UNIVERSITÁ VITA-SALUTE SAN RAFFAELE
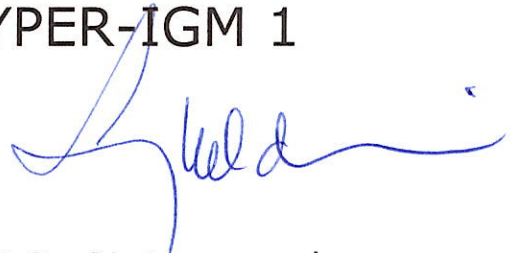
CORSO DI DOTTORATO DI RICERCA INTERNAZIONALE IN MEDICINA MOLECOLARE

CURRICULUM IN GENE AND CELL THERAPY

GENOME INTEGRITY ASSESSMENT OF EDITED CD4+ LYMPHOCYTES FOR THE TREATMENT OF HYPER-IGM 1

DoS: Prof Luigi Naldini

Second Supervisor: Prof Luigi D. Notarangelo

Tesi di DOTTORATO di RICERCA di Daniele Canarutto

matr. 015459

Ciclo di dottorato XXXV

SSD MED/38 MED/03 MED/15 BIO/11 BIO/18

Anno Accademico 2021/2022

RELEASE OF PHD THESIS

Il sottoscritto/ *I, the undersigned*   Daniele Canarutto

Matricola / *registration number*   015459

nato a/ *born in*   Moncalieri (TO)

il/*on*   06/06/1989

autore della tesi di Dottorato di ricerca dal titolo / *author of the PhD Thesis titled*

GENOME INTEGRITY ASSESSMENT OF CD4+ EDITED LYMPHOCYTES FOR THE TREATMENT OF HYPER-IGM 1

NON AUTORIZZA la Consultazione della tesi per 12 mesi /DO NOT AUTHORIZE *the public*

release of the thesis for 12 *months*
a partire dalla data di conseguimento del titolo e precisamente / *from the PhD thesis date,*

specifically

Dal / *from* ......3..../.....3..../.......2023 Al / *to* ......3..../.....3..../......2024

Poiché /*because*:

☐ l'intera ricerca o parti di essa sono potenzialmente soggette a brevettabilità/ *The whole project or parts of it may be the subject of a patent application;*

☐ ci sono parti di tesi che sono già state sottoposte a un editore o sono in attesa di pubblicazione/ *Parts of the thesis have been or are being submitted to a publisher or are in the press;*

X la tesi è finanziata da enti esterni che vantano dei diritti su di esse e sulla loro pubblicazione/ *the thesis project is financed by external bodies that have rights over it and its publication.*

E' fatto divieto di riprodurre, in tutto o in parte, quanto in essa contenuto /
reproduction of the thesis in whole or in part is forbidden

Data /Date ....3-2-23............. Firma /Signature ....

DECLARATION

This thesis has been:
- composed by myself and has not been used in any previous application for a degree. Throughout the text I use both 'I' and 'We' interchangeably.

  - has been written according to the editing guidelines approved by the University.

Permission to use images and other material covered by copyright has been sought and obtained. For the following image/s (not applicable), it was not possible to obtain permission and is/are therefore included in thesis under the "fair use" exception (Italian legislative Decree no. 68/2003).

All the results presented here were obtained by myself, except for:

1. ***Karyotype analysis*** *(*Figure 2*C-D* Figure 8*C) was performed in collaboration with Marianna Paulis and Anna Villa (IRCCS Humanitas Research Hospital, Rozzano, IT)*
2. ***Editing of CD4+ T-cells with IDLV*** *(*Figure 2*B) was performed in collaboration with Claudia Asperti and Elisabetta Rovelli (Process Development Lab, SR-TIGET, Milan, IT)*
3. ***Functional analysis of CD4+ T-cells.*** *CD40 flow-citometry staining assay was developed by Simona Porcellini (Anna Villa's Lab, SR-TIGET, Milan, IT) upon my suggestion. The HEK-Blue assay was initially set up by myself and later carefully optimized by Simona Porcellini. Results shown in* Figure 7 *were obtained in collaboration with Simona Porcellini and Claudia Asperti (Process Development Lab, SR-TIGET, Milan, IT).*

All sources of information are acknowledged by means of reference.

Aknowkledgements

Abstract

We have been developing a CD4+ T-cell gene correction strategy for the treatment of Hyper-IgM1 (HIGM1), a rare combined immunodeficiency, with Cas9 and a donor template. In parallel with designing a first-in-human clinical trial and developing the manufacturing process, we have been addressing the genome integrity of edited cells. Having previously addressed nuclease off-target activity, we initially focused on the on-target DNA double strand break (DSB), uncovering the presence of large deletions early after editing, that were counter-selected in culture. We then validated and exploited optical mapping, a novel technology that does not require DNA amplification, to assess genome-wide integrity. We documented the presence of recurring on-target large integrations in up to 13-21% of selected cells, compatible with integration of one or more templates. Edited cells could be enriched as expected by the gene correction strategy, indicating that large on-target integration events did not appear to compromise functionality of the cells, as also indicated by suitable functional assays. In summary, edited CD4+ T-cells have a satisfactory genome integrity profile. Large on-target deletions and integrations are possible outcomes of combining nucleases with corrective DNA templates.

Table of contents

Acronyms and Abbreviations

AAV6	Adenoviral-associated virus 6
Cas9	CRISPR associated protein 9
CD154	Cluster of differentiation 154, also known as CD40L
CD40	Cluster of differentiation 40
CD40L	CD40 ligand (protein)
CD40LG	CD40 ligand (gene)
CNV	Copy number variation
CRISPR	Clustered regularly interspaced short palindromic repeats
ddPCR	Droplet digital PCR
DMSO	Dimethyl sulfoxide
DSB	Double strand break
DP	Drug product
GFP	Green fluorescent protein
GLP	Good Laboratory Practice
gRNA	Guide RNA
HIGM1	Hyper IgM Syndrome 1
HDR	Homology directed repair
HSCT	Haematopoietic stem cell transplantation
HSPC	Haematopoietic stem and progenitor cell
IDLV	Integrase defective lentiviral vector
IRES	Internal ribosome entry site
∆NGFR	Low-affinity nerve growth factor receptor
LTR	Long terminal repeats
PGK	Phosphoglycerate kinase
RNP	Ribonucleoprotein
SA	Splice acceptor sequence
SD	Splice donor sequence
SV	Structural variant
SR-TIGET	San Raffaele Telethon Institute for Gene Therapy
TCR	T-cell receptor
VAF	Variant allele frequency
VCN	Vector copy number
WT	Wild type

List of Figures

List of Tables

1. Introduction

Hyper-IgM1 (HIGM1, OMIM #3082) is a rare X-linked combined immunodeficiency caused by mutations in the CD40 ligand (*CD40LG)* gene. *CD40LG* encodes for CD40 ligand (CD40L), also known as CD154 (Tangye *et al*, 2020). The biological and pathophysiological role of CD154 is best understood in regulating the interactions of CD4+ T-cells with cognate cells. While absent from the surface of resting CD4+ cells, CD154 is promptly displayed upon activation, in a tightly regulated manner. Binding of CD154 to its ligand CD40, the latter present on target cells, is required to induce class-switch recombination and somatic hypermutation of B cells. This leads to the characteristic immunoglobulin phenotype of normal to high IgM levels and low levels of the other immunoglobulin classes (Notarangelo *et al*, 2014). The CD40L/CD40 axis has also been implicated in the maturation of myeloid dendritic cells (Cabral-Marques *et al*, 2012; Ferris *et al*, 2020), activation of macrophages (Cabral-Marques *et al*, 2017), and granulopoiesis (Yazdani *et al*, 2019; Cabral-Marques *et al*, 2018).

Patients affected by HIGM1 are vulnerable to infections, particularly of the respiratory and gastrointestinal tract, and to opportunistic pathogens, most commonly *Pneumocistis jiroveci* and *Cryptosporidium parvum*. Neutropenia, liver disease, and increased cancer susceptibility are other debilitating features of the syndrome (Yazdani *et al*, 2019). Median survival is 25 years (de la Morena *et al*, 2017). Hygienic measures, immunoglobulin supplementation, antimicrobial prophylaxis and treatment of infections are the mainstays of treatment. Haematopoietic stem cell transplantation (HSCT) can be a curative option, but is a high risk procedure in patients ≥10 years of age, with pre-existing organ damage, *Cryptosporidium* infection, >2 year from the diagnosis, and without a fully matched donor available (Ferrua *et al*, 2019).

HIGM1 cannot be treated with gene addition, as constitutive expression of *CD40LG* leads to abnormal lymphoid proliferation (Brown *et al*, 1998; Sacco *et al*, 2000). Instead, restoration of physiological gene expression from its native locus can be achieved by gene editing approaches (Kuo *et al*, 2018; Hubbard *et al*, 2016; Vavassori *et al*, 2021b; Ferrari *et al*, 2021). Specifically, Elisabetta Mercuri has shown comparable therapeutic efficacy in treating HIGM1 mice with wild type (WT) T-cells or haematopoietic stem cells (HSPCs), while Valentina Vavassori has developed a Cas9, adenoviral-associated virus 6 (AAV6) based gene editing strategy that restores regulated expression of the gene (Vavassori *et al*, 2021b).

Based on this work and additional experimental data, San Raffaele Telethon Institute for Gene Therapy (SR-TIGET) and Genespire have decided to pursue clinical

translation of gene editing for HIGM1. Both autologous HSPCs and T-cells can in principle be gene edited for therapeutic benefit; a choice to privilege the development of a CD4+ T-cell platform was made taking into account safety and efficacy considerations, a rationale shared also by others in the field (Fox *et al*, 2022). T-cells can be efficiently edited, selected, and infused even without prior administration of chemotherapy or lymphodepleting drugs, and are considered to be less incline to transformation given their more differentiated phenotype. In the pursuit of translating T-cell therapy for HIGM1 we applied for the orphan drug designation of autologous peripheral blood-derived CD4+ T-cells CRISPR-edited at the CD40LG locus with SpyFi Cas9 and an integrase defective lentiviral vector (IDLV) donor template (EMA/OD/0000079230), that received positive opinion by the European Medicines Agency Committee for Orphan Medicinal Products in April 2022 (https://www.ema.europa.eu/en/documents/minutes/ minutes-comp-meeting-11-13-april-2022_en.pdf).

Anticipating clinical translation, we drafted a clinical trial accounting for the patients' characteristics, ongoing therapies, and the drug product profile, outlined in Appendix 2: Trial design synopsis. While gene editing products that have undergone a DNA double strand break (DSB) have already entered clinical testing (Frangoul *et al*, 2020), little is known about preservation of the genome integrity of the drug product before and after infusion, and there is no consensus on the relevant characterization that must be performed. This is in stark contrast for example to established gene addition products, for which vector copy number (VCN), percentage of transduction and integration site analysis are established means for the characterization of genetic manipulation outcomes (Calabria *et al*, 2022).

Indeed, a growing body of evidence has been highlighting the potential genotoxic effects of gene editing (Sasu et al, 2022; Boutin et al, 2021; Ferrari et al, 2022; Nahmad et al, 2022; Leibowitz et al, 2021). Here, we aimed to characterize the genome integrity profile of the drug product (DP), beyond previously published assays of on-target integration of the donor and functional expression of the corrected gene (Vavassori *et al*, 2021b). *CD40LG* editing strategy is a privileged environment to study genome integrity, as hemizygosity limits the number of combinatorial outcomes that otherwise emerge from simultaneous delivery of two DSBs on sister chromatids. As off-target activity of the nuclease has previously been shown to be below the limit of detection (Vavassori *et al*, 2021b), we focused on events occurring on-target, both in terms of molecular and functional readouts, and genome-wide, striving to be as unbiased as possible in evaluating unexpected genomic outcomes.

This work has been performed in parallel with the development of the manufacturing process of the drug product by the SR-TIGET Process Development Laboratory (PDL). Whenever possible, assays were performed on experimental samples representative of the latest PDL development status, in order to best reflect the profile of the final drug product.

2. Aim of the work

The aim of this translational work is assessing the genome integrity of CD4+ T-cells upon gene editing at the *CD40LG* locus with Cas9 and a corrective donor template, in order to generate data on the predicted safety of the editing procedure and develop relevant safety assays for drug product characterization and patient follow-up in the framework of a clinical trial, summarized in Appendix 2: Trial design synopsis.

3. Results

3.1 Large on-target deletions are counter selected in culture

The proof-of-concept of editing CD4+ T-cells for the treatment of HIGM1 has been recently published (Vavassori *et al*, 2021b) and updated (Vavassori *et al*, 2021a). Briefly, CD4+ T-cells were selected from a buffy coat or leukapheresis by immunomagnetic selection and then activated by CD3 and CD28 stimulation, in the presence of IL7 and IL15 (Figure 1A). On day 2 or 3 cells were electroporated with Cas9 ribonucleoprotein targeting *CD40LG* intron 1 and transduced with a viral vector carrying a bicistronic template encoding for the CD40LG corrective sequence functionally coupled with a low-affinity nerve growth factor receptor (ΔNGFR) selector, which is exploitable for the enrichment of correctly edited cells. Notably, we have also modelled enrichment of correctly edited CD4+ T-cells in the murine model of the disease, showing efficacy benefits in terms of immune reconstitution (Appendix Figure 9 and Vavassori et al, 2021a).

For initial experiments an AAV6 was used delivery vehicle of the corrective template, and ΔNGFR expression was placed under the control of a constitutive promoter (Figure 1B) (Vavassori *et al*, 2021b). However, during the development of the manufacturing process IDLV was chosen instead (Figure 2A). In fact, we showed in separate studies that IDLV appears to have a more reassuring safety profile as compared to AAV6 as a delivery vehicle for homology driven repair (HDR) editing in HSPCs, with comparable efficiency (Ferrari *et al*, 2022).

