Universiteit Leiden
The Netherlands

Soft genome editing based on CRISPR nickases: it takes one break to tango

Wang, Q.

Citation
Wang, Q. (2024, June 26). *Soft genome editing based on CRISPR nickases: it takes one break to tango*. Retrieved from https://hdl.handle.net/1887/3765882

Version:	Publisher's Version
License:	Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden
Downloaded from:	https://hdl.handle.net/1887/3765882

Note: To cite this publication please use the final published version (if applicable).

Chapter 2

Expanding the editable genome and CRISPR-Cas9 versatility using DNA cutting-free gene targeting based on *in trans* paired nicking

Xiaoyu Chen[1,†], Francesca Tasca[1,†], Qian Wang[1], Jin Liu[1], Josephine M. Janssen[1], Marcella D. Brescia[1], Milena Bellin[2], Karoly Szuhai[1], Josefin Kenrick[3], Richard L. Frock[3] and Manuel A.F.V. Gonçalves[1]

[1]Department of Cell and Chemical Biology, Leiden University Medical Center, Einthovenweg 20, 2333 ZC Leiden, the Netherlands;
[2]Department of Anatomy and Embryology, Leiden University Medical Center, Einthovenweg 20, 2333 ZC Leiden, the Netherlands;
[3]Department of Radiation Oncology, Division of Radiation and Cancer Biology, Stanford University School of Medicine, 269 Campus Dr. Stanford, CA 94305, USA
[†]These authors contributed equally

Nucleic Acids Research, 2020, Vol.48, No.2, ***974–995***
doi:10.1093/nar/gkz1121

ABSTRACT

Genome editing typically involves recombination between donor nucleic acids and acceptor genomic sequences subjected to double-stranded DNA breaks (DSBs) made by programmable nucleases (e.g. CRISPR–Cas9). Yet, nucleases yield off-target mutations and, most pervasively, unpredictable target allele disruptions. Remarkably, to date, the untoward phenotypic consequences of disrupting allelic and non-allelic (e.g. pseudogene) sequences have received scant scrutiny and, crucially, remain to be addressed. Here, we demonstrate that gene-edited cells can lose fitness as a result of DSBs at allelic and non-allelic target sites and report that simultaneous single-stranded DNA break formation at donor and acceptor DNA by CRISPR–Cas9 nickases (*in trans* paired nicking) mostly overcomes such disruptive genotype-phenotype associations. Moreover, *in trans* paired nicking gene editing can efficiently and precisely add large DNA segments into essential and multiple-copy genomic sites. As shown herein by genotyping assays and high-throughput genome-wide sequencing of DNA translocations, this is achieved while circumventing most allelic and non-allelic mutations and chromosomal rearrangements characteristic of nuclease-dependent procedures. Our work demonstrates that *in trans* paired nicking retains target protein dosages in gene-edited cell populations and expands gene editing to chromosomal tracts previously not possible to modify seamlessly due to their recurrence in the genome or essentiality for cell function.

INTRODUCTION

Genome editing based on homology-dependent and homology-independent DNA repair pathways activated by programmable nucleases permits modifying specific chromosomal sequences in living cells (1). Importantly, these genetic changes can span from single base pairs to whole transgenes (2). However, the genomic double-stranded DNA breaks (DSBs) required for DNA repair activation inevitably yield complex and unpredictable genetic structural variants. These by-products result from the fact that DSBs (targeted or otherwise) are substrates for prevalent non-homologous end joining (NHEJ) pathways and other error-prone recombination processes (3). These processes can trigger local (4) and genome-wide mutations and rearrangements, in the form of insertions and deletions (indels), duplications and/or translocations (5–10). Likewise insidious, targeted DSBs at homologous alleles can result in the assembly of unstable dicentric chromosomes through head-to-head inversional translocations (10). Finally, the engagement of donor DNA with target and off-target DSBs often leads to inaccurate and random chromosomal insertion events, respectively (2,11). This is especially so when donor DNA is presented in target cell nuclei as free-ended double-stranded recombination substrates (11–13).

The unpredictability of genome editing outcomes is naturally aggravated whenever nuclease target sites are located in (i) coding sequences, especially those associated with essentiality and haploinsufficiency, (ii) overlapping *trans*-acting or *cis*-acting sequences and (iii) multiple-copy sequences, such as those in paralogs and pseudogenes. To date, genotypic and phenotypic consequences resulting from editing these three types of genomic regions have received limited examination and remain to be addressed.

Single-stranded DNA breaks (SSBs) made by programmable sequence-specific and strand-specific nucleases (nickases) are intrinsically less disruptive than DSBs as they do not constitute canonical NHEJ substrates (14–17). In this regard, CRISPR–Cas9 nickases consisting of guide RNAs (gRNAs) and Cas9 proteins with either their RuvC or HNH nuclease domains disabled (e.g. $Cas9^{D10A}$ and $Cas9^{H840A}$, respectively), are particularly appealing programmable nicking enzymes (18–20). Indeed, similarly to their cleaving counterparts, CRISPR-Cas9 nickases target DNA consisting of a protospacer adjacent motif (PAM; NGG in *Streptococcus pyogenes* SpCas9) and a sequence complementary to the 5'-terminal 20 nucleotides (nts) of the gRNA (spacer) (18,21). Pairs of CRISPR–Cas9 nickases are commonly used to induce site-specific DSBs through coordinated nicking at opposite target DNA strands. This dual nicking strategy can significantly improve the specificity of DSB formation as SSBs made at off-target sites are, for the most part, faithfully repaired (22,23). However, genome editing based on paired CRISPR–Cas9 nickases remains prone to mutagenesis and chromosomal rearrangements due to the ultimate creation of DSBs (12,22,23).

The non-disruptive character of genome editing based on targeted chromosomal SSBs offers the possibility for seamlessly modifying a broad range of genomic sequences, including those that encode functional protein motifs or essential proteins or that are present in genomic tracts with high similarity

to DNA located elsewhere in the genome. Unfortunately, chromosomal SSBs are, *per se*, poor stimuli for genome editing via precise homology-directed DNA repair (HDR), even in instances in which single base pairs are due to be inserted at a target site (14–17,24).

Here, we sought to determine whether chromosomal regions previously not possible to edit in an efficient and seamless manner could in fact be modified as such. In particular, we hypothesized that *in trans* paired nicking, comprising coordinated SSB formation at donor and acceptor HDR substrates by CRISPR–Cas9 nickases, permits expanding the 'editable genome', i.e. the genomic space amenable to operative DNA editing. Recently, it has been demonstrated that this genetic engineering principle achieves precise HDR-mediated genomic insertions, from a few base pairs (12,25) to whole transgenes (12), without provoking the competing NHEJ pathway. However, the performance of *in trans* paired nicking at coding sequences of endogenous genes, in particular those associated with haploinsufficiency and essentiality, is unknown. To date, equally unknown is the performance of genome editing approaches based on repairing SSBs versus DSBs at these coding sequences using donor plasmids. By targeting exons in the *H2A.X variant histone gene* (*H2AX*) and the *POU class 5 homebox 1 gene* (*POU5F1* or *OCT4*), whose products are essential for the DNA damage response and stem cell pluripotency, respectively, we demonstrate that in contrast to DSB-dependent strategies, in trans paired nicking achieves precise gene editing while disrupting neither functional motifs nor allelic or non-allelic homologous DNA. Moreover, after adapting linear amplification-mediated high-throughput genome-wide translocation sequencing (HTGTS) (10,26) for the detection of SSB-initiated translocations, we found that CRISPR-SpCas9 nickases greatly reduce large-scale chromosomal rearrangements when compared to their nuclease counterparts. Finally, *PARP1* gene targeting experiments showed that, also in instances in which a target gene is not associated with haploinsufficiency or essentiality, in trans paired nicking achieves accurate HDR-mediated gene knock-ins without mutagenizing unmodified alleles, and hence, without reducing target protein dosages.

MATERIALS AND METHODS

Cells

Human cervix carcinoma HeLa cells and human embryonic kidney 293T (HEK293T) cells (both from American Type Culture Collection) were cultured in Dulbecco's modified Eagle's medium (DMEM; ThermoFisher Scientific; Cat. No.: 41966029) supplemented with 5% (v/v) and 10% (v/v), respectively, fetal bovine serum ultra-low endotoxin (FBS; biowest; Cat. No.: S1860500). The HeLa cells, authenticated before by karyotyping analysis (11), were used for gene editing experiments. The HEK293T cells were used for assembling lentiviral vector LV.Cre particles and orthogonal HTGTS analyses. The generation and characterization of the human induced pluripotent stem cells (iPSCs) used in this work (LUMC0020iCTRL) were detailed elsewhere (27). In the current study, these cells were further characterized by COBRA-FISH karyotyping. The iPSCs were cultured in feeder-free Essential 8 Medium (E8; ThermoFisher Scientific; Cat. No.: A1517001) supplemented with 25 U ml^{-1} penicillin and 25 μg ml^{-1} of streptomycin (ThermoFisher Scientific; Cat. No.: 15140122). The iPSCs were kept in wells of six-well plates (Greiner Bio-One; Cat. No.: 662160) coated for 1 h at room with Vitronectin Recombinant Human Protein (VTN-N; ThermoFisher Scientific; Cat. No.: A14700) diluted 1:100 to a final concentration of 5 ng ml^{-1} in Dulbecco's phosphate-buffered saline, no calcium, no magnesium (DPBS; ThermoFisher Scientific; Cat. No.: 14190094). When ready for sub-culturing, to let cell-cell dissociation occur, the iPSCs were first washed with DPBS solution and then incubated with 0.5 mM ethylenediaminetetraacetic acid (EDTA; Invitrogen Cat. No.: 15575020) in DPBS at 37°C and room temperature for 4 and 1 min, respectively. After the removal of the EDTA solution, the cells were seeded in new wells of 24-well plates coated with VTN-N and containing E8 medium supplemented with a 1:200 dilution of RevitaCell (ThermoFisher Scientific; Cat. No.: A2644501). The cells used in this study were mycoplasma free and were kept at 37°C in a humidified-air atmosphere with 5% CO_2 (iPSCs) or 10% CO_2 (HeLa and HEK293T cells).

Recombinant DNA

The expression plasmids AU26_pCAG.Cas9 and AU28_pCAG.Cas9^{D10A} encoding cleaving Cas9 and nicking Cas9^{D10A} enzymes, respectively, have been described previously (12). The control plasmid gRNA_Cloning Vector (Addgene #41824) and the *OCT4*-targeting donor construct eGFP-PGK-Puro (Addgene #31937), herein named pgRNA^{Empty} and pDonor^{OCT4}, respectively, were also described before (20,28). The annotated maps and nucleotide sequences of donor constructs AX74_pDonor$^{OCT4.TS}$,

AX66_pDonor$^{OCT4.1TS}$, AZ44_pDonor^{H2AX}, AZ25_pDonor$^{H2AX.TS}$, AW77_pDonor^{PARP1} and AW69_pDonor$^{PARP1.TS}$ are available in pages 1-14 of the **Supplementary Information**. The annotated maps
and nucleotide sequences of the *S. pyogenes* gRNA-expressing plasmids AZ34_pgRNA$^{H2AX.1}$, AZ35_pgRNA$^{H2AX.2}$, AM70_pgRNA^{PARP1}, AX33_pgRNA$^{OCT4.1}$, AX34_pgRNA$^{OCT4.2}$ are available in pages 15-24 of the **Supplementary Information.** The annotated map and nucleotide sequence of the Cre-expressing lentiviral vector construct BC17_pLV.Cre is available in pages 25-27 of the **Supplementary Information**. The constructs used in the experiments for identifying CRISPR-SaCas9 nucleases inducing HTGTS bait DSBs at RAG1 were BA15_pCAG.SaCas9.rBGpA (29), AV85_pSa-gRAG1.1, AV86_pSa-gRAG1.2, AV87_pSa-gRAG1.3, AP65_pSa-gAAVS1. With the exception of BA15_pCAG.SaCas9.rBGpA (29), all these constructs are described in pages 28–33 of the **Supplementary Information**. The plasmid BPK2660 (Addgene #70709) served as a negative control as it encodes an irrelevant, non-targeting, Staphylococcus aureus gRNA, herein named Sa-gNT (30). Moreover, after BsmBI digestion, BPK2660 also served as an isogenic cloning vector for the insertion of annealed oligonucleotides corresponding to the spacers of S. aureus gRNAs; Sa-gRAG1.1, Sa-gRAG1.2, Sa-gRAG1.3 and Sa-gAAVS1. Plasmids encoding *S. aureus* CRISPR components used for inducing universal HTGTS bait DSBs (i.e., BA15_pCAG.SaCas9.rBGpA and AV85_pSa-gRAG1.1), were combined with constructs AV62_pCAG.Cas9.rBGpA, AB65_pCAG.Cas9^{D10A}.rBGpA and gRNA_AAVS1-T2 (20) expressing *S. pyogenes* CRISPR elements for triggering test HTGTS prey DNA lesions in the form of *AAVS1*-targeted DSBs or SSBs. The latter plasmid (Addgene #41818) encodes an *AAVS1*-targeted gRNA, herein dubbed gAAVS1. The annotated maps and nucleotide sequences of AV62_pCAG.Cas9.rBGpA and AB65_pCAG.Cas9^{D10A}.rBGpA are described in pages 34-39 of the **Supplementary Information**. The full sequences and annotated maps of the plasmids applied in the *AAVS1* gene targeting experiments; AV15_pCAG.Cas9.gRNA^{S1}, AV44_pCAG.Cas9^{D10A}.gRNA^{S1}, AV13_pCAG.Cas9.gRNA^{NT}, AV11_pDonor.EP^{S1} (Addgene #100296) and AV09_pDonor.EP$^{S1.TS}$ (Addgene #100297) are available elsewhere (12).

HeLa and HEK293T cell transfections

HeLa and HEK293T cells were seeded in the tissue culture vessels indicated in **Supplementary Tables S1–S6**. The next day, transfections started by adding a 1 mg ml^{-1} 25 kDa linear polyethyleneimine (PEI, Polysciences) solution (pH 7.4) to each plasmid mixture diluted in 50 μl of 150 mM NaCl (Merck). The cell numbers, the amounts of PEI and DNA (in ng) as well as the compositions of each of the DNA mixtures corresponding to the different transfection reactions are specified in **Supplementary Tables S1–S6**. After the addition of PEI, the transfection reactions were immediately and vigorously vortexed for 10 s, after which, DNA-PEI complexes were allowed to form for 15 min at room temperature. The resulting DNA-PEI complexes were subsequently added directly into the culture media of the target cells and, after 6 h, the transfection media were substituted by regular culture media. Whenever appropriate, reporter-directed flow cytometry was performed at 3 days post-transfection to determine the transfection efficiencies. In the gene targeting experiments, cell populations were then sub-cultured for at least 2 weeks to eliminate episomal donor DNA templates, after which, reporter-directed flow cytometry was used to quantify the frequencies of stably transfected cells.

Human induced pluripotent stem cell (iPSCs) transfections

The iPSCs were first seeded in wells of 24-well plates (Greiner Bio-One) that had been previously coated with VTN-N (ThermoFisher Scientific) as indicated above. The next day, the iPSC culture media were refreshed at least 2 h prior to transfection. Transfections were initiated by adding the appropriate plasmid mixtures together with Lipofectamine Stem Transfection Reagent (ThermoFisher Scientific, Cat. No.: STEM00003) to 50 μl of Opti-MEM medium (Gibco; Cat. No.: 31985-047) in 1.5-ml sterile Eppendorf tubes (**Supplementary Tables S7** and **S8**). After mixing by pipetting, the transfection reactions were incubated at room temperature for 10 min and were then added into the culture media of the target iPSCs (**Supplementary Tables S7** and **S8**). The media were replaced 24 h later and, at 2–3 days post-transfection, the iPSCs were transferred into a new culture well and were subsequently expanded in wells of 6-well plates (Greiner Bio-One) for 5–7 days in the presence of 0.5 μg ml−1 puromycin in StemFlex Medium (ThermoFisher Scientific, Cat. No.: A3349401) containing 25 U ml^{-1} penicillin and 25 μg ml^{-1} of streptomycin. Parallel cultures of mock-transfected iPSCs served as negative controls. At the end of the selection period, puromycin-resistant iPSC colonies were identified by using the leukocyte alkaline phosphatase kit (Sigma-Aldrich; Cat. No.: 86R-1KT) for detecting enzymatic activity from the

pluripotency marker alkaline phosphatase. Cultures of puromycin-resistant iPSC populations and individual randomly selected iPSC colonies were also expanded, collected and cryopreserved for further analyses. The iPSC genomic DNA samples used for orthogonal HTGTS analyses were generated by nucleofecting iPSCs with constructs expressing SaCas9:Sa-gRAG1.1 and SpCas9:gAAVS1 or SaCas9:Sa-gRAG1.1 and SpCas9^{D10A}:gAAVS1. Nucleofection of iPSCs with plasmids expressing only the SaCas9:Sa-gRAG1.1 complexes needed for generating bait DSBs served as an orthogonal HTGTS assay control (**Supplementary Table S9**). The iPSC nucleofections were performed in a Nucleofector 2b-device (Lonza) using Amaxa Human Stem Cell Nucleofector Kit 2 (Lonza; Cat. No.: VPH-5022). A total amount of 8 μg of DNA diluted in 10 μl of Milli-Q water were added to 100 μl of nucleofection buffer containing 2 × 10^6 iPSCs. After gentle mixing, the cell suspensions were transferred to the device-tailored cuvettes and immediately subjected to the nucleofection program B-016, selected for human embryonic stem cells. Next, the iPSCs were transferred to wells of 6-well plates (Greiner Bio-One) containing 2 ml of pre-warmed E8 medium (ThermoFisher Scientific; Cat. No.: A1517001) supplemented with a 1:100 dilution of RevitaCell (ThermoFisher Scientific; Cat. No.: A2644501). After an overnight incubation period, the culture medium was replenished and, at 3 days post-nucleofection, genomic DNA was extracted. Finally, genomic DNA samples were subjected to T7 endonuclease I (T7EI)-based genotyping assays directed at *RAG1* and *AAVS1* alleles and, subsequently, orthogonal HTGTS analyses was performed as described below.

