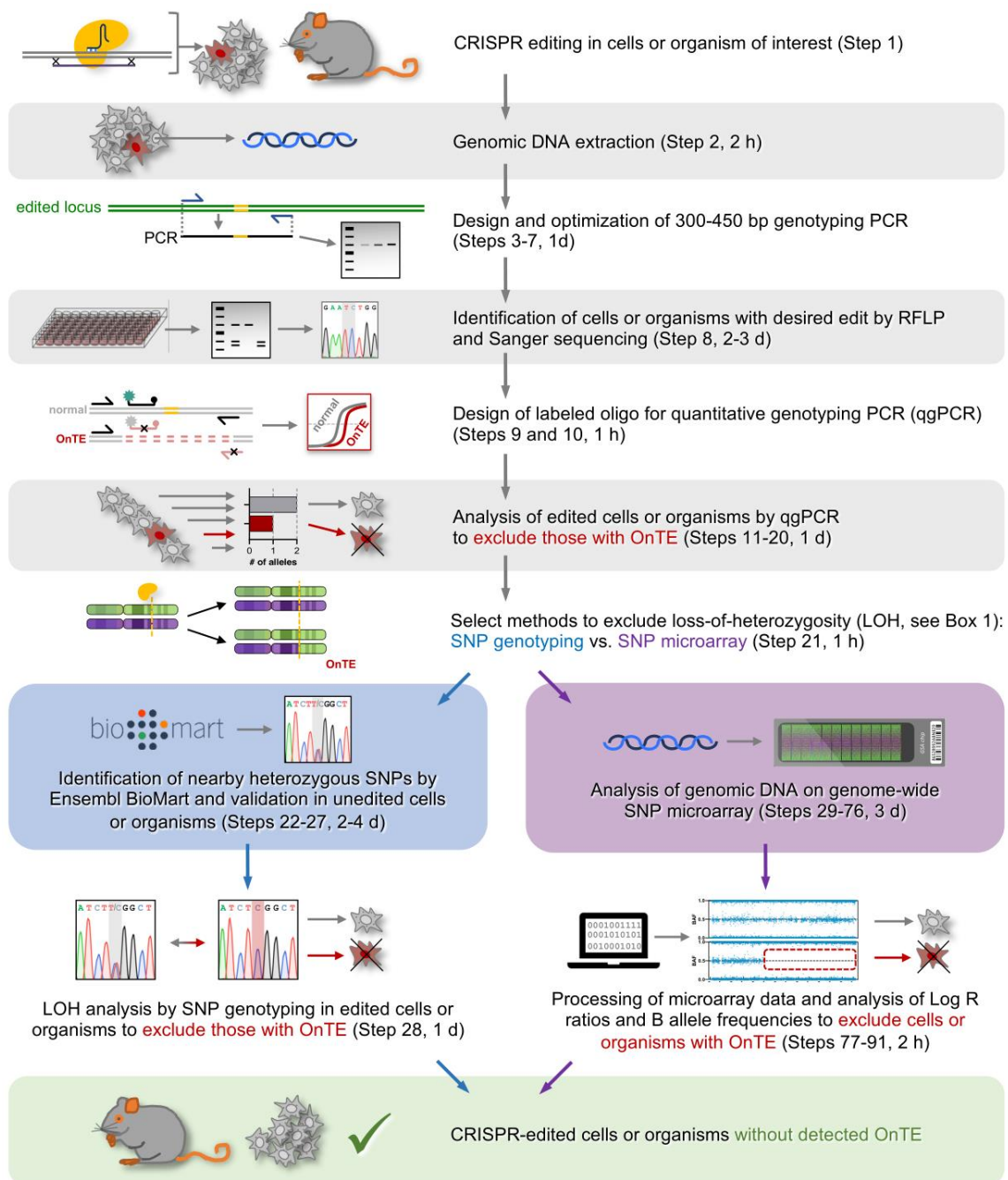


Fig. 2 | Overview of the procedure for simple and reliable OnTE detection after CRISPR genome editing.

Controls

OnTE are always detected by comparing gDNA from edited to the 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 normalization factor to exclude clones that have acquired copy number changes by genome editing. For LOH 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 zygosity after editing.

qgPCR design

Performing qgPCR requires applying a qPCR assay amplifying a region of around 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 qPCR assay for the edited gene. Amplifying a reference gene known to be present in two copies in the genome enables normalizing the gene of interest to an endogenous control. Thus, C_t value differences caused by different amounts of genomic DNA in the reaction tube can be normalized.