As previously shown in other contexts (Boutin *et al*, 2021; Nahmad *et al*, 2022), we hypothesized that the DNA DSB could result in long range deletions. Moving from a perspective of clinical translation, we reasoned that deletions affecting *CD40LG* alone would be neutral to patients already carrying a non-functional copy of the gene. We chose therefore to investigate the presence of deletions affecting the closest annotated genes neighboring *CD40LG*, i.e. *LNC00892*, located towards the centromere, and *ARHGEF6*, located towards the telomere. *MECP2,* located on the Xq telomere was chosen as additional target to assess loss of the Xq arm (Figure 1C). In this context, whereby a ddPCR assay is used to measure a copy number variation (CNV) in a heterogeneous population - rather than a clonal one -, precision is somewhat limited by stochastic error. Nevertheless, we could detect a significant loss of amplicons mapping onto *LNC00892*, indicating deletions spanning at least 7 kb in the direction of the centromere in edited cells (Figure 1D-E). The CNV nadir was 2-4 days after editing.
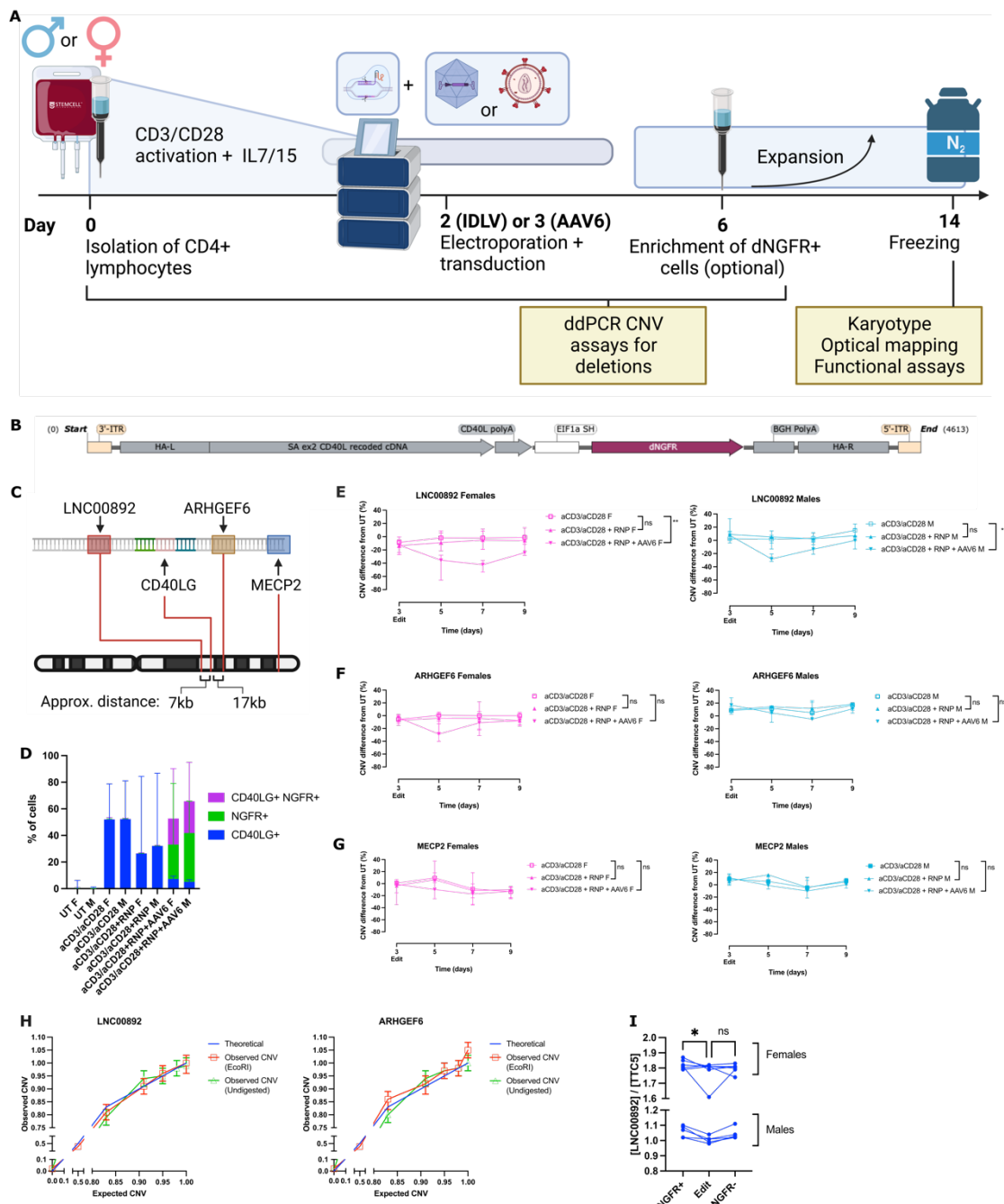
Figure 1: Process overview and on-site large deletions in edited CD4+ T-cells.

A: Overview of the gene editing process. CD4+ T-cells were immunomagnetically selected from buffy coats or leukaphereses, and stimulated with anti-CD3/anti-CD28 beads or transact until day 2 (IDLV process) or 3 (AAV6 process). T-cells were then electroporated with Cas9 ribonucleoprotein targeting CD40LG intron 1 and a donor template carried by either an AAV6 or an IDLV. On day 6 cells could undergo immunomagnetic enrichment for ΔNGFR. Cells were then expanded and frozen after 7-8 days. Created with BioRender.com. B: AAV6 donor template used for deletions experiments. ITR inverted terminal repeats. HA-L homology arm left, SA splce acceptor, SH short, HA-R homology arm right. C: Schematic representation of ddPCR CNV assays' positions on chromosome X with reference to the CD40LG locus. D: Editing efficiency by flow cytometry at day 6-9. N=6 male biological replicates + 6 female biological

Cells harboring deletions were counterselected in culture, as the CNV difference between edited cells and untreated ones progressively decreased, and six days after editing little difference could be seen between the different experimental conditions. Of note, differences appeared to be of greater magnitude in female cells. This was likely due to the presence of twice as many targets for the DNA DSB and greater tolerance to genetic disruptions in presence of the second copy of the X chromosome; another contributing factor could be different assay performance depending on baseline target amplicon abundance. Differences in ARHGEF6 and MECP2 were minor, possibly due to the greater distance from the DNA DSB (approximately 17 kb for *ARHGEF6* vs 7 kb for *LNC00892*) but cannot be excluded with certainty, especially in females (Figure 1F-G).

Based on these results, we reasoned that on-target deletions were a relevant genomic integrity concern, worthy of assessment in the drug product. In order to pursue Good Laboratory Practice (GLP) validation of these assays we thus aimed to generate suitable positive controls. Given the hurdles faced in generating a deletion in the same genomic region (see pages 18 - 26) we screened a panel of restriction enzymes that could digest only the target amplicon while sparing the reference one (i.e. *TTC5*) in a banked SKW 6.4 clone (#9), with known karyotype (Appendix Figure 8C). We found MseI to be suitable, allowing for transfer of the assay to SR-TIGET GLP facility (Figure 1H), along with optimized reaction parameters.

We wondered whether beside potential advantages in terms of efficacy (see Appendix Figure 8 and Vavassori et al, 2021a), enrichment of edited cells with an IRES-coupled ∆NGFR selector could also correspond to a reduction in the number of deletions involving *LNC00892* (Figure 1I). This hypothesis, while hard to demonstrate due to technical limitations - both of the assay and the selection procedures - appeared to be true, even if of very little magnitude especially if compared to the spontaneous purging of deletions observed simply by culturing the cells.

3.2 Edited CD4+ T-cells have a normal karyotype

Karyotype on 100 metaphases is currently the clinical gold standard for assessing large scale acquired genomic events. For this reason, we decided to perform conventional karyotype analysis on three biological replicates of CD4+ lymphocytes edited by the PDL.
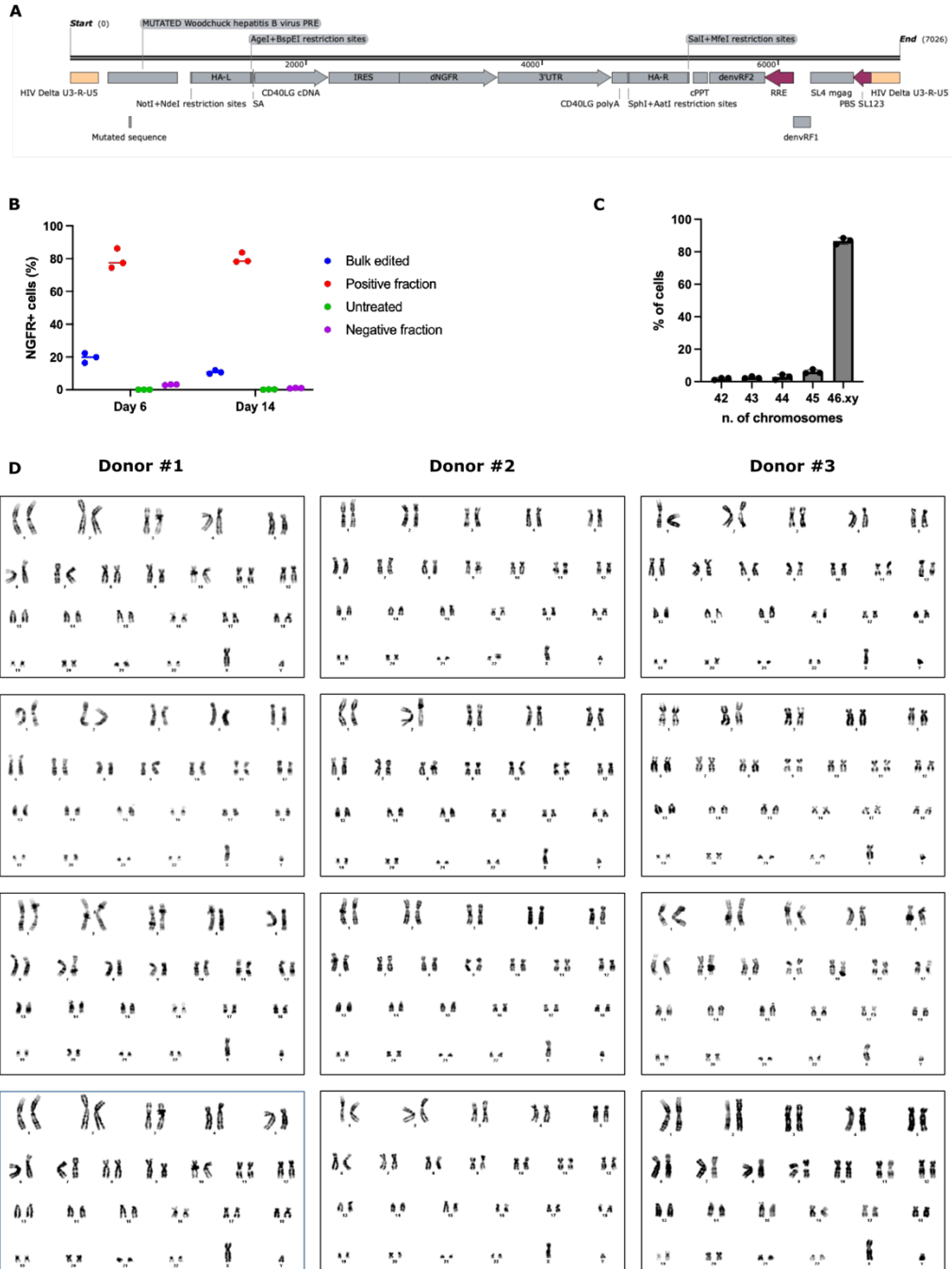
Figure 2: Karyotype of edited CD4+ T-cells.

A IDLV donor construct used to edit CD4+ T-cells. HA-L left homology arm. IRES internal ribosome entry sequence. HA-R right homology arm. C: ΔNGFR expression by flow cytometry

*in CD4+ T-cells at day 6 (day of immunomagnetic enrichment of ΔNGFR+ cells) and day 14 (end of process). **D**: Results of karyotype analysis performed on three bulk edited biological replicates 14 days after editing and representative images (**E**, n=92, 95 and 91 metaphases respectively for donor 1, 2, and 3).*

Editing was performed at the *CD40LG* locus with the donor IDLV illustrated in (Figure 2A), and cells were analyzed at the end of the manufacturing process (Figure 1A). While the manufacturing protocol already foresaw enrichment of edited cells (Figure 2B), we chose to analyze bulk edited cells to maximize sensitivity, reasoning that deleterious events (i.e. loss of the X-chromosome arm) could be purged out during ΔNGFR selection. Karyotype analysis (Figure 2C-D), was considered within normal limits, and anomalies – which were not recurrent – were attributed to sample preparation. Supported by these results, we did not analyze the positive fraction. Instead, we investigated whether optical mapping, a more sophisticated technology for the assessment of structural and copy number variations could be fit for the purpose of better assessing genome integrity, beyond the limits of conventional karyotyping. We initially focused on the ΔNGFR-enriched fraction, which we expected to be the least impacted by adverse genomic events and better suited for infusion into HIGM1 patients.

3.3 Assessment of optical mapping sensitivity

Generation of control samples

Optical mapping is a novel technology for unbiased genome wide assessment of structural variations (SVs) and copy number variations (CNVs) commercialized by Bionano Genomics (San Diego, California). First, the DNA is extracted to preserve the integrity of long fragments, then the DNA is labeled with a fluorescent marker recognizing the DLE-1 sequence CTTAAG, thus creating a restriction map. Linearized DNA molecules are scanned by a dedicated instrument, and images are converted into digital objects called "molecules". Molecules files are then assembled into consensus genome maps, called "assemblies".

Assemblies may then be compared against each other and/or a reference to detect and annotate SVs. CNVs are instead detected by assessing variations in normalized region coverage. The Annotated Rare Variant bioinformatic pipeline was chosen for data analysis in lieu of the De Novo pipeline, as it was better tailored to assess rare events. According to the company, the bioinformatic pipeline could detect SVs of length greater than 5000 bp at 1% frequency with a target coverage of 1600X. SVs smaller than 5000 bp could not be detected with this pipeline, mostly due to computational burdens, while sensitivity decreases to around 10% for events of 250-500 kb (personal communication).
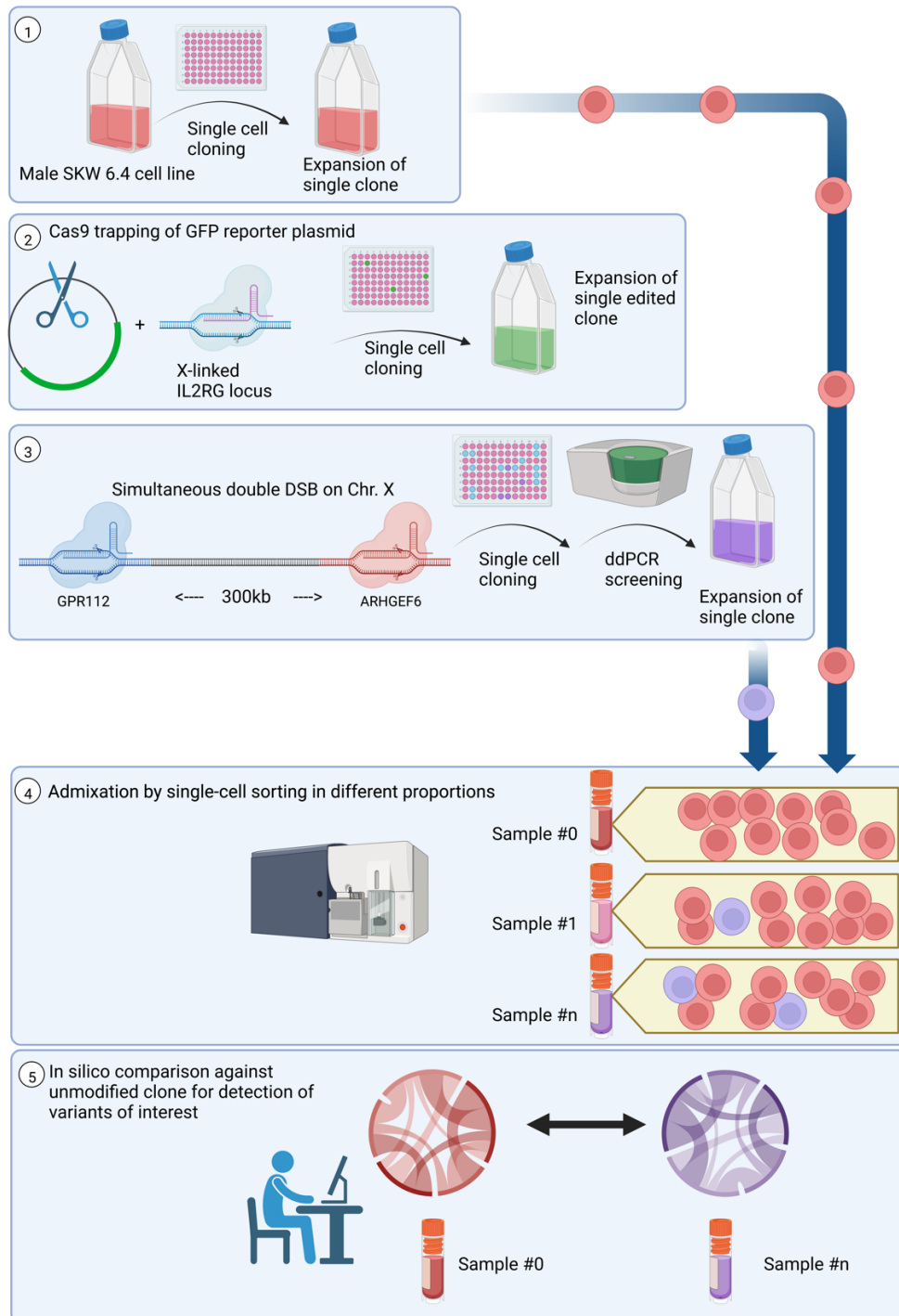
Figure 3: Assessment of optical mapping sensitivity.

