Title

High-efficiency transgene integration by homology-directed repair in human primary cells using DNA-PKcs inhibition

Journal

Nature Biotechnology, 42(5)

ISSN

1087-0156

Authors

Selvaraj, Sridhar
Feist, William N
Viel, Sebastien
et al.

Publication Date

2024-05-01

DOI

10.1038/s41587-023-01888-4

nature biotechnology

Article

https://doi.org/10.1038/s41587-023-01888-4

High-efficiency transgene integration by homology-directed repair in human primary cells using DNA-PKcs inhibition

Received: 29 June 2022

Accepted: 28 June 2023

Published online: 3 August 2023

Check for updates

Sridhar Selvaraj[1,2], William N. Feist[1,2], Sebastien Viel[1,2,3,4], Sriram Vaidyanathan[1,2,5,6], Amanda M. Dudek[1,2], Marc Gastou[1,2], Sarah J. Rockwood[1,2], Freja K. Ekman[1,2], Aluya R. Oseghale[1,2], Liwen Xu[1,2], Mara Pavel-Dinu[1,2], Sofia E. Luna[1,2], M. Kyle Cromer[1,2,7], Ruhi Sayana[1,2], Natalia Gomez-Ospina[1,2] & Matthew H. Porteus[1,2]

Therapeutic applications of nuclease-based genome editing would benefit from improved methods for transgene integration via homology-directed repair (HDR). To improve HDR efficiency, we screened six small-molecule inhibitors of DNA-dependent protein kinase catalytic subunit (DNA-PKcs), a key protein in the alternative repair pathway of non-homologous end joining (NHEJ), which generates genomic insertions/deletions (INDELs). From this screen, we identified AZD7648 as the most potent compound. The use of AZD7648 significantly increased HDR (up to 50-fold) and concomitantly decreased INDELs across different genomic loci in various therapeutically relevant primary human cell types. In all cases, the ratio of HDR to INDELs markedly increased, and, in certain situations, INDEL-free high-frequency (>50%) targeted integration was achieved. This approach has the potential to improve the therapeutic efficacy of cell-based therapies and broaden the use of targeted integration as a research tool.

Genome editing is a method to change the nucleotide sequence of a cell with single-nucleotide precision. The most well-developed method of genome editing is using an engineered nuclease to create a site-specific DNA double-strand break (DSB) in the genome[1,2]. CRISPR–Cas9 is the most widely used engineered nuclease system for genome editing due to its ease of design, high activity and high specificity. CRISPR–Cas9 genome editing is a two-component system consisting of a guide RNA (gRNA) that is complementary in sequence to the genomic target site, which recruits the Cas9 nuclease to create a precise DSB[3–7]. There are redundant cellular mechanisms to repair the induced DSB[8]. DSBs can be repaired by the non-homologous end joining (NHEJ) pathway, a generally accurate form of repair that occasionally can result in small insertions/deletions (INDELs) often of a single nucleotide at the break site[9,10]. If the DSB is repaired by microhomology-mediated end joining (MMEJ), small to large INDELs are created[11–13]. Finally, homology-directed repair (HDR) involves the use of either endogenous or exogenous homologous donor template sequence to precisely repair the DSB. After providing an exogenous donor template, the HDR pathway can be harnessed for targeted integration of single-nucleotide changes or changes of several thousand base pairs[8,14]. HDR-mediated genome editing is currently the most flexible method of creating the widest variety of changes to the genome of a cell and has now entered clinical studies[15].

One of the hallmarks of nuclease-based genome editing is that in a population of cells, a mixture of INDELs and HDR-mediated targeted

[1]Department of Pediatrics, Stanford University, Stanford, CA, USA. [2]Institute for Stem Cell Biology and Regenerative Medicine, Stanford University, Stanford, CA, USA. [3]Immunology Department, Lyon Sud University Hospital, Pierre-Bénite, France. [4]International Center of Research in Infectiology, Lyon University, INSERM U1111, CNRS UMR 5308, ENS, UCBL, Lyon, France. [5]Center for Gene Therapy, Abigail Wexner Research Institute, Nationwide Children's Hospital, Columbus, OH, USA. [6]Department of Pediatrics, The Ohio State University, Columbus, OH, USA. [7]Department of Surgery, University of California, San Francisco, San Francisco, CA, USA. e-mail: mporteus@stanford.edu

insertions occurs. Moreover, while HDR-to-INDEL ratios of 1:1 or greater can be achieved, the frequency of INDELs is often greater than that of HDR[8,16]. An important advance in genome editing would be to develop methods in which HDR is significantly more frequent than INDELs, and, ideally, INDEL-free targeted integration might be achieved. During genome editing, HDR-based targeted integration has been shown to outcompete MMEJ-based INDELs but not NHEJ-based INDELs[16,17]. This is likely due to the NHEJ pathway being active in all phases of the cell cycle, while HDR and MMEJ are active in S and G2 phases of the cell cycle[18,19]. Thus, different methods for inhibition of the NHEJ pathway are being explored for enhancing the frequency of HDR-based gene targeting/targeted integration[20–22].

The NHEJ pathway is initiated following the binding of Ku70/Ku80 protein to the ends of the DSB[11,23]. Ku70/Ku80 then recruits the catalytic subunit of DNA-dependent protein kinase (DNA-PKcs) to form the DNA-dependent protein kinase complex. The DNA-PKcs is activated through autophosphorylation and is critical for progression of the NHEJ pathway[24]. Subsequently, the XRCC4, XLF and DNA ligase IV complex is recruited to ligate the broken ends[25]. Inhibition of Ku or DNA ligase IV through short hairpin RNA/short interfering RNA or small molecules has been shown to moderately improve HDR efficiency by reducing the levels of NHEJ[26–29]. In addition, peptide-based inhibition of 53BP1 (a protein that promotes the NHEJ pathway) can also result in increased frequencies of HDR-mediated genome editing[30]. A recent study showed that introduction of a catalytically inactive mutation in DNA-PKcs leads to high levels of HDR-based gene targeting following gene editing with single-stranded oligodeoxynucleotide (ssODN) donor and Cas9 gRNA in human induced pluripotent stem cells (iPSCs)[31]. This enhancement in gene targeting could also be recapitulated with transient inhibition of DNA-PKcs using the small-molecule compound M3814 (refs. 16,31). Although, M3814 is a potent DNA-PKcs inhibitor, it has been shown to be less selective than other small-molecule DNA-PKcs inhibitors[32]. At the concentrations used for gene targeting, M3814 can inhibit various kinases of the PI3K family and the mTOR kinase, potentially causing cellular toxicity (Fig. 1a)[16,31]. Thus, DNA-PKcs inhibitors with higher specificity might result in further improvements in gene targeting efficiencies with mitigated cellular toxicity.

We screened various selective small-molecule DNA-PKcs inhibitors for enhancing gene targeting efficiency using Cas9 ribonucleoprotein (RNP) and adeno-associated virus serotype 6 (AAV6) HDR donor delivery in human PSCs[33]. From this screening, AZD7648 was identified as the most potent compound for enhancing gene targeting efficiency. AZD7648 was more potent than M3814 for improving gene targeting and has also been shown to be a more selective DNA-PKcs inhibitor than M3814 (Fig. 1a)[32]. AZD7648 treatment significantly enhanced gene targeting efficiency at different genomic loci for integration of both short and long sequences in various therapeutically relevant human primary cells, such as PSCs, hematopoietic stem and progenitor cells (HSPCs), T cells and human bronchial epithelial cells (HBECs). With AZD7648 treatment, we achieved high levels of gene targeting with low to no INDELs in most cases. Allelic gene targeting efficiency reached as high as 90% in some cases. Remarkably, AZD7648 could turn a gRNA that generated very few INDELs when delivered as an RNP alone without AZD7648 into a gRNA that generated ~50% targeted integrations without detectable INDELs when combined with AAV6 HDR donor and AZD7648. The use of AZD7648 resulted in flipping of the HDR-to-INDEL ratio from 1:1 to >5–100:1 in most circumstances. Thus, AZD7648 treatment can broadly enhance the ex vivo targeted integration frequency in primary cells, thereby improving the efficacy of ex vivo gene-targeted cell-based therapies and expanding the application of targeted integration for research purposes.

Results

AZD7648 enhances gene targeting efficiency in PSCs

The small-molecule DNA-PKcs inhibitors AZD7648 and VX984 have better selectivity profiles than M3814, as described previously[32] (Fig. 1a). To identify the most potent and selective DNA-PKcs small-molecule inhibitor for enhancing gene targeting efficiency, we performed a screen at the *CCR5* locus to knock-in two stop codons using the RNP and AAV6 gene editing platform in PSCs[33] (Fig. 1b). For this screen, we tested various previously described DNA-PKcs inhibitors, AZD7648 (refs. 32,34), M3814 (ref. 35), VX984 (ref. 36), KU57788 (ref. 37), BAY8400 (ref. 38) and LTURM34 (ref. 39), at three different concentrations (1, 0.1 and 0.01 μM). AZD7648, M3814, VX984 and BAY8400 improved gene targeting efficiency at a concentration of 1 μM. At 0.1 μM, only AZD7648 significantly enhanced gene targeting efficiency more than fourfold compared to no treatment (Fig. 1c and Extended Data Fig. 1a). Thus, AZD7648 is more potent than the other DNA-PKcs inhibitors for improving gene targeting efficiency. Next, we performed concentration gradient (1, 0.5, 0.25 and 0.1 μM) and time course (4, 8, 12 and 24 h) analyses for gene targeting with AZD7648 treatment. From this analysis, we found that between 0.25 and 0.5 μM was the optimal concentration and 24 h was the optimal incubation time for improving gene targeting (Fig. 1d,e and Supplementary Fig. 1a,b). Notably at the optimal concentration (0.5 and 0.25 μM) for gene targeting, AZD7648 was highly specific for DNA-PKcs inhibition (Fig. 1a).

CCR5 is a safe harbor genomic locus that can be used for integration of large gene cassettes with exogenous promoters to safely overexpress genes of interest for therapeutic and research purposes[40–43]. Thus, we tested whether DNA-PKcs inhibitor treatment can improve gene targeting efficiency at the *CCR5* locus for integration of large gene cassettes. We compared the effects of AZD7648 and other DNA-PKcs inhibitors on gene targeting efficiency at the *CCR5* locus for the knock-in of a multikilobase sequence (2.6 kb; Fig. 1f). We found that AZD7648 significantly enhanced allelic gene targeting efficiency at 1 and 0.1 μM, while M3814 and VX984 were effective only at 1 μM. With AZD7648 treatment, allelic gene targeting efficiency reached 85%, which is a 2.4-fold increase compared to that observed in untreated control cells. BAY8400 treatment significantly improved gene targeting at 0.1 μM (Fig. 1g). Next, we assessed cell viability at different time points (24, 48 and 72 h) after *CCR5* gene targeting with and without treatment with different DNA-PKcs inhibitors (AZD7648, M3814, VX984 and BAY8400 at 0.5 and 0.25 μM). AZD7648 treatment showed the highest improvement in gene targeting efficiency (Extended Data Fig. 1b). Cell viability analysis showed that AAV6 transduction alone led to high levels of toxicity, which has been reported in previous studies on AAV transduction in PSCs[44]. Among the gene-targeted cells, treatment with DNA-PKcs inhibitors led to slightly higher toxicity at 48 h and 72 h than observed in untreated cells. BAY8400 was found to be more toxic than other DNA-PKcs inhibitors at all three time points (Extended Data Fig. 1c). Next, we assessed biochemically whether treatment with AZD7648 effectively inhibits DNA-PKcs activity without affecting the function of kinases belonging to the PI3K family. For DNA-PKcs activity, we assessed autophosphorylation at Ser 2056 (ref. 32), and for PI3K function, we assessed phosphorylation of AKT at Ser 473 (ref. 45). We used bleomycin to induce broadscale DNA damage to enhance the detection of DNA-PKcs activity[46]. Western blotting showed that PSCs treated with bleomycin for 2 h exhibited autophosphorylation of DNA-PKcs (Ser 2056), while AZD7648 co-treatment inhibited this autophosphorylation. RNP/AAV6 gene editing did not result in detectable DNA-PKcs autophosphorylation due to minimal activation of the protein because the high-fidelity Cas9 RNP creates few breaks in the genome at any given point. Importantly, AZD7648 treatment did not affect phosphorylation of AKT (Ser 473) by PI3K family of kinases (Extended Data Fig. 2a). We also confirmed that AZD7648 treatment improved gene targeting efficiency in PSCs used for this biochemical analysis (Extended Data Fig. 2b).

To confirm the maintenance of pluripotency in gene-targeted PSCs, we assessed the expression of pluripotency markers and the ability to differentiate into three germ layers following gene targeting at the *CCR5* locus. Flow cytometry analysis showed that AZD7648

https://doi.org/10.1038/s41587-023-01888-4

a Specificity of different DNA-PKcs inhibitors

Targets	M3814 (IC_{50})	AZD7648 (IC_{50})	VX984 (IC_{50})
DNA-PKcs	43 nM	91.3 nM	115 nM
ATM	>30 μM	17.93 μM	>30 μM
ATR	>30 μM	>29.77 μM	>30 μM
PI3Kα	0.8 μM	>8.03 μM	>20 μM
PI3Kβ	0.17 μM	>30 μM	>30 μM
PI3Kγ	1.59 μM	1.37 μM	7.1 μM
PI3Kδ	0.35 μM	>30 μM	>30 μM
mTOR	0.55 μM	>30 μM	>30 μM

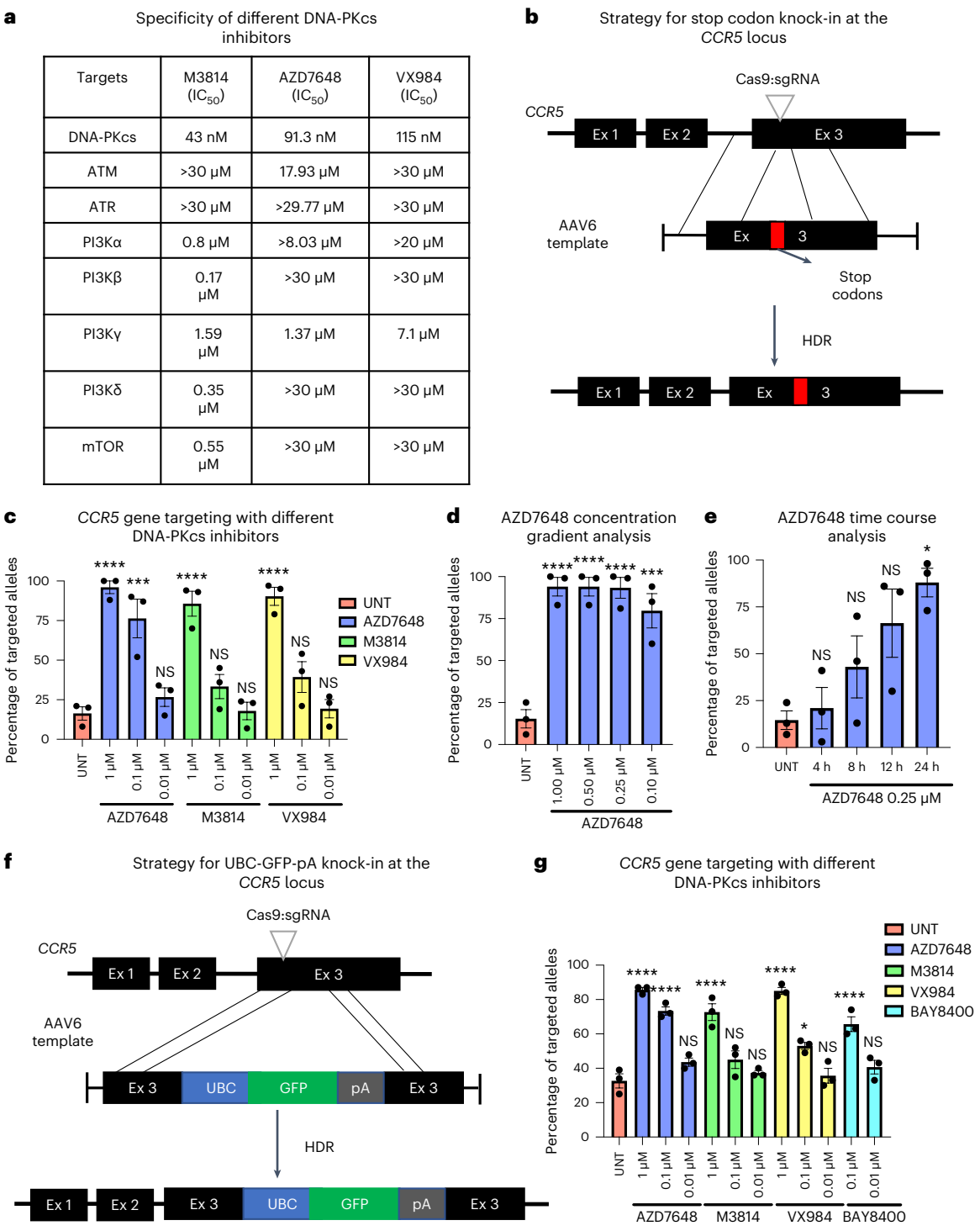


Fig. 1 | AZD7648 is the most potent DNA-PKcs inhibitor for improving gene targeting in PSCs. a, Table showing the half-maximal inhibitory concentration (IC_{50}) values of different small-molecule DNA-PKcs inhibitors (M3814, AZD7648 and VX984) against DNA-PKcs, ATM, ATR, mTOR and various PI3K family kinases in A549 cells (data are adapted from a previous study)[32]. **b**, Schematic of the gene targeting strategy to introduce two stop codons at the *CCR5* locus using Cas9 RNP and AAV6 gene editing. **c**, Allelic gene targeting efficiency after gene editing at the *CCR5* locus (**b**) with different concentrations (1, 0.1 and 0.01 μM) of DNA-PKcs inhibitors (AZD7648, M3814 and VX984) compared to untreated cells (UNT) as measured by inference of CRISPR edits (ICE) analysis 4 d after gene editing ($n = 3$). **d**, Concentration gradient analysis of AZD7648 treatment (1, 0.5, 0.25 and 0.1 μM) for gene targeting at the *CCR5* locus (**b**). The percentage of targeted alleles was measured by ICE analysis 4 d after gene editing ($n = 3$). **e**, Time course analysis of AZD7648 treatment (4, 8, 12 and 24 h) for gene targeting at the *CCR5* locus (**b**). The percentage of targeted alleles was measured by ICE analysis 4 d after gene editing ($n = 3$). **f**, Schematic for gene targeting at the *CCR5* locus for knock-in of UBC-GFP-bGH-pA sequence using Cas9 RNP and AAV6 gene editing. **g**, Gene targeting at the *CCR5* locus (**f**) with different concentrations (1, 0.1 and 0.01 μM) of DNA-PKcs inhibitors (AZD7648, M3814, VX984 and BAY8400). The percentage of targeted alleles was measured by droplet digital PCR (ddPCR) analysis 5 d after gene editing ($n = 3$). All data in **c**–**e** and **g** are shown as mean ± s.e.m. and were compared by one-way analysis of variance and Tukey's multiple comparisons test; $^{*}P < 0.05$; $^{***}P < 0.001$; $^{****}P < 0.0001$; NS, not significant.

treatment led to gene targeting in more than 90% of cells, which was a 2.5-fold increase compared to that observed in untreated control cells (Extended Data Fig. 2c). Analysis of the allelic distribution of wild-type (WT), INDEL and HDR frequencies showed that AZD7648 treatment improved HDR frequency 3-fold, with a corresponding increase in HDR-to-INDEL ratio from 0.3 to ~8 (~24-fold increase; Extended Data Fig. 2d). Interestingly, we found that treatment with AZD7648 led to a minor fraction of small to large deletions and a predominant increase in the frequency of WT in non-targeted alleles, while the remaining alleles showed a 1-base pair (bp) insertion. This 1-bp insertion was the only type of INDEL observed in non-AZD7648-treated genome-edited cells (Supplementary Fig. 2a). We found that gene-targeted PSCs with and without AZD7648 treatment maintained the expression of different pluripotency markers (Extended Data Fig. 3a). Differentiation of gene-targeted PSCs into three germ layers (ectoderm, mesoderm and endoderm) was also unaffected by AZD7648 treatment, further confirming the maintenance of pluripotency (Extended Data Fig. 3b).

We performed single-cell cloning of PSCs gene targeted at the *CCR5* locus and treated with different concentrations of AZD7648 (Extended Data Fig. 3c) and assessed the genotypes of the single-cell clones. We found that treatment with 0.5, 0.25 and 0.1 μM AZD7648 led to the generation of 90%, 80% and 60% of clones with biallelic knock-in, respectively, while in untreated cells, only 30% of clones had biallelic knock-in (Extended Data Fig. 3d). This result confirms that AZD7648 treatment leads to an increase in biallelic gene targeting. We then tested whether AZD7648 pretreatment before gene editing could further improve gene targeting efficiency. PSCs were pretreated with AZD7648 for 24 h before gene targeting and were then either left untreated or were continued on AZD7648 treatment after nucleofection. Allelic gene targeting efficiency was unaffected by pretreatment with AZD7648 with or without treatment after nucleofection compared to the corresponding control cells. This indicates that pretreatment has no effect, and treatment after editing is essential for improving gene targeting efficiency (Extended Data Fig. 3e).

We then tested the effect of AZD7648 treatment on improving gene targeting across different genomic loci in PSCs. We first tested AZD7648 treatment for editing sickle cell disease (SCD) and cystic fibrosis (CF) mutations at *HBB* and *CFTR* loci, respectively, using the previously described Cas9 RNP and AAV6 donor reagents in WT PSCs[47–49] (Extended Data Fig. 4a,b). At both concentrations (0.5 and 0.25 μM) tested, AZD7648 treatment improved gene targeting efficiencies at *HBB* and *CFTR* loci, with corresponding decreases in the levels of INDELs. Notably, AZD7648 treatment improved the mean ratio of HDR to INDELs from 5 to 26 at the *HBB* locus and from 2 to 13 at the *CFTR* locus (Extended Data Fig. 4a,b and Supplementary Fig. 2b,c). We then tested the effects of AZD7648 and M3814 treatment on the ability to knock-in a multikilobase sequence (2.6 kb) at the *HBB* locus. Although allelic gene targeting in untreated cells was quite high at 50%, treatment with AZD7648 or M3814 improved gene targeting by ~1.4-fold at the higher concentration (2 μM), and only AZD7648 showed a significant improvement at the lower concentration (0.5 μM; Extended Data Fig. 4c). Non-gene-targeted alleles in untreated cells showed a 9-bp deletion as the major INDEL, while there was also a small fraction of insertions and other deletions, resembling the pattern in the RNP-only control. AZD7648 treatment led to an increase in the frequency of WT alleles, and the remaining alleles had a small fraction of the 9-bp deletion (Supplementary Fig. 3a). We then assessed whether gene replacement efficiency could be improved with AZD7648 treatment at the *HBA1* locus using the previously described gene editing strategy[50]. AZD7648 treatment improved *HBA1* gene replacement efficiency by around fourfold at both high and low concentrations (2 and 0.5 μM), while M3814 was effective only at the higher concentration (2 μM; Extended Data Fig. 4d). Non-gene-targeted alleles in untreated cells showed a 1-bp insertion as the major INDEL, while there was also a small fraction of alleles with deletions, resembling the RNP-only control. AZD7648 treatment led to an increase in the frequency of WT alleles, a small fraction of alleles with deletions and few to no alleles with the 1-bp insertion (Supplementary Fig. 3b).

In conclusion, these results confirm that AZD7648 is the current most potent DNA-PKcs inhibitor for enhancing gene targeting efficiency across different genomic loci in PSCs without affecting pluripotency.