Orthogonal HTGTS sample preparation

Transfections for generating genomic DNA samples for orthogonal HTGTS analyses were carried out in HEK293T cells and iPSCs (**Supplementary Tables S1** and **S9**, respectively). The genomic DNA was isolated at 36 h post-transfection as described before (31). In brief, the cells were collected by centrifugation and resuspended in lysis buffer consisting of 200 mM NaCl, 10 mM Tris–HCl pH 7.4, 2 mM EDTA pH 8.0, 0.2% (w/v), sodium dodecyl sulphate (SDS) and freshly added proteinase K (Thermo Fisher Scientific; Cat. No.: #EO0491) at a final concentration of 200 ng ml^{-1}. After an overnight incubation period at 56°C, the DNA was precipitated by adding isopropanol (1:1) and immediate mixing of the aqueous and organic phases. Next, the DNA was transferred to a new Eppendorf tube containing 1 ml of 70% (v/v) ethanol. The DNA was next pelleted by centrifugation at 13 000 × g for 5 min at 4°C, and dissolved in TE buffer (10 mM Tris–HCl pH 8.0; 1 mM EDTA pH 8.0) for at least 2 h at 56°C. Before orthogonal HTGTS analyses, genomic DNA samples were subjected to T7EI-based genotyping assays. These assays permitted assessing bait and prey chromosomal DNA breaks at *RAG1* and *AAVS1* alleles, respectively, in HEK293T and iPSC cell populations. To this end, the *RAG1* and *AAVS1* target regions were PCR-amplified with KOD Hot Start DNA Polymerase (Merck Millipore; Cat. No.: 71086–3) and GoTaq G2 Flexi DNA Polymerase (Promega; Cat. No.: M7805) using the PCR mixtures indicated in **Supplementary Tables S10** and **S11**, respectively. The PCR primers and cycling parameters used to amplify *RAG1* and *AAVS1* DNA are specified in **Supplementary Tables S12** and **S13**, respectively. Indels generated by NHEJ-mediated DSB repair were detected by exposing *RAG1* and *AAVS1* amplicons to T7EI (Biolabs; Cat. No.: M0302L) as below indicated. Transfections for selecting Sa-gRNAs inducing universal HTGTS bait DSBs at *RAG1* were performed on HeLa cells and HEK293T cells (**Supplementary Table S2**). At 3 days post-transfection, indel formation at the target gene was assessed by T7EI-based genotyping assays as below indicated. To this end, genomic DNA was extracted by using the DNeasy Blood & Tissue kit (Qiagen; Cat. No.: 69506) according to the manufacturer's instructions. Next, the RAG1 target region in HeLa and HEK293T cells was PCR-amplified with KOD Hot Start DNA Polymerase (Merck Millipore). The PCR mixtures, primers and cycling parameters are indicated in **Supplementary Tables S10**, **S12** and **S13**, respectively. The construct expressing *S. aureus* gRNA Sa-gRAG1.1 was selected to induce bait DSBs at *RAG1* in orthogonal HTGTS experiments in HEK293T cells and iPSCs (**Supplementary Tables S1** and **S9**, respectively).

Gene targeting and gene tagging experiments

Transfections for *AAVS1* gene targeting experiments were done in HEK293T cells and iPSCs (**Supplementary Tables S3** and **S8**, respectively) using as donors plasmids AV11_pDonor.EP^{S1} (Addgene #100296) and AV09_pDonor.EP$^{S1.TS}$ (Addgene #100297) (12). The former differs from the latter in that is has its targeting module flanked by gAAVS1 target sites. The targeting modules of these donors consist of sequences homologous to the AAVS1 locus framing expression units encoding both puromycin N-acetyltransferase and EGFP. In these experiments, these donors were combined with plasmids AV15_pCAG.Cas9.gRNA^{S1}, AV44_pCAG.Cas9^{D10A}.gRNA^{S1} and AV13_pCAG.Cas9.gRNA^{NT}

which co-express SpCas9 proteins and gRNAs (12). At 3 days post-transfection, the transfection efficiencies were determined by EGFP-directed flow cytometry. Subsequently, the cells were sub-cultured for 14 days, for the removal of episomal donor templates, after which stable transfection frequencies were established via EGFP-directed flow cytometry. In addition, stably transfected cells present in long-term HEK293T cell cultures were selected for by incubation with 3 μg ml^{-1} of puromycin (InvivoGen; Cat. No.: 58582) during 9 days. The distribution of EGFP expression levels in the resulting puromycin-resistant populations was assessed by EGFP-directed flow cytometry. Transfections for tagging H2AX and PARP1 proteins were performed on HeLa cells (**Supplementary Tables S4** and **S5**, respectively). Transfections of HeLa cells for *OCT4* gene targeting (**Supplementary Table S6**), were assessed by colony-formation assays. To this end, at approximately 2 weeks post-transfection, the cells were counted and seeded at a density of 10^5 cells per 60 mm × 15 mm culture dishes (Greiner Bio-One; Cat. No.: 628160). After a 17-day exposure period to 1 μg ml^{-1} of puromycin (InvivoGen), HeLa cell colonies were identified by Giemsa staining.

Determining genome-wide off-target effects by orthogonal HTGTS analysis

The orthogonal HTGTS analyses were done in a blind fashion on genomic DNA samples isolated from HEK293T cells and iPSCs. Genomic DNA samples from the former and latter cell types were generated as described above using the transfection mixtures specified in **Supplementary Tables S1** and **S9**, respectively. The reagents and procedures for HTGTS analysis have been detailed elsewhere (10,31). In brief, 25 μg of genomic DNA was used for each sample. Samples were sheared using a Bioruptor (Diagenode) with a circulating temperature of 4°C, on a low power setting: 2 × 30 s pulses interspaced by a 60 s cool down period. The biotinylated RAG1A/B – F1 primer (10) was used for LAM-PCR (31), and ssDNA products were enriched on streptavidin-coated magnetic beads (ThermoFisher Scientific; Cat. No.: 65002) prior to ligation of bridge adapters (10,31). Barcoded RAG1A/B – F2 I5 and AP2 I7 primers (10) were used for the nested PCR. P5–I5 and P7–I7 primers (31) were used in the final PCR. The resulting amplicons between 500 bp to 1 kb were separated and gel extracted (Qiagen; Cat. No.: 28706). Phusion polymerase (ThermoFisher Scientific; Cat. No.: F530L) was used in all PCR steps and the blocking enzyme step was omitted. HTGTS libraries were run on a Bioanalyzer (Agilent 2100) prior to MiSeq 2 × 250 bp sequencing (Illumina; Cat. No.: MS-102-2003). Pooled sequence reads were demultiplexed and trimmed according to predetermined molecular barcodes and adapter sequences; each library was subjected to bait/prey alignments (hg19), filtering, and post-pipeline analysis as described (31). Significantly enriched translocation sites in sequence read libraries from individual experiments were identified using MACS2 as previously described (10). Translocation hotspots were called if such enriched translocation sites were statistically significant in the majority of the independent replicate experiments.

Characterization of genome-modifying events by clonal analysis

$EGFP^+$ and $mCherry^+$ HeLa cells generated after PARP1 and H2AX gene editing, respectively, were sorted at 2–3 weeks post-transfection as single cells or as whole populations with the aid of a BD FACSAria III flow cytometer (BD Biosciences). The single cell-derived clones were seeded in wells of 96-well plates (Greiner Bio-One) and were grown in HeLa culture medium supplemented with 50 U ml^{-1} penicillin, 50 μg ml^{-1} of streptomycin and, to increase their cloning efficiency, 50 μM α-thioglycerol and 20 nM bathocuproine disulfonate (both from Sigma-Aldrich) (32). Next, conventional and junction PCR analyses were performed on chromosomal DNA from individual clones, each of which representing a specific genome-modifying event. The PCR screening of the $mCherry^+$ HeLa cell clones was done with the GoTaq G2 Flexi DNA Polymerase system (Promega; Cat. No.: M7808) using the PCR mixtures and cycling parameters indicated in **Supplementary Tables S14** and **S15**, respectively. The screening of the $EGFP^+$ HeLa cell clones was performed with the reagents and protocol provided in the Phire Tissue Direct PCR Master Mix kit (ThermoFisher Scientific, Cat. No.: F170L). The PCR mixtures and cycling parameters used for these analyses are also indicated in the **Supplementary Tables S14** and **S15**, respectively.

Characterization of genome-modifying events in iPSCs by clonal analysis

Puromycin-resistant iPSC colonies derived from *OCT4* targeting experiments using $pDonor^{OCT4}$ and $pDonor^{OCT4.TS}$, were picked from 6-well plates and transferred into wells of 96-well plates by applying a standard 'cut-and-paste' technique. The resulting iPSC clones, each of which representing an individual genome-modifying event, were first cultured in StemFlex Medium (ThermoFisher Scientific) containing

25 U ml^{-1} penicillin and 25 μg ml^{-1} of streptomycin supplemented with Revitacell (ThermoFisher Scientific). Next, the iPSC clones were expanded and adapted to E8 medium (ThermoFisher Scientific) in wells of 24-well plates (Greiner-BioOne). The junction PCR screening for detecting and characterizing genome-modifying events in iPSCs was done on total genomic DNA purified from iPSC clones using the reagents and protocol provided in the Phire Tissue Direct PCR Master Mix kit (ThermoFisher Scientific). The PCR mixtures and cycling parameters applied for these analyses are indicated in the **Supplementary Tables S14** and **S15**, respectively.

Characterization of iPSC clones by COBRA-FISH analysis

Combined binary ratio labelling (COBRA) multicolour FISH-based molecular karyotyping (COBRA-FISH) was carried out on native and gene-edited iPSC lines essentially as detailed elsewhere (33). In brief, glass coverslips containing metaphase spreads air-dried for at least 24 h were incubated with 100 μg ml^{-1} RNase A (Roche; Cat. No.: 10154105103) in 2× saline-sodium citrate (SSC; Sigma-Aldrich; Cat. No.: S0902) at 37 °C for 10 min, followed by incubation with 0.005% pepsin (Sigma-Aldrich; Cat. No.: P0525000) in 0.1 M HCl for 5 min at 37 °C and fixation with 1% formaldehyde (Merck; Cat. No.: 1.03999.1000) in PBS pH 7.4 at room temperature for 10 min. The specimens were dehydrated through a series of incubations in 70–90–100% ethanol solutions, 3 min each, followed by air drying. The probe mix containing the paint mixes covering all chromosomes was dissolved in hybridization mixture, denatured and let hybridize in a moist chamber for 72 h. After hybridization, the glass coverslips were washed in 2× SSC and 0.1% Tween-20 (Promega, Cat. No.: PRH5152), then in 50% formamide (Merck; Cat. No.: 1.09684.1000), 2× SSC pH 7.0 solution at 44°C followed by incubation in 0.1× SSC at 60°C. Each washing step was performed twice for 5 min. The specimens were then dehydrated through a series of incubations in 70–90–100% ethanol solutions, air-dried and embedded in Citifluor AF1/DAPI (400 ng ml^{-1}) solution (Aurion; Cat. No.: E17970). Stained chromosomes were visualised using a Leica DMRA fluorescence microscope (Leica, Wetzlar, Germany) and images were captured with the aid of a CoolSnap HQ2 camera (Photometrics, Tucson, USA). For image processing and karyotyping ColorProc, an in-house developed software tool, was used. A detailed protocol of the whole procedure has been published elsewhere (33).

Reverse transcriptase PCR analysis

Analysis of *H2AX* transcripts in $mCherry^+$ cells subjected to standard, *in trans* paired nicking and paired breaking gene editing, using either $gRNA^{H2AX.1}$ or $gRNA^{H2AX.2}$, was done as follows. Total RNA was extracted with the aid of the NucleoSpin RNA kit (Macherey-Nagel) essentially as specified by the manufacturer after adding 350 μl of RA1 buffer and 3.5 μl of β-mercaptoethanol (Merck). Reverse transcription on 1 μg of total RNA was performed at 50°C for 1 h with 200 ng of random primers, 0.2 mM dNTPs, 1× First-Strand Buffer, 5 mM dithiothreitol, and 200 U of SuperScript III Reverse Transcriptase (all from ThermoFisher Scientific). Next, 1-μl cDNA aliquots were subjected to PCR amplifications with the GoTaq G2 Flexi DNA Polymerase system (Promega; Cat. Nr.: M7808) using 0.4 μM of primer #1444 (5'-CAACGACGAGGAGCTCAACA-3'), 0.4 μM of primer #1508 (5'-GGCGGTGGTGGCCCTTAAAA-3'), 1 mM $MgCl_2$, 0.4 mM dNTPs, 1× GoTaq Flexi buffer and 1.25 U TaqI and H_2O to a final volume of 25 μl. Cycling parameters are specified in **Supplementary Table S16**. To serve as internal controls, 1-μl cDNA aliquots were also subjected to *GAPDH*-directed PCR amplifications with the GoTaq G2 Flexi DNA Polymerase system (Promega; Cat. No.: M7808) using, in this case, 0.4 μM of primer #119 (5'-AGCCACATCGCTCAGACACC-3') and 0.4 μM of primer #120 (5'-GTACTCAGCGCCAGCATCG-3'). Cycling parameters are specified in **Supplementary Table S16**. Finally, 10 μl PCR samples corresponding to *H2AX* and *GAPDH* transcripts were electrophoresed through a 2% (w/v) agarose gel in 1× TAE buffer.

Detection of indels by targeted amplicon sequencing

Target site genotyping of HeLa cell populations containing unmodified cells mixed with cells generated by gene editing involving standard, paired breaking or in trans paired nicking was performed as follows. PCR products spanning $gRNA^{H2AX.1}$ and $gRNA^{H2AX.2}$ target sites were amplified from total cellular DNA extracted from cells at two different timepoints by using the reagents and protocol provided in the DNeasy Blood & Tissue kit (Qiagen; Cat. No.: 69506). The cycling parameters and PCR mixture composition used for amplifying the *H2AX* target region are specified in **Supplementary Tables S16** and **S17**, respectively. *H2AX*-specific PCR products amplified from unmodified HeLa cell populations served as controls. Next, the amplicons corresponding to untagged *H2AX* alleles were extracted

following the QIAEX II Gel Extraction Kit (Qiagen Cat. No.: 20021) and were subjected to Sanger sequencing for determining indel frequencies and distributions with the aid of the ICE software https://ice.synthego.com/#/ (34).

Characterization of *PARP1* alleles in gene-edited cell populations

EGFP$^+$ HeLa cells resulting from *PARP1* gene tagging experiments using *in trans* paired nicking and standard gene editing protocols, were sorted with the aid of a BD FACSAria III flow cytometer (BD Biosciences). Next, total genomic DNA from these EGFP$^+$ populations and from unmodified HeLa cells was extracted by using the DNeasy Blood & Tissue Kit (Qiagen; Cat. No.: 69506), according to the manufacturer's instructions. The various DNA samples were subsequently subjected to PCR amplifications with two different primer pairs (i.e. primer pair A and B). Milli-Q water served as negative controls. The cycling parameters and PCR mixture compositions that were applied are indicated in **Supplementary Tables S16** and **S17**, respectively. Indels at *PARP1* alleles were detected by exposing amplicons to the mismatching-sensing T7EI (Biolabs) as below indicated. The presence of a 121-bp *PARP1* deletion in EGFP$^+$ HeLa cells generated through standard gene editing was established by direct Sanger sequencing of the low-molecular-weight species (241-bp) resulting from PCR with the primer pair B (**Supplementary Table S17**). Finally, the amplicons spanning the SpCas9-induced composite mutations were cloned using the TA cloning protocol (ThermoFisher Scientific Cat. No.: K1214) and were subsequently subjected to Sanger sequencing.