1. *A subclone (in red) was generated from a bulk SKW 6.4 cell line.* **2.** *A linearized plasmid constitutively expressing GFP was integrated into the IL2RG locus by trapping in a Cas9 induced DSB. A single clone containing the correct insert was characterized and expanded* **3.** *The same GFP+ clone was cut simultaneously at two 300-kb spaced loci (GPR112 and ARHFEF6), respectively upstream and downstream the CD40LG locus on chromosome X. A single cell clones was derived again after this modification, characterized, and expanded.* **4.** *The clone of interest (purple) was admixed in known proportion with the unmodified clone (step 1, in red) by single cell sorting.* **5.** *Samples admixed at different proportions (pink-purple) were compared against their unmodified counterpart derived in step 1 (in red). Created with BioRender.com.*

Figure 4: Generation of SKW 6.4 clones with genomic rearrangements.

A: Digestion strategy of PGK-GFP reporter plasmid. KpnI and ScaI were used to linearize the plasmid. HA-L homology arm left. hPGK no SA human phophoglycerate kinase promoter. GFP green fluorescent protein. pA polyadenylation sequence. HA-R homology arm right. Note that homology arms are not homologous to IL2RG but rather the AAVS1 locus. *B*: Gel

*purification of digested band, before and after excision from the gel. **C**: GFP copies in SKW 6.4 clones. Green clones were considered to be harboring a single GFP integration; remaining were discarded. "Bis" indicates technical duplicate. Bars indicate 95% confidence interval of CNV estimate. Bars are color coded according to deviation from expected value of 1. **D**: Amplification of IL2RG intron 1 in selected SKW 6.4 clones. HD healthy donor control. #9, #13, #19 and #23 indicate single cell clones. **E-F**: Sequencing of the 5' (E) and 3' (F) genome-plasmid junction. Purple=genome, orange=plasmid, pink=regions deleted from the plasmid. **G**: Alignment of genomic-plasmid junctions on the IL2RG locus (solid red bars within red arrows). A 35 bp deletion is highlighted in magenta. **H**: Reconstruction of the 6.9 kb insertion. Sequenced junctions are shown by dark red arrows. **I**: Large genomic rearrangements detected in selected clones. Bars indicate ratios between ddPCR CNV assays located near on chromosome X, near the centromere (ARHGEF9), on chromosome X, near the telomere (MECP2), within the rearranged region comprised within the Cas9 target sites (ARHGEF6 and LNC00892), chromosome 14 (TTC5), chromosome 19 (EPS8L1), chromosome 1 (EIF2C1). Bars indicate 95% confidence interval of the copy number.*

We decided to validate the sensitivity of the Annotated Rare Variant pipeline on custom biological samples, in order to be confident with the subsequent analysis of gene edited T-cell samples. To this end, we decided to generate a cell line simultaneously harboring two SVs: a first SV close to the lower limit of detection (5 kb), and another one close to the upper limit of detection (250-300 kb). Both events were generated on a hemizygous chromosome X to facilitate data interpretation, generation and validation of the control samples, and exemplify as much as possible gene editing of CD4+ T-cells at the *CD40LG* locus. The clone harboring the two desired SVs was then precisely admixed with the unmodified one in different predetermined proportions, each corresponding a certain variant allele frequency (VAF). In silico comparison of an admixed sample against the unmodified one was then performed to accurately assess technology performance at the desired VAF. Experiment design is detailed in Figure 3.

Initially, a single cell clone was re-derived from a freshly acquired commercial male SKW 6.4 cell line. Integrity of the X-chromosome in the single clone was confirmed with a combination of three ddPCR CNV assays spanning *ARHGEF9*, a centromeric gene, *ARHGEF6*, located on the q arm, and *MECP2*, located at the telomeric end of the q arm (Appendix Figure 8A). One subclone (#2) with no apparent CNV variations of the X chromosome was then chosen for subsequent modifications.

A 9 kb plasmid containing green fluorescent protein (GFP) under the transcriptional control of the human Phosphoglycerate kinase (PGK), with no homology with the genomic target site, was cut and linearized as illustrated in Figure 4A-B, and trapped into the first intron of *IL2RG* using Cas9 ribonucleoprotein (RNP) repurposing a published strategy for the correction of *IL2RG* mutations (Schiroli *et al*, 2019). GFP+ cells were sorted one week after editing, and single cell sub-cloning was performed by sorting one week thereafter. GFP copy number in 23 clones was assessed by ddPCR, to exclude those harboring multiple integrations (Figure 4C). On-

target integration was confirmed by absent amplification of the cut site (Figure 4D) and sequencing of the genome-plasmid junctions. Clone #23 was found to have the desired outcome: a 5' junction with a 457 bp plasmid deletion (mapping onto the wpre sequence) and a 3' junction with the deletion of 9 nucleotides on the genomic side and 21 on the plasmid side (Figure 4E-F), corresponding to a 6.9 kb insertion in *IL2RG* intron 1, along with a 35 bp on-target genomic deletion (Figure 4G-H).

Generation of a large rearrangement in selected clone #23 was pursued by simultaneously cutting the X-chromosome in two loci, *GPR112* and *ARHGEF6*, located upstream and downstream the *CD40LG* locus, and situated about 300 kb from each other (Figure 3). Around 274 single-cell derived clones were screened by ddPCR CNV assays mapping within the targeted region. While no clone containing a 300 kb deletion could be found, 6 clones with relevant genomic rearrangements could be identified using the same ddPCR assays, coupled with others mapping onto the centromeric and telomeric end of the Xq arm, and other reference chromosomes (Figure 4I). Clone #2E04, apparently carrying an additional copy of chromosome 14, and a duplication/triplication of the genomic region containing *LNC00892* and *ARHGEF6* was chosen for further analysis.

Karyotype analysis was performed on samples #23 (with plasmid integration, Figure 4E-H), #2E04 (also carrying a large genomic rearrangement, Figure 4I), as well as two other clones (#9, #11) that had been generated for other purposes and derived from the same original SKW 6.4 cell line, as detailed in Table 1 (representative images are reported in Appendix Figure 8C). The karyotype of sample #2 could not be analyzed for technical issues but could be inferred from the other samples.

The following samples were thus generated admixing single cells from clone #2E04 and clone #2: sample #0 (100% cells from clone #2), sample #1 (90% cells from clone #2 and 10% cells from clone #2E04), sample #5 (97.5% cells from clone #2 and 2.5% cells from clone #2E04), along with others that were not ultimately analyzed further.

Table 1: Karyotype of selected SKW 6.4 clones

Sample	SKW 6.4	#9	#11	#2	#23	#2E04
Derived from		SKW 6.4	SKW 6.4	SKW 6.4	#2	#23
Cells counted, n		23	15	15	15	20
Cells karyotyped, n		17	12	Not done	10	15
Karyotype	Not done	2n: 46,XY, der(21)t(21;?), der(22)t(22:?)	2n: 45,X0, der(21)t(21;?)	der(21)t(21; ?)	2n: 46,XY, der(21)t(21;?)	2n: 47,XY, der(21)t(21;?), +14
Notes	Original bulk cell line	Male karyotype with two chromosome translocations: one chromosome 21 and one chromosome 22 both translocated to unidentified extra chromosomal segments in nearly all the metaphases analyzed	45,X0 karyotype with one of the two chromosome 21 translocated to an unidentified extra chromosomal segment in all the metaphases analyzed.	Presumed to be 46,XY, der(21)t(21; ?)	Male karyotype with one of the two chromosome 21 translocated to an unidentified extra chromosomal segment in nearly all the metaphases analyzed (>90%).	Male karyotype with a trisomy 14 and one of the two chromosome 21 translocated to an unidentified extra chromosomal segment in nearly all the metaphases analyzed (>90%).

Analysis of control samples by optical mapping

To assess platform performance in detecting events occurring at 10% allele frequency we performed comparison of sample #1 against sample #0 with the Variant Annotation Pipeline, as summarized in Figure 5A. The integration of the GFP reporter plasmid #2733 (Figure 4H) was correctly detected as a 6.9 kb insertion with a VAF of 0.16, with high confidence (.99) (Figure 5B). The position of the new label present in the insertion was consistent with the position of the DLE-1 restriction sequence CTTAAG. The rearrangement occurring at the CD40LG locus was called by two distinct albeit overlapping events: a 149 kb duplicated inversion (confidence -1, VAF 0.12) and a 238 kb inversion (confidence 0.83, VAF 0.12), the former being located more towards the centromere and the latter towards the telomere. Manual alignment of the relative label position between the two events indicated the length of the event be approximately 296 kb (Figure 5C-D).

Other significant detected events were an isolated 3q chromosome gain, with a fractional copy number of .5, also present in the parental cell line (Appendix Figure 8D). The puzzling nature of an apparently centromere-less 3q arm gain was interpreted in light of a translocation with chromosome 21 [t(3;21)(q22.1;p11.2)], that had been filtered out due to masking, confidence (0.02) ad VAF (0.02), which also matched the derivative chromosome 21 observed by karyotype (Appendix Figure 8E). The chromosome 14 aneuploidy was also detected at the expected copy number of 2.11 (score 0). Other events are detailed in Appendix Table 3.

We then repeated the same analysis, this time comparing sample #5 against sample #0, to detect sensitivity in detecting events occurring at 2.5% allele frequency. Insertion of plasmid #2733 was detected as a 6.9 kb insertion, with VAF 0.04 and confidence .99. The large rearrangement was filtered out due for failed chimeric score. Other SVs and CNVs that emerged from the analysis are listed in Appendix Table 4.

In summary, we validated performance of optical mapping in detecting events on a hemizygous X-chromosome with relevant biological controls: a relatively short insertion was detected with high confidence at 2.5% allele frequency, while a large 296 kb rearrangement could be detected at 10% allele frequency. Masked regions were found to be a limit of the technology, and a number of event calls appeared to be false positives (Appendix Table 3 and Appendix Table 4).
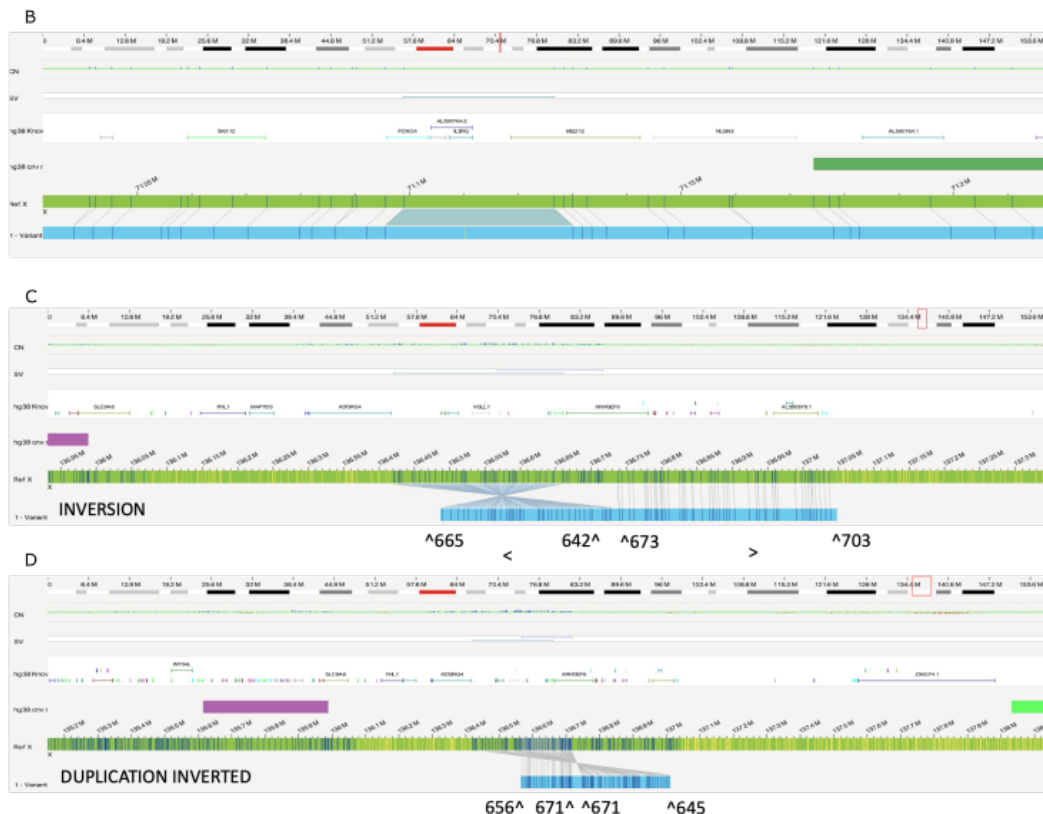
Figure 5: Detection of events at 10% allele fraction by optical mapping.