AZD7648 improves gene targeting efficiency in HSPCs

We compared the effects of various DNA-PKcs inhibitors (AZD7648, M3814, VX984 and BAY8400) at two different concentrations (2 μM and 0.5 μM) on improving gene targeting at the *CCR5* locus in primary human CD34$^+$ HSPCs. Allelic gene targeting efficiency and cell viability counts indicated that treatment with 0.5 μM inhibitor was less toxic than treatment with 2 μM inhibitor, with similar gene targeting efficiencies observed for all compounds. At 0.5 μM, AZD7648 showed the highest improvement in gene targeting efficiency compared to the other compounds (Extended Data Fig. 5a). We then determined the optimal concentration of AZD7648 treatment with a concentration gradient analysis (1–0.01 μM) for gene targeting at the *CCR5* locus. We found that gene targeting efficiency remained high between 0.25 and 0.5 μM, which is similar to the results from PSCs, and we used 0.5 μM for all subsequent gene targeting experiments in HSPCs (Extended Data Fig. 5b). In non-gene-targeted alleles, we found that AZD7648 treatment at concentrations of 1, 0.5 and 0.25 μM led to an increase in the frequency of WT alleles and a decrease in the frequency of alleles with 1-bp insertions compared to untreated cells (Supplementary Fig. 3c).

We then tested the effect of AZD7648 on gene targeting across different therapeutically relevant genomic loci in HSPCs. AZD7648 treatment at the optimal 0.5 μM concentration improved allelic gene targeting efficiency by around threefold, with a concomitant sevenfold increase in HDR-to-INDEL ratio at the *CCR5* safe harbor locus for knock-in of a multikilobase sequence (2.6 kb; Fig. 2a). To confirm that gene targeting with AZD7648 does not affect the ability of HSPCs to differentiate into different hematopoietic lineages, we performed a colony-forming unit (CFU) assay in vitro. For this, we used *CCR5* gene-targeted HSPCs (RNP + AAV6) with and without AZD7648 treatment. As controls, we used mock-, AAV6-only- and RNP-only-treated cells with and without AZD7648 treatment and assessed colony formation. As has been previously described, the RNP/AAV6 system caused a decrease in the total number of colonies without a change in distribution[50–53], but AZD7648 treatment did not cause further decreases in total colony number nor a change in the colony type distribution (Fig. 2b and Extended Data Fig. 5c). We then determined the genotypes of the gene-targeted HSPC-derived erythroid burst-forming unit (BFU-E) and CFU granulocytes and monocytes (CFU-GM) colonies from the CFU assay and found that AZD7648 treatment increased the frequency of colonies with biallelic HDR knock-in by more than sevenfold in both colony types (9% to 67% increase in BFU-E and 9% to 78% increase in CFU-GM; Fig. 2c). Thus, AZD7648 treatment increases the frequency of biallelic HDR knock-in without affecting the differentiation potential of HSPCs.

We assessed whether AZD7648 treatment improved gene targeting in long-term hematopoietic stem cells (LT-HSCs) and multipotent progenitors (MPPs). For this, we gene targeted CD34$^+$ HSPCs at the *CCR5* locus with and without AZD7648 treatment and sorted the LT-HSC and MPP populations by fluorescence-activated cell sorting (Supplementary Fig. 4a). AZD7648 treatment improved the frequency of gene-targeted cells in LT-HSC and MPP populations by more than 2.5-fold, as assessed by flow cytometry for green fluorescent protein (GFP; Supplementary Fig. 4b). Allelic gene targeting efficiency was improved by almost fourfold with AZD7648 treatment in LT-HSC and MPP populations, with a concomitant increase in HDR-to-INDEL ratio (Fig. 2d). Thus, AZD7648 treatment improves gene targeting in

https://doi.org/10.1038/s41587-023-01888-4

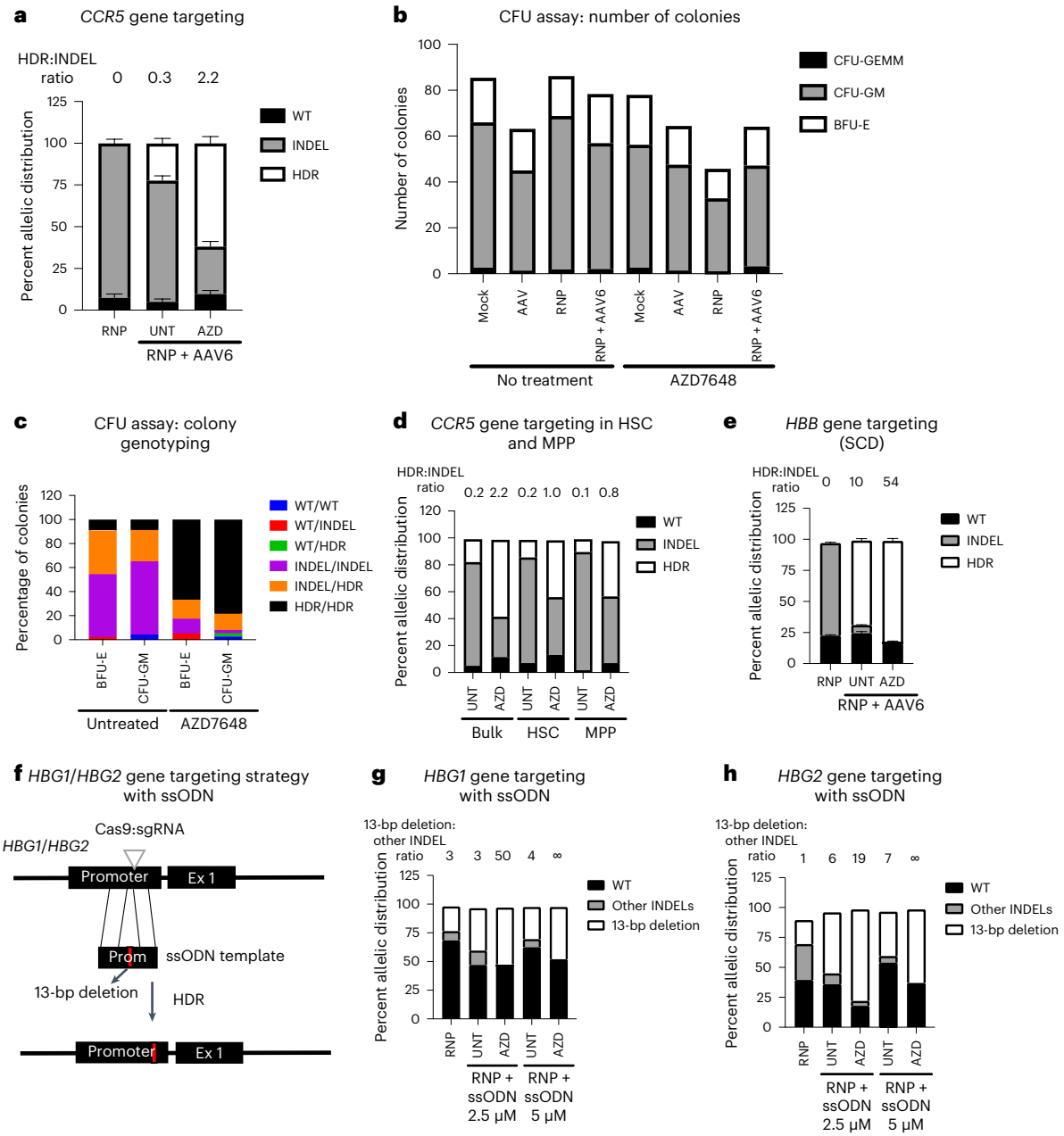


Fig. 2 | AZD7648 improves gene targeting in HSPCs. a, Allelic distribution of WT, INDEL and HDR frequencies following gene editing at the *CCR5* locus for knock-in of the UBC-GFP-bGH-pA sequence using RNP/AAV6 gene editing with and without AZD7648 treatment (0.5 μM; $n = 6$). HDR frequency was measured by ddPCR analysis, and WT and INDEL frequencies were measured by ICE analysis. Mean HDR-to-INDEL ratios are represented above the bars. **b**, A CFU assay was performed on mock-, RNP-, AAV6- and RNP + AAV6-treated HSPCs following gene editing at the *CCR5* locus with or without treatment with 0.5 μM AZD7648. The plot shows the distribution of the mean absolute number of multipotent granulocyte, erythroid, macrophage and megakaryocyte progenitor cells (CFU-GEMM), CFU-GM and BFU-E colonies ($n = 2$). **c**, Single-cell colonies (BFU-E and CFU-GM) from RNP + AAV6-treated cells with or without AZD7648 treatment from **b** were genotyped to assess WT, INDEL and HDR frequencies. The plot shows the distribution of the percentage of BFU-E and CFU-GM colonies with different genotypes ($n = 1$). **d**, Allelic distribution of mean WT, INDEL and HDR frequencies following gene editing at the *CCR5* locus for knock-in of the UBC-GFP-bGH-pA sequence using RNP/AAV6 with and without AZD7648 (0.5 μM) treatment in bulk, LT-HSC and MPP populations ($n = 2$). HDR frequency was measured by ddPCR analysis, and WT and INDEL frequencies were measured by ICE analysis. Mean HDR-to-INDEL ratios are represented above the bars. **e**, Allelic distribution of WT, INDEL and HDR frequencies following gene editing at the *HBB* locus for editing the SCD mutation (E6V) using RNP/AAV6 with and without AZD7648 (0.5 μM) treatment as measured by ICE analysis ($n = 4$). Mean HDR-to-INDEL ratios are represented above the bars. **f**, Schematic for gene targeting at the *HBG1* and *HBG2* promoters to introduce a 13-bp deletion using Cas9 RNP- and ssODN-based donor delivery. **g**,**h**, Allelic distribution of mean WT, 13-bp deletion and other INDEL frequencies at *HBG1* (**g**) and *HBG2* (**h**) loci following gene editing (**f**) as measured by ICE analysis ($n = 2$). The mean ratios of the 13-bp deletion to other INDELs are represented above the bars. All data in **a** and **e** are shown as mean ± s.e.m.

phenotypic LT-HSCs, which are essential for achieving sustained clinical benefit in potential therapeutic applications.

We then tested AZD7648 treatment for editing the SCD mutation at the *HBB* locus[47,48]. *HBB* gene targeting efficiency was improved with AZD7648 treatment to nearly INDEL-free levels with a fivefold (10 to 54) increase in HDR-to-INDEL ratio (Fig. 2e and Supplementary Fig. 5a). Next, we tested the ability of AZD7648 treatment to improve gene replacement at the *HBA1* locus with *HBB* gene sequence, a previously

reported therapeutic gene editing strategy for β-thalassemia (Extended Data Fig. 5d)[50]. We found that gene replacement efficiency improved by about 1.4-fold with AZD7648 treatment with a 2-fold increase in HDR-to-INDEL ratio (Extended Data Fig. 5e). This improvement in gene targeting was only modest possibly because of the use of split homology arms, limiting the improvement in frequency of HDR-based targeted integration. In non-gene-targeted alleles, we found that AZD7648 treatment led to an increase in the frequency of WT alleles and a decrease in the frequency of alleles with 1-bp insertions compared to no treatment (the same as observed at the *CCR5* locus, described above; Supplementary Fig. 5b). Cell viability was assessed in HSPCs after gene targeting at the *CCR5*, *HBA1* and *HBB* loci. Similar to PSCs, we found some toxicity with transduction of AAV6 donors alone in HSPCs, as described previously[54]; but importantly, gene targeting with AZD7648 treatment resulted in only a small drop in viable cell count at 72 h after gene targeting at all three loci compared to mock-treated cells (Extended Data Fig. 5f).

We then tested whether gene targeting with AZD7648 treatment affected the off-target activity of Cas9 RNP gene editing. For this, we assessed the previously characterized top off-target sites for *CCR5* (OT39)[42], *HBB* (OT1)[48] and *HBA1* (OT1)[50] gRNAs in gene-targeted HSPCs using next-generation sequencing (NGS). At the OT1 site for *HBB* gRNA, we found that 1% of NGS reads had INDELs in untreated cells, while AZD7648-treated cells had about 3% of reads with INDELs. At the OT39 site for *CCR5* gRNA, AZD7648-treated cells showed 0.03% of reads with INDELs, while the frequency of reads with INDELs was below the detection limit in untreated cells. At the OT1 site for *HBA1* gRNA, AZD7648-treated cells had INDELs in around 0.5% of reads, while untreated cells had INDELs in 0.24% of reads. Interestingly, this increase in off-target activity observed with AZD7648 treatment was due to an increase in the frequency of deletion INDELs at all three assessed off-target sites, and this is likely due to use of the MMEJ pathway[16] (Extended Data Fig. 6a). Blocking of specific off-target breaks is possible, however, if there were biologically important reasons to do so[55,56]. We next assessed whether on-target activity of Cas9 RNP is affected by AZD7648 treatment in the absence of AAV6 donor template. For this, we used *CCR5* and *HBB* gRNAs and assessed the pattern of INDELs following AZD7648 treatment with and without AAV6 donor template. At the *CCR5* locus, we found an increase in the frequency of WT and deletion INDELs with AZD7648 treatment in RNP-only-treated cells. With the addition of AAV6 donor template to introduce two stop codons, we found that AZD7648 treatment enhanced HDR efficiency without affecting the pattern of INDELs in non-gene-targeted alleles (Extended Data Fig. 6b and Supplementary Fig. 5c,d). At the *HBB* locus, we found that AZD7648 treatment in RNP-only-treated cells eliminated the small fraction of alleles with insertion INDELs and increased the frequency of WT alleles. The addition of AAV6 donor template with AZD7648 treatment led to an increase in HDR without affecting the pattern of INDELs in non-targeted alleles (Extended Data Fig. 6c and Supplementary Fig. 6a,b). Thus, gene editing with AZD treatment increases the frequency of alleles with WT sequence and deletion INDELs without the HDR donor template.

To confirm whether AZD7648 treatment is also relevant for gene targeting using ssODN-based donor template delivery, we tested SCD mutation editing with Cas9 RNP and ssODN donor in WT HSPCs[57]. With the lowest concentration of ssODN donor tested (2.5 μM), AZD7648 treatment improved gene targeting frequency from 49% to 65% with a corresponding 2.7-fold increase in the HDR-to-INDEL ratio. At the highest concentration of ssODN, the gene targeting frequency was improved to a lesser extent with AZD7648 treatment, but there was a higher increase in HDR-to-INDEL ratio due to an increase in the frequency of WT alleles (Extended Data Fig. 6d and Supplementary Fig. 6c,d). Next, we further tested the applicability of AZD7648 treatment with ssODN donor for gene targeting at *HBG1* and *HBG2* to activate fetal hemoglobin expression (a therapeutic strategy for β-hemoglobinopathies)[58]. For this, we tested the feasibility of using an ssODN donor to introduce a naturally occurring 13-bp deletion mutation in the promoters of *HBG1* and *HBG2*, which is associated with a benign condition called hereditary persistence of fetal hemoglobin[59,60] (Fig. 2f). We used a previously reported gRNA near the target site, which, when delivered with Cas9 as RNP in HSPCs, led to a low frequency of the precise 13-bp deletion. This likely occurred through MMEJ repair of the DSB due to the presence of a 6-bp microhomology region around the cut site[61]. In accordance with the previous study, we observed an ~20% frequency of the 13-bp deletion in both *HBG1* and *HBG2* promoters with Cas9 RNP delivery. Inclusion of the ssODN donor with Cas9 RNP (untreated) resulted in a 2- to 2.5-fold increase in the frequency of the 13-bp deletion due to an HDR-based deletion. AZD7648 treatment improved the 13-bp deletion frequency further by about 1.3- to 1.5-fold compared to no treatment, with a concomitant increase in the ratio of the 13-bp deletion to other INDELs by several fold (Fig. 2g,h and Supplementary Fig. 7a–c). Thus, these results confirm the applicability of AZD7648 treatment for improving gene targeting across different genomic loci with either AAV6- or ssODN-based HDR donors in HSPCs.

AZD7648 improves gene targeting with low-activity gRNAs

We tested whether AZD7648 treatment can improve gene targeting efficiency while using a seemingly inactive gRNA, which creates very low levels of INDELs following Cas9 gRNA RNP delivery. For this, we tested gene targeting at the *STING1* locus to introduce a point mutation (V155M) associated with an autoinflammatory disease called STING-associated vasculopathy with onset in infancy (SAVI[62]; Fig. 3a). Following a gRNA screening near the mutation site in exon 5 of the *STING1* gene, we identified a seemingly inactive gRNA (sg5) that generates only 1% INDELs and a highly active gRNA (sg3) that generates around 70% INDELs following RNP delivery. Gene targeting with the seemingly inactive gRNA improved from ~1% to ~50% with AZD7648 treatment with no INDELs in PSCs. With the active gRNA, HDR frequency was improved by 2.3-fold with no INDELs following AZD7648 treatment in PSCs (Fig. 3b and Supplementary Fig. 8a,b). To test the feasibility of building a primary cell SAVI model, we next performed *STING1* gene targeting in human $CD34^+$ HSPCs. With the seemingly inactive gRNA, *STING1* gene targeting efficiency in HSPCs improved by almost tenfold (from 6% to 59%) with no INDELs following AZD7648 treatment. Gene targeting with the active gRNA improved gene targeting efficiency by 2.5-fold in HSPCs with a corresponding increase in HDR-to-INDEL ratio from 1 to 20 (Fig. 3c and Supplementary Fig. 8c).

Next, we assessed gene targeting at the *CCR5* locus using low-activity gRNAs sg1 and sg4 with AZD7648 treatment in PSCs and HSPCs (Supplementary Fig. 9a). sg1 and sg4 RNP-only gene editing resulted in 17.5% and 40% INDELs, respectively, in PSCs, while editing with the high-activity gRNA sg11 resulted in 90% INDELs. Gene targeting with AZD7648 treatment resulted in INDEL-free allelic gene targeting efficiencies of 75% and 80% with sg1 and sg4 gRNAs, respectively, for the introduction of two stop codons, while the HDR efficiency was less than 20% in non-AZD7648-treated cells (Fig. 3d and Supplementary Fig. 9b,c). In HSPCs, sg1 and sg4 RNP-only gene editing resulted in 0% and 11.5% INDELs, respectively, while sg11 RNP showed 85% INDELs. Gene targeting with AZD7648 treatment yielded INDEL-free HDR at efficiencies of 47.5% and 79%, respectively, which are four- to fivefold higher than that of untreated control cells for the introduction of two stop codons in HSPCs (Fig. 3e and Supplementary Fig. 9d). To confirm whether this finding is applicable for the insertion of large gene cassettes, we tested gene targeting at the *CCR5* locus for knock-in of a multikilobase sequence (2.6 kb) in HSPCs. With sg1, AZD7648 treatment resulted in an INDEL-free HDR efficiency of 25%, while untreated cells had an HDR efficiency of only 3.5%. With sg4, AZD7648 treatment yielded an HDR efficiency of 50%, a fivefold increase compared to that observed in untreated cells. Gene targeting using the high-activity gRNA sg11 with AZD7648 treatment resulted in an HDR efficiency of

https://doi.org/10.1038/s41587-023-01888-4

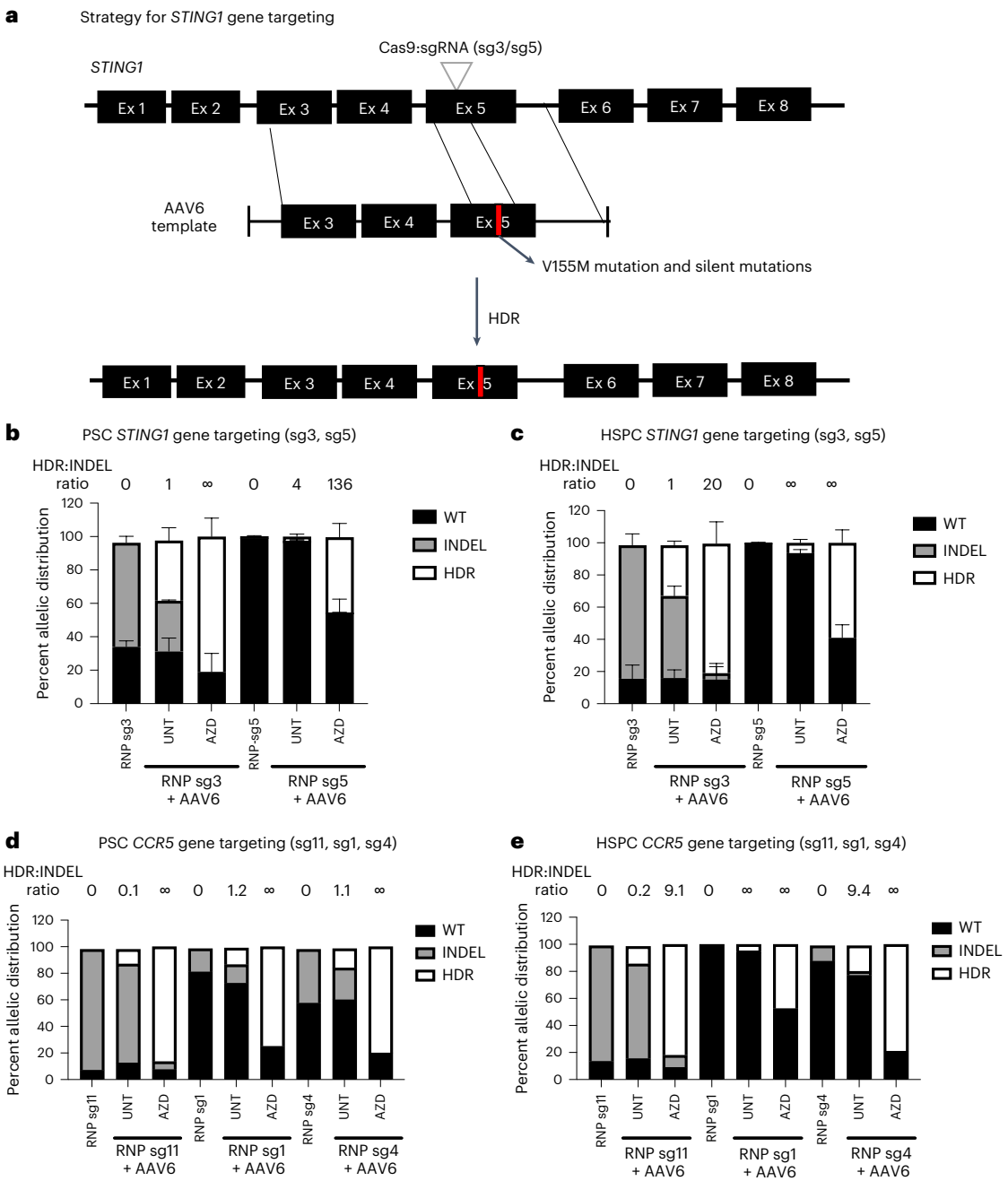


Fig. 3 | AZD7648 treatment improves gene targeting with seemingly inactive and low-activity gRNAs. a, Schematic for the gene targeting strategy at the *STING1* locus (exon 5) to introduce a point mutation (V155M) associated with SAVI disease using RNP (sg3 and sg5)/AAV6 gene editing. sg5 is a seemingly inactive gRNA, and sg3 is an active gRNA. **b**, Allelic distribution of WT, INDEL and HDR frequencies following gene editing at the *STING1* locus (**a**) using sg3 and sg5 gRNAs in PSCs with and without AZD7648 treatment (0.5 μM) as measured by ICE analysis ($n = 3$). **c**, Allelic distribution of WT, INDEL and HDR frequencies following *STING1* gene editing using sg3 and sg5 gRNAs in human HSPCs with and without AZD7648 treatment (0.5 μM) as measured by ICE analysis ($n = 3$). **d**, PSCs were gene targeted at the *CCR5* locus for the introduction of two stop codons using a high-activity gRNA (sg11) and two low-activity gRNAs (sg1 and sg4). The allelic distribution of mean WT, INDEL and HDR frequencies following *CCR5* gene editing using sg11, sg1 and sg4 gRNAs in human PSCs with and without AZD7648 treatment (0.5 μM) as measured by ICE analysis ($n = 2$) is shown. **e**, HSPCs were gene targeted at the *CCR5* locus for the introduction of two stop codons using sg11, sg1 and sg4 gRNAs. The allelic distribution of mean WT, INDEL and HDR frequencies following *CCR5* gene editing using sg11, sg1 and sg4 gRNAs in HSPCs with and without AZD7648 treatment (0.5 μM) as measured by ICE analysis ($n = 2$) is shown. Mean HDR-to-INDEL ratios are represented above the bars in **b**–**e**, and a ratio of ∞ denotes an INDEL frequency of zero. All data in **b** and **c** are shown as mean ± s.e.m.