Identification and *in silico* analyses of *H2AX* and *OCT4* gRNAs

The number and distribution of candidate off-target sites for CRISPR complexes was probed by using publicly available algorithms (35,36). The UCSC Genome Browser (Assembly GRCh38/hg38) was used to display all canonical *S. pyogenes* CRISPR-SpCas9 gRNAs in and around the target sequences for tagging H2AX and OCT4. The tracks of the UCSC Genome Browser displayed in the present study are available through the links: https://genome.ucsc.edu/s/mafvg/hg38 Chen Tasca et al C-terminus H2AX CRISPR Zoom, https://genome.ucsc.edu/s/mafvg/hg38 Chen Tasca et al C-terminus CRISPR Zoom and https://genome.ucsc.edu/s/mafvg/hg38_Chen_Tasca_et_al_OCT4_CRISPR_1.5X. The computing of the predicted performance of each CRISPR-SpCas9 complex was made by a combination of algorithms in the crispor.org tool (36). The tracks for chained self-alignments and repeating elements are presented in full mode with the former depicting alignments of the human genome with itself after filtering out the redundant chromosomal positions that map to each other. As specified in the UCSC Genome Browser (Assembly GRCh38/hg38) website, the chained self-alignments and repeating elements tracks were generated with the aid of Blastz (37) and RepeatMasker (http://www.repeatmasker.org/), respectively.

Production and purification of lentiviral vector particles

The vesicular stomatitis virus glycoprotein G (VSV-G)-pseudotyped lentiviral vector LV.Cre was generated according to previously detailed protocols (38,39). In brief, 17 × 10^6 HEK293T cells were seeded per 175-cm^2 culture flask (Greiner Bio-One). The next day, the cells were transfected by adding to 19 ml of regular HEK293T cell culture medium, 1 ml of a 150 mM NaCl solution containing a mixture of 30 µg of DNA composed of lentiviral vector shuttle, packaging, and VSV-G-pseudotyping plasmids at a ratio of 2:1:1 (size-normalized for molecule copy number) and 90 µl of 1 mg ml^{-1} PEI solution (25 kDa PEI, Polysciences). The shuttle, packaging and pseudotyping constructs used were, BC17_pLV.Cre (Supplementary Information), psPAX2 (Addgene #12260) and pLP/VSVG (Invitrogen). The HEK293T cells were incubated overnight in a total 20-ml transfection mixture, after which, this transfection medium was removed and replaced by fresh DMEM supplemented with 5% FBS. At 3 days post-transfection, producer-cell conditioned media containing released vector particles were collected and the cellular debris were removed by centrifugation and filtration using 0.45-µm pore-sized HT Tuffryn membrane filter (Pall Life Sciences; Cat. No. PN4184). The resulting clarified supernatants were gently added onto 5-ml 20% (v/v) sucrose cushions in 35.8-ml polyallomer tubes (Beckman Coulter; Cat. No.: 326823). After ultracentrifugation (15,000 rpm for 2 h at 4°C) in an Optima LE-80K centrifuge (Beckman Coulter) using the SW28 rotor, vector-containing pellets were resuspended in 400 µl of ice-cold PBS pH 7.4 supplemented with 1% (w/v) bovine serum albumin. The vector particle titer of the purified LV.Cre stock was shown to be 31589 ng $p24^{gag}$ ml^{-1} after employing the RETROTEK HIV-1 p24 antigen ELISA kit reagents and protocol (ZeptoMetrix, Cat. No.: 0801111).

Quantification of *OCT4* gene targeting frequencies in iPSCs

Puromycin-resistant iPSCs resulting from *OCT4* gene editing via single nicking, *in trans* paired nicking, standard and paired breaking protocols, were seeded in wells of 24-well plates (Greiner Bio-One) at a density of 30 000 cells per well. The next day, LV.Cre was added to the iPSCs in a total volume of 500 µl of culture medium at a multiplicity-of-infection of 10 vector particles per cell. The frequency of iPSCs expressing OCT4::EGFP assembled via Cre-mediated recombination was determined by flow cytometry at 9 days and 18 days post-transduction.

Characterization of iPSCs with *OCT4* gene-edited alleles

Gene edited iPSCs expressing OCT4::EGFP after coupling *in trans* paired nicking to Cre-mediated recombination, were sorted through a BD FACSAria III flow cytometer (BD Biosciences) as single cell-deposited clones or as polyclonal populations. Both the OCT4::EGFP$^+$ clones and the OCT4::EGFP$^+$ cell populations were deposited in StemFlex Medium (ThermoFisher Scientific; Cat. No.: A3349401) containing 25 U ml^{-1} penicillin and 25 µg ml^{-1} of streptomycin (ThermoFisher Scientific) supplemented with Revitacell (ThermoFisher Scientific). The medium of the iPSC clones was replenished every other day. The medium was refreshed every day when the wells of 96-well plates (Greiner Bio-One) contained visible clusters of viable cells. These cell colonies were further expanded into wells of 48-well plates (Greiner Bio-One) and subsequently into wells of 24-well plates (Greiner Bio-One). Finally, they were expanded and adapted to grow in E8 medium. The OCT4::EGFP$^+$ iPSC clones and iPSC polyclonal populations were subsequently subjected to OCT4/EGFP dual-colour confocal microscopy and flow cytometry assays. Finally, the pluripotency of iPSCs was assessed after applying differentiation protocols and confocal microscopy analyses as detailed under the section 'Differentiation of iPSCs'.

Confocal microscopy analyses

Cells seeded in glass coverslips were fixed in 2% or 4% (v/v) paraformaldehyde (PFA) and were permeabilized in 0.5% (w/v) Triton X-100 in Tris-buffered saline (TBS) pH 7.6 (50 mM Tris–HCl pH 7.6, 150 mM NaCl) at room temperature for 5–10 min (**Supplementary Table S18**). Subsequently, the cells were incubated for 1 h to 2 h with blocking Antibody Diluting Solution (Abdil) consisting of TBS, Triton X-100, 2% (w/v) bovine serum albumin and 0.1% sodium azide. In-between each fixation, permeabilization and blocking steps, the specimens were washed three times for 5 min at room temperature with 0.1% Triton X-100 in TBS. The primary antibodies were diluted in Abdil (**Supplementary Table S18**) and were added to the cells for 1 h at room temperature. After three 5-min washes with 0.1% Triton X-100 in TBS, the cells were incubated with fluorochrome-conjugated secondary antibodies diluted in Abdil for 30 min to 1 h in the dark at room temperature (**Supplementary Table S18**). Next, the specimens were subjected to three 5-min washes with 0.1% Triton X-100 in TBS and were mounted in ProLong Gold Antifade Mounting reagent containing DAPI (ThermoFisher Scientific; Cat. No.: P36931) or in VECTASHIELD Antifade Mounting Medium (VECTOR; Cat. No.: H-1000). Before the addition of the latter mounting medium, the specimens were incubated for 5 min in the dark with the DNA staining reagent DAPI (Invitrogen Cat. No.: R37606) diluted 1:1000 in TBS. Finally, fluorescence microscopy was carried out with an upright Leica SP8 confocal microscope (Leica Microsystems) equipped with Leica hybrid detectors, HyD (Leica Microsystems).

Differentiation of iPSCs

The culturing of clumps of iPSCs on glass coverslips coated with VTN-N triggered the spontaneous differentiation of iPSCs along the three embryonic germ layers. In brief, iPSCs were treated with PBS-EDTA for 1 min at 37°C and were subsequently gently dissociated into large cell clumps by scrapping. The resulting cell clumps were then cultured in suspension for 24 h on low-attachment plates at 37°C. Next, the iPSCs were seeded on coverslips coated with VTN-N in Essential 8 medium (ThermoFisher Scientific, #A1517001) supplemented with Revitacell (ThermoFisher Scientific, Cat. #A2644501). The day after, the medium was changed to DMEM/F12 growth medium (Gibco Cat. #31331–028) containing 20% FBS (Biowest Cat. #S1860–500). The DMEM/F12 medium was replenished every 2–3 days. After 3 weeks under differentiation conditions, the iPSCs were processed for immunofluorescence confocal microscopy for the detection of markers characteristic of the endoderm, mesoderm and ectoderm lineages (**Supplementary Table S19**). The markers corresponding to the three embryonic germ layers that were tested were α-fetoprotein (AFP), forkhead box protein A2 (FOXA2), α-smooth muscle actin (α-SMA), endothelial cell adhesion molecule-1 (CD31), and tubulin β3 class III (TUBB3).

T7 endonuclease I-based genotyping assays

Genotyping assays based on the mismatch-sensing T7EI enzyme were performed for detecting indels at target sequences of CRISPR complexes located at human *PARP1*, *RAG1* and *AAVS1* alleles and at off-target chromosomal positions located in the human *OCT4* pseudogenes *POU5F1P4* and *POU5F1P5*. For the latter assays, the genomic DNA of puromycin-resistant iPSC populations grown after *OCT4*-targeting experiments was extracted by using the DNeasy Blood & Tissue Kit and protocol (Qiagen, Cat. No.: 69506). The GoTaq G2 Flexi DNA Polymerase system (Promega; Cat. No.: M7808) was subsequently applied to amplify the *POU5F1P4* and *POU5F1P5* genomic sequences. The cycling parameters and PCR mixture compositions are specified in **Supplementary Tables S16** and **S17**, respectively. Next, the resulting amplicons were subjected to the thermocycling procedure indicated in **Supplementary Table S20** after which, 10-µl samples were incubated at 37 °C for 17 min with 1.5 µl 10× NEBuffer 2, 0.5 µl (5U) T7EI (New England Biolabs; Cat. No.: M0302) and 3 µl of Milli-Q water. Samples that were not treated with T7EI provided for negative controls. Finally, after agarose gel electrophoresis, untreated and T7EI-treated amplicons were analysed by using the Gel-Doc XR+ system and the ImageLab 4.1 software (both from Bio-Rad).

Flow cytometry

The frequencies of cells expressing H2AX::mCherry, EGFP::PARP1, OCT4::EGFP and EGFP were determined by using a BD LSR II flow cytometer (BD Biosciences). Parental unmodified cells or cells corresponding to experimental negative controls were used to establish the thresholds corresponding to background fluorescence. At least 10 000 viable single cells were analysed per sample. Data were analysed with the aid of FlowJo 10.5.0 software (Tree Star).

Western blotting

After two washes with ice-cold PBS pH 7.4, sorted EGFP::PARP1^{+} and EGFP::PARP1^{-} HeLa cells that had been exposed to standard gene editing or in trans paired nicking procedures were collected from wells of six-well plates by adding 250 µl of lysis RIPA buffer (Pierce Cat. No.: 89900) supplemented with a protease inhibitor cocktail (cOmplete Mini, Sigma-Aldrich Cat. No.: 11836153001). Untreated HeLa cells were taken along as negative controls. The cell lysates were subsequently passed thrice through a 1 ml syringe with a 26 GA 3/8 0.45 × 10 needle (BD Plastipak Cat. No.: 300015) and spun at 14 000 RPM for 5 min at 4°C in an Eppendorf 5424 centrifuge. The protein concentrations in the resulting supernatants were determined by using the BCA Protein Assay Kit (ThermoFisher Scientific Cat. No.: 23225) according to the manufacturer's instructions. Next, 15 µg of protein were diluted in 4× sample buffer and 20× reducing agent (both from Bio-Rad Cat. No. 161-0791 and 161-0792, respectively) and incubated at 95°C for 5 min. Protein samples were loaded in a 7% SDS-PAGE gel. After electrophoreses, the proteins were transferred to a PVDF membrane (Millipore Immobilon Cat. No.: IPVH00010) and were blocked overnight in TBS with 0.05% (v/v) Tween 20 (TBST, ThermoFisher Scientifc Cat. No.: 28358) supplemented with 5% (w/v) Elk milk (Campina). Next, the membrane was incubated with PARP1 polyclonal antibody (Thermo Fisher, Cat. No.: PA5-34803) diluted 1:5,000 in blocking buffer or with α/β tubulin antibody (Cell Signalling Cat. No.: CST 2148) diluted 1:5000 in blocking buffer. After an overnight incubation period at 4°C, the membranes were washed in TBST and incubated for 4 h at 4°C with an anti-rabbit IgG secondary antibody conjugated to horseradish peroxidase (IgG-HRP; Santa Cruz Cat. No.: sc-2004) diluted 1:1,000 in TBST. Proteins were detected by using horseradish peroxidase substrate Pierce ECL2 (Pierce Cat. No.: 80196) following the manufacturer's specifications and Super RX-N X-ray film (Fujifilm).

Statistical analyses

With the exception of genomic DNA samples used for assessing genome-wide off-target effects of CRISPR complexes by orthogonal HTGTS analyses, the researchers were not blinded to sample allocation. Statistical analyses were performed on data sets derived from a minimum of three biological replicates done on different days. These data were analyzed by using the GraphPad Prism 8.0.1 software. The statistical significances were calculated with the tests indicated in the figure legends. P values lower than 0.05 were considered to be statistically significant.

RESULTS

Distinct prevalence of genome-wide rearrangements after Cas9 versus Cas9^{D10A} delivery

Genome-wide off-target effects of programmable nucleases are commonly assessed by high-throughput sequencing of exogenous DNA tags 'trapped' at two-ended DSB termini or, more recently, in situ detection of DSB repair factors (40,41). Although SSBs are mostly resolved through conservative repair processes they can in principle lead to DSBs if a replication fork advances through them and collapses (42). However, the resulting one-ended chromosomal breaks are unlikely substrates for exogenous DNA 'trapping'. Therefore, to fulfil the lack of a sensitive and unbiased genome-wide assay for comparing off-target effects triggered by programmable nucleases versus programmable nickases, we have adapted the HTGTS assay (10). In contrast to other approaches, HTGTS detects off-target effects by deep sequencing of translocations joining bait and prey DSBs made by universal and test nucleases, respectively (**Figure 1A**). In addition to taking place at *bona fide* target sites, prey DSBs can also occur at off-target sites of a specific test nuclease under examination. In adapting the HTGTS assay for comparing off-target effects induced by nucleases versus nickases, we assured that bait DSBs are exclusively made by a universal nuclease whilst prey DSBs are instead generated by either test nucleases or test nickases. To this end, we combined *S. pyogenes* SpCas9 with its ortholog *Staphylococcus aureus* Cas9 (SaCas9). In particular, test *S. pyogenes* and universal *S. aureus* CRISPR complexes were designed for generating prey DNA lesions (i.e. SSBs or DSBs) and universal bait DSBs, respectively (orthogonal HTGTS). After selecting RAG1-targeting SaCas9:Sa-gRAG1.1 complexes as inducers of bait DSBs (**Supplementary Figure S1**), HEK293T cells were exposed to these complexes together with SpCas9:gAAVS1 or SpCas9^{D10A}:gAAVS1, each cleaving or nicking, respectively, at the commonly used *AAVS1* safe-harbour locus (**Figure 1B**). As expected, genotyping assays based on the mismatch-sensing T7EI enzyme, readily revealed indels at *RAG1* and *AAVS1* in cells subjected to SaCas9:Sa-gRAG1.1 and SpCas9:gAAVS1 (**Supplementary Figure S2**). In contrast, indels were detected at *RAG1* but not at *AAVS1* in cells treated with SaCas9:Sa-gRAG1.1 and SpCas9^{D10A}:gAAVS1, confirming that the latter complex displays low mutagenicity at the target intron (**Supplementary Figure S2**) (12). Control orthogonal HTGTS read libraries generated by delivering SaCas9:Sa-gRAG1.1 alone, besides detecting a single poorly-enriched off-target site on chromosome 1, revealed a genome-wide translocation pattern consistent with previously described *S. pyogenes* SpCas9:gRNA bait libraries (**Figure 1C**, **Supplementary Figures S3** and **S4**) (10). Importantly, applying orthogonal HTGTS analyses to experimental DNA samples (**Figure 1C**, **Supplementary Figures S3** and **S4**), demonstrated that amidst cells exposed to SpCas9:gAAVS1 and SpCas9^{D10A}:gAAVS1, the former had significantly higher numbers of off-target translocation hotspots than the latter; i.e. 30.7 ± 6.4 versus 0.7 ± 0.6 recurrent hotspots, respectively (**Figure 1C** and **D** and **Supplementary Figure S4**). In addition, SpCas9:gAAVS1 yielded higher frequencies of translocation junctions per hotspot than SpCas9^{D10A}:gAAVS1 (**Figure 1E** and **Supplementary Figure S4**). It is also noteworthy that, amongst the two translocation hotspots associated with SpCas9^{D10A}:gAAVS1 activity, was that involving *RAG1* bait and *AAVS1* prey target DNA (**Figure 1C**, **Supplementary Figures S3** and **S4**). This data suggests that individual SSBs can indeed be processed into chromosomal DSBs in living mammalian cells.

Together, these data establish orthogonal HTGTS as a sensitive method for the unbiased genome-wide detection of off-target effects elicited by genomic SSBs. Importantly, these results also lend support to SpCas9^{D10A} as a genome-editing tool that diminishes allelic and non-allelic chromosomal mutations and rearrangements.

