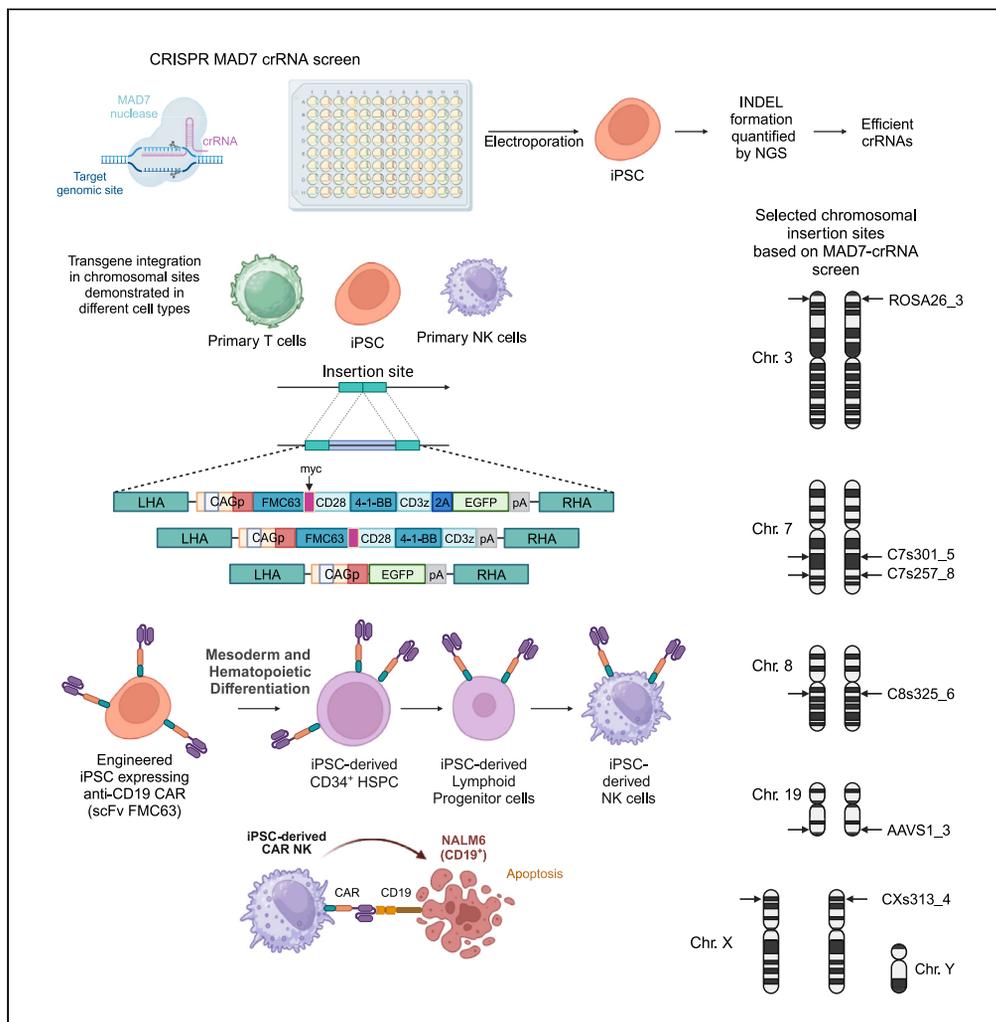


Article

CRISPR-Cas12a-integrated transgenes in genomic safe harbors retain high expression in human hematopoietic iPSC-derived lineages and primary cells



Arsenios Vlassis,
Tanja L. Jensen,
Marina Mohr, ...,
Lars K. Nielsen,
Tanya Warnecke,
Ryan T. Gill

rtg@artisancells.com

Highlights

CRISPR-MAD7 in complex with efficient crRNAs allows for up to 96% indels in human iPSCs

Demonstration of MAD7-facilitated integration of a CAR-EGFP gene in 4 safe harbors

GFP⁺ and CAR⁺GFP⁺ iPSC clones differentiated with 80%–90% efficiency to CD34⁺ HSPCs

HSPCs differentiated further to CD56⁺CD45⁺CD43⁺ NK cells with up to 16% efficiency

Vlassis et al., iScience 26, 108287
December 15, 2023 © 2023
<https://doi.org/10.1016/j.isci.2023.108287>



Article

CRISPR-Cas12a-integrated transgenes in genomic safe harbors retain high expression in human hematopoietic iPSC-derived lineages and primary cells

Arsenios Vlassis,^{1,5} Tanja L. Jensen,^{1,6} Marina Mohr,^{1,6} Dominika J. Jedrzejczyk,^{1,6} Xiangyou Meng,¹ Gergo Kovacs,⁴ Martí Morera-Gómez,¹ Andrea Barghetti,^{1,2} Sergi Muyo Abad,¹ Roland F. Baumgartner,^{1,2} Kedar N. Natarajan,⁴ Lars K. Nielsen,^{1,3} Tanya Warnecke,^{1,2} and Ryan T. Gill^{1,2,7,*}

SUMMARY

Discovery of genomic safe harbor sites (SHSs) is fundamental for multiple transgene integrations, such as reporter genes, chimeric antigen receptors (CARs), and safety switches, which are required for safe cell products for regenerative cell therapies and immunotherapies. Here we identified and characterized potential SHS in human cells. Using the CRISPR-MAD7 system, we integrated transgenes at these sites in induced pluripotent stem cells (iPSCs), primary T and natural killer (NK) cells, and Jurkat cell line, and demonstrated efficient and stable expression at these loci. Subsequently, we validated the differentiation potential of engineered iPSC toward CD34⁺ hematopoietic stem and progenitor cells (HSPCs), lymphoid progenitor cells (LPCs), and NK cells and showed that transgene expression was perpetuated in these lineages. Finally, we demonstrated that engineered iPSC-derived NK cells retained expression of a non-virally integrated anti-CD19 CAR, suggesting that several of the investigated SHSs can be used to engineer cells for adoptive immunotherapies.

INTRODUCTION

Targeted integration at specific genomic sites in mammalian cells enables regulated transgene expression without disrupting the expression of endogenous genes. Gene expression depends on cell type and developmental stage on account of differentiation-specific activation or inactivation of different parts of chromatin.^{1,2} Due to the complex transcriptional gene regulation in mammalian cells through networks of *cis*- and *trans*-regulatory elements,³ altering the default genomic architecture by integration of exogenous sequences can disrupt cell homeostasis. Specifically, unintended gene integration next to proto-oncogenes or tumor suppressor genes can result in oncogenesis, which is not compatible with clinical applications.⁴

Since the first demonstrations of successful reprogramming of somatic murine⁵ and human fibroblasts⁶ to iPSCs, and of the comparability between human iPSCs and human embryonic stem cells (hESCs),⁷ collectively called human pluripotent stem cells (hPSCs), extensive research has focused on genomic engineering of PSCs. Steering endogenous gene expression or transgene integration in PSCs, followed by directed differentiation to somatic cells and tissues, is intensively investigated with the goal of developing advanced regenerative therapies.^{8,9} The first clinical applications have relied on viral vectors and transposons for semi-random transgene integration that resulted in uncontrolled insertion, heterogeneous expression, and even oncogenesis.^{4,10} Thus, targeted gene integrations at so-called genomic “safe harbor” sites that do not cause genetic instability or transcriptome or cell cycle perturbations will be ultimately required for clinical applications. Several criteria for selection of candidate SHSs in the human genome have been described.^{8,11}

Currently, targeted chromosomal integration of exogenous sequences is usually achieved by homology-directed repair (HDR) of CRISPR nuclease-induced DNA double-strand breaks using HDR-donor repair templates (HDRTs). While the most commonly used genome-editing CRISPR system utilizes Cas9,^{12–14} Cas12a nucleases have recently received attention¹⁵ due to their broader protospacer-adjacent motif specificity and lower off-target activity than Cas9.^{16,17} ErCas12a nuclease (MAD7) has shown high genome-editing efficiency in a wide range of organisms including human cells^{18,19} and is a leading alternative to Cas9 for clinical cell therapy applications.^{20,21} Transfection by electroporation (EP) is currently the most efficient and broadly utilized method for delivering CRISPR ribonucleoprotein (RNP) complexes and HDRTs into human

¹Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Kemitorvet 220, 2800 Kongens Lyngby, Denmark

²Artisan Bio, 363 Centennial Parkway, Suite 310, Louisville, CO 80027, USA

³Australian Institute for Bioengineering and Nanotechnology (AIBN), The University of Queensland, Brisbane, QLD 4072, Australia

⁴DTU Bioengineering, Technical University of Denmark, Søtofts Plads 224, 2800 Kongens Lyngby, Denmark

⁵Senior author

⁶These authors contributed equally

⁷Lead contact

*Correspondence: rtg@artisancells.com

<https://doi.org/10.1016/j.isci.2023.108287>



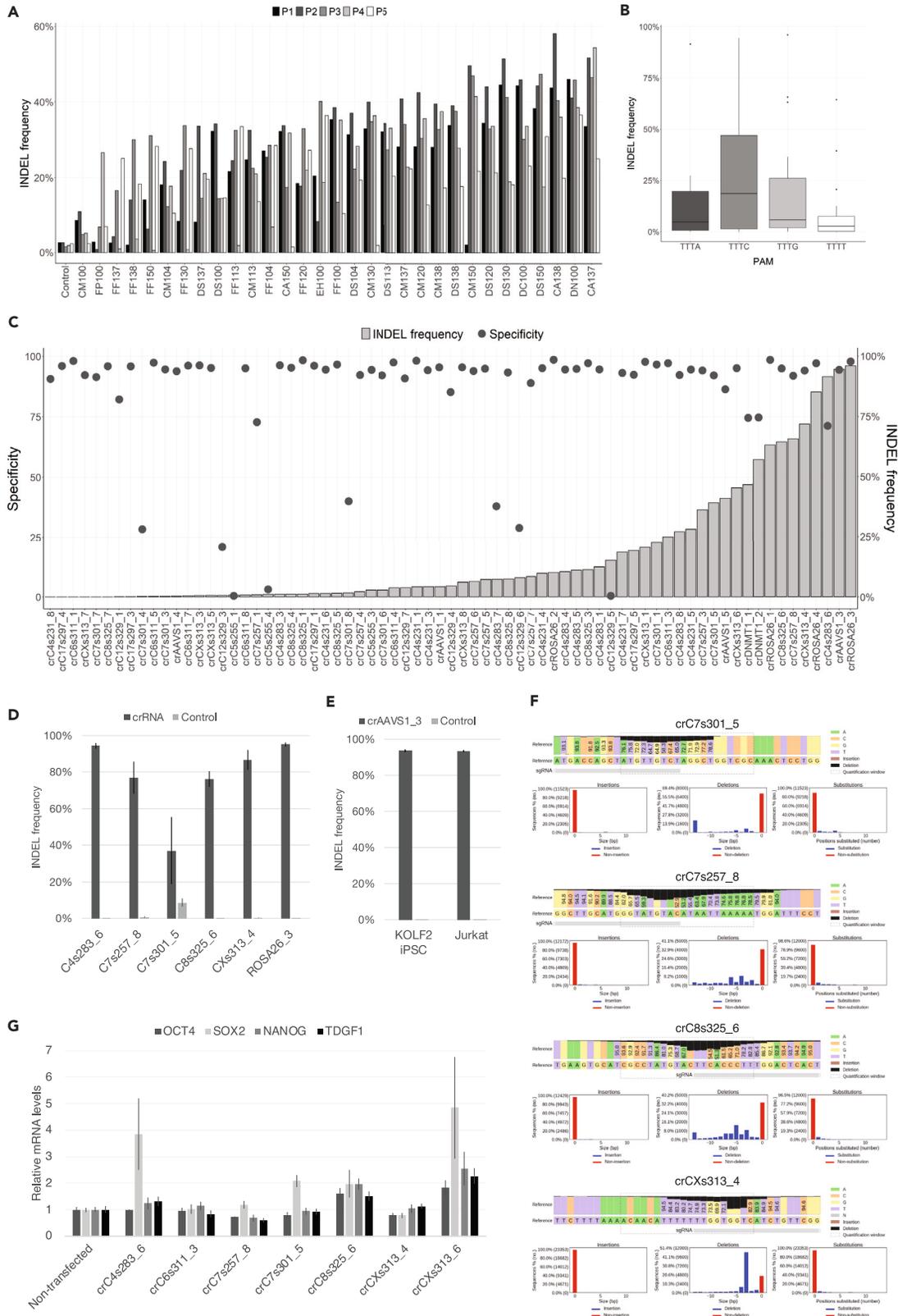


Figure 1. Optimization of electroporation conditions, selection of SHS, and validation

(A) Screen for optimal EP conditions for KOLF2-C1 iPSCs: 31 programs × 5 buffers (155 EP conditions), using MAD7^{3×NLS} and crDNMT1 (biological replicas n = 1; technical replicas t = 1). Indel frequency (analyzed by targeted NGS) as a function of EP programs.

(B) Indel frequency of SHS crRNAs (from C) as a function of PAM sequence. Boxplot generated by ggplot2. Error bars represent standard deviation (SD) of indel frequency among crRNAs that have the same PAM sequence. Dots beyond the end of the whiskers represent outliers.

(C) Indel frequency as a function of SHS crRNAs. crRNA specificity score is depicted as dots. KOLF2 were electroporated in buffer P3, using program CA137, with MAD7^{3×NLS} and SHS crRNA RNPs (n = 1; t = 1). See also Figure S1E for a biological replica of SHS crRNA screen using MAD7^{1×NLS}.

(D) Validation of indel frequency as a function of top crRNAs (n = 3–4; t = 1–2). Error bars represent SD among biological replicas.

(E) Validation of crAAVS1_3 indel frequency in iPSCs and Jurkat (n = 1; t = 3). Error bars represent SD among technical replicas.

(F) Representative CRISPResso2-derived nucleotide percentage quilt around 4 selected SHS crRNAs and histograms depicting the size and frequency of insertions, deletions, and substitutions. Not the entire sequence of crRNA is shown in the quilt because the quantification window is set to ±10 bases around the MAD7 cut site (+1 base at the 3' end of crRNA, as for Cas12a).

(G) Relative mRNA expression levels of non-differentiation markers *OCT4*, *SOX2*, *NANOG*, and *TDGF1* in samples transfected with the indicated crRNAs were calculated by the 2^{-ΔΔCt} method compared to *GAPDH* housekeeping gene and are shown here relative to a non-transfected control sample. Mean ± SD from qPCR technical quadruplicates is shown (n = 1, t = 4). See also Figure S1.

cells.^{22,23} In stem cells, HDR insertion of short sequences using single-stranded oligodeoxynucleotides (ssODNs) can result in efficiencies as high as 70% of all mutagenic events,²⁴ while insertion rates of long DNA HDRTs still remain low.^{25,26} Use of anionic nanoparticles for RNP and HDRT delivery showed improved insertion efficiencies in a wide range of cells,²⁷ while adding enhancers of the HDR-dependent DNA repair, or inhibitors of the non-homologous end-joining (NHEJ) repair, or a combination of these, increases gene integration in iPSC to 27%.²⁶

One of the first sites for transgene integration in hESCs demonstrating stable GFP (green fluorescent protein) expression during passaging and differentiation, located in chromosome 12 at position q23.1 (called “Envy”), was identified after random integration of a human β-actin promoter-driven GFP plasmid.²⁸ Prior to the discovery of targeted nucleases, transgene expression was carried out by semi-random retroviral/lentiviral integration into the genome or by exploiting the site specificity of adeno-associated virus-2 (AAV2).^{9,29} The integration sites in clones with stable expression were retrieved by DNA sequencing followed by validation. This procedure has led to the discovery of site Thal5.10-2 (in chromosome 1: Chr.1 188,083,272) in thalassemia iPSCs.⁹ Other sites were described using the promiscuous phage-derived integrases (Chromosome 13 q32)³⁰ or searching for recognition sequences of the mCrel homing endonuclease.³¹

Using advanced site-specific nucleases (TALENs, Zinc-finger nucleases, and latest CRISPR), several studies have demonstrated stable expression of transgenes at the three most established integration sites, *AAVS1*, *ROSA26*, and *CCR5* in hPSCs and their differentiated progeny.^{32–35} The *AAVS1* (AAV2 site-1) locus, first identified as a hotspot for integration of adeno-associated virus,³⁶ is within the first intron of the ubiquitously expressed gene *PPP1R12C* (Chr.19). Although the consequences of the disruption of this gene following integration of exogenous sequences are not fully investigated yet, *AAVS1* remains a preferred SHS for research studies^{11,37} and has currently been tested for CD19-chimeric antigen receptor (CAR) expression in pre-clinical applications.^{11,38} Another intragenic locus (*CCR5*) has been studied as a safe harbor^{33,34} after the discovery that people carrying a homozygous 32 bp deletion in *CCR5* gene (chemokine C-C-motif receptor 5 gene) that leads to a non-functional receptor were resistant to HIV-1 infection.³⁹ The human *ROSA26* locus (within long noncoding RNA [lncRNA] *THUMPD3-AS1*, Chr.3), the human homolog of the mouse *Rosa26* locus, was initially characterized in hESCs upon transgene integration using linearized HDRTs without a targeting nuclease.⁴⁰ Transgene expression from *ROSA26* persisted in multiple PSC-derived lineages, while the mono- or biallelic loss of the endogenous locus *THUMPD3-AS1* had no detectable effects on the differentiation potential of stem cells.^{35,40} Very recently, bioinformatic screens rationally identified additional sites that satisfy several criteria for safe harbors in human cells^{31,41} and used CRISPR/Cas9-targeted genome integration of reporter genes into some predicted sites in human cell lines. Further, Aznauryan et al., 2022,⁴¹ validated the safety of two SHSs, Rogi1 (GSH1; genomic safe harbor 1; location Chr.1 195,338,589) and Rogi2 (GSH2; location Chr.3 22,720,711) by showing that there are no significant genome-wide transcriptomic alterations following transgene integrations in these sites in human primary T cells and primary fibroblasts. Altogether, discovery and validation of *bona fide* SHSs are instrumental for human cell genomic engineering intended for safe cell therapy products and will enable multiple transgene integrations, including reporter genes, CARs, cytokines, growth factors, and safety switches.

RESULTS

Optimization of MAD7-RNP editing of iPSC

Transfection with RNP complexes enables high-efficiency genome editing with reduced off-target activity.^{16,27,42} In this study, we used purified MAD7 protein with a codon-optimized sequence containing a single, triple, or quadruple nuclear localization signal (NLS) and determined its editing activity in KOLF2 iPSCs by targeted next-generation sequencing (NGS) analysis (Figures 1A and S1A–S1C). Using MAD7^{3×NLS} in complex with CRISPR RNA (crRNA) targeting *DNMT1* locus,¹⁵ we tested 155 different EP conditions on the Lonza Nucleofector 96-well system and identified two conditions (solutions P3 and P4 with program CA137) that resulted in high editing efficiency (45%–50%) without affecting cell growth (Figures 1A and S1A–S1C).