54% (Extended Data Fig. 7a). In non-gene-targeted alleles, AZD7648 treatment resulted in a small fraction of deletion INDELs, which were absent in untreated cells for sg11 and sg4. For sg1, all non-gene-targeted alleles were found to be WT alleles with and without AZD7648 treatment (Supplementary Fig. 10a).

We then tested whether this finding would be applicable to gene targeting at exon 1 of the *IL2RG* locus using previously characterized high-activity (sg1) and low-activity (sg5, sg6 and sg7) gRNAs in HSPCs[51]. For this gene targeting, we designed two different AAV6 HDR donor template vectors with split homology arms (one vector for sg1 and sg6

and the other vector for sg1, sg5 and sg7; Extended Data Fig. 7b,c). sg1 RNP-only editing resulted in an INDEL frequency of 94%, while editing with sg6 yielded an INDEL frequency of only 26%. Gene targeting with AZD7648 treatment yielded an HDR efficiency of around 35% with sg6 gRNA, which was more than fourfold higher than that of untreated cells. With sg1, AZD7648 treatment resulted in a gene targeting efficiency of 53%, a 2.8-fold increase compared to non-AZD7648-treated cells (Extended Data Fig. 7d). sg5 and sg7 RNP-only editing resulted in low INDEL frequencies of 11% and 2%, respectively, in HSPCs. With AZD7648 treatment, sg5 and sg7 yielded gene targeting efficiencies of around 49% and 31%, respectively, in HSPCs, and this was more than a sixfold increase compared to the HDR efficiency observed in untreated cells (Extended Data Fig. 7e). In non-gene-targeted alleles, AZD7648 treatment increased the frequency of deletion INDELs with all four gRNAs at the *IL2RG* locus and interestingly reduced the frequency of WT alleles with the three low-activity gRNAs (Supplementary Fig. 10b,c). Importantly, at all three loci (*STING1*, *CCR5* and *IL2RG*), gene targeting with seemingly inactive and low-activity gRNAs yielded higher HDR efficiencies with AZD7648 treatment than observed in untreated cells gene targeted with the corresponding high-activity gRNAs.

Thus, AZD7648 treatment can improve gene targeting dramatically even with seemingly inactive and low-activity gRNAs across different genomic loci and can achieve INDEL-free HDR in human PSCs and HSPCs.

High HDR efficiency with lower amounts of RNP and AAV6

We tested whether AZD7648 treatment allows for titrating down the amounts of Cas9 RNP and AAV6 donor without affecting gene targeting efficiency. For RNP titration in PSCs, we tested reducing Cas9 RNP by two-, five- and tenfold for gene targeting at the *CCR5* locus. Titrating down the amount of RNP in untreated cells resulted in an RNP dose-dependent decrease in gene targeting efficiency. With AZD7648 treatment, gene targeting efficiency was consistently higher at different doses of RNP. Notably, allelic gene targeting efficiency and HDR-to-INDEL ratio with AZD7648 treatment at a tenfold lower dose of RNP (0.1×) were about two- and fivefold higher, respectively, than those observed in untreated cells with a full dose of RNP (1×; Fig. 4a). We confirmed these results by measuring the frequency of gene-targeted cells by flow cytometry, which showed a similar pattern (Extended Data Fig. 8a). AZD7648 treatment increased the frequency of WT alleles and resulted in a small fraction of deletion INDELs in the non-gene-targeted alleles; this pattern was consistent with different amounts of RNP (Supplementary Fig. 11a). We then tested the effect of RNP titration in human HSPCs for gene targeting at the *STING1* locus. Similar to PSCs, gene targeting efficiency and HDR-to-INDEL ratio with lower doses of RNP were consistently higher with AZD7648 treatment in HSPCs (Fig. 4b and Supplementary Fig. 11b).

We tested the effect of titrating down the AAV6 donor multiplicity of infection (MOI) in human PSCs and HSPCs (100 to 10,000 for PSCs and 100 to 5,000 for HSPCs) for gene targeting at the *CCR5* locus. Allelic gene targeting efficiency and HDR-to-INDEL ratio with AZD7648 treatment at an AAV6 MOI of 500 were higher than those of untreated cells at the highest tested MOI of 10,000 for PSCs and 5,000 for HSPCs (Fig. 4c,d). Flow cytometry analysis for the frequency of gene-targeted cells confirmed these findings in PSCs (Extended Data Fig. 8b). With AZD7648 treatment, non-gene-targeted alleles showed an increase in WT allele frequency and a small fraction of deletion INDEL alleles, which were absent in untreated cells (Supplementary Fig. 12a). In HSPCs, the frequency of gene-targeted cells at an AAV6 MOI of 500 with AZD7648 treatment was higher than that of untreated cells at an AAV6 MOI of 5,000 (Extended Data Fig. 8c). Similar to the results in PSCs, AZD7648 treatment resulted in an increase in the frequency of WT alleles and a small fraction of alleles with deletion INDELs in non-gene-targeted alleles in HSPCs (Supplementary Fig. 12b). We further tested the effect of AZD7648 treatment for gene targeting with a lower amount of AAV6 donor at the *HBB* locus for editing the SCD mutation. We observed a similar gene targeting efficiency and a threefold higher expansion with AZD7648 treatment at an AAV6 MOI of 600 than observed when gene targeting with an AAV6 MOI of 2,500 in the absence of AZD7648 (Extended Data Fig. 8d,e). Correspondingly, we noticed a twofold decrease in p53 response as measured by *CDKN1A* expression levels, which correlated with a lower number of AAV genomes per cell (Extended Data Fig. 8f,g). Because increased p53 response has been shown to lower the in vivo engraftment potential of HSPCs[63,64], this finding could be important for enhancing the engraftment and therapeutic potential of gene-targeted HSPCs.

Thus, these results show that AZD7648 treatment allows for reducing the amounts of RNP and AAV6 and maintenance of high gene targeting efficiencies.

AZD7648 treatment improves gene targeting in T and B cells

We then compared the effect of various DNA-PKcs inhibitors (AZD7648, M3814, VX984 and BAY8400) at two different concentrations (2 μM and 0.5 μM) for improving gene targeting at the *CCR5* locus in primary human T cells. Allelic gene targeting efficiency and cell viability counts indicated that treatment with 0.5 μM inhibitor was less toxic than treatment with 2 μM inhibitor, with similar gene targeting efficiencies observed for all compounds. At 0.5 μM, AZD7648 and M3814 showed the highest improvement in gene targeting efficiency compared to the other two compounds (Extended Data Fig. 9a). Next, we further tested and optimized AZD7648 treatment for gene targeting in T cells. Following a concentration gradient analysis, we found that 0.5 μM was the optimal concentration of the compound for gene targeting at the *CCR5* safe harbor locus in T cells, similar to findings in other cell types (Fig. 5a and Extended Data Fig. 9b). Importantly, we observed a minimal drop in cell viability at 72 h after gene targeting with different concentrations of AZD7648 compared to mock and untreated control cells (Extended Data Fig. 9c). In non-gene-targeted alleles, AZD7648 treatment (1 and 0.5 μM) increased the frequency of WT alleles with a concomitant decrease in alleles with 1-bp insertion INDELs, and there was a small fraction of alleles with deletion INDELs, which were absent in untreated cells. At 0.1 μM, AZD7648 treatment did not affect the INDEL pattern in non-targeted alleles (Supplementary Fig. 13a). Next, we tested the effect of compound treatment when titrating down AAV6 donor amounts for gene targeting at the *CCR5* locus. At the different AAV6 MOIs tested (1,000–10,000), we found that allelic gene targeting efficiency with AZD7648 was consistently higher with a corresponding increase in HDR-to-INDEL ratio. Notably, the HDR-to-INDEL ratio at an MOI of 1,000 with AZD7648 treatment was nearly threefold higher than that observed in untreated control cells at an MOI of 10,000 (Fig. 5b). We also found that the frequency of gene-targeted cells remained consistently high with AZD7648 treatment at different AAV6 MOIs with very little drop in cell viability (Extended Data Fig. 9d,e). In non-gene-targeted alleles, AZD7648 treatment induced a small fraction of deletion INDELs, which were absent in untreated cells, and this pattern was largely unaffected by the amount of AAV6 (Supplementary Fig. 13b). We then tested the effect of AZD7648 treatment in T cells for the knock-in of a chimeric antigen receptor (CAR), which is therapeutically relevant for cancer immunotherapies. For this, we assessed the CD19-CAR knock-in at the *TRAC* gene locus in T cells using a previously reported gene editing strategy[65] (Fig. 5c). At two different AAV6 MOIs tested (1,250 and 2,500), we found that gene targeting with AZD7648 treatment resulted in a 1.4-fold increase in the frequency of CD19-CAR$^+$ T cells compared to untreated control cells (Fig. 5d). To confirm that gene targeting with AZD7648 treatment does not affect T cell function, we assessed the cytotoxicity of CD19-CAR$^+$ T cells against a CD19$^+$ Nalm6 leukemia cell line expressing GFP. CD19-CAR$^+$ T cells and Nalm6 cells were cocultured at a 1:1 ratio, and the cytotoxic activity of T cells was assessed by flow cytometry for GFP at 24 and 48 h. CD19-CAR$^+$ T cells generated with and without AZD7648 treatment showed similar

https://doi.org/10.1038/s41587-023-01888-4

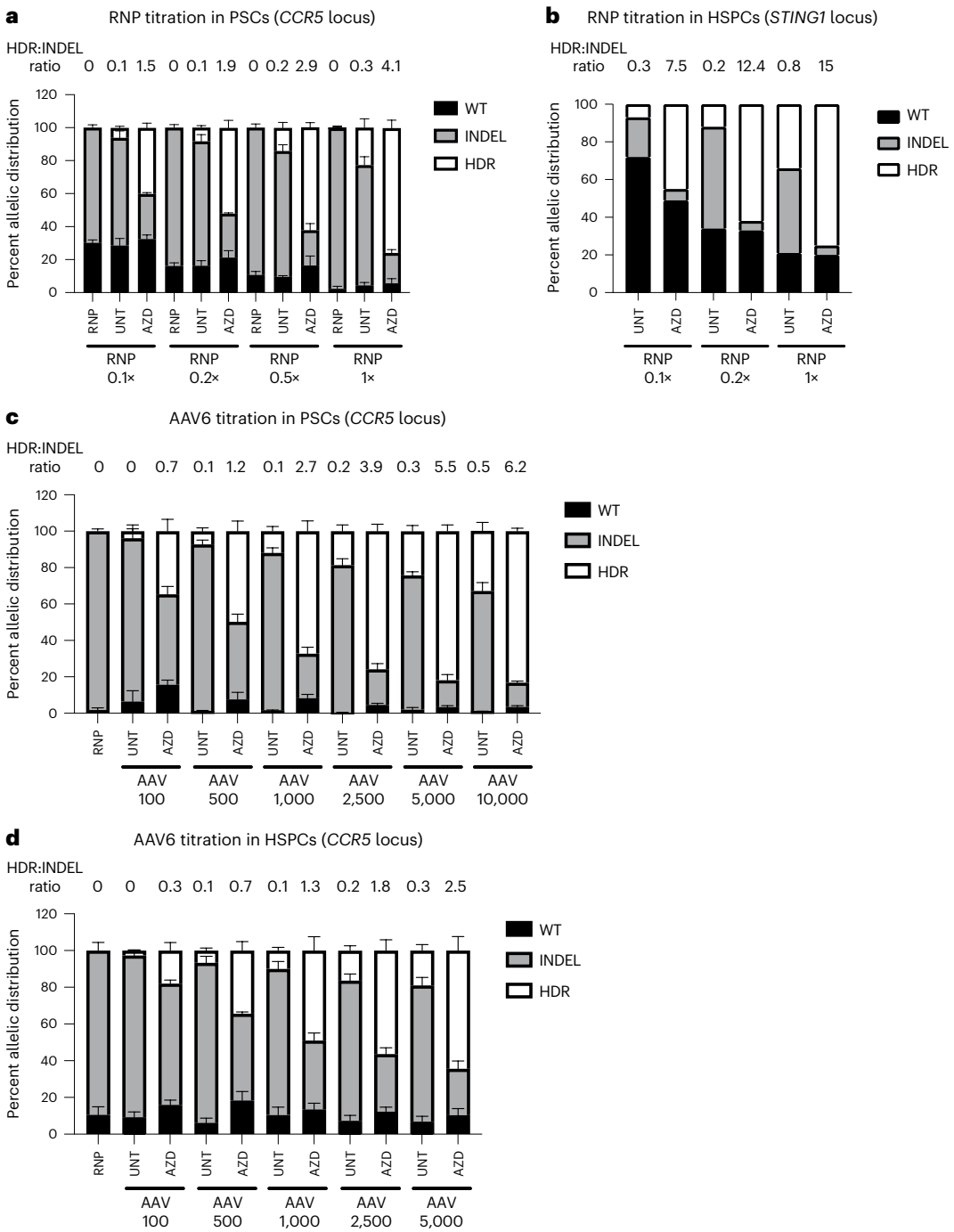


Fig. 4 | AZD7648 improves gene targeting with lower amounts of RNP and AAV6. a, Allelic distribution of WT, INDEL and HDR frequencies following gene editing at the *CCR5* locus for knock-in of the UBC-GFP-bGH-pA sequence in PSCs with or without AZD7648 treatment (0.25 μM) using different amounts of Cas9 RNP, as indicated, and a fixed amount of AAV6 donor (MOI of 2,500; $n = 3$). HDR frequency was measured by ddPCR analysis, and WT and INDEL frequencies were measured by ICE analysis. Mean HDR-to-INDEL ratios are represented above the bars. RNP 1× denotes 250 μg ml^{-1} Cas9 protein complexed with 100 μg ml^{-1} gRNA. **b**, Allelic distribution of WT, INDEL and HDR frequencies measured by ICE analysis following gene editing at the *STING1* locus (sg3) in human HSPCs for introduction of a V155M point mutation with or without AZD7648 treatment (0.5 μM) using different amounts of Cas9 RNP, as indicated, and a fixed amount of AAV6 donor (MOI of 2,500; $n = 1$). RNP 1× denotes 300 μg ml^{-1} Cas9 protein complexed with 160 μg ml^{-1} gRNA. HDR-to-INDEL ratios are represented above the bars. **c**, Allelic distribution of WT, INDEL and HDR frequencies following gene editing at the *CCR5* locus for knock-in of the UBC-GFP-bGH-pA sequence in PSCs with or without AZD7648 treatment (0.25 μM) using a fixed amount of RNP and different amounts of AAV6 donor, as indicated ($n = 3$). HDR frequency was measured by ddPCR analysis, and WT and INDEL frequencies were measured by ICE analysis. Mean HDR-to-INDEL ratios are represented above the bars. **d**, Allelic distribution of WT, INDEL and HDR frequencies following gene editing at the *CCR5* locus for knock-in of the UBC-GFP-bGH-pA sequence in human HSPCs with or without AZD7648 treatment (0.5 μM) using a fixed amount of RNP and different amounts of AAV6 donor, as indicated ($n = 3$). HDR frequency was measured by ddPCR analysis, and WT and INDEL frequencies were measured by ICE analysis. Mean HDR-to-INDEL ratios are represented above the bars. All data in **a**, **c** and **d** are shown as mean ± s.e.m.

https://doi.org/10.1038/s41587-023-01888-4

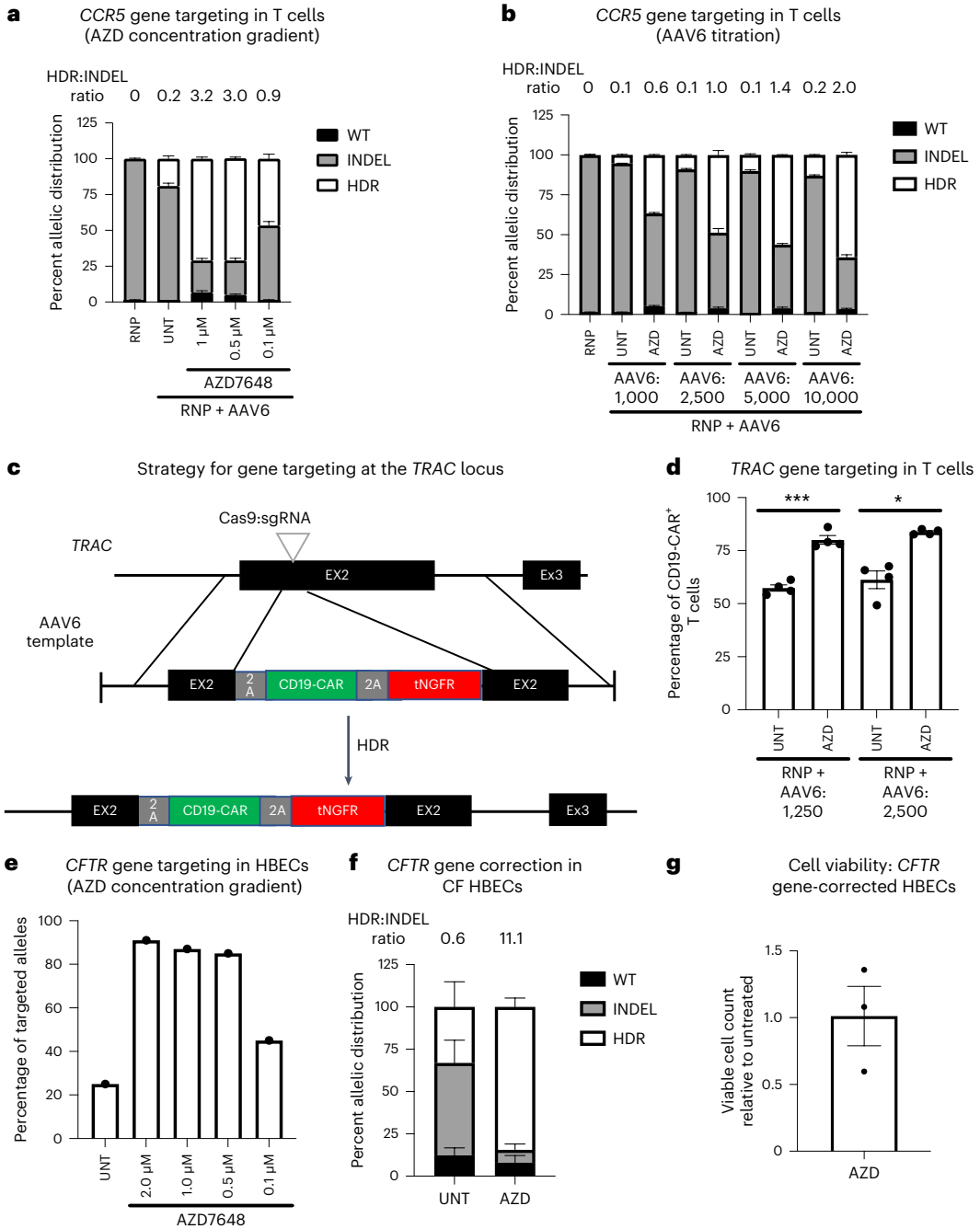


Fig. 5 | AZD7648 improves gene targeting in primary human T cells and HBECs. a, Allelic distribution of WT, INDEL and HDR frequencies following gene editing at the *CCR5* locus for knock-in of the UBC-GFP-bGH-pA sequence in T cells with different concentrations of AZD7648 (as indicated) using RNP/AAV6 gene editing ($n = 3$). HDR frequency was measured by ddPCR, and WT and INDEL frequencies were measured by ICE analysis. Mean HDR-to-INDEL ratios are represented above the bars. **b**, Allelic distribution of WT, INDEL and HDR frequencies following gene editing at the *CCR5* locus for knock-in of the UBC-GFP-bGH-pA sequence in T cells with or without AZD7648 (0.5 μM) using a fixed amount of RNP and different amounts of AAV6 donor, as indicated ($n = 3$). **c**, Schematic for gene targeting at the *TRAC* gene locus for knock-in of CD19-CAR using RNP/AAV6 gene editing. **d**, Frequency of CD19-CAR$^+$ cells measured by flow cytometry for truncated nerve growth factor receptor (tNGFR) expression following gene editing at the *TRAC* gene locus (**c**) in T cells with or without AZD7648 (0.5 μM) treatment. Gene editing was performed with a fixed amount of RNP and two different MOIs of the AAV6 donor ($n = 4$). Data were compared by *t*-test with Welch's correction; $^*P < 0.05$; $^{***}P < 0.001$. **e**, Percentage of gene-targeted alleles in WT HBECs after gene editing at the *CFTR* locus for editing the ΔF508 mutation using RNP/AAV6 with different concentrations of AZD7648, as indicated ($n = 1$). Allelic gene targeting frequency was measured by ICE analysis. **f**, Allelic distribution of WT, INDEL and HDR frequencies as measured by ICE analysis in HBECs derived from individuals with CF; HBECs were gene edited at the *CFTR* locus for the correction of the ΔF508 mutation using RNP/AAV6 with or without AZD7648 treatment (0.5 μM; $n = 4$). Mean HDR-to-INDEL ratios are represented above the bars. **g**, Viable cell count of *CFTR* gene-targeted HBECs with AZD7648 treatment (**f**) relative to that of untreated cells ($n = 3$). All data in **a**, **b**, **d**, **f** and **g** are shown as mean ± s.e.m.

cytotoxic activities, confirming that gene targeting with AZD7648 does not affect T cell function (Extended Data Fig. 9f). Thus, AZD7648 treatment can enhance the efficiency of therapeutically relevant gene targeting in T cells without affecting function.

We tested gene targeting at the *CCR5* locus in B cells with different concentrations of AZD7648 and observed a twofold increase in gene targeting efficiency with AZD7648 between the concentrations of 4 and 0.5 µM and a 1.7-fold increase at 0.1 µM. This improvement in gene targeting frequency and HDR-to-INDEL ratio in B cells was moderate at this locus compared to the other cell types tested (Extended Data Fig. 10a). In non-gene-targeted alleles, AZD7648 treatment at 4, 2, 1 and 0.5 µM showed an increase in the frequency of WT alleles and a small fraction of alleles with deletion INDELs, which were absent in untreated cells. At 0.1 µM, AZD7648 treatment did not affect the pattern of INDELs in non-gene-targeted alleles, similar to results in HSPCs and T cells (Supplementary Fig. 14a).

High-efficiency gene correction of HBECs derived from individuals with CF

We next tested the effect of AZD7648 on gene targeting in primary HBECs. First, we performed an AZD7648 concentration gradient analysis in WT HBECs for editing the CF mutation ΔF508 in exon 11 of the *CFTR* gene. Similar to other cell types, we found 0.5 µM to be the minimum optimal concentration required to improve gene targeting in HBECs, as measured by allelic gene targeting efficiency (Fig. 5e). We then tested the effect of AZD7648 on gene targeting at exon 1 of the *CFTR* gene for integration of a sequence larger than 1 kb in HBECs (Extended Data Fig. 10b). Following gene targeting, we found that AZD7648 treatment improved the frequency of GFP^+ cells by around twofold (Extended Data Fig. 10c,d). Next, we tested the effects of AZD7648 treatment on gene targeting to correct the ΔF508 mutation in HBECs derived from individuals with CF[49]. The gene correction efficiency in untreated cells was around 30%, which improved by 2.5-fold with AZD7648 treatment. Accordingly, we observed an 18.5-fold increase in the HDR-to-INDEL ratio with AZD7648 treatment, indicating a dramatic drop in INDEL frequency (Fig. 5f and Supplementary Fig. 14b). Importantly, cell viability was not affected in gene-targeted cells treated with AZD7648 compared to untreated cells (Fig. 5g). Thus, these results show the applicability of AZD7648 treatment for correction of the most common CF mutation in human-derived HBECs.

Overall, these results confirm the applicability of AZD7648 treatment for improving gene targeting across different genomic loci in a wide variety of human primary cells.