A: Comparison of Sample #1, containing 10% of cells with genomic rearrangements of the X-chromosome, against Sample #0 consisting of unmodified cells. Circos plot strata, from the most external to the most internal one represent: 1. Chromosome number 2. Conventional

*banding 3. Annotated genes 4. Structural variants, color coded 5. Copy number variations 6. Translocations (none present in current plot). The green dot on chromosome X indicates an insertion in the IL2RG locus, while the blue and purple dots indicate a rearrangement around the CD40LG locus. **B**: Integration of the reporter plasmid in the IL2RG locus. Tracks from top to bottom: 1. Banding of the X chromosome. 2. Copy number variations (green track) 3. Structural variants; green line indicates insertion. 4. Annotated genes. 5. Reference assembly (green, labels in blue). 6. Sample assembly (light blue), with labels matching the reference (dark blue) or not (yellow). **C-D**: Partially overlapping inversion (C) and Inverted duplication (D) involving the Xq arm around the CD40LG locus. Relative label position is annotated, preceded by "^".*

3.4 Optical mapping of edited CD4+ T-cells

To unbiasedly assess the occurrence of unexpected genomic events we performed optical mapping on edited CD4+ T-cells – either enriched for ΔNGFR or not - from the same three male donors on which we had performed karyotype analysis (Figure 2D-E). ΔNGFR expression was 10-12% in the bulk edited population and 78-84% in the positive fraction (Figure 2C).

Dual analyses of positive fractions, each compared against their respective untreated counterpart are reported in Figure 6A. The intended HDR outcome was not expected to be seen, being the size below the technique's threshold of 5 kb. Instead, an insertion in the *CD40LG* locus was found to be recurring in 2/3 donors. The 16.9-17 kb insert was called at a VAF of .13-.21 with high confidence (.99). Closer inspection of the insertions (Figure 6B) revealed that the most abundant event contained one additional DLE-1 label. Moreover, additional events were present, containing either an additional label in a different position, or additional regularly spaced labels. To ascertain whether the ΔNGFR-enriched fraction of donor #1 was truly devoid of insertions at the *CD40LG* locus, we checked if the event was truly absent or present but filtered by the software. In fact, 5 insertion events could be found at the same locus after filters removal: 3 had been filtered due to low molecule counts and 2 for being also apparently present also in the untreated sample. Conversely, no insertion event at the same locus was found upon manual revision of the untreated sample #1. We thus concluded that the insertion event(s) was indeed there also in donor #1, but somehow had been incorrectly filtered by the software.

We used variant allele frequency and relative molecule abundance from each sample annotated rare variant analysis to estimate the frequency of such insertion events across every donor and condition (Figure 6C). We found that 8-10 kb events were the most abundant, and that overall the frequency of insertion events increased with the percentage of ΔNGFR+ cells in the sample. We speculated that the new labels present in Figure 6B could correspond to recognition of the long terminal repeats (LTRs) present in the donor IDLV (Figure 2B), that contain a DLE-1 restriction site. Of note, the resolution of the technique was not sufficient discriminate between

26

one label and two closely (<1-2 kb) spaced ones, as could be expected in case of concatemer formation.
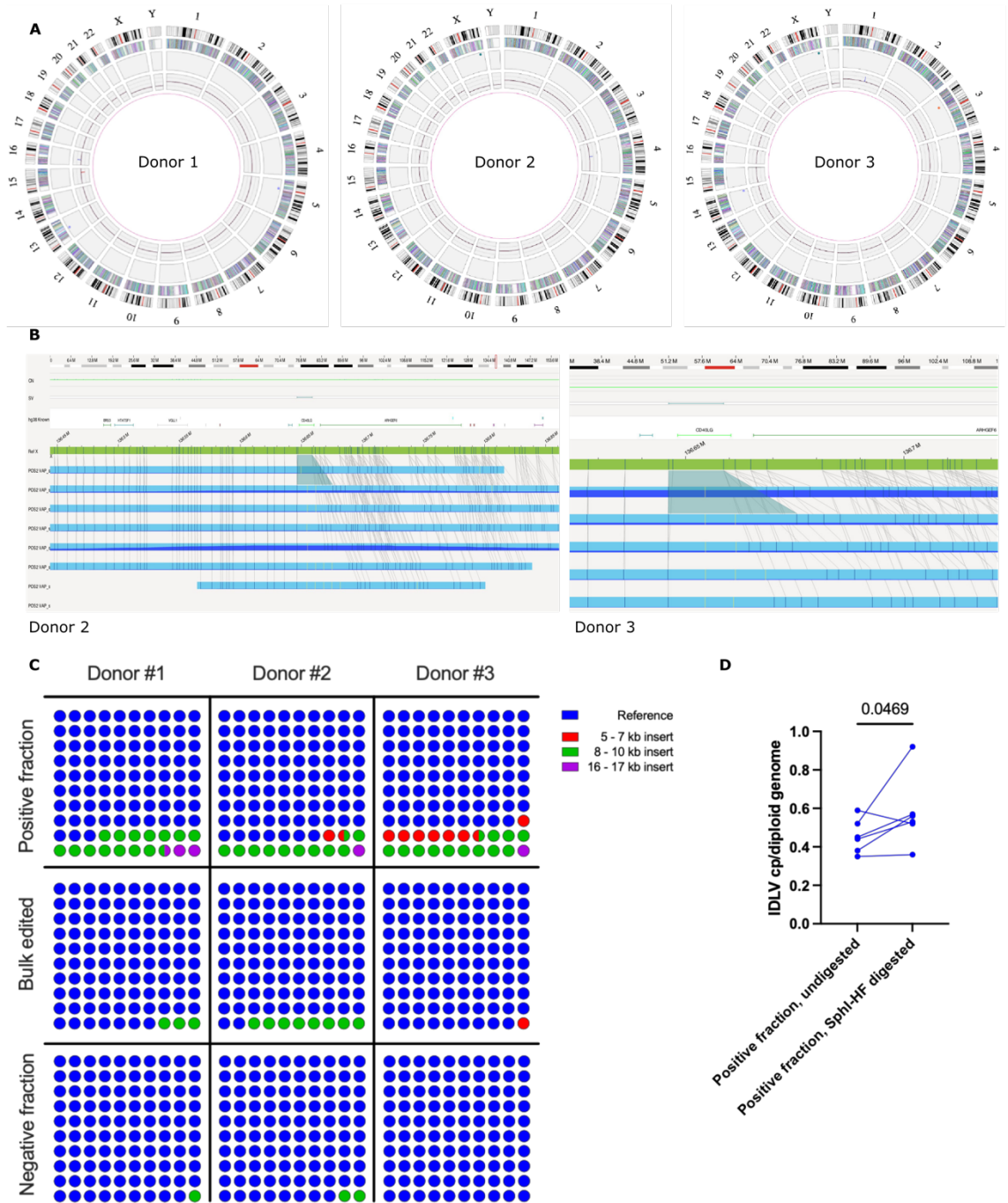
Figure 6: Optical mapping of edited CD4+ T-cells.

A: Circos plots of ΔNGFR+ CD4+ T-cells compared with each donors' cultured unedited counterpart. Circos plot strata, from the most external to the most internal one represent: 1. Chromosome number 2. Conventional banding 3. Annotated genes 4. Structural variants, color coded as in Figure 5A. 5. Copy number variations 6. Translocations (none present in current plot). The green dot on chromosome X indicates an insertion in the CD40LG locus. B: Representation of insertion events occurring around in CD40LG locus in donor#2 (left) and

*donor #3 (right). Tracks from top to bottom: 1. Banding of the X chromosome. 2. Copy number variations (green track) 3. Structural variants; green line indicates insertion (also represented by shaded trapezium). 4. Annotated genes. 5. Reference assembly (green, labels in blue). 6. Sample assemblies (light blue, shaded in dark blue according to relative abundance), with labels matching the reference (dark blue) or not (yellow). **C**: Dot plot matrix indicating the frequency of insertions at CD40LG of different size in edited samples. **D**: VCN of the IDLV in the same 3 positively selected samples and 3 additional positively selected samples from similar experiments. One-tailed Wilcoxon matched-pairs signed rank test.*

We sought to confirm this hypothesis with a ddPCR linkage assay. We reasoned that IDLVs grouped in concatemers, being close to one another, were very likely to segregate together in the same oil droplet, leading to underestimation of IDLV copies in the sample. Thus, we thought that enzymatic digestion the DNA before droplet generation could separate the tandem repats of target amplicons, allowing for independent segregation into oil droplets, uncovering the true VCN. Indeed, we found a median VCN increase of 0.1 after enzymatic digestion, indirectly confirming the presence of concatemers (Figure 6D).

Other events occurring in the bulk and ΔNGFR enriched fraction are reported in Table 2. We did not detect recurrent events, translocations, nor events occurring at the previously published non-HiFi Cas9 off-target in chr8:59861013-59861036, nor at the other potential off-targets chr5:119086385-119086409, chr6:105224111-105224135, chr8:47680153-47680175 (Vavassori *et al*, 2021b). Notably, we also did not detect on-target deletions involving LNC00892 by optical mapping at the end of the manufacturing process in any edited sample, selected or not, even without event filtering.

We reviewed CRISPOR predicted off-targets for the *CD40LG* gRNA sequence (TGGATGATTGCACTTTATCA-nGG) falling in the regions reported in Table 2 for homology with and found only one potential, low-scoring, off-target within event 11 (Appendix Table 5) using CRISPOR (Concordet & Haeussler, 2018). Sequences homologous with the CD40LG were also sought with a very relaxed approach for regions involved in events 6, 9 and 10, i.e. those not appearing to be false positives (Appendix Table 6).

Table 2: Off-target genomic events in edited CD4+ T-cells.

Dual analysis by variant annotation pipeline. Chr. Chromosome. VAF Variant allele frequency. Pos ΔNGFR enriched cell fraction. Asterisk indicates likely false positives.

#	Sample	chr:position in hg38	Event	Size (kb)	VAF	Notes
	Donor 1					
1	Bulk	9:41545212-41552664	Deletion	3	0	Present in .6% of control samples database *
2	Bulk /pos	15:75759464-77912434 (bulk) / 15:75759464-77912434 (pos)	CNV Loss	2152 (bulk) / 2153 (pos)	.5 (bulk) / .46 (pos)	Present in UT*
3	Bulk /pos	16:34944870-35898364 (bulk) / 16:34905599-35614048 (pos)	CNV Gain	953 (bulk) / 708 (pos)	.48 (bulk) / .43 (pos)	Partial CNV mask overlap, near centromere*
4	Bulk	Y:23230658-23264381	Duplication	33	.01	Masked region, overlapping DAZ2*
5	Pos	5:46874164-46890530	Inverted duplication	16	.5	Present in .6% of control samples database*
6	Pos	13:57143300-57261030	Inverted duplication	117	.5	Overlapping PRR20A-E, PRRFP. 7 molecules
	Donor 2					
7	Bulk /pos	4:91501905-92528369 (bulk) / 4:91481313-92478681 (pos)	CNV Gain	1026 (bulk) / 997 (pos)	.42 (bulk) / .41 (pos)	Present in UT*
	Donor 3					
8	Bulk /pos	1:144453902-145230161 (bulk) / 1:144416847-145264785 (pos)	CNV Gain	776 (bulk) / 847 (pos)	.60 (bulk) / .06 (pos)	Present in UT*
9	Bulk	2:57670638-57700364	Deletion	29	.02	No annotated genes
10	Pos	3:58006840-58123739	Deletion	104	.04	Overlapping FLNB, 9 molecules
11	Pos	14:98239205-98427434	Duplication	188	1	Overlapping RN7SL714P. Detected also in bulk with VAF 1 and failed chimeric score, not in negative fraction or UT*

3.5 Functional assessment of edited CD4+ T-cells

To complete safety and efficacy characterization of IDLV edited CD4+ T-cells we sought to confirm expression of CD154 with the same kinetic of its native counterpart, the gene product recapitulated the physiologic binding to CD40 and downstream signal transduction. Indeed, we found that both patient edited and healthy donor edited CD4+ T-cells were able to bind CD40 with a kinetic similar to that of healthy controls and induce intracellular signal transduction (Figure 7).
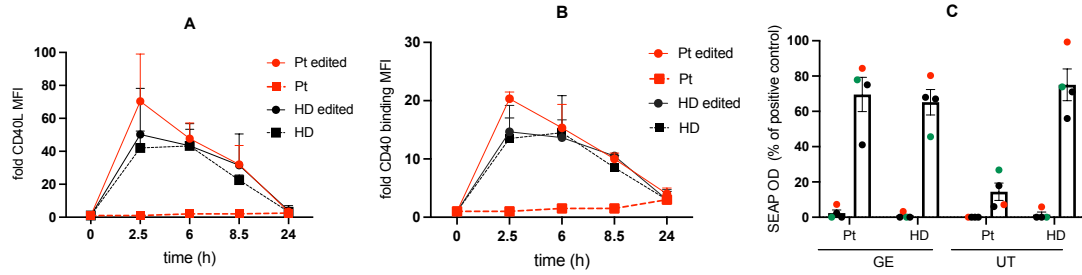
Figure 7: Functional characterization of edited CD4+ T-cells.

A Fold expression of CD40LG median fluorescence intensity (MFI) after PMA-ionomycin stimulation B Fold MFI of surface bound CD40muIgG C Relative quantification of embryonic alkaline phosphatase (SEAP) optical density (OD) in the surnatant of edited CD4+ T-cells co-cultured with HEK-Blue™ CD40L reporter cells. HEK-Blue™ cells release SEAP in the culture medium upon activation of the NF-kB pathway by CD40. Pt patient. HD healthy donor. GE gene edited. UT untreated. Ctrl control.

4. Discussion

Nucleases can be precisely programmed to deliver a DNA DSB in a specific locus, with no detectable off-target activity, and by providing a donor template it is possible to functionally correct most – if not all – mutations of a given gene (Ferrari *et al*, 2021). It is becoming increasingly clear however that the DNA DSB itself is not only repaired by NHEJ, microhomology mediated end joining (MM-EJ) or HDR. Instead, for yet unclear reasons, a DNA DSB may result in a long-range deletion stemming from the DNA DSB and propagating towards one or both ends of the chromosome. These events have likely been underestimated due to the loss of the primer binding sites often used to evaluate the frequency of NHEJ/MM-EJ and HDR.

By tiling the regions of interest near the DNA DSB we found deletion events equal to 7 kb or more, reaching the neighboring gene *LNC00892*, appearing early after editing, and that were spontaneously counter-selected over time. *ARHGEF6*, located 17 kb in the opposite direction was relatively spared, possibly simply due to the greater distance from DNA DSB. However, the trend observed in females cannot fully exclude this possibility. Indeed, the accuracy and precision of these CNV assays are somewhat limited when analyzing bulk populations, rather than clonal ones (Ferrari et al, 2022 and Fiumara*, Ferrari* et al, submitted). Nevertheless, they can be suitable for routine drug product characterization, to exclude unexpected clonal expansion stemming for example from the deletion of an oncosuppressor gene or truncation of a proto-oncogene.

In principle, both on-target and off-target genome integrity – as long as the off-target is known - can be quantitatively assessed with this kind of design, tiling the genome as much as desired. These assays have the advantage of not relying on conservation of the PCR primer binding sites, and thus do not risk underestimating the number of events. Conversely, they raise the issue of how to accurately quantify translocation events with heterogeneous junctions, arising from the joining of DNA DSBs that have undergone long range deletions.