When we developed the method in our original study, we investigated the possibility of assay artifacts by testing two different qPCR 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 iPSCs¹³. In these experiments, we did not observe conflicting results between different assay sets and therefore conclude that one qPCR 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.

We have recently shown that it is crucial not to use qPCR amplicon sizes typically recommended by design tools with a standard of maximum 150 bp for qPCR¹³. In these instances, assays failed to reliably flag clones affected by OnTE 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 qPCR. However, longer amplicons with the recommended length of 300-450 bp were sufficient to reveal OnTE non-contiguous with the cut site. This size is a good compromise between efficiency of qPCR (which decreases with increasing amplicon length), and likelihood of OnTE (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 around the edited locus that can be applied to evaluate presence of LOH. It is up to the user to decide which assay is more appropriate for

individual experiments and lab requirements (see detailed discussion of benefits and challenges in Box 1).

Nearby SNP sequencing. Performing nearby SNP sequencing for LOH analysis requires to first identify potentially heterozygous SNPs on both sides of the edited locus in the 'parent' cell line or organism using Ensembl BioMart²¹, 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 around 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 around 5-10 SNPs on each side, the genomic region can be increased or the cutoff for minor allele frequency 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 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 9 SNPs with MAF of ≥ 0.3 have to be checked to reach a 99% chance to find a heterozygous one (allele frequency according to Hardy-Weinberg = $2 \cdot 0.7 \cdot 0.3 = 42\%$, 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 detecting smaller regions of LOH. SNP microarrays can also be customized to a certain degree, which allows

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 (on-target effects) 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 (QC) workflow for each edited iPSC line³⁵ and founder animal, which is prioritized according to frequency of occurrence as well as cost and complexity of the assay:

1. **On-target effects (cells and animals):** As described in this protocol, every line or founder is screened for OnTE by qgPCR and SNP genotyping (see discussion above about applicability in animals).
2. **Chromosomal aberrations (cells and animals):** Both cutting a chromosome by CRISPR as well as 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 pre-screening for the most typical aberration by qPCR³⁶. 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 rate are chosen using CRISPOR³⁷. After editing, potential off-target effects are determined by Sanger sequencing the top 5 off-target hits from both MIT and CFD scoring algorithms³⁸⁻⁴⁰. 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-profiling³⁵, but if a lab 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 1⁴¹. Variability of alleles between individuals can be visualized by analyzing PCR products on a 2% agarose gel (wt/vol) (PCR described in Step 4, annealing temperature: 68 °C, extension time: 1 min, primers: GTCTTGTTGGAGATGCACGTGCCCTTGC, GAAACTGGCCTCCAAACACTGCCCGCG). In **animals, genetic background can be determined**. In case that 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 (like the Nnt deletion or Crb1rd8 mutation in mice⁴²⁻⁴⁴) 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 described²⁷.
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 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 (EB)-based differentiation followed by identification of expression markers for all 3 germ layers can also be performed⁴⁵.

assaying more SNPs in a region of interest. We applied the Illumina Global Screening Array v2 genotyping chip, which is an affordable option commonly used for GWAS studies²² that 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) etc. 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, so 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 lab performing CRISPR genome editing. This includes performing and optimizing PCR and qPCR, as well as 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 qPCR 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, repetitive sequence etc. 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, qPCR is not an optimal OnTE detection method if very large edits are desired, such as kilobase-sized insertions, as 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 qPCR 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 qPCR but are also too small to be detected by SNP sequencing or microarrays. For example, translocation of a small region containing the entire qPCR amplicon to another genomic position or large inversions would not affect copy number values of qPCR and zygosity in SNP microarrays. However, as the genomic cut by Cas9 is the initial trigger for the formation of OnTE, we expect that the large majority of OnTE will affect the immediate region around the target site and therefore, events escaping detection would be extremely rare.