Discussion

Despite the development of methods for improving targeted integration in human primary cells, the ratio of HDR-based targeted integration to INDELs often remains at 1 or lower, which is a limitation for the application of targeted integration for therapeutic and research purposes[8]. Small-molecule-based inhibition of DNA-PKcs has been reported as a potential method for improving the frequency of HDR-based targeted integration through transient inhibition of the NHEJ pathway[16,17,27,31]. However, there is lack of a comprehensive evaluation of various commercially available small-molecule DNA-PKcs inhibitors to identify the most potent and selective compound for improving gene targeting. This is particularly important as highly potent and selective small-molecule DNA-PKcs inhibitors have been recently developed due to the relevance of DNA-PKcs as a target for anticancer therapy[66,67]. Following a screening of these recently developed compounds, we identified AZD7648 as the most potent DNA-PKcs inhibitor for improving gene targeting efficiency. Transient inhibition of DNA-PKcs by treatment with AZD7648 for 24 h during gene editing was sufficient to enhance gene targeting efficiency and increase the ratio of HDR to INDELs across different genomic loci in various therapeutically relevant human primary cells.

DNA-PKcs is a key player in the NHEJ-based repair of spontaneous and gene editing-induced DSBs[68], and, thus, it is essential to confirm the effect of AZD7648 treatment on off-target editing as it could affect genomic stability in gene-targeted cells. To confirm the safety of gene targeting with AZD7648 treatment, we assessed off-target activity in gene-targeted HSPCs and found an increase of two- to threefold in the frequency of off-target deletion INDELs potentially due to the activation of the MMEJ pathway. Despite this increase in off-target activity, the high specificity of HiFi Cas9-based gene editing is largely maintained, as it has been found to be 20- to 35-fold more specific than gene editing with WT Cas9 (refs. 69,70). If there were a biologically important rationale, specific off-target INDELs can be blocked using a gRNA/dead Cas9 RNP to block that site[55,56]. Moreover, a previous study showed that genomic stability was largely maintained in human iPSCs with a catalytically inactive mutation in DNA-PKcs even after treatment with the DSB-inducing drug bleomycin or extensive passaging. Interestingly, this study showed that the number of genomic translocations induced by bleomycin treatment and the number of mutations accumulated with passaging were fewer in DNA-PKcs-mutant human iPSCs than in WT cells[31]. Finally, we note that these DNA-PKcs inhibitors have been given in vivo to humans as part of clinical trials without noted adverse events (NCT03907969, NCT02644278 and NCT02316197). Thus, transient (24 h) small-molecule-based inhibition of DNA-PKcs during gene editing may not affect genomic integrity in gene-targeted cells.

PSCs hold great potential for applications such as disease modeling, drug screening and cell-based therapeutics[71,72]. Gene targeting of PSCs is broadly relevant for realizing the full potential of these applications[73]. Although the feasibility of single-cell cloning allows for the purification of gene-targeted PSCs, low gene targeting efficiency often results in a long and tedious process to isolate the clones with desired edits. With AZD7648 treatment, we consistently achieved 50% to 90% allelic gene targeting efficiency in PSCs with low to no INDELs across different genomic loci. This high gene targeting efficiency will make it easier to screen and isolate biallelic gene-targeted single-cell PSC clones.

Therapeutic gene editing of HSPCs for autologous cell therapy holds the potential for treatment of currently incurable diseases of the blood and immune system[47,50,51,74]. Gene-targeted HSPC-based autologous cell therapy has already entered clinical trials for SCD (NCT04819841)[48]. However, xenograft studies assessing the engraftment potential have shown that gene-targeted HSPCs possess a lower long-term engraftment capacity, which could potentially limit the therapeutic benefit[42,47,57,75]. One possible explanation for this could be the lower levels of gene targeting in quiescent stem cells with long-term engraftment capacity. Because AZD7648 treatment significantly enhances targeted integration frequency in both bulk HSPCs and LT-HSCs, it may also increase the frequency of gene-targeted cell engraftment, especially because AZD7648 allows lower amounts of both AAV6 and Cas9 RNP to be used while maintaining high levels of gene targeting. In addition, this enhanced gene targeting approach can also be useful for research applications, such as disease modeling with HSPCs, without the need for selection and enrichment of the targeted cells.

AZD7648 treatment also improved gene targeting in other primary cell types, such as T cells, B cells and HBECs, which shows the broad relevance of this approach across different cell types. Gene targeting in T cells is highly relevant for developing cancer immunotherapies[76], and here, we have shown that AZD7648 treatment can improve T cell gene targeting at therapeutically relevant *CCR5* (ref. 77) and *TRAC* loci[65,78]. HBECs are being considered as a potential cell source for developing autologous cell therapies to treat CF[49,79,80]. Gene targeting with AZD7648 treatment improved the gene correction efficiency of the most common CF mutation at the *CFTR* locus (ΔF508) by several fold in HBECs[49]. Thus, this gene targeting approach can improve the

therapeutic efficacy of potential ex vivo gene-targeted autologous cell-based therapies for CF.

One potential challenge associated with Cas9/gRNA-based gene editing is to find an active gRNA that creates high rates of INDELs near the genomic target site for achieving high-efficiency gene targeting[81]. Here, we have shown that application of AZD7648 treatment for gene editing with seemingly inactive and low-activity gRNAs can result in a high gene targeting efficiency of ~50% in both human PSCs and HSPCs. Using a low-activity gRNA, high frequencies of gene targeting were achieved with low to no INDELs. Thus, this gene targeting approach could expand the number of gRNAs that can be used for gene targeting due to the possibility of rescuing seemingly inactive gRNAs. This will be particularly important for contexts where there is limited availability of gRNA target sites near the genomic region of interest or when it is important to achieve zero or minimal INDELs as part of the process.

Furthermore, we found that AZD7648 treatment allows for titrating down the amount of Cas9 RNP and AAV6 donor without compromising the gene targeting efficiency. Reducing the AAV6 dose for gene targeting can help reduce the toxicity associated with AAV transduction in primary cells[44,54]. Although the off-target activity associated with Cas9 gRNA gene editing is significantly reduced through the use of the high-fidelity version of Cas9 nuclease[69], the use of lower amounts of high-fidelity Cas9 and gRNA for gene editing could further reduce off-target activity. Thus, titrating down the amounts of Cas9 RNP and AAV6 together to define the optimal amounts for gene targeting with AZD7648 could help achieve high gene targeting efficiency with reduced toxicity and off-target effects.

In conclusion, we have developed a gene targeting approach for achieving high gene targeting efficiency for precise knock-in of short to large sequences with low to no INDELs across different genomic loci in various human primary cells. This approach will broaden the application of targeted integration in human cells as a tool for therapeutic and research applications.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41587-023-01888-4.

References

1. Carroll, D. Genome engineering with targetable nucleases. *Annu. Rev. Biochem.* **83**, 409–439 (2014).
2. Gaj, T., Gersbach, C. A. & Barbas, C. F. III. ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. *Trends Biotechnol.* **31**, 397–405 (2013).
3. Horvath, P. & Barrangou, R. CRISPR/Cas, the immune system of bacteria and archaea. *Science* **327**, 167–170 (2010).
4. Doudna, J. A. & Charpentier, E. Genome editing. The new frontier of genome engineering with CRISPR–Cas9. *Science* **346**, 1258096 (2014).
5. Hsu, P. D., Lander, E. S. & Zhang, F. Development and applications of CRISPR–Cas9 for genome engineering. *Cell* **157**, 1262–1278 (2014).
6. Cong, L. et al. Multiplex genome engineering using CRISPR/Cas systems. *Science* **339**, 819–823 (2013).
7. Mali, P. et al. RNA-guided human genome engineering via Cas9. *Science* **339**, 823–826 (2013).
8. Nambiar, T. S., Baudrier, L., Billon, P. & Ciccia, A. CRISPR-based genome editing through the lens of DNA repair. *Mol. Cell* **82**, 348–388 (2022).
9. Betermier, M., Bertrand, P. & Lopez, B. S. Is non-homologous end-joining really an inherently error-prone process? *PLoS Genet.* **10**, e1004086 (2014).
10. Kanaar, R., Hoeijmakers, J. H. & van Gent, D. C. Molecular mechanisms of DNA double strand break repair. *Trends Cell Biol.* **8**, 483–489 (1998).
11. Chang, H. H. Y., Pannunzio, N. R., Adachi, N. & Lieber, M. R. Non-homologous DNA end joining and alternative pathways to double-strand break repair. *Nat. Rev. Mol. Cell Biol.* **18**, 495–506 (2017).
12. Sfeir, A. & Symington, L. S. Microhomology-mediated end joining: a back-up survival mechanism or dedicated pathway? *Trends Biochem. Sci.* **40**, 701–714 (2015).
13. Seol, J. H., Shim, E. Y. & Lee, S. E. Microhomology-mediated end joining: good, bad and ugly. *Mutat. Res.* **809**, 81–87 (2018).
14. Doudna, J. A. The promise and challenge of therapeutic genome editing. *Nature* **578**, 229–236 (2020).
15. Porteus, M. H. A new class of medicines through DNA editing. *N. Engl. J. Med.* **380**, 947–959 (2019).
16. Fu, Y. W. et al. Dynamics and competition of CRISPR–Cas9 ribonucleoproteins and AAV donor-mediated NHEJ, MMEJ and HDR editing. *Nucleic Acids Res.* **49**, 969–985 (2021).
17. Tatiossian, K. J. et al. Rational selection of CRISPR–Cas9 guide RNAs for homology-directed genome editing. *Mol. Ther.* **29**, 1057–1069 (2021).
18. Hustedt, N. & Durocher, D. The control of DNA repair by the cell cycle. *Nat. Cell Biol.* **19**, 1–9 (2016).
19. Xue, C. & Greene, E. C. DNA repair pathway choices in CRISPR–Cas9-mediated genome editing. *Trends Genet.* **37**, 639–656 (2021).
20. Ray, U. & Raghavan, S. C. Modulation of DNA double-strand break repair as a strategy to improve precise genome editing. *Oncogene* **39**, 6393–6405 (2020).
21. Sun, W. et al. Strategies for enhancing the homology-directed repair efficiency of CRISPR–Cas systems. *CRISPR J.* **5**, 7–18 (2022).
22. Bischoff, N., Wimberger, S., Maresca, M. & Brakebusch, C. Improving precise CRISPR genome editing by small molecules: is there a magic potion? *Cells* **9**, 1318 (2020).
23. Fattah, F. et al. Ku regulates the non-homologous end joining pathway choice of DNA double-strand break repair in human somatic cells. *PLoS Genet.* **6**, e1000855 (2010).
24. Frit, P., Ropars, V., Modesti, M., Charbonnier, J. B. & Calsou, P. Plugged into the Ku-DNA hub: the NHEJ network. *Prog. Biophys. Mol. Biol.* **147**, 62–76 (2019).
25. Zhao, B. et al. The essential elements for the noncovalent association of two DNA ends during NHEJ synapsis. *Nat. Commun.* **10**, 3588 (2019).
26. Chu, V. T. et al. Increasing the efficiency of homology-directed repair for CRISPR–Cas9-induced precise gene editing in mammalian cells. *Nat. Biotechnol.* **33**, 543–548 (2015).
27. Robert, F., Barbeau, M., Ethier, S., Dostie, J. & Pelletier, J. Pharmacological inhibition of DNA-PK stimulates Cas9-mediated genome editing. *Genome Med.* **7**, 93 (2015).
28. Maruyama, T. et al. Increasing the efficiency of precise genome editing with CRISPR–Cas9 by inhibition of nonhomologous end joining. *Nat. Biotechnol.* **33**, 538–542 (2015).
29. Pinder, J., Salsman, J. & Dellaire, G. Nuclear domain 'knock-in' screen for the evaluation and identification of small molecule enhancers of CRISPR-based genome editing. *Nucleic Acids Res.* **43**, 9379–9392 (2015).
30. Canny, M. D. et al. Inhibition of 53BP1 favors homology-dependent DNA repair and increases CRISPR–Cas9 genome-editing efficiency. *Nat. Biotechnol.* **36**, 95–102 (2018).
31. Riesenberg, S. et al. Simultaneous precise editing of multiple genes in human cells. *Nucleic Acids Res.* **47**, e116 (2019).
32. Fok, J. H. L. et al. AZD7648 is a potent and selective DNA-PK inhibitor that enhances radiation, chemotherapy and olaparib activity. *Nat. Commun.* **10**, 5065 (2019).

33. Martin, R. M. et al. Highly efficient and marker-free genome editing of human pluripotent stem cells by CRISPR–Cas9 RNP and AAV6 donor-mediated homologous recombination. *Cell Stem Cell* **24**, 821–828 (2019).
34. Goldberg, F. W. et al. The discovery of 7-methyl-2-[(7-methyl[1,2,4]triazolo[1,5-apyridin-6-yl)amino]-9-(tetrahydro-2*H*-pyran-4-yl)-7,9-dihydro-8*H*-purin-8-one (AZD7648), a potent and selective DNA-dependent protein kinase (DNA-PK) inhibitor. *J. Med. Chem.* **63**, 3461–3471 (2020).
35. Zenke, F. T. et al. Pharmacologic inhibitor of DNA-PK, M3814, potentiates radiotherapy and regresses human tumors in mouse models. *Mol. Cancer Ther.* **19**, 1091–1101 (2020).
36. Khan, A. J. et al. VX-984 is a selective inhibitor of non-homologous end joining, with possible preferential activity in transformed cells. *Oncotarget* **9**, 25833–25841 (2018).
37. Hardcastle, I. R. et al. Discovery of potent chromen-4-one inhibitors of the DNA-dependent protein kinase (DNA-PK) using a small-molecule library approach. *J. Med. Chem.* **48**, 7829–7846 (2005).
38. Berger, M. et al. BAY-8400: a novel potent and selective DNA-PK inhibitor which shows synergistic efficacy in combination with targeted alpha therapies. *J. Med. Chem.* **64**, 12723–12737 (2021).
39. Morrison, R., Al-Rawi, J. M., Jennings, I. G., Thompson, P. E. & Angove, M. J. Synthesis, structure elucidation, DNA-PK and PI3K and anti-cancer activity of 8- and 6-aryl-substituted-1-3-benzoxazines. *Eur. J. Med. Chem.* **110**, 326–339 (2016).
40. Liu, R. et al. Homozygous defect in HIV-1 coreceptor accounts for resistance of some multiply-exposed individuals to HIV-1 infection. *Cell* **86**, 367–377 (1996).
41. Papapetrou, E. P. & Schambach, A. Gene insertion into genomic safe harbors for human gene therapy. *Mol. Ther.* **24**, 678–684 (2016).
42. Gomez-Ospina, N. et al. Human genome-edited hematopoietic stem cells phenotypically correct mucopolysaccharidosis type I. *Nat. Commun.* **10**, 4045 (2019).
43. Scharenberg, S. G. et al. Engineering monocyte/macrophage-specific glucocerebrosidase expression in human hematopoietic stem cells using genome editing. *Nat. Commun.* **11**, 3327 (2020).
44. Brown, N., Song, L., Kollu, N. R. & Hirsch, M. L. Adeno-associated virus vectors and stem cells: friends or foes? *Hum. Gene Ther.* **28**, 450–463 (2017).
45. Tsuchiya, A., Kanno, T. & Nishizaki, T. PI3 kinase directly phosphorylates Akt1/2 at Ser 473/474 in the insulin signal transduction pathway. *J. Endocrinol.* **220**, 49–59 (2014).
46. Chen, J., Ghorai, M. K., Kenney, G. & Stubbe, J. Mechanistic studies on bleomycin-mediated DNA damage: multiple binding modes can result in double-stranded DNA cleavage. *Nucleic Acids Res.* **36**, 3781–3790 (2008).
47. Dever, D. P. et al. CRISPR/Cas9 β-globin gene targeting in human haematopoietic stem cells. *Nature* **539**, 384–389 (2016).
48. Lattanzi, A. et al. Development of β-globin gene correction in human hematopoietic stem cells as a potential durable treatment for sickle cell disease. *Sci. Transl. Med.* **13**, eabf2444 (2021).
49. Vaidyanathan, S. et al. High-efficiency, selection-free gene repair in airway stem cells from cystic fibrosis patients rescues CFTR function in differentiated epithelia. *Cell Stem Cell* **26**, 161–171 (2020).
50. Cromer, M. K. et al. Gene replacement of α-globin with β-globin restores hemoglobin balance in β-thalassemia-derived hematopoietic stem and progenitor cells. *Nat. Med.* **27**, 677–687 (2021).
51. Pavel-Dinu, M. et al. Gene correction for SCID-X1 in long-term hematopoietic stem cells. *Nat. Commun.* **10**, 1634 (2019).
52. Romero, Z. et al. Editing the sickle cell disease mutation in human hematopoietic stem cells: comparison of endonucleases and homologous donor templates. *Mol. Ther.* **27**, 1389–1406 (2019).
53. Pavani, G. et al. Ex vivo editing of human hematopoietic stem cells for erythroid expression of therapeutic proteins. *Nat. Commun.* **11**, 3778 (2020).
54. Dudek, A. M. & Porteus, M. H. Answered and unanswered questions in early-stage viral vector transduction biology and innate primary cell toxicity for ex-vivo gene editing. *Front. Immunol.* **12**, 660302 (2021).
55. Rose, J. C. et al. Suppression of unwanted CRISPR–Cas9 editing by co-administration of catalytically inactivating truncated guide RNAs. *Nat. Commun.* **11**, 2697 (2020).
56. Coelho, M. A. et al. CRISPR GUARD protects off-target sites from Cas9 nuclease activity using short guide RNAs. *Nat. Commun.* **11**, 4132 (2020).
57. DeWitt, M. A. et al. Selection-free genome editing of the sickle mutation in human adult hematopoietic stem/progenitor cells. *Sci. Transl. Med.* **8**, 360ra134 (2016).
58. Cavazzana, M., Antoniani, C. & Miccio, A. Gene therapy for β-hemoglobinopathies. *Mol. Ther.* **25**, 1142–1154 (2017).
59. Gilman, J. G. et al. Distal CCAAT box deletion in the A γ globin gene of two black adolescents with elevated fetal A γ globin. *Nucleic Acids Res.* **16**, 10635–10642 (1988).
60. Chen, F. et al. High-frequency genome editing using ssDNA oligonucleotides with zinc-finger nucleases. *Nat. Methods* **8**, 753–755 (2011).
61. Traxler, E. A. et al. A genome-editing strategy to treat β-hemoglobinopathies that recapitulates a mutation associated with a benign genetic condition. *Nat. Med.* **22**, 987–990 (2016).
62. Liu, Y. et al. Activated STING in a vascular and pulmonary syndrome. *N. Engl. J. Med.* **371**, 507–518 (2014).
63. Piras, F. et al. Lentiviral vectors escape innate sensing but trigger p53 in human hematopoietic stem and progenitor cells. *EMBO Mol. Med.* **9**, 1198–1211 (2017).
64. Ferrari, S. et al. Efficient gene editing of human long-term hematopoietic stem cells validated by clonal tracking. *Nat. Biotechnol.* **38**, 1298–1308 (2020).
65. Wiebking, V. et al. Genome editing of donor-derived T-cells to generate allogenic chimeric antigen receptor-modified T cells: optimizing αβ T cell-depleted haploidentical hematopoietic stem cell transplantation. *Haematologica* **106**, 847–858 (2021).
66. Mohiuddin, I. S. & Kang, M. H. DNA-PK as an emerging therapeutic target in cancer. *Front. Oncol.* **9**, 635 (2019).
67. Hu, S. et al. Small molecule DNA-PK inhibitors as potential cancer therapy: a patent review (2010–present). *Expert Opin. Ther. Pat.* **31**, 435–452 (2021).
68. Yue, X., Bai, C., Xie, D., Ma, T. & Zhou, P. K. DNA-PKcs: a multi-faceted player in DNA damage response. *Front. Genet.* **11**, 607428 (2020).
69. Vakulskas, C. A. et al. A high-fidelity Cas9 mutant delivered as a ribonucleoprotein complex enables efficient gene editing in human hematopoietic stem and progenitor cells. *Nat. Med.* **24**, 1216–1224 (2018).
70. Cromer, M. K. et al. Comparative analysis of CRISPR off-target discovery tools following ex vivo editing of $CD34^+$ hematopoietic stem and progenitor cells. *Mol. Ther.* **31**, 1074–1087 (2023).
71. Shi, Y., Inoue, H., Wu, J. C. & Yamanaka, S. Induced pluripotent stem cell technology: a decade of progress. *Nat. Rev. Drug Discov.* **16**, 115–130 (2017).
72. Yamanaka, S. Pluripotent stem cell-based cell therapy—promise and challenges. *Cell Stem Cell* **27**, 523–531 (2020).
73. Hockemeyer, D. & Jaenisch, R. Induced pluripotent stem cells meet genome editing. *Cell Stem Cell* **18**, 573–586 (2016).

74. Sweeney, C. L. et al. Correction of X-CGD patient HSPCs by targeted CYBB cDNA insertion using CRISPR/Cas9 with 53BP1 inhibition for enhanced homology-directed repair. *Gene Ther.* **28**, 373–390 (2021).
75. Genovese, P. et al. Targeted genome editing in human repopulating haematopoietic stem cells. *Nature* **510**, 235–240 (2014).
76. June, C. H., O'Connor, R. S., Kawalekar, O. U., Ghassemi, S. & Milone, M. C. CAR T cell immunotherapy for human cancer. *Science* **359**, 1361–1365 (2018).
77. Sather, B. D. et al. Efficient modification of *CCR5* in primary human hematopoietic cells using a megaTAL nuclease and AAV donor template. *Sci. Transl. Med.* **7**, 307ra156 (2015).
78. Eyquem, J. et al. Targeting a CAR to the *TRAC* locus with CRISPR/Cas9 enhances tumour rejection. *Nature* **543**, 113–117 (2017).
79. Berical, A., Lee, R. E., Randell, S. H. & Hawkins, F. Challenges facing airway epithelial cell-based therapy for cystic fibrosis. *Front. Pharmacol.* **10**, 74 (2019).
80. Vaidyanathan, S. et al. Targeted replacement of full-length CFTR in human airway stem cells by CRISPR–Cas9 for pan-mutation correction in the endogenous locus. *Mol. Ther.* **30**, 223–237 (2022).
81. Wilson, L. O. W., O'Brien, A. R. & Bauer, D. C. The current state and future of CRISPR–Cas9 gRNA design tools. *Front. Pharmacol.* **9**, 749 (2018).

Methods

Ethics statement

Umbilical cord blood-derived HSPCs were obtained from Binns Program for Cord Blood Research with approval from the Stanford Institutional Review Board Committee under protocol 33813. All donors provided informed consent, and participant information was deidentified. CF HBECs were obtained from lungs explanted during transplantation under protocol 13-1396 approved by the Committee on the Protection of the Rights of Human Subjects at the University of North Carolina at Chapel Hill. All donors provided informed consent.

DNA-PKcs inhibitors

Small-molecule DNA-PKcs inhibitors used in this study are all commercially available. AZD7648 (S8843) and M3184 (S8586) were from Selleck Chemicals. VX984 (HY-19939S), KU57788 (HY-11006), LTURM34 (HY-101667) and BAY8400 (HY-132293) were from MedChemExpress. All DNA-PKcs inhibitors were resuspended in DMSO to make either 2 mM or 4 mM stocks and were diluted in cell culture medium to make up the indicated final concentrations for gene targeting.

Cas9 and gRNA used for genome editing

High-fidelity Cas9 (ref. 69) purchased from Aldevron (SpyFi Cas9, 9214) was used for all genome editing experiments. gRNAs used for genome editing were purchased from either Synthego or TriLink Biotechnologies. gRNAs were chemically modified to include 2′-*O*-methyl-3′-phosphorothioate at the first and last three nucleotides[82]. Genomic target sites for the different gRNAs used are listed in Supplementary Table 1.