Discovery of SHSs

Using the current human genome assembly GRCh38, we defined a genomic region ±1 kb around 35 mCrel putative recognition sites, 20 nucleotide long, that were previously described as potential SHSs.³¹ These 2 kb regions were used for searching crRNAs for

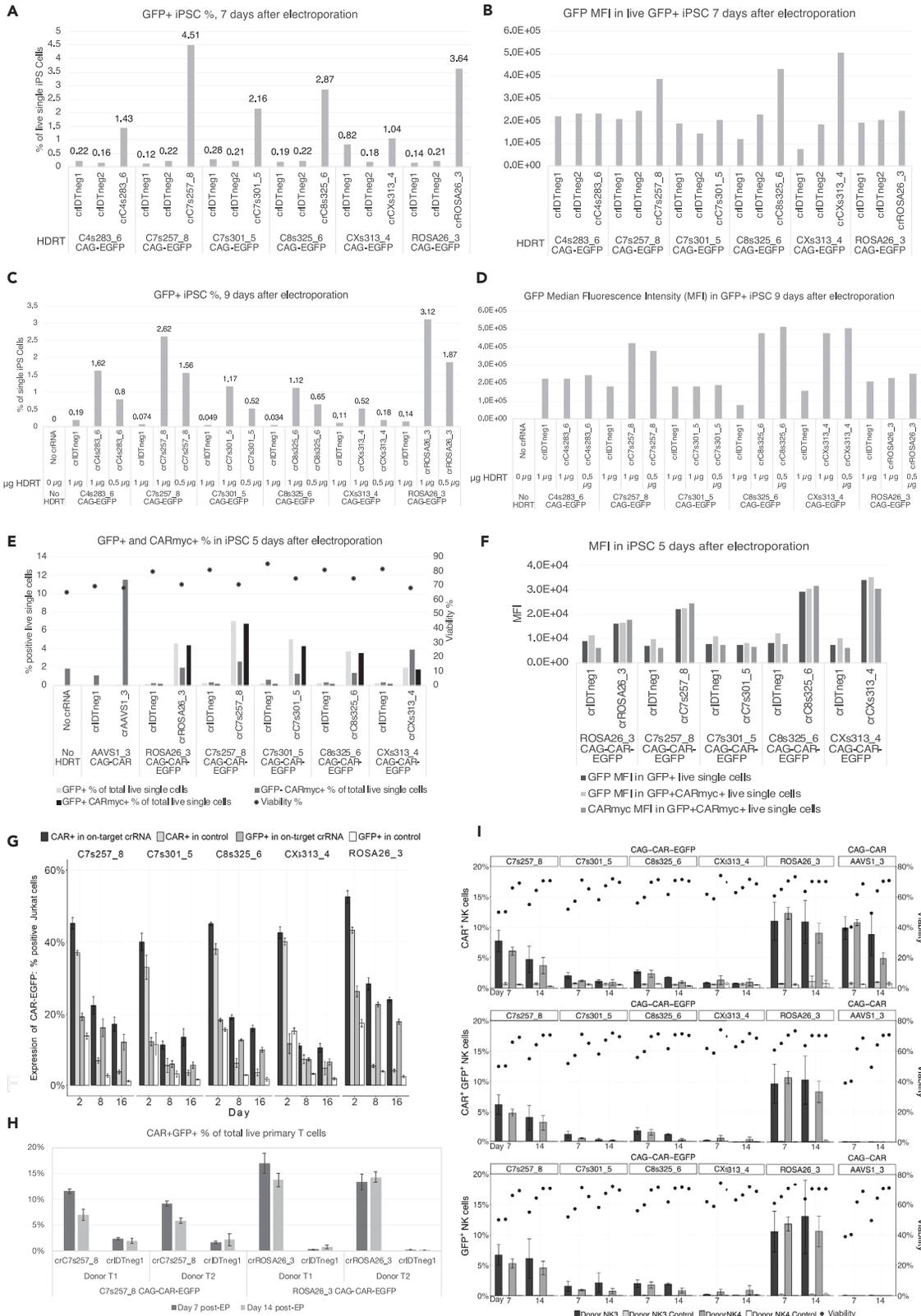


Figure 2. Transgene integrations into SHS in iPSCs, primary T and NK cells, and Jurkat

(A) Quantification of GFP⁺ live iPSC cells upon insertion of GFP into SHSs and ROSA26, 7 days after electroporation (post-EP). EP conditions: 2×10⁵ iPSC, buffer P3, program CA137, 100 pmol RNP (100:125 pmol MAD7^{4×NLS}:crRNA), 1 μg (~0.45 pmol) of CAG-EGFP linear dsDNA HDRT. Non-targeting crRNAs crIDTneg1 and crIDTneg2 were used to assess the non-HDR-mediated insertion and expression of HDRT.

(B) GFP median fluorescence intensity (MFI) in GFP⁺ cells from (A).

(C) Comparison of integration efficiency using 0.5 μg (~0.23 pmol) vs. 1 μg (~0.45 pmol) of CAG-EGFP linear dsDNA HDRT in iPSCs, 9 days post-EP.

(D) GFP MFI in GFP⁺ cells from (C).

(E) Quantification of CAR⁺GFP⁺ double-positive, CAR⁺GFP⁻ single-positive, and total GFP⁺ iPSCs 5 days post-EP with 0.5 pmol CAG-CAR-EGFP HDRT or AAVS1_3 CAG-CAR HDRT. EP conditions: P3-CA137; 2×10⁵ iPSC, 50 pmol RNP (50:60 pmol MAD7^{4×NLS}:crRNA), 100 μg PGA (poly-L-glutamic acid) as RNP carrier, and 0.5 pmol corresponding linear dsDNA HDRT per reaction.

(F) GFP MFI and CAR^{+/yc} MFI in GFP⁺/CAR⁺GFP⁺ iPSCs that are shown in (E).

(G) Expression of CAG-CAR-EGFP (as CAR⁺GFP⁺ double-positive Jurkat cells) inserted into SHSs and ROSA26 in Jurkat over time (2, 8, 16 days post-EP). EP conditions: buffer SF, program CA137; 2×10⁵ Jurkat cells, 50 pmol RNP (50:75 pmol MAD7^{4×NLS}:crRNA, mixed with 100 μg PGA) and 0.5 pmol CAG-CAR-EGFP linear dsDNA HDRT per reaction. Control is IDTneg2 crRNA. Data are represented as mean ± SD of 3 technical replicas per sample from 1 biological replica (n = 1; t = 3); 2nd biol. replica shown in Figure S2D.

(H) Transfection of CAR-GFP in primary T cells and quantification of GFP⁺CAR⁺ live primary T cells on days 7 and 14 post-EP. T cells from two donors: T1 and T2, were electroporated in parallel; mean and SD from 3 technical replicas per crRNA-HDRT combination is shown (n = 2; t = 3). Control crRNA is IDTneg1. EP conditions: P3-EH115. More details in STAR Methods.

(I) Expression of integrated CAR-GFP at SHS and ROSA26, and of CAR at AAVS1 in primary NK cells isolated from two donors NK3 and NK4 (n = 2; t = 3), at days 7 and 14 post-EP. Mean and SD from 3 technical replicas per crRNA-HDRT combination per donor are shown. Control is IDTneg1 crRNA. EP conditions: P3-EN138. More details in STAR Methods. See also Figure S2.

Cas12a/MAD7 later in the pipeline. Then, we examined whether these 2 kb areas overlap with any transcription units, and if not, whether there are transcription units at 50 kb and 300 kb distances (Figures S1D). Because the genome annotation is constantly updated, we found that some sites did not match the commonly accepted SHS distance criteria^{8,11} any longer and there were overlaps. First, we excluded all sites that overlapped or resided within any gene in GRCh38. Also, the sites 233 (Chr.6) and 303 (Chr.2) did not fulfill the criteria: (2) >300 kb from any microRNA (miRNA)/other functional small RNAs and (3) >50 kb from any 5' gene end anymore, despite both sites scoring highest (score 7) among the SHSs of Pellenz et al. 2019.³¹ Thus, 22 sites were excluded, except for sites 311 and 303 because of their high score by Pellenz et al. 2019, and >300 kb distance from cancer-related genes. Site 303 was eventually excluded from further characterization due to low-specificity crRNAs within its 2 kb region. Sites that were closer than 50 kb from >2 genes were also excluded (4 sites). By this exclusion method, we narrowed down to ten putative SHS loci, of which two are closer than 300 kb from any cancer-related genes (sites 257 and 313). However, 257 was selected anyway because it is in open chromatin, and 313 because it is on X chromosome. We adapted the following naming system for the SHSs: C4s231, C4s283, C5s255, C6s311, C7s257, C7s301, C8s325, C12s329, C17s297, and CXs313 (Chromosome; site ID from Pellenz et al. 2019³¹). Next, we designed 88 unique crRNAs for MAD7 (Table S2) within the aforementioned 2 kb SHS-search regions (Table S1), and in a ±0.5 kb window around published Cas9 single guide RNAs (sgRNAs) for AAVS1 and ROSA26.³⁵ Activity of crRNAs was screened in KOLF2 iPSC and determined by targeted NGS analysis (primers in Table S3) that revealed ten highly active crRNAs (≥40% insertion or deletion [indel] frequency) targeting seven loci: AAVS1, ROSA26, C4s283, CXs313, C7s257, C8s325, and C7s301 (Figures 1C and S1E). We validated the efficiency of the seven top-targeting crRNAs (AAVS1_3, ROSA26_3, C4s283_6, C7s257_8, C8s325_6, CXs313_4, and C7s301_5) in iPSC (Figure 1D), and of AAVS1_3 further in Jurkat (Figure 1E), as well as the targeting pattern of four crRNAs (Figure 1F). Because the specificity of crC4s283_6, calculated using Hsu et al. (2013) algorithm,⁴³ was lower than 75% (Figure 1C), this crRNA was only used in initial tests. To confirm the specificity of the selected crRNAs, we have performed a whole-genome *in silico* analysis to predict potential off-targets of the selected crRNAs (Table S6). The analysis determines potential off-targets when there are 0–4 nucleotide mismatches and shows that off-target rates for all selected crRNAs are very low, except for crC4s283_6. Next, we investigated the effect of the transfected crRNAs on the pluripotency of iPSCs by measuring the mRNA levels of the transcription factors OCT4, SOX2, and NANOG, and of the membrane protein gene *TDGF1*, involved in Nodal signaling (Figure 1G). The qPCR data showed that the activity of the top-targeting crRNAs had no effect on any of the transcription factor levels that would suggest loss of pluripotency, which was verified by immunofluorescence staining for OCT4 protein and the SSEA-4 marker (Figures S1F).

Transgene integration into SHSs

Next, we used KOLF2 to evaluate transgene integration and protein expression efficiencies into SHSs using purified MAD7^{4×NLS} and highly efficient crRNAs (Figures 2A–2F and S2A–S2C). For each reaction, we transfected linear double-stranded DNA (dsDNA) HDRT, where the expression cassette is flanked by 500 bp homology arms, together with the respective RNP. All the cassettes contain the strong synthetic CAG (Cytomegalovirus, β-Actin, β-Globin) promoter (Table S4) that drives expression of either the EGFP gene alone (enhanced green fluorescent protein; hereafter GFP denotes the protein, and EGFP the gene), or the FMC63CAR^{44,45} (an anti-CD19 CAR; hereafter CAR) alone, or FMC63CAR and EGFP, where the two CDSs (coding sequences) are separated by a 2A peptide (hereafter CAR-EGFP denotes the transgene and CAR-GFP the co-expressed proteins) (Table S5). Seven days post-electroporation (post-EP; or else post-transfection [PT]) of EGFP HDRT, after the transient expression of GFP was diminished, we observed stable GFP integration and expression of 3.6% at ROSA26_3 (Figure 2A). Similar insertion of EGFP was observed at C7s257_8, while insertion at C8s325_6 resulted in 20% lower efficiency (2.9%) compared to ROSA26_3. Integration efficiency at the other SHSs was overall lower compared to ROSA26_3 (Figures 2A–2D and

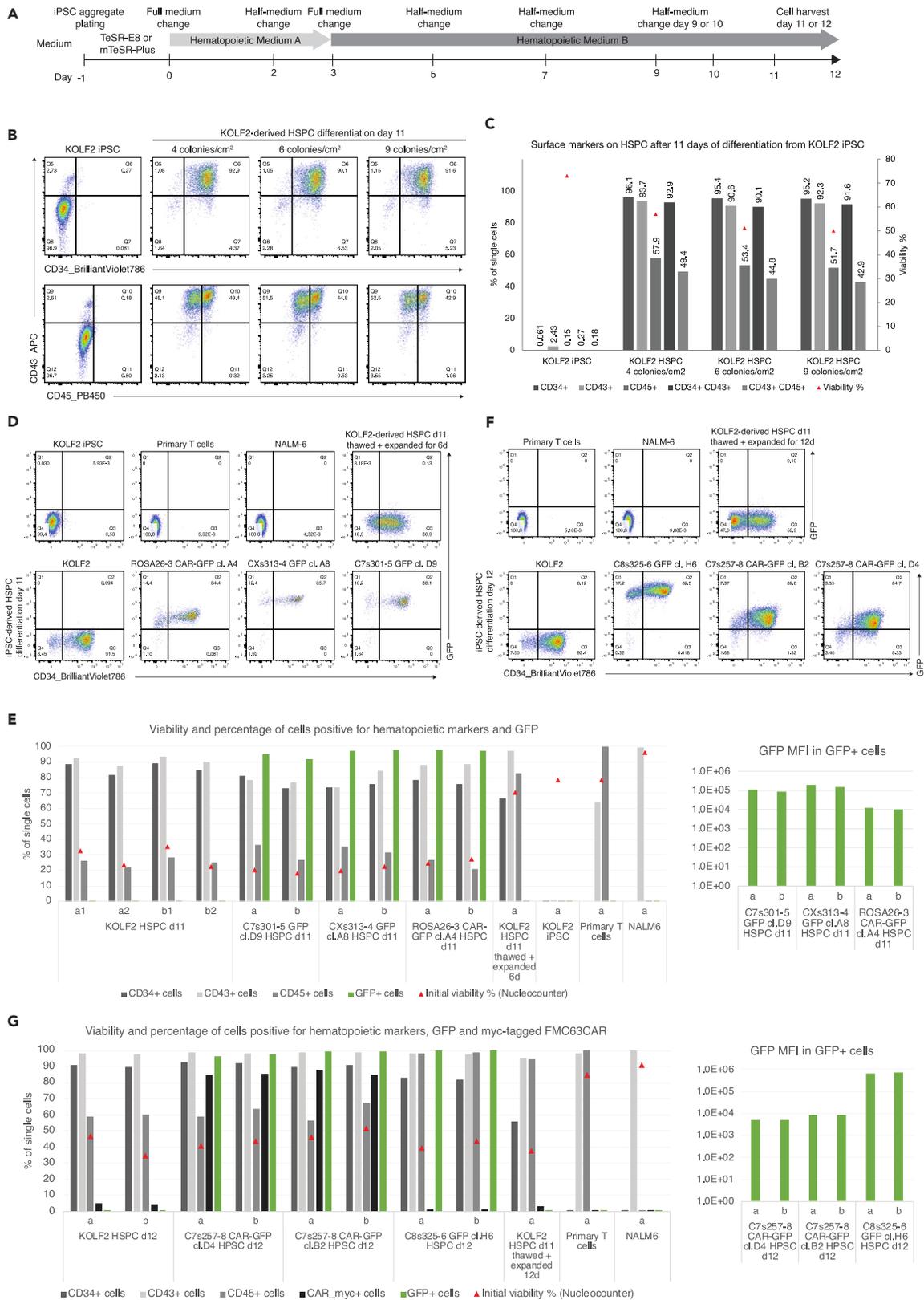


Figure 3. Hematopoietic differentiation of monoclonal iPSC lines with GFP/CAR-EGFP transgenes at SHSs

(A) Differentiation overview (2D protocol; details in STAR Methods).

(B) HSPCs derived from KOLF2. Initial seeding at different densities resulted in 4, 6 or 9 colonies/cm². HSPCs from each initial seeding category were pooled from 3 to 4 technical replica wells (t = 4, pooled), stained with viability marker/or antibodies recognizing the indicated key cell-surface markers of hematopoietic differentiation (CD34, CD43, CD45) in pairs, and analyzed by flow cytometry.

(C) Quantification of positive cells from panel (B) upon double staining with CD34 and CD43 antibodies or CD43 and CD45 antibodies.

(D) Monoclonal iPSC lines ROSA26_3 CAR-GFP (cl. A4), CXs313_4 GFP (cl. A8), or C7s301-5 GFP (cl. D9), along with parental KOLF2 were differentiated to HSPCs and harvested on day 11. Cells were stained with CD34, CD43, and CD45 antibodies, separately and pairwise. Staining controls: primary T cells, NALM6, iPSCs, and maturing HSPCs that were derived from iPSCs in a previous experiment.

(E) Quantification of marker expression and GFP MFI in single-antibody-stained samples from D. "a" and "b" in sample name denote two technical replicas (t = 2).

(F) The monoclonal iPSC lines C8s325_6 GFP (cl. H6) or C7s257_8 CAR-GFP (cl. B2 & D4) were differentiated to HSPCs and harvested on day 12. Cells were stained with CD34, CD43, and CD45 antibodies, separately and pairwise. Staining controls: primary T cells, NALM6, and maturing iPSC-derived HSPCs.

(G) Quantification of marker expression and GFP MFI in single-antibody-stained samples from F. "a" and "b" in sample name denote two technical replicas (t = 2). See also Figures S3–S5.

S2A). Next, we evaluated the integration and expression of CAR and CAR-EGFP HDRT at the SHSs and observed ~40% higher insertion efficiencies of CAR-EGFP at C7s257_8 (6.5%) compared to ROSA26_3 (Figures 2E, 2F, S2B, and S2C), whereas there were comparable CAR and CAR-EGFP insertion efficiencies of up to 4% into sites C8s325_6, C7s301_5, and CXs313_4. CAR and CAR-GFP have probably faster turnover than sole GFP, as their transient expression is diminished already from day 5 (Figures 2E, S2B, and S2C).

Subsequently, we evaluated SHS-specific transgene insertion and expression efficiencies in Jurkat, and primary T and natural killer (NK) cells. First, we transfected Jurkat cells with 0.5 pmol CAR-EGFP HDRTs and measured protein expression on days 2, 8, and 16 post-EP (Figure 2G). After the transient gene expression subsided on day 8 PT, we observed the highest CAR-EGFP integration efficiencies at ROSA26_3, yielding 28% CAR⁺ and 22.5% GFP⁺ cells. Compared to ROSA26_3, insertion efficiencies at C7s257_8 were lower by ~20%, at C8s325_6 by ~40%, and at CXs313_4 and C7s301_5 by ~70%. Although the transgene expression decreased over time, gene expression at ROSA26_3, C7s257_8, and C8s325_6 stabilized 16 days PT, yielding 18%, 12%, and 10% GFP⁺ cells, respectively. Transfections of Jurkat cells with EGFP HDRT at C7s257_8 and C8s325_6 sites resulted in 15% GFP⁺ cells 20 days PT (Figure S2D).

Next, we transfected primary T cells with 0.25 pmol (0.8 μg) CAR-EGFP HDRTs at C7s257_8 and ROSA26_3. Here, we observed ~17% insertion of CAR-EGFP at ROSA26_3, while at C7s257_8 it did not exceed 12% (Figure 2H). However, at day 14 PT, we observed <40% reduced protein expression at both sites (Figure S2E). Finally, after identifying a suitable EP program (Figure S2F), we transfected primary NK cells with 0.5 pmol CAR-EGFP HDRT at ROSA26_3 and at the four studied SHSs, or CAR at AAVS1_3 (Figure 2I). Again, CAR-EGFP integrated with the highest efficiency at ROSA26_3 (~10% CAR⁺GFP⁺ double-positive cells), followed by C7s257_8 (~5% CAR⁺GFP⁺). Expression of GFP was stable at 14 days PT, while the number of CAR⁺ cells decreased over time. Insertions at the remaining three SHSs occurred with <2.5% efficiencies. CAR was integrated into AAVS1_3 at intermediate efficiency compared to CAR-EGFP efficiencies into other sites, although direct comparison cannot be made. In conclusion, the results indicate that all four sites are potentially suited as SHSs for targeted transgene insertion, but with different applicability depending on the cell type.

Stability of targeted transgene integration during differentiation

Next, we investigated the stability of transgene integration at the SHSs in monoclonal iPSC-derived lineages. Following isolation of GFP⁺ cells by fluorescence-activated cell sorting (FACS) and establishment of monoclonal GFP⁺/CAR⁺GFP⁺ iPSC lines (Figures S3A–S3D), we examined transgene insertion into the targeted sites using an in-out PCR approach⁴⁶ (Figures S5A–S5E); we selected iPSC clones with correct iPSC colony morphology (Figures S3D) and genome integration pattern (Figures S5A–S5E) and verified their non-differentiated status by qPCR measurement of OCT4, SOX2, and NANOG mRNA levels and by immunostaining of OCT4 and SOX2 proteins followed by flow cytometry (Figure S3). The mRNA expression levels of the non-differentiation markers OCT4, SOX2, and NANOG in the selected GFP⁺/CAR⁺GFP⁺ iPSC lines are similar to the parental KOLF2 iPSC (Figure S3E), whereas the mRNA levels of the mesoderm-related transcription factor HANDB1 are very low in the iPSC samples compared to a mesoderm sample (Figure S3F). In addition, we demonstrated the ability of the selected iPSC lines to differentiate to mesoderm, marked by strong reduction of OCT4 and SOX2 expression, measured by flow cytometry (Figures S3A–S3C). Mesoderm induction reduced overall the median GFP fluorescence intensity (MFI), with strongest effect in lines C7s257_8 CAR-EGFP cl. B2 and D4 and CXs313_4 EGFP cl. A8.