***In trans* paired nicking minimizes disruptive genotype-phenotype associations**

Earlier *AAVS1* gene targeting experiments in HeLa cells and human induced pluripotent stem cells (iPSCs) demonstrated that DSB-dependent gene editing approaches yield more inaccurate and random donor DNA insertions than *in trans* paired nicking (12). Besides augmenting genotype-phenotype unpredictability (e.g. via insertional mutagenesis), random chromosomal DNA integration results in transgene expression variegation due to chromosomal positional effects (11-12). Similar *AAVS1* gene targeting experiments performed in HEK293T cells support these previous findings (11-12) by showing that heterogeneous transgene expression is prevalent in cell populations subjected to donor plasmids and DSB-forming nucleases (**Supplementary Figure S5**).

Tagging endogenous proteins with fluorescent reporters is a frequent goal of genome editing endeavours, including for establishing live-cell screening systems or studying cellular processes in a spatiotemporal fashion. However, the need for targeting gene termini limits the availability of gRNAs with potentially high activities and/or specificities. The presence of functional motifs further limits gRNA

design as, often, HDR-mediated knock-in of one allele is accompanied by NHEJ-induced knockout of the other allele creating functional gene-dose imbalances. The gRNA availability issue becomes extreme in cases where target sequences (coding or otherwise) are not unique in the genome. These sequences are in fact dubbed 'impossible to target' in the CRISPR tracks of the UCSC Genome Browser and are defined as having at least one identical copy in the genome (43). Thus, as challenging targets for comparing the performance of SpCas9 versus SpCas9^{D10A}, we sought to tag housekeeping *H2AX* and cell type-specific *OCT4* alleles with live-cell reporters. The difficulty in tagging these genes stems from the fact that H2AX function depends on a C-terminal SQ phosphorylation motif (44) that restricts gRNA selection in this coding region and *OCT4* termini share 100% sequence identity with sequences found in four autosomal pseudogenes that prevents the identification of *OCT4*-specific gRNAs.

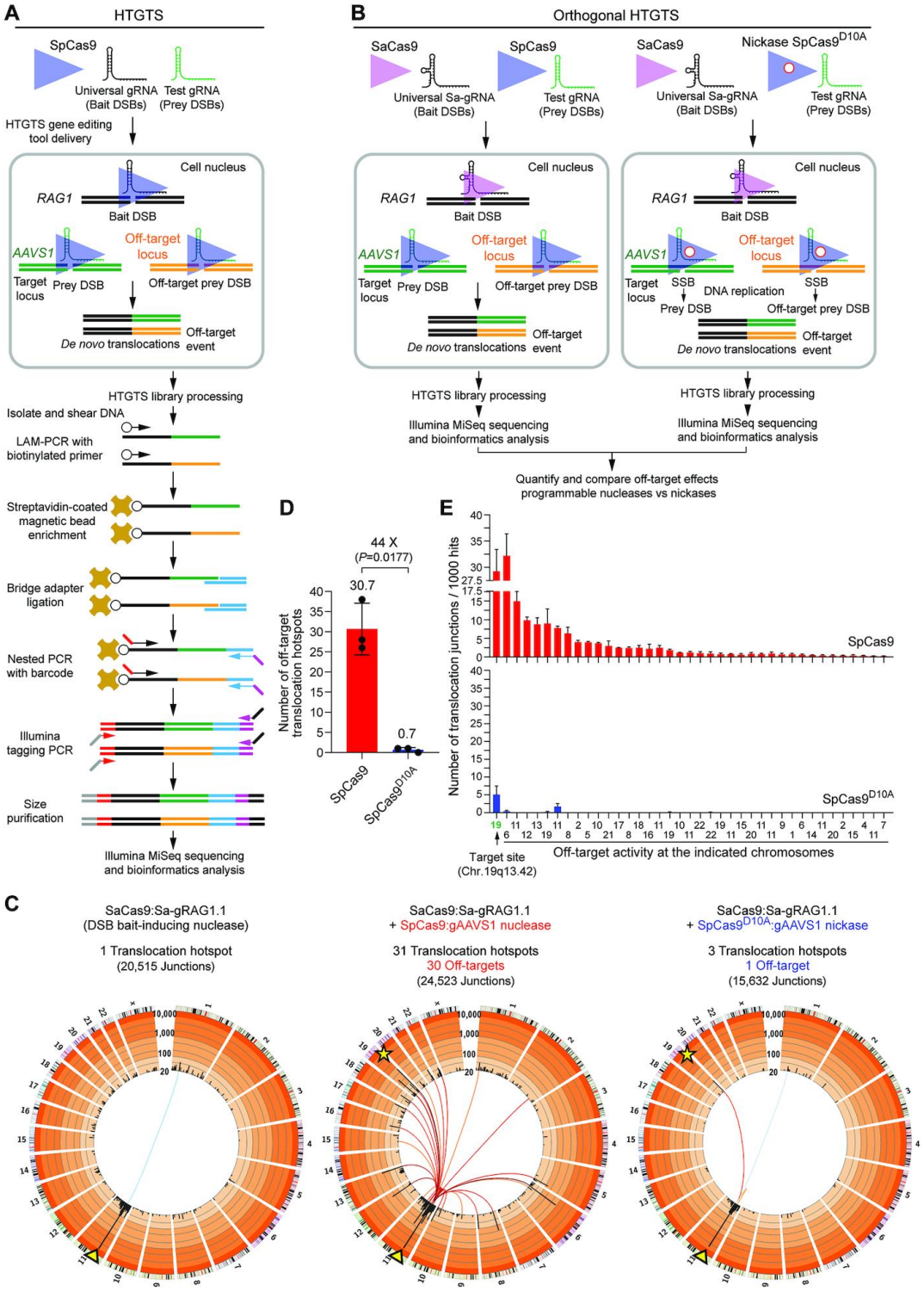

Figure 1. Comparing off-target effects triggered by cleaving versus nicking CRISPR complexes. (**A**) Diagram of the HTGTS pipeline for detecting SpCas9-induced off-target effects. Cells are exposed to *S. pyogenes* CRISPR complexes containing universal and test gRNAs that induce bait and prey DSBs at *RAG1* and target loci, respectively. The prevalence and distribution of off-target hotspots conferred by test gRNAs are determined by an HTGTS pipeline comprising next-generation sequencing of translocations between *RAG1* and off-target DNA (black and orange lines, respectively). (**B**) Diagram of the orthogonal HTGTS pipeline for detecting SpCas9^{D10A}-induced off-target effects. Orthogonal HTGTS assays make use of *S. aureus* and *S. pyogenes* CRISPR complexes for generating bait DSBs at *RAG1* and either prey DSBs or nicks at target loci, respectively. The orthogonality (i.e. lack of cross-talk) between gRNAs and Cas9 proteins from these CRISPR systems avoids nicking at *RAG1* and cleaving at off-target sites of test SpCas9^{D10A}:gRNA complexes (right panel). Further, exchanging SpCas9^{D10A} by SpCas9 in parallel orthogonal HTGTS assays permits comparing side-by-side genomic disruptions inflicted by cleaving versus nicking CRISPR complexes (left panel). Original and orthogonal HTGTS assays share the same downstream library processing and bioinformatics analysis steps. (**C**) Cumulative orthogonal HTGTS analyses (i.e. Circos plots) from three biological replicates. Arrowheads on chromosome 11 indicate the location of the SaCas9:Sa-gRAG1.1 universal bait DSB for all sequence read libraries; stars on chromosome 19 mark the AAVS1 target site of test *S. pyogenes* CRISPR complexes. Blue-graded lines from bait DSBs at the *RAG1* locus indicate bait-related off-targets whereas red-graded lines indicate test gAAVS1-related translocation hotspots from the activity of S. pyogenes CRISPR complexes at target and off-target sites. Hotspots are established only when significantly enriched translocation sites are present in the majority of independent HTGTS replicate experiments (n≥2). Black bars represent 5 Mb bins across each chromosome and enrichment is displayed on a custom color coded log scale by order of magnitude. (**D**) Number of gAAVS1 off-target translocation hotspots in SpCas9 and SpCas9^{D10A} sequence read libraries. Significance was calculated with paired two-tailed Student's *t* tests. (**E**) Relative frequencies of junctions per gAAVS1 translocation hotspot in SpCas9 and SpCas9^{D10A} sequence read libraries. Individual experimental values and respective Circos plots are shown in **Supplementary Figures S3** and **S4**, respectively. Bars and error bars in panels **D** and **E** indicate mean ± S.D., respectively (n = 3 independent biological replicates).

H2AX gene editing experiments were initiated by transfecting HeLa cells with plasmids expressing cleaving SpCas9:gRNA or nicking SpCas9^{D10A}:gRNA complexes containing gRNA$^{H2AX.1}$ or gRNA$^{H2AX.2}$ (**Figure 2A**). The transfection mixtures included donor constructs pDonor^{H2AX} or pDonor$^{H2AX.TS}$. The latter differs from the former in that it has the *H2AX*-specific gRNA target sites flanking the targeting module consisting of 'homology arms' and a mCherry reporter tag (**Figure 2A** and **B**). After delivering these tools, we sought to access the efficiency and precision of gene editing involving (i) DSBs on target DNA (standard), (ii) DSBs on target and donor DNA (paired breaking; DSB^2), (iii) SSBs on target DNA (single nicking) and (iv) SSBs on target and donor DNA (*in trans* paired nicking; Nick^2) (**Figure 2B**). The efficiency and precision of *H2AX* gene editing was ascertained by combining flow cytometric quantification of mCherry$^+$ cells with molecular analysis of randomly isolated mCherry$^+$ clones, each of which, representing an individual genome-modifying event. Importantly, we exploited the fact that the mCherry-tagged intronless *H2AX* gene in donor plasmids behaves as an autonomous reporter unit (**Figure 2C**, top panel) to avoid biased selection of cells harbouring targeted exogenous DNA chromosomal insertions (targeted integrants). The frequencies of transiently and stably transfected cells were determined by flow cytometry before and after episomal templates had been eliminated through sub-culturing (**Figure 2C**, top and bottom panel, respectively). This analysis revealed that, for both gRNAs used, *in trans* paired nicking yielded ~4-fold higher percentages of stably transfected cells than those resulting from the single nicking approach (**Figure 2C**, bottom panel). The robust enhancement on the frequencies of genetically modified cells achieved by *in trans* paired nicking over those resulting from the single nicking strategy is consistent with previous experiments targeting introns (12). Hence, in addition to supporting initial theoretical models postulating nicked DNA partners as homologous recombination substrates (45), these results further stress the limited utility of the single nicking approach. The paired breaking strategy led to the highest frequencies of stably transfected cells (**Figure 2C**, bottom panel). However, it is worth noting that the attendant free-ended donor DNA templates created in cellula by SpCas9-mediated cutting (paired breaking) are prone to yielding complex genome-modifying events, i.e., off-target and inaccurately targeted chromosomal insertions, including concatemeric and HDR-independent integrants (2,11–13). Indeed, although genetically modified cells expressed tagged *H2AX* transcripts independently of the gene editing procedure used (**Supplementary Figure S6**), junction PCR screens of randomly selected mCherry$^+$ clones readily revealed that paired breaking yielded the least precisely targeted integrants when compared to standard and *in trans* paired nicking (**Figure 2D** and **Supplementary Figure S7**). Notably, untagged *H2AX* alleles in mCherry$^+$ clones exposed to SpCas9 and SpCas9^{D10A} had varying and uniform sizes, respectively (**Supplementary Figure S7**). These results support recent findings indicating that, in addition to short indels, SpCas9 can induce gross structural variants at target sequences, such as, large insertions and deletions (4,10). To further characterize these collateral gene-editing events, nucleotide sequencing of *H2AX* alleles was done in mCherry$^+$ clones modified through either standard or *in trans* paired nicking procedures. This target site genotyping analysis confirmed the presence of a range of indel footprints in mCherry$^+$ cells obtained via standard gene editing (**Supplementary Figure S8**). In contrast, untagged

H2AX alleles remained intact in mCherry^{+} cells generated through *in trans* paired nicking (**Supplementary Figure S8**), with the respective tagged *H2AX* alleles expressing the H2AX::mCherry fusion protein in the nuclei of the respective cell populations (**Figure 2E**).

For further assessing the accuracy and mutagenicity of the different gene editing strategies (**Figure 2D** and **Supplementary Figure S7**, respectively), we randomly selected mCherry^{+} clones from cultures initially exposed to the gRNA with the fewest predicted off-target sites, i.e., gRNA$^{H2AX.2}$ (**Supplementary Figure S9**). Interestingly, gRNA$^{H2AX.2}$ directs SpCas9 and SpCas9^{D10A} to cut and nick, respectively, within the codons of the previously mentioned SQ phosphorylation motif whose integrity is crucial for H2AX function (**Figures 2A** and **Supplementary Figure S9**). In this regard, it is worth noting that reduced H2ax dosages in heterozygous *H2ax*$^{+/-}$ knockout mice have uncovered pleiotropic haploinsufficiency phenotypes (46). For instance, embryonic fibroblasts from these *H2ax*$^{+/-}$ mice present growth kinetic curves that are in-between those of wild type and homozygous *H2ax*$^{-/-}$ mice (46). Thus, we next compared the fitness of human cells whose *H2AX* loci had been edited by either *in trans* paired nicking or DSB-dependent gene editing approaches. To this end, populations of mCherry^{+} cells were mixed with a small fraction of unmodified cells (i.e. 5%) and were subsequently monitored by flow cytometry upon serial sub-culturing rounds. Such cell competition settings demonstrated a fitness loss (i.e. growth disadvantage) specifically in cells that had undergone standard and paired breaking gene editing after SpCas9:gRNA$^{H2AX.2}$ delivery (**Figure 2F**). This loss-of-fitness phenotype correlated with the time-dependent disappearance of cells harboring *H2AX* indels disabling the SQ phosphorylation target motif (**Supplementary Figure S10**). We also performed competition experiments in which edited cells had initially been exposed to gRNA$^{H2AX.1}$ instead of gRNA$^{H2AX.2}$. Although displaying a higher potential for off-target effects than gRNA$^{H2AX.2}$, gRNA$^{H2AX.1}$ has a lower change of disrupting the SQ protein motif (**Figure 2A** and **Supplementary Figure S9**). In this case, we observed neither the replacement of edited cells by unedited cells (**Supplementary Figure S11**) nor the elimination of cells with DSB-derived *H2AX* indels (**Supplementary Figure S12**). Thus, in contrast to a process of 'purification' from mutations at the cost of gene-edited cell loss, there was instead, gene-edited cell maintenance at the cost of a 'fixation' of mutations in the populations subjected to SpCas9:gRNA$^{H2AX.1}$ complexes (**Supplementary Figures S11** and **S12**). Importantly, reminiscent of the previous sequencing of *H2AX* alleles in individual clones (**Supplementary Figure S8**), the population-level *H2AX* genotyping assays further confirmed the non-disruptive character of *in trans* paired nicking by revealing the striking dominance of gene edited cells lacking *H2AX* mutations at both time points analysed, independently of the gRNA used (**Supplementary Figures S10** and **S12**, bottom D panels). Taken together, these data indicate that the loss-of-fitness phenotype seen in SpCas9:gRNA$^{H2AX.2}$-treated cells (**Figure 2F**) is attributable to functional H2AX haploinsufficiency caused by NHEJ-mediated disruption of the SQ post-translational modification motif (**Supplementary Figures S8** and **S10**).

***In trans* paired nicking minimizes mutagenesis within coding sequences of target alleles**

PARP1, like H2AX, is also involved in DNA repair, however, functional redundancies with other PARP family members are reported (46,47). Tagging PARP1 with EGFP after delivering conventional pDonor^{PARP1} or target site-containing pDonor$^{PARP1.TS}$, together with cleaving Cas9:gRNA^{PARP1} or nicking Cas9^{D10A}:gRNA^{PARP1} complexes (**Figure 3A**), revealed that *in trans* paired nicking and standard gene editing led to higher frequencies of stably transfected cells than those reached by using the single nicking approach (**Figure 3B**). Importantly, junction PCR screens of randomly isolated EGFP^{+} clones confirmed accurate DNA targeting events in cell populations subjected to *in trans* paired nicking and standard gene editing (**Figure 3C**). Moreover, cell competition experiments involving tracking mixtures of unedited and *PARP1*-edited cells provided no evidence for cell-fitness losses in each of the EGFP::PARP1-expressing populations (**Figure 3D**). Despite this, we sought to characterize EGFP::PARP1^{+} and EGFP::PARP1^{-} cell populations obtained through in trans paired nicking versus standard gene editing (**Figure 4A**). In addition to the typical small indels established after NHEJ-mediated DSB repair, the EGFP::PARP1^{+} cell fraction generated through standard gene editing contained large *PARP1* deletions (**Figures 4B** and **4C**). Of note, small indels were even detected in the EGFP::PARP1^{-} cell fraction isolated from cultures subjected to standard gene editing (**Figure 4C**). Sequence analysis of *PARP1* target DNA in EGFP::PARP1^{+} cells identified a 121-bp deletion mixed with shorter deletions of varying sizes (**Figure 4D** and **Figure 4E**, respectively). These structural variants are reminiscent of those detected in mCherry^{+} cells that had been exposed to cleaving *H2AX*-specific CRISPR-Cas9 complexes (**Supplementary Figures S7** and **S8**), and further support the data indicating

that targeted DSBs can trigger gross genomic alterations (4,10). In contrast, *PARP1* structural variants consisting of large deletions and small indels were detected neither in EGFP::PARP1^{+} nor EGFP::PARP1^{-} cell fractions generated after *in trans* paired nicking (**Figures 4B-D**).