AAV6 vector construction, production and purification

For construction of the AAV transfer plasmid, pAAV-MCS plasmid (Agilent) backbone was used. pAAV-MCS plasmid was digested with NotI-HF enzyme (New England Biolabs), and sequences of the homology arms and insert were cloned into the backbone using NEBuilder HiFi DNA Assembly master mix (New England Biolabs, E2621L). After cloning, transfer plasmids were sequence verified and purified using a PureLink Expi Endotoxin-Free Maxi Plasmid Purification kit (Thermo Fisher Scientific, A31217). AAV6 vectors were either produced in-house or acquired through Vigene or Signagen. For in-house production, 293T cells were seeded in five to ten 150-mm dishes at 10 million cells per dish. After 24 h, each dish with 293T cells was transfected with 22 µg of packaging/helper plasmid, pDGM6 (a gift from D. Russell (University of Washington), Addgene plasmid 110660) and 6 µg of the transfer plasmid in 1 ml of OptiMEM I (Gibco, 31985088) using PEI (Polysciences, 23966-1). Seventy-two hours after transfection, AAV6 was purified using an AAVpro purification kit (Takara, 6666), as per the manufacturer's instructions. AAV6 titer was determined by ddPCR, as per the manufacturer's instructions, using the previously validated primer/probe set[83] listed in the Supplementary Table 1.

PSC culture and genome editing

Gene editing experiments were performed with three PSC lines: H9 embryonic stem cells (ESCs; WiCell) and the previously described 1205-4 and 1208-2 iPSC lines[84]. PSCs were maintained under feeder-free conditions on Matrigel-coated (Corning, 354277) plates in mTeSR1 medium (STEMCELL Technologies, 85850). For gene editing, 24 h before nucleofection, PSCs were pretreated with 10 µM Y27632 (Cayman Chemical, 10005583). For each nucleofection, RNP complex was prepared with 5 µg of Cas9 protein and 2 µg of gRNA and incubated for 15 min at room temperature. PSCs were dissociated into single cells using Accutase (Innovative Cell Technologies, AT104). Five hundred thousand cells were resuspended with the RNP complex diluted in 20 µl of P3 primary cell nucleofector solution (Lonza, V4XP-3032). Resuspended cells were added to 1 well of a 16-well Nucleocuvette Strip (Lonza), and nucleofection was performed in a 4D-Nucleofector (Lonza) using the program CA137. Nucleofected PSCs were plated in mTeSR1 medium supplemented with 10 µM Y27632 at a density of 100,000 cells per well in 48-well plates. AAV6 vector and DNA-PKcs inhibitors were added to the medium at the indicated concentrations. After 24 h of incubation, the existing medium was removed, and fresh mTeSR1 supplemented with 10 µM Y27632 was added to the cells. From the following day, PSCs were maintained in mTeSR1 without Y27632 until passaging. Gene targeting efficiency was analyzed 4–6 d after editing.

For single-cell cloning, gene-targeted PSCs were plated at a density of 250 cells per well of a six-well plate in mTeSR1 medium supplemented with 1× CloneR2 (STEMCELL Technologies, 100-0691) and incubated at 37 °C for 2 d. After 2 d, medium was switched to fresh mTeSR1 medium supplemented with 1× CloneR2. Two days later, the cells were switched to and maintained in mTeSR1 medium with daily medium changes. At days 7–10, single-cell colonies were picked by scraping and propagated individually. Gene targeting was assessed in single-cell clones by ddPCR, as mentioned above, and by PCR amplification of the region spanning the knock-in to determine the frequency of mono- and biallelic and non-targeted clones.

Cell viability measurement

An MTT assay was used for measuring gene-edited PSC viability. Gene-edited PSCs were plated in 96-well plates, and cell viability was assessed at 24, 48 and 72 h after editing. For this, 0.5 mg ml^{-1} MTT (MedChemExpress, HY-15924) diluted in growth medium was added to the cells, and the plates were incubated at 37 °C for 2 h. After this incubation, MTT was removed, and cells were lysed using 100 µl of lysis buffer (0.1 N HCl and 0.5% SDS in isopropanol). After lysis, absorbance was measured at 570 nm with 650 nm as the reference using a SpectraMax M3 plate reader (Molecular Devices). Absorbance values were used to calculate cell viability of gene-edited cells as a percentage relative to mock-treated cells. For measuring the viability of HSPCs, T cells and HBECs, cell suspensions were mixed with Trypan Blue (Gibco, 15250061), and viable cell counts were determined using a TC10 cell counter (Bio-Rad) at 3–5 d after gene editing. Viable cell counts are represented as percent relative to mock or untreated cells as indicated.

Pluripotency marker analysis

Mock, RNP-only and gene-targeted PSCs were assessed for the expression of the pluripotency markers SSEA4, OCT3/OCT4, SOX2 and NANOG. For SSEA4 expression analysis, dissociated PSCs were stained with anti-SSEA4 (BioLegend, 330407), and flow cytometry was used to measure the percentage of $SSEA4^+$ cells. OCT3/OCT4, SOX2 and NANOG expression was assessed following immunofluorescence staining of fixed PSCs. For this, PSCs were fixed with 4% paraformaldehyde (Electron Microscopy Sciences, 15710) for 20 min at room temperature. Cells were then permeabilized with 0.3% Triton X-100 (Thermo Fisher Scientific, 85111) for 20 min at room temperature and incubated with a blocking solution of 3% bovine serum albumin (BSA; Gold Biotechnology, A-420-10) in PBS for 1 h at room temperature. After blocking, cells were incubated with primary antibody diluted 1:200 in 3% BSA and incubated at 4 °C overnight. The following day, cells were washed once with PBS and incubated for 45 min at room temperature with secondary antibody and DAPI (Santa Cruz Biotechnology, sc-3598) diluted at 1:500 and 1:5,000, respectively, in 3% BSA. Primary and secondary antibodies used are listed in Supplementary Table 1. After this incubation, cells were washed once with PBS and stored in fresh PBS at 4 °C until imaging. Stained cells were imaged using a BZ-X710 microscope (Keyence), and images were analyzed using ImageJ software (v1.53, NIH). Areas of Alexa Fluor 647 and DAPI staining were measured following manual thresholding, and the ratio of these areas was used to determine the percentage of cells positive for the corresponding markers. The percentage values were capped at 100% to account for staining artifacts.

Trilineage differentiation of PSCs and analysis

Gene-targeted PSCs (H9 ESCs and 1205-4 iPSCs) were assessed for three germ layer differentiation potential using a STEMdiff Trilineage Differentiation kit (STEMCELL Technologies, 05230), as per the manufacturer's instructions. Following differentiation, cells were dissociated with Accutase and stained with antibodies to respective extracellular markers. Cells were then fixed and permeabilized using an eBioscience Foxp3/Transcription Factor Staining Buffer Set (Thermo Fisher Scientific, 00-5523-00) and stained with antibodies (listed in Supplementary Table 1) to intracellular markers following the manufacturer's instructions. Flow cytometry analysis was used to measure the percentage of cells double positive for the corresponding markers of the three germ layers.

ddPCR analysis of targeted integration

Gene targeting efficiency for the knock-in of multikilobase sequences was measured by ddPCR analysis. Genomic DNA was extracted from cells using QuickExtract DNA Extraction Solution (QE; Lucigen, QE09050). Pellets of 100,000 to 200,000 cells were resuspended in 40 to 50 µl of QE solution. This cell suspension was incubated at 65 °C for 6 min and then at 100 °C for 10 min. The quick-extracted genomic DNA was digested with either BamHI-HF or HindIII-HF (New England Biolabs) at 37 °C for 1 to 2 h. One to two microliters of digested genomic DNA was used as a template for the ddPCR reaction with 10 µl of ddPCR supermix for probes (no dUTP; Bio-Rad, 1863025), 0.5 µl each of fluorophore-labeled target and reference primer/probe assays in a total volume of 20 µl. ddPCR droplets were generated in the droplet generator (Bio-Rad) with 70 µl of droplet generation oil (Bio-Rad, 1863005) and 20 µl of sample, as per the manufacturer's instructions. Forty microliters of the generated droplet sample was used for PCR with the following cycle: 95 °C for 10 min; 50 cycles of 94 °C for 30 s, 60 °C for 30 s and 72 °C for 2 min and 98 °C for 10 min. After PCR, droplets were analyzed using a QX200 Droplet Digital PCR reader (Bio-Rad), and Quantasoft software (v1.7.4.0917) was used to determine the number of copies of the target and reference DNA in the droplets. Allelic gene targeting efficiency was calculated as the percentage of target DNA copies relative to that of the reference. Target primer/probe assays were designed such that one primer anneals to the insert, and the other primer anneals outside the homology arm (In-Out PCR). Sequences of the different primer/probe assays used are listed in Supplementary Table 1.

ICE analysis for quantification of gene editing

The ICE CRISPR analysis tool (Synthego) was used to determine the allelic distribution of WT, INDEL and HDR frequencies for gene targeting involving knock-in of short sequences and to calculate the frequency of WT and INDEL alleles for gene targeting involving the knock-in of multikilobase sequences. Quick-extracted genomic DNA from mock and gene-edited samples were used as a template for PCR to amplify the gene-edited region using PrimeSTAR GXL DNA Polymerase (Takara, R050A). Following PCR, amplicons were detected on agarose gels and purified from the gel using a GeneJet Gel Extraction kit (Thermo Fisher Scientific, K0692). Purified DNA was Sanger sequenced through MCLAB or GENEWIZ. Sequencing chromatograms were analyzed through the ICE tool using the mock sample sequence as the reference following the manufacturer's instructions. Sequences of the primers used for PCR and sequencing are listed in Supplementary Table 1.

Western blotting

Cells were lysed with RIPA buffer (Thermo Fisher Scientific, PI89900) supplemented with Halt Protease and Phosphatase Inhibitor Cocktail (Thermo Fisher Scientific, PI78440) after incubation for 15 min at 4 °C. Protein concentration was determined using a BCA assay kit (Thermo Fisher Scientific, PI23227) following the manufacturer's instructions. Cell lysates were denatured by adding 1× Laemmli sample buffer (Bio-Rad, 1610747) and heating at 100 °C for 5 min. Total protein (20 µg) was loaded onto 4–15% Mini-PROTEAN TGX precast protein gels (Bio-Rad, 4561084), and electrophoresis was performed. Following electrophoresis, protein was transferred to a PVDF membrane (Bio-Rad, 1620177). The membrane was blocked in 5% dry milk in Tris-buffered saline with 0.1% Tween 20 (TBST) for 1 h at room temperature and incubated with primary antibody diluted in 5% BSA in TBST at 4 °C overnight. The following day, membranes were washed three times with TBST for 5 min per wash and incubated with horseradish peroxidase-conjugated secondary antibody diluted in 5% dry milk for 45 min at room temperature. After secondary antibody incubation, membranes were washed three times with TBST for 5 min per wash. For detection, membranes were incubated in SuperSignal West Femto Maximum Sensitivity Substrate (Thermo Fisher Scientific, 34095) for 5 min, and blots were imaged on a ChemiDoc imager using Image Lab software. Primary and secondary antibodies used are listed in Supplementary Table 1.

$CD34^+$ HSPC culture and genome editing

Human $CD34^+$ HSPCs were isolated from cord blood (provided by the Binns Program for Cord Blood Research). HSPCs were cultured in either StemSpan SFEM II (STEMCELL Technologies, 09655) or SCGM (Cellgenix, 20802) medium supplemented with 100 ng ml^{-1} stem cell factor (PeproTech, 300-07), 100 ng ml^{-1} thrombopoietin (PeproTech, 300-18), 100 ng ml^{-1} FLT3-ligand (PeproTech, 300-19), 100 ng ml^{-1} interleukin-6 (IL-6; PeproTech, 200-06), 20 U ml^{-1} penicillin, 20 mg ml^{-1} streptomycin (Cytiva, SV30010) and 35 nM UM171 (APExBIO, A89505). HSPCs were cultured at a cell density of 0.25 million–0.5 million cells per ml of growth medium at 37 °C, 5% O_2 and 5% CO_2. For gene editing of HSPCs, RNP complex was prepared with 6 µg of Cas9 and 3.2 µg of gRNA, incubated for 15 min at room temperature and diluted in 20 µl of P3 primary cell nucleofector solution (Lonza, V4XP-3032). Cells (0.5 million–1 million) were resuspended in 20 µl of the RNP nucleofection solution, and nucleofection was performed in the 4D-Nucleofector (Lonza) using the program DZ-100. Nucleofected HSPCs were plated at the cell density mentioned above with the indicated amounts of AAV6 and AZD7648. After a 24-h incubation, cells were switched to fresh growth medium without AAV6 and AZD7648. Gene targeting efficiency was assessed at 3 d after gene editing. For ssODN donor-based gene targeting experiments, 200-bp-long ssODN Alt-R HDR Donor oligonucleotides (IDT) were used. ssODN was delivered into the cells through nucleofection. ssODN was added to the nucleofection solution with RNP at a concentration of 2.5 or 5 µM, and HSPCs were nucleofected and cultured as described earlier.

CFU assay and genotyping of clones

For CFU assays, 250 to 500 HSPCs were plated in SmartDish six-well plates (STEMCELL Technologies, 27370) containing MethoCult H4434 Classic (STEMCELL Technologies, 04444). After 14 d, colonies were counted and scored using a STEMvision automated counter (STEMCELL Technologies) to determine the number of BFU-E, CFU-M, CFU-GM and CFU-GEMM colonies. To determine the ratio of monoallelic to biallelic integration, individual colonies of each type were picked, and genomic DNA was extracted with QE. ddPCR was used to determine the clones with mono- versus biallelic integration, as described earlier. WT and INDEL frequencies in the clones were determined by ICE analysis.

FACS of LT-HSC and MPP populations

Cord blood-derived human $CD34^+$ HSPCs were gene targeted at the *CCR5* locus for knock-in of the UBC-GFP-bGH-pA sequence with and without AZD7648 treatment. Two days after gene targeting, HSPCs were stained with a cocktail of antibodies listed in Supplementary Table 1 after blocking with Human TruStain FcX (BioLegend, 422301).

Cell viability was assessed using a LIVE/DEAD Fixable Near-IR Dead Cell Stain kit (Thermo Fisher Scientific, L10119). LT-HSC and MPP cells were isolated by FACS for $CD34^+CD90^+Lin^-CD45RA^-CD38^-$ (LT-HSCs) and $CD34^+CD90^-Lin^-CD45RA^-CD38^-$ (MPP cells) in a FACS Aria II sorter, part of the FACS Core Facility in the Stanford Institute for Stem Cell Biology and Regenerative Medicine. Allelic gene targeting efficiency in LT-HSC and MPP populations was measured by ddPCR analysis, as described earlier. FACS data analysis was performed using FlowJo (10.8.1 for Mac) software.

NGS analysis for off-target effects

For NGS analysis, we used the commercially available Amplicon-EZ NGS service from Azenta. Previously characterized top off-target sites for *CCR5* (OT39), *HBB* (OT1) and *HBA1* (OT1) gRNAs were analyzed. PCR was used to amplify a 350- to 450-bp region encompassing the off-target sites from genomic DNA of gene-targeted HSPCs. Sequences of the primers used for PCR and the Illumina adapters added to the primers are listed in Supplementary Table 1.

Following PCR, the amplicons were purified using a Genejet PCR purification kit (Thermo Fisher Scientific, K0701). Samples were normalized to a concentration of 20 ng μl^{-1}, and a 25-µl volume was submitted to the Amplicon-EZ NGS service (Azenta). DNA library preparations, sequencing reactions and bioinformatics analyses were conducted at Azenta. DNA library preparation was performed using an NEBNext Ultra DNA Library Prep kit following the manufacturer's recommendations (New England Biolabs). DNA amplicons were indexed and enriched by limited cycle PCR. DNA libraries were validated on an Agilent TapeStation (Agilent Technologies), quantified using a Qubit 2.0 fluorometer (Invitrogen) and multiplexed in equal molar mass. Pooled DNA libraries were loaded on an Illumina instrument according to the manufacturer's instructions. Samples were sequenced using a 2 × 250-bp paired-end configuration. Image analysis and base calling were conducted with Illumina Control Software on the Illumina instrument. Raw Illumina reads were checked for adapters and quality via FastQC. The raw Illumina sequence reads were trimmed of their adapters and nucleotides with poor quality using Trimmomatic v.0.36. Paired sequence reads were then merged to form a single sequence if the forward and reverse reads were able to overlap. The merged reads were aligned to the reference sequence, and variant detection was performed using the Azenta proprietary Amplicon-EZ program. INDEL frequency at the target site was quantified and plotted as the frequency of reads with insertions and deletions.

AAV6 genome quantification

AAV6 genomes in cells after gene editing were quantified by ddPCR. Gene-targeted HSPCs were collected at days 2, 4 and 6 after gene editing, along with non-gene-edited controls. ddPCR analysis was performed as mentioned above on genomic DNA to measure the AAV genomes, and the values were normalized to that of the reference locus *ZEB2*. Sequences of primers/probes used for ddPCR are listed in Supplementary Table 1.

Quantification of *CDKN1A* expression levels

CDKN1A expression levels were quantified by reverse transcription-ddPCR analysis. Total RNA was extracted from cell pellets using an RNeasy Plus Micro kit following the manufacturer's recommendations (Qiagen, 74034). cDNA was synthesized from total RNA using iScript RT Supermix following the manufacturer's instructions (Bio-Rad, 1708841). Synthesized cDNA was used to measure *CDKN1A* mRNA levels normalized to the *TBP* housekeeping control gene by ddPCR. PrimeTime predesigned quantitative PCR assays (IDT) were used for ddPCR amplification of *CDKN1A* (assay ID Hs.PT.58.40874346.g) and *TBP* (assay ID Hs.PT.58v.39858774). The following ddPCR cycling conditions were used: 95 °C (10 min); 45 cycles (2 °C ramp rate) of 94 °C (30 s), 60 °C (1 min) and then 98 °C (10 min). Normalized *CDKN1A* expression levels in gene-edited cells were plotted relative to those of the non-gene-edited controls.

T cell culture and genome editing

Leukocyte reduction system chambers from healthy donors (Stanford Blood Center) were used for the isolation of T cells. Peripheral blood mononuclear cells isolated on a Ficoll density gradient were used for obtaining T cells using a $CD4^+$ T Cell Isolation kit (Miltenyi, 130-096-533). For CD19-CAR targeting experiments, isolated αβ T cells were used. T cells were cultured in X-VIVO 15 medium (Lonza, 04-418Q) supplemented with 5% human AB serum (Sigma, H3667) and 100 IU ml^{-1} recombinant human IL-2 (PeproTech, 200-02) at 37 °C, ambient O_2 and 5% CO_2. Medium changes were performed every 2–4 d, and cultured cells were maintained at a target density of 0.5 million–1 million cells per ml unless otherwise indicated. T cells were activated with Dynabeads Human T Cell Activator (Gibco, 11161D) for 72–96 h, and beads were removed before nucleofection. For gene targeting, gRNA was complexed with Cas9 at a molar ratio of 2.5:1 (gRNA:protein) and nucleofected in P3 primary cell nucleofector solution (Lonza) into activated T cells using a 4D-Nucleofector (Lonza) in 16-well cuvette strips. One million activated T cells were used per nucleofection using the program EO-115. Cells were resuspended directly after nucleofection in 80 µl of complete T cell medium and diluted to the target density[65]. For gene targeting, cells were incubated within 15 min after electroporation with AAV6 for transduction at the indicated MOI with or without AZD7648. After 3–4 h, the suspension was diluted with complete medium to reach the target cell concentration of 1 million cells per ml. After a 24-h incubation, cells were switched to fresh medium. Gene targeting was analyzed 3–4 d after gene editing.

CAR T cell cytotoxicity assay

CD19-directed CAR T cells generated in the presence or absence of AZD7648 were challenged with target CD19-expressing GFP^+ Nalm6 leukemia cells. Effector and target cells were cocultured at a ratio of 1:1 in RPMI medium supplemented with 10% bovine growth serum. The cytotoxic effect of the CAR T cells or the depletion of target leukemia cells was monitored daily over 2 d by measuring the levels of GFP^+ cells by flow cytometry using a Beckman Coulter Accuri (Accuri C6 software) or CytoFLEX flow cytometer.

B cell culture and genome editing

Primary human B cells were isolated from leukocyte reduction system chambers obtained from the Stanford Blood Center via negative selection using a human B Cell Isolation kit II (Miltenyi Biotec, 130-091-151) according to manufacturer's instructions. Cells were cultured in Iscove's modified Dulbecco's medium (IMDM; Thermo Fisher Scientific, 12440053) supplemented with 10% bovine growth serum (Hyclone, SH30541.03HI), 1% penicillin–streptomycin (Cytiva, SV30010), 55 µM 2-mercaptoethanol (Sigma-Aldrich, M3148), 50 ng ml^{-1} IL-2 (Peprotech, 200-02), 50 ng ml^{-1} IL-10 (Peprotech, 200-10), 10 ng ml^{-1} IL-15 (Peprotech, 200-15), 100 ng ml^{-1} recombinant human MEGACD40L (Enzo Life Sciences, ALX-522-110-C010) and 1 µg ml^{-1} CpG oligonucleotide 2006 (Invivogen, tlrl-2006-1) at a density of 1 million cells per ml (ref. 85). B cells were cultured at 37 °C, 5% CO_2 and ambient oxygen levels.

For genome editing, gRNA targeting *CCR5* was complexed with Cas9 protein at a 2.5:1 (Cas9:gRNA) molar ratio at room temperature for 20 min. B cells were nucleofected 4–5 d after thawing using a Lonza 4D-Nucleofector (program EO-117) using 1 million cells per condition. Immediately following nucleofection, cells were incubated with AAV6 donor vector (UBC-GFP) at an MOI of 25,000 and different concentrations of AZD7648 in 100 µl of basal IMDM in a 96-well plate for 3–4 h (ref. 86). Cells were then replated at 1 million cells per ml in complete B cell activation medium. Approximately 24 h after nucleofection, cells

were replated in fresh medium to remove AZD7648. Gene targeting was assessed at 3 d after editing.

Cell culture and genome editing of HBECs

HBECs were obtained from Lonza, the Primary Airway Cell Biobank at McGill University, the Cystic Fibrosis Foundation Cell Bank or the University of North Carolina. HBECs were cultured in Pneumacult Ex-Plus medium (STEMCELL Technologies, 05040) at 3,000–10,000 cells per cm^2 in tissue culture flasks. Cells were cultured at 37 °C, 5% O_2 and 5% CO_2. Medium was supplemented with 10 µM ROCK inhibitor (Santa Cruz Biotechnology, Y27632, sc-281642A).

For genome editing, HBECs were resuspended at a density of 5 million cells per ml in OptiMEM I. Nucleofection was performed using Lonza 4D 16-well Nucleocuvette Strips (Lonza, V4XP-3032). Six micrograms of high-fidelity Cas9 and 3.2 µg of gRNA (molar ratio of 1:2.5) were complexed at room temperature for 10 min and mixed with 20 µl of cell suspension in OptiMEM I. Cells were electroporated using the program CA137. AAV6 was added at an MOI of 100,000, and AZD7648 was added at the indicated concentrations.

Statistical analysis

Microsoft Excel (v16.7 for Mac) was used for processing raw data. GraphPad Prism 9 software was used for all statistical analysis.

Reporting summary

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

Data availability

NGS data for off-target analysis have been deposited in the National Center for Biotechnology Information's Sequence Read Archive database (accession number PRJNA982854)[87] and can be accessed at http://www.ncbi.nlm.nih.gov/bioproject/982854. Sequences of the gRNA, PCR primers, ddPCR primers and probes have been included in Supplementary Table 1. Raw data for ICE analysis and screenshots of Sanger sequencing chromatograms are provided in the Supplementary Information. Source data are provided with this paper.