qPCR and SNP genotyping is a crucial quality control measure to increase reliability of CRISPR-based studies by excluding edited cells and organisms with OnTE. However, it is important to keep in mind that while our tools are useful to exclude OnTE, 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 performing OnTE detection 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 analysed organisms, which is not always given 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 qPCR/bulk SNP genotyping, such as targeted deep sequencing. For this reason, prevention of unwanted OnTE is currently mainly performed by backcrossing of edited animals, which separates alleles with OnTE 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 RNA²³ or by accelerated proteasomal degradation of Cas9²⁴. Li et al.²⁵ recently reported dramatic decrease of 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 broader 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 in a variable genetic background. While 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 labs, 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 (like many common rat models or diversity outbred (DO) mice²⁶, 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, https://scicrunch.org/resolver/CVCL_RM92, https://web.expasy.org/cellosaurus/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 2X 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)
- 50x TAE Buffer (Thermo Fisher Scientific, cat. no. B49)
- SYBR Safe DNA Gel Stain (Invitrogen, cat. no. S33102) **! CAUTION** SYBR Safe has possible mutagenic effects. Wear full PPE and handle with care in a designated separate working area.
- Gene ruler 100 bp plus DNA ladder (Thermo Fisher Scientific, cat. no. SM0323)

qPCR

- PrimeTime Gene Expression Master Mix (IDT, cat. no. 1055770)
- human TaqMan Copy Number Reference Assays (e.g. 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 (e.g. mouse Tfr1 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 qPCR (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 3' 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 qPCR 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+ v2 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 100% (Merck, cat. no. I9516)
- Formamide, 99% (Carl Roth, cat. no. 6749.1) ! **CAUTION** Formamide is toxic, use under 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 qPCR, use plates that fit to the qPCR machine (e.g. 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 1250 µl, long (Sarstedt, cat. no. 70.1186.210)
- 0.8 ml storage plate (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 to 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 (0.1-2.5µl; Eppendorf, cat. no. 312000011)
- Pipette (2-20µl; Eppendorf, cat. no. 312000038)
- Pipette (20-200µl; Eppendorf, cat. no. 312100082)
- Pipette (100-1000µl; 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)
- Multi-Sample BeadChip Alignment Fixture (Illumina)
- Robot BeadChip Alignment Fixture (Illumina)
- Robot Tip Alignment Guide-G (Illumina)
- Hyb chambers, gaskets and inserts (Illumina)
- Te-Flow LCG flow-through chambers, with black frames, LCG spacers, LCG glass back plates and clamps (Illumina)
- Wash dishes and racks (Illumina)
- Hybridization Oven (Illumina)
- iScan scanner with software (Illumina, <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 V4.3.10 (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 v2.0.5 with genotyping module v2.0.5. (Illumina, <https://emea.illumina.com/techniques/microarrays/array-data-analysis-experimental-design/genomestudio.html>)
- *iScan Control* software v4.0.0 (Illumina, <https://emea.support.illumina.com/downloads/iscan-control-software-release-notes.html>)

Reagent Setup

Preparation of qgPCR probe. Centrifuge the tube 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 for up to two 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 two years.

TAE buffer solution for electrophoresis. Dilute 50x TAE buffer to a 1x working solution using deionized water.

Agarose gels for electrophoresis. Prepare 2% (wt/vol) agarose solution in 1x TAE buffer. Microwave solution in a shatter proof glass container without the lid until agarose has dissolved. Add 10,000x 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 one week.

Procedure

CRISPR editing in cells or organism of interest • TIMING 1-3 month, depending on cell type and organism

- 1 Perform genome editing in cells or organism of interest according to published efficient CRISPR protocols (e.g.²⁷⁻³¹). Briefly, this includes designing a suitable guide RNA (gRNA) for the targeted locus and expressing the gRNA 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.

Genomic DNA extraction • TIMING 2 h

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

▲CRITICAL Stored genomic DNA (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 k – 1 M 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 1,000g 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 genomic DNA 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 H₂O To 10 µl –

- 5 Run a gradient PCR using the following thermocycler program with annealing temperatures varying between 50-70 °C in 4 °C steps (6 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 etc. is used for the PCR reaction, thermocycling conditions might need to be adjusted according to manufacturer's instructions.

Step	Temperature	Time	No. of cycles
1 Initial Denaturation	94 °C	5 min	1×
2 Denaturation	94 °C	30 s	
3 Annealing	50-70 °C	30 s	35×
4 Extension	68 °C	30 s	
5 Final Extension	68 °C	5 min	1×
6 Hold	10 °C	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% agarose (wt/vol) 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 RFLP 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 (RFLP) assay (as described e.g. here^{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 quantitative genotyping PCR (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 should 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 to the specs 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 performing duplex qPCR reactions (i.e. two targets

are amplified in a single well), 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 qPCR to exclude those with OnTE • 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 UV-treated clean hood for pipetting the qPCR.