Safety precautions dictate unanticipated outcomes be also investigated with the least possible degree of bias. While 100-metaphase karyotype remains the necessary first step, its sensitivity is hardly sufficient for this context. Optical mapping promises unbiased, amplification-free detection of rare genomic SVs and CNVs greater than 5 kb, with the obvious downside of not providing sequence information. Indeed, we confirmed sensitivity of the platform in detecting a hemizygous 5 kb insertion at 2.5% allele frequency and a large hemizygous genomic rearrangement of about 300 kb at 10% allele frequency. As for other technologies, masked regions and filtering are

intrinsic limits and potential caveats. We noticed that integration of karyotype and optical mapping results can be helpful in detection of events that cannot be solved separately by each platform.

Both karyotype analysis and optical mapping of IDLV-edited CD4+ T-cells at the end of the manufacturing process yielded reassuring results, with no detection of on-target deletions nor translocations with the off-target previously detected with non-HiFi Cas9. As we were unable to generate a SKW 6.4 clone with a deletion in the same region, and deletion events were purged over time, we speculate that cells harboring large deletions in this region are not viable. In principle, deletions of heterogeneous length could still be present at low frequency and not be detected by optical mapping; still, even if present, cells harboring them have not undergone a clonal expansion reaching the limit of detection of the technology.

Interestingly, we were able to further detail the heterogeneity of repair outcomes of the DNA DSB, while also estimating their relative frequency. Imprecise HDR, whereby a longer than expected template is incorporated into the genome appears to be a relatively common event. Our data suggests that these integrations may likely be IDLV templates containing LTRs, often arranged in concatemers. It is presently unknown in which direction each vector is integrated, or what is the nature of each genomic-vector junction, be it HDR, trapping or other imprecise events. However, a significant fraction of the junctions between *CD40LG* intron 1 and the (first) template are functional, as indicated by the expression of ΔNGFR, and the possibility of enriching for this surface marker also cells with long integration events. Furthermore, we confirmed appropriate binding to CD40 – with a kinetic similar to healthy donors, and downstream signal transduction. Characterization of these large scale events is a clear advantage of the long read throughput of optical mapping, combined with the presence of the DLE-1 target sequence within the LTR. Given the currently prohibitive costs of routine application of optical mapping on each drug product, we reason that this technology is best fitted for non-routine product characterization, that can guide in the choice of custom, simpler ddPCR and functional assays.

We cannot draw definitive conclusions concerning non-recurring genomic events found in edited CD4+ T-cells. Most appear to be false positives, either found in the untreated controls (#2, #7, #8), in healthy subjects (#1, #5), falling into masked regions (#3, #4) or at unreasonably high VAF (#11). A low-scoring CRISPOR-predicted off-target could only be found in the region involved in event #11, and its low predicted efficiency is hard to reconcile with the VAF of 1 observed in both bulk and ΔNGFR enriched fraction. For events #6, #9, and #10 it is currently impossible to determine whether the remaining are truly consequences of editing, but the

likelihood appears to be low, given the absence of predicted/validated off-targets in the region and the very lax similarity with the gRNA. Of note, we found a significant number of false positive events with a high VAF, not compatible with the sample admixing rate, when comparing edited SKW 6.4 cells with their unmanipulated counterpart. Of note, while SVs are filtered by the dual analysis pipeline, CNVs are not (Bionano Genomics personal communication).

In summary, CD4+ T-cells can be edited at the *CD40LG* locus with a donor IDLV template. Integration of the donor template may be imprecise, but this does not appear to impair the functionality of the drug product, and no other recurrent events were detected. For the purpose of clinical translation for the treatment of HIGM1 we believe we have generated a reasonable body of data to guide the development of appropriate routinely applicable genome integrity assays, further consolidating the rationale of our ddPCR assay tiling strategy (Ferrari *et al*, 2022).

From a more general perspective, large on-target concatemers are a novel finding that must be accounted for when designing integration strategies. For instance, one might question the opportunity of tandem integrations of a strong promoter or of a full protein-coding sequence, potentially nullifying (or greatly hampering) the precise gene regulation that is to be achieved with gene editing.

A first limitation of this work is intrinsic to the analysis of bulk, heterogeneous populations. Novel single cell DNA ddPCR assays and sequencing studies may possibly elucidate more granularly the nature of genomic outcomes following gene editing. Second, different vectors were used in the various stages of the project, as it evolved in parallel with the strategic choices taken during the development of the drug manufacturing process. Differences between AAV6 and IDLV in generating on-target deletions and other complex rearrangements early after editing are very relevant questions that are yet to be answered.

5. Materials and methods

Culture of SKW 6.4 cells

SKW6.4 (ATCC-TIB-215TM lot 5082527) human B lymphoblasts were cultured in RPMI 1640 medium (Corning), 2% glutamine, 100 IU/ml penicillin, 100 µg/ml streptomycin, 10-20% heat-inactivated fetal bovine serum (FBS, Euroclone).

Enzymatic digestion and purification

40 ug of plasmid DNA were digested overnight at 37°C with 100 U of Kpn1-HF (New England Biolabs) and 100 U of Sca1-HF (New England Biolabs) in rCutsmart buffer 1X. The desired band was excised from 1% agarose gel and purified with NucleoSpin Gel and PCR Clean-up kit (Macherey-Nagel) according to manufacturer's instructions.

DNA Extraction

Genomic DNA was extracted with QIAamp DNA Micro Kit (QIAGEN) or KingFisher Flex (Thermo Scientific) according to manufacturers' instructions.

ddPCR assays

For digital droplet PCR (ddPCR) analyses we analyzed 5–50 ng of gDNA per reaction with the QX200 Droplet Digital PCR System (Bio-Rad). The following Bio-Rad commercial assays were used: ARHGEF9 dHsaCP2506327/dHsaCP2506706; ARHGEF6 dHsaCNS170742226; MECP2 dHsaCP1000579; EIF2C1 dHsaCP000002; EPS8L1 dHsaCNS737827302. The following ThermoFisher commercial assays were used: LNC00892 Hs04111584_cn, ARHGEF6 Hs00889052_cn.

The following non-commercial assays were also used:

Application	Primer	Sequence
GFP	Forward	CAGCTCGCCGACCACTA
	Reverse	GGGCCGTCGCCGAT
	Probe	CCAGCAGAACACCCCC (MGBNFQ/FAM)
TELO	Forward	GGCACACGTGGCTTTTCG
	Reverse	GGTGAACCTCGTAAGTTTATGCAA
	Probe	TCAGGACGTCGAGTGGACACGGTG (TAMRA/VIC)
IDLV	Forward	TCACTCCCAACGAAGACAAGATC
	Reverse	GAGTCCTGCGTCGAGAGAG

Human TTC5 dHsaCP2506733 or dHsaCP2506310 (Bio-Rad) was used for normalization. Thermal conditions for commercial ddPCR assays were 95°C x 10', (94°C x 30'', 60°C x 1') x 40 cycles, 98°C x 10'.

To estimate the frequency of concatemers, samples were digested or not with SphI-HF in rCutsmart buffer (R3182S, New England Biolabs) for 1 hour at 37°C and then heat inactivated for 20 minutes at 80 °C. ddPCR assay was performed in parallel on digested and undigested samples using IDLV primers with EvaGreen reagents (mastermix and oil). The concentration of TELO amplicons was used as reference. Thermal conditions were 95°C x 5', (95°C x 30", 63°C x 1') x 40 cycles, 4°C x 5', 90°C x 5'.

Copy numbers and confidence intervals were calculated with QXManager version 1.2 (Bio-Rad), assuming two copies of normalizer per genome.

Editing of SKW 6.4 cell line

Briefly, SKW6.4 were edited at IL2RG intron 1 as previously described (Schiroli *et al*, 2019), using electroporation code for K562 and SF Cell Line 4D-Nucleofector™ X Kit S (Lonza).

gRNAs targeting ARHGEF6 and GPR112 were designed with CRISPOR (Concordet & Haeussler, 2018). The following gRNAs were used to target *ARHGEF6*: #1 CTCTCCCCAAAAGCCGTCAA(AGG), #2 GTATAATAATTCTTGGTAAG(TGG), #3 TACCAAGAATTATTATACTG(TGG). The following gRNAs were used to target *GPR112*: #1 AAATAGCGTTGCCCCATACA(CGG), #2 TGTATACTTGCTCGGGTGAT(GGG), #3 CCCATACACGGTCCTGAAAC(TGG). Each gRNA targeting *ARHGEF6* was electroporated with each gRNA targeting *GPR112* and vice versa. Single cell clones were assessed for CNV of the genes contained within the targeted region (*LNC00892* or *ARHGEF6*).

Sorting and freezing of SKW 6.4 clones

Admixation of clone #2E04 and clone #2 was performed by counting single cells with BD FACSAria Fusion (BD Biosciences), as TC20 Automated Cell Counter (Bio-Rad) coefficient of variation was experimentally determined to be too high (>15%, see Appendix Figure 8B). Cells, eluted in 1.5 ml of sheath fluid, were sorted in cryovials containing 1,5 ml of heat inactivated fetal bovine serum. Dimethyl sulfoxide (DMSO, D8418, Sigma) was added to a final concentration of 10%. Cells were frozen with KryoPlaner 560-16 Software version v5.26 (Planer Limited, Sunbury-On-Thames, UK) and stored in nitrogen vapors.

Culturing and gene editing of CD4+ T-cells with AAV6

Culture and editing was performed as previously described (Vavassori *et al*, 2021b), with the following modifications. White blood cells were isolated from buffy coats using Human Lympholyte® Cell Separation Media (Euroclone) and SepMate™-50 (Stemcell Technologies), according to the latter's instructions. CD4+ lymphocytes cells were isolated by negative selection with CD4+ T Cell Isolation Kit, human (Miltenyi Biotec). Cells were cultured X-Vivo + 100 IU/ml penicillin + 100 µg/ml streptomycin (Lonza, cat. no. DE17-602E), IL-15 5 ng/ml (Peprotech), IL-7 5 ng/ml (Peprotech), and activated with Dynabeads™ Human T-Activator CD3/CD28 for T Cell Expansion and Activation (ThermoFisher Scientific).

Culturing and gene editing of CD4+ T-cells with IDLV

Editing with IDLV was performed in the SR-TIGET process development laboratory to be as representative as possible of the drug substance manufacturing process. Briefly, CD4+ T-cells were isolated by immunomagnetic positive selection from buffy coats of healthy donors with Buffy Coat Straight from CD4+ (Miltenyi Biotec). CD4+ T-cells were cultured in X-Vivo 15 (w/out phenol red, Lonza) supplemented with 0,5% Human Serum Albumin (Baxalta), 1% Pen/Strep (Lonza), IL7 100 U/ml and IL15 200 U/ml (both from Miltenyi Biotec). and stimulated with T Cell TransAct (Miltenyi Biotec cat- no. 130-111-160) at $2x10^6$ cells/ml. On day 2 cells were transduced at $1x10^6$ cells/ml with CD40LG-IRESNGFR-IDLV at MOI 40 and after 8±1 hours electroporated with SpyFi™ high fidelity Cas9 complexed with a single guide RNA (gRNA) targeting the first intron of *CD40LG* (Vavassori *et al*, 2021b)*.* On day 6 cells were enriched by immunomagnetic anti-NGFR beads (CD271 beads, Mitenyi Biotec) and subsequently cultured until day 13-14 in X-VIVO 15 w/out phenol red 5% Human Serum (Pan Biotech) 1% Pen/Strep (Lonza), IL7 100 U/ml, IL15 200 U/ml and IL2 50 U/ml (all from Miltenyi Biotec). Cells were then frozen in saline solution 7% Human Serum Albumin (Baxalta) 5 % DMSO with a controlled temperature rate freezing device, KryoPlaner 560-16 Software version v5.26 (Planer Limited, Sunbury-On-Thames, UK) and stored in nitrogen vapors.

Sequencing

Primers for sequencing of IL2RG intron 1 have been previously reported (*Schiroli et al*, 2017). The following primers were used for amplification and sequencing:

Application	Primer	Sequence
SKW 6.4 clone #23 5' junction	Forward	CACCCTCTGTAAAGCCCTGG
	Reverse	AAGACGGCATGGGGTTGGGT
SKW 6.4 clone #23 3' junction	Forward	TTCATTCAGCTCCGGTTCCC
	Reverse	TCATCTCCCCTCAACCGACT

Chromosome Preparation and Analysis

Karyotype analysis was performed by the Stem Cell Lab, IRCCS Humanitas Research Hospital, Rozzano, IT by dr. Marianna Paulis. Chromosome analysis was done on slide preparations of cell suspensions. Briefly, cell line cultures or CD4+ T-cells 48h after stimulation were treated with KaryoMAX Colcemid solution (Thermo Fisher Scientific) at a final concentration of 0.1 µg/ml for 2 h at 37°C and the cell suspension was then centrifuged (1000 rpm for 10 min). After hypotonic treatment with 0.075 M KCl and fixation in methanol:acetic acid (3:1 v/v), the cell suspension was dropped onto a slide and air dried. Chromosome counts and karyotype analyses were done on metaphases stained with Vectashield mounting medium with DAPI (Vector Laboratories) for banding (equivalent of Q-banding).

Images were captured using an Olympus BX61 Research Microscope equipped with a cooled CCD camera and analyzed with Applied Imaging Software CytoVision (CytoVision Master System with mouse karyotyping). To exclude the presence of aneuploidy or chromosome rearrangements, analysis of at least 90/100 metaphases was performed for each sample. Banding: DAPI banding (Q-banding). Resolution: 350/400 band level.

Optical mapping

Samples were shipped to Bionano Genomics (Clermont-Ferrand, France) in dry ice for DNA extraction and DLE-1 optical mapping, labelling the sequence CTTAAG. Cells in freezing media were thawed 2 min at 37° C and DNA was extracted according to Bionano Prep SP Frozen Cell Pellet DNA Isolation Protocol v2 - 30398, Rev B, from the SP Blood & Cell Culture DNA Isolation Kit. DNA was labeled according to Bionano Prep Direct Label and Stain (DLS) Protocol - 30206, Rev G and loaded on the chip the day after.

Target coverage was 1600X (effective coverage 1411.77X-1491.58X) for all samples except sample 0 (effective coverage 397X), which was run with lower target coverage. The Rare Variant Pipeline (RVP) and Variant Calling were executed on Bionano Solve software (v3.7). Reporting and direct visualization of structural variants was done on Bionano Access (v1.7), using Hg38 as reference genome. Each sample was compared against its unedited, cultured counterpart in paired analysis. The following filters were used: 1. SV confidence: recommended. 2. Non masked SV only 3. SV in less or equal to 1% of the paired control db samples. 4. SV found in self molecules 5. SV in less than or equal to 1% of paired control db samples 6. SV not found in paired control assemblies 7. SV nor found in paired control molecules 8. CNV filters: all variants, recommended confidence, min size 500000 bp, non masked only 9. Aneuploidy: all, recommended confidence scores. Events were annotated after filtering. The frequency of insertions of different lengths at the *CD40LG* locus was estimated from the unpaired rare variant analysis by multiplying the VAF for the relative molecule abundance of each event.

