CRISPR/Cas9 deletions induce adverse on-target genomic effects leading to functional DNA in human cells

Keyi Geng[1], Lara G. Merino[1], Linda Wedemann[1], Aniek Martens[1], Małgorzata Sobota[1], Jonas Nørskov Søndergaard[1], Robert J. White[2], Claudia Kutter[1]*

[1] Department of Microbiology, Tumor, and Cell Biology, Karolinska Institute, Science for Life Laboratory, Sweden
[2] University of York, Department of Biology, York, United Kingdom
* To whom correspondence should be addressed. Tel: +46 (0) 70 4933896. Email: claudia.kutter@ki.se

KEYWORDS
unintended CRISPR/Cas9 edits, genomic duplication-inversion-integration, droplet-based target enrichment, long-read sequencing, *de novo* sequence assembly

LIST OF ABBREVIATIONS
Δ, deletion; BP, breakpoint; Cas9, CRISPR-associated protein 9; CRISPR, clustered regularly interspaced short palindromic repeats; DC, daughter cell; DSB, double-strand break; *E. coli*, *Escherichia coli*; gRNA, guide RNA; H3K4me3, histone 3 lysine 4 trimethylation; HAP1, chronic myeloid leukemia cell line; HepG2, hepatocellular carcinoma; LRS, long-read sequencing; MMEJ, microhomology-mediated end joining; NHEJ, nonhomologous end-joining; ONT, Oxford Nanopore Technology; PAM, protospacer adjacent motif; Pol III, RNA polymerase III; tRNA, transfer RNA

ABSTRACT
The CRISPR/Cas9 system is widely used to permanently delete genomic regions by inducing double-strand breaks via dual guide RNAs. However, consequences of Cas9 deletion events have not been fully investigated. To characterize Cas9-induced genotypic abnormalities in human cells, we utilized an innovative droplet-based target enrichment approach followed by long-read sequencing and coupled it to customized *de novo* sequence assembly. We here describe extensive genomic disruptions by Cas9, involving a genomic duplication and inversion of the target region as well as integrations of exogenous DNA at the double-strand break sites. Although these events altered the genomic composition of the on-target region, we found that the aberrant DNA fragments are still functional, marked by active histones and bound by RNA polymerase III. Our findings broaden the consequential spectrum of the Cas9 deletion system, reinforce the necessity of meticulous genomic validations and rationalize extra caution when interpreting results from a deletion event.

INTRODUCTION

The CRISPR/Cas9 system has revolutionized genome engineering approaches. Various toolsets have been developed, which allow loss-of-function perturbations of functional genomic elements in a simple and efficient manner. To modify the genome, the components of the CRISPR/Cas9 system are exogenously introduced into human cells (1–3). In the presence of a guide RNA (gRNA) that is complementary to the target site and next to a protospacer adjacent motif (PAM), the expressed Cas9 endonuclease induces a double-strand break (DSB) at a specific genomic site. If an exogenous homologous DNA template is supplied, the DSB is repaired by the homology-directed repair to introduce specific mutations or insertions of desired sequences (4). Alternatively, microhomology-mediated end joining (MMEJ) can be triggered by short (5 to 25 bp) overlapping sequences that allow for recombination at the DSB (5). In contrast, in the absence of a homologous template, the cell uses non-homologous end-joining (NHEJ) to resolve the DSB. This usually results in small deletions or insertions of a few nucleotides (1). Instead of introducing single nucleotide or short sequence deletions, longer DNA regions can be excised from the genome by using dual gRNAs that flank the target region and guide Cas9 to induce two DSBs (6, 7). Through this approach, a target region of about 1 Mb was successfully removed from the genome in mouse cells (8). In addition, a variety of genomic regions have been modified using the dual gRNA system. For example, the genomic manipulation of enhancers (9, 10), chromatin loop anchors (11, 12), protein-coding genes (7, 13) and noncoding genes (14) revealed disease-associated functional elements and genes. Although a broad range of functional sequences has been successfully deleted from the genome, transfer RNA (tRNA) genes have not been targeted by using the dual gRNA system in human cells.

tRNA genes are transcribed by RNA polymerase III (Pol III) and marked by active histones, such as histone 3 lysine 4 trimethylation (H3K4me3) (15–19). tRNAs represent one of the most abundant RNA types. These circa 73 nt long RNA molecules act as physical adapter molecules during translation, in which the anticodon regions of a tRNA molecule recognize the complementary codon of a messenger RNA and adds an amino acid to the growing polypeptide chain. tRNAs have a wide spectrum of additional functions. For example, a tRNA molecule can be cleaved into tRNA halves or fragments, which impair gene regulation transcriptionally or post-transcriptionally and contribute to diseases, such as cancers and neurodegenerative disorders (20–26). Furthermore, the tRNA gene locus itself functions in genome organization by acting as insulator (27).

In order to investigate tRNA gene functionality, we deleted two tRNA genes from the genomes of human haploid chronic myeloid leukemia (HAP1) and hyperploid hepatocellular carcinoma (HepG2) cells using the CRISPR/Cas9 system with dual gRNAs. By applying the droplet-based target enrichment method (Xdrop) (28) followed by Oxford Nanopore Technology (ONT) long-read sequencing (LRS) (29), we uncovered unexpected genomic alterations at the target site. Although Cas9 induced genomic cuts and our target region was successfully excised, we found that the deleted region got duplicated, inverted in tandem orientation and reintegrated at the original genomic locus in HAP1 cells. In HepG2 cells, we also observed a duplication event but in divergent orientation. Furthermore, additional sequences originating from the CRISPR/Cas9-gRNA transfection vector and the *Escherichia coli* (*E. coli*) genome got integrated downstream of the duplicated target locus. Despite our initial attempt to delete tRNA genes to impair their expression, the aberrant genomic modifications at the original locus resulted in actively transcribed tRNA genes in HAP1 and HepG2 cells. This highlights that CRISPR/Cas9-based genomic engineering can result in undesired on-target effects. Our work also underscores the complexity of human DNA repair mechanisms in the presence of the powerful prokaryotic Cas9 nuclease and raises serious concerns when studying the functionality of modified genomes. We therefore recommend

examining modified cells beyond Sanger sequencing and PCR validations since unintended genomic alterations may occur and could be missed by standard verification methods.

RESULTS

The target region remained detectable and functional in Cas9 deletion clones

To test the effectiveness of CRISPR/Cas9 for deleting tRNA genes, we focused on a pair located on the human chromosome 17 (**Fig. 1A**). Since tRNA genes belong to one of the biggest multi-copy gene families (15), in which individual gene family members are identical in sequence composition, we designed gRNAs mapping to the unique 5' and 3' flanking regions. That enabled Cas9 to excise an 870bp long genomic fragment with two tRNA genes (**Fig. 1B**). We transfected HAP1 and HepG2 cells with two Cas9 plasmids carrying one of the two gRNAs and a small size plasmid to enhance transfection efficiency (Method, **Fig. S1A**). Since the Cas9 vector contains the puromycin resistance gene, we used antibiotic selection to identify positively transfected cells and propagated further single cell-derived clones. To validate the clonal deletion of the target region, we performed a PCR with flanking region-specific primers. The size of the PCR product was indicative of a successful deletion when inspected by agarose gel electrophoresis and its sequence content was verified by Sanger sequencing (**Fig. 1B, Fig. S1A**). Overall, 94 HAP1 and 90 HepG2 single cell-derived clones were generated. Among them, 5 HAP1 and 17 HepG2 clones contained a deletion (**Fig 1C, Fig S1B-D**). We focused our subsequent analysis on two validated deletion clones, referred to as HAP1 Δ72 and HepG2 Δ15 (**Fig. 1C, Fig. S1B**).

To study the consequences of tRNA gene deletions, we either utilized published Pol III ChIP-seq data in HepG2 unmodified cells (30) or profiled the genome-wide binding of H3K4me3 and Pol III in both deletion and control clones. We only considered reads with high mapping quality (Methods). Surprisingly, we observed binding of H3K4me3 and Pol III to DNA sequences within the target region although at a lower magnitude in HAP1 and HepG2 deletion clones when compared to their respective control clones (**Fig. 1D**). To explain this observation, we considered several possibilities that can lead to imperfect genomic Cas9 deletions in cells. Previous reports described that applications of the dual gRNA system (I) could lead to heterozygous deletion clones (8) whereby one allele carries the deletion and the other allele is unmodified, (II) mutations could have occurred at the PAM sequences (44, 45) preventing the recruitment of Cas9 and its cutting activity at the gene locus, (III) the Cas9 cleavage might be unsynchronized at the two DSB sites (46), in which one DSB would have been repaired before the induction of the second DSB or the target sequence could have been excised and then either (IV) inverted or (V) duplicated (8, 13, 47–49). Scenarios I to IV can be excluded since we would have detected a PCR product corresponding to the size of the unmodified allele (ca. 1,200 bp) (**Fig. 1C, Fig. S1C-D**). We can also exclude scenario V since we found ChIP-seq reads spanning the excision site in the control but not in the deletion clones (**Fig. 1E**). This prompted us to consider alternative consequences occurring as a result of Cas9 genome engineering.