Finally, dual-colour confocal microscopy showed that, regardless of the gene editing methodology, EGFP-tagged PARP1 localized properly in cell nuclei (**Figure 5A**). Tellingly, however, western blot analysis revealed that contrary to EGFP::PARP1^{+} cells resulting from *in trans* paired nicking, EGFP::PARP1^{+} cells derived from standard gene editing suffered a substantial depletion of the endogenous, untagged, PARP1 protein (**Figure 5B**). This data is consistent with the high prevalence of *PARP1* structural variants in EGFP::PARP1^{+} cells initially treated with pDonor^{PARP1} and Cas9:gRNA^{PARP1} (**Figures 4B-E**).

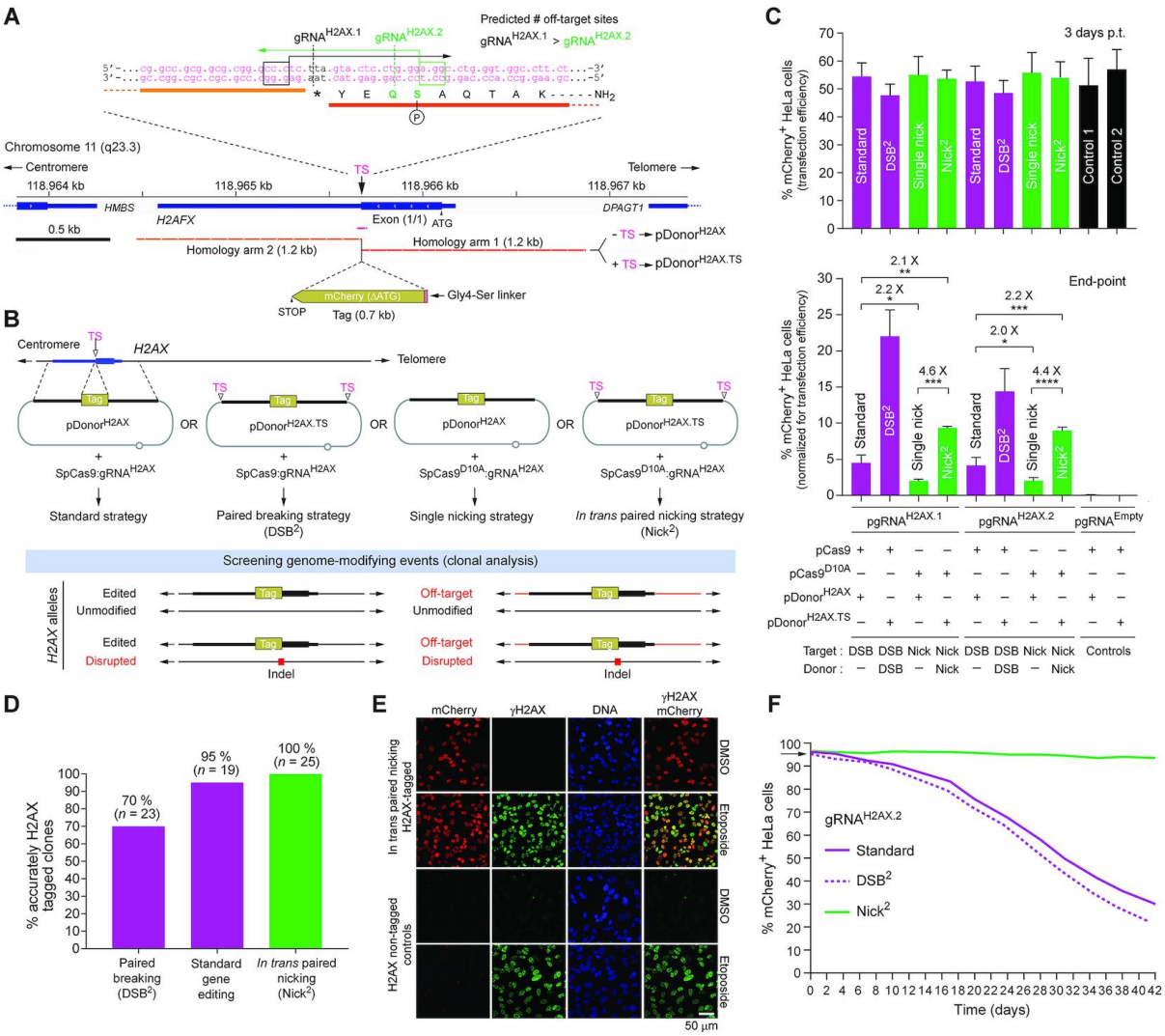

Figure 2. Homology-directed *H2AX* gene editing based on cleaving or nicking CRISPR complexes. (**A**) Diagram of the *H2AX* genomic region. The gRNA$^{H2AX.1}$ and gRNA$^{H2AX.2}$ target sites (TS) are highlighted by horizontal arrows and boxed nucleotides (PAMs). The H2AX post-translationally phosphorylated serine residue 129 is marked with a circled P. The donor plasmids pDonor^{H2AX} and pDonor$^{H2AX.TS}$ contain as targeting module *H2AX* sequences ('arms of homology') flanking a *mCherry* tag. (**B**) Schematics of *H2AX* gene editing strategies. Standard and paired breaking gene editing involve DSB formation at the genomic TS or at this TS and those in the donor DNA, respectively. Single nicking and *in trans* paired nicking gene editing comprise SSB formation at the genomic TS or at this TS and those in the donor DNA, respectively. Wanted and unwanted (red icons) genome-modifying events are depicted. (**C**) Quantification of transiently and stably transfected human cells. Flow cytometry was done on HeLa cell cultures co-transfected with the indicated plasmids. Top and bottom graphs, frequencies of mCherry^{+} cells at early and late time points after transfection (3 days and 2 weeks, respectively). Data are presented as mean ± s.e.m. of four independent biological replicates. Significance between the indicated datasets was calculated with one-way ANOVA followed by Tukey's test for multiple comparisons; *P< 0.05; **P<0.01; ***P< 0.001; ****P< 0.0001. (**D**) Assessing *H2AX* gene editing accuracy. The frequencies of precisely targeted mCherry+ clones were determined through junction PCR screens (**Supplementary Figure S7**). (**E**) Confocal microscopy analysis of *H2AX* gene-edited cells. HeLa cells genetically modified by *in trans* paired nicking were

subjected to direct and indirect fluorescence microscopies for detecting, respectively, mCherry and H2AX phosphorylated at Ser-126 (γH2AX). Prior to microscopy, the cells were incubated with a DNA damaging antitumor agent (etoposide) or with vehicle (DMSO). Nuclei were stained with DAPI. (**F**) Competition experiments comprising unedited and H2AX edited cells. Long-term cultures of cells expressing H2AX::mCherry (95% at t = 0 days) mixed with unedited cells (5% at t = 0 days) were monitored by flow cytometry. H2AX tagging was done through standard, paired breaking (DSB_2), or in trans paired nicking ($Nick_2$) gene editing using $gRNA^{H2AX.2}$.

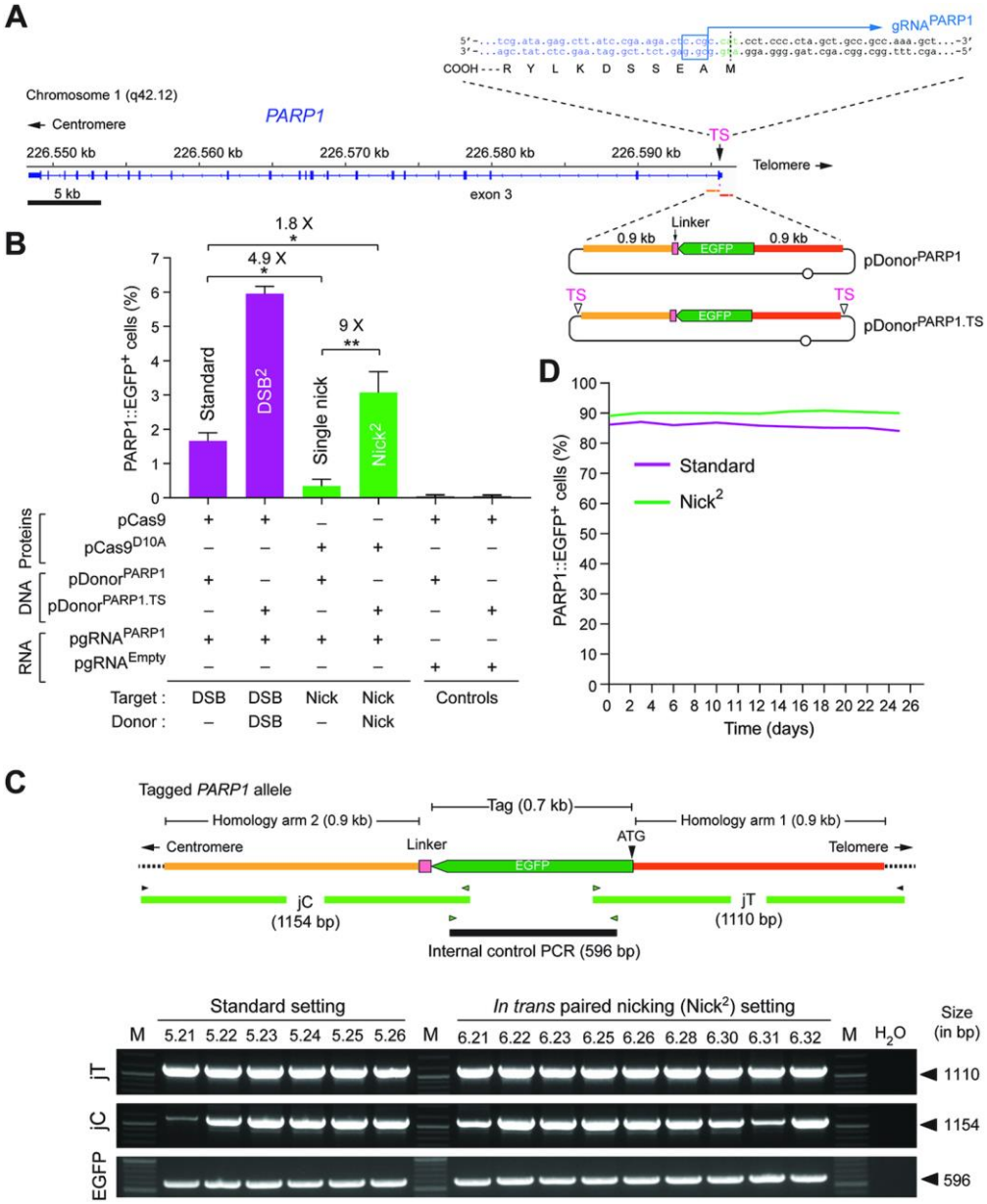

Figure 3. Homology-directed *PARP1* gene editing based on cleaving or nicking CRISPR complexes. (**A**) Diagrams of *PARP1* and *PARP1*-tailored gene editing tools. The gRNAPARP1 target site (TS) is indicated by the horizontal arrow and boxed nucleotides (PAM). The vertical dashed line marks the SpCas9:$gRNA^{PARP1}$ cleaving position. The N-terminal PARP1 amino acids are drawn next to their respective codons. The donor constructs $pDonor^{PARP1}$ and $pDonor^{PARP1.TS}$ have as targeting module *PARP1* sequences ('arms of homology') flanking a *EGFP* tag. The latter construct has, in addition, TS sequences flanking the targeting module. (**B**) Quantification of genetically modified human cells. Flow cytometry of HeLa cell cultures co-transfected with the indicated plasmids. Data are presented as mean ± S.D. of three independent biological replicates. Significance between the indicated datasets was calculated with one-way ANOVA followed by Tukey's test for multiple comparisons; **P*< 0.05; ***P*< 0.01. (C) Molecular characterization of human cells genetically modified through standard versus in trans paired nicking gene editing at *PARP1*. Top panel, Junction PCR assay for assessing *PARP1* gene tagging. Amplicons diagnostic for HDR-derived centromeric and telomeric junctions between exogenous DNA and *PARP1* (jC and jT, respectively) are depicted. Amplicons specific for *EGFP* served as internal controls (EGFP). Bottom panel, Junction PCR analysis on genomic DNA from $EGFP^+$ HeLa cell clones retrieved from cultures co-transfected with pCas9, $pDonor^{PARP1}$ and $pgRNA^{PARP1}$ (Standard setting) or with $pCas9^{D10A}$, $pDonor^{PARP1.TS}$ and $pgRNA^{PARP1}$ (*In trans* paired nicking setting). H_2O, PCR sample containing nuclease-free water instead of genomic DNA. Lanes M, GeneRuler DNA Ladder Mix molecular weight marker. (D) Competition experiment involving unedited and *PARP1* edited cells. Long-term cultures of HeLa cells expressing EGFP-tagged PARP1 mixed with unedited cells were monitored by flow cytometry. Green and magenta lines, $EGFP^+$ cells generated by *in trans* paired nicking and standard gene editing, respectively.

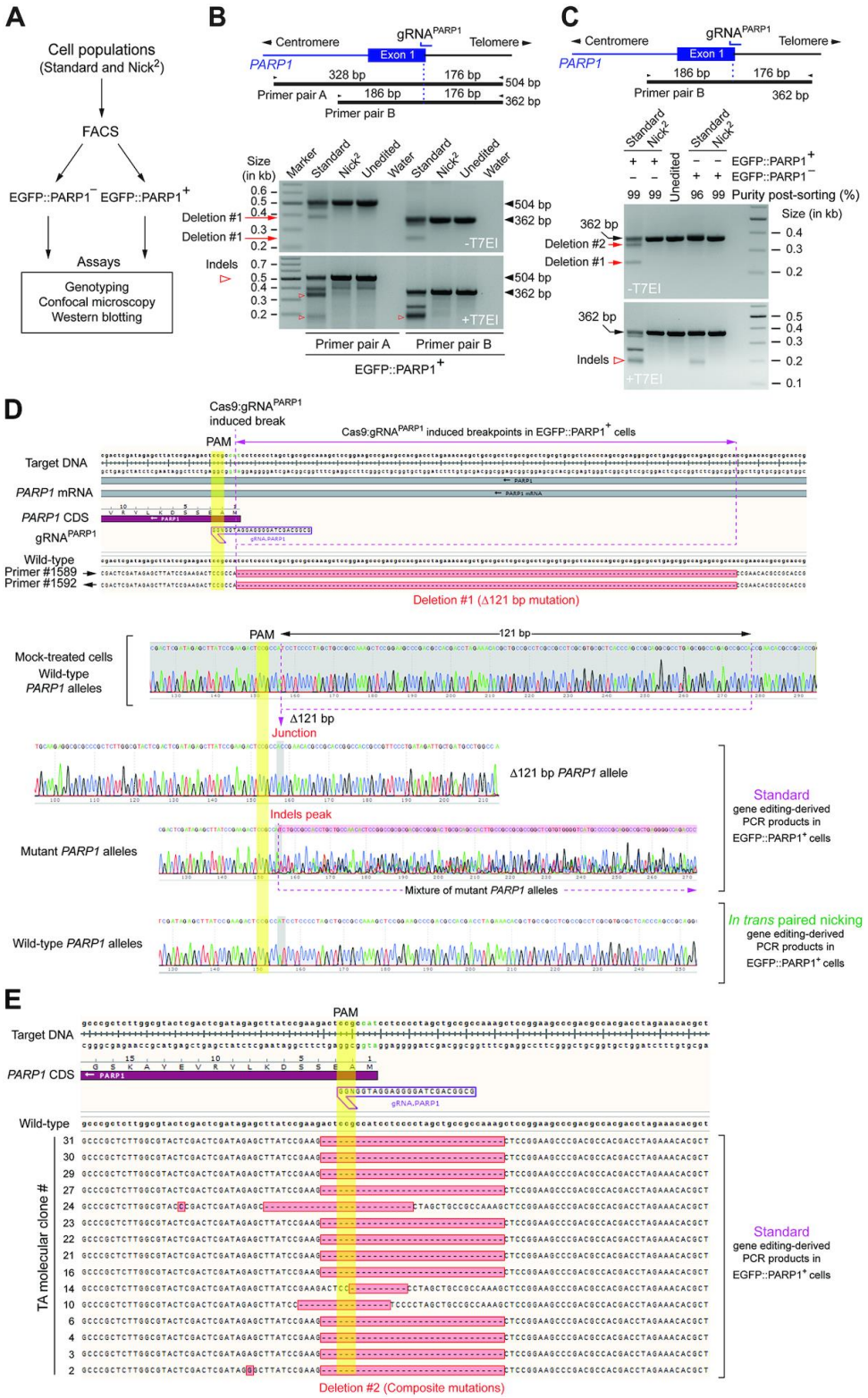

Figure 4. Characterization of *PARP1* alleles in cell populations subjected to standard versus *in trans* paired nicking gene editing. (**A**) Overview of the experimental design. HeLa cell populations subjected to SSB-mediated in trans paired nicking and DSB-mediated standard gene editing were sorted in their respective EGFP::PARP1^{-} and EGFP::PARP1^{+} populations. Each of these cell fractions was next characterized at the DNA and protein levels by the indicated assays. (**B** and **C**) Examination of *PARP1* mutagenesis after gene editing based on DSBs versus SSBs. Untreated and T7EI-treated PCR products spanning the gRNA^{PARP1} target site provided evidence for large deletions and small indels, respectively, in EGFP::PARP1^{+} cells generated by standard gene editing (panel **B**). Indels were equally detected in EGFP::PARP1^{-} cells exposed to standard gene editing (panel **C**). DNA species diagnostic for SpCas9:gRNA^{PARP1}-induced deletions and indels are marked with arrows and open arrowheads, respectively. (**D**) Sequence analysis of the *PARP1* target region in gene edited cells. Top panel, Sanger sequencing of the low molecular weight amplicons shown in panel **B** (-T7EI, primer pair **B**) with forward and reverse primers revealing the presence of a 121-bp deletion at target sequences in EGFP::PARP1^{+} cells that underwent standard gene editing. The *PARP1* proximal deletion

breakpoint coincides with the predicted SpCas9:gRNA^{PARP1} cleaving position. Bottom panel, chromatograms corresponding to *PARP1* alleles in EGFP::PARP1^{+} cells engineered by standard gene editing and in trans paired nicking. Chromatograms corresponding to wild-type *PARP1* and to the 121-bp *PARP1* deletion are also displayed. (**E**) Characterization of additional *PARP1* deletion products. The *PARP1* species with a molecular weight between unmodified and 121 bp-deleted alleles (Deletion #2) presented various mutations as determined by TA cloning and sequence analysis.