Then, we proceeded to differentiation of iPSCs into hematopoietic stem and progenitor cells (HSPCs) following a 2D (2-dimensional) differentiation protocol (Figures 3A and S4A). Non-adherent cells were harvested on day 11–12 of differentiation, and expression of CD34, CD43, and CD45 markers was analyzed by immunostaining and flow cytometry. Analysis of non-engineered iPSC-derived cells revealed up to 96%, 93%, and 50% of CD34⁺, CD43⁺, and CD45⁺ cells, respectively (Figures 3B and 3C), with the median cell viability of 51%. The highest cell viability and expression of markers in the differentiated cells were achieved at the initial iPSC density of 4 colonies/cm², whereas density of 9 colonies/cm² led to reduced viability and the fraction of CD45⁺ cells by 10% (Figure 3C). Similarly, we differentiated iPSC clones with GFP inserted into C7s301_5 (cl. D9), or CXs313_4 (cl. A8), or CAR-GFP inserted into ROSA26_3 (cl. A4) into HSPCs. CD34⁺ cells were up to 80% of the total differentiated cells, CD43⁺ were up to 90%, and CD45⁺ were up to 40% (Figures 3D, 3E, and S4C). Simultaneous analysis of some iPSC lines and their HSPC product showed higher MFI in the iPSC stage (Figure S4B). Of note, although hematopoietic differentiation through embryoid body (EB) formation for 12 days resulted in lower hematopoietic efficiencies and higher variation than the 2D protocol,

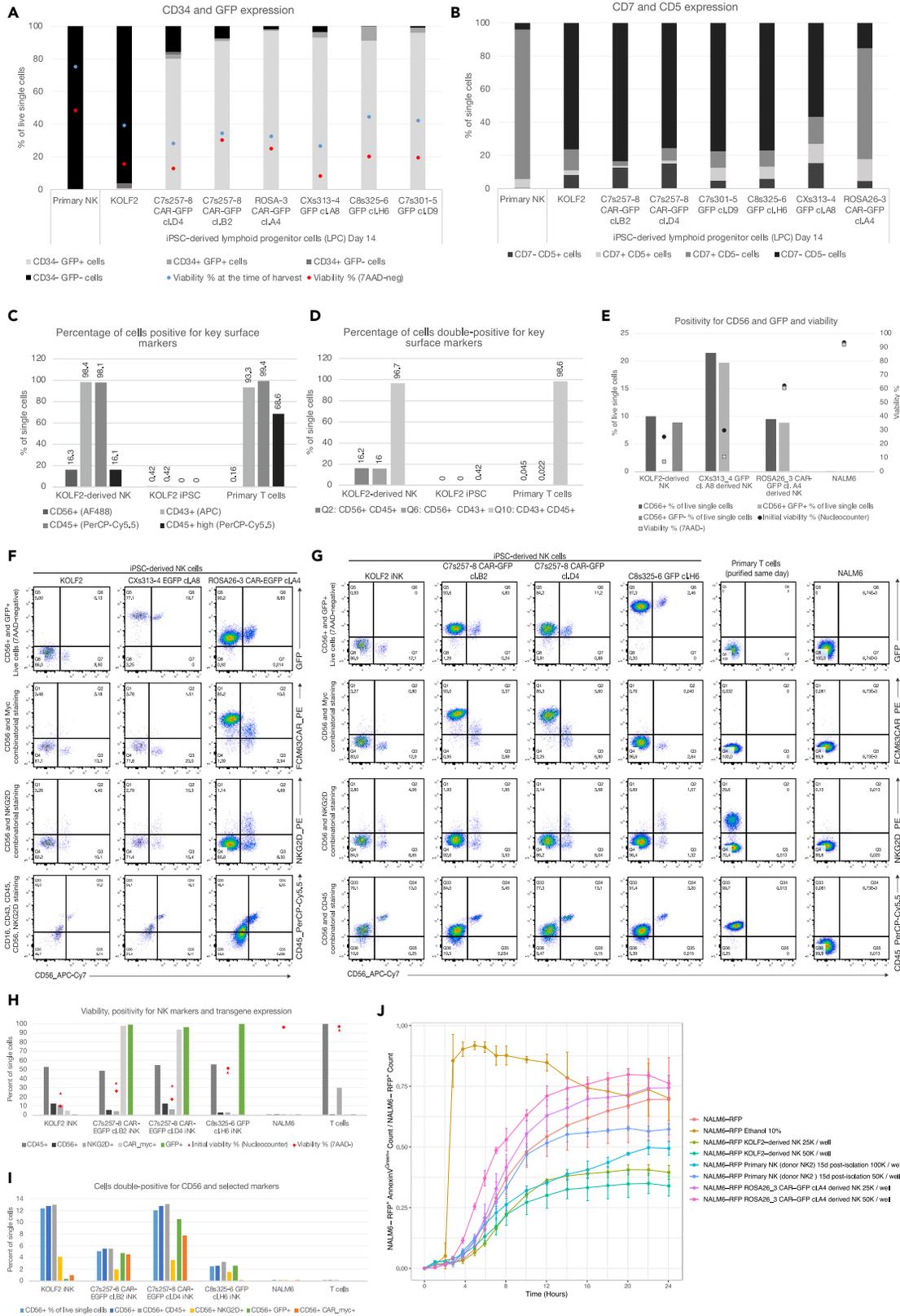


Figure 4. Differentiation of monoclonal iPSC lines to lymphoid progenitor cells (LPCs) and NK cells

(A and B) Parental and edited iPSC clones were differentiated to LPCs through initial CD34⁺ HSPC stage. HSPCs were reseeded and LPCs harvested on day 14 and stained with the pairwise antibody combinations indicated in Figure S6A. Quantification of cell subtypes from pairwise antibody staining. Data are from biological replica (n) = 1 and technical replica (t) = 1 per sample.

(C and D) Parental KOLF2 iPSCs were differentiated to NK cells through the HSPC and LPC stages. The indicated samples obtained relatively sufficient NK number for analysis. Data are related to Figure S6C (n = 1; t = 1).

(E and F) Parental KOLF2 iPSCs and iPSC lines CXs313_4 GFP (cl. A8) and ROSA26_3 CAR-GFP (cl. A4) were differentiated to NK cells which were analyzed by immunostaining and flow cytometry (n = 1; t = 1).

(G) Parental KOLF2 iPSCs and the iPSC clones C7s257 CAR-GFP (cl. B2 and cl. D4) and C8s325 GFP (cl. H6) were initially differentiated to CD34⁺ HSPCs (see Figures 3F and 3G), followed by differentiation to NK cells (abbreviated here iNK) (n = 1; t = 1). Cells were harvested at 29 days and stained with antibodies in pairwise combinations. CD56⁺CD45^{bright} are considered as NK cells.

(H and I) Quantification of cells positive for key markers (n = 1; t = 1). Samples are from same experiment as in panel (G).

(J) Cell apoptosis assay using 10⁴ NALM6-RFP⁺ (expressing red fluorescent protein; RFP) as target cells. The apoptosis reagent IncuCyte Annexin V Green was added (final dilution 1:200) into the target cells 1 h after seeding, followed by immediate addition of the indicated effector cells in different ratios (K = 10³ cells). Effector cells were primary NK cells from donor NK2 (15 days after isolation from blood and kept in culture), KOLF2-derived NK cells and NK derived from iPSC line ROSA26_3 CAR-GFP (cl. A4). As positive control, absolute ethanol was added to 3 replicate wells to a final concentration 10%, 2 h after initial target cell seeding. NK Basal Medium was used for all cell suspensions in the assay. Each treatment consists of 3 technical replica wells. The graph shows the average ratio of NALM6 RFP⁺AnnexinV⁺/NALM6 RFP⁺ and SD of 3 technical replicas (t = 3). The assay was done as one biological replica (n = 1). Cells were imaged by IncuCyte system at 20× magnification by non-adherent cell-by-cell format. See also Figure S6.

judging from the percentage of CD34⁺ cells in the dissociated EB cells, expression of GFP within each line remained unchanged and at similar levels as in the 2D hematopoietic protocol (Figures S5F and S5G).

Afterward, we tested two clones with CAR-EGFP at C7s257_8 (cl. B2 and D4), and one with EGFP at C8s325_6 (cl. H6). KOLF2-derived CD34⁺ HSPCs, NALM6 acute lymphoblastic B cell leukemia cell line, and primary T cells were controls for staining; NALM6 were CD34⁺CD43⁺CD45⁻, and primary T cells were CD34⁺CD43⁺CD45^{bright} (Figures 3F, 3G, and S4D). Optimizing the initial iPSC seeding method resulted in up to 94% CD34⁺ engineered HSPCs (Figures 3F and 3G), which retained the expression of the transgenes (CAR-EGFP or EGFP) (Figures 3D–3G and S4D). In addition, this resulted in 55%–84% CD45⁺ cells and >97% CD43⁺ cells, with median cell viability of 40%. Of note, KOLF2-derived HSPCs that were cryopreserved on day 11 of hematopoietic differentiation, and then thawed and re-cultivated showed a gradual decrease in CD34 marker from 81% at day 6 to 53% at day 12 of re-cultivation, indicating a progressive maturation of iPSC-derived HSPCs (Figures 3D–3G, S4C, and S4D). These data indicate that most iPSC-clone-derived differentiated cells are CD34⁺CD43⁺ HSPCs that retain the expression of the inserted transgene (EGFP or CAR-EGFP).

In summary, we have demonstrated that ROSA26_3, CXs313_4, C7s301_5, C7s257_8 and C8s325_6 sites, all support the expression of transgenes throughout iPSC differentiation to HSPCs.

Differentiation to NK cells through the lymphoid progenitor cell (LPC) stage

We further differentiated iPSC-derived HSPC to NK cells^{47–49} under feeder-free conditions. At day 14 of differentiation, LPCs were harvested for analysis and reseeded. Patterns of expression of the key lymphoid markers CD7 and CD5 revealed that only a small fraction of LPCs, approx. 10%, expressed CD7 alone (CD7⁺CD5⁻), which is a pattern that characterizes mature primary NK cells (Figures 4A, 4B, S6A, and S6B). Similarly, a small fraction of LPCs was either CD5⁺CD7⁻ or CD5⁺CD7⁺, whereas most cells were CD43⁺ that retained the expression of GFP or CAR-GFP, and only an insignificant fraction of cells still expressed the HSPC marker CD34 after 14–15 days of lymphoid differentiation (Figures 4A, 4B, S6A, and S6B).

Then, we further differentiated the non-edited KOLF2-derived LPCs for 14–16 days to NK cells, which were then stained for CD43, CD45, the lineage-specific NK marker CD56, and the stimulatory receptor NKG2D. We observed low differentiation efficiency to NK cells, with up to 16% CD56⁺ from parental KOLF2, albeit 97% of differentiated cells were CD43⁺CD45⁺ (Figures 4D and S6C). CD56⁺ cells expressed CD45 brightly (CD45^{bright}) (Figures 4C and S6C).

Similarly, we differentiated engineered iPSCs with transgenes integrated into CXs313_4 (EGFP) and ROSA26_3 (CAR-EGFP). Overall, we observed low cell viability (mean 28%) and <16% CD56⁺ cells in the product. Nevertheless, all engineered CD56⁺ cells retained GFP expression (Figures 4E, 4F, and S6D). Afterward, clones B2 and D4 with CAR-EGFP at C7s257_8, and clone H6 with EGFP at C8s325_6, were differentiated to NK cells through the HSPC and LPC stages. Again, at the NK level, we observed low cell viability (mean 29.5%), with 12% CD56⁺ cells in clone D4, whereas clones B2 and H6 yielded 5% and 2.5% CD56⁺ cells, respectively (Figures 4G–4I). Moreover, we observed that 30%–50% of CD56⁺ cells expressed NKG2D. Even though the yields of CD56⁺ cells were low, 93%–100% of the CD56⁺ cells expressed GFP from C7s257_8 or C8s325_6. However, we observed reduced levels of CAR expression from C7s257_8 in CD56⁺ cells compared to the CD56⁻ fraction of differentiated cells of the same clone, with only 63%–77% of CD56⁺ cells being CD56⁺CAR⁺ (Figures 4G–4I). Similarly, we observed reduced CAR expression at ROSA26_3, with 88% of CD56⁺ cells being CD56⁺CAR⁺ (Figure 4F). To test whether our CAR transgene is functional, we have performed cell apoptosis assay of ROSA26_3 CAR-GFP NK cells against the CD19⁺ NALM6 line expressing red fluorescent protein (RFP) NALM6-RFP⁺ (Figure 4J). Although, NALM6-RFP undergo apoptosis spontaneously in the medium, in which we performed the assay in order to keep the iPSC-derived NK cells viable, there is a clear increase of apoptotic NALM6-RFP in the presence of NK cells expressing CAR-GFP from ROSA26_3 site, already 4 h after addition of the effector cells, especially at effector to target ratio 5:1 (5 · 10⁴ NKs: 10⁴

NALM6-RFP). Addition of KOLF2 NK or primary NK cells obviously does not promote apoptosis to higher levels than what the NALM6 spontaneously undergo.

Altogether, we demonstrated that integrated transgenes into four chromosomal sites C7s257, C8s325, C7s301, and CXs313 are expressed in diverse cell types (iPSCs, Jurkat, primary T and NK cells), and transgene expression was retained in substantial levels during differentiation from iPSCs, to HSPCs, LPCs, and NK cells. Overall, C7s257 site was active in most tested cell types, whereas insertions into C7s301, C8s325, and CXs313 have not occurred with high efficiencies in T and NK cells. Furthermore, although transgenes might be exposed to silencing during differentiation, we observed that anti-CD19-CAR NK cells derived from ROSA26_3 CAR-GFP iPSCs are active against NALM6.

DISCUSSION

Here, we aimed to discover safe genomic locations in the human genome that enable high transgene expression during cell propagation and differentiation and resist silencing. The extensively used SHS AAVS1 lies within an intron of the protein-coding gene *PPP1R12C* and the promoters ENSR00001026290 and ENSR00000111596, while the ROSA26 site lies within the RNA gene *THUMPD3-AS1* and the promoter ENSR00000148275, which render these sites potentially unsafe for cell therapy applications. Therefore, there is a need for more genomic locations to host transgenes for safer and multiplexed genome engineering, in particular, (1) when integration of synthetic circuits or apoptotic switches into therapeutic cells is required to enable elimination of the transplant,⁵⁰ or (2) when immune cells need to express several receptors that recognize different antigens and fine-tune their activity,⁵¹ or (3) when a receptor consists of several modules.⁵²

We identified four sites, C7s257, C7s301, C8s325, and CXs313, and validated the efficiency of targeting crRNAs. As per the latest human genome browser,⁵³ the crRNA C7s257_8 is targeting an intron of the protein-coding gene *ZC3HAV1* and the promoter flanking region ENSR00001132248. Our results indicate that transgenes inserted into C7s257_8 may undergo silencing over time, as we observed in primary T cells. While C7s301_5 does not overlap with any protein-coding genes or regulatory elements, it is proximal to an H3K9me3 stretch in PSC lines, which is an epigenetic marker of the transcriptionally inactive heterochromatin and could potentially lead to gene silencing. Likewise, C8s325_5 does not overlap with any genes or regulatory elements, but it is 90 bp away from the enhancer ENSR00000857919, which is active in PSC lines, and 3 kb from an H3K27me3 stretch, a marker of facultative heterochromatin. Finally, while CXs313_4 does not overlap with any genes or regulatory elements either, it is surrounded by enhancers ENSR00001158268 and ENSR00001291056 and is proximal to the promoter flanking region ENSR00001158269 that is active in PSC lines. Its location on the X chromosome excludes biallelic insertions in cells from male donors, such as KOLF2 cells, while in female cells, monoallelic insertions could potentially be silenced in some cells due to X chromosome inactivation. These conclusions are summarized in [Table 1](#).

We validated NHEJ editing efficiency in KOLF2, and HDR efficiency in various cell types, using MAD7 together with crRNAs that demonstrate high on-target activity and low probability of off-targets. We achieved ~15% insertion rates into C8s325 and C7s257, and up to 10% into C7s301 and CXs313 in Jurkat. Transgene integration at C7s257 occurred at up to 10% in primary T and NK cells, whereas the other crRNAs were not as efficient. In iPSCs, the overall integration efficiencies were low, possibly due to iPSC resistance to undergo DNA recombination.⁵⁴ Using inhibitors of Rho/ROCK kinase before and after EP to avoid apoptosis led to insertion efficiencies of up to 8%. The NHEJ editing efficiency of crRNAs varied, with ~95% for crROSA26_3 and crAAVS1_3, whereas the efficiencies of crC7s257_8 and crC8s325_6 were approximately 75%. An open chromatin environment at C7s257 may account for the relatively high integration efficiencies and transgene expression levels, both comparable to ROSA26. We observed that by halving the amount of HDRT, the percentage of GFP⁺ cells also dropped by approximately 50% at all sites. However, MFI of GFP largely remained unchanged regardless of the amount of dsDNA ([Figure 2D](#)), which suggests that it is the intrinsic properties of genomic sites that define the levels of expression of the inserted transgenes.

We differentiated non-engineered KOLF2 iPSCs and engineered monoclonal iPSC lines that bear transgene SHS-targeted integrations into CD34⁺ HSPCs at high efficiencies (up to 96%) and achieved further differentiation to NK cells under feeder-free conditions as reported previously,⁴⁸ although at considerably low efficiencies (up to 16%). Further improvements of differentiation protocols are required to obtain better expansions in feeder-free media.^{47,48,55} We achieved expressions of GFP and FMC63CAR in monoclonal iPSC lines that harbor the CAG-driven expression cassette at all sites. Upon differentiation to CD34⁺ HSPCs, both GFP and CAR remained homogeneously and highly expressed in all differentiated HSPCs. Further differentiation of engineered cells with CAR-GFP at C7s257 or ROSA26 to NK resulted in reduced CAR expression in the CD56⁺ cell fraction compared to the CD56⁻ fraction, whereas the levels of GFP were comparable in both populations. Nevertheless, ROSA26_3 CAR⁺ NK cells promoted apoptosis of target NALM6 cells. However, the reduction of CAR expression was more pronounced from C7s257 than from ROSA26. For the remaining sites C7s301, C8s325, and CXs313, we obtained iPSC clones with genome-integrated GFP, whose expression persisted throughout differentiation.

Hematopoietic stem cells (HSCs) are the pivot of hematopoiesis, being the sole multipotent stem cells that can both self-renew and give rise to all blood cell types. Further differentiation leads to multipotent progenitors (MPPs) with no self-renewal ability but full differentiation potential. HSCs and MPPs are collectively called hematopoietic stem and progenitor cells (HSPCs). Finally, lineage-restricted progenitors can only give rise to certain cell types. For example, lymphocytes (B, NK, and T cells) are derived from the common LPCs, which are derived from the lymphoid multipotent progenitors.⁵⁶ Optimizing the production of HSPCs from iPSCs in combination with genome engineering of PSCs can improve the design of off-the-shelf HSPCs and alleviate the need for compatible bone marrow donors for transplantations. Furthermore, CAR-engineered primary or iPSC-derived NK cells can be used as allogeneic immune cells against cancer without entailing the graft-versus-host disease that CAR T cells are more prone to cause.⁵⁷

Table 1. Characteristics of the 5 selected SHSs we tested for CRISPR-assisted integrations

Site ID	Transgene Integration efficiency	Expression of GFP transgene (GFP MFI compared to ROSA26 GFP MFI)	Advantages	Disadvantages	Reference
C4s283_6	1.5% in iPSCs	100% in iPSCs	In open chromatin	crC4s283_6 has specificity score <75%	ENSEMBL Genome Browser GRCh38.p13
C7s257_8	2.5%–4.5% in iPSCs 12%–15% in Jurkat cells 4%–10% in primary T/NK cells	200% in iPSCs	In open chromatin	Within an intron of the protein-coding gene <i>ZC3HAV1</i> and the promoter flanking region ENSR00001132248. <300 kb away from cancer-related genes.	ENSEMBL Genome Browser GRCh38.p13 Pellenz et al. 2019 ³¹
C7s301_5	1.2%–2.2% in iPSCs 5%–10% in Jurkat cells Very low in primary T/NK cells	Up to 100% in iPSCs	No overlap with any protein-coding genes or regulatory elements	Proximal to heterochromatic H3K9me3 in PSC lines, could potentially lead to transgene silencing.	ENSEMBL Genome Browser GRCh38.p13
C8s325_6	1.1%–2.9% in iPSCs 10%–15% in Jurkat cells 1%–2% in primary T/NK cells	200%–250% in iPSCs	No overlap with any protein-coding genes or regulatory elements	90 bp away from the enhancer ENSR00000857919, which is active in PSC lines, and 3 kb from an H3K27me3 stretch, a marker of facultative heterochromatin.	ENSEMBL Genome Browser GRCh38.p13
CXs313_4	0.5%–1% in iPSCs 5%–10% in Jurkat cells Very low in primary T/NK cells	250% in iPSCs	No overlap with any protein-coding genes or regulatory elements	Surrounded by enhancers ENSR00001158268 and ENSR00001291056 and is proximal to the promoter flanking region ENSR00001158269 that is active in PSC lines. <300 kb away from cancer-related genes. On the X chromosome, which excludes biallelic insertions in male-origin cells.	ENSEMBL Genome Browser GRCh38.p13 Pellenz et al. 2019 ³¹

Limitations of the study

Although our results show differences of integration efficiencies into the SHSs among different cell types, additional studies are required to comprehensively elucidate the connection between the observed differences in gene expression and chromatin activity at the studied SHSs among cell types. It is possible that transgenes are transitionally silenced at some of the reported sites during terminal differentiation to NK cells, whereas the persistent high levels of GFP may be the result of slower turnover of GFP protein. Similar reduction of CAR-GFP at C7s257 was observed in edited T and NK cells, during cell culture from day 7 to 14 PT. Using an unstable version of CAG-driven GFP fused to a degron motif⁵⁸ would enable monitoring of expression changes of GFP during differentiation with higher accuracy.³⁵ Although KOLF2 is listed as a healthy control cell line by the Human Induced Pluripotent Stem Cell Initiative (HipSci), a recent whole-genome sequencing analysis revealed variants in four out of five iPSC control lines from HipSci.⁵⁹ In particular, KOLF2 cell line is heterozygous for mutations in two genes, *COL3A1* and *ARID2*, that are likely pathogenic and associated with the autosomal dominant (AD) type IV Ehlers-Danlos syndrome, and Coffin-Siris syndrome, respectively.^{25,59} Therefore, we cannot exclude the possibility of these SNPs affecting differentiation to certain cell types. Finally, we performed a minimal characterization of the FMC63CAR⁺ iPSC-derived NK cells with CAR-EGFP integrated in ROSA26_3, but not in other sites and not in primary T or NK cells. Therefore, additional studies are required to fully characterize CAR integration into the sites we are reporting here in terms of cellular function and expression stability of transgenes.