▲ **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 qPCR and yielded samples with only one allele at the same targeted locus, fresh gDNA from these samples can be added as additional positive control.

11 Determine DNA concentration of genomic DNA samples using a spectrophotometer. Prepare 9.5 ul of a 10 ng/ul dilution of gDNA with DNase-free H₂O for triplicate reactions (see Step 12) in a new microcentrifuge tube (or 6.5 ul for duplicate reactions).

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

12 *Prepare the following qPCR 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 triplicates. Volumes provided are for one reaction in one well; scale up all components as necessary and add around 10% extra volume to account for small pipetting errors.

Components	Amount	Final concentration
2× PrimeTime Gene Expression Master Mix	7.5 µl	1×
20× qPCR assay (see reagent setup)	0.75 µl	1×
20× TaqMan Copy Number Reference Assay	0.75 µl	1×
DNase-free H ₂ O	3 µl	–

▲ **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 duplicates to save material cost.

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

14 Pipette 15 ul of qPCR 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 1,000g 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 qPCR reaction on a quantitative PCR thermocycler using the following program.

Step	Temperature	Time	No. of cycles
1 Polymerase activation	95 °C	5 min	1×
2 Denaturation	95 °C	15 s	40×
3 Annealing/Extension	60 °C	1 min	
4 Hold	10 °C	Until ready to process	1×

16 Examine amplification plot of qPCR reactions. It should show a baseline at the beginning, a logarithmic amplification phase (usually between Cycle 20-30 with the threshold for C_t value determination, see Fig. 3 c, d) and a plateau towards the end of the run. Discard samples with abnormal amplification plots for analysis. Detailed discussion on baselines and amplification plots can be found elsewhere¹⁶.

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 higher than 0.2.

From our experience, C_t values of qPCR assays for target genes may vary but are usually around 25-26. C_t values for TERT are usually between 27-28. 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 results³².

? 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):

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

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

? TROUBLESHOOTING

19 Analyze qPCR products by agarose gel electrophoresis: Load one entire 15 μ l reaction together with 2 μ l DNA ladder on a 2% agarose (wt/vol) gel with SYBR Safe (10000 \times). 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 Tfrca at 91 bp).

? TROUBLESHOOTING

20 Edited single cell clones/animals with 2 alleles in the qPCR reaction and one single PCR product on the agarose gel do not contain OnTE such as large deletions, large insertions or other complex rearrangements. Exclude all other clones or animals.

Select methods to exclude loss-of-heterozygosity (LOH) • TIMING 1 h

21 After qPCR, 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 advantages and disadvantages in Box 1 or perform both assays in parallel.

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 100 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 minor allele frequency of ≥ 0.3 under general variant filters.

▲ CRITICAL STEP The frequency of genomic variation strongly varies between different loci. We recommend performing the analysis with around 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 cut-off for the global minor allele frequency 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 ₂ O	To 50 µl	

- 25 Run the following program on a thermocycler:

Step	Temperature	Time	No. of cycles
1 Initial denaturation	94 °C	2 min	1×
2 Denaturation	94 °C	30 s	
3 Annealing	60 °C (50-70 °C)	30 s	35×
4 Extension	68 °C	1 min/1,000 bp	
5 Final extension	68 °C	5 min	1×
6 Hold	10 °C	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 OnTE