References

82. Hendel, A. et al. Chemically modified guide RNAs enhance CRISPR–Cas genome editing in human primary cells. *Nat. Biotechnol.* **33**, 985–989 (2015).
83. Aurnhammer, C. et al. Universal real-time PCR for the detection and quantification of adeno-associated virus serotype 2-derived inverted terminal repeat sequences. *Hum. Gene Ther. Methods* **23**, 18–28 (2012).
84. Khan, T. A. et al. Neuronal defects in a human cellular model of 22q11.2 deletion syndrome. *Nat. Med.* **26**, 1888–1898 (2020).
85. Hung, K. L. et al. Engineering protein-secreting plasma cells by homology-directed repair in primary human B cells. *Mol. Ther.* **26**, 456–467 (2018).
86. Rogers, G. L. et al. Optimization of AAV6 transduction enhances site-specific genome editing of primary human lymphocytes. *Mol. Ther. Methods Clin. Dev.* **23**, 198–209 (2021).
87. Selvaraj, S. et al. High-efficiency transgene integration by homology-directed repair in human primary cells using DNA-PKcs inhibition. *Sequence Read Archive* http://www.ncbi.nlm.nih.gov/bioproject/982854 (2023).

Acknowledgements

We thank the Binns Program for Cord Blood Research for providing cord blood $CD34^+$ HSPCs. We thank the FACS Core Facility at the Stanford Institute of Stem Cell Biology and Regenerative Medicine for access to the FACS machines. We would like to thank the Taube and Koret Foundation for funding support. We thank F. Suchy and the laboratory of H. Nakauchi at Stanford University for help with ddPCR primer/probe design.

Author contributions

S.S., W.N.F., S. Viel, S. Vaidyanathan, A.M.D., M.G., S.J.R., F.K.E., A.R.O., L.X., M.P.-D. and S.E.L. designed and performed experiments and analyzed data. M.K.C. provided reagents for experiments. R.S. performed experiments and analyzed data. N.G.-O. and M.H.P. supervised the project. S.S. and M.H.P wrote the initial and final drafts of the manuscript with input from other authors.

Competing interests

M.H.P. is a member of the scientific advisory board of Allogene Therapeutics and a member of the board of directors of Graphite Bio. S.S. and M.H.P. are inventors on intellectual property related to this work.

Additional information

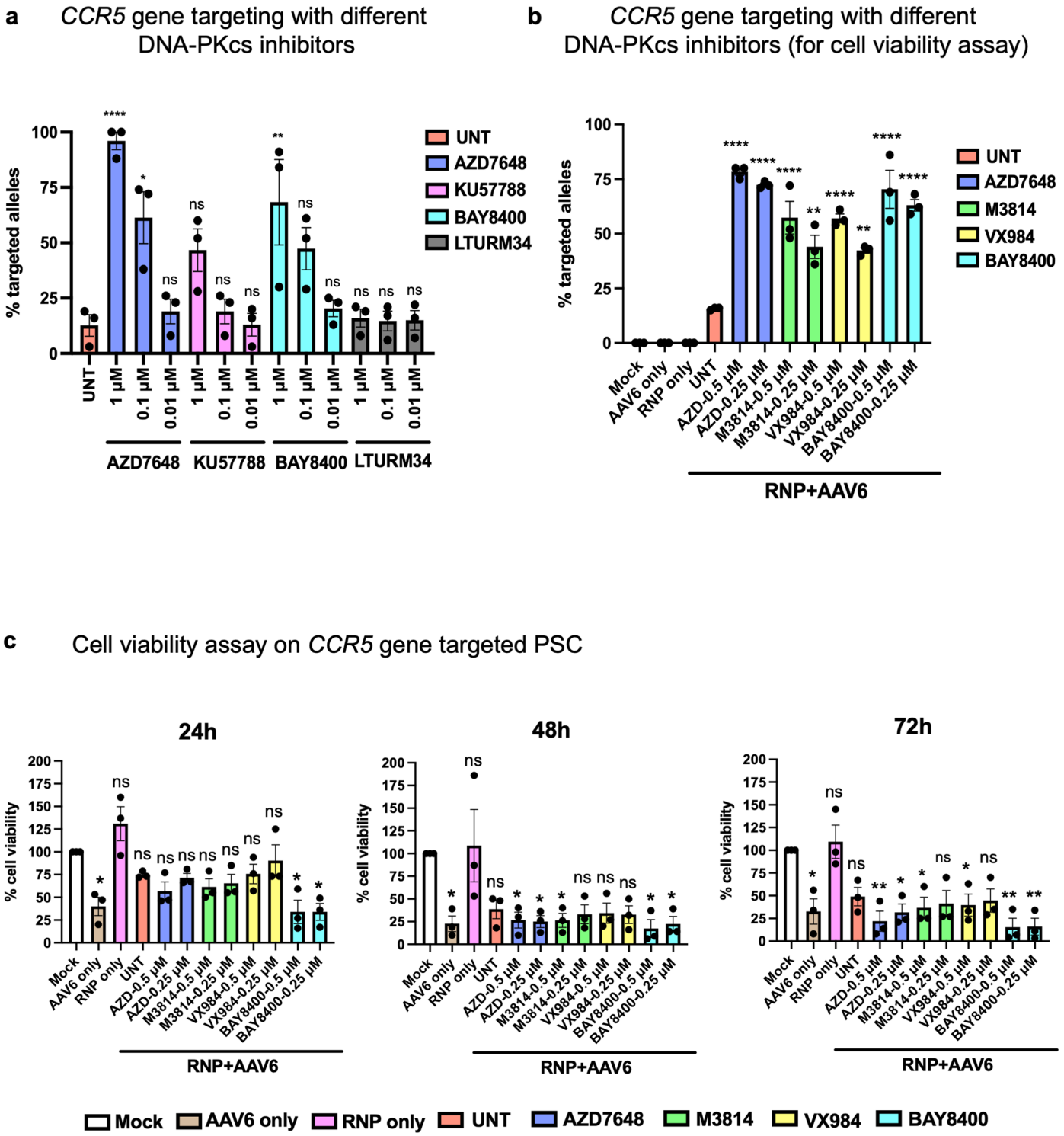
Extended data is available for this paper at https://doi.org/10.1038/s41587-023-01888-4.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41587-023-01888-4.

Correspondence and requests for materials should be addressed to Matthew H. Porteus.

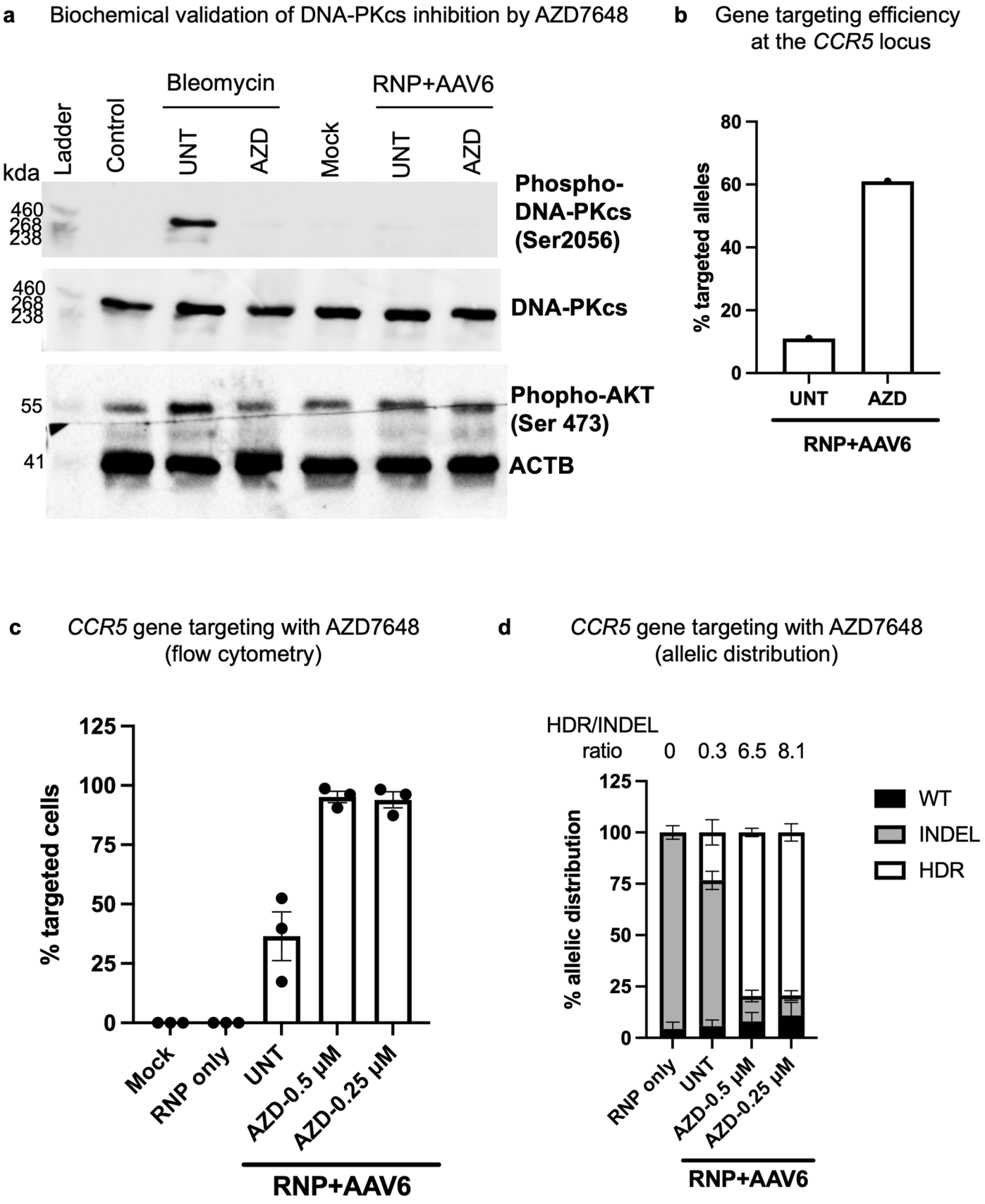
Peer review information *Nature Biotechnology* thanks the anonymous reviewers for their contribution to the peer review of this work.

https://doi.org/10.1038/s41587-023-01888-4



Extended Data Fig. 1 | Comparison of different DNA-PKcs inhibitors for gene targeting with AZD7648. a. Allelic gene targeting efficiency post gene editing at the *CCR5* locus for introduction of two stop codons with different concentrations of DNA-PKcs inhibitors (AZD7648, KU57788, LTURM34 and BAY8400) in comparison with the untreated cells (UNT) as measured by ICE analysis at 4 days post-gene editing (n = 3). **b**. Allelic gene targeting efficiency at the *CCR5* locus for knock-in of UBC-GFP-bGHpA sequence with two different concentrations (0.5 and 0.25 µM) of various DNA-PKcs inhibitors (AZD7648, M3814, VX984 and BAY8400) as measured by ddPCR analysis at 3 days post-gene editing (n = 3). **c**. MTT cell viability assay on gene targeted PSCs (**b**) at 24 h, 48 h and 72 h post-gene editing represented as percent cell viability normalized to the mock cells (n = 3). All data in **a, b** and **c** are shown as mean ± SEM. Data in **a, b** and **c** were compared with one-way ANOVA and Tukey's multiple comparisons test, *P < 0.05, **P < 0.01, ***P < 0.001, ****P < 0.0001 and ns denotes not significant.

https://doi.org/10.1038/s41587-023-01888-4

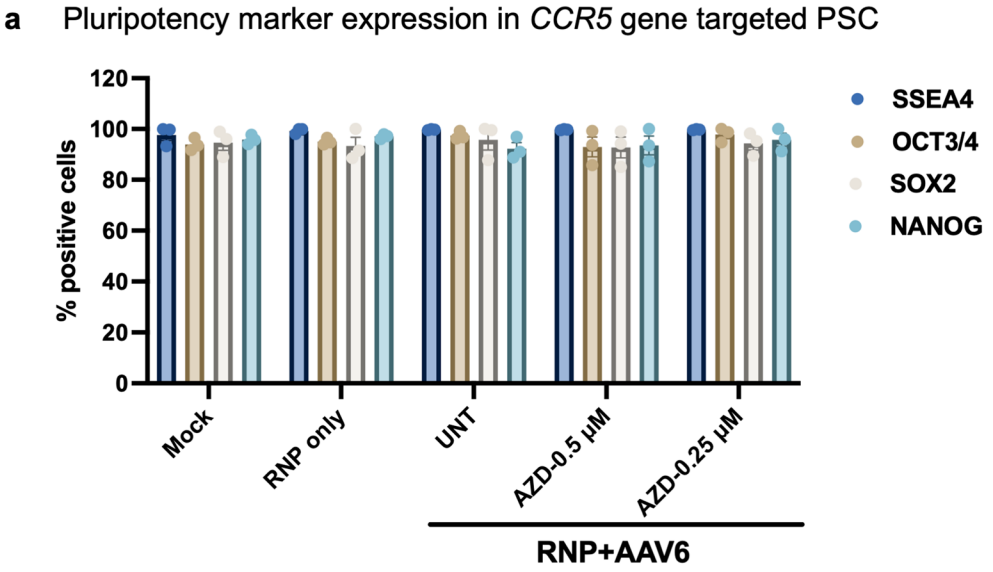


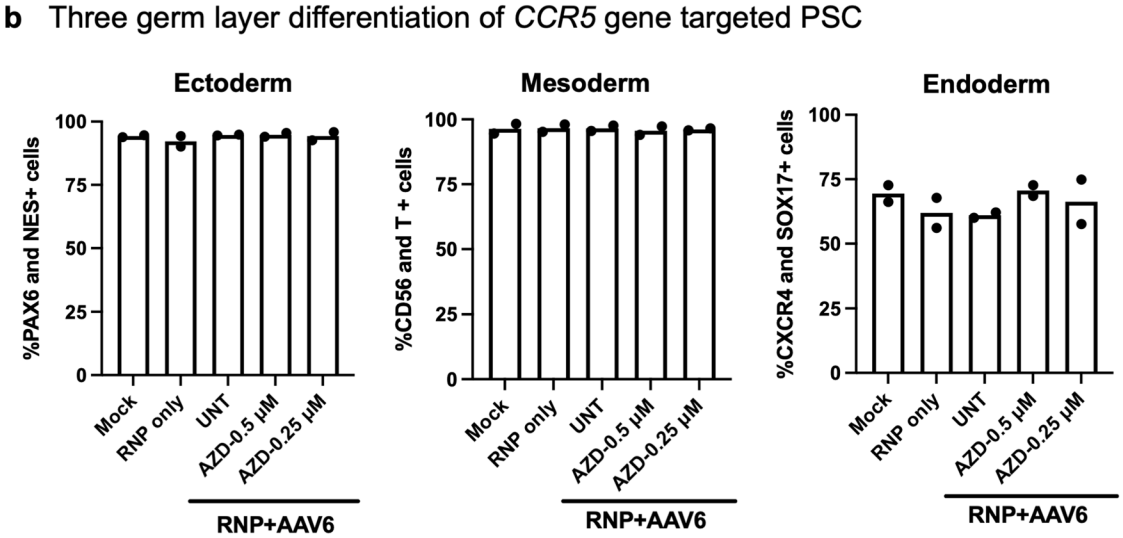
Extended Data Fig. 2 | See next page for caption.

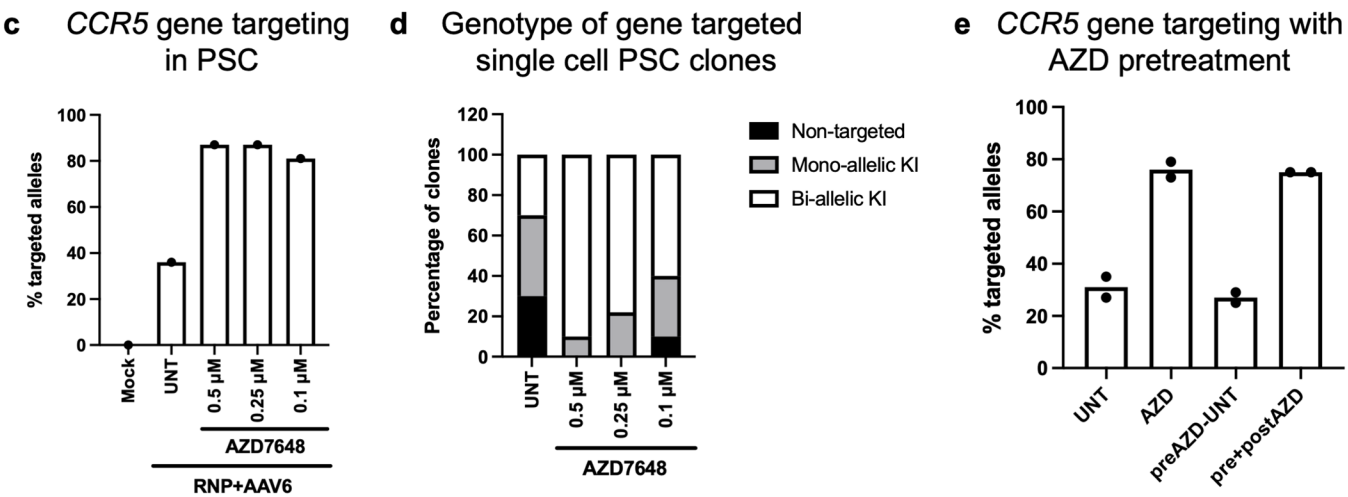
https://doi.org/10.1038/s41587-023-01888-4

Extended Data Fig. 2 | Biochemical validation of DNA-PKcs inhibition and *CCR5* gene targeting with AZD7648 in PSC. a. Western blot analysis for phospho DNA-PKcs (pSer2056), DNA-PKcs and phospho AKT (pSer473) in PSCs treated with bleomycin ± AZD for 2 h or gene targeted with and without AZD treatment at 2 h post gene editing. ACTB is used as a loading control. Control denotes sample not treated with bleomycin. Mock denotes control sample nucleofected without RNP. UNT denotes sample either treated with bleomycin or RNP-AAV6 gene editing and AZD denotes AZD7648 treated sample with either bleomycin or RNP-AAV6 gene editing treatment as indicated. **b**. Allelic gene targeting efficiency of PSCs gene targeted at the *CCR5* locus for knock-in of UBC-GFP-bGHpA sequence (**a)** with and without AZD treatment at 72 h post gene editing as measured by ddPCR (n = 1). **c**. Percentage of gene targeted cells at the *CCR5* locus for knock-in of UBC-GFP-bGHpA sequence with two different concentrations of AZD7648 (0.5 and 0.25 μM) in comparison with the untreated cells (UNT) as measured by flow cytometry for GFP at 5-days post-gene editing (n = 3). Mock and RNP only cells were used as negative controls. **d**. Allelic distribution of WT, INDEL, HDR frequencies in *CCR5* gene edited PSCs (**c**) (n = 3). HDR frequency was measured by ddPCR analysis, WT and INDEL frequencies were measured using ICE analysis. Mean HDR to INDEL ratio is represented above the bars. All data in **c** and **d** are shown as mean ± SEM.

https://doi.org/10.1038/s41587-023-01888-4





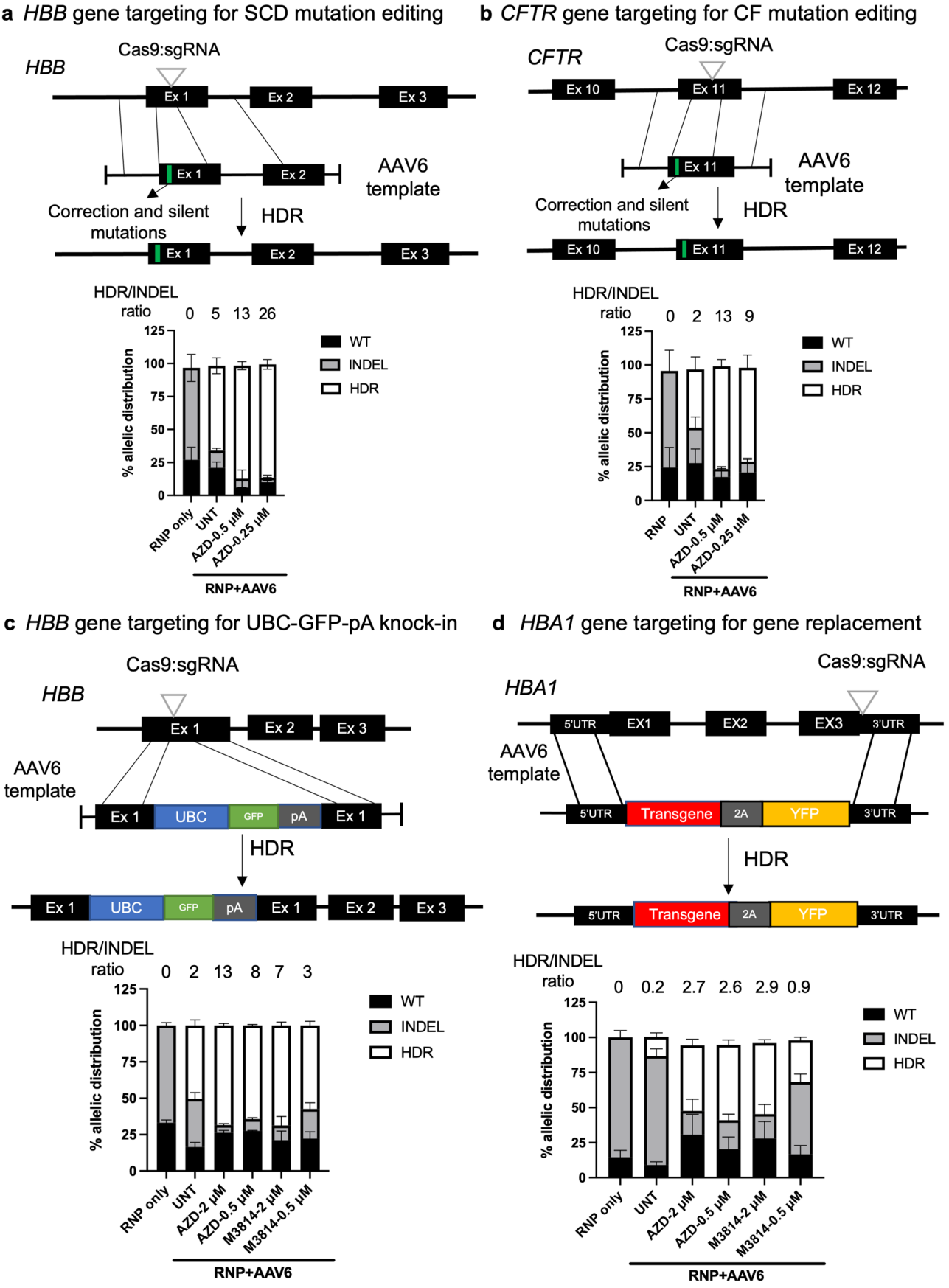


Extended Data Fig. 3 | See next page for caption.

https://doi.org/10.1038/s41587-023-01888-4

Extended Data Fig. 3 | Pluripotency, trilineage differentiation and single cell cloning analysis of *CCR5* gene targeted PSCs. a. Percentage of cells positive for various markers of pluripotency (SSEA4, OCT3/4, SOX2 and NANOG) in the *CCR5* gene edited PSCs (n = 3). SSEA4 expression was measured by flow cytometry analysis. OCT3/4, SOX2 and NANOG expression was assessed by quantification of immunofluorescence staining images for corresponding markers in fixed PSCs and normalized to the total cell count measured through DAPI staining (n = 3). All data are shown as mean ± SEM. **b**. *CCR5* gene edited PSCs were differentiated into the three germ layers. Mean frequency of differentiated cells as assessed by flow cytometry for the expression of corresponding markers for ectoderm (PAX6 and NES), mesoderm (CD56 and T) and endoderm (CXCR4 and SOX17) (n = 2). **c**. PSCs were gene targeted at the *CCR5* locus with and without different concentrations of AZD7648 (0.5, 0.25 and 0.1 μM) for single cell cloning analysis. Allelic gene targeting efficiency was measured using ddPCR analysis (n = 1). **d**. Gene targeted PSCs (**c**) were subjected to single cell cloning and the frequency of clones with mono-, bi-allelic and no gene targeting was measured using ddPCR and PCR analysis. For each condition, 9-10 clones were picked and analyzed (n = 1). **e**. Mean allelic gene targeting efficiency in PSC at the *CCR5* locus for the knock-in of UBC-GFP-bGHpA sequence with pre-treatment only (preAZD-UNT), post-treatment (AZD) only and pre+post treatment (pre+post AZD) with AZD7648 (0.5 μM) (n = 2). For pre-treatment, cells were treated with AZD7648 for 24 hours before gene targeting and for post-treatment, cells were treated with AZD7648 for 24 hours post gene editing.