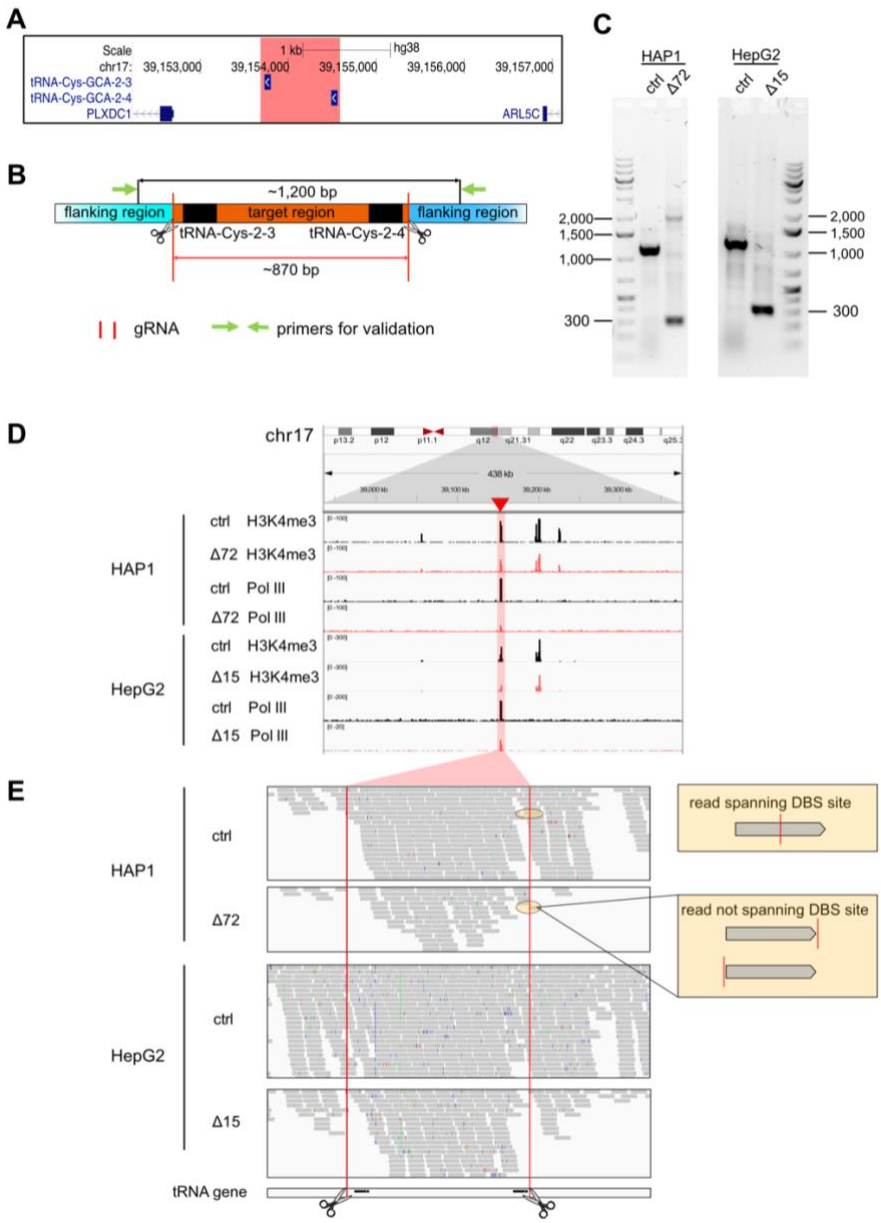

Fig 1. The target region remained functional in deletion clones. (A) Genome browser view shows the genomic location of our target locus (red). Arrows denote directionality of gene transcription. **(B)** Illustration of the design strategy of Cas9 dual gRNA deletion strategy. Cas9 cut sites (red horizontal lines and scissors) and primers for validation (green arrows) are indicated. The target region (orange) containing two target tRNA genes (black) is around 870 bp. The size of the PCR product in control clones is around 1,200 bp. **(C)** Agarose gel separates PCR products by size to validate deletion clones when using validation specific primers (B). Selected DNA marker bands [in bp] are depicted. **(D)** Genome browser view demonstrates normalized ChIP-seq reads for H3K4me3 and Pol III covering the target loci (highlighted in red) in the HAP1 control (ctrl) and Δ72 as well as HepG2 ctrl and Δ15 clones. **(E)** Alignment tracks show individual H3K4me3 ChIP-seq reads of the deleted and surrounding region as in (D). DSB sites (red lines and scissors) and tRNA gene locations (black) are indicated. Examples of reads spanning (top) or not spanning (bottom) DSB site are illustrated (yellow box).

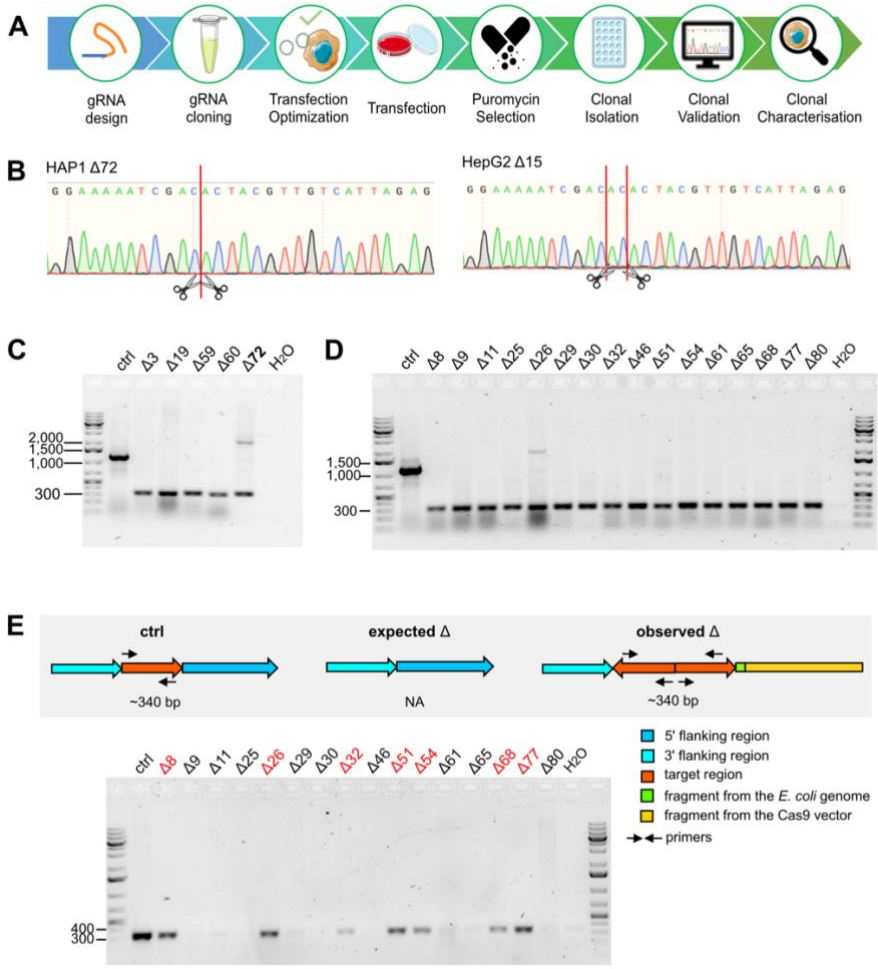

Fig S1. Validation of the deletion in HAP1 and HepG2. (A) Schematic illustration of the workflow to obtain single cell-derived deletion clones. **(B)** Chromatograms confirm deletions in the HAP1 Δ72 and HepG2 Δ15 clone. DSB sites induced by CRISPR/Cas9 are shown by red lines and scissors. **(C-E)** Agarose gels confirm the size of the obtained PCR products to validate deletion events (320bp) using primers annealing to the flanking regions of the target sites (Fig. 1B) in HAP1 (C) and HepG2 (D-E) clones. Additional primers to distinguish unmodified control, expected and observed deletion events are used (E). Marker bands specify DNA size in bp. In C, the HAP1 clone with duplicated target regions is bolded. In D, HepG2 Δ15 deletion clone is not shown here but in Fig. 1C. In E, deletion clones confirmed in D with a reinserted target region (340bp) are indicated in red.