● **TIMING 1 d**

- 28 Using the same assay, analyze edited cell lines to check if SNP zygosity has 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 genomic DNA 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 v2 genotyping chip for analysis in human cells, which provides reasonably dense and well-balanced genome-wide coverage at a relatively inexpensive price (about US\$50). A chip suitable for mouse could be the GGP GIGA-MUGA-24 that 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 μ l DNA on a 0.8% (wt/vol) agarose gel. Add 4 μ l gDNA (30-100 ng/ μ l) to a 96-well 0.8 ml microplate (MIDI plate).
▲CRITICAL STEP Microarray genotyping workflows of the leading chip manufacturers Illumina and Thermo Fisher Scientific require high-molecular DNA as input since all protocols start with a whole-genome amplification. Accurate quantification is crucial so UV measurement methods are not recommended, as RNA or extraction contaminants may lead to an overestimation of the quantity. Instead, methods using fluorophores specific for double-stranded DNA like 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 μ l MA1 and 4 μ l 0.1 N NaOH into each well using a multichannel pipette, seal plate with adhesive foil and vortex at 1500 rpm 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 μ l MA2 and 38 μ l 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 x g briefly and incubate the plate in the Illumina Hybridization Oven at 37 °C for 20-24 hours.
- 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 x g for 1 min at room temperature. Remove cap mat, add 25 μ l FMS using a multichannel pipette with filter tips, seal the plate again with the same cap mat, vortex 1 min at 1500 rpm and centrifuge at 280 x g for 1 min.
- 39 Place plate on the 37 °C heat block for 1 hour.
▲CRITICAL STEP Incubation time must not exceed 2 hours.
- 40 *Precipitation.* Thaw PM1 at room temperature, remove the cap mat from the MIDI plate and add 50 μ l PM1 to each well with a multichannel pipette. Seal the plate with the same cap mat, vortex at 1500 rpm for 1 min and centrifuge at 280 x g for 1 min.
- 41 Incubate the plate in the 37 °C heat block for 5 min followed by centrifugation at 280 x g for 1 min. Remove the cap mat and add 155 μ l 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 at 4 °C and centrifuge the plate at 3,000 x g for 20 min.
▲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 hours to air dry the pellets.
▲CRITICAL STEP Keep plate inverted all the time to avoid contamination until pellets are dry. Do not dry pellets longer than 2 hours 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 to 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 seconds and incubate the plate in the Illumina Hybridization Oven at 48 °C for 1 hour.
- 47 Remove the plate from the oven, vortex at 1700 rpm for 1 min and centrifuge at 280 x g for 1 min at room temperature.
 - **PAUSE POINT** The plate can be stored at -20 °C for up to 24 hours 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_his_assay_protocol_user_guide*. You need one chamber for one 96-well plate, i.e., four bead chips. Pipette 400 µl PB2 in each of the 8 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 x g for 1 min. 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 approximately 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 hours.
- 56 Prepare the XC4 reagent by adding 330 ml absolute ethanol. Shake the bottle vigorously for 15 min and leave it on the lab 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 99%, 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 lab 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% 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 hours.
- 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 10 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 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 4 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 approximately 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 OnTE • **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, 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 single-nucleotide polymorphism (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, etc.; 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 to save analysis files and a project name; the sample sheet and raw data files should be provided in a *Data Repository* folder, 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 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 x-axis and, after selecting individual sample in *Columns*, select *B Allele Freq* as 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 x-axis and *Log R ratio* as 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 around 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; poor DNA quality or integrity; locus difficult to amplify by standard PCR	(i) Increase amounts of gDNA in the PCR reaction. (ii) Check for gDNA degradation by gel electrophoresis. Genomic DNA 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 260/280 ratio using a spectrophotometer. If required, further purify gDNA by ethanol precipitation. (iii) 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. (iv) Perform a nested PCR, which can improve sensitivity and specificity.
7, 27	Sanger sequencing does not work	Poor DNA quality; inappropriate amounts of DNA; locus difficult to sequence	(i) Purify PCR product using recommended kit. (ii) Check for presence of PCR product by gel electrophoresis after purification. (iii) Accurately determine DNA concentration by spectrophotometer or fluorometer measurements and dilute to concentration recommended by Sanger sequencing provider. (iv) Use primer for sequencing lying inward of PCR product. (v) Perform TOPO cloning of PCR products into a plasmid vector. Note: Abundance of alleles may vary in the TOPO clones. Heterozygous events will therefore be more difficult to detect.
17	Standard deviation of replicates is above 0.2	Unequal amounts of reaction mix in each well; unequal detection by qPCR machine	(i) Thoroughly mix samples before pipetting. (ii) Ensure accurate pipetting for each sample, which can be achieved by not re-using pipette tips for the same sample or using low retention pipette tips. If available, electronic pipettes can also help to improve precision of pipetting. (iii) Run qPCR machine maintenance to ensure equal detection of all wells, if problem persists, avoid areas with unequal detection, which can occur especially towards the edges of the plate.
17	C _t values are higher than 30	Too little amount of gDNA within the reaction; inhibitors of PCR reaction present	(i) Increase gDNA concentration. (ii) Purify gDNA using recommended kit.
17	No template control gives a signal	Contamination of reaction mix	(i) Repeat qPCR using all fresh reagents and filter tips for pipetting. (ii) If available, using a dedicated and/or UV-treated area for qPCR reactions can further minimize risk of contamination.
17	No amplification detectable in all or some samples	Issues with primers, probe, template, buffers or qPCR machine	See detailed info on qPCR basics in <i>Real-time PCR Handbook</i> (https://www.thermofisher.com/content/dam/LifeTech/global/Forms/PDF/real-time-pcr-handbook.pdf), and troubleshooting in protocol by Weissenborn et. al, 2010 ³³ for detailed discussion and suggestions.
18	Copy number values strongly	Varying integrities of gDNA samples;	(i) Repeat gDNA extraction of all samples together and make sure to treat all samples exactly the same.