STAR★METHODS

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

- **KEY RESOURCES TABLE**
- **RESOURCE AVAILABILITY**
 - Lead contact
 - Materials availability
 - Data and code availability
- **EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS**
 - Cell cultures
- **METHOD DETAILS**
 - Design of SHS crRNA library
 - Nuclease expression and purification
 - RNP formation
 - iPSC electroporation (EP)
 - Genomic DNA (gDNA) extraction
 - NGS-primer design and pair-end amplicon sequencing
 - Validation of crRNA INDEL efficiency by NGS
 - Optimization of electroporation conditions for iPSC
 - RNA extraction and RT-qPCR analyses
 - Construction of plasmids carrying the HDR templates
 - Preparation of linear dsDNA HDR templates
 - Fluorescence-activated cell sorting (FACS)
 - Genomic PCR for confirming transgene insertions
 - Electroporation of Jurkat cells
 - Electroporation of primary T cells
 - Electroporation of primary NK cells
 - Optimization of iPSC differentiation to CD34⁺ HSPC
 - iPSC differentiation to CD34⁺ HSPC by the “2D protocol”
 - Hematopoietic differentiation through embryoid bodies (EB)
 - Differentiation of HSPC to lymphoid progenitors and NK cells
 - Flow cytometry
 - Phenotyping of iPSC-derived HSPC, LPC and NK
 - Immunostaining of non-differentiation markers for imaging
 - Immunostaining of non-differentiation markers for flow cytometry
 - Annexin V apoptosis assay by IncuCyte
- **QUANTIFICATION AND STATISTICAL ANALYSIS**
 - NGS data analysis
 - Flow cytometry data analysis
 - Statistical analysis

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.isci.2023.108287>.

ACKNOWLEDGMENTS

We are grateful to Sophia Julia Häfner, Biotech Research and Innovation Center, University of Copenhagen, Denmark, for providing an initial batch of KOLF2 iPSC cells, transferring initial knowledge and providing important guidance on iPSC culture and iPSC gene editing, which was crucial for the initiation of this project. We thank Brian Tate Weinert for fruitful discussions during meetings, Anna Oliver-Almirall and Sara Pereira-Schöning for assistance with cloning of some of the plasmids, Sanne Schoffelen for assistance with the MAD7 protein purifications, and Johanne G. Høyer and Katrine Zeeberg for providing isolated primary T cells, which were used in some of the experiments of the present study, and for assistance during NK-cell purification. We would also like to thank Vijayalakshmi Kandasamy and the other members of DNA Foundry at NNF Center for Biosustainability for guidance on NGS library preparation and QC, and for the operation of the Illumina MiSeq sequencer. The graphical abstract has been created with [BioRender.com](https://www.biorender.com) under a subscription that confers publication rights. This work was mainly supported by the Novo Nordisk Foundation (RTG., NNF17OC0026988). The research in the KNN lab is supported by the Villum Young Investigator grant (VYI #00025397).

AUTHOR CONTRIBUTIONS

A.V. performed most of the CRISPR experiments in iPSCs: designed CRISPR reagents (crRNAs and HDRTs), iPSC electroporation, clonal selection by FACS and differentiation and prepared the manuscript. All experiments were performed at NNF Center for Biosustainability (NNFCB), Technical University of Denmark (DTU). Cloning was performed by A.V. partially at the BioInnovation Institute, Copenhagen, DK, and partially at NNFCB. A.V., occasionally assisted by X.M. and M.M.G., performed plasmid cloning and plasmid preps and isolated linear dsDNA HDRTs for all electroporation experiments. T.L.J. and A.V. performed the optimization of electroporation (Lonza scan) for iPSC. A.V. designed the SHS crRNA library. A.V. and A.B. designed NGS primers and aligned SHS to the genome to find overlaps with transcriptional and regulatory units. A.V. and T.L.J. performed SHS crRNA screens in iPSCs and prepared targeted amplicon NGS libraries for indel detection. M.M. and A.B. performed NGS data analysis by CRISPResso2.⁶⁰ M.M. performed *in silico* off-target analysis for crRNAs. T.L.J. and G.K. purified RNA and carried out RT-qPCR for non-differentiation markers and selected differentiation markers. D.J.J. validated the indel efficiency of selected crRNAs (amplification, purification, indexing, and amplicon library QC) in genomic DNA samples from electroporation experiments carried out by A.V.; A.V., assisted by X.M., performed FACS, maintenance and cryopreservation of iPSC lines, differentiation to HSPCs and NK, and subsequent antibody staining and cytometry. A.V. performed in-out (flanking) PCR for validation of genomic integrations in iPSC clones used for differentiation. A.V. isolated and cultured human primary NK cells with the help of acknowledged colleagues. A.V. electroporated primary NK cells and performed flow cytometry assays. D.J.J. electroporated primary T cells and performed cytometry assays, assisted by M.M. D.J.J., M.M.G., and A.V.; D.J.J. and A.V. electroporated Jurkat cells, and A.V. analyzed the cells by flow cytometry. S.M.A. and A.V. visualized data and made graphs by R/Tidyverse. R.T.G., T.W.L., A.V., and R.F.B. conceived the initial project strategy. K.N.N. provided guidance during the revision. L.K.N. provided guidance on differentiation strategy choice. R.T.G., M.M., L.K.N., and A.V. reviewed and edited the manuscript.

DECLARATION OF INTERESTS

R.T.G., T.W., and A.V. are inventors on a patent that has been filed by the Technical University of Denmark (DTU) and has been licensed to Artisan Bio (patent #WO2023137233A2, published 2023-07-20, based on the U.S. Provisional Patent Application No. 63/300,244, filed January 17, 2022). R.T.G., T.W., A.B., and R.F.B. have financial interests in Artisan Bio. L.K.N. is a member of the advisory board of Artisan Bio.

Received: September 19, 2022

Revised: April 23, 2023

Accepted: October 18, 2023

Published: October 27, 2023

REFERENCES

- Dixon, J.R., Jung, I., Selvaraj, S., Shen, Y., Antosiewicz-Bourget, J.E., Lee, A.Y., Ye, Z., Kim, A., Rajagopal, N., Xie, W., et al. (2015). Chromatin architecture reorganization during stem cell differentiation. *Nature* 518, 331–336. <https://doi.org/10.1038/nature14222>.
- Bernstein, B.E., Mikkelsen, T.S., Xie, X., Kamal, M., Huebert, D.J., Cuff, J., Fry, B., Meissner, A., Wernig, M., Plath, K., et al. (2006). A Bivalent Chromatin Structure Marks Key Developmental Genes in Embryonic Stem Cells. *Cell* 125, 315–326. <https://doi.org/10.1016/j.cell.2006.02.041>.
- ENCODE Project Consortium, Birney, E., Stamatoyannopoulos, J.A., Guigó, R., Guigó, R., Gingeras, T.R., Margulies, E.H., Weng, Z., Snyder, M., Dermitzakis, E.T., et al. (2007). Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. *Nature* 447, 799–816. <https://doi.org/10.1038/nature05874>.
- Hacein-Bey-Abina, S., Von Kalle, C., Schmidt, M., McCormack, M.P., Wulffraat, N., Leboulch, P., Lim, A., Osborne, C.S., Pawliuk, R., Morillon, E., et al. (2003). LMO2-Associated Clonal T Cell Proliferation in Two Patients after Gene Therapy for SCID-X1. *Science* 302, 415–419. <https://doi.org/10.1126/science.1088547>.
- Takahashi, K., and Yamanaka, S. (2006). Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors. *Cell* 126, 663–676. <https://doi.org/10.1016/j.cell.2006.07.024>.

6. Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K., and Yamanaka, S. (2007). Induction of Pluripotent Stem Cells from Adult Human Fibroblasts by Defined Factors. *Cell* 131, 861–872. <https://doi.org/10.1016/j.cell.2007.11.019>.
7. Choi, J., Lee, S., Mallard, W., Clement, K., Tagliazucchi, G.M., Lim, H., Choi, I.Y., Ferrari, F., Tsankov, A.M., Pop, R., et al. (2015). A comparison of genetically matched cell lines reveals the equivalence of human iPSCs and ESCs. *Nat. Biotechnol.* 33, 1173–1181. <https://doi.org/10.1038/nbt.3388>.
8. Sadelain, M., Papapetrou, E.P., and Bushman, F.D. (2011). Safe harbours for the integration of new DNA in the human genome. *Nat. Rev. Cancer* 12, 51–58. <https://doi.org/10.1038/nrc3179>.
9. Papapetrou, E.P., Lee, G., Malani, N., Setty, M., Riviere, I., Tirunagari, L.M.S., Kadota, K., Roth, S.L., Giardina, P., Viale, A., et al. (2011). Genomic safe harbors permit high β -globin transgene expression in thalassemia induced pluripotent stem cells. *Nat. Biotechnol.* 29, 73–78. <https://doi.org/10.1038/nbt.1717>.
10. Hacein-Bey-Abina, S., Garrigue, A., Wang, G.P., Soulier, J., Lim, A., Morillon, E., Clappier, E., Caccavelli, L., Delabesse, E., Beldjord, K., et al. (2008). Insertional oncogenesis in 4 patients after retrovirus-mediated gene therapy of SCID-X1. *J. Clin. Invest.* 118, 3132–3142. <https://doi.org/10.1172/JCI35700>.
11. Papapetrou, E.P., and Schambach, A. (2016). Gene insertion into genomic safe harbors for human gene therapy. *Mol. Ther.* 24, 678–684. <https://doi.org/10.1038/mt.2016.38>.
12. Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A., and Charpentier, E. (2012). A Programmable Dual-RNA-Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science* 337, 816–821. <https://doi.org/10.1126/science.1225829>.
13. Mali, P., Yang, L., Esvelt, K.M., Aach, J., Guell, M., DiCarlo, J.E., Norville, J.E., and Church, G.M. (2013). RNA-guided human genome engineering via Cas9. *Science* 339, 823–826. <https://doi.org/10.1126/science.1232033>.
14. Cong, L., Ran, F.A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P.D., Wu, X., Jiang, W., Marraffini, L.A., and Zhang, F. (2013). Multiplex genome engineering using CRISPR/Cas systems. *Science* 339, 819–823. <https://doi.org/10.1126/science.1231143>.
15. Zetsche, B., Gootenberg, J.S., Abudayyeh, O.O., Slaymaker, I.M., Makarova, K.S., Essletzbichler, P., Volz, S.E., Joung, J., Van Der Oost, J., Regev, A., et al. (2015). Cpf1 Is a Single RNA-Guided Endonuclease of a Class 2 CRISPR-Cas System. *Cell* 163, 759–771. <https://doi.org/10.1016/j.cell.2015.09.038>.
16. Kleinstiver, B.P., Tsai, S.Q., Prew, M.S., Nguyen, N.T., Welch, M.M., Lopez, J.M., McCaw, Z.R., Aryee, M.J., and Joung, J.K. (2016). Genome-wide specificities of CRISPR-Cas Cpf1 nucleases in human cells. *Nat. Biotechnol.* 34, 869–874. <https://doi.org/10.1038/nbt.3620>.
17. Tran, M.H., Park, H., Nobles, C.L., Karunadharma, P., Pan, L., Zhong, G., Wang, H., He, W., Ou, T., Crynen, G., et al. (2021). A more efficient CRISPR-Cas12a variant derived from Lachnospiraceae bacterium MA2020. *Mol. Ther. Nucleic Acids* 24, 40–53. <https://doi.org/10.1016/j.omtn.2021.02.012>.
18. Price, M.A., Cruz, R., Bryson, J., Escalettes, F., and Rosser, S.J. (2020). Expanding and understanding the CRISPR toolbox for *Bacillus subtilis* with MAD7 and dMAD7. *Biotechnol. Bioeng.* 117, 1805–1816. <https://doi.org/10.1002/bit.27312>.
19. Liu, Z., Schiel, J.A., Maksimova, E., Strezoska, Z., Zhao, G., Anderson, E.M., Wu, Y., Warren, J., Bartels, A., Van Brabant Smith, A., et al. (2020). ErCas12a CRISPR-MAD7 for Model Generation in Human Cells, Mice, and Rats. *Cris. J.* 3, 97–108. <https://doi.org/10.1089/crispr.2019.0068>.
20. Diem, M.C. (2021). Engineered Stem Cells Power Expansive Range of Cancer Therapies (Biopharma Deal), p. 4985. <https://www.nature.com/articles/d43747-021-00013-1>.
21. *Inscripta-Aldebron* (2022). *Aldebron Licenses the Manufacturing of a Type-V CRISPR Nuclease from Inscripta*.
22. Kim, S., Kim, D., Cho, S.W., Kim, J., and Kim, J.S. (2014). Highly efficient RNA-guided genome editing in human cells via delivery of purified Cas9 ribonucleoproteins. *Genome Res.* 24, 1012–1019. <https://doi.org/10.1101/gr.171322.113>.
23. Roth, T.L., Puig-Saus, C., Yu, R., Shifrut, E., Carnevale, J., Li, P.J., Hiatt, J., Saco, J., Krystofinski, P., Li, H., et al. (2018). Reprogramming human T cell function and specificity with non-viral genome targeting. *Nature* 559, 405–409. <https://doi.org/10.1038/s41586-018-0326-5>.
24. Bruntraeger, M., Byrne, M., Long, K., and Bassett, A.R. (2019). Editing the Genome of Human Induced Pluripotent Stem Cells Using CRISPR/Cas9 Ribonucleoprotein Complexes. In *Methods in Molecular Biology*, pp. 153–183. https://doi.org/10.1007/978-1-4939-9170-9_11.
25. Skarnes, W.C., Pellegrino, E., and McDonough, J.A. (2019). Improving homology-directed repair efficiency in human stem cells. *Methods* 164–165, 18–28. <https://doi.org/10.1016/j.ymeth.2019.06.016>.
26. Riesenberger, S., and Maricic, T. (2018). Targeting repair pathways with small molecules increases precise genome editing in pluripotent stem cells. *Nat. Commun.* 9, 2164. <https://doi.org/10.1038/s41467-018-04609-7>.
27. Nguyen, D.N., Roth, T.L., Li, P.J., Chen, P.A., Apathy, R., Mamedov, M.R., Vo, L.T., Tobin, V.R., Goodman, D., Shifrut, E., et al. (2020). Polymer-stabilized Cas9 nanoparticles and modified repair templates increase genome editing efficiency. *Nat. Biotechnol.* 38, 44–49. <https://doi.org/10.1038/s41587-019-0325-6>.
28. Costa, M., Dottori, M., Ng, E., Hawes, S.M., Sourris, K., Jamshidi, P., Pera, M.F., Elefanti, A.G., and Stanley, E.G. (2005). The hESC line Envy expresses high levels of GFP in all differentiated progeny. *Nat. Methods* 2, 259–260. <https://doi.org/10.1038/nmeth748>.
29. Smith, J.R., Maguire, S., Davis, L.A., Alexander, M., Yang, F., Chandran, S., Ffrench-Constant, C., and Pedersen, R.A. (2008). Robust, Persistent Transgene Expression in Human Embryonic Stem Cells Is Achieved with AAVS1-Targeted Integration. *Stem Cell.* 26, 496–504. <https://doi.org/10.1634/stemcells.2007-0039>.
30. Liu, Y., Thyagarajan, B., Lakshminath, U., Xue, H., Lieu, P., Fontes, A., MacArthur, C.C., Scheyhing, K., Rao, M.S., and Chesnut, J.D. (2009). Generation of Platform Human Embryonic Stem Cell Lines That Allow Efficient Targeting at a Predetermined Genomic Location. *Stem Cell. Dev.* 18, 1459–1472. <https://doi.org/10.1089/scd.2009.0047>.
31. Pellenz, S., Phelps, M., Tang, W., Hovde, B.T., Sinit, R.B., Fu, W., Li, H., Chen, E., and Monnat, R.J. (2019). New Human Chromosomal Sites with “Safe Harbor” Potential for Targeted Transgene Insertion. *Hum. Gene Ther.* 30, 814–828. <https://doi.org/10.1089/hum.2018.169>.
32. Hockemeyer, D., Soldner, F., Beard, C., Gao, Q., Mitalipova, M., Dekelver, R.C., Katibah, G.E., Amora, R., Boydston, E.A., Zeitler, B., et al. (2009). Efficient targeting of expressed and silent genes in human ESCs and iPSCs using zinc-finger nucleases. *Nat. Biotechnol.* 27, 851–857. <https://doi.org/10.1038/nbt.1562>.
33. Lombardo, A., Genovese, P., Beausejour, C.M., Colleoni, S., Lee, Y.-L., Kim, K.A., Ando, D., Urnov, F.D., Galli, C., Gregory, P.D., et al. (2007). Gene editing in human stem cells using zinc finger nucleases and integrase-defective lentiviral vector delivery. *Nat. Biotechnol.* 25, 1298–1306. <https://doi.org/10.1038/nbt1353>.
34. Lombardo, A., Cesana, D., Genovese, P., Di Stefano, B., Provasi, E., Colombo, D.F., Neri, M., Magnani, Z., Cantore, A., Lo Riso, P., et al. (2011). Site-specific integration and tailoring of cassette design for sustainable gene transfer. *Nat. Methods* 8, 861–869. <https://doi.org/10.1038/nmeth.1674>.
35. Bertero, A., Pawlowski, M., Ortmann, D., Snijders, K., Yiangou, L., Cardoso de Brito, M., Brown, S., Bernard, W.G., Cooper, J.D., Giacomelli, E., et al. (2016). Optimized inducible shRNA and CRISPR/Cas9 platforms for *in vitro* studies of human development using hPSCs. *Development* 143, 4405–4418. <https://doi.org/10.1242/dev.138081>.
36. Kotin, R.M., Linden, R.M., and Berns, K.I. (1992). Characterization of a preferred site on human chromosome 19q for integration of adeno-associated virus DNA by non-homologous recombination. *EMBO J.* 11, 5071–5078. <https://doi.org/10.1002/j.1460-2075.1992.tb05614.x>.
37. Mohr, M., Damas, N., Gudmand-Hoyer, J., Zeeberg, K., Jedrzejczyk, D., Vlassis, A., Morera-Gómez, M., Pereira-Schoning, S., Puš, U., Oliver-Almirall, A., et al. (2023). The CRISPR-Cas12a Platform for Accurate Genome Editing, Gene Disruption, and Efficient Transgene Integration in Human Immune Cells. *ACS Synth. Biol.* 12, 375–389. <https://doi.org/10.1021/acssynbio.2c00179>.
38. Zhang, J., Hu, Y., Yang, J., Li, W., Zhang, M., Wang, Q., Zhang, L., Wei, G., Tian, Y., Zhao, K., et al. (2022). Non-viral, specifically targeted CAR-T cells achieve high safety and efficacy in B-NHL. *Nature* 609, 369–374. <https://doi.org/10.1038/s41586-022-05140-y>.
39. Liu, R., Paxton, W.A., Choe, S., Ceradini, D., Martin, S.R., Horuk, R., MacDonald, M.E., Stuhlmann, H., Koup, R.A., and Landau, N.R. (1996). Homozygous defect in HIV-1 coreceptor accounts for resistance of some multiply-exposed individuals to HIV-1 infection. *Cell* 86, 367–377. [https://doi.org/10.1016/S0092-8674\(00\)80110-5](https://doi.org/10.1016/S0092-8674(00)80110-5).
40. Irion, S., Luche, H., Gadue, P., Fehling, H.J., Kennedy, M., and Keller, G. (2007). Identification and targeting of the ROSA26 locus in human embryonic stem cells. *Nat. Biotechnol.* 25, 1477–1482. <https://doi.org/10.1038/nbt1362>.
41. Aznauryan, E., Yermanos, A., Kinzina, E., Devaux, A., Kapetanovic, E., Milanova, D., Church, G.M., and Reddy, S.T. (2022). Discovery and validation of human genomic safe harbor sites for gene and cell therapies. *Cell Rep. Methods* 2, 100154. <https://doi.org/10.1016/j.crmeth.2021.100154>.