Applying Xdrop in deletion clones confirmed the genomic remodeling of the target region

The Xdrop technology has recently been applied to validate CRISPR/Cas9 genome modifications (28, 38). In order to investigate the on-target editing outcomes in our Cas9 deletion clones, we applied the Xdrop approach to enrich for sequences containing our CRISPR/Cas9-targeted genomic region in the HAP1 Δ72 and HepG2 Δ15 deletion clones. Subsequently, we identified the sequence content of the Xdrop-enriched molecules by ONT LRS. On average, we obtained 217,000 and 179,000 reads of a median size of 4,600 and 5,200 bp in HAP1 and HepG2, respectively. Read coverage at the target loci indicated sufficient enrichment (**Fig. S2A-B**). By aligning both corrected and raw reads to the human reference genome (Methods), we observed sharp declines in coverage at the two DSB sites and no read spanned the two DSB sites in the HAP1 and HepG2 deletion clones (**Fig. S2A**). This supported our conclusion drawn from our ChIP-seq data that the target region was altered in our two deletion clones.

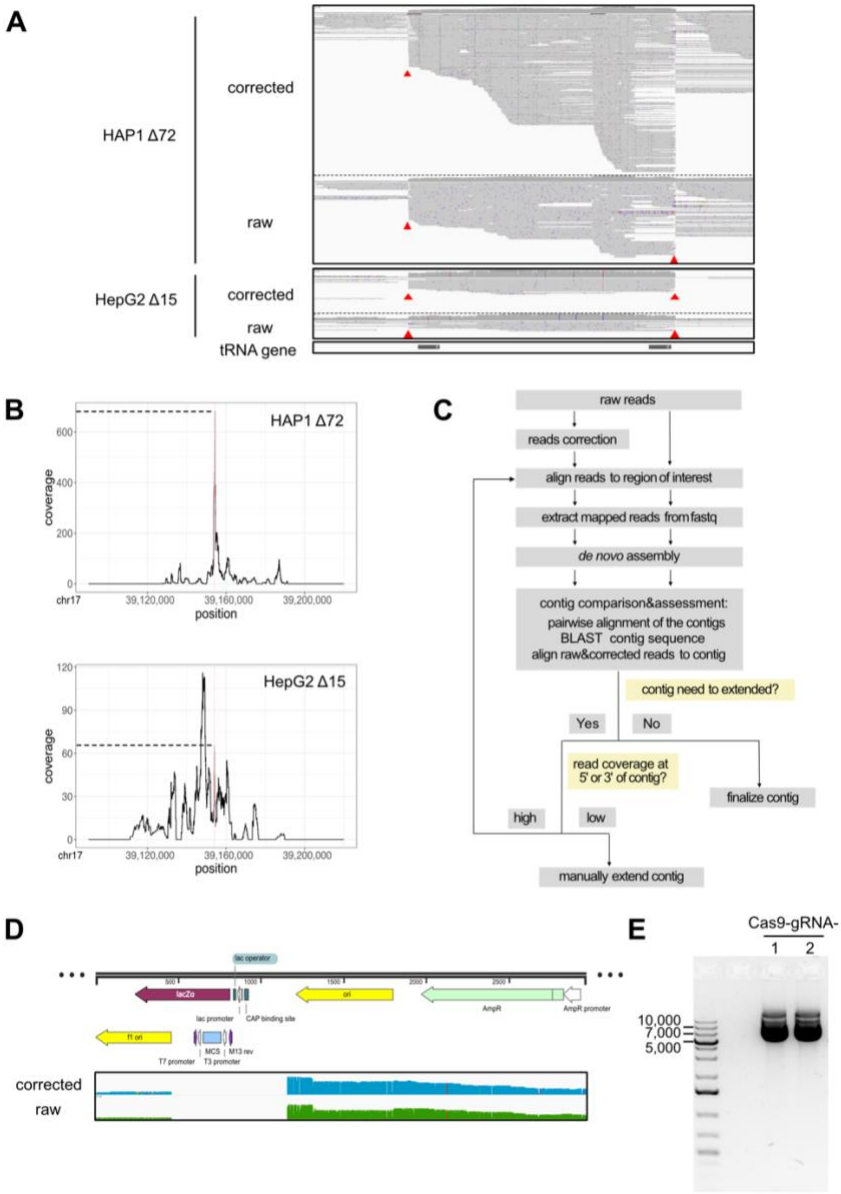

Fig S2. Complex on-target genomic alterations are revealed by Xdrop. (A) Alignment tracks display Xdrop corrected and raw reads at the target loci in HAP1 Δ72 (top) and HepG2 Δ15 (bottom) deletion clones. DSB sites (red triangles) and tRNA genes (black) are shown. **(B)** Line graphs show coverage when aligning raw reads obtained for the HAP1 Δ72 (top) and HepG2 Δ15 (bottom) deletion clones to the human reference genome. The position was defined by the coordinates from the human reference genome. Red lines highlight the location and dashed lines indicate the number of reads covering the target regions. **(C)** *De novo* assembly-based analysis workflow. **(D)** Coverage tracks display Xdrop-LRS corrected and raw reads aligning to the pBlueScript vector. Sequence annotations and directions of gene transcription are labelled in accordance to the pBlueScript vector map. **(E)** Agarose gel confirms the size of the Cas9-gRNA-1 and -2 vectors (around 9,170 bp) after plasmid purification.

A *de novo* assembly-based approach revealed a duplication, inversion and local insertion of the target region in the HAP1 deletion clone

In order to decipher the underlying alterations that caused a rearrangement of the CRISPR/Cas9-targeted region, we employed a customized *de novo* sequence assembly

approach (**Fig. S2C**). For this, we assembled contigs containing the target region and connected them with sequences of the respective flanking regions for each of the two deletion clones.

In our HAP1 Δ72 deletion clone, we identified contigs with three distinct breakpoints that deviated from the human reference genome composition (**Fig. 2A**). One breakpoint (BP2) connected two units of our deleted target region in tandem orientation. This finding is indicative of a duplication event of our original target fragment. The other two breakpoints (BP1 and BP3) connected the duplicated fragments with the 5' and 3' flanking region of our original target region in inverse orientation. This suggested an unexpected event in which the deleted fragment of our target region got duplicated, inverted and inserted at the original gene locus (**Fig. 2A**). To eliminate any confounding effects from our analysis workflow, we inspected both raw and corrected read alignment tracks. We found multiple long reads spanning the contig and several additional reads covering the individual breakpoints confirming on-target genomic alteration events (**Fig. 2A**).

In addition, we performed an independent validation of our *de novo* sequence assembly by designing three sets of primers specific for the flanking and/or target region (**Supplemental table 1**). The different primer combinations and size of PCR products enabled us to discern the genomic composition of our target locus in our unmodified, expected and actually observed deletion clones (**Fig. 2B**). This PCR-based result supported our Xdrop-LRS findings and our conclusion that the target region got duplicated, inverted and inserted in the HAP1 Δ72 deletion clone.