https://doi.org/10.1038/s41587-023-01888-4

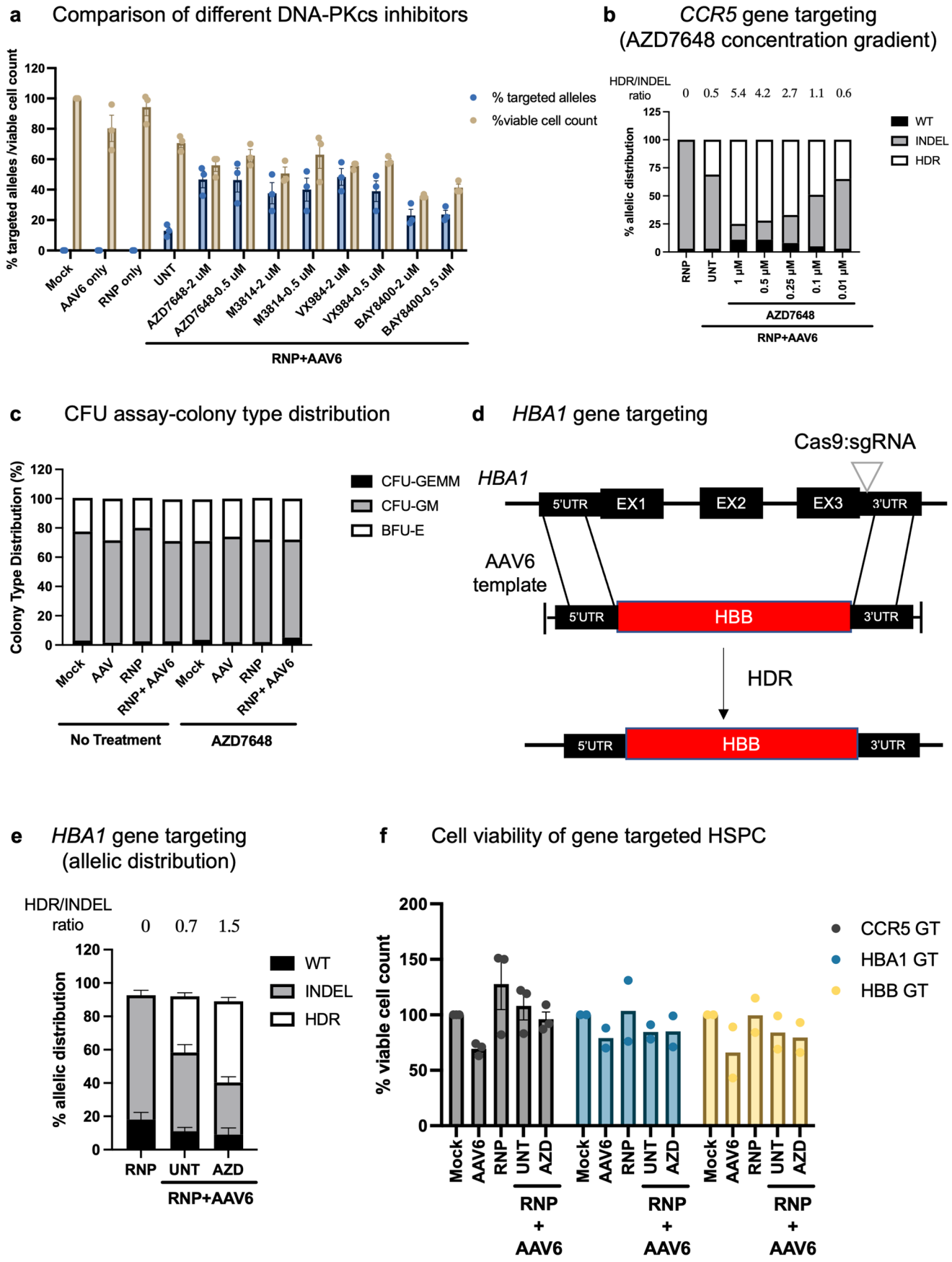


Extended Data Fig. 4 | See next page for caption.

https://doi.org/10.1038/s41587-023-01888-4

Extended Data Fig. 4 | AZD7648 treatment improves gene targeting at *HBB*, *CFTR* and *HBA1* loci in PSC. a. Schematic for gene targeting at the *HBB* locus to edit SCD (E6V) mutation in exon 1 using RNP/AAV6 gene editing (upper panel). Allelic distribution of WT, INDEL and HDR frequencies following gene editing at the *HBB* locus with and without AZD7648 treatment as measured by ICE analysis (lower panel) (n = 3). Mean HDR to INDEL ratio is represented above the bars. **b**. Schematic for gene targeting at the *CFTR* locus to edit CF disease mutation (ΔF508) in exon 11 using RNP/AAV6 gene editing (upper panel). Allelic distribution of WT, INDEL and HDR frequencies following gene editing with and without AZD7648 treatment at the *CFTR* locus as measured by ICE analysis (lower panel) (n = 3). Mean HDR to INDEL ratio is represented above the bars. **c**. Schematic for gene targeting at the *HBB* locus for knock-in of UBC-GFP-bGHpA sequence using RNP/AAV6 gene editing (upper panel). Allelic distribution of WT, INDEL and HDR frequencies following gene editing at the *HBB* locus with and without AZD7648 and M3814 (2 and 0.5 μM) treatment (lower panel) (n = 3). HDR efficiency was measured through ddPCR analysis. WT and INDEL frequencies were measured by ICE analysis. Mean HDR to INDEL ratio is represented above the bars. **d**. Schematic for gene replacement at the *HBA1* locus for knock-in of transgene-2A-YFP sequence using RNP/AAV6 gene editing (upper panel). Allelic distribution of WT, INDEL and HDR frequencies following gene editing at the *HBA1* locus with and without AZD7648 and M3814 (2 and 0.5 μM) treatment (lower panel) (n = 3). HDR efficiency was measured through ddPCR analysis. WT and INDEL frequencies were measured by ICE analysis. Mean HDR to INDEL ratio is represented above the bars. All data in **a, b, c** and **d** are shown as mean ± SEM.

https://doi.org/10.1038/s41587-023-01888-4

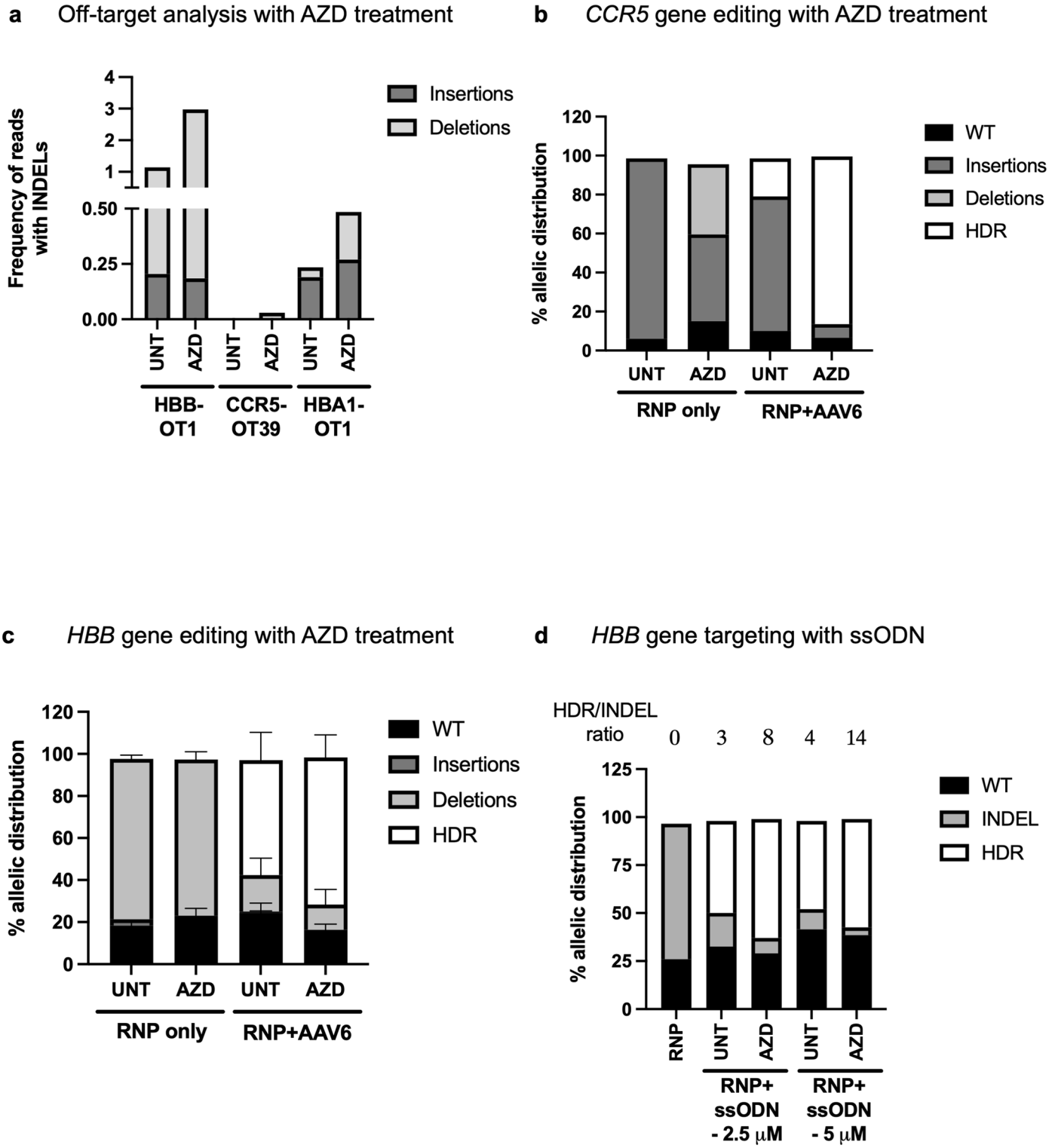


Extended Data Fig. 5 | See next page for caption.

https://doi.org/10.1038/s41587-023-01888-4

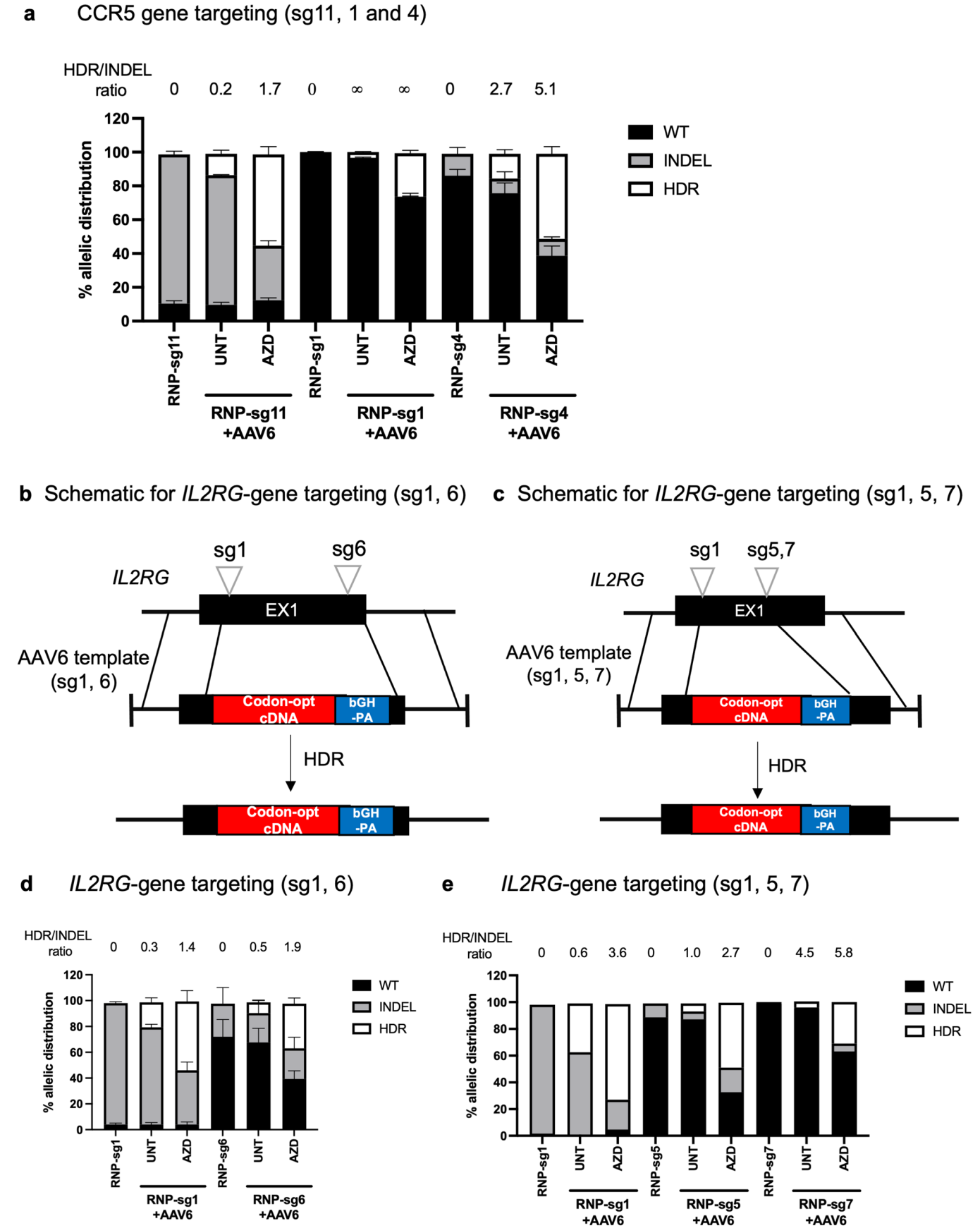
Extended Data Fig. 5 | Gene targeting in HSPCs at *CCR5* and *HBA1* loci with AZD7648 treatment. a. Comparison of different DNA-PKcs inhibitors for gene targeting at the *CCR5* locus for knock-in of UBC-GFP-bGHpA sequence at two different concentrations (2 and 0.5 μM) in HSPCs. Allelic gene targeting efficiency was measured by ddPCR analysis and cell viability was assessed by measuring live cell count and normalizing the number to the Mock sample (n = 3). All data are shown as mean ± SEM. **b**. Allelic distribution of WT, INDEL and HDR frequencies following gene editing at the *CCR5* locus for knock-in of UBC-GFP-bGHpA sequence with different concentrations of AZD7648 as indicated (n = 1). HDR frequency was measured using ddPCR, WT and INDEL frequencies were measured using ICE analysis. **c**. CFU assay was performed on Mock, RNP, AAV6 and RNP + AAV6 treated HSPCs following gene editing at the *CCR5* locus with or without AZD7648 (0.5 μM) treatment. Plot shows the distribution of the mean percentages of CFU-GEMM (multi-potent granulocyte, erythroid, macrophage, megakaryocyte progenitor cells), CFU-GM (colony forming unit-granulocytes and monocytes) and BFU-E (erythroid burst forming units) colonies (n = 2). **d**. Schematic for gene replacement at the *HBA1* locus to replace *HBA1* with *HBB* sequence using RNP/AAV6 gene editing. **e**. Allelic distribution of WT, INDEL and HDR frequencies following gene editing at the *HBA1 locus* with or without AZD7648 treatment (0.5 μM) (n = 3). HDR efficiency was measured through ddPCR analysis. WT and INDEL frequencies were measured by ICE analysis. Mean HDR to INDEL ratio is represented above the bars. All data are shown as mean ± SEM. **f**. Mean percentage of viable cell count of HSPCs at 3 days post-gene targeting (GT) at *CCR5* (n = 4), *HBB* (n = 2) and *HBA1* (n = 2) loci in untreated (UNT) and AZD7648 (AZD, 0.5 μM) treated cells. Mock, AAV only and RNP only treated cells were included as controls. Viable cell counts were measured at 72 h post-gene editing and plotted as percentage relative to the mock cell count. For *CCR5* gene targeting, cell viability is shown as mean ± SEM.

https://doi.org/10.1038/s41587-023-01888-4



Extended Data Fig. 6 | Off-target analysis, gene targeting at *CCR5* and *HBB* loci in HSPCs. a. HSPCs gene targeted at *HBB, CCR5* and *HBA1* loci with and without AZD7648 treatment (0.5 μM) were assessed for off-target activity at the top off-target sites (OT1 for *HBB*, OT39 for *CCR5* and OT1 for *HBA1*) through next generation sequencing (NGS). Mean frequency of reads with insertion and deletion INDELs is shown following SNP/INDEL detection analysis of the NGS data (n = 2). **b.** HSPCs were gene edited with Cas9 RNP at the *CCR5* locus with and without AZD7648 treatment in the presence or absence of AAV6 donor template (for knock-in of two stop codons). Allelic distribution of mean frequencies of WT, insertions, deletions and HDR were determined by ICE analysis (n = 2). **c.** HSPCs were gene edited with Cas9 RNP at the *HBB* locus with and without AZD7648 treatment in the presence or absence of AAV6 donor template (for SCD mutation editing). Allelic distribution of frequencies of WT, insertions, deletions and HDR were determined by using ICE analysis (n = 3). All data are shown as mean ± SEM. **d.** Allelic distribution of mean WT, INDEL and HDR frequencies following gene editing at the *HBB* locus for editing SCD mutation using RNP and ssODN donor with and without AZD7648 treatment (0.5 μM) (n = 2). ssODN donor was tested at two different concentrations as indicated (2.5 and 5 μM).

https://doi.org/10.1038/s41587-023-01888-4

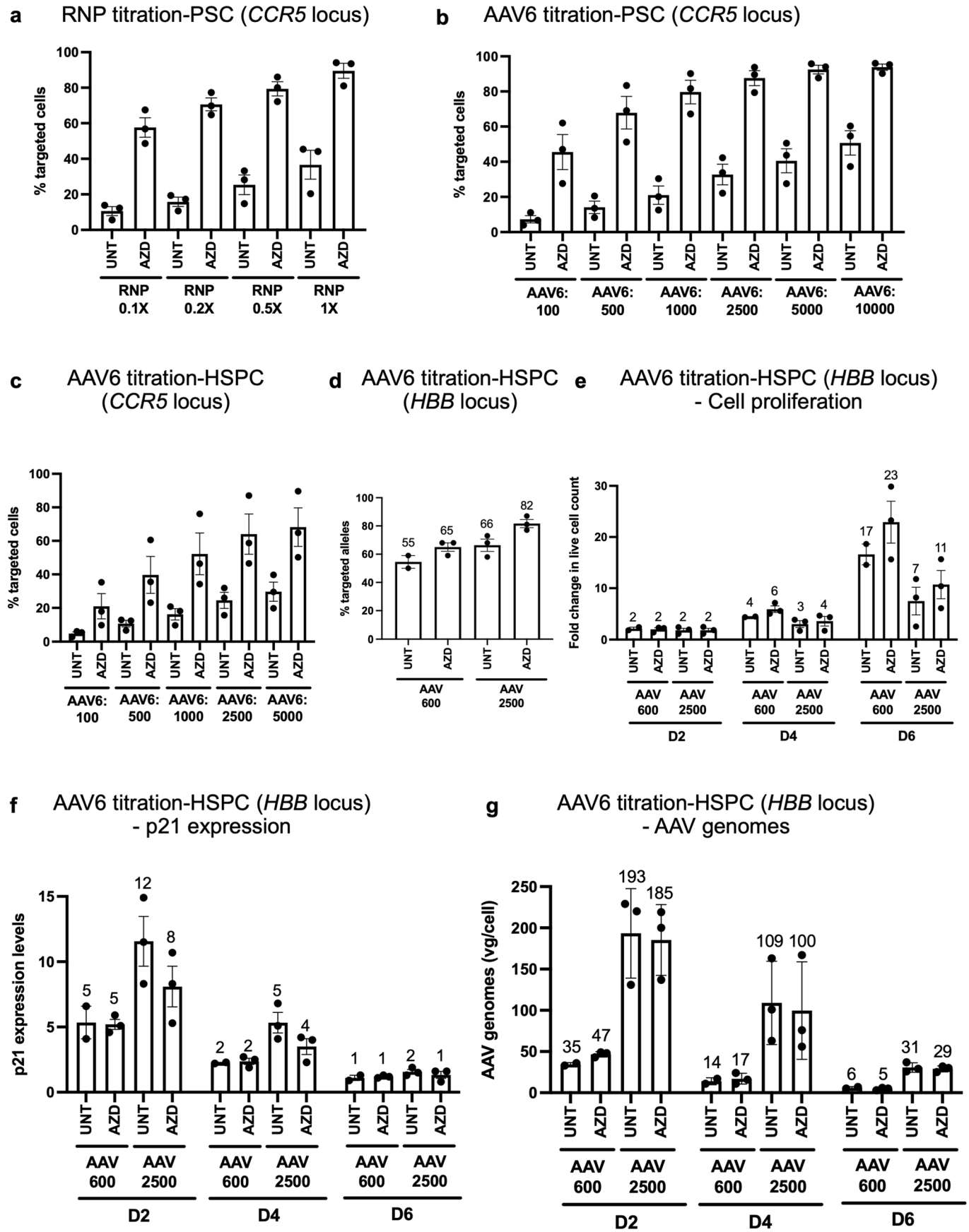


Extended Data Fig. 7 | See next page for caption.

https://doi.org/10.1038/s41587-023-01888-4

Extended Data Fig. 7 | Gene targeting at the *CCR5* and *STING1* loci in HSPCs using seemingly low activity gRNAs. a. HSPCs were gene targeted at the *CCR5* locus for knock-in of UBC-GFP-bGH-pA sequence using high activity (sg11) and low activity gRNAs (sg1 and 4). Allelic distribution of WT, INDEL and HDR frequencies following gene editing at the *CCR5* locus with and without AZD7648 treatment (n = 3). HDR frequency was determined by ddPCR analysis. WT and INDEL frequencies were determined by ICE analysis. Mean HDR to INDEL ratio is represented above the bars. All data are shown as mean ± SEM. **b**. Schematic for gene targeting at the *IL2RG* locus for knock-in of codon-optimized cDNA and bGHpA in exon 1 using RNP/AAV6 gene editing with a high activity gRNA (sg1) and a low activity gRNA (sg6). **c**. Schematic for gene targeting at the *IL2RG* locus for knock-in of codon-optimized cDNA and bGHpA in exon 1 using RNP/AAV6 gene editing with a high activity gRNA (sg1) and two low activity gRNAs (sg5, 7). **d**. Allelic distribution of WT, INDEL and HDR frequencies in HSPCs gene targeted at the *IL2RG* locus with or without AZD7648 treatment (0.5 μM) using sg1 and 6 gRNAs (n = 3) (**b**). HDR efficiency was measured through ddPCR analysis. WT and INDEL frequencies were measured by ICE analysis. Mean HDR to INDEL ratio is represented above the bars. All data are shown as mean ± SEM. **e**. Allelic distribution of mean WT, INDEL and HDR frequencies in HSPCs gene targeted at the *IL2RG* locus with or without AZD7648 treatment (0.5 μM) using sg1, 5 and 7 (n = 2) (**c**). HDR frequency was measured through ddPCR analysis. WT and INDEL frequencies were measured by ICE analysis. Mean HDR to INDEL ratio is represented above the bars.