	vary between samples and are not close to values of '1' or '2' allele numbers	unequal amounts of gDNA in each sample; restricted accessibility of qPCR assay to gDNA	(ii) 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. (iii) Pre-digestion 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.
18	NHEJ clones have copy number values lower than 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; sample has OnTE at edited locus	(i) Unspecific amplification occurring in multiple samples could affect the efficiency of target amplification and therefore cause unreliable results. Adjusting qPCR conditions like input gDNA amount or primer concentrations or eventually exchanging primers should be tested. (ii) Multiple bands in a single sample might be caused by OnTE 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; an inbred organism was used	(i) Expand the region in which the zygosity of SNPs is investigated by doubling the size of the analyzed genomic region and lowering the cut-off for the global minor allele frequency by 0.1. In principle, there is no limit on 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. (ii) Perform SNP microarray to obtain genome-wide SNP profile. (iii) Perform editing in cell line or animal strain with higher genomic variation. (iv) LOH analysis is not necessary for inbred organisms.
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 .
80	GenomeStudio fails to load certain samples	Incorrect Samplesheet	Using a text editor, check in the Samplesheet if the right directory is provided and 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 organism of interest: 1-3 month, depending on organism
Step 2, Genomic DNA 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 and Sanger sequencing: 2-3 d

Steps 9-10, Design of labeled oligo for quantitative genotyping PCR (qgPCR): 1 h
Steps 11-20, Analysis of edited cells or organisms by qgPCR to exclude those with OnTE: 1 d
Step 21, Select methods to exclude loss-of-heterozygosity (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 OnTE: 1 d

Steps 29-76, Analysis of genomic DNA 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 OnTE: 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 on-target effects (OnTE) 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 disease³⁴. 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 qPCR analysis. The underlying principle of the qPCR 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 threshold cycle (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 OnTE will have the same C_t value as the edited control. Samples with OnTE 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 OnTE at varying frequencies, ranging from 17% to 57% depending on the editing system (HDR and plasmid vs. NHEJ and RNP, respectively). These clones need to be excluded for further analysis. In general, OnTE frequency may also strongly vary based on the edited locus or organism.

After successful exclusion of OnTE like large deletions or complex rearrangements by qPCR, 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 v2 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, 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 since such cases are likely 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.