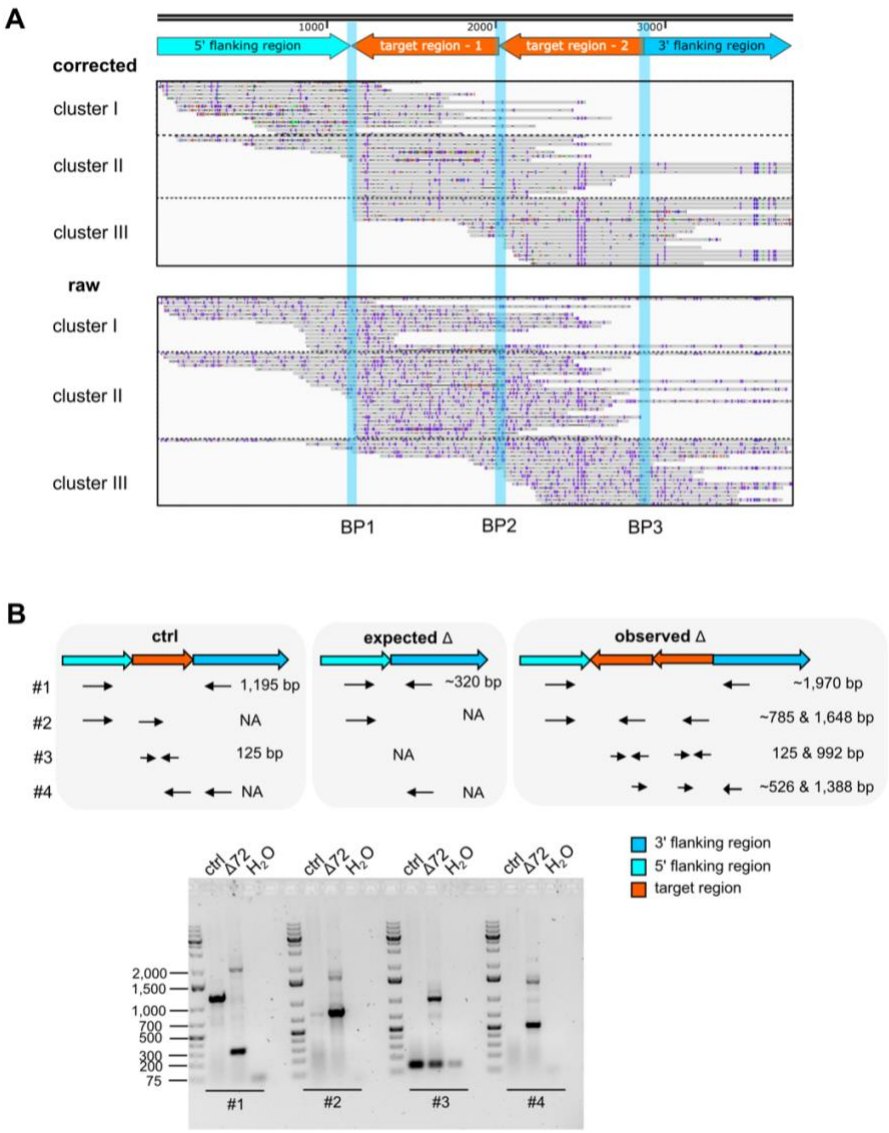

Fig 2. A duplication, inversion and local insertion of target-derived fragments occurred in the HAP1 Δ72 deletion clone. (A) Genome browser view shows Xdrop corrected (top) and raw (bottom) reads aligned against the assembled contig. Arrows (top) show genomic orientation and approximate size of the flanking (5' light and 3' dark blue) and target (orange) regions within the contig. Representative aligned reads for cluster I, II and III support three different breakpoints (BP1-3, blue vertical lines) within the contig. **(B)** Schematic illustration of the primer design strategy to validate contigs by PCR and the expected PCR amplicon length (top) representing scenarios in an unmodified control, expected and observed deletion event. Agarose gel (bottom) confirms the size of the obtained PCR products. Marker bands specify DNA size in bp.

The target sequence got duplicated, inverted and co-integrated with exogenous DNA fragments in the HepG2 deletion clone

Similar to the HAP1 deletion clone, we employed our customized *de novo* sequence assembly approach to investigate potential genomic aberrations of the target region in the HepG2 Δ15 deletion clone. The assembly of our LRS reads revealed four breakpoints (**Fig. 3A**). Similar to the HAP1 Δ72 deletion clone, we found that a duplication of our target region occurred in the HepG2 Δ15 deletion clone. Breakpoint 2 (BP2) connected these two units of the target sequence but in divergent orientation. This suggested that only one of the two units got inverted,

which was further confirmed by BP1 which linked the 5' flanking region of the target to the duplicated target region itself. Resolving the events at the 3' cut site revealed that exogenous DNA fragments were integrated at the DSB site. We searched for high sequence similarity of these fragments (Methods) and found that an approximately 200 bp long fragment aligned perfectly to a part of the ribonuclease R (rnr) gene in the *E. coli* genome (**Fig. 3A**). BP3 joined the duplicated target region and this 200 bp *E. coli* genome fragment. In addition, a more than 6,000 bp long fragment mapped to the CRISPR/Cas9 vector that we used for cellular transfections. BP4 connected the *E. coli* fragment to the CRISPR/Cas9 vector sequence. Reads mapping to the CRISPR/Cas9 vector included sequences of the U6 promoter, the gRNA cloning site, gRNA scaffold, chicken β-actin promoter and the first half of the Cas9 gene (**Fig. 3B**). Remarkably, we found reads carrying the sequence of gRNA-1, which we synthesized to induce the DSB at the 5' flanking region but there were no reads containing the sequence of gRNA-2, suggesting that parts of the Cas9-gRNA-1 but not -2 vector were integrated. At a lower coverage, we also noted reads mapping to sequences of the poly(A) signal, f1 origin of replication (f1 ori), ampicillin resistance gene and a second ori (**Fig. 3B**). Reads mapping to the breakpoints were found in both the raw and corrected reads supporting an accurate contig assembly (**Fig. 3A-B**).

Since we co-transfected the CRISPR/Cas9 and the pBlueScript vector to increase cell transfection efficiency (3), we inspected whether sequences of the pBlueScript vector could have integrated. We aligned the LRS data to the pBlueScript vector and found reads mapping to sequences encoding the f1 ori, ampicillin resistance gene and a second ori (**Fig. S2D**). This part of the pBlueScript vector serves as backbone in many commonly used plasmids. It is therefore not surprising that vector sequences were in part identical. Since we did not detect any reads mapping to the pBlueScript-specific region located between the f1 ori and the second ori, we concluded that the vector sequence that integrated into the HepG2 target region originated from the CRISPR/Cas9 and not from the pBlueScript vector.

In sum, in the HepG2 Δ15 deletion clone, we observed duplication of the target region, inversion of one copy and integration at the target region together with sequences from the *E. coli* genome and the CRISPR/Cas9 vector.

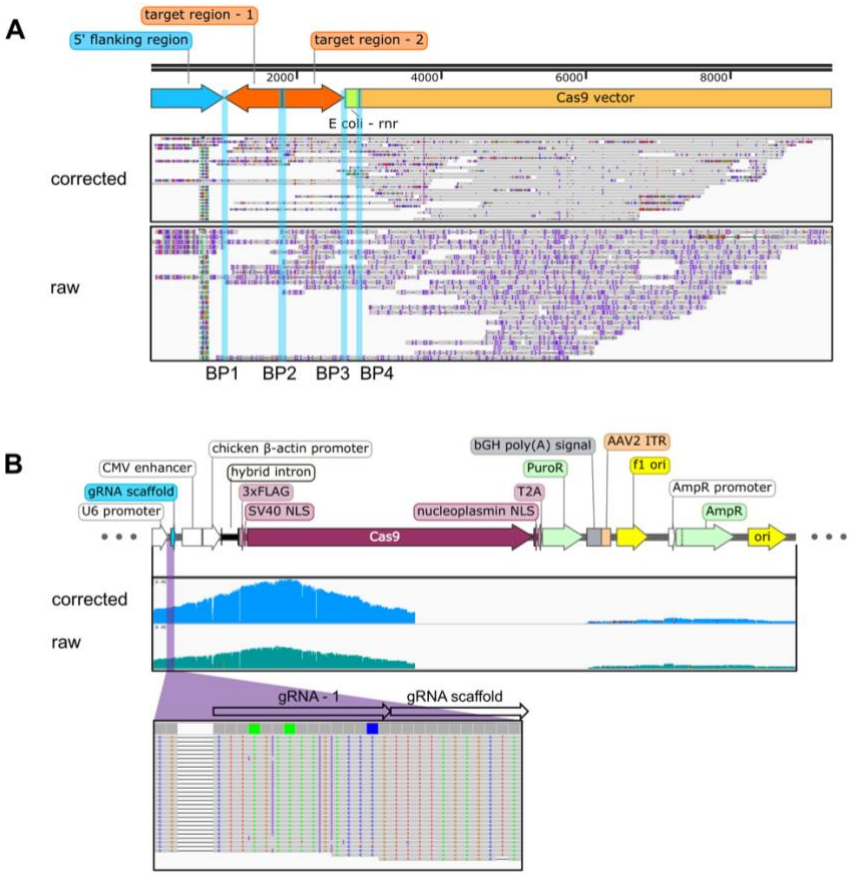

Fig 3. A duplication, inversion of target-derived fragments and integration exogenous DNA fragment arose in the HepG2 Δ15 deletion clone. (A) Genome browser view shows Xdrop corrected (top) and raw (bottom) reads aligned against the assembled contig supporting four different breakpoints (BP1-4, blue vertical lines) within the contig. Arrows (top) show genomic orientation and approximate size of the flanking (blue) and target (orange) regions as well as sequences of the *E. coli* genome (green) and Cas9-gRNA-1 transfection vector (yellow) present within the contig. **(B)** Coverage tracks display Xdrop-LRS corrected and raw reads aligning to the Cas9 vector (top). Sequence annotations and directions of gene transcription are labelled in accordance to the Cas9 vector map. The gRNA-1 cloning site and reads mapping to this region are highlighted (purple horizontal line) and magnified (bottom).