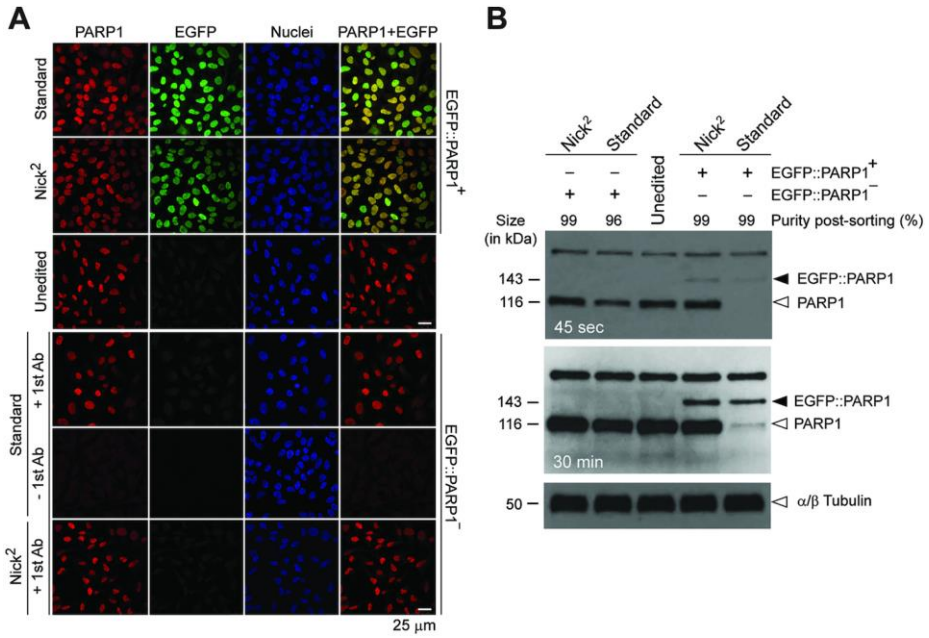

Figure 5. Examination of PARP1 protein status after gene editing triggered by DSBs versus SSBs. (**A**) Confocal microscopy analysis of HeLa cells expressing untagged and EGFP-tagged PARP1. Confocal microscopy of EGFP::PARP1^{+} and EGFP::PARP1^{-} cells confirming co-localization of PARP1 and EGFP in the nuclei of the former cell populations engineered by in trans paired nicking or standard gene editing. Nuclei were counterstained with DAPI. Unedited HeLa cells served as negative controls. Specimens of EGFP::PARP1^{-} cells not incubated with the primary PARP1-specific antibody (-1st Ab) provided for an additional staining control. (**B**) Western blot analysis of HeLa cells expressing untagged and EGFP-tagged PARP1. Western blotting of EGFP::PARP1^{+} and EGFP::PARP1^{-} cells exposing a striking reduction in the amounts of endogenous PARP1 antigen exclusively in EGFP::PARP1^{+} cells generated through standard DSB-dependent gene editing (open arrowhead). Properly sized EGFP::PARP1 fusion products were detected in both EGFP::PARP1^{+} cell populations (solid arrowhead). Unedited HeLa cells served as negative controls. α/β Tubulin antigens served as internal protein loading controls.

In trans paired nicking achieves seamless editing of essential iPSC genomic sequences

The OCT4 transcription factor is essential for human embryogenesis (49) and for the genetic circuitry underpinning pluripotent stem cell states (50,51). For these reasons, it is a coveted gene-editing target. Yet, especially at its termini, *OCT4* shares substantial homology with several of its pseudogenes (**Figure 6A** and **B**). These multiple-copy sequences make the identification of suitable gRNAs hard or impossible (**Figure 6A** and **Supplementary Figure S13**). Hence, we next sought to compare the performance of the different gene editing strategies in a challenging gene-editing model involving tagging *OCT4* at its last exon using gRNAs that lack *OCT4* specificity. To this end, HeLa cells and iPSCs were transfected with conventional pDonor^{OCT4} or target site-modified pDonor$^{OCT4.TS}$, each mixed with plasmids coding for SpCas9:gRNA$^{OCT4.1}$ or SpCas9^{D10A}:gRNA$^{OCT4.1}$ (**Figure 6B**). Colony-formation assays showed that, when compared to single nicking and standard gene editing approaches, in trans paired nicking comprising SSB formation at *OCT4* and donor templates led to higher numbers of puromycin-resistant colonies regardless of the cell type (**Figure 6C**). Similar results were obtained in independent iPSC transfections in which an additional gRNA was included (**Supplementary Figure S14**). Crucially, genomic DNA analysis of randomly isolated iPSC colonies readily revealed that in trans paired nicking achieved a much higher precision in *OCT4* targeting than the DSB-dependent approaches (**Supplementary Figure S15A** and **S15B**). Multicolour FISH-based molecular karyotyping (COBRA-FISH) revealed that neither iPSCs subjected to *in trans* paired nicking nor iPSCs exposed to the DSB-dependent protocols harboured overt chromosomal rearrangements (n = 6; **Figure 7A**). Possibly, this outcome is the result of a strong selection against iPSCs that had initially been exposed to multiple DSBs. Related with this, robust mutagenesis at gRNA$^{OCT4.1}$ target sites located in off-target chromosomal locations (**Figure 7B**) was readily detected in iPSC populations subjected to DSB-dependent gene editing (**Figure 7C**). The

fact that gRNA target sequences in *OCT4* pseudogenes overlap with coding cellular genes, further compounds the genotype of SpCas9:gRNA$^{OCT4.1}$-treated cells (**Figure 7B** and **C**).

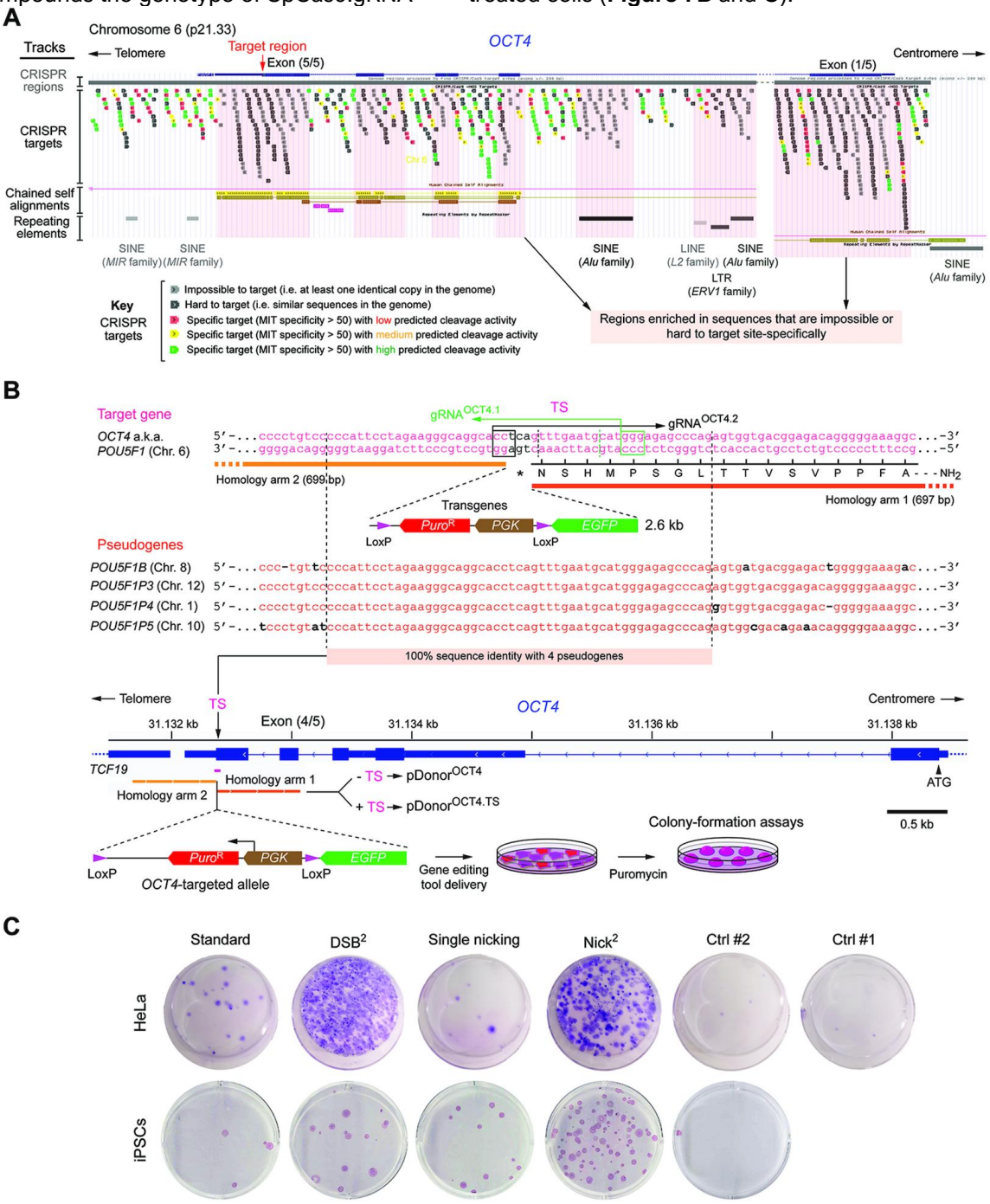

Figure 6. Homology-directed *OCT4* gene editing based on cleaving or nicking CRISPR complexes. (**A**) The *OCT4* genomic region. All potential *S. pyogenes* CRISPR-SpCas9 target sites, as defined by 20-mer spacers and canonical NGG PAMs, are colour-coded according to their predicted target site specificity and activity (CRISPR targets track). Genomic features sharing full or partial sequence identity with *OCT4* are highlighted as duplications and repeats (chained self-alignments and repeating elements tracks, respectively). Tracks annotations were retrieved from the UCSC Genome Browser, Assembly GRCh38/hg38. (**B**) The *OCT4* target region. The *OCT4* terminal nucleotides are drawn in relation to similar sequences present in its pseudogenes and in donor plasmids pDonor^{OCT4} and pDonor$^{OCT4.TS}$. The former and latter constructs lack and contain, respectively, gRNA target sites (TS) flanking the targeting module. The target sites are indicated by horizontal arrows and boxed nucleotides (PAMs). Donor constructs are built to knock-in a floxed positive-selection cassette plus an *EGFP* reporter into *OCT4* loci. The Cre-mediated excision of the selection cassette generates a traceable OCT4::EGFP fusion product exclusively in accurately targeted iPSCs. (**C**) *OCT4* gene editing. Colony-formation assays for detecting stably transfected cells. iPSCs and HeLa cells were co-transfected with conventional pDonor^{OCT4} or target site-modified pDonor$^{OCT4.TS}$ templates each mixed with constructs expressing SpCas9:gRNA$^{OCT4.1}$ or SpCas9^{D10A}:gRNA$^{OCT4.1}$. After puromycin selection, alkaline phosphatase and Giemsa staining identified genetically modified colonies of iPSCs and HeLa cells, respectively.

The generation of DSBs at *OCT4* pseudogenes (**Figure 7C**) raises the possibility for the insertion of *OCT4*-targeting donor DNA at these off-target genomic positions due to the partial homology between them and donor DNA (**Supplementary Figure S15C**). A junction PCR assay devised to investigate this possibility did not detect donor DNA insertions at *OCT4* pseudogenes in puromycin-resistant iPSC clones (n = 22) randomly isolated from cultures subjected to *in trans* paired nicking (**Supplementary Figure S15C** and **D**).

Previous experiments in pluripotent stem cells (i.e. human embryonic stem cells and iPSCs) revealed that *in trans* paired nicking using SpCas9^{D10A}:gAAVS1 complexes yields higher gene targeting frequencies than those achieved by standard gene-editing involving SpCas9:gAAVS1 (12). Similar *AAVS1* gene targeting experiments performed in the iPSC line used in the current study were consistent with these earlier findings (**Supplementary Figure S16**). To investigate whether chromosomal rearrangements were detectable in these iPSCs soon after their exposure to CRISPR complexes, we performed orthogonal HTGTS analysis on cell populations exposed to SaCas9:Sa–gRAG1.1 alone or together with SpCas9:gAAVS1 or SpCas9^{D10A}:gAAVS1 complexes (**Supplementary Figure S17**). The orthogonal HTGTS assay detected translocations exclusively in iPSCs that had been co-treated with SaCas9:Sa-gRAG1.1 and SpCas9:gAAVS1 nucleases (**Supplementary Figures S18** and **S19**). When compared with the orthogonal HTGTS experiments performed in aneuploid HEK293T cells (**Figure 1C**, **Supplementary Figures S3** and **S4**), the overall lower frequencies of translocations detected in iPSCs might have resulted from their diploid character and/or lower exposure to CRISPR complexes (compare **Supplementary Figures S2** with **Supplementary Figure S17**). Crucially, in line with the orthogonal HTGTS experiments in HEK293T cells, this data support that SpCas9^{D10A} nickases trigger less chromosomal rearrangements than their SpCas9 counterparts, in this case, in diploid iPSCs (**Supplementary Figures S18** and **S19**).

To complement the characterization of gene-edited iPSCs (**Figure 7** and **Supplementary Figure S15**), we set-up a quantitative specificity assay in which Cre-mediated OCT4::EGFP assembly reports on precise gene editing in iPSCs (**Figure 8A**). The results from this functional genetic assay confirmed the strikingly different *OCT4* targeting levels achieved by nicking versus cleaving CRISPR complexes. In particular, in contrast to the single nick-dependent and DSB-dependent approaches, induction of SSBs at acceptor and donor DNA results in efficient targeted gene editing in viable iPSCs (**Figure 8B**). Our results suggest that exposing iPSCs to nicking as opposed to cleaving CRISPR complexes overcomes a strong negative selection against *OCT4*-edited iPSCs. These results are in agreement with previous experiments showing that even very few DSBs, including those made by SpCas9 nucleases, can significantly reduce the division and survival rates of PSCs (12,52–54). Finally, dual-colour confocal microscopy and flow cytometry analyses confirmed proper EGFP tagging of the endogenous OCT4 protein in iPSCs subjected to in trans paired nicking, at both the population and clonal levels (**Figure 8C** and **D**, respectively). Importantly, these OCT4::EGFP-expressing iPSCs were equally capable of differentiating along the three embryonic germ layers (**Figure 8E** and **Supplementary Figure S20**).

In conclusion, unwarranted genotypes and deleterious phenotypic traits created by CRISPR-SpCas9 nucleases during gene knock-in procedures are mostly avoided by using *in trans* paired nicking genome editing.

DISCUSSION

There are some concerns regarding the application of genome editing technologies. This is especially so when these applications are directed towards biotechnologies and genetic therapies (55). In part these concerns stem from the fact that, regardless of their specificity, programmable nucleases generate DSBs that are prone to large-scale and small-scale mutagenic events (4–10). In this regard, programmable nuclease-induced DSBs are particularly problematic, hence avoided, at multiple-copy sequences and/or at sequences needed for proper cell functioning or overall viability. As corollary, DSB-dependent genome editing substantially limits the editable genome. Moreover, in mammalian diploid cells, nuclease-induced homologous chromosome rearrangements (10) and allelic mutations potentiate cell transformation events and gene-dose unbalances, respectively. Equally insidious are the recent findings that DSB-induced nonsense mutations can trigger transcriptional compensatory mechanisms that further confound genotype-phenotype associations (56–58).