Functional assays

Cryopreserved CD4+ T-cells were thawed; after an overnight resting in X-Vivo™ 15 medium (Lonza) with 5% Human Serum (Pan Biotech), 100 U/ml Pen-Strep cells were activated using phorbol myristate acetate (PMA, 10 ng/mL Sigma) plus ionomycin (500 ng/mL, Sigma) for 5 hours and surface expression of CD40L was followed over a 2.5/24-hours time course. T-cells were stained with a panel of antibodies including monoclonal mouse anti-human CD271, CD3 (BD Biosciences), CD4 (Miltenyi), CD154 (Biolegend) or CD40muIgG fusion protein (Vinci Biochem). CD40L fold expression or CD40 binding were calculated as follows: median fluorescence intensity (MFI) of the stimulated sample / MFI of the unstimulated sample.

To evaluate transduction across the CD40L/CD40 axis T-cells, from both healthy donor and patient were co-cultured with HEK-Blue™ CD40L cells (InvivoGen, cat. No. hkb-cd40). This commercial cell line is engineered in order to allow detection of bioactive and functional CD40L. CD40L-CD40 interaction leads to the activation of NF-kB pathway which results in embryonic alkaline phosphatase (SEAP) production and secretion in the supernatant. Briefly, cells were stimulated with PMA/Ionomycin or left unstimulated. After 2.5 hours, CD4+ T-cells were harvested, washed and co-cultured in 1:1 ratio with HEK-Blue™ CD40L cells for 24 hours. Recombinant human CD40L (0.1 μg/ml InvivoGen, cat. rcyec-hcd40l) was used as positive control. Unstimulated T-cells were used as negative control. Given that HEK cells also respond

to IL-1β and TNF-α through NF-kB pathway, neutralizing antibodies anti hIL-1β (10 μg/ml Invivogen, cat. No. mabg-hIl1b-3) and anti-hTNF-α (10 μg/ml InvivoGen, cat. No. htnfa-mab1) were added to the co-culture in order to block SEAP production potentially derived from IL-1β and TNF-α receptor stimulation. After 24 hrs, secreted SEAP was measured in the supernatant by QUANTI-Blue solution (InvivoGen cat. no. rep-qbs), a colorimetric enzyme assay developed to determine alkaline phosphatase in biological samples. The absorbance was read by Omega reader (BMG Labtech) at 650 nm.

Flow-citometry

50-500k cells were incubated with conjugated antibodies for 10 minutes at 4°C and then washed in Dulbecco's Phosphate-Buffered Saline (Corning) + 2% heat inactivated FBS (Euroclone). At least 20000 events were acquired with BD FACSCanto II (Becton Dickinson). Data was analyzed with FCS Express 7 Research (De Novo Software) or FlowJo software (Tree Star, Inc.).

Antibodies

Target	Fluorochrome	Vendor	Clone	Cat. No.	Dilution
CD3	PeCy7	Biolegend	HIT3a	300316	5:100
CD3	FITC	BD Pharmingen	SK7	345763	1:33
CD4	PB	BD Pharmingen	RPA-T4	558116	1:100
CD4	Viogreen	Miltenyi biotec	REA 623	130-113-230	1:100
NGFR	APC	Miltenyi biotec	ME20.4-1.H4	130-113-418	2:100
NGFR	AF647	BD Pharmingen	C40-1457	560326	1:25
CD154	PE	Invitrogen	24-31	12-548-42	5:100
CD154	PE	Biolegend	24-31	310806	1:50
7-AAD Viability Staining Solution	7-AAD	Biolegend	-	420404	1:100
CD40muIgG fusion protein	PE	Vinci Biochem	-	ANC-504-050	1:50

Statistical analysis

Non-parametric statistical analyses were performed only on experimental data with at least 5 replicates, using Prism 9 (GraphPad Software). Details are reported in each corresponding figure legend.

6. References

Boutin J, Rosier J, Cappellen D, Prat F, Toutain J, Pennamen P, Bouron J, Rooryck C, Merlio JP, Lamrissi-Garcia I, *et al* (2021) CRISPR-Cas9 globin editing can induce megabase-scale copy-neutral losses of heterozygosity in hematopoietic cells. *Nat Commun* 12: 4922

Brown MP, Topham DJ, Sangster MY, Zhao J, Flynn KJ, Surman SL, Woodland DL, Doherty PC, Farr AG, Pattengale PK, *et al* (1998) Thymic lymphoproliferative disease after successful correction of CD40 ligand deficiency by gene transfer in mice. *Nat Med* 4: 1253–1260

Cabral-Marques O, Arslanian C, Ramos RN, Morato M, Schimke L, Soeiro Pereira PV, Jancar S, Ferreira JF, Weber CW, Kuntze G, *et al* (2012) Dendritic cells from X-linked hyper-IgM patients present impaired responses to Candida albicans and Paracoccidioides brasiliensis. *J Allergy Clin Immunol* 129: 778–786

Cabral-Marques O, França TT, Al-Sbiei A, Schimke LF, Khan TA, Feriotti C, da Costa TA, Junior OR, Weber CW, Ferreira JF, *et al* (2018) CD40 ligand deficiency causes functional defects of peripheral neutrophils that are improved by exogenous IFN-ɤ. *J Allergy Clin Immunol* Mar 5.: pii: S0091-6749(18)30318-X.

Cabral-Marques O, Ramos RN, Schimke LF, Khan TA, Amaral EP, Barbosa Bomfim CC, Junior OR, Franca TT, Arslanian C, Correia Lima JDC, *et al* (2017) Human CD40 ligand deficiency dysregulates the macrophage transcriptome causing functional defects that are improved by exogenous IFN-ɤ. *J Allergy Clin Immunol* 139: 900–912

Calabria A, Spinozzi G, Cesana D, Benedicenti F, Pais G, Scala S, Lidonnici MR, Scaramuzza S, Albertini A, Esposito S, *et al* (2022) Adaptive Routes of Hematopoietic Stem Cell Differentiation to Disease Conditions and Age in Gene Therapy Patients. *Blood* 140: 1904–1905

Chang C-CH, Bryce CL, Shneider BL, Yabes JG, Ren Y, Zenarosa GL, Tomko H, Donnell DM, Squires RH & Roberts MS (2018) Accuracy of the Pediatric End-stage Liver Disease Score in Estimating Pretransplant Mortality Among Pediatric Liver Transplant Candidates. *JAMA Pediatr* 172: 1070

Ciceri F, Bonini C, Stanghellini MTL, Bondanza A, Traversari C, Salomoni M, Turchetto L, Colombi S, Bernardi M, Peccatori J, *et al* (2009) Infusion of suicide-gene-engineered donor lymphocytes after family haploidentical haemopoietic stem-cell transplantation for leukaemia (the TK007 trial): a non-randomised phase I-II study. *Lancet Oncol* 10: 489–500

Concordet J-P & Haeussler M (2018) CRISPOR: intuitive guide selection for CRISPR/Cas9 genome editing experiments and screens. *Nucleic Acids Res* 46: W242–W245

Doran SL, Stevanović S, Adhikary S, Gartner JJ, Jia L, Kwong MLM, Faquin WC, Hewitt SM, Sherry RM, Yang JC, *et al* (2019) T-Cell Receptor Gene Therapy for Human Papillomavirus–Associated Epithelial Cancers: A First-in-Human, Phase I/II Study. *J Clin Oncol* 37: 2759–2768

Ferrari S, Jacob A, Cesana D, Laugel M, Beretta S, Varesi A, Unali G, Conti A, Canarutto D, Albano L, *et al* (2022) Choice of template delivery mitigates the genotoxic risk and adverse impact of editing in human hematopoietic stem cells. *Cell Stem Cell* 29: 1428-1444.e9

Ferrari S, Vavassori V, Canarutto D, Jacob A, Castiello MC, Javed AO & Genovese P (2021) Gene Editing of Hematopoietic Stem Cells: Hopes and Hurdles Toward Clinical Translation. *Front Genome Ed* 3: 1–15

Ferris ST, Durai V, Wu R, Theisen DJ, Ward JP, Bern MD, Davidson JT, Bagadia P, Liu T, Briseño CG, *et al* (2020) cDC1 prime and are licensed by CD4+ T cells to induce anti-tumour immunity. *Nature* 584: 1–6

Ferrua F, Galimberti S, Courteille V, Slatter MA, Booth C, Moshous D, Neven B, Blanche S, Cavazzana M, Laberko A, *et al* (2019) Hematopoietic stem cell transplantation for CD40 ligand deficiency: results from an EBMT/ ESID-IEWP-SCETIDE-PIDTC study. *J Allergy Clin Immunol* 143: 2238–2253

Fox TA, Houghton BC & Booth C (2022) Gene Edited T Cell Therapies for Inborn Errors of Immunity. *Front Genome Ed* 4

Frangoul H, Altshuler D, Cappellini MD, Chen Y-S, Domm J, Eustace BK, Foell J, de la Fuente J, Grupp S, Handgretinger R, *et al* (2020) CRISPR-Cas9 Gene Editing for Sickle Cell Disease and β-Thalassemia. *N Engl J Med*

Hubbard N, Hagin D, Sommer K, Song Y, Khan I, Clough C, Ochs HD, Rawlings DJ, Scharenberg AM & Torgerson TR (2016) Targeted gene editing restores regulated CD40L function in X-linked hyper-IgM syndrome. *Blood* 127: 2513–2522

June CH & Sadelain M (2018) Chimeric antigen receptor therapy. *N Engl J Med* 379: 64–73

Kuo CY, Long JD, Campo-Fernandez B, de Oliveira S, Cooper AR, Romero Z, Hoban MD, Joglekar A v., Lill GR, Kaufman ML, *et al* (2018) Site-Specific Gene Editing of Human Hematopoietic Stem Cells for X-Linked Hyper-IgM Syndrome. *Cell Rep* 23: 2606–2616

de la Morena MT, Leonard D, Torgerson TR, Cabral-Marques O, Slatter M, Aghamohammadi A, Chandra S, Murguia-Favela L, Bonilla FA, Kanariou M, *et al* (2017) Long-term outcomes of 176 patients with X-linked hyper-IgM syndrome treated with or without hematopoietic cell transplantation. *J Allergy Clin Immunol* 139: 1282–1292

Nahmad AD, Reuveni E, Goldschmidt E, Tenne T, Liberman M, Horovitz-Fried M, Khosravi R, Kobo H, Reinstein E, Madi A, *et al* (2022) Frequent aneuploidy in primary human T cells after CRISPR–Cas9 cleavage. *Nat Biotechnol* 40: 1807–1813

Notarangelo LD, Giliani S & Plebani A (2014) CD40 and CD40 ligand deficiencies. In *Primary immunodeficiency diseases*, Ochs HD Smith CIE & Puck JM (eds) pp 324–342. Oxford: Oxford University Press

Sacco MG, Ungari M, Catò EM, Villa A, Strina D, Notarangelo LD, Jonkers J, Zecca L, Facchetti F & Vezzoni P (2000) Lymphoid abnormalities in CD40 ligand transgenic mice suggest the need for tight regulation in gene therapy approaches to hyper immunoglobulin M (IgM) syndrome. *Cancer Gene Ther* 7: 1299–1306

Sasu BJ, Opiteck GJ, Gopalakrishnan S, Kaimal V, Furmanak T, Huang D, Goswami A, He Y, Chen J, Nguyen A, *et al* (2022) Detection of chromosomal alteration after infusion of gene-edited allogeneic CAR T cells. *Mol Ther*

Schiroli G, Conti A, Ferrari S, della Volpe L, Jacob A, Albano L, Beretta S, Calabria A, Vavassori V, Gasparini P, *et al* (2019) Precise Gene Editing Preserves Hematopoietic Stem Cell Function following Transient p53-Mediated DNA Damage Response. *Cell Stem Cell* 24: 551-565.e8

Schiroli G, Ferrari S, Conway A, Jacob A, Capo V, Albano L, Plati T, Castiello MC, Sanvito F, Gennery AR, *et al* (2017) Preclinical modeling highlights the therapeutic potential of hematopoietic stem cell gene editing for correction of SCID-X1. *Sci Transl Med* 9

Tangye SG, Al-Herz W, Bousfiha A, Chatila T, Cunningham-Rundles C, Etzioni A, Franco JL, Holland SM, Klein C, Morio T, *et al* (2020) Human Inborn Errors of Immunity: 2019 Update on the Classification from the International Union of Immunological Societies Expert Committee. *J Clin Immunol* 40: 24–64

Tebas P, Stein D, Tang WW, Frank I, Wang SQ, Lee G, Spratt SK, Surosky RT, Giedlin MA, Nichol G, *et al* (2014) Gene editing of CCR5 in autologous CD4 T cells of persons infected with HIV. *N Engl J Med* 370: 901–10

Vavassori V, Mercuri E, Marcovecchio G, Castiello MC, Canarutto D, Asperti C, Jacob A, Albano L, Fontana E, Scanziani E, *et al* (2021a) Towards Clinical Translation

of Hematopoietic Cell Gene Editing for Treating Hyper-IgM Type 1. *Blood* 138: 3978–3978

Vavassori V, Mercuri E, Marcovecchio G, Castiello MC, Schiroli G, Albano L, Fontana E, Annoni A, Capo V, Margulies C, *et al* (2020) Modeling, optimization and comparative efficacy of HSC- and T-cell based editing strategies for treating Hyper IgM syndrome. *Mol Ther* 4S1: 405–6

Vavassori V, Mercuri E, Marcovecchio GE, Castiello MC, Schiroli G, Albano L, Margulies C, Buquicchio F, Fontana E, Beretta S, *et al* (2021b) Modeling, optimization, and comparable efficacy of T cell and hematopoietic stem cell gene editing for treating hyper-IgM syndrome. *EMBO Mol Med* 13: e13545

Yazdani R, Fekrvand S, Shahkarami S, Azizi G, Moazzami B, Abolhassani H & Aghamohammadi A (2019) The hyper IgM syndromes: Epidemiology, pathogenesis, clinical manifestations, diagnosis and management. *Clin Immunol* 198: 19–30