42. Kim, D., Kim, J., Hur, J.K., Been, K.W., Yoon, S.H., and Kim, J.S. (2016). Genome-wide analysis reveals specificities of Cpf1 endonucleases in human cells. *Nat. Biotechnol.* **34**, 863–868. <https://doi.org/10.1038/nbt.3609>.
43. Hsu, P.D., Scott, D.A., Weinstein, J.A., Ran, F.A., Konermann, S., Agarwala, V., Li, Y., Fine, E.J., Wu, X., Shalem, O., et al. (2013). DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat. Biotechnol.* **31**, 827–832. <https://doi.org/10.1038/nbt.2647>.
44. Nicholson, I.C., Lenton, K.A., Little, D.J., Decorso, T., Lee, F.T., Scott, A.M., Zola, H., and Hohmann, A.W. (1997). Construction and characterisation of a functional CD19 specific single chain Fv fragment for immunotherapy of B lineage leukaemia and lymphoma. *Mol. Immunol.* **34**, 1157–1165. [https://doi.org/10.1016/S0161-5890\(97\)00144-2](https://doi.org/10.1016/S0161-5890(97)00144-2).
45. Kochenderfer, J.N., Feldman, S.A., Zhao, Y., Xu, H., Black, M.A., Morgan, R.A., Wilson, W.H., and Rosenberg, S.A. (2009). Construction and preclinical evaluation of an anti-CD19 chimeric antigen receptor. *J. Immunother.* **32**, 689–702. <https://doi.org/10.1097/CJI.0b013e3181ac6138>.
46. Bak, R.O., Dever, D.P., and Porteus, M.H. (2018). CRISPR/Cas9 genome editing in human hematopoietic stem cells. *Nat. Protoc.* **13**, 358–376. <https://doi.org/10.1038/nprot.2017.143>.
47. Cichocki, F., Bjordahl, R., Gaidarova, S., Mahmood, S., Abujarour, R., Wang, H., Tuininga, K., Felices, M., Davis, Z.B., Bendzick, L., et al. (2020). iPSC-derived NK cells maintain high cytotoxicity and enhance *in vivo* tumor control in concert with T cells and anti-PD-1 therapy. *Sci. Transl. Med.* **12**, eaaz5618. <https://doi.org/10.1126/scitranslmed.aaz5618>.
48. Lupo, K.B., Moon, J.-I., Chambers, A.M., and Matosevic, S. (2021). Differentiation of natural killer cells from induced pluripotent stem cells under defined, serum- and feeder-free conditions. *Cytotherapy* **23**, 939–952. <https://doi.org/10.1016/j.jcyt.2021.05.001>.
49. Knorr, D.A., Ni, Z., Hermanson, D., Hexum, M.K., Bendzick, L., Cooper, L.J.N., Lee, D.A., and Kaufman, D.S. (2013). Clinical-Scale Derivation of Natural Killer Cells From Human Pluripotent Stem Cells for Cancer Therapy. *Stem Cells Transl. Med.* **2**, 274–283. <https://doi.org/10.5966/sctm.2012-0084>.
50. Labanieh, L., Majzner, R.G., and Mackall, C.L. (2018). Programming CAR-T cells to kill cancer. *Nat. Biomed. Eng.* **2**, 377–391. <https://doi.org/10.1038/s41551-018-0235-9>.
51. MacKay, M., Afshinnekoo, E., Rub, J., Hassan, C., Khunte, M., Baskaran, N., Owens, B., Liu, L., Roboz, G.J., Guzman, M.L., et al. (2020). The therapeutic landscape for cells engineered with chimeric antigen receptors. *Nat. Biotechnol.* **38**, 233–244. <https://doi.org/10.1038/s41587-019-0329-2>.
52. Cho, J.H., Collins, J.J., and Wong, W.W. (2018). Universal Chimeric Antigen Receptors for Multiplexed and Logical Control of T Cell Responses. *Cell* **173**, 1426–1438.e11. <https://doi.org/10.1016/j.cell.2018.03.038>.
53. ENSEMBL (2022). *Genome Assembly: GRCh38.p13*.
54. Zwaka, T.P., and Thomson, J.A. (2003). Homologous recombination in human embryonic stem cells. *Nat. Biotechnol.* **21**, 319–321. <https://doi.org/10.1038/nbt788>.
55. Li, Y., Hermanson, D.L., Moriarity, B.S., and Kaufman, D.S. (2018). Human iPSC-Derived Natural Killer Cells Engineered with Chimeric Antigen Receptors Enhance Anti-tumor Activity. *Cell Stem Cell* **23**, 181–192.e5. <https://doi.org/10.1016/j.stem.2018.06.002>.
56. Cedar, H., and Bergman, Y. (2011). Epigenetics of haematopoietic cell development. *Nat. Rev. Immunol.* **11**, 478–488. <https://doi.org/10.1038/nri2991>.
57. Liu, E., Marin, D., Banerjee, P., Macapinlac, H.A., Thompson, P., Basar, R., Nassif Kerbauy, L., Overman, B., Thall, P., Kaplan, M., et al. (2020). Use of CAR-Transduced Natural Killer Cells in CD19-Positive Lymphoid Tumors. *N. Engl. J. Med.* **382**, 545–553. <https://doi.org/10.1056/nejmoa1910607>.
58. Li, X., Zhao, X., Fang, Y., Jiang, X., Duong, T., Fan, C., Huang, C.-C., and Kain, S.R. (1998). Generation of Destabilized Green Fluorescent Protein as a Transcription Reporter. *J. Biol. Chem.* **273**, 34970–34975. <https://doi.org/10.1074/jbc.273.52.34970>.
59. Hildebrandt, M.R., Reuter, M.S., Wei, W., Tayebi, N., Liu, J., Sharmin, S., Mulder, J., Lesperance, L.S., Brauer, P.M., Mok, R.S.F., et al. (2019). Precision Health Resource of Control iPSC Lines for Versatile Multilineage Differentiation. *Stem Cell Rep.* **13**, 1126–1141. <https://doi.org/10.1016/j.stemcr.2019.11.003>.
60. Clement, K., Rees, H., Canver, M., Gehrke, J.M., Farouni, R., Hsu, J.Y., Cole, M.A., Liu, D.R., Joung, K.J., Bauer, D.E., et al. (2019). CRISPResso2 provides accurate and rapid genome editing analysis. *Nat. Biotechnol.* **37**, 224–226. <https://doi.org/10.1038/s41587-019-0032-3>.
61. Kilpinen, H., Goncalves, A., Leha, A., Afzal, V., Alasoo, K., Ashford, S., Bala, S., Bensaddek, D., Casale, F.P., Culley, O.J., et al. (2017). Common genetic variation drives molecular heterogeneity in human iPSCs. *Nature* **546**, 370–375. <https://doi.org/10.1038/nature22403>.
62. Wagner, J., Pfannenstiel, V., Waldmann, A., Bergs, J.W.J., Brill, B., Huenecke, S., Klingebiel, T., Rödel, F., Buchholz, C.J., Wels, W.S., et al. (2017). A two-phase expansion protocol combining interleukin (IL)-15 and IL-21 improves natural killer cell proliferation and cytotoxicity against rhabdomyosarcoma. *Front. Immunol.* **8**, 676. <https://doi.org/10.3389/fimmu.2017.00676>.
63. Magee, M., Vujanovic, L.N., Butterfield, L.H., and Vujanovic, N.L. (2010). Isolation, culture and propagation of natural killer cells. In *Natural Killer Cells* (Elsevier), pp. 125–135. <https://doi.org/10.1016/B978-0-12-370454-2.00009-0>.
64. Jedrzejczyk, D.J., Poulsen, L.D., Mohr, M., Damas, N.D., Schoffelen, S., Barghetti, A., Baumgartner, R., Weinert, B.T., Warnecke, T., and Gill, R.T. (2022). CRISPR-Cas12a nucleases function with structurally engineered crRNAs: Synthetic trAcrRNA. *Sci. Rep.* **12**, 12193. <https://doi.org/10.1038/s41598-022-15388-z>.
65. Rojek, J.B., Basavaraju, Y., Nallapareddy, S., Bulté, D.B., Baumgartner, R., Schoffelen, S., Grav, L.M., Goletz, S., and Pedersen, L.E. (2023). Expanding the CRISPR toolbox for Chinese hamster ovary cells with comprehensive tools for Mad7 genome editing. *Biotechnol. Bioeng.* **120**, 1478–1491. <https://doi.org/10.1002/bit.28367>.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
<i>Antibodies</i>		
Mouse anti-human CD34, clone 8G12, Brilliant Violet 786	BD Biosciences	Cat#744741; RRID:AB_2742449
Mouse anti-human CD45, clone HI30, Pacific Blue (PB450)	STEMCELL Technologies	Cat#60018PB
Mouse anti-human CD45, clone HI30, PerCP-Cy5.5	STEMCELL Technologies	Cat#60018PS.1
Mouse anti-human CD45, clone HI30, PerCP-Cy5.5	BD Biosciences	Cat#564105; RRID:AB_2744405
Mouse anti-human CD45, clone HI30, PerCP	BioLegend	Cat#304026; RRID:AB_893337
Mouse anti-human CD43, clone CD43-10G7, APC	STEMCELL Technologies	Cat#60085AZ
Mouse anti-human CD5, clone UCHT2, Alexa Fluor 700	BioLegend	Cat#300631; RRID:AB_2632670
Mouse anti-human CD7, clone CD7-6B7, PE	BioLegend	Cat#343106; RRID:AB_1732011
Mouse anti-human CD38, clone HIT2, PerCP	BioLegend	Cat#303519; RRID:AB_893315
Mouse anti-human CD16, clone 3G8, Brilliant Violet 421	BioLegend	Cat#302038; RRID:AB_2561578
Mouse anti-human CD56 (NCAM), clone HCD56, Alexa Fluor 488	BioLegend	Cat#318312; RRID:AB_604102
Mouse anti-human CD56 (NCAM), clone HCD56, APC/Cyanine7	BioLegend	Cat#318332; RRID:AB_10896424
Mouse anti-Human CD314 (NKG2D), clone 1D11, PE	BD Biosciences	Cat#557940; RRID:AB_396951
Mouse anti-Myc tag antibody, clone 9E10, PE	Abcam	Cat#ab72468; RRID:AB_2148456
Mouse anti-human Oct3/4 (POU5F1), clone 40/Oct-3, Alexa Fluor 555	BD Biosciences	Cat#560306; RRID:AB_1645312
Mouse anti-human SOX2, clone O30-678, Alexa Fluor 647	BD Biosciences	Cat#562139; RRID:AB_10897844
Mouse Anti-human SSEA-4, clone MC-813-70, non-conjugated	STEMCELL Technologies	Cat#60062; RRID:AB_2721031
Mouse anti-human OCT4 (OCT3), clone 3A2A20, non-conjugated	STEMCELL Technologies	Cat#60093; RRID:AB_2801346
Goat anti-mouse IgG3 cross-adsorbed secondary antibody, polyclonal, Alexa Fluor 488	Thermo Fisher Scientific	Cat#A-21151; RRID:AB_2535784
Goat anti-mouse IgG2b cross-adsorbed secondary antibody, polyclonal, Alexa Fluor 568	Thermo Fisher Scientific	Cat#A-21144; RRID:AB_2535780
Mouse IgG3 κ , Isotype control, clone MG3-35	STEMCELL Technologies	Cat#60073.1
Mouse IgG2b κ , Isotype control, clone MPC-11	STEMCELL Technologies	Cat#60072.1

(Continued on next page)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Bacterial and virus strains		
TOP10 chemically competent <i>E.coli</i>	Thermo Fisher Scientific	Cat#C404003
One Shot Mach1 T1 Phage-Resistant Chemically Competent <i>E.coli</i>	Thermo Fisher Scientific	Cat#C862003
Biological samples		
Sample NK5: Human Peripheral Blood NK Cells by negative selection; 5x10 ⁶ cells/cryovial	STEMCELL Technologies	Cat#70036
Chemicals, peptides, and recombinant proteins		
Phusion Hot Start II High-Fidelity PCR Master Mix	Thermo Fisher Scientific	Cat#F565S
IDTE pH 7.5 (1X TE Solution)	Integrated DNA Technologies (IDT)	Cat#11-05-01-15
Matrigel hESC-Qualified Matrix, LDEV-free	Corning	Cat#354277
DMEM/F-12	Thermo Fisher Scientific	Cat#31331028 or #11039021
TeSR-E8	STEMCELL Technologies	Cat#05990
mTeSR-Plus	STEMCELL Technologies	Cat#05825; New Cat#100-0276
Y-27632 (RHO/ROCK-pathway inhibitor)	STEMCELL Technologies	Cat#72304
DMSO (Hybri-Max)	Merck-Sigma	Cat#D2650-5X5ML
DPBS, without Mg ²⁺ /Ca ²⁺	Merck-Sigma	Cat#D8537-500ML
ReLeSR	STEMCELL Technologies	Cat#05872
Accutase	STEMCELL Technologies	Cat#07920
CloneR	STEMCELL Technologies	Cat#05888 or #05889
KnockOut Serum-Replacement	Thermo Fisher Scientific	Cat#10828028
PGA; Poly-L-glutamic acid sodium salt, x=200, MW=30,000kDa	Alamanda-Polymers	Cat#CAS#26247-79-0
TrypLE-Select	Thermo Fisher Scientific	Cat#12563-011
MycoAlert PLUS	Lonza	Cat#LT07-705
QuickExtract (QE) Solution	Lucigen	Cat#QE09050
Nuclease-free water (not DEPC-treated)	Thermo Fisher Scientific	Cat#AM9938
AMPure XP beads	Beckman-Coulter	Cat#A63880
KAPA HiFi-Hot-Start polymerase Ready-Mix	Roche	Cat#0795893500
InFusion HD cloning kit	TaKaRa Bio	Car#639650
InFusion Snap-Assembly kit	TaKaRa Bio	Cat#638948
HiSpeed plasmid maxiprep kit	Qiagen	Cat#12663
Sodium acetate 3M pH 5.2	Merck-Sigma	Cat#567422-100ML
CryoStor CS10	STEMCELL Technologies	Cat#07930
FreSR-S (cryopreservation medium for stem cells as single cells)	STEMCELL Technologies	Cat#05859
M3814 (Nedisertib; DNA-PK inhibitor)	SelleckChem	Cat#S8586
UM729	STEMCELL Technologies	Cat#72332
DAPI dilactate	Thermo Fisher Scientific	Cat#D3571
ImmunoCult-XF T-Cell Medium	STEMCELL Technologies	Cat#10981
Human AB serum (heat-inactivated at 56°C for 30 minutes)	Merck-Sigma	Cat#H4522-100ML
IL-2	STEMCELL Technologies	Cat#78036.3
IL-15	STEMCELL Technologies	Cat#78031.1
IL-21	STEMCELL Technologies	Cat#78082

(Continued on next page)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
IL-7	STEMCELL Technologies	Cat#78053.1
RPMI 1640 Medium (ATCC modification)	Thermo Fisher Scientific	Cat#A1049101
Fetal Bovine Serum, qualified, heat inactivated	Thermo Fisher Scientific	Cat#16140071
Penicillin-Streptomycin (10,000 U/ml)	Thermo Fisher Scientific	Cat#15140122
BbvCI restriction enzyme	New England Biolabs	Cat#R0601L
XhoI restriction enzyme	New England Biolabs	Cat#R0146L
EcoRI-HF restriction enzyme	New England Biolabs	Cat#R3101S
PvuII-HF restriction enzyme	New England Biolabs	Cat#R3150L
FspI restriction enzyme	New England Biolabs	Cat#R0135L
PvuII-HF	New England Biolabs	Cat#R3151S
SeaPlaque GTG Agarose	Lonza	Cat#50110
SYBR Safe DNA gel stain. 10000X concentrate in DMSO	Thermo Fisher Scientific	Cat#S33102
NucleoSpin gel and PCR clean-up kit	Macherey-Nagel	Cat#740609
QIAprep-Spin Miniprep kit	Qiagen	Cat#27106
4% formaldehyde solution in DPBS	Merck-Sigma	Cat#252549-500ML
Critical commercial assays		
P3 Primary Cell 96-well Nucleofector Kit - 1x96-well plate or 10x96-well plates	Lonza	Cat#V4SP-3096 or #V4SP-3960
Primary Cell Nucleofector Optimization Kit	Lonza	Cat#V4SP-9096
Nextera XT Index Kit v2	Illumina	Cat# Set-A: FC-131-2001, Set-B: FC-131-2002, Set-C: FC-131-2003, or Set-D: FC-131-2004
MiSeq Reagent Kit v2, 300 Cycles	Illumina	Cat#MS-102-2022
MiSeq Reagent Kit v2, 500 Cycles	Illumina	Cat#MS-102-2003
Qubit dsDNA-High-Sensitivity	Thermo Fisher Scientific	Cat#Q32854
Qubit ssDNA Assay	Thermo Fisher Scientific	Cat#Q10212
RNeasy Mini RNA-purification kit	Qiagen	Cat#74104
RNeasy Micro RNA-purification kit	Qiagen	Cat#74004
ProtoScript First-Strand cDNA-Synthesis	New England Biolabs	Cat#E6300S
High-Capacity cDNA Reverse Transcription Kit	Applied Biosystems	Cat#4368814
PowerUp SYBR-Green Mastermix	Thermo Fisher Scientific	Cat#A25742
STEMdiff Hematopoietic differentiation kit	STEMCELL Technologies	Cat#05310
STEMdiff NK differentiation kit	STEMCELL Technologies	Cat#100-0170
StemSpan NK generation kit	STEMCELL Technologies	Cat#09960
Zombie Violet	BioLegend	Cat#423113
7-AAD (7-Aminoactinomycin D)	Thermo Fisher Scientific	Cat#A1310
IncuCyte Annexin-V Green reagent for apoptosis	Sartorius	Cat#4642
EasySep NK-isolation kit	STEMCELL Technologies	Cat#17955
EasySep T-cell-isolation kit	STEMCELL Technologies	Cat#17951
ImmunoCult human CD3/CD28/CD2 T-Cell-Activator	STEMCELL Technologies	Cat#10990
SF Cell Line 96-well Nucleofector Kit	Lonza	Cat#V4SC-2096 or #V4SC-2960
CytoFLEX daily QC beads	Beckman Coulter	Cat#B53230
Via1-Cassettes	Chemometec	Cat#941-0012

(Continued on next page)

<i>Continued</i>		
REAGENT or RESOURCE	SOURCE	IDENTIFIER
<i>Deposited data</i>		
NGS data of genome editing using SHS crRNA in complex with MAD7; generated in this study	European Nucleotide Archive (ENA)	Accession Number ENA: PRJEB67344
Sequences and annotation of plasmids generated in this study	Zenodo	Zenodo: https://doi.org/10.5281/zenodo.8374369
<i>Experimental models: Cell lines</i>		
Human induced pluripotent cell line KOLF2	HipSci; Wellcome Trust Sanger Institute; Hinxton; United Kingdom.	HPSI0114i-kolf_2; RRID:CVCL_AE29
Human induced pluripotent cell line KOLF2, clone C1	HipSci; Wellcome Trust Sanger Institute; Hinxton; United Kingdom.	HPSI0114i-kolf_2-C1; RRID:CVCL_9S58
NALM-6, Adult B-cell acute lymphoblastic leukemia line	DSMZ-German Collection of Microorganisms and Cell Cultures GmbH	NALM-6; DSMZ ACC-128; RRID:CVCL_0092
Jurkat, Childhood T-cell acute lymphoblastic leukemia line	DSMZ-German Collection of Microorganisms and Cell Cultures GmbH	Jurkat; DSMZ ACC 282; RRID:CVCL_0065
<i>Oligonucleotides</i>		
OCT4 qPCR forward primer	IDT	CGAAAGAGAAAGCGAACCCAG
OCT4 qPCR reverse primer	IDT	AACCACACTCGGACCACATC
SOX2 qPCR forward primer	IDT	ACACCAATCCCATCCACACT
SOX2 qPCR reverse primer	IDT	CCTCCCAGGTTTTCTCTGT
NANOG qPCR forward primer	IDT	CAAAGGCAAACAACCCACTT
NANOG qPCR reverse primer	IDT	TCTGCTGGAGGCTGAGGTAT
GAPDH qPCR forward primer	IDT	AACGGATTGGTCTGATTGG
GAPDH qPCR reverse primer	IDT	CTTCCCGTTCTCAGCCTTG
<i>Recombinant DNA</i>		
Sequences and annotation of plasmids generated in this study	Zenodo	https://doi.org/10.5281/zenodo.8374369
Gene Fragments for plasmid assembly	Twist Bioscience	https://www.twistbioscience.com/products/genes?tab=fragment
gBlocks for plasmid assembly	Integrated DNA Technologies (IDT)	https://eu.idtdna.com/site/order/gblockentry
<i>Software and algorithms</i>		
Benchling	Biology Software (2019-2022)	Retrieved from https://biosustain.benchling.com/ ; RRID:SCR_013955
ggplot2	https://github.com/tidyverse/ggplot2	ggplot2 (RRID:SCR_014601)
CRISPResso2	https://github.com/pinellolab/CRISPResso2	http://crispresso2.pinellolab.org ; RRID:SCR_024503
FlowJo	BD Biosciences	https://www.flowjo.com/solutions/flowjo/ ; RRID:SCR_008520
<i>Other</i>		
Lonza 4D Nucleofector with 96-well Shuttle Add-on	Lonza	Core Unit: AAF-1002B; X Unit: AAF-1002X; 96-well Shuttle Add-on: AAM-1001S
IncuCyte S3 Live-Cell Analysis System	Essen Bioscience (Sartorius)	https://www.sartorius.com/en/products/live-cell-imaging-analysis/live-cell-analysis-instruments/s3-live-cell-analysis-instrument ; RRID:SCR_023147
CytoFLEX S V4-B2-Y4-R3 Flow Cytometer	Beckman Coulter	https://www.beckman.com/flow-cytometry/research-flow-cytometers/cytoflex-s/c09766

(Continued on next page)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
QuantStudio5	Applied-Biosystems	qPCR instrument
CFX96 C1000 Touch Thermal Cycler	BioRad	qPCR instrument
MiSeq	Illumina	https://www.illumina.com/systems/sequencing-platforms/miseq.html ; RRID:SCR_016379
Registration information for KOLF2 iPSC (HPSI0114i-kolf_2)	Human Pluripotent Stem Cell Registry	https://hpscereg.eu/cell-line/WTSIi018-B
Registration information for KOLF2-C1 iPSC (HPSI0114i-kolf_2-C1)	Human Pluripotent Stem Cell Registry	https://hpscereg.eu/cell-line/WTSIi018-B-1

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Ryan T. Gill (rtg@artisancells.com).