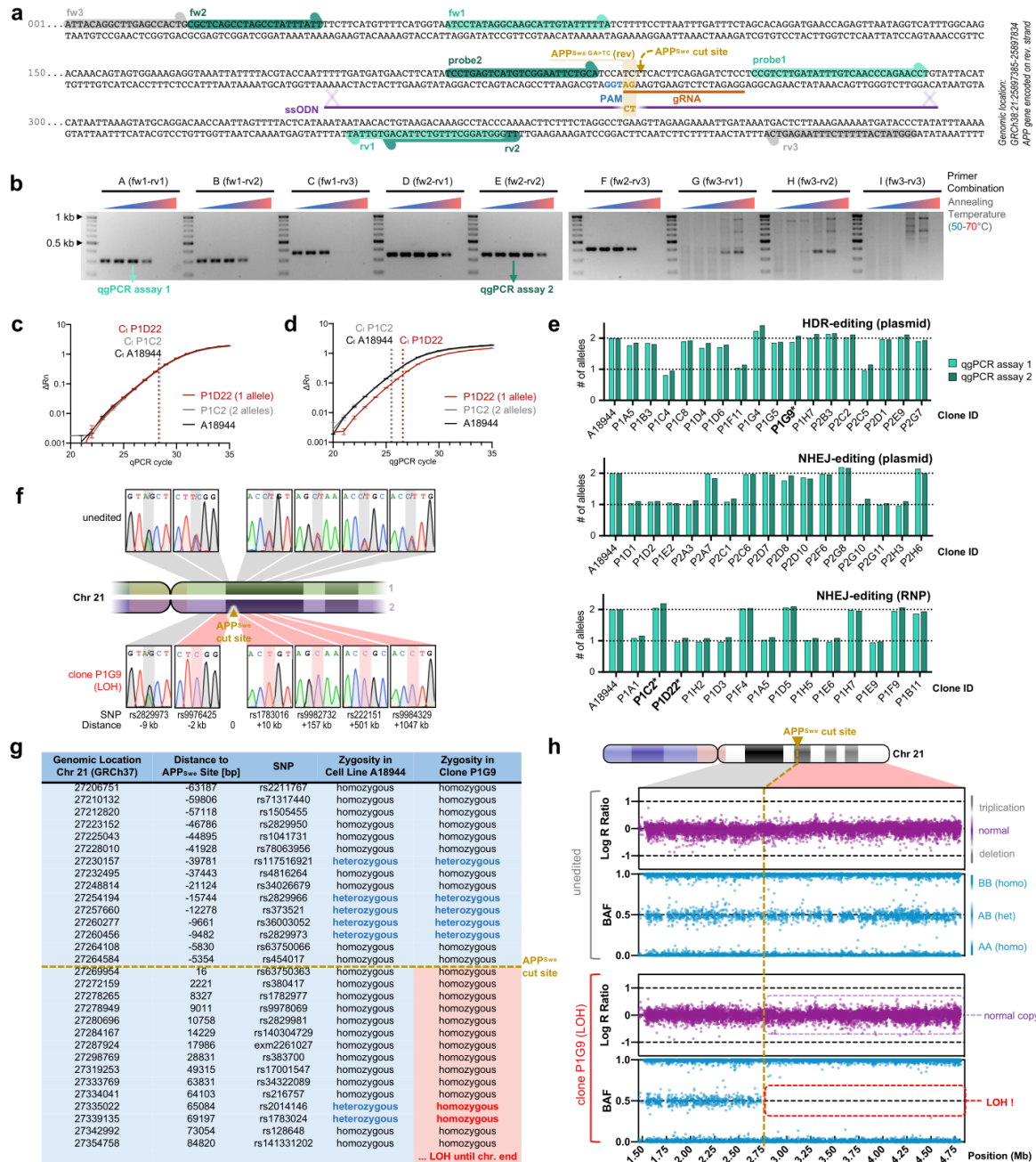


Fig. 3 | Anticipated results for OnTE detection, example of CRISPR editing at the APP^{Swe} locus in human iPSCs. a, Overview of CRISPR editing and positions of primers and probes for two independent qPCR assays at the APP^{Swe} locus. ssODN: single-stranded oligodeoxynucleotide, fw: forward, rv: reverse. **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% agarose gel (wt/vol). Two primer combinations (A and E) were selected for the qPCR assays 1 and 2 shown in panel e). **c**, qPCR amplification plot of human TERT reference gene for two NHEJ-RNP-edited clones and the unedited 'parent' line A18944. (ΔR_n = normalized reporter fluorescence signal - baseline) Data are represented as means \pm SEM (n=2). **d**, qPCR amplification plot of APP^{Swe} qPCR assay 1 for two NHEJ-RNP-edited clones and the unedited A18944. Note shift in ΔR_n for clone P1D22 indicating OnTE. (ΔR_n = normalized reporter fluorescence signal - baseline) Data are represented as means \pm SEM (n=2). **e**, qPCR 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-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 v2 genotyping chip and zygosity in unedited cell line A18944 and LOH clone P1G9 (HDR-plasmid editing). **h**, SNP microarray analysis determining Log R ratios and B allele frequencies (BAF) 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 (https://scicrunch.org/resolver/CVCL_RM92) has been validated by fingerprinting, pluripotency assays and karyotyping, and for absence of mycoplasma. Figure panels e (upper two graphs), f and h are modified from Ref. 13.

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 (ForTraGmbH) (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., 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.

Competing Financial Interests

The authors declare no competing interests.