On-target genomic alterations can occur frequently

In order to estimate the approximate frequency of Cas9-induced genomic alterations, we tested whether on-target insertion events occurred in our other confirmed HAP1 and HepG2 deletion clones. For HAP1, we used primers recognizing the 5' and 3' flanking sequences of the target region (**Fig. 1B, Supplemental table 1**). We inspected five HAP1 clones of which four contained the expected homozygous deletion (320 bp) and only the HAP1 Δ72 deletion clone showed an extra band (~2,000 bp), which is indicative of the unexpected on-target effect (**Fig. 4A, Fig. S1C**). In HepG2, we tested Δ15 and 16 additional clones in which the target region was deleted (**Fig. 1C, Fig. S1D**). Since our LRS contig suggested an integration of genomic sequences larger than 7,900 bp, we could not quantify the frequency of on-target events by using primers spanning the 5' and 3' flanking regions. We therefore designed a pair of primers annealing to the target region. Although we had confirmed the deletion in these 16 HepG2 clones, we still detected the target region in seven deletion clones (**Fig. 4A, Fig. S1E**). However, we cannot determine the full extent of the on-target genomic alterations in these clones without additional Xdrop-LRS validations. In sum, aberrant genomic changes at the on-target locus can occur at varying frequencies in HAP1 (20%) and HepG2 (47%) clones (**Fig. 4A**).

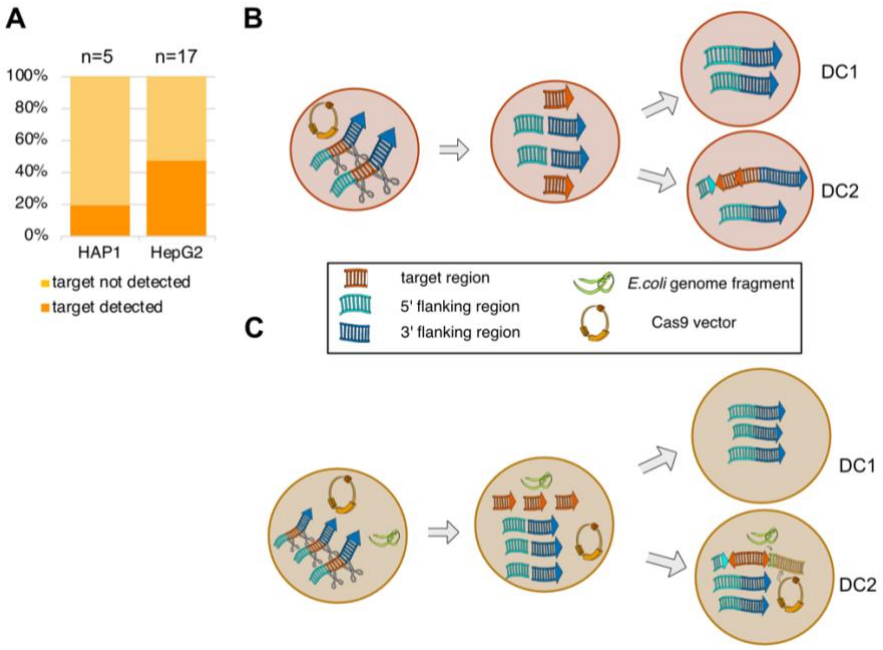

Fig 4. The frequency of on-target genomic alterations and a proposed model. (A) Stacked bar plot indicates the frequency of on-target genomic alterations in the validated HAP1 (n=5) and HepG2 (n=17) deletion clones. **(B-C)** Hypothetical model of on-target genomic alterations in HAP1 (B) and HepG2 (C). Cas9 caused DSBs cleaving the target region from the genome. Fragments were inverted and reinserted in one daughter cell (DC). In HepG2, additional exogenous DNA sequences from the *E. coli* genome and the Cas9 vector carrying gRNA-1 were integrated into the HepG2 genome downstream of the reinserted target-derived fragments.

DISCUSSION

We confirmed genomic on-target alterations in HAP1 and HepG2 cell clones by applying Xdrop-based enrichment for our target region followed by long-read sequencing and *de novo* sequence assembly. Although our genomic region of interest was cleaved by Cas9, we showed that the genomic fragment deriving from the target region was not eliminated from the nucleus. Instead, the cleaved fragment got duplicated, inverted and locally inserted into the genomes of HAP1 and HepG2 cells. In addition, we detected an integration of exogenous DNA fragments in HepG2 cells.

Based on our results, we reasoned that the HAP1 Δ72 deletion clone was heterogenous for the on-target genomic alteration (**Fig. 4B**). HAP1 is a near-haploid cell line but the haploid state is unstable and cells can rapidly become diploid (50). Although the exact chronologic order of events that occurred in this cell clone cannot be reconstructed, the observed genotype can be explained if the HAP1 Δ72 deletion clone derived from a diploid HAP1 cell. We speculate that after transfection of the CRISPR/Cas9 vector, Cas9 and the two gRNAs were efficiently expressed. Cas9 got directed via the dual gRNAs to cut the target-derived fragment from each of the two alleles in the diploid HAP1 cell. In the following cell division, one daughter cell (DC1) used NHEJ to repair the DSBs resulting in a deletion. In contrast, the second daughter cell (DC2) employed MMEJ leading to the ligation and inversion of the two fragments at the original cut site during the repair process (48, 51). We postulate that the orientation of the fragment insertion was stochastic and can occur in either tandem (as in the HAP1 Δ72 deletion clone) or divergent (as in the HepG2 Δ15 deletion clone) direction (**Fig. 4B-C**). This results in a heterogenous cell population, which also explains why, despite the duplication, ChIP-seq signals of Pol III and H3K4me3 over the target region were lower in cells deriving from the HAP1 Δ72 clone when compared to the control clone (**Fig. 1D**). We

hypothesize that if a genomic locus or gene product deriving from a deleted region is advantageous for cell proliferation or survival, selective pressures will favor the propagation of cells containing the inserted target region over time.

Similar to our model for the HAP1 cell clone, we consider that the HepG2 Δ15 deletion clone was heterogenous (**Fig. 4C**). HepG2 has a hyperploid karyotype with trisomy of chromosome 17 (52). Therefore, we speculate that two of the target-derived fragments were inserted into one of the three alleles in one daughter cell. Unlike HAP1, we found that a large part of the CRISPR/Cas9 transfection vector including the gRNA sequence targeting the genomic region of interest was inserted. Interestingly, a fragment originating from the *E. coli* genome was inserted into the HepG2 genome as well. Since we used *E. coli* to produce large amounts of plasmid DNA, it is possible that bacterial over-lysis during plasmid preparation caused a release of *E. coli* genomic DNA, which got fragmented during plasmid purification. Small *E. coli* fragments can still bind to the purification column that selectively captures DNA of smaller size and get co-eluted with the plasmid DNA. We tested for any DNA contamination in our purified plasmid preparation by agarose gel electrophoreses. Although we were unable to visualize any DNA indicative of contamination (**Fig. S2E**), we cannot rule out that very low amounts of fragmented *E. coli* genomic DNA remained undetected.

Latest CRISPR engineering efforts have been focused on preventing potential on- or off-target effects. For example, the observed on-target effects described in this study might be mitigated by using Cas9-directed proteasomal degradation. That would limit the residence time of the Cas9 protein at the cut site and therefore enable accessibility of the DNA repair machinery and faster DSB repair (53, 54).

In conclusion, Xdrop technology combined with *de novo* long-read sequence assembly unexpectedly revealed complex genomic alterations that accompanied CRISPR/Cas9 deletions. These genomic rearrangements cannot be studied by using approaches based on PCR genotyping, Sanger sequencing or standard alignments to the reference genome. By using our approach, we found that Cas9 can induce duplications, inversions and local insertions of target-derived functional fragments as well as integration of exogenous DNA fragments. Although it has been reported that a genomic region of interest can be either duplicated (48) or inverted (51) when using the dual gRNA system, our study provided evidence that a duplication and an inversion, along with an integration of exogenous DNA fragments, can occur at the same time. Furthermore, we also demonstrated for the first time that despite these alterations, the target-derived fragments were still functional and can confound mechanistic interpretations. These findings present a new instance of unintended CRISPR/Cas9 editing events that can be easily overlooked and profoundly affect conclusions drawn from experimental read-outs.