Materials availability

Plasmids sequences have been deposited at the general-purpose repository Zenodo and are publicly available as of the date of publication. Zenodo: <https://doi.org/10.5281/zenodo.8374369>.

Plasmids and cell lines generated in this study are registered and stored at the NNF Center for Biosustainability. Requests for these materials should be directed to and will be fulfilled by the [lead contact](#), Ryan T. Gill.

Data and code availability

NGS data

NGS data of genome editing using SHS crRNA in complex with MAD7, have been deposited at the European Nucleotide Archive (ENA) with the Accession Number: ENA: PRJEB67344, which is also listed in the [key resources table](#).

Plasmid sequences

The sequences of plasmids generated in this study have been deposited at Zenodo and are publicly available as of the date of publication: Zenodo: <https://doi.org/10.5281/zenodo.8374369>. DOI also listed in the [key resources table](#).

Code

This paper does not report original code.

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Cell cultures

Feeder-free iPSC cultures

Few frozen vials of the clone KOLF2-C1 were kindly provided by Dr. Sophia Julia Häfner (Biotech Research and Innovation Centre, University of Copenhagen, Denmark) and were used in initial experiments from March to November 2019. Later on, we used the iPSC line KOLF2 (HPSI0114i-kolf_2)^{24,25} which was derived by reprogramming skin fibroblasts from a male individual. HPSI0114i-kolf_2 was generated at Wellcome Trust Sanger Institute, UK, as part of the Human iPSC Initiative (HipSci) project.⁶¹ The registration summary of the HPSI0114i-kolf_2 can be found at <https://hpscereg.eu/cell-line/WTSIi018-B>. We expanded the polyclonal HPSI0114i-kolf_2 (passage P21 upon receipt, cryopreserved 11.02.2019) and cryopreserved the passages P24 and P25, which we used for all other experiments (see “KOLF2 initial thawing, expansion, and cryopreservation” section of *STAR Methods*). Before seeding, tissue-culture (TC)-treated vessels were coated with Matrigel (Corning #354277), diluted at a batch-specific ratio (provided by the manufacturer) in ice-cold DMEM/F-12 (Thermo-Fisher #31331028 or #11039021) and placed at 37°C for 0.5-1 hour to allow polymerization. Coated culture vessels were stored at 4°C for ≤ 1 week and before use, they were let reach room temperature (RT) for ≥ 15 min. For cultivation, TeSR-E8 (STEMCELL #05990) was initially used, based on Skarnes et al. 2019²⁵ and medium was changed daily. After March 2020, TeSR-E8 was substituted by mTeSR-Plus, a stabilized serum-free medium (STEMCELL #05825, new catalog number #100-0276) after a passaging step with ReLeSR, and mTeSR-Plus was changed every 2-3 days. The culture medium was supplemented with Y-27632 (RHO/ROCK-pathway inhibitor) (STEMCELL #72304) at 10 μM (from 10 mM stock in DMSO (Sigma #D2650-5X5ML) for 24 hrs (or 48 hrs when confluence remained low) every time upon thawing, and each time the iPSC culture was dissociated to single cells. For routine passaging of iPSC as aggregates, cells were washed with DPBS without Mg²⁺/Ca²⁺ (Thermo-Fisher #14190144, #14190094, or Merck-Sigma #D8537-500ML), then non-enzymatic reagent ReLeSR (STEMCELL #05872) 105 μl/cm² (1 ml/well of 6-well plate, 8 ml/T75 flask) was added for 1 min at RT followed by aspiration of 80-90% of its volume. Thereafter, the thin layer of ReLeSR was let work 4 min at RT for TeSR-E8, and 6 min RT for mTeSR-Plus cultures. Then, culture medium was added at 105 μl/cm² (1 ml/well of 6-well plate, 8 ml/T75) and iPSC aggregates of the correct diameter 50-200 μm were created by tapping for 30-60 seconds. Suspension was spun at 120g for 3 min, aggregate pellet was dislodged gently and resuspended in culture medium

without Y-27632 and fractions of the volume were seeded on Matrigel-coated vessels in medium without Y-27632, creating several splitting ratios. In general, Y-27632 was not used during routine passaging with ReLeSR. Optimal splitting ratio for KOLF-2 from a culture with confluence 75-80% is 1:6-1:9. For single-cells procedures, iPSC were washed with DPBS and dissociated with TrypLE-Select (Thermo-Fisher #12563-011) or Accutase (STEMCELL #07920), rinsed 1:10 with culture medium to remove excess enzyme, and cultured in medium with 10 μ M Y-27632. In latest experiments, we used CloneR (STEMCELL #05888) to enhance survival before and after electroporation or sorting (FACS). iPSC culture media contained no antibiotics, except during and after FACS until clone cryopreservation. All handling of iPSC was conducted under sterile conditions in BSC-II (biosafety cabinets type-II). TC-treated vessels used for iPSC cultures: 25cm² (T25; Greiner #690175) and 75cm² flasks (T75; Greiner #658175) with filter lid, 96-well plates (Corning #3595), 24-well (Corning #3524), 12-well (Corning #3513), 6-well (Greiner #657160 or Corning #3516).

KOLF2 initial thawing, expansion, and cryopreservation

KOLF2 polyclonal iPSC (Culture Collections; Public Health England; Line-ID: HPSI0114i-kolf_2, Biosample-ID: SAMEA2547615, Acc#77650100, Lot#19A018, passage P21 upon receipt, cryopreserved 11.02.2019), were thawed in water-bath and cultured in two wells of a Matrigel-coated 6-well plate in TeSR-E8 containing 10 μ M Y-27632 for 2 days. Cells were dissociated with ReLeSR and expanded into T75 flasks on day 4 and on day 8. Growth was monitored continuously by IncuCyte S3 Live-Cell Analysis System (Essen-BioScience). KOLF2 of passage number P24 (P_{initial}+3) were cryopreserved in barcoded Matrix cryotubes (Thermo-Fisher #3741) as aggregates using cryopreservation medium: 90% v/v KnockOut-Serum-Replacement (Thermo-Fisher #10828028), 10% v/v DMSO (Sigma #D2650-5X5ML). Cells were tested and found negative for Mycoplasma at the time of banking (MycoAlert-PLUS, Lonza #LT07-705). Thawing of each cryovial of banked KOLF2 was done into two wells of a 6-well plate in TeSR-E8 medium supplemented with 10 μ M Y-27632 for the first 24hrs. In latest experiments, after the first passaging with ReLeSR, aggregates were transitioned to mTeSR-Plus medium.

Jurkat and NALM6 cell culture

Jurkat human T-cell leukemia cell-line (DSMZ #ACC282) was cultured in RPMI-1640 medium ATCC-modification (Thermo-Fischer #A1049101) with 10% heat-inactivated fetal-bovine-serum (FBS) (Thermo-Fischer #10500064) supplemented with 1% (100 U/ml) penicillin-streptomycin (Thermo-Fischer #15140122). Cells were cultured at 37°C in 5% CO₂ incubators and maintained at a density of 0.5-1.5×10⁶ cells/ml. 24 hrs before electroporation, we passaged at 0.1×10⁶ cells/ml. Cell culture media supernatant was periodically tested for mycoplasma contamination. NALM6 (DSMZ #ACC-128) were cultured in the same medium as described above for Jurkat cells, except that cell density was maintained at 1-2 × 10⁶ cells/ml and cells were split at ratios 1:2 to 1:3 every 3 days.

Isolation and culture of primary T cells

Research with human primary T-cells was performed in accordance with the Declaration of Helsinki. Human peripheral blood was obtained from healthy adults after informed consent (Technical University of Denmark DTU-Rigshospitalet National Hospital approval BC-40). No personal information regarding the donors has been disclosed to us. T cells were isolated from the PBMC population by immunomagnetic negative selection using the EasySep T-cell-isolation kit (STEMCELL #17951), then activated with 25 μ l ImmunoCult hCD3/CD28/CD2 T-Cell-Activator (STEMCELL #10990) per ml of ImmunoCult-XF T-Cell Medium (STEMCELL #10981) containing 12.5 ng/ml IL-2, 5 ng/ml IL-7, and 5 ng/ml IL-15 (STEMCELL #78036.3, 78053.1, 78031.1, respectively) and seeded at 1.0×10⁶ cells/ml until transfection 48 hrs later. Cells were maintained at 37°C in 5% CO₂ incubators. We used primary T cells isolated from two donors (donors T1 and T2) for CAG-CAR-EGFP transgene insertions in SHS shown in [Figures 2H](#) and [S2E](#). Samples from T cell cultures frequently prepared in our lab from different donors were used as negative controls during immunostaining of HSPC and NK markers.

Isolation and culture of primary NK cells

Research with human primary NK was performed in accordance with the Declaration of Helsinki. Peripheral blood from healthy adults was obtained after informed consent (DTU-Rigshospitalet approval BC-40). No personal information regarding the donors has been disclosed to us. NK cells were isolated from PBMCs by immunomagnetic negative selection using EasySep NK-isolation kit (STEMCELL #17955) and cultivated in non-treated flasks (Greiner #658195) at 1.5-2.0×10⁶ live cells/ml in NK Complete Medium:⁶² 95% v/v ImmunoCult-XF (STEMCELL #10981), 5% v/v heat-inactivated (h.i.) human AB-serum (Merck-Sigma #H4522-100ML), 100 U IL-2/ml (equivalent to 2.5 ng/ml) and 10 ng IL-15/ml. The NK Basal Medium consists of 95% v/v ImmunoCult-XF and 5% v/v h.i. human AB-serum, without interleukins. Human AB-serum was heated to 56°C for 30 minutes to inactivate the complement cascade. Cells were cultured in humidified incubators at 37°C, 5% CO₂. NK cells grow partially adherent.⁶³ 3 days after isolation, a complete medium exchange was performed as follows: the entire supernatant of the culture was collected and centrifuged (300g, RT, 5 min). In the meantime, 12.5 ml of fresh NK-medium was added to the adherent NK. The centrifuged cell pellet was resuspended in 12.5 ml NK-medium and added to the adherent part. On day 6, the non-adherent NK cells were centrifuged, then resuspended in NK-Expansion-Medium: 95% v/v ImmunoCult-XF T-cell medium, 5% v/v h.i. hAB serum, 100 U/ml IL-2, 10 ng/ml IL-15, and 25 ng/ml IL-21 (STEMCELL #78082) and transferred to a new flask, whereas the adherent NK remained in the original flask and medium was exchanged with NK-Expansion-Medium. A fraction of the non-adherent NK cells started to adhere in the new flask gradually. On day 8, NK cells were harvested for electroporation. We used primary NK cells isolated from donor NK1 and NK2 for the transfections shown in [Figure S2F](#) (test of two electroporation programs) and from two donors NK3 and NK4 for CAR and CAR-EGFP transgene insertions in SHS shown in [Figure 2I](#). Additionally, NK cells from NK2 were used 15 days after isolation from blood in the co-culture experiment shown in [Figure 4J](#). Finally, primary NK

cells Sample NK5 (Human Peripheral Blood NK Cells by negative selection; 5×10^6 cells/cryovial; STEMCELL Technologies #70036; lot #190482404C) were used as a control during immunostaining of lymphoid progenitors and NK markers shown in Figures 4A, 4B, and S6A.

METHOD DETAILS

Design of SHS crRNA library

We searched for CRISPR RNAs (crRNAs) in selected genomic regions ± 1 kb that surround 20 nucleotide-long putative recognition sequences of the homing endonuclease mCrel, which were reported by Pellenz et al. 2019³¹ to satisfy several criteria for safe-harbor sites described before¹¹ (Figure S1D; Table S1). Namely, the criteria that the 20-nt mCrel sequences satisfy according to Pellenz et al. 2019 using the UCSC Genome Browser tracks as references, were: (1) >300 kb from any cancer-related gene on allOncogenes list, (2) >300 kb from any miRNA/other functional small RNAs (sno/miRNA), (3) >50 kb from any 5' gene end (RefSeq), (4) >50 kb away from any replication origin (Repli-seq peaks), (5) >50 kb away from any ultra-conserved element (VISTA enhancers), (6) Low transcriptional activity (no mRNA ± 25 kb), (7) Not in copy number variable region (segmental dups), (8) In open chromatin (DNase HS) (DHS signal ± 1 kb) and (9) Unique (one copy in the genome; tested by BLAST). After we converted the genomic coordinates of the selected 2 kb regions to the current human genome assembly (GRCh38 (hg38)), we designed 21-nt crRNA spacers for AsCas12a/LbCas12a nucleases with PAM specificity TTTN, using the CRISPR-RNA finding algorithm on Benchling [Biology Software (2019); Retrieved from <https://biosustain.benchling.com>]. crRNA composition settings: 0-80% for each nucleotide base. Off-target (specificity) score was calculated with the algorithm of Hsu et al. 2013: higher score indicates higher specificity. By default, Benchling skips (masks) repeats and low-complexity regions during off-target search. We designed 88 crRNAs for two control loci (hROSA26, AAVS1) and eleven previously non-characterized loci (C: chromosome, S: site): five crROSA26, five crAAVS1, five crC2s303, eight crC4s231, eight crC4s283, five crC5s255, eight crC6s311, eight crC7s257, eight crC7s301, eight crC8s325, seven crC12s329, five crC17s297 and eight crCXs313. We initially designed crRNAs with specificity score ≥ 90 using default search parameters. Because our SHS library targets intergenic regions, we eventually removed masking, to allow alignments with repeated and low-complexity sequences. The re-calculated specificity of 13 out of 88 crRNAs (crC2s303_1-5, crC4s283_7, crC5s255_1 & 4, crC7s301_4 & 8, crC12s329_3, 5 & 6) was <70 and crRNAs were excluded from subsequent selection. Alt-R Cas12a-crRNAs were ordered from Integrated DNA Technologies (IDT) in a 96-well plate (Table S2) and were resuspended in IDTE pH 7.5 (IDT #11-05-01-15) at 50 μ M. For subsequent characterization experiments, 10 nmol of each selected crRNA was resuspended in IDTE pH 7.5 at 100 μ M. As controls, we used crDNMT1_1 (20 nt) and crDNMT1_2 (21 nt) targeting the DNMT1 locus.¹⁵ crDNMT1 sequences (see also Table S2) are: CTGATGGTCCATGTCTGTTA(C).

Nuclease expression and purification

MAD7^{1xNLS}, MAD7^{3xNLS} and MAD7^{4xNLS} nuclease versions were produced in house. MAD7 protein expression and purification in batches has been carried out as described in Jedrzejczyk et al. 2022⁶⁴ and Rojek et al. 2023.⁶⁵ In detail: Three plasmids for expression of MAD7 were designed by Jedrzejczyk et al. 2022: 1xNLS (pNic28-EcMad7-1xNLS), 3xNLS (pNic28-EcMad7-3xNLS) and 4xNLS (pNic28-EcMad7-4xNLS). MAD7 versions were polyhistidine-tagged to allow protein purification. *E. coli* BL21 star (DE3) competent cells (Thermo-Fisher Scientific #C600003) were transformed with one expression plasmid encoding the respective MAD7 gene. 2xYT medium supplemented with kanamycin was inoculated with a single colony and incubated overnight at 37°C. The culture was diluted in 1-2 L of 2 x YT medium to OD₆₀₀ = 0.1 and grown at 37°C to OD₆₀₀ = 0.6. At this point, the culture was cooled on ice for 15–20 min. Next, IPTG (isopropyl thiogalactoside) was added in the final concentration of 0.2 mM, and protein was expressed overnight (18-20 h) at 18°C. Cells were harvested by centrifugation (5000g, 18°C, 15 min) and resuspended in lysis buffer (20 mM Tris, 500 mM NaCl, and 10 mM imidazole, pH = 8.0) supplemented with cOmplete, EDTA-free protease inhibitor cocktail (Roche #04693132001). After resuspension, Benzonase nuclease (Merck #E1014, ≥ 250 units/ μ l, 10 μ l per 40 ml lysate) and lysozyme (Merck #10837059001) (1 mg/ml lysate) were added and the cell suspension was placed on ice for 30 min. Cells were disrupted on an Avestin EmulsiFlex C-5 homogenizer (15000-20000 psi), and insoluble cell debris removed by centrifugation (15000 g, 4°C, 15 min). All subsequent chromatography steps were carried out at 10°C. The cleared lysate was loaded on a 5 ml HisTrap FF column (Cytiva #17525501). The resin was washed with 10 column volumes of wash buffer (20 mM Tris, 500 mM NaCl, and 20 mM imidazole, pH = 8.0) and the his-tagged protein was eluted with 10 column volumes of elution buffer (20 mM Tris, 500 mM NaCl, and 250 mM imidazole, pH = 8.0). Fractions containing the protein (typically 13.5 ml) were pooled and diluted to 25 ml in dialysis buffer (250 mM KCl, 20 mM HEPES, 1 mM DTT, and 1 mM EDTA, pH = 8.0). The sample was dialyzed against 1 L of dialysis buffer at 10°C using a dialysis membrane tubing with a molecular weight cut-off of 6-8 kDa (Spectra/Por standard grade regenerated cellulose, 23 mm wide). The dialysis buffer was replaced after 1-2 h and dialysis continued overnight. The next day, the dialyzed sample was diluted two-fold in 10 mM HEPES (pH = 8.0) and immediately loaded on a 5 ml HiTrap Heparin HP column (Cytiva #17040601), pre-equilibrated with buffer A (20 mM Hepes, 150 mM KCl, pH = 8.0). The column was washed with 2 column volumes of buffer A and the protein eluted using a linear gradient 0-50% of buffer B (20 mM Hepes, 2 M KCl, pH = 8.0) over 12 column volumes. Fractions containing the protein were pooled (typically 10-15 ml) and concentrated to 2 ml using a centrifugal filter unit Amicon Ultra-15 with 30000 MWCO (Millipore #UFC9030) by centrifugation at 4°C. A final size-exclusion chromatography step was performed by injecting the sample on a 120 ml Superdex 200 pg preparative-SEC column (Cytiva #28989335) with 50 mM sodium phosphate, 300 mM NaCl, 0.1 mM EDTA, pH = 7.5 as separation buffer. Fractions of interest were pooled and concentrated by centrifugal filtration (Amicon Ultra-15, 30000 MWCO; centrifugation at 4°C) to at least 20 mg/ml (concentration was determined by measuring absorbance at 280 nm on a NanoDrop 2000, Thermo-Fisher Scientific) with a percent solution extinction coefficient (Abs 0.1%) of MAD7). The concentrated protein solution was supplemented with glycerol (20% v/v, final concentration) and DTT (1 mM final concentration), snap-frozen in liquid nitrogen and stored at -80°C. Approximately, 20 mg of MAD7 was isolated per liter of *E. coli* culture.