https://doi.org/10.1038/s41587-023-01888-4

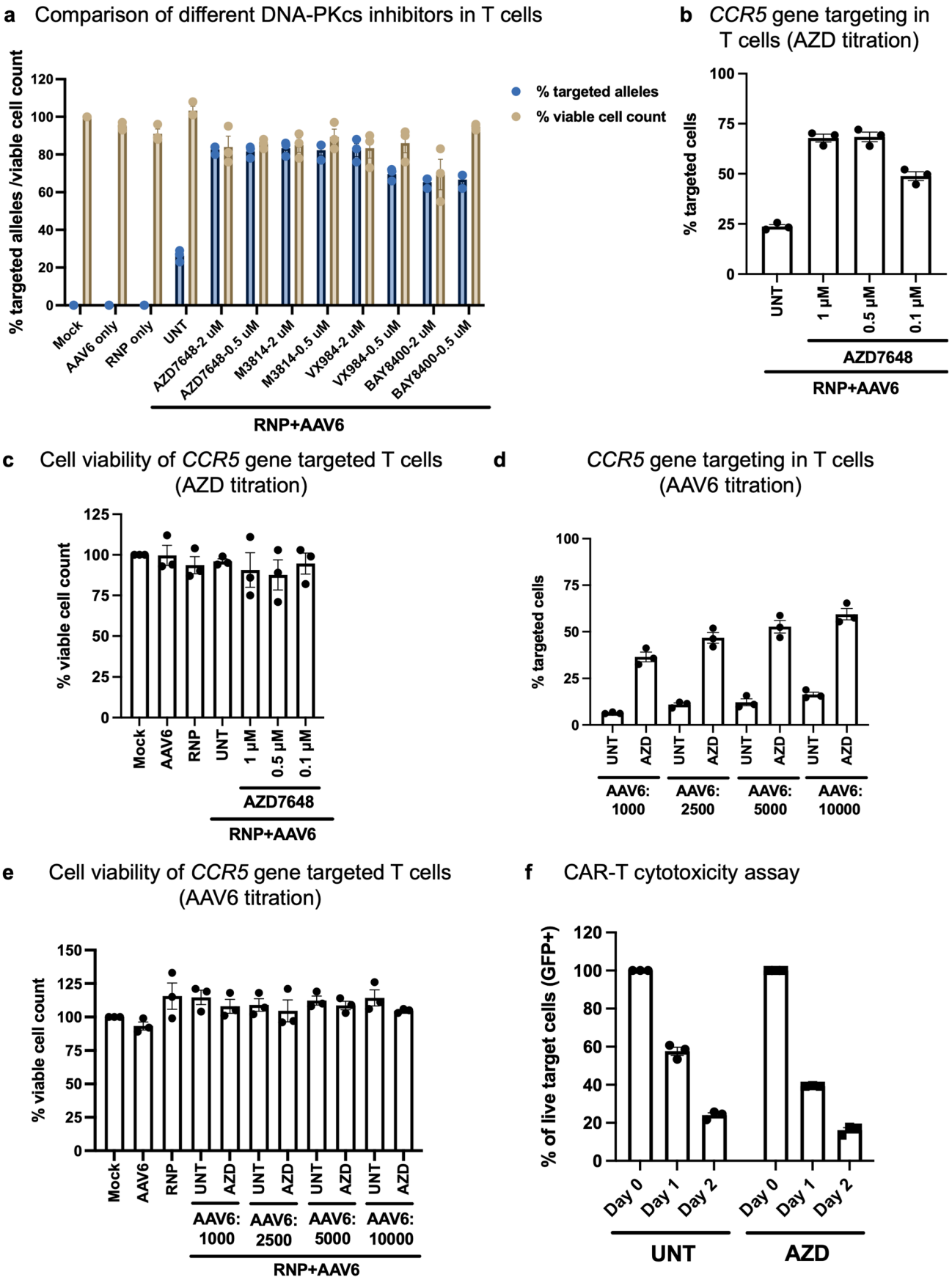


Extended Data Fig. 8 | See next page for caption.

https://doi.org/10.1038/s41587-023-01888-4

Extended Data Fig. 8 | AZD7648 improves gene targeting with lower amounts of RNP/AAV6 in PSC and HSPC. a. Frequency of gene targeted cells measured by flow cytometry for GFP following gene editing at the *CCR5* locus for knock-in of UBC-GFP-bGH-pA sequence with or without AZD7648 treatment (0.25 μM) using varying amounts of Cas9-RNP as indicated and fixed amount of AAV6 donor (MOI:2500) in PSCs (n = 3). RNP-1X denotes 250 μg/ml of Cas9 protein complexed with 100 μg/ml of gRNA. **b**. Frequency of gene targeted cells measured by flow cytometry for GFP following gene editing at the *CCR5* locus for knock-in of UBC-GFP-bGH-pA sequence in PSCs with or without AZD7648 treatment (0.25 μM) using a fixed amount of RNP and varying amounts of AAV6 donor (n = 3). **c**. Frequency of gene targeted cells measured by flow cytometry for GFP following gene editing at the *CCR5* locus for knock-in of UBC-GFP-bGH-pA sequence in HSPCs with or without AZD7648 treatment (0.5 μM) using a fixed amount of RNP and varying amounts of AAV6 donor (n = 3). **d**. Frequency of gene targeted alleles measured by ICE analysis following gene editing at the *HBB* locus for editing the SCD mutation in HSPCs with or without AZD7648 treatment (0.5 μM) using a fixed amount of RNP and varying amounts of AAV6 donor (n = 3). **e**. Fold change in live cell count at D2, 4 and 6 post gene editing relative to D0 of HSPCs gene targeted at the *HBB* locus (**d**) (n = 3). **f**. p21 (*CDKN1A*) expression levels in gene targeted HSPCs **(d)** at D2, 4 and 6 post gene editing represented as values relative to non-gene edited control cells (n = 3). p21 expression levels were measured by RT-ddPCR and normalized to the levels of the house keeping gene, TBP. **g**. Number of AAV genomes in gene targeted HSPCs (**d**) at D2, 4 and 6 post gene editing represented as viral genomes (vg) per cell (n = 3). AAV genomes were quantified using ddPCR amplifying the AAV ITR and the values were normalized to a reference locus, *ZEB2*. Mean values are represented above the bars. All data are shown as mean ± SEM.

https://doi.org/10.1038/s41587-023-01888-4

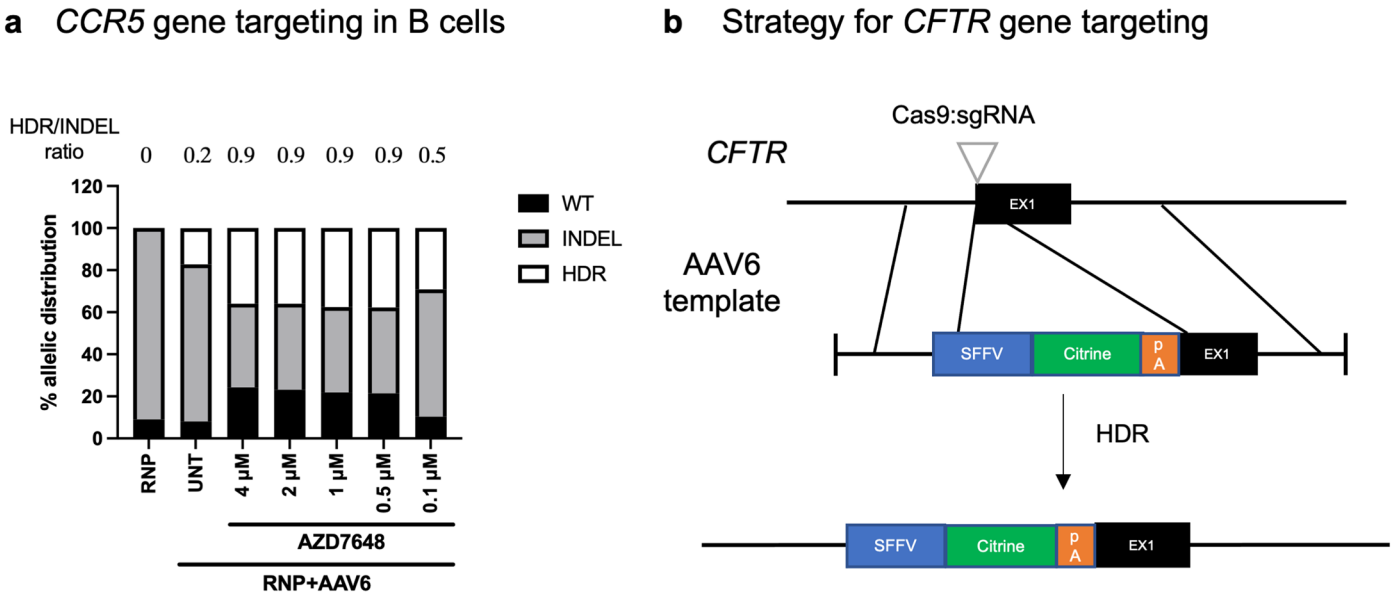


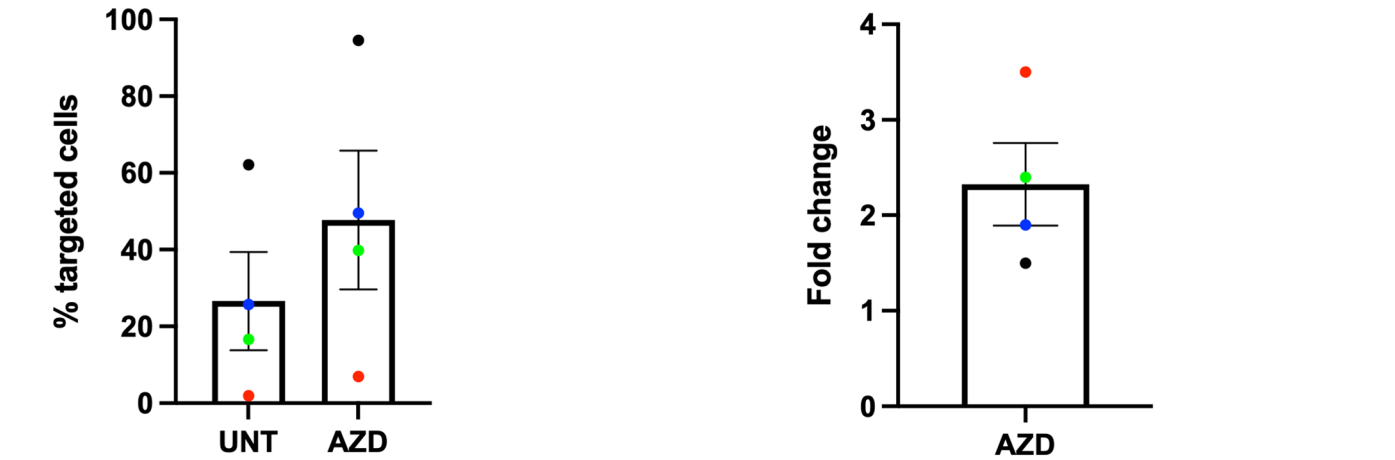
Extended Data Fig. 9 | See next page for caption.

https://doi.org/10.1038/s41587-023-01888-4

Extended Data Fig. 9 | AZD7648 improves gene targeting in T cells. **a**. Comparison of different DNA-PKcs inhibitors for gene targeting at the *CCR5* locus for knock-in of UBC-GFP-bGHpA sequence at two different concentrations (2 and 0.5 μM) in T cells. Allelic gene targeting efficiency was measured by ddPCR analysis and cell viability was assessed by measuring live cell counts and normalizing the number to the Mock sample (n = 3). **b**. Frequency of gene targeted cells as measured by flow cytometry for GFP expression following UBC-GFP-bGH-pA sequence knock-in at the *CCR5* locus in T cells with varying concentrations of AZD7648 using RNP/AAV6 gene editing (n = 3). **c**. Percentage of viable cell count for the *CCR5* gene targeted T cells (**b**) at 72 h post gene editing (n = 3) relative to the mock cells. **d**. Frequency of gene targeted cells as measured by flow cytometry for GFP following UBC-GFP-bGH-pA sequence knock-in at the *CCR5* locus in T cells with or without AZD7648 (0.5 μM) treatment using a fixed amount of RNP and varying amounts of AAV6 donor (n = 3). **e**. Percentage of viable cell count for the *CCR5* gene targeted T cells (**d**) at 72 h post gene editing (n = 3) relative to the mock cells. **f**. T cells were gene targeted at the *TRAC* locus for knock-in of CD19-CAR with and without AZD treatment (n = 3). Engineered CD19 CAR T cells were challenged with GFP+ Nalm6 leukemia target cells in co-culture at an effector to target ratio of 1:1 for 72 hours. Potency of the CAR T cell cytotoxicity activity was monitored by the residual percentage of GFP+ target cells by flow cytometry at 24 h and 48 h post challenge. All data are shown as mean ± SEM.

https://doi.org/10.1038/s41587-023-01888-4





Extended Data Fig. 10 | AZD7648 improves gene targeting in B cells and HBECs. a. Allelic distribution of mean WT, INDEL and HDR frequencies following gene editing at the *CCR5* locus for knock-in of UBC-GFP-bGH-pA sequence in B cells with varying concentrations of AZD7648 using RNP/AAV6 gene editing (n = 2). HDR frequency was measured by ddPCR. WT and INDEL frequencies were measured through ICE analysis. Mean HDR to INDEL ratio is represented above the bars. **b**. Schematic for gene targeting at exon 1 of the *CFTR* gene for knock-in of SFFV-Citrine-pA sequence using RNP/AAV6 gene editing. **c**. Frequency of gene targeted cells as measured by flow cytometry for Citrine expression following *CFTR*-exon 1 gene editing (**b**) with or without AZD7648 treatment (0.5 µM) in HBECs (n = 4). **d**. Fold change in the frequency of *CFTR* gene targeted HBECs (**c**) with AZD7648 (0.5 µM) treatment relative to the untreated cells (n = 4). All data in **c** and **d** are shown as mean ± SEM.

nature portfolio

Corresponding author(s): Matthew Porteus

Last updated by author(s): Jun 21, 2023

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Microsoft Excel for Mac (v16.7) was used for data collection. ddPCR data was collected using QuantaSoft software (v1.7.4.0917, Biorad). FACS/flow cytometry data was collected through respective instrument software (FACS diva 8 for FACS Aria II and Accuri C6 for Accuri). Western blot and gel imaging data were collected using Image lab 2 software (Biorad). STEMvision automated counter (STEMCELL Technologies) to determine the number of BFU-E, CFU-M, CFU-GM, and CFU-GEMM colonies following CFU assay.
Data analysis	Graph pad 9 and Microsoft Excel for Mac (v16.7) were used for data analysis. WT, INDEL and HDR frequency analysis was done through ICE analysis online tool (https://ice.synthego.com/) on Sanger sequencing samples. FACS data was analyzed through FlowJo 10.8.1 for Mac. Illumina Control, Fast QC and Trimmomatic v.0.36 Softwares was used for base scaling, quality check and adapter trimming respectively for NGS. AZENTA's proprietary Amplicon-EZ program was used for SNP/INDEL detection analysis of the NGS data. ImageJ (v1.53) was used for quantification of immunofluorescence staining images.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

NGS dataset for off-target analysis is available at National Center for Biotechnology Information's Sequence Read Archive database (accession code PRJNA982854) which can be accessed through the following link, https://www.ncbi.nlm.nih.gov/sra/PRJNA982854. Source data are provided with this paper. Source data for the figures is provided in source data-file1 and for the extended data figures is provided in source data-file 2. Sequences of the guide RNA, PCR primers, ddPCR primers and probes have been included in Supplementary Table 1. Raw data for ICE analysis and screenshots of Sanger sequencing chromatograms are provided in the supplementary figures.

Human research participants

Policy information about studies involving human research participants and Sex and Gender in Research.

Reporting on sex and gender	Primary cell samples used in this study were de-identified. So we do not have information on the gender of the donors.
Population characteristics	Primary cell samples used in this study were de-identified. So we do not have information on the population characteristics of the donors.
Recruitment	Primary HSPC, B and T cells were derived from healthy donors and there was no selection bias involved. CF patient-derived HBECs were obtained from lungs explanted during transplantation.
Ethics oversight	Stanford IRB committee approved the research on HSPC and the protocol number is 33813. CF patient HBECs were obtained from lungs explanted during transplantation under protocol #13-1396 approved by the Committee on the Protection of the Rights of Human Subjects at the University of North Carolina at Chapel Hill.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Most experiments were performed in 3 different pluripotent stem cell lines and primary cells from 3 different donors for HSPC, T cells and HBECs. Some experiments were performed in 2 different pluripotent stem cell lines and primary cells from 2 donors for HSPC and B cells. The sample size was determined based on the previous studies (Martin and Fowler et al, Nat Comm 2020,
Data exclusions	No data were excluded from the analyses
Replication	All attempts at replication was successful. Most experiments were performed in 2 to 3 biological replicates
Randomization	Sample allocation was random throughout the study
Blinding	The off-target analysis was done through Azenta, Inc NGS service and the samples were blinded when submitted for this analysis. Sanger sequencing was done through Mclab or Azenta and the samples were blinded when they were submitted for sequencing. The remaining experiments couldn't be performed with blinding as they were comparison between with and without a single small molecule treatment.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☐	☒ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used

Alexa Fluor 647 conjugated anti-SSEA4 antibody (Biolegend, cat: 330407). OCT3/4 (cat: sc-5279), SOX2 (cat: sc-365823) and NANOG (cat: sc-293121) were all from Santa Cruz Biotechnology. Alexa Fluor-647 conjugated anti-mouse secondary antibody (Thermo Fisher Scientific, cat: A21235). PAX6 (BD Biosciences, cat: BDB562249), NES (Biolegend, cat: 656805), CD56 (Biolegend, cat: 318305), T (R&D systems, cat: IC2085A), CXCR4 (Biolegend, cat: 306505) and SOX17 (BD Biosciences, cat: BDB562594). Anti-beta-Actin antibody (BioLegend, cat: 643802), Anti-DNA-PKcs antibody (Cell Signaling Technology, cat: 38168S), Anti-DNA-PKcs (phospho-Ser2056) antibody (Abcam, cat: ab18192) and Anti-AKT (phosphor-Ser473) antibody (Cell Signaling Technology, cat: 4060S). Anti-mouse IgG-HRP antibody (Cytiva, cat: NA931) and Anti-Rabbit IgG-HRP antibody (Cytiva, cat: NA934). APC anti-CD34 (BioLegend, cat: 343510), BV785 anti-CD90 (BioLegend, cat: 328142), BV650 anti-CD38 (BioLegend, cat: 356619), BV605 anti-CD45RA (BioLegend, cat: 304134) and BV510 anti-Lineage cocktail (BioLegend, cat: 348807)

Validation

Alexa Fluor 647 conjugated anti-SSEA4 antibody (Biolegend, cat: 330407) has been validated for flow cytometry in human NCCIT (human teratocarcinoma cell line) and cited in 25 publications according to the manufacturer's website. OCT3/4 (SCBT, cat: sc-5279) has been validated for immunostaining in human adrenal gland tissue and cited in 2439 publications according to manufacturer's website. SOX2 (SCBT, cat: sc-365823) has been validated for immunostaining in human bronchus tissue and cited in 264 publications according to manufacturer's website. NANOG (SCBT, cat: sc-293121) has been validated for immunostaining in NTERA-2 (pluripotent human embryonal carcinoma cell line) and cited in 122 publications according to manufacturer's website. Alexa Fluor-647 conjugated anti-mouse secondary antibody (Thermo Fisher Scientific, cat: A21235) has been validated for immunostaining and flow cytometry with anti-mouse primary antibodies and cited in 1279 publications according to manufacturer's website. PAX6 (BD Biosciences, cat: BDB562249) has been validated for flow cytometry in human ESC derived neural stem cells by the manufacturer. NES (Biolegend, cat: 656805) has been validated for flow cytometry in human umbilical vein endothelial cells and cited in 3 publications according to manufacturer's website. CD56 (Biolegend, cat: 318305) has been validated for flow cytometry in human peripheral blood lymphocytes and cited in 37 publications according to manufacturer's website. T (R&D systems, cat: IC2085A) has been validated for flow cytometry in human iPSC derived mesoderm and cited in 4 publications according to manufacturer's website. CXCR4 (Biolegend, cat: 306505) has been validated for flow cytometry in human peripheral blood lymphocytes and cited in 38 publications according to manufacturer's website. SOX17 (BD Biosciences, cat: BDB562594) has been validated for flow cytometry in human ESC derived endoderm by the manufacturer. Anti-beta-Actin antibody (BioLegend, cat: 643802) has been validated for western blot in HeLa cells (immortalized human cervix adenocarcinoma epithelial cells) and cited in 22 publications according to manufacturer's website. Anti-DNA-PKcs antibody (Cell Signaling Technology, cat: 38168S) has been validated for western blot in M059K cells (human glial cell line from malignant glioblastoma) and cited in 15 publications according to manufacturer's website. Anti-DNA-PKcs (phospho-Ser2056) antibody (Abcam, cat: ab18192) has been validated for western blot in HeLa cells (immortalized human cervix adenocarcinoma epithelial cells) and cited in 166 publications according to manufacturer's website. Anti-AKT (phosphor-Ser473) antibody (Cell Signaling Technology, cat: 4060S) has been validated for western blot in PC-3 cells (human adenocarcinoma cell line) and cited in 12083 publications according to manufacturer's website. Anti-mouse IgG-HRP antibody (Cytiva, cat: NA931) has been validated for western blot against anti-mouse primary antibodies in 128 publications according to citeab website. Anti-Rabbit IgG-HRP antibody (Cytiva, cat: NA934) has been validated for western blot against anti-rabbit primary antibodies in 152 publications according to citeab website. APC anti-CD34 (BioLegend, cat: 343510) has been validated for flow cytometry in human peripheral blood leukocytes and cited in 18 publications according to manufacturer's website. BV785 anti-CD90 (BioLegend, cat: 328142) has been validated for flow cytometry by the manufacturer in human erythroleukemia cell line. BV650 anti-CD38 (BioLegend, cat: 356619) has been validated for flow cytometry in human peripheral blood lymphocytes and cited in 8 publications according to manufacturer's website. BV605 anti-CD45RA (BioLegend, cat: 304134) has been validated for flow cytometry in human peripheral blood lymphocytes and cited in 30 publications according to manufacturer's website. BV510 anti-Lineage cocktail (BioLegend, cat: 348807) has been validated for flow cytometry in human peripheral blood leukocytes and cited in 10 publications according to manufacturer's website.

Eukaryotic cell lines

Policy information about cell lines and Sex and Gender in Research

Cell line source(s)

H9 ES cells were acquired from WiCell. 1205-4 and 1208-2 iPS cell lines were kindly provided by Sergiu Pasca lab at Stanford University. 293T cells were acquired fro ATCC.

Authentication

H9 ES cells were authenticated by STR profiling at WiCell. 1205-4 and 1208-2 iPS cell lines were authenticated through SNP arrays by the Sergiu Pasca lab (Khan et al, Nature Medicine, 2020). 293T cells were authenticated by STR profiling at ATCC.

Mycoplasma contamination

All cell lines were tested for mycoplasma and found to be negative

Commonly misidentified lines
(See ICLAC register)

No commonly misidentified cell lines were used.

Flow Cytometry

Plots

Confirm that:

☒ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☒ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☒ All plots are contour plots with outliers or pseudocolor plots.

☒ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation	Cord-blood derived human CD34+ HSPCs were blocked with Human TruStain FcX and treated with LIVE/DEAD™ Fixable Near-IR Dead Cell Stain for 5 minutes in FACS buffer. Following this, cells were treated with a cocktail of antibodies: APC anti-CD34 (BioLegend, cat: 343510), BV785 anti-CD90 (BioLegend, cat: 328142), BV650 anti-CD38 (BioLegend, cat: 356619), BV605 anti-CD45RA (BioLegend, cat: 304134) and BV510 anti-Lineage cocktail (BioLegend, cat: 348807) and incubated for 20 minutes on ice. The cells were then washed with PBS and then resuspended in FACS buffer.
Instrument	FACS Aria II sorter (BD biosciences)
Software	FACS Diva software 8 was used for data collection. Data analysis was performed using FlowJo 10.8.1 for Mac.
Cell population abundance	The samples were more than 95% pure as determined by post-sort flow cytometry analysis
Gating strategy	FSC-A/SSC-A plot was used to gate for the starting cell population. Singlets were selected with FSC-H/FSC-W and SSC-H/SSC-W plots. Gating was done to determine positive/negative populations based on the negative control cells for each markers.

☒ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.