MATERIAL & METHODS

Cell culture

HAP1 cells were obtained from Horizon Discovery and grown in Isveco Modified Dulbecco Medium (Hyclone) supplemented with 10% Fetal Bovine Serum (Hyclone) and 1% penicillin/streptavidin (Sigma). HepG2 cells were obtained from American Type Culture Collection (ATCC, Rockville, MD) and grown in Dulbecco Modified Eagle Medium (Sigma) supplemented with 10% FBS and 1% Penicillin-Streptomycin. Cells were cultured in T75 flasks at 37°C and 5% atmospheric CO_2. HAP1 and HepG2 cells were propagated by splitting 1/10 every two days and 1/4 every three days, respectively. Upon splitting, the medium was aspirated, cells were washed with phosphate buffered saline (Sigma) and detached with 2mL of a trypsin-EDTA solution (Sigma). Trypsin was subsequently inactivated with a minimum of 3-fold surplus of culture medium before a cell fraction was passaged. Both cell lines have a certified genotype.

Plasmid construction

gRNAs were designed and assessed using the design tool from the Zhang's lab (https://zlab.bio/guide-design-resources). Each gRNA was separately cloned into the two BbsI restriction sites of pSpCas9(BB)-2A-Puro (px459) (2).

Transfection and generation of single cell-derived clones

To enhance transfection efficiency, the short-size pBlueScript vector was co-transfected with large-size CRISPR/Cas9 vectors (3). For generating HAP1 and HepG2 deletion clones, two Cas9 vectors (Cas9-gRNA-1 and -2) were used. For creating control clones, no gRNA sequence was cloned into the CRISPR/Cas9 vector. About 160,000 HAP1 cells per well were plated in 6-well plate and on the next day transfected using TurboFectin 8.0 (OriGene) according to the manufacturer's instructions. Roughly 100,000 HepG2 cells were electroporated using the NEON electroporation system (Invitrogen) as previously described (3). After 24h (HAP1) to 48h (HepG2), cells were then selected by adding 2µg/mL puromycin to the cell culture media for two days. Afterwards, cells were allowed to recover in normal medium. Single-cell clones were hand-picked, grown and expanded.

Genomic DNA extraction

HAP1 and HepG2 cells were lysed in 400µL lysis buffer (0.5% SDS, 0.1M NaCl, 0.05M EDTA, 0.01M Tris HCl, 200µg/mL proteinase K) and incubated at 55ºC overnight. After which 200µL of 5M NaCl was added, and the samples were vortexed and incubated on ice for 10min. After centrifugation (15,000×g, 4ºC, 10min), 400µL of the supernatant was transferred to a new tube, mixed with 800µL of 100% EtOH and incubated on ice for at least 10min. Genomic DNA was pellet by centrifugation (18,000×g, 4ºC, 15min), subsequently washed once with 70% EtOH and resuspended in nuclease-free water.

PCR for genotyping

PCR was performed with Taq polymerase PCR (New England Biolabs) according to the manufacturer's manual. Briefly, 100 to 300ng of genomic DNA was used as template DNA. The 25µL PCR reaction consisting of 1× standard Taq reaction buffer, 200µM dNTPs, 0.2µM primers, and 1U Taq DNA polymerase. After thorough mixing of the PCR reaction, the subsequent PCR was performed using an initial denaturation step of 5 min at 95°C followed by 35 cycles of 30sec at 95°C, 30sec at 55 to 60°C, 60sec per 1kb at 68°C and a final extension step of 5min at 68°C in a thermocycler (Applied Biosystems) with preheated lids. All PCR reactions were hold at 4°C. PCR products were assessed by gel electrophoresis on a 1.2% agarose gel.

Primer design

Primers were designed using NCBI Primer-BLAST with default parameters (https://www.ncbi.nlm.nih.gov/tools/primer-blast/index.cgi?LINK_LOC=BlastHome). To ensure primer sequence specificity, we chose "Genomes for selected organisms" and "*Homo Sapiens*". Primers with the highest specificity and GC-content of 40 to 60% were selected.

Sanger sequencing

The PCR band indicative for the homozygous deletion (320 bp) was excised from the agarose gel and purified using Zymoclean Gel DNA Recovery Kit (ZymoResearch) according to the manufacturer's instructions. The purified DNA (75-150ng) and the flanking sequence-specific primers (Supplemental Table 1) were sent for Sanger sequencing (Eurofins, Mix2Seq). Chromatograms are available under Supplementary Files (Figshare).

Chromatin immunoprecipitation followed by sequencing (ChIP-seq)

ChIP-seq experiments were performed as previously described (30). Briefly, HAP1 and HepG2 cells were fixed in 1% formaldehyde, lysed, sonicated and then incubated with Pol III antibodies recognizing antigen POLR3A (18) or H3K4me3 antibodies (05-1339, Millipore). Immunoprecipitated DNA was used to generate sequencing libraries using the Takara SMARTer ThruPLEX DNA-seq Kit according to the manufacture's protocol. Library size distribution and quality were assessed by an Agilent Bioanalyzer instrument using high-sensitivity DNA chips. KAPA-SYBR FAST qPCR kit (Roche) was used to quantify the libraries. The sequencing run was performed with the NextSeq 500/550 High Output v2 kit (Illumina) for 81 cycles, single-end on an Illumina Nextseq500 platform.

ChIP-seq data analysis

Quality of reads were assessed using FastQC (31). Reads were aligned to the human reference genome hg38 using BWA (32). PCR duplicates and reads mapping to ENCODE blacklist (https://sites.google.com/site/anshulkundaje/projects/blacklists) were removed using SAMtools (33) and NGSUtils (34). Bedgraph files were generated using deepTools (35). UCSC (http://genome.ucsc.edu) (36) or IGV (37) was used for visualization of the genomic loci or bam and bedgraph files. To avoid multi-mapping issues, we only considered reads with a mapping quality higher than 30 when inspecting peaks located at the target region.

Xdrop

High-molecular-weight genomic DNA from HAP1 Δ72 and HepG2 Δ15 deletion clone was extracted using Quick-DNA kit (Zymo Research) and shipped to Samplix Services (Denmark) for Xdrop® DNA enrichment as described (28, 38) followed by Oxford Nanopore Technology (ONT) long-read sequencing (LRS). Briefly, the Xdrop target enrichment workflow allows to capture and enrich for a region of interest of up to 100kb by targeting a short detection sequence. For the enrichment of the specific target region, detection primers were designed using the online primer design tool from Samplix to target a 120bp detection sequence approximately mid-way between the Cas9 cut sites. At Samplix services, DNA quality was checked using the Tapestation™ System (Agilent Technologies Inc.), using Genomic DNA ScreenTape according to the manufacturer's instructions. The DNA was further purified using HighPrep™ PCR Clean-up Bead System according to the manufacturer's instructions (MAGBIO Genomics) with the following changes: Bead-to-sample ratios were 1:1 (v:v) and elution was performed by heating the sample in the elution buffer for 3min at 55°C before separation on the magnet. The samples were eluted in 20µL 10 mM Tris-HCl (pH 8). Purified DNA samples were quantified by Quantus (Promega Inc.) Fluorometer™, according to the manufacturer's instructions. PCR reagents, primers, as well as 5ng and 10ng purified DNA of the HAP1 Δ72 and HepG2 Δ15 deletion clones, respectively, were partitioned in droplets and subjected to droplet PCR (dPCR) using the detection primers (see above). The dPCR droplets were then sorted by fluorescence-activated cell sorting (FACS). The isolated droplets were broken, and DNA was again partitioned in droplets and amplified by multiple displacement amplification in droplets (dMDA) reactions. After amplification DNA was isolated and quantified. The MinION sequencing platform from ONT was used to generate LRS data from the dMDA samples as described by the manufacturer's instructions (Premium whole genome amplification protocol (SQK-LSK109) with the Native Barcoding Expansion 1-12 (EXP-NBD104) and 13-24 (EXP-NBD114)). In short, 1.5µg amplified DNA of each sample was T7 Endonuclease I treated followed by size selection, end-repair, barcoding, and adaptor ligation. After library generation, the samples were loaded onto a MinIon flow cell 9.4.1 (20fmol) and run for 16h under standard conditions as recommended by the manufacturer (Oxford Nanopore

Inc.). Generated raw data (FAST5) was subjected to base-calling using Guppy v. 3.4.5 with high accuracy and quality filtering to generate FASTQ sequencing data.