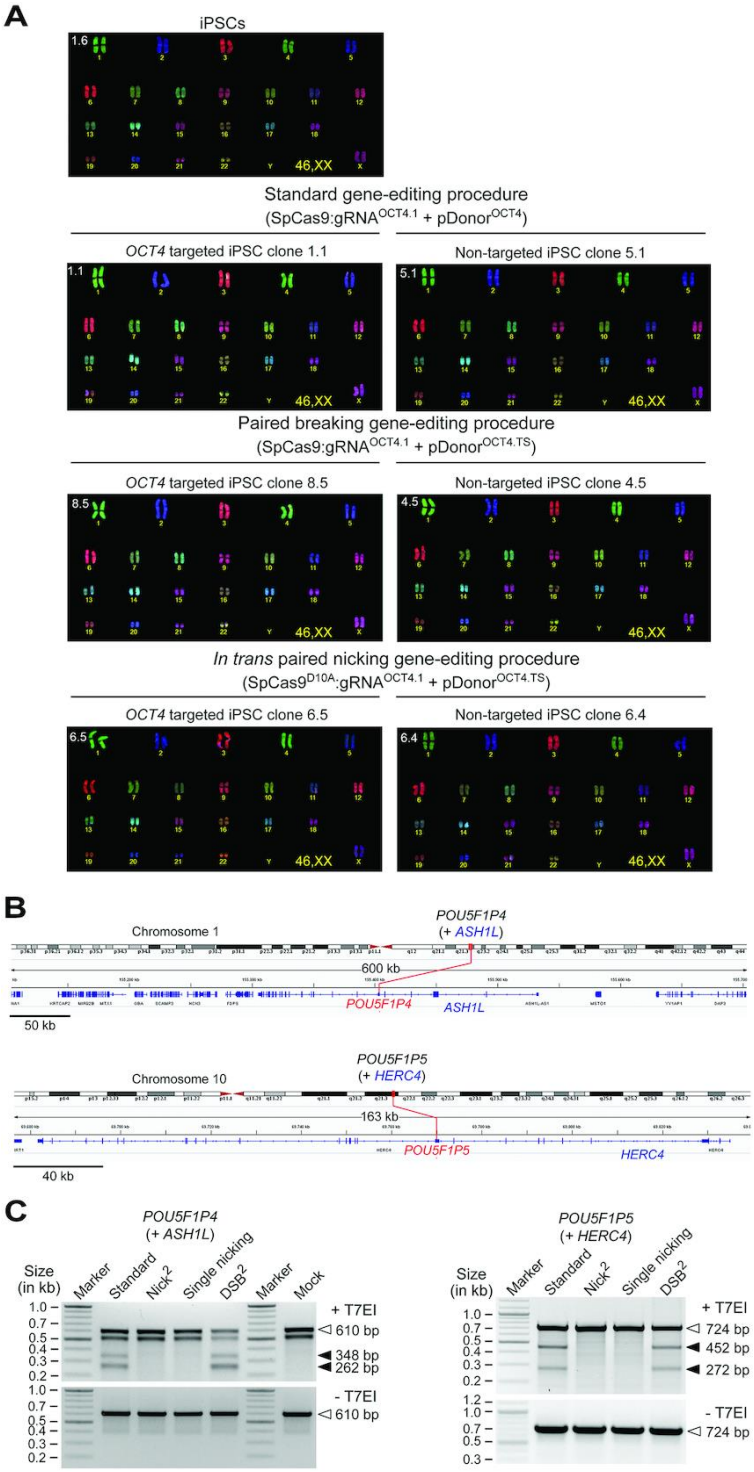

Figure 7. Characterization of iPSCs after *OCT4* gene editing using cleaving versus nicking CRISPR complexes. (**A**) Karyotyping of genetically modified iPSC clones. Overview of COBRA-FISH analysis of parental iPSCs and individual targeted and non-targeted clones showing a seemingly normal diploid karyotype (46,XX). Each clone was isolated after adding puromycin to iPSC populations subjected to the indicated gene editing strategies. (**B**) Chromosomal and genomic coordinates of *POU5F1P4* and *POU5F1P5*. The former and latter *OCT4* pseudogenes overlap with nucleotide sequences from *ASH1L* (ASH1-like histone lysine methyltransferase) and *HERC4* (HECT and RLD domain containing E3 ubiquitin protein ligase 4), respectively. *ASH1L* codes for a member of the trithorax group of transcriptional activators and is ubiquitously expressed in over 25 tissues; *HERC4* belongs to the HERC family of ubiquitin ligases and is ubiquitously expressed in over 25 tissues. As a result, indels generated at *OCT4* pseudogenes inevitably create additional genotypic complexity in target cell populations whose, cell type-specific, phenotypic consequences are difficult to predict and assess. (**C**) Comparing genome-disrupting events at *OCT4* gRNA target sites located at off-target chromosomal positions. T7EI-based genotyping assays were performed on DNA from puromycin-resistant iPSC populations expanded after *OCT4*-targeting experiments involving the indicated gene editing procedures. T7EI-specific products diagnostic for mutant alleles generated by NHEJ-mediated DSB repair are pinpointed by closed arrowheads;

products corresponding to intact alleles are instead indicated by open arrowheads in untreated and T7EI-treated samples. Marker, GeneRuler DNA Ladder Mix molecular weight marker.

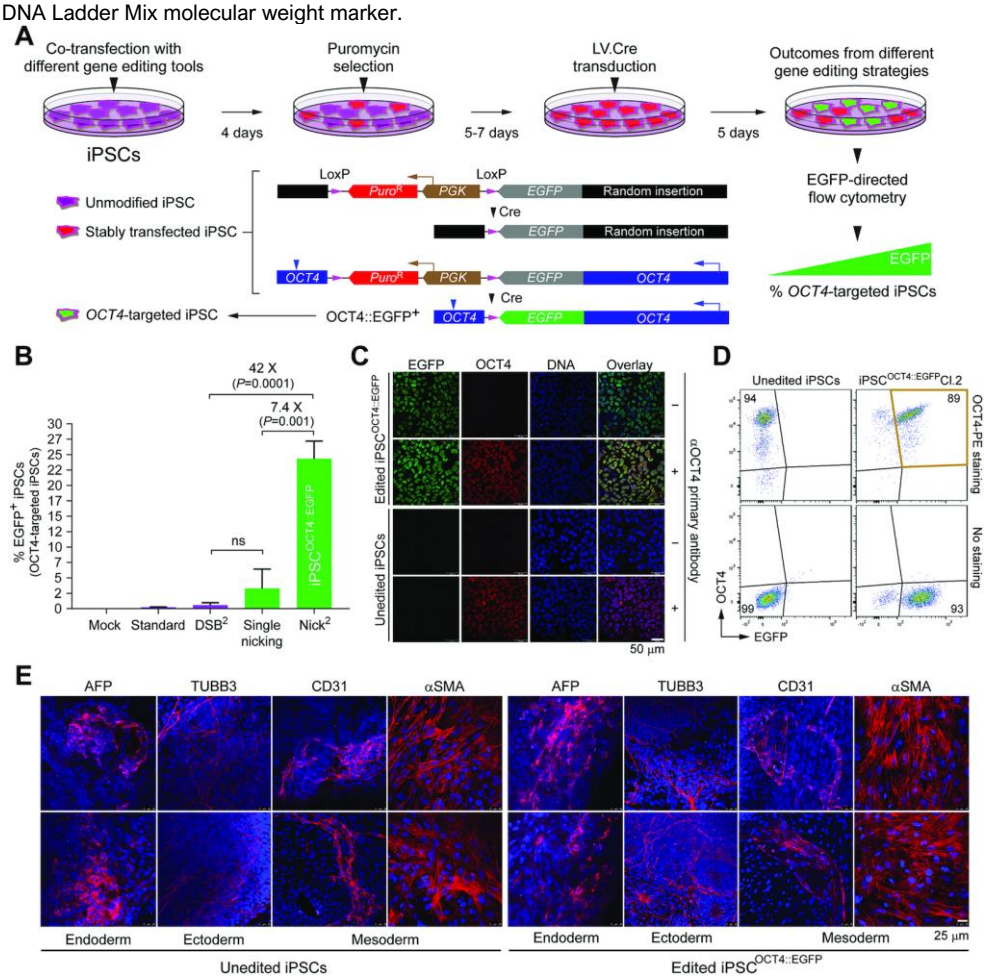

Figure 8. Comparing the accuracy of OCT4 gene editing after delivering cleaving versus nicking CRISPR complexes. (**A**) Genetic assay for determining *OCT4* targeting frequencies. iPSCs co-transfected with plasmid combinations corresponding to each of the four different gene editing strategies, were sequentially exposed to puromycin and Cre recombinase. *OCT4*-targeted iPSCs expressing Cre-derived OCT4::EGFP fusion products report accurate genome-modifying events. The Cre recombinase was delivered by transducing iPSCs with lentiviral vector LV.Cre at a multiplicity-of-infection of 10 physical particles per cell. (**B**) Comparing the performance of *OCT4* gene editing strategies in iPSCs. The frequencies of *OCT4*-targeted iPSCs expressing OCT4::EGFP were determined by EGFP-directed flow cytometry. Data are shown as mean ± S.D. of independent biological replicates. Significance was calculated with two-tailed Student's *t* tests (n = 3); ns, non-significant. (**C**) Confocal microscopy analysis of *OCT4* edited iPSCs. OCT4::EGFP-expressing iPSCs engineered through in trans paired nicking and Cre delivery (iPSC$^{OCT4::EGFP}$) were subjected to indirect and direct fluorescence microscopies for detecting OCT4 and EGFP, respectively. Nuclei were stained with DAPI. Nuclear localization of OCT4::EGFP is highlighted by the merging of the three fluorescence signals. Unedited iPSCs served as negative controls. iPSC and iPSC$^{OCT4::EGFP}$ populations that were not incubated with the OCT4-specific primary antibody served as staining controls. (**D**) Flow cytometric analysis of *OCT4* edited iPSCs. Flow cytometry of iPSC clone 2 isolated from an iPSC$^{OCT4::EGFP}$ population confirming OCT4 and EGFP co-labelling (coloured quadrant). Unedited iPSCs served as controls. Cultures of parental iPSCs and iPSC$^{OCT4::EGFP}$ clone 2 that were not exposed to the PE-conjugated OCT4 antibody were used as staining controls. (**E**) Testing multi-lineage differentiation capacity of iPSC populations expressing OCT4::EGFP. Immunofluorescence microscopy analysis of iPSC$^{OCT4::EGFP}$ cells differentiated into cellular lineages representative of endoderm, ectoderm and mesoderm. Unedited iPSCs served as differentiation controls. Markers for each germ layer are indicated. Nuclei were stained with DAPI.

We report that concomitant SSB formation at target and donor DNA by CRISPR-SpCas9 nickases elicits accurate and non-disruptive gene editing, including at loci associated with haploinsufficiency and essentiality. This DSB-free *in trans* paired nicking approach prevented the loss of gene-edited cells due to the disruption of a functional protein motif or a pluripotency supporting gene in iPSCs. The observed difficulty in isolating iPSCs edited at *OCT4* after CRISPR-SpCas9 delivery is in line with the essentiality of this gene in safeguarding stem cell phenotypes (49–51) and with earlier experiments showing that gene targeting frequencies at *OCT4* are very low. Indeed, gene editing of iPSCs using TALENs and the herein used pDonor^{OCT4} construct, did not yield any correctly targeted clone (0/48) (28). In another study, gene editing of human embryonic stem cells deploying SpCas9 and donor templates containing the same 'homology arms' of pDonor^{OCT4}, resulted in only 8 correctly targeted clones (8/288) (59). In

contrast to these studies, viable and correctly targeted iPSC clones were readily isolated after targeting *OCT4* with pDonor$^{OCT4.TS}$ and SpCas9^{D10A} (21/22) (**Supplementary Figure S15B**). Importantly, in trans paired nicking gene editing introduces a low mutagenic load into target cell populations by minimizing NHEJ-mediated chromosomal disruption of allelic and non-allelic target sequences, such as those in *OCT4* and its pseudogenes, respectively. These multiple-copy gRNA target sites, are likely to have exacerbated the difficulty in isolating *OCT4*-targeted iPSCs after SpCas9 delivery (**Figure 8B** and **Supplementary Figure S15B**) as pluripotent stem cells are particularly prone to DSB-induced cell cycle arrest and apoptosis (12,52–54). There are other experimental data linking detrimental genome editing outcomes to target sequences associated with copy number variations. In particular, genome-wide CRISPR-SpCas9 library screens have demonstrated that DSBs mapping in amplified genomic regions create false-positive hits of gene essentiality in cancer cell lines (60,61).

Notwithstanding the fact that nicking CRISPR complexes are significantly less mutagenic than their cleaving counterparts at both target and off-target sites, they can nonetheless trigger DNA disruptions if, for example, an advancing replication fork collapses after hitting the SSB product (42). In the present work, by using orthogonal HTGTS assays, we have provided experimental evidence for such events in mammalian cells (**Figure 1C** and **Supplementary Figures S3** and **S4**). These events should be most problematic at off-target sites. In this regard, it will be worth investigating whether *in trans* paired nicking is amenable to RNA-guided nickases built on high-specificity Cas9 scaffolds (62).

Although the *OCT4* edited iPSC clones analysed lacked donor DNA insertions at SSB-susceptible *OCT4* pseudogenes (**Supplementary Figure S15D**), unwanted knock-ins at genomic regions exhibiting high homology with donor DNA constitute a possible limitation of *in trans* paired nicking. Therefore, whenever possible, this risk should be minimized by avoiding SSB formation at such potential off-target regions and/or reducing the extent of homology between them and donor DNA (63). Conversely, assuring SSB formation at donor DNA and multiple-copy homologous sequences might offer the prospect for co-editing these recurrent regions in the genome without attendant large-scale chromosomal mutations and rearrangements.

In conclusion, HDR-mediated gene editing through *in trans* paired nicking offers high specificity and low mutagenicity, which, as a result, mostly preserves cellular genotypes and phenotypes. Moreover, the coordinated nicking of donor and acceptor HDR templates boosts the versatility of CRISPR-based gene editing by substantially enlarging the fraction of candidate gRNAs that can become operational, regardless of their *a priori* specificity profiles. The seamless and scarless character of *in trans* paired nicking should be particularly beneficial in instances in which precise and predictable genetic interventions are crucial. Examples include modelling or rescuing disease traits in stem cells (64) and functionally dissecting genomic sequences by multiplexed knock-in of donor DNA libraries (65). Finally, *in trans* paired nicking might expand the 'editable genome' to different types of repetitive elements shedding light on this large and variegated portion of the functionally unknown genomic 'dark matter' (66).

DATA AVAILABILITY

All data generated and analysed in this study are included in the article and its supplementary files. Additional raw datasets are available from the corresponding author on reasonable request. The Gene Expression Omnibus (GEO) datasets corresponding to the orthogonal HTGTS analyses are available via accession code GSE135064.

SUPPLEMENTARY DATA

Supplementary Data are available online at doi: 10.1093/nar/gkz1121.

ACKNOWLEDGEMENTS

The authors thank Martijn Rabelink, Steve Cramer and Hans Vrolijk (Department of Cell and Chemical Biology, LUMC, Leiden, the Netherlands) for performing p24^{gag} ELISA titrations of lentiviral vector particles, assistance with the research logistics and making available the ColorProc source code, respectively. The authors also thank Elisa Giacomelli and Dorien Ward-van Oostwaard for sharing their iPSC culture protocols and Christian Freund (Department of Anatomy and Embryology, LUMC, Leiden, the Netherlands) for making available the Nucleofector 2b-device.

FUNDING

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 765269; and from the Dutch Prinses Beatrix Spierfonds (W.OR16-13); X.C. and Q.W. receive support from Ph.D. fellowships from the China Scholarship Council-Leiden University Joint Scholarship Programme.
Conflict of interest statement. None declared.