RNP formation

RNPs were prepared by mixing crRNA and purified nuclease MAD7^{1×NLS}, or MAD7^{3×NLS}, or MAD7^{4×NLS} and allowed to form for 20-60 min at RT. In the first SHS-crRNA screen, 4 μl crRNA 50 μM (200 pmol) was mixed with 1 μl MAD7^{3×NLS} 174.8 μM (174.8 pmol). In the second SHS-crRNA screen, 4 μl crRNA 50 μM (200 pmol) was mixed with 2 μl MAD7^{1×NLS} 63 μM (126 pmol). In subsequent experiments, 100 or 50 pmol MAD7^{4×NLS} (improved version of MAD7) was complexed with stoichiometrically 1.5X crRNA (150 or 75 pmol), and the RNP master-mix was aliquoted to technical-replicas. For T-cell and NK-cell experiments (Figures 2H and 2I), and for the latest Jurkat and iPSC experiments (Figures 2E-2G), 1 μl of a 100 μg/μl aqueous solution of 15-50 kDa poly-L-glutamic acid (PGA, Alamanda-Polymers CAS #26247-79-0) was added to crRNAs, followed by addition of MAD7^{4×NLS} protein. Poly-L-glutamic acid (PGA) is a negatively charged polymer that has been described to stabilize the RNPs, and thereby increase the editing efficiency by approximately twofold.²⁷

iPSC electroporation (EP)

Cells growing in TesR-E8 or mTeSR-Plus (iPSC culture medium) on Matrigel-coated T75-flasks or 6-well plates and monitored by IncuCyte were harvested at 70-80% confluence. The cell layer was washed with DPBS and dissociated with TrypLE-Select (25 μl/cm²) for 5-10 min at 37°C. Since September 2020, Accutase (STEMCELL #07920) at 53 μl/cm² (0.5 ml/well of 6-well plates or 4 ml/T75-flask) replaced TrypLE, which improved dissociation to single cells in shorter time: 4 min 37°C (with addition of 1-2 min, when necessary). For 6-well: single-cell suspension was created by pipetting 5x with a P1000 tip and transferred in 15-ml tubes containing 4.5 ml iPSC culture medium. Cells were centrifuged at 200g for 4 min and pellet was dislodged gently before addition of 2 ml iPSC culture medium with 10 μM Y-27632. Viability and cell counts were measured using Via1 cassettes on Nucleocounter-NC200 (Chemometec). iPSC viability before EP was >70% in all experiments. For iPSC electroporation, the P3 Primary Cell 96-well Nucleofector Kit (Lonza #V4SP-3096) was used. Pre-treatment with fresh iPSC culture medium containing Y-27632 for at least 2 hrs before harvest for EP and 4 min dissociation time increased pre-EP viability >80% and survival post-EP. Cells were re-centrifuged (200g, 4 min) and pellet was resuspended in P3 Nucleofector solution. For transgene integration, we used linear dsDNA HDRT, instead of plasmids that are more stable, to reduce HDR-independent transient transgene expression. Per reaction, 2×10⁵ iPSC were resuspended in 20 μl P3, mixed with x μl RNP and y μl dsDNA HDRT (x+y= 5 μl) and electroporated using program CA137. Immediately after EP, 80 μl Y-27632-containing iPSC medium was added to cells for recovery, and each sample was split in two (10⁵ cells) or four fractions (5×10⁴ cells) and plated in Matrigel-96-well plates in TesR-E8/mTeSR-Plus with 10 μM Y-27632. Since January 2021, Y-27632 was substituted by CloneR (STEMCELL #05888) in the iPSC culture medium after electroporation to enhance survival, which also resulted in higher transgene-insertion rates (6-12% for certain sites; Figures 2E and S2C).

Genomic DNA (gDNA) extraction

48 hrs post-EP, iPSC were washed with DPBS and 20-50 μl QuickExtract (QE) Solution (Lucigen #QE09050) was added on the cell layer. The plates were sealed with DNase/RNase-free adhesive cover (Applied-Biosystems #4360954 or BioRad-Microseal-B #MSB1001) and plates were heat-treated in a Thermomixer-C (Eppendorf #13527550), with a plate Thermoblock (Eppendorf #13597550) covered with Thermotop (Eppendorf #13527630) to prevent condensation. Program: 65°C 15 min, centrifugation 2250g at 4°C for 5 min, Extract transferred into PCR plate, followed by heating to 68°C, 10 min, and 95°C for 10 min. gDNA was stored at -20°C, and before amplification, it was diluted 1:10 in nuclease-free water (Thermo-Fisher #AM9938).

NGS-primer design and pair-end amplicon sequencing

For 150 bp pair-end, the PCR product must be ≤250 bp including the primers. Furthermore, it should surround the cut site in an equal distance from each end. We used Primer3 to search for primers and Primer-BLAST for checking specificity. Upon excluding an area ±50 bp around the cut site, where no primers should be picked, we let Primer3 pick primers within 100 bp at each side of the 100 bp excluded area. We restricted PCR product to 200-250 bp. For 250 bp pair-end NGS, we restricted PCR product to 450 bp.

Genomic DNA was amplified by PCR using Phusion-High-Fidelity-2X-Mastermix (Thermo-Fisher #531L) or Phusion-HotStart-II-High-Fidelity-2X-Mastermix (Thermo-Fisher #F-565) using specific primers pairs for each amplicon (Table S3) at final concentration 0.5 μM each primer. Program: [98°C_3min] × 1, [98°C_10s | 72°C_10s (D -1°C/cycle) | 72°C_30s] × 17, [98°C_10s | 72°C_30s] × 30, [72°C_7min] × 1, [4°C_∞]. Amplicons were pooled or not, then diluted to 50 μl final volume and purified with AMPure-XP beads (Beckman-Coulter #A63880) using sample to bead ratio 1:1 v/v, because some of the amplicons are <500 bp and some 533 bp (including the Nextera overhangs; Table S3). DNA was eluted from the beads with nuclease-free water and analyzed on 2% agarose E-gels (Thermo-Fisher). In the second PCR, the P5 and P7 clustering sequences followed by unique pairs of i5 and i7 indexes from the Nextera XT Index-Kit-v2 (Illumina Set-A: FC-131-2001, Set-B: FC-131-2002, Set-C: FC-131-2003, or Set-D: FC-131-2004) were added to the amplicons by PCR using polymerase KAPA HiFi Hot-Start Ready-Mix (Roche #0795893500). Program: [95°C_3min] × 1, [95°C_30s | 55°C_30s | 72°C_30s] × 8, [72°C_5min] × 1, [4°C_∞]. The amplified products were purified with AMPure-XP beads with sample to bead ratio 1:1.8 and analyzed on E-gels. DNA concentration of each indexed sample was measured by Qubit dsDNA-High-Sensitivity (Thermo-Fisher #Q32854). DNA was normalized to 30 nM with Tris-HCl 10 mM pH 8.5. Quality of amplicon libraries was tested by Bioanalyzer-DNA-High-Sensitivity assay (Agilent). Amplicon libraries from the SHS-crRNA screen were sequenced by Illumina MiSeq v2 pair-end 2×250 bp. De-multiplexed FASTQ files were obtained from Illumina-BaseSpace. Data was analyzed using CRISPResso2 and visualized by R-Tidyverse.

Validation of crRNA INDEL efficiency by NGS

Genomic DNA (gDNA) from RNP-transfected samples (without HDRT) from several HDR-integration experiments was prepared by cell lysis 48 hrs post-EP using QuickExtract. gDNA was amplified and indexed in two-step PCR. Amplicons were sequenced by Illumina NGS pair-end 2×250 bp in MiSeq instrument. N ≥ 3 biological replicas in iPSC (n = 4 for C7s301_5) and 1-2 technical replicas per biological replica for each targeting crRNA. Average of technical replicas was calculated first, and then average and standard deviation (SD) of biological replicas. Control non-edited genomic DNA was obtained from either of these samples: cells treated with control RNP (crIDTneg1 or crIDTneg2), or non-electroporated cells, or electroporated cells without substrate (no RNP, no HDRT). Average indel % of all control non-edited samples was collectively calculated as Control.

Optimization of electroporation conditions for iPSC

Initial optimization of electroporation program for RNPs in small scale was done using 3 selected programs (CA137, CD118 and DN100). KOLF2 iPSC (cl. C1) were electroporated with 100 pmol RNP (100 pmol MAD7^{1×NLS} in complex with 150 pmol of the indicated guide crRNA) in P3 primary cell electroporation buffer. 10 miscellaneous crRNAs were tested (shown in Table S7). gDNA was quick-extracted and amplified by NGS primers. Amplicons were analyzed by Illumina MiSeq v2 2×150 bp (pair-end 150 bp reads). crCD19ex2, crCD90_1, crDNMT1 and crTRAC_43 were represented by 2 technical replicas (t = 2) whereas the rest of the samples were done as single replicas (t = 1). To identify the most optimal EP conditions, we transfected KOLF2-C1 with MAD7^{3×NLS}:crDNMT1 RNPs using combinations of 31 EP programs and 5 solutions: P1-P5 (Primary-Cell Optimization Kit Lonza #V4SP-9096). RNP master mix was prepared by mixing 100 pmol crDNMT1_2/reaction with 150 pmol MAD7^{3×NLS}/reaction. Before RNP formation, DTT was added to MAD7 at 1 μM concentration. 2×10⁵ KOLF2-C1 cells/reaction were electroporated and seeded at 10⁵ cells/well in 96well-plates (Corning #3595). Of all 155 combinations, P2-CA138, P4-CA137, P2-CA137 and P2-DS130 resulted in >50% indels (Figure S1C). Of the top-ranking conditions with editing efficiency >45%, only P3-CM150, P4-CA137 and P3-CA137 allowed cells to grow >50%, measured as a relative increase from the initial cell confluence (Figure S1C). However, upon inspection of individual IncuCyte images, cells in P2-CA138, P4-CA137, P2-CA137 and P2-DS130 appeared sick right after EP. Overall, cell growth after 2 days was <4-fold, less than the expected, due to cell death and high initial seeding density. Altogether, P3-CM150 and P3-CA137 are the best conditions. We performed all subsequent experiments using P3-CA137.

RNA extraction and RT-qPCR analyses

Two days post-EP, the iPSC layer was washed with DPBS, the plate sealed and stored until RNA extraction at -80°C. Total RNA was extracted using the RNeasy Mini (Qiagen #74104). cDNA was synthesized with ProtoScript First-Strand cDNA-Synthesis (NEB #E6300S). cDNA concentration was measured using Qubit ssDNA Assay (Thermo-Fisher #Q10212) and diluted to 10 ng/μl. 10 ng cDNA was used to quantify the expression of pluripotency-associated genes (*OCT4*, *SOX2*, *NANOG*, and *TDFG1*). All evaluated genes were analyzed in quadruplicates by qRT-PCR on the Quantstudio5 instrument (Applied-Biosystems) using PowerUp SYBR-Green Mastermix (Thermo-Fisher #A25742). The reactions were conducted using the following parameters: 50°C for 2 min, 95°C for 2 min and 40 cycles at 95°C for 15 s, 59°C for 15 sec, 72°C for 1 min. To assess the levels of pluripotency and selected differentiating markers in the engineered iPSC lines, total RNA was extracted using the RNeasy Micro Kit (Qiagen #74004). cDNA was synthesized using High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems #4368814), cDNA concentration was evaluated with ThermoFisher Nanodrop^{One} instrument and diluted to 10 ng/μl, then qPCR analysis were performed with PowerUp SYBR-Green Master Mix using BioRad CFX96 C1000 Touch Thermal Cycler, where all evaluated genes were analyzed in quadruplicates. The reactions were conducted using the following parameters: 50°C for 2 min, 95°C for 5 min and 40 cycles at 95°C for 15 s, 56°C for 15 sec, 72°C for 1 min. Gene expression levels relative to *GAPDH* housekeeping gene were calculated by the 2^{-ΔΔCt} method. Primers used for RT-qPCR of non-differentiation markers are shown in KRT and Table S9.

Construction of plasmids carrying the HDR templates

The pTwist-Ampicillin-High-Copy backbone was linearized by PCR of 2 pg plasmid with primers recognizing its pTwistUni9R/F adaptors. The expression cassettes were initially constructed with CMV-promoter and SV40-polyA and designed to be flanked by a 500 bp left homology arm (LHA) and a 500 bp right homology arm (RHA), homologous to the 1 kb chromosomal region that surrounds the cut site of each of the 8 selected crRNAs. 18 bp of the genomic region including part of the PAM (2-3 nt) and part of the spacer (15-16 nt) were omitted from the homology arms to avoid recutting by MAD7. The parts CMV-enhancer^{1-177bp}, CMV-enhancer-promoter^{178-508bp}_EGFP_SV40-polyA and the LHA/RHA for each site were ordered as gBlocks (IDT) or GeneFragments (Twist Biosciences). The LHA for crROSA26_3 and RHA for crROSA26_4 were produced by PCR using KOLF2 gDNA as a template. gBlocks and GeneFragments were amplified by PCR using primers with 20-30 nt homology overhangs suitable for Gibson assembly and polymerase Phusion HS-II-HF (Thermo-Fisher #F565S). LHA fragments contain the Uni9F at 5' end and RHA the Uni9R at 3' end. All fragments were flanked by restriction sites. Each plasmid was constructed by assembling 5 different parts (pTwist^{Amp}, LHA, CMV-enhancer^{1-176bp}, CMV-enhancer-promoter^{178-508bp}_EGFP_SV40-polyA, and RHA) using InFusion-HD-cloning kit (Takara-Bio #639650), and was transfected in TOP10 chemically competent *E.coli* (Thermo-Fisher #C404003) and plated on LB-agar with ampicillin 100 μg/ml. Colonies were screened by colony PCR and positive clones were cultured in LB-ampicillin 37°C overnight. Plasmids were purified using QIAprep-Spin (Qiagen #27106). Positive clones were verified by Sanger sequencing.

Because CMV promoter did not drive efficient expression of GFP in iPSC, as also reported before,³⁵ CAG-promoter-driven plasmids with bGH(bovine growth hormone)-polyA were constructed. The chimeric antigen receptor is a tagged version of an anti-CD19 CAR, consisting

of: the scFv FMC63,^{44,45} the myc peptide sequence, the CD28 transmembrane domain (123 bp), a 126 bp cytoplasmic domain of 4-1-BB co-stimulatory receptor and the CD3 ζ domain. The CAR-GFP plasmids contain the sequences of both FMC63^{CAR-CD28-41BB-CD3 ζ} and EGFP separated by a 2A peptide sequence. The myc-tag allows CAR⁺ cell detection by staining with anti-myc antibodies. Construction of CAG plasmids was carried out in three steps: Step-1: the expression cassettes CAG-CAR-bGHpA, CAG-EGFP-bGHpA, and CAG-CAR-2A-EGFP-bGHpA (Table S5) were isolated by gel purification of previously constructed plasmids digested with EcoRV (NEB #R3195L) and SphI (NEB #R3182L). PvuI (NEB #R3150L) or FspI (NEB #R0135L) were added, to destroy the Ampicillin^R gene of the backbone. Step-2: the part of the original plasmid containing the RHA, pTwist^{Amp} backbone and LHA for each site, was amplified from 2 ng plasmid by temperature-gradient PCR (BioRad-C1000 Thermal-Cycler), $T_{\text{annealing}}=60^{\circ}\text{C}-65^{\circ}\text{C}$, using Phusion HF-HS-II (Thermo-Fisher #F-5655), primers with 30-38 nt overhangs homologous to the cassettes CAG-CAR-bGHpA, CAG-EGFP-bGHpA, and CAG-CAR-2A-EGFP-bGHpA (Table S9) and 0%, 3% or 5% DMSO. $\geq 3\%$ DMSO was necessary to amplify all RHA-pTwist^{Amp}-LHA backbones. After PCR, the remaining parental methylated pTwist^{Amp}-CMV-EGFP templates were eliminated by DpnI (Thermo-Fisher #FD1703) and the PCR products were gel purified. Step-3: 35 fmol of CAG-CAR or CAG-EGFP were assembled with 35 fmol linearized RHA-pTwist^{Amp}-LHA using InFusion-HD kit and transformed in Mach-1 competent *E. coli* (Thermo-Fisher #C862003). For CAG-CAR-EGFP constructs, parts were assembled at 22 fmol using InFusion Snap-Assembly kit (TaKaRa #638948) and transformed in TOP10 competent *E. coli*. Colonies were cultured on LB-ampicillin, 37°C overnight. Plasmid DNA was purified using QIAprep-Spin. Positive clones were identified by Sanger sequencing. Plasmid sequences have been deposited at Zenodo (see [key resources table](#)).

Preparation of linear dsDNA HDR templates

CAG-driven GFP, CAR or CAR-GFP plasmids (construction methods in Supplementary) were propagated overnight in TOP10 *E. coli* and purified using HiSpeed Plasmid-Maxiprep (Qiagen #12663). Excision of linear dsDNA HDRTs was performed by BbvCI (or EcoRI, or BsrDI) and XhoI, whose recognition sites exist at the 5' of LHA and at 3' of RHA respectively, at 0.5-2.0 enzyme Units/ μg plasmid. For AAVS1_3 plasmids, PvuII-HF (NEB #R3151S) was used together with XhoI. To facilitate the separation of the desired band upon agarose electrophoresis, PvuI or FspI was added to destroy the Amp^R gene of the backbone. The combinations of restriction enzymes used for excision of linear HDRTs from plasmids, as well as the sizes of the HDRTs and molarity calculations are shown in Table S10. Note that BbvCI, which is of the enzymes designed to cut at the 5' of LHA for C8s325_6, cuts also the RHA of C8s325_6 at its 280 nt position. This means that all the HDRTs designed for integration at C8s325_6 have truncated RHA consisting of 280 bp instead of 500 bp. The other sites that was added to the 5' of LHA for C8s325_6, BclI, does not cut methylated DNA. We propagated our plasmids in *dam*⁺*dcm*⁺ *E. coli*. Methylase-negative *E. coli* strains are available, if a 500 bp RHA for C8s325_6 is required in future studies. For digestions, 100 μg plasmid were digested overnight at 37°C and after heat inactivation of restriction enzymes at 65°C for 30 min (note that not all enzymes are able to inactivate), the digested DNA was resolved in 1% agarose gels containing SYBR Safe DNA gel stain (1x) (Thermo Fisher #S33102), at 80 Volts for 45min RT in Tris-Acetate EDTA buffer (Thermo Fisher #B49). Ladder: 25 μl GeneRuler 1 kb. SeaPlaque GTG (low-melting) agarose (Lonza #50110) was generally preferred over the available standard agarose (Bionordika #BN-50004) for this preparative electrophoresis. After the gels were solidified, they were placed at 4°C for 0.5-1 hr before electrophoresis. Each gel was visualized using iBright FL1500 Imaging System (ThermoFisher) and the desired agarose band was excised and purified using NucleoSpin (Macherey-Nagel #740609). DNA concentration was measured on Nanodrop (Thermo-Fisher). DNA was precipitated overnight at -20°C after addition of 1:10 v/v sodium acetate 3M pH 5.2 (Merck-Sigma #567422-100ML) and 2.5x v/v cold absolute ethanol (-20°C). followed, washing of the DNA pellet and centrifugation at 13000g for 10 min. DNA was pelleted by centrifugation at 13000g, 4°C, 30 min, washed twice with 70% ethanol (-20°C), and air-dried under BSC-II for 5-15 min before resuspending in sterile IDTE pH 7.5. DNA concentration was measured in 1:25 dilutions by Nanodrop.

Fluorescence-activated cell sorting (FACS)

5 ml CloneR (10x) and 500 μl penicillin-streptomycin (10000 U/ml; 100x; Thermo-Fisher #15140122) were added to 45 ml complete mTeSR-Plus or TeSR-E8 to prepare iPSC-Cloning-Medium. iPSC-Cloning-Medium was added 100 μl /well to Matrigel-pre-coated 96-well plates. SONY-SH800 sorter was loaded with a single-use 130 μm sorting chip and calibrated with beads (SONY #LE-B3001); Settings: Lasers 405 nm and 488 nm switched on; single-cell sorting quality; indexed sorting, sample pressure 5-6, 800-1200 eps. GFP⁺ or CAR⁺GFP⁺ or parental iPSC growing in 6-well plates were single-cell harvested and resuspended in 2 ml iPSC-Cloning-Medium, roughly leading to 10⁶ live cells/ml suspension and viability was measured (Via1-cassette). Single GFP⁺ cells were sorted into individual wells of 96-plates and incubated at 37°C, 2 days, followed by a complete iPSC-Cloning-Medium change. 4 days after sorting, medium containing pen-strep but no CloneR was used and exchanged daily (TeSR-E8) or every 2-3 days (mTeSR-Plus) until colonies were ready to be picked. 8-10 days after FACS, colonies were scanned by IncuCyte (4x). Wells with colonies were selected and re-scanned by IncuCyte (10x/20x magnification, 4 images/well) to assess cell morphology and distinguish clones with high-stem-cell-quality: Criteria: GFP expression, round-smooth colony periphery, round cells, tight cell connections, prominent nucleolus-nucleus versus cytoplasm. 10-14 (≤ 15) days after FACS, cells from selected clones were harvested by Accutase (30-50 μl /well) and expanded in 12-well or 6-well plates in medium with 10 μM Y-27632 and pen-strep for 48 hrs, followed by changes with medium containing pen-strep (no Y-27632), daily (TeSR-E8) or every 2 days (mTeSR-Plus) until cryopreservation. Clones were cryopreserved either as aggregates using CryoStor-CS10 (STEMCELL #07930) (1 well of 6-well plate in 2 cryovials) or as single cells (10⁶ cells/ml) using FreSR-S (STEMCELL #05859).