Xdrop data analysis

FASTQ reads were first corrected using Canu (39), followed by SACRA (40) to further identify and split chimeric reads. Reads were aligned to either the human reference genome (hg38) or the region of interest with minimap2 (41). For *de novo* sequence assembly-based approach (**Fig. S2C**), the sequences of mapped reads were extracted from FASTQ files based on the mapped reads' IDs with SAMtools and SeqKit (42). *De novo* sequence assembly was performed with these mapped reads using Canu or Raven (43). The output contigs were compared and assessed by pairwise alignment with Needle (https://www.ebi.ac.uk/Tools/psa/emboss_needle/), MegaBLAST against NCBI standard database of nucleotide collection (https://blast.ncbi.nlm.nih.gov/Blast.cgi) and manual inspection of the alignment quality and coverage of reads to the assembled contigs. If the flanking sequence of the target region needed to be extended, reads mapping to the 5' or 3' end of the contig were used for the second round of a *de novo* sequence assembly. If the read coverage was lower than the requirement of the assemblers, a manual extension was required. To do so, the soft clipped sequences of the reads mapping to the 5' or 3' end of the contig were visualized in IGV, manually compared and summarized. After the contig was successfully assembled, both corrected reads (using Canu and SACRA corrections) and raw reads were aligned to the contig. The contigs were assessed by visualization of aligned reads supporting breakpoints of the contig using IGV. The finalized contigs are available under Supplementary Files (Figshare)

DATA & CODE AVAILABILITY

The datasets used or generated during the current study are available in the ArrayExpress repository, under accession numbers:
Pol III and H3K4me3 ChIPseq data (E-MTAB-4046, E-MTAB-10651)
Xdrop ONT long-read sequencing data (E-MTAB-10652)
Scripts used for bioinformatics analyses are available on GitHub: https://github.com/KeyiG/Cas9_ontarget_alteration.git

ACKNOWLEDGEMENTS

We are grateful to the Samplix Services team, especially Christoffer Rozenfeld, for the fast and thorough delivery of enrichment data and support with data analysis. We would like to thank the laboratories of Marc Friedländer, Vicent Pelechano and Claudia Kutter, as well as Kristoffer Sahlin, Philip Ewels and Remi-Andre Olsen for helpful feedback regarding to the data analysis and John Svetoft for the graphical design of the model. We are grateful for the free usage of the SnapGene plasmid and contig viewer (https://www.snapgene.com/snapgene-viewer/) and free SMART Medical ART images (https://smart.servier.com).

FUNDING

This work was supported by the Samplix Xdrop Grant Program (KG); Chinese Scholarship Council (201700260271 KG, CK); Knut & Alice Wallenberg foundation (KAW 2016.0174, CK); Ruth & Richard Julin foundation (2017–00358, 2018–00328, 2020-00294 CK); SFO-SciLifeLab fellowship (SFO_004, CK); Swedish Research Council (2019-05165, CK); Lillian Sagen & Curt Ericsson research foundation (2021-00427, CK); Gösta Milton's research foundation (2021-00527, CK); Swedish National Infrastructure for Computing (SNIC) at UPPMAX (SNIC 2017/7-261, SNIC 2020/16-223, SNIC 2020/15-292, SNIC 2017/7-154 and uppstore2018110).

AUTHOR CONTRIBUTIONS

CK and KG conceptualized the project. KG, LGM, LW, AM, MS and JNS performed the laboratory experiments. KG did the analysis and visualized the data. RJW contributed reagents. KG and CK acquired funding. KG and CK wrote the original draft. All authors contributed to the review and editing process.

CONFLICT OF INTEREST

The authors declare no conflict of financial and non-financial interests.

REFERENCES

1. Sander,J.D. and Joung,J.K. (2014) CRISPR-Cas systems for editing, regulating and targeting genomes. *Nat. Biotechnol.*, **32**, 347–350.
2. Ran,F.A., Hsu,P.D., Wright,J., Agarwala,V., Scott,D.A. and Zhang,F. (2013) Genome engineering using the CRISPR-Cas9 system. *Nat. Protoc.*, **8**, 2281–2308.
3. Søndergaard,J.N., Geng,K., Sommerauer,C., Atanasoai,I., Yin,X. and Kutter,C. (2020) Successful delivery of large-size CRISPR/Cas9 vectors in hard-to-transfect human cells using small plasmids. *Commun. Biol.*, **3**, 1–6.
4. Greene,E.C. (2016) DNA sequence alignment during homologous recombination. *J. Biol. Chem.*, **291**, 11572–11580.
5. McVey,M. and Lee,S.E. (2008) MMEJ repair of double-strand breaks (director's cut): deleted sequences and alternative endings. *Trends Genet.*, **24**, 529–538.
6. Yang,H., Wang,H., Shivalila,C., Cheng,A., Cell,L.S.- and 2013, undefined One-step generation of mice carrying reporter and conditional alleles by CRISPR/Cas-mediated genome engineering. *Elsevier*.
7. Maddalo,D., Manchado,E., Concepcion,C.P., Bonetti,C., Vidigal,J.A., Han,Y.C., Ogrodowski,P., Crippa,A., Rekhtman,N., Stanchina,E. De, *et al.* (2014) In vivo engineering of oncogenic chromosomal rearrangements with the CRISPR/Cas9 system. *Nature*, **516**, 423–428.
8. Canver,M.C., Bauer,D.E., Dass,A., Yien,Y.Y., Chung,J., Masuda,T., Maeda,T., Paw,B.H. and Orkin,S.H. (2014) Characterization of genomic deletion efficiency mediated by clustered regularly interspaced palindromic repeats (CRISPR)/cas9 nuclease system in mammalian cells. *J. Biol. Chem.*, **289**, 21312–21324.
9. Mansour,M.R., Abraham,B.J., Anders,L., Berezovskaya,A., Gutierrez,A., Durbin,A.D., Etchin,J., Lee,L., Sallan,S.E., Silverman,L.B., *et al.* (2014) An oncogenic super-enhancer formed through somatic mutation of a noncoding intergenic element. *Science (80-.).*, **346**, 1373–1377.
10. Zhang,X., Choi,P.S., Francis,J.M., Imielinski,M., Watanabe,H., Cherniack,A.D. and Meyerson,M. (2016) Identification of focally amplified lineage-specific super-enhancers in human epithelial cancers. *Nat. Genet.*, **48**, 176–182.
11. Guo,Y., Perez,A.A., Hazelett,D.J., Coetzee,G.A., Rhie,S.K. and Farnham,P.J. (2018) CRISPR-mediated deletion of prostate cancer risk-associated CTCF loop anchors identifies repressive chromatin loops. *Genome Biol.*, **19**, 1–17.
12. Guo,Y., Xu,Q., Canzio,D., Shou,J., Li,J., Gorkin,D.U., Jung,I., Wu,H., Zhai,Y., Tang,Y., *et al.* (2015) CRISPR Inversion of CTCF Sites Alters Genome Topology and Enhancer/Promoter Function. *Cell*, **162**, 900–910.
13. Antoniani,C., Meneghini,V., Lattanzi,A., Felix,T., Romano,O., Magrin,E., Weber,L., Pavani,G., Hoss,S. El, Kurita,R., *et al.* (2018) Induction of fetal hemoglobin synthesis by CRISPR/Cas9-mediated editing of the human b-globin locus. *Blood*, **131**, 1960–1973.
14. Zhu,S., Li,W., Liu,J., Chen,C.H., Liao,Q., Xu,P., Xu,H., Xiao,T., Cao,Z., Peng,J., *et al.* (2016) Genome-scale deletion screening of human long non-coding RNAs using a paired-