REFERENCES

1. Chandrasegaran,S. and Carroll,D. (2016) Origins of programmable nucleases for genome engineering. *J. Mol. Biol*., **428**, 963-989.
2. Maggio, I. and Gonçalves,M.A. (2015) Genome editing at the crossroads of delivery, specificity, and fidelity. *Trends Biotechnol*., **33**, 280-291.
3. Chang,H.H.Y., Pannunzio,N.R., Adachi,N. and Lieber,M.R. (2017) Non-homologous DNA end joining and alternative pathways to double-strand break repair. *Nat. Rev. Mol. Cell. Biol*., **18**, 495-506.
4. Kosicki,M., Tomberg,K. and Bradley,A. (2018) Repair of double-strand breaks induced by CRISPR-Cas9 leads to large deletions and complex rearrangements. *Nat. Biotechnol*. 2018, **36**, 765-771.
5. Cradick,T.J., Fine, E.J., Antico,C.J. and Bao,G. (2013) CRISPR/Cas9 systems targeting β-globin and CCR5 genes have substantial off-target activity. *Nucleic Acids Res*., **41**, 9584-9592.
6. Fu,Y., Foden,J.A., Khayter,C., Maeder,M.L., Reyon,D., Joung,J.K. and Sander,J.D. (2013) High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nat. Biotechnol*., **31**, 822-826.
7. Cho,S.W., Kim,S., Kim,Y., Kweon,J., Kim,H.S., Bae,S. and Kim,J.S. (2014) Analysis of off-target effects of CRISPR/Cas-derived RNA-guided endonucleases and nickases. *Genome Res*., **24**, 132-141.
8. Lin,Y., Cradick,T.J., Brown,M.T., Deshmukh,H., Ranjan,P., Sarode,N., Wile,B.M., Vertino,P.M., Stewart,F.J. and Bao,G. (2014) CRISPR/Cas9 systems have off-target activity with insertions or deletions between target DNA and guide RNA sequences. *Nucleic Acids Res*., **42**, 7473-7485.
9. Kuscu,C., Arslan,S., Singh,R., Thorpe,J. and Adli,M. (2014) Genome-wide analysis reveals characteristics of off-target sites bound by the Cas9 endonuclease. *Nat. Biotechnol*., **32**, 677-683.
10. Frock,R.L., Hu,J., Meyers,R.M., Ho,Y.J., Kii,E. and Alt,F.W. (2015) Genome-wide detection of DNA double-stranded breaks induced by engineered nucleases. *Nat. Biotechnol*., **33**, 179-186.
11. Holkers,M., Maggio,I., Henriques,S.F., Janssen,J.M., Cathomen,T. and Gonçalves MA. (2014) Adenoviral vector DNA for accurate genome editing with engineered nucleases. *Nat. Methods*., **11**, 1051-1057.
12. Chen,X., Janssen,J.M., Liu,J., Maggio,I., 't Jong,A.E., Mikkers,H.M., Gonçalves,M.A. (2017) In trans paired nicking triggers seamless genome editing without double-stranded DNA cutting. *Nat. Commun*., **8**, 657.
13. Zhang,J.P., Li,X.L., Li,G.H., Chen,W., Arakaki,C., Botimer,G.D., Baylink,D., Zhang,L., Wen,W., Fu,Y.W. *et al*. (2017) Efficient precise knockin with a double cut HDR donor after CRISPR/Cas9-mediated double-stranded DNA cleavage. *Genome Biol*., **18**, 35.
14. McConnell Smith,A., Takeuchi,R., Pellenz,S., Davis,L., Maizels,N., Monnat,R.J.Jr. and Stoddard,B.L. (2009) Generation of a nicking enzyme that stimulates site-specific gene conversion from the I-AniI LAGLIDADG homing endonuclease. *Proc. Natl. Acad. Sci. USA*., **106**, 5099-5104.
15. Metzger,M.J., McConnell-Smith,A., Stoddard,B.L. and Miller,A.D. (2011) Single-strand nicks induce homologous recombination with less toxicity than double-strand breaks using an AAV vector template. *Nucleic Acids Res*., **39**, 926-935.
16. Ramirez,C.L., Certo,M.T., Mussolino,C., Goodwin,M.J., Cradick,T.J., McCaffrey,A.P., Cathomen,T., Scharenberg,A.M. and Joung,J.K. (2012) Engineered zinc finger nickases induce homology-directed repair with reduced mutagenic effects. *Nucleic Acids Res*., **40**, 5560-5568.
17. Wang,J., Friedman,G., Doyon,Y., Wang,N.S., Li,C.J., Miller,J.C., Hua,K.L., Yan,J.J., Babiarz,J.E., Gregory,P.D. *et al*. (2012) Targeted gene addition to a predetermined site in the human genome using a ZFN-based nicking enzyme. *Genome Res*., **22**, 1316-1326.

18. Jinek,M., Chylinski,K., Fonfara,I., Hauer,M., Doudna,J.A. and Charpentier,E. (2012) A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science*, **337**, 816-821.
19. Cong,L., Ran,F.A., Cox,D., Lin,S., Barretto,R., Habib,N., Hsu,P.D., Wu,X., Jiang,W., Marraffini,L.A. *et al*. (2013) Multiplex genome engineering using CRISPR/Cas systems. *Science*, **339**, 819-823.
20. Mali,P., Yang,L., Esvelt,K.M., Aach,J., Guell,M., DiCarlo,J.E., Norville,J.E. (2013) and Church,G.M. RNA-guided human genome engineering via Cas9. *Science*, **339**, 823-826.
21. Doudna,J.A. and Charpentier,E. (2014) Genome editing. The new frontier of genome engineering with CRISPR-Cas9. *Science*, **346**, 1258096.
22. Mali,P., Aach,J., Stranges,P.B., Esvelt,K.M., Moosburner,M., Kosuri,S., Yang,L. and Church,G.M. (2013) CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. *Nat. Biotechnol*., **31**, 833-838.
23. Ran,F.A., Hsu,P.D., Lin,C.Y., Gootenberg,J.S., Konermann,S., Trevino,A.E., Scott,D.A., Inoue,A., Matoba,S., Zhang,Y. *et al*. (2013) Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. *Cell*, **154**, 1380-1389.
24. van Nierop,G.P., de Vries,A.A., Holkers,M., Vrijsen,K.R. and Gonçalves,M.A. (2009) Stimulation of homology-directed gene targeting at an endogenous human locus by a nicking endonuclease. *Nucleic Acids Res*., **37**, 5725-5736.
25. Nakajima,K., Zhou,Y., Tomita,A., Hirade,Y., Gurumurthy,C.B. and Nakada,S. (2018) Precise and efficient nucleotide substitution near genomic nick via noncanonical homology-directed repair. *Genome Res*., **28**, 223-30.
26. Paulsen B.S., Mandal P.K., Frock R.L., Boyraz B., Yadav R, Upadhyayula S., Gutierrez-Martinez P., Ebina W., Fasth A., Kirchhausen T. *et al*. (2017) Ectopic expression of RAD52 and dn53BP1 improves homology-directed repair during CRISPR–Cas9 genome editing. *Nat. Biomed. Eng.*, **1**, 878–888.
27. Zhang,M., D'Aniello,C., Verkerk,A.O., Wrobel,E., Frank,S., Ward-van Oostwaard,D., Piccini,I., Freund,C., Rao,J., Seebohm,G. *et al*. (2014) Recessive cardiac phenotypes in induced pluripotent stem cell models of Jervell and Lange-Nielsen syndrome: disease mechanisms and pharmacological rescue. *Proc. Natl. Acad. Sci. USA*., **111**, E5383-E5392.
28. Hockemeyer,D., Wang,H., Kiani,S., Lai,C.S., Gao,Q., Cassady,J.P., Cost,G.J., Zhang,L., Santiago,Y., Miller,J.C. *et al*. (2011) Genetic engineering of human pluripotent cells using TALE nucleases. *Nat. Biotechnol*., **29**, 731-734.
29. Chen,X., Liu,J., Janssen,J.M. and Gonçalves,M.A. (2017) The chromatin structure differentially impacts high-specificity CRISPR-Cas9 nuclease strategies. *Mol. Ther. Nucleic Acids*, **8**, 558-563.
30. Kleinstiver,B.P., Prew,M.S., Tsai,S.Q., Nguyen,N.T., Topkar,V.V., Zheng,Z. and Joung,J.K. (2015) Broadening the targeting range of Staphylococcus aureus CRISPR-Cas9 by modifying PAM recognition. *Nat. Biotechnol*., **33**, 1293-1298.
31. Hu,J., Meyers,R.M., Dong,J., Panchakshari,R.A., Alt,F.W. and Frock,R.L. (2016) Detecting DNA double-stranded breaks in mammalian genomes by linear amplification-mediated high-throughput genome-wide translocation sequencing. *Nat. Protoc*., **11**, 853-871.
32. Brielmeier,M., Béchet,J.M., Falk,M.H., Pawlita,M., Polack,A. and Bornkamm,G.W. (1998) Improving stable transfection efficiency: antioxidants dramatically improve the outgrowth of clones under dominant marker selection. *Nucleic Acids Res*., **26**, 2082-2085.
33. Szuhai,K. and Tanke,H.J. (2006) COBRA: combined binary ratio labeling of nucleic-acid probes for multi-color fluorescence in situ hybridization karyotyping. *Nat. Protoc*., **1**, 264-275.
34. Hsiau,,T., Maures,T., Waite,K., Yang,J., Kelso,R., Holden,K. and Stoner,R. (2018) Inference of CRISPR Edits from Sanger Trace Data. biorXiv, 10.1101/251082v2, 14-January-2018, pre-print: not peer-reviewed.
35. Bae,S., Park,J. and Kim,J.S. (2014) Cas-OFFinder: a fast and versatile algorithm that searches for potential off-target sites of Cas9 RNA-guided endonucleases. *Bioinformatics*, **30**, 1473-1475.
36. Doench,J.G., Fusi,N., Sullender,M., Hegde,M., Vaimberg,E.W., Donovan,K.F., Smith,I., Tothova,Z., Wilen,C., Orchard,R. *et al*. (2016) Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. *Nat. Biotechnol*., **34**, 184-191.
37. Schwartz,S., Kent,W.J., Smit,A., Zhang,Z., Baertsch,R., Hardison,R.C., Haussler,D. and Miller,W. (2003) Human-mouse alignments with BLASTZ. *Genome Res*., **13**, 103-107.

38. Pelascini,L.P., Janssen,J.M. and Gonçalves,M.A. (2013) Histone deacetylase inhibition activates transgene expression from integration-defective lentiviral vectors in dividing and non-dividing cells. *Hum Gene Ther*., **24**, 78-96.
39. Pelascini,L.P. snf Gonçalves,M.A. (2014) Lentiviral vectors encoding zinc-finger nucleases specific for the model target locus HPRT1. *Methods Mol. Biol*., **1114**, 181-199.
40. Gabriel,R., von Kalle,C. and Schmidt,M. (2015) Mapping the precision of genome editing. *Nat. Biotechnol*. , **33**, 150-152.
41. Wienert,B., Wyman,S.K., Richardson,C.D., Yeh,C.D., Akcakaya,P., Porritt,M.J.., Morlock,M., Vu,J.T., Kazane,K.R., Watry,H.L. *et al*. (2019) Unbiased detection of CRISPR off-targets in vivo using DISCOVER-Seq. *Science* , **364**, 286-289.
42. Kuzminov,A. (2001) Single-strand interruptions in replicating chromosomes cause double-strand breaks. *Proc. Natl. Acad. Sci. USA* , **98**, 8241-8246.
43. Casper,J., Zweig,A.S., Villarreal,C., Tyner,C., Speir,M.L., Rosenbloom,K.R., Raney,B.J., Lee,C.M., Lee,B.T., Karolchik,D. *et al*. (2018) The UCSC Genome Browser database: 2018 update. *Nucleic Acids Res*., **46**, D762-D769.
44. Foster,E.R. and Downs,J.A. (2005) Histone H2A phosphorylation in DNA double-strand break repair. *FEBS J*. , **272**, 3231-3240.
45. Holliday,R. (1964) A mechanism for gene conversion in fungi. *Genet. Res*. **5**:283–304.
46. Celeste,A., Difilippantonio,S., Difilippantonio,M.J., Fernandez-Capetillo,O., Pilch,D.R., Sedelnikova,O.A., Eckhaus,M., Ried,T., Bonner,W.M. and Nussenzweig,A. (2003) H2AX haploinsufficiency modifies genomic stability and tumor susceptibility. *Cell*., **114**, 371-383.
47. Hanzlikova,H., Gittens,W., Krejcikova,K., Zeng,Z. and Caldecott,K.W. (2017) Overlapping roles for PARP1 and PARP2 in the recruitment of endogenous XRCC1 and PNKP into oxidized chromatin. *Nucleic Acids Res*., **45**, 2546-2557.
48. Ronson,G.E., Piberger,A.L., Higgs,M.R., Olsen,A.L., Stewart,G.S., McHugh,P.J., Petermann,E. and Lakin ND. (2018) PARP1 and PARP2 stabilise replication forks at base excision repair intermediates through Fbh1-dependent Rad51 regulation. *Nat. Commun*. , **9**, 746.
49. Fogarty,N.M.E., McCarthy,A., Snijders,K.E., Powell,B.E., Kubikova,N., Blakeley,P., Lea,R., Elder,K., Wamaitha,S.E., Kim,D. *et al*. (2017) Genome editing reveals a role for OCT4 in human embryogenesis. *Nature*, **550**, 67-73.
50. Takahashi,K., Tanabe,K., Ohnuki,M., Narita,M., Ichisaka,T., Tomoda,K. and Yamanaka S. (2007) Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell*, **131**, 861-872.
51. Yilmaz,A., Peretz,M., Aharony,A., Sagi,I. and Benvenisty,N. (2018) Defining essential genes for human pluripotent stem cells by CRISPR-Cas9 screening in haploid cells. *Nat. Cell Biol*., **20**, 610-619.
52. Liu,J.C., Lerou,P.H. and Lahav,G. (2014) Stem cells: balancing resistance and sensitivity to DNA damage. *Trends Cell. Biol*., **24**, 268–274.
53. Ihry,R.J., Worringer,K.A., Salick,M.R., Frias,E., Ho,D., Theriault,K., Kommineni,S., Chen,J., Sondey,M., Ye,C. et al. (2018) p53 inhibits CRISPR-Cas9 engineering in human pluripotent stem cells. *Nat. Med*., **24**, 939-946.
54. Haapaniemi,E., Botla,S., Persson,J., Schmierer,B. and Taipale,J. (2018) CRISPR-Cas9 genome editing induces a p53-mediated DNA damage response. *Nat. Med*., **24**, 927-930.
55. Smalley,E. (2018) As CRISPR-Cas adoption soars, summit calls for genome editing oversight. *Nat. Biotechnol*., **36**, 4852018.
56. El-Brolosy,M.A., Kontarakis,Z., Rossi,A., Kuenne,C., Günther,S., Fukuda,N., Kikhi,K., Boezio,G.L.M., Takacs,C.M., Lai,S.L. *et al*. (2019) Genetic compensation triggered by mutant mRNA degradation. *Nature*, **568**, 193-197.
57. Ma,Z., Zhu,P., Shi,H., Guo,L., Zhang,Q., Chen,Y., Chen,S., Zhang,Z., Peng,J. and Chen,J. (2019) PTC-bearing mRNA elicits a genetic compensation response via Upf3a and COMPASS components. *Nature*, **568**, 259-263.
58. Tuladhar,R., Yeu,Y., Piazza,J.T, Tan,Z., Clemenceau,J.R., Wu,X., Barrett,Q., Herbert,J., Mathews,D.H., Kim,J. *et al*. (2019) CRISPR-Cas9-based mutagenesis frequently provokes on-target mRNA misregulation. *Nat. Commun*., **10**, 4056.
59. Zhu,Z., Verma,N., González,F., Shi,Z.D. and Huangfu,D. (2015) A CRISPR/Cas-mediated selection-free knockin strategy in human embryonic stem cells. *Stem Cell Rep*., **4**, 1103-1111.

60. Munoz,D.M., Cassiani,P.J., Li,L., Billy,E., Korn,J.M., Jones,M.D., Golji,J., Ruddy,D.A., Yu,K., McAllister,G. *et al*. (2016) CRISPR screens provide a comprehensive assessment of cancer vulnerabilities but generate false-positive hits for highly amplified genomic regions. *Cancer Discov*., **6**, 900-913.
61. Gonçalves,E., Behan,F.M., Louzada,S., Arnol,D., Stronach,E.A., Yang,F., Yusa,K., Stegle,O., Iorio,F. and Garnett,M.J. (2019) Structural rearrangements generate cell-specific, gene-independent CRISPR-Cas9 loss of fitness effects. *Genome Biol*., **20**, 27.
62. Chen,X. and Gonçalves,M.A. (2018) DNA, RNA, and Protein Tools for Editing the Genetic Information in Human Cells. *iScience*, **6**, 247-263.
63. Deng, C. and Capecchi,M.R. (1992) Reexamination of gene targeting frequency as a function of the extent of homology between the targeting vector and the target locus. *Mol. Cell Biol*., **12**, 3365-3371.
64. Bellin,M., Marchetto,M.C., Gage,F.H. and Mummery,C.L.(2012) Induced pluripotent stem cells: the new patient? *Nat. Rev. Mol. Cell. Biol*., **13**, 713-726.
65. Findlay,G.M., Boyle,E.A., Hause,R.J., Klein,J.C. and Shendure,J. (2014) Saturation editing of genomic regions by multiplex homology-directed repair. *Nature*, **513**, 120-123.
66. Sedlazeck,F.J., Lee,H., Darby,C.A. and Schatz,M.C. (2018) Piercing the dark matter: bioinformatics of long-range sequencing and mapping. *Nat. Rev. Genet*., **19**, 329-346.