Appendix 1: Supplementary data
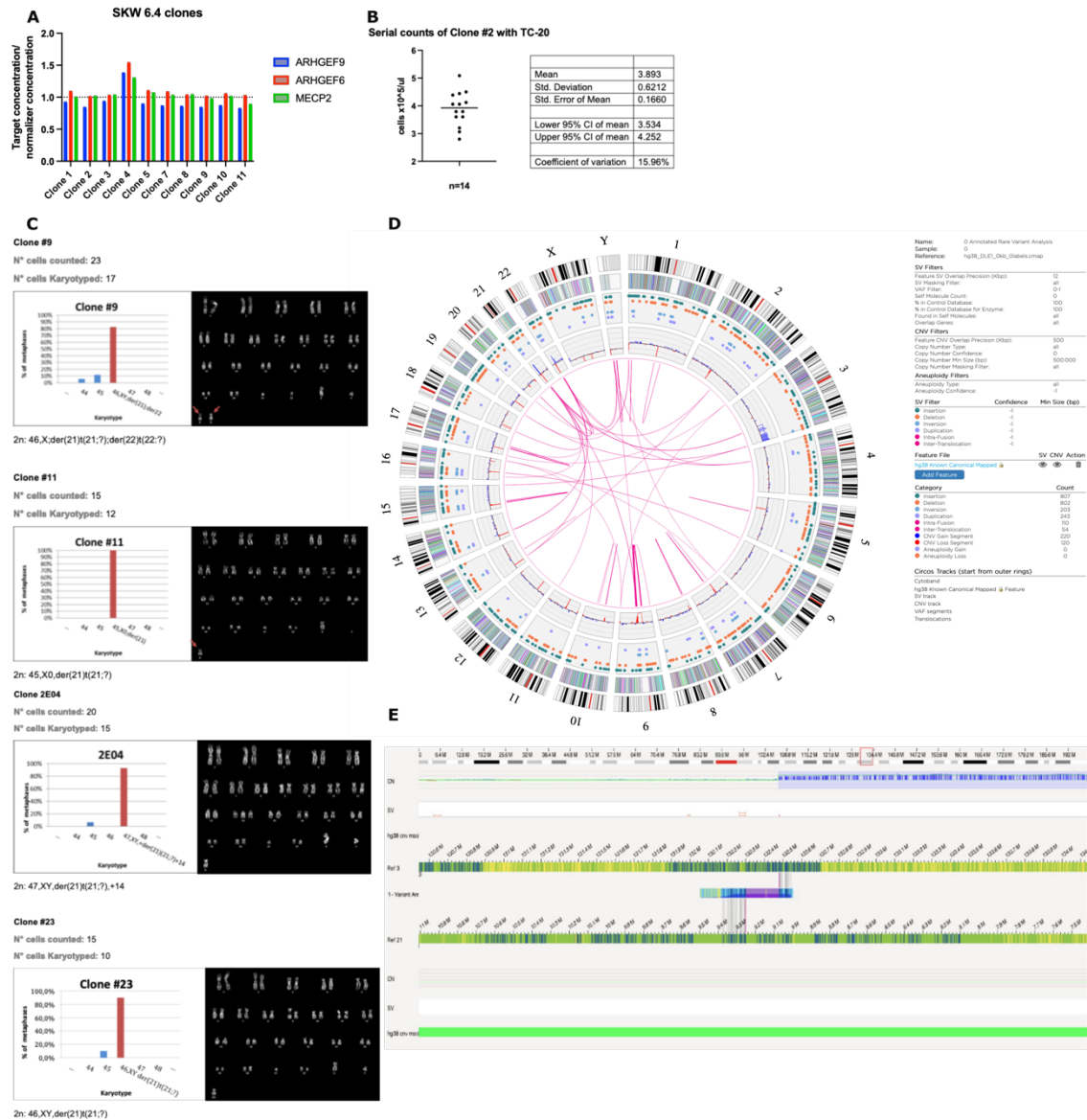
Figure 8: Characterization of SKW 6.4 clones.

A X-chromosome ploidy assessment in single SKW 6.4 clones. The relative number of X-chromosomes and the integrity of the X-chromosome with respect to chromosome 14 was estimated by ddPCR assays using MECP2, ARHGEF9 and ARHGEF96 targets and TTC5 as normalizer. B Estimation of TC-20 coefficient of variation in counting SKW 6.4 cells. The same sample was serially counted without dilutions nor trypan blue. C Karyotype of selected SKW 6.4 clones. D Unfiltered rare variant analysis of sample #0 E Translocation between chromosome 3q and 21p in SKW 6.4 sample #0. The translocation event t(3;21)(q22.1;p11.2) is shown by the blue and purple assembly; vertical red lines indicate the closest labels on each chromosome. Note masking of this region of chromosome 21 (solid green bar at the bottom) and the increase in copy number on chromosome 3 (CN track, in blue).
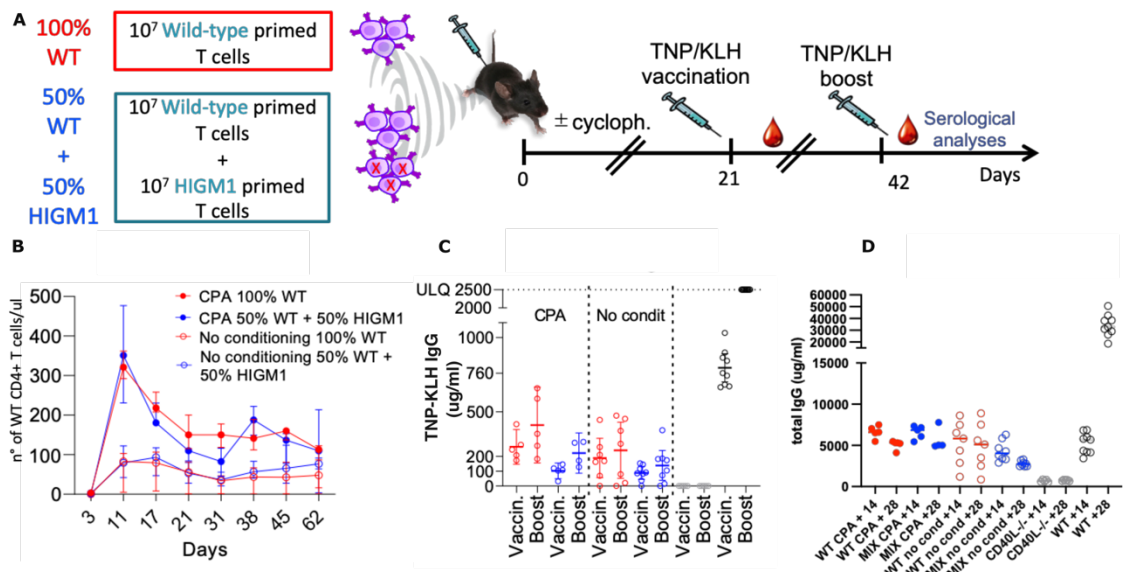
Figure 9: Competition between wild type and CD40LG deficient CD4+ T-cells.

A CD4+ T-cells from HIGM1 and wild type (WT) mice were primed in vivo against 2,4,6, Trinitrophenyl Keyhole Limpet Hemocyanin (TNP-KLH). Fourteen days after vaccination CD4+ T-cells were isolated from the spleen and activated in vitro with anti-CD3/antiCD28 beads as previously reported (Vavassori et al, 2021b). Recipient HIGM1 mice either received 300 mg/kg of cyclophosphamide (CPA) or not, and were transplanted with 10^7 primed WT CD4+ T-cells alone or 10^7 primed WT CD4+ T-cells admixed with 10^7 primed HIGM1 CD4+ T-cells. Mice were vaccinated 21 days and 42 days after CD4+ T cell infusion. Serum was collected 14 days after the first vaccination and 7 days after the second one. B Absolute engraftment levels of WT CD4+ T-cells by flow cytometry C-D TNP-KLH specific (C) and total IgG (D) production after the first vaccination (+14 days) and after boosting (+28 days) in the different subgroups. Grey circles indicate untreated HIGM1 mice, black circles untreated WT mice.

Table 3: Other events detected by optical mapping in sample #1.

Asterisk indicate overlapping events. Conf. Confidence. FCN fractional copy number. VAF variant allele frequency. Del deletion. Asterisk indicates events in the same region

Event	ID	Location	Event	Length (kb)	Conf.	FCN/ VAF	Found in sample #0
SV	7716*	chr16:78,406,139 - 78,411,527	Del	3.2	.99	/.6	CNV loss and 5 deletion events in chr16:78,284,533 – 78,900,325
SV	7700*	chr16:78,406,139 - 78,411,527	Del	3.2	.99	/.6	
SV	7706*	chr16:78,406,139 - 78,411,527	Del	3.1	.99	/.6	
CNV	*	Chr16:78,326,043- 78,764,981	Loss	438		1/	
SV	142	chr1:13,164,480 - 13,211,240	Del	3.5	.99	/.32	No
CNV	-	3q	Gain	Long	1	.5/ 0.38-0.8	Yes
CNV	25	chr1:143,278,152- 144,085,065	Loss	806	1	.5/	No
CNV	435	chrX:94,174,555 - 98,840,689	Loss	4666	1	1/	Yes
Chr 14 Aneuploidy					0	2.1/	No

Table 4: Other events detected by optical mapping in sample #5.

Asterisk indicate event overlap. FCN fractional copy number. VAF variant allele frequency. Del deletion. Ins insertion. Dup-inv inverted duplication

Event	ID	Location	Event	Length (kb)	Conf.	FCN/ VAF	Found in sample #0
SV	4645	chr7:74,377,433 - 74,437,190	Del	3.1	.95	/.48	N
SV	1744	chr2:192,671,029 - 192,697,109	Del	3.5	.99	/.02	N
SV	4758	chr7:102,619,401 - 102,691,518	Ins	3.7	.96	/.22	N
SV	14179	chrY:24,845,042 - 24,865,121	Dup-inv	20	-1	/.11	N
CNV	350	chr18:56,825,942 - 59,633,030	Gain	2807	1	/.14	N
CNV	145	chr6:51,724,024 - 54,342,312	Gain	22618	1	/.14	Y
CNV	238	chr10:113,438,902 - 114,828,458	Gain	1389	1	/.14	N
CNV	329	chr16:78,297,427 - 78,901,720	Loss	604	1	/.52	Y

CNV	196	chr8:126,925,696 - 127,805,714	Gain	880	1	/.18	Y
CNV	416	chrX:94,200,708 - 98,852,326	Loss	4651	1	/1	Y
CNV	120	chr4:183,734,942 - 184,938,665	Gain	1203	1	/.14	N
CNV	42	chr2:73,182,362 - 74,897,930	Gain	1715	1	/.09	N
CNV	5	chr1:40,667,623 - 41,773,766	Gain	1106	1	/.16	N
CNV	233	chr10:86,785,007 - 90,063,010	Gain	3278	1	/.12	N
CNV	2303	chr3:41,795,966 - 44,099,939	Gain	2303	.99	/.10	N

Table 5: CRISPOR predicted off-targets

List of off-targets predicted by CRISPOR (Concordet & Haeussler, 2018) on chromosomes listed in Table 2. The off-target overlapping an event found by optical mapping is in bold.