guide RNA CRISPR-Cas9 library. *Nat. Biotechnol.*, **34**, 1279–1286.
15. Dieci,G., Fiorino,G., Castelnuovo,M., Teichmann,M. and Pagano,A. (2007) The expanding RNA polymerase III transcriptome. *Trends Genet.*, 10.1016/j.tig.2007.09.001.
16. White,R.J. (2011) Transcription by RNA polymerase III: More complex than we thought. *Nat. Rev. Genet.*, **12**, 459–463.
17. Schmitt,B.M., Rudolph,K.L.M., Karagianni,P., Fonseca,N.A., White,R.J., Talianidis,I., Odom,D.T., Marioni,J.C. and Kutter,C. (2014) High-resolution mapping of transcriptional dynamics across tissue development reveals a stable mRNA-tRNA interface. *Genome Res.*, **24**, 1797–1807.
18. Kutter,C., Brown,G.D., Gonçalves,Å., Wilson,M.D., Watt,S., Brazma,A., White,R.J. and Odom,D.T. (2011) Pol III binding in six mammals shows conservation among amino acid isotypes despite divergence among tRNA genes. *Nat. Genet.*, **43**, 948–957.
19. Ottenburghs,J., Geng,K., Suh,A. and Kutter,C. (2021) Genome Size Reduction and Transposon Activity Impact tRNA Gene Diversity While Ensuring Translational Stability in Birds. *Genome Biol. Evol.*, **13**.
20. White,R.J. (2005) RNA polymerases I and III, growth control and cancer. *Nat. Rev. Mol. Cell Biol.*, **6**, 69–78.
21. Yeganeh,M. and Hernandez,N. (2020) RNA polymerase III transcription as a disease factor. *Genes Dev.*, **34**, 865–882.
22. Ishimura,R., Nagy,G., Dotu,I., Zhou,H., Yang,X.L., Schimmel,P., Senju,S., Nishimura,Y., Chuang,J.H. and Ackerman,S.L. (2014) Ribosome stalling induced by mutation of a CNS-specific tRNA causes neurodegeneration. *Science (80-.).*, **345**, 455–459.
23. Gomez-Roman,N., Grandori,C., Eisenman,R.N. and White,R.J. (2003) Direct activation of RNA polymerase III transcription by c-Myc. *Nature*, **421**, 290–294.
24. Gingold,H., Tehler,D., Christoffersen,N.R., Nielsen,M.M., Asmar,F., Kooistra,S.M., Christophersen,N.S., Christensen,L.L., Borre,M., Sørensen,K.D., *et al.* (2014) A Dual Program for Translation Regulation in Cellular Proliferation and Differentiation. *Cell*, **158**, 1281–1292.
25. Goodarzi,H., Nguyen,H.C.B., Zhang,S., Dill,B.D., Molina,H. and Tavazoie,S.F. (2016) Modulated expression of specific tRNAs drives gene expression and cancer progression. *Cell*, **165**, 1416–1427.
26. Blanco,S., Dietmann,S., Flores,J. V, Hussain,S., Kutter,C., Humphreys,P., Lukk,M., Lombard,P., Treps,L., Popis,M., *et al.* (2014) Aberrant methylation of t RNA s links cellular stress to neuro-developmental disorders . *EMBO J.*, **33**, 2020–2039.
27. Raab,J.R., Chiu,J., Zhu,J., Katzman,S., Kurukuti,S., Wade,P.A., Haussler,D. and Kamakaka,R.T. (2012) Human tRNA genes function as chromatin insulators. *EMBO J.*, 10.1038/emboj.2011.406.
28. Madsen,E.B., Höijer,I., Kvist,T., Ameur,A. and Mikkelsen,M.J. (2020) Xdrop: Targeted sequencing of long DNA molecules from low input samples using droplet sorting. *Hum. Mutat.*, 10.1002/humu.24063.
29. Mikheyev,A.S. and Tin,M.M.Y. (2014) A first look at the Oxford Nanopore MinION sequencer. *Mol. Ecol. Resour.*, **14**, 1097–1102.
30. Rudolph,K.L.M., Schmitt,B.M., Villar,D., White,R.J., Marioni,J.C., Kutter,C. and Odom,D.T. (2016) Codon-Driven Translational Efficiency Is Stable across Diverse Mammalian Cell States. *PLoS Genet.*, **12**.
31. FastQC: a quality control tool for high throughput sequence data – ScienceOpen.
32. Li,H. and Durbin,R. (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, **25**, 1754–1760.
33. Li,H., Handsaker,B., Wysoker,A., Fennell,T., Ruan,J., Homer,N., Marth,G., Abecasis,G. and Durbin,R. (2009) The Sequence Alignment/Map format and SAMtools.

Bioinformatics, **25**, 2078–2079.
34. Breese,M.R. and Liu,Y. (2013) NGSUtils: A software suite for analyzing and manipulating next-generation sequencing datasets. *Bioinformatics*, **29**, 494–496.
35. Ramírez,F., Ryan,D.P., Grüning,B., Bhardwaj,V., Kilpert,F., Richter,A.S., Heyne,S., Dündar,F. and Manke,T. (2016) deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res.*, **44**, W160–W165.
36. Kent,W.J., Sugnet,C.W., Furey,T.S., Roskin,K.M., Pringle,T.H., Zahler,A.M. and Haussler, a. D. (2002) The Human Genome Browser at UCSC. *Genome Res.*, **12**, 996–1006.
37. Robinson,J.T., Thorvaldsdóttir,H., Winckler,W., Guttman,M., Lander,E.S., Getz,G. and Mesirov,J.P. (2011) Integrative genomics viewer. *Nat. Biotechnol.*, **29**, 24–26.
38. Blondal,T., Gamba,C., Møller Jagd,L., Su,L., Demirov,D., Guo,S., Johnston,C.M., Riising,E.M., Wu,X., Mikkelsen,M.J., *et al.* (2021) Verification of CRISPR editing and finding transgenic inserts by Xdrop indirect sequence capture followed by short- and long-read sequencing. *Methods*, 10.1016/j.ymeth.2021.02.003.
39. Koren,S., Walenz,B.P., Berlin,K., Miller,J.R., Bergman,N.H. and Phillippy,A.M. (2017) Canu: Scalable and accurate long-read assembly via adaptive κ-mer weighting and repeat separation. *Genome Res.*, **27**, 722–736.
40. Impact of Chimera-less Long Reads on Metagenomics of Human Gut Viromes Treated With Multiple Displacement Amplification (2020) 10.21203/RS.3.RS-58640/V1.
41. Li,H. (2018) Minimap2: Pairwise alignment for nucleotide sequences. *Bioinformatics*, **34**, 3094–3100.
42. Shen,W., Le,S., Li,Y. and Hu,F. (2016) SeqKit: A cross-platform and ultrafast toolkit for FASTA/Q file manipulation. *PLoS One*, **11**, e0163962.
43. Vaser,R. and Šikić,M. (2020) Raven: A de novo genome assembler for long reads. *bioRxiv*, 10.1101/2020.08.07.242461.
44. Mali,P., Yang,L., Esvelt,K.M., Aach,J., Guell,M., DiCarlo,J.E., Norville,J.E. and Church,G.M. (2013) RNA-guided human genome engineering via Cas9. *Science (80-.).*, **339**, 823–826.
45. Sternberg,S.H., Redding,S., Jinek,M., Greene,E.C. and Doudna,J.A. (2014) DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. *Nature*, **507**, 62–67.
46. Bosch-Guiteras,N., Uroda,T., Guillen-Ramirez,H.A., Riedo,R., Gazdhar,A., Esposito,R., Pulido-Quetglas,C., Zimmer,Y., Medová,M. and Johnson,R. (2021) Enhancing CRISPR deletion via pharmacological delay of DNA-PKcs. *genome.cshlp.org*, 10.1101/gr.265736.120.
47. Binda,C.S., Klaver,B., Berkhout,B. and Das,A.T. (2020) CRISPR-Cas9 Dual-gRNA attack causes mutation, excision and inversion of the HIV-1 proviral DNA. *Viruses*, **12**, 330.
48. Shou,J., Li,J., Liu,Y. and Wu,Q. (2018) Precise and Predictable CRISPR Chromosomal Rearrangements Reveal Principles of Cas9-Mediated Nucleotide Insertion. *Mol. Cell*, **71**, 498-509.e4.
49. Kraft,K., Geuer,S., Will,A.J., Chan,W.L., Paliou,C., Borschiwer,M., Harabula,I., Wittler,L., Franke,M., Ibrahim,D.M., *et al.* (2015) Deletions, inversions, duplications: Engineering of structural variants using CRISPR/Cas in mice. *Cell Rep.*, **10**, 833–839.
50. Olbrich,T., Mayor-Ruiz,C., Vega-Sendino,M., Gomez,C., Ortega,S., Ruiz,S. and Fernandez-Capetillo,O. (2017) A p53-dependent response limits the viability of mammalian haploid cells. *Proc. Natl. Acad. Sci. U. S. A.*, **114**, 9367–9372.
51. Li,J., Shou,J., Guo,Y., Tang,Y., Wu,Y., Jia,Z., Zhai,Y., Chen,Z., Xu,Q. and Wu,Q. (2015) Efficient inversions and duplications of mammalian regulatory DNA elements and gene clusters by CRISPR/Cas9. *J. Mol. Cell Biol.*, **7**, 284–298.
52. Zhou,B., Ho,S.S., Greer,S.U., Spies,N., Bell,J.M., Zhang,X., Zhu,X., Arthur,J.G., Byeon,S., Pattni,R., *et al.* (2019) Haplotype-resolved and integrated genome analysis of the cancer

cell line HepG2. *Nucleic Acids Res.*, **47**, 3846–3861.
53. Richardson,C.D., Ray,G.J., DeWitt,M.A., Curie,G.L. and Corn,J.E. (2016) Enhancing homology-directed genome editing by catalytically active and inactive CRISPR-Cas9 using asymmetric donor DNA. *Nat. Biotechnol.*, **34**, 339–344.
54. Brinkman,E.K., Chen,T., de Haas,M., Holland,H.A., Akhtar,W. and van Steensel,B. (2018) Kinetics and Fidelity of the Repair of Cas9-Induced Double-Strand DNA Breaks. *Mol. Cell*, **70**, 801-813.e6.