Reporting Summary

Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

Data that support the findings of this study are available in Mendeley Data, doi: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. *Nature Protocols* **15**, 1–28 (2020).
8. Gkazi, S. A. Quantifying CRISPR off-target effects. *Emerging Topics in Life Sciences* **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**, 15464–10 (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 Research* **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 Meth* **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. *Nature Protocols* **1**, 1559–1582 (2006).
17. Bi, C. *et al.* Long-read Individual-molecule Sequencing Reveals CRISPR-induced Genetic Heterogeneity in Human ESCs. *bioRxiv* (2020). doi:10.1101/2020.02.10.942151
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. 1–18 (2019). doi:10.1186/s13059-019-1720-5
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 Research* **41**, D21–7 (2019).
 22. Illumina. Infinium Global Screening Array-24 v3.0 BeadChip. <https://www.illumina.com/products/by-type/microarray-kits/infinium-global-screening.html> 1–6 (2020). Available at: <https://www.illumina.com/products/by-type/microarray-kits/infinium-global-screening.html>. (Accessed: 22nd July 2020)
 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* 1–12 (2020). doi:10.1038/s41467-020-18391-y
 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. *Nature Protocols* **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. *Nature Protocols* **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 in Biotechnology* **37**, 761–774 (2019).
 33. Weissenborn, S. J., Wieland, U., Junk, M. & Pfister, H. Quantification of beta-human papillomavirus DNA by real-time PCR. *Nature Protocols* **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 Reports* **7**, 998–1012 (2016).
 37. Concordet, J.-P. & Haeussler, M. CRISPOR: intuitive guide selection for CRISPR/Cas9 genome editing experiments and screens. *Nucleic Acids Research* **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**, 4063–14 (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 quantitative 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 Reports* **8**, 1340–1353 (2017).

Supplementary Table 1. Primer, sgRNA and ssODN sequences.

Primers/probes for PCR optimization and qgPCR assays (Fig. 1b, 3a, b, d, e)		
Purpose	Name	Sequence 5'-3'
PCR optimization, qgPCR 1	fw1	ATCCTATAGGCAAGCATTGTATTTTAA
PCR optimization, qgPCR 1	rv1	GGGTAGGCTTTGTCTTACAGTGTAT
qgPCR 1	probe1	56-FAM/CCGTCTTGA/ZEN/ TATTTGTCAACCCAGAACCT/3IABkFQ
PCR optimization, qgPCR 2	fw2	CGCTCAGCCTAGCCTATTTATT
PCR optimization, qgPCR 2	rv2	TTGGGTAGGCTTTGTCTTACAG
qgPCR 2	probe2	56-FAM/TCCTGAGTC/ZEN/ ATGTCGGAATTCTGCA/3IABkFQ
PCR optimization	fw3	ATTACAGGCTTGAGCCACTG
PCR optimization	rv3	GGGTATCATTTTCTTTAAGAGTCA
Primers for nearby SNP genotyping (Fig. 1f, 3f)		
Locus	Name	Sequence 5'-3'
APP ^{Swe} _-9kb	rs2829973_fw+Sanger	GCGGGCATTITTTCACTCTAA
APP ^{Swe} _-9kb	rs2829973_rv	ACCTGAACACAGGGAGTTGC
APP ^{Swe} _-2kb	rs9976425_fw	TGCTATTGCACATGTAACAGACT
APP ^{Swe} _-2kb	rs9976425_rv	GGGTAGGCTTTGTCTTACAGTGTAT
APP ^{Swe} _-2kb	rs9976425_Sanger	CAAGGTCAGGAGGTCGAGAG
APP ^{Swe} _+10kb	rs1783016_fw+Sanger	TTGAATCAAATCCTTTGCTGT
APP ^{Swe} _+10kb	rs1783016_rv	CAAAATGTGAAGCTGCCTTCT
APP ^{Swe} _+157kb	rs9982732_fw+Sanger	TGCCATTATACCCCCACAAT
APP ^{Swe} _+157kb	rs9982732_rv	CAAAATGGCATCCAAAACCT
APP ^{Swe} _+501kb	rs222151_fw+Sanger	CCAAATATCCAGGTGCCTTC
APP ^{Swe} _+501kb	rs222151_rv	AGCAATGCAAGAGCAGCCTA
APP ^{Swe} _+1047kb	rs9984329_fw+Sanger	AATGCCAAACGCAGAAAACCT
APP ^{Swe} _+1047kb	rs9984329_rv	GTGGGAGAAGTCCCAGAGA
Gene editing components (Fig. 3a)		
Purpose	Strand	Sequence 5'-3' (*bold = cut site, blue = intended mutation, <u>underlined</u> = NGG)
APP ^{Swe} sgRNA	-	GGAGATCTCTGAAGTGAAGAT <u>GG</u>
APP ^{Swe} ssODN for HDR	-	CAGGTTCTGGGTTGACAAATATCAAGACGGAGGAGAT CTCTGAAGTGAAT <u>CTGG</u> ATGCAGAATTCCGACATGACT CAGGATATGAAGTTCATCATCAAAA