Genomic PCR for confirming transgene insertions

Using Quick-Extracted gDNA from monoclonal iPSC lines, and a primer that binds on the chromosome outside the 500 bp homology arm of our HDRT, and a primer that binds inside the transgene, we amplified (Flanking In-Out PCR) a region 0.8-1.6 kb around either the LHA or RHA, that can only exist if the transgene is correctly inserted in the intended site and does not exist in the non-edited parental genome or the HDRT itself. For LHA, the transgene primer binds either in the CMV-enhancer part or the chicken-beta-actin-promoter part of the CAG, whereas for the RHA, the transgene primer binds inside the GFP gene (see [Table S9](#)). We PCR-confirmed CAG-EGFP and CAG-CAR-EGFP expressing clones. GFP expression was monitored continuously by IncuCyte and was deemed stable.

Electroporation of Jurkat cells

Jurkat were initially ([Figure S2D](#)) electroporated with these conditions: 0.5 μg (~225 fmol) CAG-EGFP linear dsDNA HDRT with 100 pmol RNP (100 pmol MAD7^{4xNLS} : 125 pmol crRNA) using Lonza Nucleofector-4D buffer-program combination: SF-CA137, which we have previously validated (Mohr et al. 2022).³⁷ CAG-CAR-GFP HDRTs (in [Figure 2G](#)) were electroporated at 0.5 pmol (1.35-1.67 μg) together with 50 pmol RNP (75 pmol crRNA, mixed with 100 μg PGA (15-50 kDa) and 50 pmol MAD7^{4xNLS}) using SF-CA137. In general, just before EP, Jurkat cells were centrifuged at 200g, RT, 5 min, then resuspended in SF buffer at 10⁶ cells/ml (2 \times 10⁵ cells/20 μl SF). Cells were mixed with HDRT and the appropriate RNP (three technical replicas), transferred to the nucleofection plate and electroporated with CA137. After EP, 80 μl /well RPMI-1640 (supplemented with FBS and pen-strep) was added to the cells for recovery. 50 μl of the total 105 μl transfected cell suspension was transferred to each of two flat-bottom non-culture-treated 96-well plates (Falcon #351172) already containing 150 μl /well fresh RPMI-1640 (10% FBS, pen-strep). Cells were incubated at 37°C, 5% CO₂ for 48 hrs before measurement of viability, and GFP and CAR expression by flow cytometry.

Electroporation of primary T cells

RNP: 75 pmol crRNA were mixed with 160 μg PGA and 50 pmol MAD7^{4xNLS} at RT for 15-20 min and stored at 4°C until cells were prepared. 48 hrs after isolation, T cells from two donors were harvested (300g, RT, 5 min) and resuspended at 50 \times 10⁶ cells/ml (10⁶ cells/20 μl /reaction) in P3 solution. Cells were mixed with HDRT, then with RNP, and were electroporated with program EH115. Each reaction was performed in three replicas. After transfection, 80 μl pre-warmed cultivation medium without IL-2 was added for recovery. Cells were seeded at a density of 2.5 \times 10⁶ cells/ml in non-treated 96-well plates (Falcon #351172) containing 150 μl /well pre-warmed medium with 12.5 ng IL-2/ml. Viability was measured 24 hrs post-EP, after which the cells were reseeded in fresh cultivation medium containing IL-2 12.5 ng/ml with or without splitting every 2-4 days until flow cytometry measurements. CAR or CAR-GFP insertion efficiency as the percentage of CAR⁺ or CAR⁺GFP⁺ was measured 7 and 14 days post-transfection.

Electroporation of primary NK cells

Linear dsDNA CAG-CAR-GFP HDRTs for integration into four SHS and ROSA26_3, along with CAG-CAR for AAVS1_3 were used. RNP formation: 60 pmol crRNA were mixed with 100 μg PGA (15-50 kDa) and 50 pmol MAD7^{4xNLS}, incubated at RT for 15-20 min and stored at 4°C until cells were prepared. Expanded NK cells, 8 days after isolation, were harvested as follows: non-adherent NK cells were collected and centrifuged (300g, RT, 5 min). 7 ml Mg/Ca-free DPBS was added on the still adherent NK cells in T75 and set at 4°C for 10 min. This resulted in complete detachment and breaking up of clumps.⁶³ The detached cells in DPBS were pooled with the pelleted non-adherent cells and total cells were counted (Via1 cassette, NC200). The required number of cells was pelleted by centrifugation (300g, RT, 5 min) and re-suspended at 50 \times 10⁶ cells/ml (10⁶ cells/20 μl /reaction) in P3 solution. Cells were mixed with 500 fmol dsDNA HDRT and this suspension was then immediately mixed with the corresponding RNP, followed by EP with program EN138. Each reaction was performed in three technical replicas. In a preliminary experiment ([Figure S2F](#)), program EN138 was shown to perform better for NK cells compared to EH115 (T-cell program). After transfection, 80 μl pre-warmed NK-Expansion-Medium containing M3814 at 2 μM final concentration was added to the NK cells for recovery (RT, 10 min), similarly to the procedure described by Mohr et al. 2022³⁷ for homologous recombination in primary T cells. M3814, also known as Nedisertib, is a highly potent and selective inhibitor of DNA-PK and thereby of NHEJ repair branch (SelleckChem #S8586). Then, cells were transferred onto two replica 96-well non-treated plates (Falcon #351172) containing 150 μl /well pre-warmed NK-Expansion-Medium containing M3814 at 2 μM . Cells were seeded at a density of 2.5 \times 10⁶ cells/ml and incubated at 37°C in 5% CO₂. 24 hrs post-transfection, medium was changed partially by carefully removing 150 μl old medium/well and adding 150 μl fresh NK-Expansion-Medium without M3814. Viability assay was carried out 48 hrs post-transfection. On day 5 post-EP, cells from the other plate were split into two replica plates and medium was replaced by fresh NK-Expansion-Medium. CAR and CAR-GFP insertion efficiency was measured 7 and 14 days after transfection.

Optimization of iPSC differentiation to CD34⁺ HSPC

The STEMdiff-Hematopoietic kit (STEMCELL #05310) was used for differentiation to CD34⁺ HSPC according to manufacturer's instructions, unless otherwise stated in the [results](#) section. As the number of seeded iPSC aggregates and, hence, the number of resulting hematopoietic colonies is critical for the differentiation, we pursued a validation of this range. We seeded KOLF2 aggregates at different numbers (40, 50, 60, 80, 100 and 150 aggregates/well); 4 replica wells. On Day 0, we selected for differentiation only the wells with 16-40 colonies/well (4-10 colonies/cm²). All wells with 40 and 150 aggregates resulted <15 or >40 colonies, respectively, and were discarded. On Day 11 or 12, we harvested the non-adherent cells (HSPC) and we stained with anti-CD34, CD43 and CD45 antibodies pairwise and assessed the efficiency of

differentiation as the percentage of CD34⁺ live cells compared to total live harvested cells. 3-4 wells initially seeded with the same density were pooled before staining to ensure sufficient cell numbers. In [Figures 3B](#) and [3C](#), cells were stained with anti-CD34_BV786 and anti-CD43_APC simultaneously (upper row); or with anti-CD43_APC, anti-CD45_PB450 and Isotype-Control_BV786 simultaneously (lower row). A fraction of CD34⁺ HSPC was further differentiated or cryopreserved in CryoStor-CS10 (STEMCELL #07930).

iPSC differentiation to CD34⁺ HSPC by the “2D protocol”

The STEMdiff-Hematopoietic kit (STEMCELL #05310) was used for differentiation to CD34⁺ HSPC according to manufacturer's instructions, unless otherwise stated in the [results](#) section. One day before differentiation (Day -1), iPSC are seeded as aggregates, 100-200 μm in size, on Matrigel-coated 12-well plates at specified densities so that there are a few interspersed colonies after one day (4-10 iPSC colonies/ cm^2 or else 16-40 colonies/well). iPSC culture medium is replaced by Hematopoietic-Medium-A on Day 0 to induce a mesoderm transition within 3 days. Medium-A was replaced by Hematopoietic-Medium-B on Day 3, thereafter half-volume medium changes were performed. Days 3-5: the mesoderm colony transform into a disc-shaped colony with a dense center and loose periphery and from day 5 onwards, the center becomes spongy and hollow while the entire colony increases in height ([Figure S4A](#)). After Day 7 and until harvest on Day 11 or 12, round hematopoietic stem and progenitor cells emerge in suspension from the underlying adherent supporting cell layer, while the center of the colony has a 3D gelatinous structure. Cells in suspension containing the HSPC are harvested on Day 11 or 12. When harvest is on Day 11, the last half-medium change was performed on Day 9 instead of Day 10.

Hematopoietic differentiation through embryoid bodies (EB)

We differentiated parental KOLF2 and SHS-engineered iPSC lines to embryoid bodies for 12 days using the STEMdiff NK differentiation kit (STEMCELL #100-0170) according to the instructions in two experiments. We generated single-cell suspensions from KOLF2 and monoclonal iPSC lines C7s301_5 GFP cl. D9, CXs313_4 GFP cl. A8 and ROSA26_3 CAR-GFP cl. A4 in EB-formation medium (EB-medium A + Y-27632 (10 μM final conc.) and seeded 3.5×10^6 live single cells/well into AggreWell-400 6-well plates (STEMCELL # 34425) previously rinsed with Anti-Adherence Rinsing Solution (STEMCELL #07010). Media changes were performed regularly according to the instructions of the kit. EB-medium B was used from day 3 until day 12. Fluorescence or regular imaging of EBs was performed regularly. On day 5, EBs were dislodged from the AggreWell plate and replated in a non-tissue culture-treated 6-well plate (Falcon #351146) according to the instructions of the kit. On day 12 of differentiation, EBs were harvested and dissociated according to the instructions of the STEMdiff NK kit, by sequential incubation with Collagenase type-II (STEMCELL #07418) and TrypLE Select. In one of the experiments, half EBs per sample were dissociated with Accutase and half with collagenase and TrypLE for comparison reasons. Samples of dissociated EB cells were stained with monoclonal fluorescent antibodies in BioLegend staining buffer and analyzed by flow cytometry. Because the % of CD34⁺ cells in the dissociated EBs was very low compared to KOLF2-derived HSPC previously differentiated using the 2D protocol for 12 days, which we used as a positive control for CD34 staining, we did not proceed to isolation of CD34⁺ cells using EasySep Human CD34 Positive Selection Kit II, recommended by the kit. Thus, the EB differentiation experiments ended at day 12 with analysis only.

Differentiation of HSPC to lymphoid progenitors and NK cells

For further differentiation of HSPC, we used the StemSpan-NK-generation kit (STEMCELL #09960) or the STEMdiff-NK kit (STEMCELL #100-0170), according to manufacturer's instructions, unless otherwise stated in the [results](#) section. The first part of the STEMdiff-NK kit that generates CD34⁺ from embryoid bodies (EB) did not work as efficiently as desired and was substituted by the STEMdiff-Hematopoietic kit, as described above. We seeded CD34⁺ HSPC in non-tissue treated plates coated with Lymphoid-Expansion-Coating Reagent at 5×10^4 live cells/ml using Lymphoid-Expansion-Medium and performed frequent half-media changes according to the protocol. At day 14 of lymphoid cell expansion, we harvested cells and reseeded at 5×10^4 live cells/ml in tissue-culture, non-coated plates in NK-Differentiation-Medium containing 1 μM UM729 (STEMCELL #72332). A fraction of the remaining lymphoid progenitors was cryopreserved in CryoStor-CS10, and the rest was stained pairwise for: CD5, CD7, CD43, CD45, and CD34, as we expected lymphoid progenitors to progressively lose the CD34 marker. After 14-15 days of differentiation from lymphoid to NK stage (in total 28-30 days from HSPC to NK), we harvested NK cells and stained with the lineage-specific marker CD56, the hematopoietic markers CD43, CD45, and the stimulatory receptor NKG2D.

Flow cytometry

Flow cytometric assessments were carried out on a 4-laser (Violet-Blue-Yellow-Red) CytoFLEX-S (Beckman-Coulter) in 96-well plates. CytoFLEX was calibrated daily using beads (Beckman-Coulter #B53230) and cleaned after each run. $1-2 \times 10^4$ single-cell events were recorded per sample. In differentiation experiments, cell yields in each step were low, leading to <5000 single cells in many cases. Therefore, cytometry dot-plots are shown in all figures that show differentiation. After cell harvest, viability and cell counts were measured with Via1 cassettes (Chemometec #941-0012). For fluorescence immunostaining, cells in suspension were collected in V-bottom plates (Greiner #651161), centrifuged 300g, RT, 5 min, and resuspended in a bovine calf serum-containing Staining Buffer (Biolegend #420201) or a BSA-containing Staining Buffer (BD #554657). In the meantime, appropriate antibodies individually or in combinations were diluted with specified ratios (shown in [Table S8](#)) in same staining buffer. Cells were pelleted by centrifugation 300g, RT, 5 min and resuspended in 50 μl antibody mixtures and incubated at RT for 1 hour. Thereafter, cells were washed with staining buffer three times. Viability assessment was performed with either 7-AAD (Thermo-Fisher #A1310) diluted 1:1000 in the last washing step or Zombie Violet (BioLegend #423113) diluted 1:500 or 1:1000 in Ca/Mg-

free DPBS in the very beginning of the procedure before stained with antibodies. For antibody ID, see [key resources table](#) and for dilutions see [Table S8](#).

Phenotyping of iPSC-derived HSPC, LPC and NK

CD34 density is highest on the membrane of early HSPC and decreases as cells mature. CD34 is absent from fully differentiated blood cells (mature peripheral blood lymphocytes, monocytes, granulocytes, and platelets). CD45 is expressed on most blood cells except mature erythrocytes, platelets, and plasma (B) cells. CD43 is a general hematological marker that is expressed on most human blood cells, including thymocytes, peripheral T cells, NK cells, monocytes, granulocytes, some B cells plasma cells, and hematopoietic progenitor cells. HSPC, LPC and NK samples and controls were, as a rule, stained pairwise with antibodies of selected fluorophores with minimal emission spectral overlap to avoid spillover, and therefore compensation was not necessary. NALM6 and primary T- and NK-cells were used as controls. 7AAD was not used together with other fluorophores (PE, PerCP, APC) that emit within 580-750 nm, to avoid spillover and false-positive detection as dead cells.

Immunostaining of non-differentiation markers for imaging

After transfection, cells were seeded into Matrigel-coated 96 well plates. 72 hrs after seeding, transfected cells were washed with DPBS once and fixed with 4% formaldehyde solution in DPBS (Merck Sigma #252549-500ML) for 15 RT, washed twice with DPBS and stored at 4°C until staining (can be stored up to 2 weeks at 4°C). The day of staining, the cell layer was permeabilized with Permeabilization solution (DPBS 0.5% v/v Triton-X100) for 5 min at RT. Permeabilization solution was removed, followed by washing twice with Wash solution (DPBS 0.1% v/v Tween-20) and then blocking with Blocking/Staining solution (DPBS, 0.1% v/v Tween-20, 5% w/v BSA) for 20 min at RT. Cell layer was then incubated with monoclonal antibodies ([Table S8](#)) in Blocking/Staining solution overnight at 4°C. Controls were incubated with isotype controls or no primary antibody. Next day it was washed 3 times with Wash solution, then incubated with secondary antibodies diluted in Blocking/Staining solution for 30 min RT. Equal volume of DAPI 2X concentrated (300 µg/ml or approx. 600 nM) in DPBS was added and incubated for 3-5 min RT. Cell layer was finally washed 3 times with Wash solution and imaged by IncuCyte S3 at phase-contrast, green (excitation 440–480 nm; emission 504–544 nm) and red (excitation 565–605 nm; emission 625–705 nm) at 20X magnification. DAPI was only used for verification of nucleic staining at a Leica fluorescent microscope, as IncuCyte cannot image DAPI.

Immunostaining of non-differentiation markers for flow cytometry

Cells were harvested as single cells by Accutase and rinsed with DPBS. Cells were fixed at a maximum concentration of 2×10^6 cells/ml of fixative solution (4% formaldehyde solution in DPBS; Merck Sigma #252549-500ML) for 15 min at room temperature (RT). Fixed cells were pelleted by centrifugation at 1000g for 4-5 min and then washed 2 times with DPBS (1 ml/2 million cells) and stored at 4°C for 1 week (up to 2 weeks). On the day of staining, fixed cells were spun down to remove DPBS and permeabilized with 1 ml PBS 0.5% Triton-X100 per 2×10^6 cells for 5 min at RT. Permeabilized cells were pelleted by centrifugation at 1000g for 4-5 min and then washed 2 times with 1 ml (/2 million permeabilized cells) of wash buffer (PBS 0.1% Tween) at RT. Cell pellet was resuspended in 1 ml blocking solution (PBS-T 1% BSA) and incubated for 30 min at RT (conditions of blocking and staining were optimized for each cell line and antibody). During this time, the cell suspension was aliquoted into wells of a V-bottom plate: 150 µl/well which corresponds to 300K cells per aliquot and total 6 wells. 150 µl was kept aside as the non-stained ctrl. The plate at centrifuged at 1000g for 4 min and the cell pellets were incubated with antibody mixes in PBS-T 1% BSA (50 µl/well) for 1 hour at RT followed by washing 3 times with PBS 0.1% Tween at RT (100 µl/well). Optional: addition of 1:2000 of 2X concentrated DAPI, dilactate (Thermo-Fisher #D3571) of stock 5 mg/ml (10.9 mM) in dH₂O. Final DAPI conc. 1 µg/ml (2.725 µM) at RT for 5 min. Stained cells were analyzed by flow cytometry on CytoFLEX-S.

Annexin V apoptosis assay by IncuCyte

NALM6-RFP⁺ cells were used as target cells. Prior to the assay, NALM6-RFP⁺ cells were cultured in RPMI-1640 10% h.i. FBS with penicillin-streptomycin. Viability and cell number were measured by NC200 Via1 cassettes (Chemometec #941-0012). 10^4 live cells were added per well. The apoptosis reagent IncuCyte Annexin V Green (Sartorius #4642) was added (final dilution 1:200) into the target cells 1 hour after seeding, followed by immediate addition of the effector cells in different ratios (25K, 50K or 100K; K = 10^3 cells). Effector cells were primary NK cells from donor NK2 (15 days after isolation from blood), KOLF2-derived NK and NK derived from ROSA26_3 CAR-GFP cl. A4 iPSCs. As positive control, absolute ethanol was added to 3 replicate wells containing NALM6-RFP⁺ but not effector cells to a final concentration 10%, 2 hrs after initial target cell seeding. NK Basal Medium was used for all cell suspensions in the assay. Each treatment consists of 3 technical replica wells. The average ratio of NALM6 RFP⁺AnnexinV⁺ / NALM6 RFP⁺ and SD of the 3 technical replicas is shown. Done in one biological replica. Cells were imaged by IncuCyte system at 20X magnification by non-adherent cell-by-cell format.

QUANTIFICATION AND STATISTICAL ANALYSIS

NGS data analysis

An initial quality assessment of the obtained reads was performed with FastQC. The sequencing data were aligned and analyzed with the CRISPResso2 software.⁶⁰ For the INDEL frequency analysis, the CRISPRessoBatch command was used with the parameters `-cleavage_offset`

1-quantification_window_size 10-quantification_window_center 1-expand_ambiguous_alignments. Sequence modification rates from the CRISPResso2 software output were analyzed in MS Excel.

Flow cytometry data analysis

Analysis of flow cytometry files was performed with FlowJo software (BD Biosciences). Gating strategy: FSC-A vs SSC-A density plot with axes in linear scale for selection of cell population, then FSC-A vs FSC-H plot with axes in linear scale for selection of single cells. Finally single cells were analyzed in dot plots for each fluorescent channel vs SSC-A or in histogram plots. Whenever a viability dye was used (Zombie Violet or 7-AAD), it is mentioned in the figure captions and the gating strategy was: FSC vs SSC plot for selection of all cellular events, then the ZombieViolet-negative or 7-AAD-negative events were gated and analyzed in FSC-A vs FSC-H plot for selection of single cells. No compensation was applied because the selected fluorophores used in combination do not overlap in their emission spectra.

Statistical analysis

Flow cytometry data tables exported from FlowJo were further analyzed in MS Excel. The average percentage of positive cells and its standard deviation from technical replicas are shown in graphs in main and supplementary figures. Mean \pm SD from replicate wells was calculated in MS Excel. No analysis of significance was performed as it was not in the scope of the study to analyze the significance of difference between safe harbor sites. Relative mRNA levels were calculated by the $2^{-\Delta\Delta C_t}$ method compared to *GAPDH* housekeeping gene and then calculated relative to a control sample, which is indicated in figure legends and the main text. Mean \pm SD from qPCR quadruplicates was calculated in MS Excel. Validation of indel efficiency of the selected SHS crRNAs was done in 3-4 biological replicas by NGS analysis, data analysis was done by CRISPResso2, and graphs were created in MS Excel.