Off Target sequence	mismatchPos	mismatchCount	MIT off-target s	CFD off-target	chrom	start	end	strand	Description
AGGATAAATCCACTTTATCATGG	*....*.*.*.........	4	0,797205949	0,32	chr1	438807	438829	-	intergenic:RP4-669L17.2-OR4F29
AGGATAAATCCACTTTATCATGG	*....*.*.*.........	4	0,797205949	0,32	chr1	673786	673808	-	intergenic:MIR6723-OR4F16
TGGATCATTACACTAAATCAAGG	......*...*.....**....	4	0,036748005	0,334928229	chr1	10769872	10769894	+	intron:CASZ1
TGCATGATGGCACTTTGTTAGGG	..*.....*....*.*.	4	0,115680928	0,019392372	chr1	20334078	20334100	-	intron:VWA5B1
TGGATGATGGCACTGTTTTAAGG	.........*.....*.*.*.	4	0,028412066	0,001904762	chr1	38255484	38255506	+	intergenic:LINC01343-RNU6-753P
TAGATGTTTGCAGTTTGTCATGG	..*....*....*..*..	4	0,161125676	0,008908371	chr1	57459357	57459379	+	intergenic:DAB1-AL137855.1
TGGTTGATTGCATTTTCTTAGGG	...*......*.....*.*.	4	0,074311256	0,036281366	chr1	93665795	93665817	-	intron:BCAR3
TCCATGATTGCCTTTATAATAG	.**..........*...	3	0,208155556	0,029585799	chr1	103044925	103044947	-	intron:COL11A1
TAGATGATTTCACTTTTTTAAGG	.*.....*.....*.*.	4	0,176849269	0,042307692	chr1	194525816	194525838	-	intergenic:RNU6-983P-AL357932.1
TGGATCATTGCATTGAATCATGG	......*...*.**....	4	0,015441344	0,021694215	chr1	216709592	216709614	-	intron:ESRRG
TGGATGATTCACATTTATCACGG	........****........	4	0,127005026	0,030364373	chr1	218398197	218398219	+	intron:TGFB2
GGGAAGATGGCACTTTATCATGA	*....*.............	3	0,326554149	0,020560156	chr1	232923990	232924012	+	intergenic:AL122003.1-NTPCR
TGGAACATTCCACTTTTTCAAGG	...**..*.......*..	4	0,322464761	0,018181818	chr1	236687436	236687458	+	intron:ACTN2
TGTATGATTGCGCTTACTCCAGG	..*...........**..*	4	0,04310965	0,018230433	chr13	23689595	23689617	+	intergenic:TNFRSF19-AL161422.1
TGAATAATTGAACTATATCATGG	..*.*..*....*...*..	4	0,133372593	0,325657895	chr13	28666078	28666100	+	intron:POMP
TTTATGATAGCACTTTATTAGGG	..**.....*......*..	4	0,300469943	0,136363636	chr13	41663694	41663716	-	intron:VWA8
TGCATTATTGCACTGTTTCAAGG	..*.*......*....*...	4	0,092519726	0,001709402	chr13	42036020	42036042	-	intergenic:VWA8-AS1-DGKH
TGAAAAATTGCACTTTATCAGAG	..*.**............	3	0,276778266	0,097222222	chr13	58667772	58667794	-	intergenic:RNY4P29-AL359262.1
TCTATGATTGAAATTTATCAAGG	..**.......*.......	4	0,302995647	0,099852071	chr13	82352115	82352137	+	intergenic:RP11-452B18.2-RNU6-67P
TGGATGAGACCTCTTTATCAGGG	.......***.*.*....	4	0,361291854	0,090793651	chr13	105957287	105957309	-	intergenic:SNORA25-AL603632.1
TGCATCATTGCACTTTAGCCTGG	..*..*........*.*	4	0,077196351	0,019866497	chr13	112615007	112615029	-	intergenic:RP11-88E10.4-ATP11AUN
TGGATGATTTGACTTTATTAAGG	..........**......*.	3	0,430276044	0,09375	chr14	25964357	25964379	-	intron:RP11-314P15.2
TGGGTGATGCCACTGTATCAGGG	....*...*......*....	4	0,215770049	0,007738095	chr14	31129332	31129354	+	exon:HECTD1
TAGATAATTGCCCTCTTTCAGGG	..*..*.....*.*....	4	0,119911766	0,060728745	chr14	43124854	43124876	+	intergenic:CTD-2307P3.1-RP11-305B6.3
TGGTTGGTTGCACATTATGAAGG	..*..*......*.....*..	4	0,050756291	0,034759358	chr14	84384003	84384025	+	intergenic:RP11-353P15.1-RNU6-976P
AGGATGATTGCAGTTTATGAAGG	*..........*....*...	3	0,436194915	0,017045455	chr14	89575128	89575150	-	intergenic:RP11-33N16.3-FOXN3-AS2
TGGATGATTCCACTTACACAAGG	........*....***..	4	0,016316208	0,042780749	chr14	91197621	91197643	-	intron:C14orf159
TGCATGCTTGCACTTGACCAGGG	**..*..*........*.*.**	**4**	**0,035946325**	**0,024963018**	**chr14**	**98370560**	**98370582**	**-**	**intergenic:RN7SL714P-RP11-1082A3.1**
TGTACTATTGCACTTTATCCTGG	..*..**.............*	4	0,393858933	0,075757576	chr15	48366299	48366321	-	intergenic:DUT-FBN1
TTGATGATTTCACTATTTCATGG	.*.....*.....*.*..	4	0,150462235	0,039912281	chr15	57003515	57003537	-	intergenic:AC010999.1-TCF12
TGAAAGTTTGCACTTTATCAGAG	..*.*.*...........	3	0,326828148	0,042534722	chr15	61754840	61754862	+	intergenic:AC018618.1-RP11-16B9.1
TAGAAGATTACACTGTATCAGGG	..*.*..*......*....	4	0,371023101	0,01974359	chr15	81018476	81018498	-	intergenic:MESDC1-RP11-775C24.4
TGCATGGTGGCACTTATCAAGG	..*..*.....*.......	4	0,087716364	0,048019208	chr15	101135450	101135450	+	intergenic:RP11-424I19.1-RP11-424I19.2
AGGATAAATCCACTTTATCTGG	*....*.*.*.........	4	0,797205949	0,32	chr15	101888801	101888823	+	intergenic:OR4F15-OR4F4
TGCACCATTGCACTTTATCCTGG	..*.**.............*	4	0,393858933	0,059599491	chr16	28199608	28199630	-	intergenic:RNY1P10-XPO6
TGAATGTTTGCACTTTATGGAGG	..*......*........**	4	0,140060793	0,03136489	chr2	40122852	40122874	+	intron:SLC8A1-AS1
TGAATGATTGCAATTCAGCATGG	..*......*..*..*..	4	0,02036783	0,064102564	chr2	157908967	157908989	-	intergenic:RP11-546J1.1-RNU6-436P
TGTATGCTTTCACTTTAACAGGG	..*...*.......*.*..	4	0,192480029	0,118489583	chr2	212805275	212805297	-	intergenic:AC093381.2-AC079610.1
GGGGAGGATTGCAATTTATCAGAG	*...*........*......	3	0,230140845	0,061042983	chr2	226945024	226945046	+	intergenic:RHBDD1-SNORA48
TGCGTGATTGCATTTTGTCATGG	..*........*.....*..	4	0,220828507	0,02969457	chr2	236644950	236644972	+	intergenic:ACKR3-AC011286.1
TGGAAGAATGCACTTAATCCAGG	....*.*......*....*	4	0,113563	0,082644628	chr3	37543301	37543323	-	intergenic:ITGA9-RNU6-73P
GGGGAGGATTGCATTTCATCAGGG	*...*........*....	4	0,105393	0,28568116	chr3	52024699	52024721	+	intergenic:RPL29-DUSP7
TGCATGATTCCACTTTCTTAAGG	..*....*.....*.*.	4	0,174373379	0,012530456	chr3	108782809	108782831	-	intergenic:RETNLB-TRAT1
TAGAGGTTTGCTCTTTATCAGGG	..*.*.*.....*......	4	0,480774398	0,078974359	chr3	126063449	126063471	-	intergenic:RP11-158I23.1-SLC41A3
TGTATATTTGTACTTTATCATGG	..*..**..*........	4	0,308645779	0,067307692	chr3	148500067	148500089	+	intergenic:RP11-501O2.5-AGTR1
TGGATGCATGCACCCTATCAAGG	....**.....**.......	4	0,037226836	0,027272727	chr3	166514172	166514194	+	intergenic:BCHE-LINC01326
TGCATGATTGAAGTTTAACAGGG	..*.....*....*.*..	4	0,065721776	0,026223776	chr3	194168784	194168806	-	intergenic:RN7SL215P-RP11-513G11.4
TCCATGCATGCACTTTATCAAGG	..**....**.........	4	0,919204167	0,093195266	chr4	9858381	9858403	+	intron:SLC2A9
TGGGTGATAGCGCTATATCAAGG	....*...*.*.*....	4	0,115264782	0,242663951	chr4	84959094	84959116	-	intron:WDFY3
TGGATGACTGAACTTAACTGGG	........*..*.....*.*	4	0,068185438	0,3	chr4	104853543	104853565	-	intergenic:AC004053.2-RP11-556I14.2
ATCATGATGGCACTTACCAAGG	***..........*...	4	0,305988667	0,104895105	chr4	110829918	110829940	+	intergenic:LINC01438-AC024198.1
TGGATGATTTTATTTTATGAGGG	...........**.*......*..	4	0,089151706	0,021875	chr4	171118207	171118229	-	intergenic:RP11-344G13.1-MIR6082
GGGATGAATCCACTTGATCATGG	*......*.....*.*..	4	0,250819	0,105953177	chr4	178494950	178494972	-	intergenic:RNA5SP173-SNORD65
TGTATGATTCCACTTTTACAAGG	..*..*.........**..	4	0,108498991	0,017777778	chr4	188483352	188483374	-	intron:LINC01060
TGGATGATTCAACTTTACCATGG	..........**...*..	3	0,267727316	0,128571429	chr5	5238031	5238053	-	intron:ADAMTS16
TGGATATTTGCACTTTATCCAGG	.........*.......*.*	4	0,050766209	0,066287879	chr5	12701065	12701087	-	intergenic:LINC01194-RP11-419C19.2
AAGATGATTGCACTTTATTTGGG	**................**	4	0,219695863	0,234319527	chr5	20293534	20293556	+	intergenic:CDH18-RP11-420O16.1
AGAATGAGATGCACTTCATCATAG	*..*...*.....*...	4	0,259894908	0,13308642	chr5	25323722	25323744	+	intergenic:RP11-184E9.2-RP11-184E9.3
TTGAAGCTTGCATTTTATCATGG	..*.*.*.....*......	4	0,378170105	0,097443182	chr5	28460442	28460464	-	intergenic:CTD-2061E9.1-RNU6-909P
TGGATGATTCCATTCTAGCAAGG	..........*...*.*..	4	0,025554987	0,025454545	chr5	36040652	36040674	-	intron:UGT3A2
TGGATAATAGCACTGTTTCATGG	....*.*...*......*...	4	0,054569206	0,006190476	chr5	42128611	42128633	+	intergenic:Y_RNA-RP11-112L7.1
TGTATGATTGACTTTTATCAAGG	..*......*.*.*.*.*	4	0,070816487	0,011655012	chr5	54545480	54545502	+	exon:SNX18
TTCATGAATTCACTTTATCATAG	.**...*..........*...	4	0,2479025	0,041245791	chr5	58086092	58086114	-	intergenic:RP11-478P10.1-CTD-2037L6.2
TGGAGGATTGAATTTTCTCAAGG	....*....*....*.*..	4	0,124300022	0,059294118	chr5	67796701	67796723	+	intron:RP11-434D9.1
TGGATGGTGCCAATTTATCAAGG	......*.**..*.......	4	0,203023764	0,067226891	chr5	115134074	115134096	+	exon:TRIM36
TAGATGATAGCATTCTATCAAGG	.*......*...*.*....	4	0,095256253	0,15	chr5	140069601	140069623	+	intergenic:NRG2-LINC01024
TGGATGATTTGCATTTTGCATGGG	............*.**....	4	0,011018566	0,024019118	chr5	157912910	157912932	+	intergenic:CLINT1-RNU2-48P
TGCATCATTGCACTTTAGCCTGG	..*..*........*.*	4	0,077196351	0,019866497	chr5	161452950	161452972	-	intron:GABRB2
TGGTTCATTGCTCTTTATCTTGG	..*...*......*.....*.*	4	0,196530015	0,086776859	chr5	161737526	161737548	+	intergenic:GABRA6-GABRA1
AGGATAAATCCACTTTATCATGG	*....*.*.*.........	4	0,797205949	0,32	chr5	181380136	181380158	+	intergenic:OR4F3-Gap
TTCATTGTTGCACTTTATCACGG	.**...*.*..........	4	0,531673791	0,115178939	chr9	9674381	9674381	-	intron:PTPRD
TGAATAATTGCACTTAACCAAGG	..*..*......*.*..	4	0,031841181	0,292207792	chr9	20949076	20949098	+	intron:FOCAD
TTGGTGATTCCACTTTATCAAGA	..*.*..*...........	3	0,492236287	0,01104798	chr9	21785544	21785566	-	intergenic:RN7SL151P-MTAP/RP11-145E5.5
TGGATGATAGAAATTTTCAGGG	........*.*.*....*	4	0,068963652	0,035714286	chr9	40362240	40362262	+	intergenic:FAM95B1/RP11-146D12.2-AC129778.2
TGGATGATAGAAATTTTTCAGGG	........*.*.*....*	4	0,068963652	0,035714286	chr9	64502300	64502322	+	intergenic:ANKRD20A4-AL359955.1
TGGATGATAGAAATTTTTCAGGG	........*.*.*....*.	4	0,068963652	0,035714286	chr9	66014059	66014081	-	intergenic:FOXD4L4-RP11-15E1.5
AGGTTGATTCCACTTTATGATGG	*.*..*........*..	4	0,485227553	0,031818182	chr9	96662647	96662669	-	intergenic:AAED1-RP11-535M15.1
TGGATGCTTGTACTGTATCAAGG	......*.*.....*....	3	0,27147665	0,030730769	chr9	115811244	115811266	+	intergenic:RP11-284G10.1-LINC00474
TGCATGATTGTACTCTAGCATGG	..*......*...*.*..	4	0,045512754	0,010758472	chr9	122416784	122416806	-	exon:RP11-542K23.7
TGGAACATTAAACTTTATCAGGG	....**....**........	4	0,422106805	0,238636364	chr9	130576333	130576355	-	intergenic:ASS1-FUBP3
AGGATAAATCCACTTTATCATGG	*....*.*.*.........	4	0,797205949	0,32	chr9	138332742	138332764	+	intergenic:FO082796.1-Gap
ATGATAATTGCACTATATCAGGG	**...*......*.....	4	0,243723101	0,368421052	chrY	11938438	11938460	+	intergenic:AC134882.3-RN7SL702P

48

Table 6: Sequences sharing homologies to CD40LG gRNA in selected regions.

Listed sequences were found with SnapGene ver 6.2 allowing a mismatch or gap insertion for every 4 bases with TGGATGATTGCACTTTATCA(nGG). Potential PAM sequences are shown in bold. For sequences adjacent to nGG, nAG or nGA PAM sites mismatches are shown in red, () indicates a gap, and insertions are underlined.

Location on hg38	Homologous sequences
Chr2:57670638-57700364	TTCATGAAGCCTTTATCA-AAT
	AGGAGATTATACTTCATCA-CAA
	TGGATTTGCATTAATCT-TTG
	TGAGATATGCATTTATTA-CAA
	TATAAACAACATCCA-CAC
	TGGTTTATTTCACTTAATA-TAA
	AGGATGATGACTTTGACA-CCG
Chr3:58006840-58123739	TGGAATTACACTTTTCT-CTT
	AGCATGAATGCCTTTGTCA-AAA
	TATATGATTCCATTTATA-TAA
	TGGCTGATGCACTGGACA-GTG
	TGAATGATTGTATTTATGG-CCT
	TAGATGAGACACTTTTCA-GAT
	TGGACGAATGCTTTTCA-CAT
	GGGAGATTGCATCTCTACA-AAT
	TAATTATTGCCTTTATA-GCC
	TGGATAGACTGGACTTTTTCA-AGT
	TGAGATTCTCTTTTTCA-ATG
	TGGTTGAATGGGCTTTATCT-CAT
	TAGATGATTACCTATTCA-GAT
	TGGATTTTGACATTATTCA-CAA
	TGG()T()()TTGCATTTTATCA-**GAG**
	TAATGAATGCACTTAACA-CTA
	TGGATATTAGTTCCTTATCA-GAT
	TGGTTTGCCTTTTCA-CGC
	GGGATAATTCCACTTTAACC-CTC
Chr13:57143300-57261030	TGGATGCTTTCCTTGACCA-TGC (x4)
	TCGATGATTAGATTAATCA-ATC (x4)
	TGGTTGATTGATTAATCT-AAT (x4)
	TGTATGATTTTTCTTATCA-ATT
	TGAAATTGCATTTATAA-AAA
	TGCAAGTATTCATTTTATCA-TTT
	TGAATGCTT()CA()TTTAGCA-**AGA**
	TGGCT<u>G</u>GATT()CACTTTCTCC-**AGG**
	TGATCATTGTCTTTATCA-GAA
	TGTTGATTACAATTTACCA-GCT
	TGG()TG()TTTCA()TTTATTA-**TAG**
	TGGATATTAGCCCTTTGTCA-GAT

50	TGTGATAGACTGTTATCA-TTT
	TGGTGATCAGGTTATCA-GAT
	TGGATGAGTTCACATACA-ACC
	TGGAAGTGATTCATTTATTA-TCG
	TGGCGATATTCCCTTTATCA-TTT
	TGGATATTAGCCCTTTGTCA-GAT
	TGTATGATCTCACTTATAT()A-TGG
	TGATTATAACTTTATCA-CTA
	TTGATCATTCAATGTTATCA-GAT
	TTGATGTTGCAGATTTACTCA-TTT

Appendix 2: Trial design synopsis

Protocol Title

A Single Arm, Phase I/II, Open Label Clinical Study of Autologous CD4+ T-Cells Edited Ex-Vivo at the CD40LG Locus by CRISPR/Cas9 and IDLV-based vector in Participants with X-linked Hyper IgM Syndrome Type 1 (HIGM1).

Proposed product name

Autologous peripheral blood CD4+ T-cells CRISPR-edited at the CD40LG locus for the treatment of Hyper IgM Type 1 (HIGM1).

Protocol acronym

TIGET EDI-T-HIGM1

Short Title

Study Number XXX, A Clinical Study to Evaluate the Use of a CD40LG Gene Edited T-Cells in Participants with HIGM1.

Abbreviations

AAV6	Adeno-associated virus type 6
ACT	Abnormal clonal proliferation
AE	Adverse event
DP	Drug product
DLT	Dose-limiting toxicities
HIGM1	X-linked Hyper IgM Syndrome Type 1
HSCT	Hematopoietic stem cell transplantation
IDLV	Integrase defective lentivirus
L.O.Q.	Limit of quantification
MFI	Mean fluorescence intensity
ΔNGFR	C-terminal truncated low-affinity NGFR receptor

1.1 Rationale

Patients with HIGM1 characteristically present a high susceptibility to bacterial, intracellular and opportunistic pathogens (mainly *Pneumocystis jiroveci* and *Cryptosporidium spp*.), and may develop biliary tract and liver disease, neutropenia, autoimmunity and malignancies. In the past, long-term survival has been poor with conservative therapy only, based on chronic immunoglobulin supplementation and antimicrobial prophylaxis. Median survival time from diagnosis of 25 years has been reported (de la Morena *et al*, 2017).

Currently, allogeneic hematopoietic stem cell transplantation (HSCT) is the only curative treatment available. Recent data show that best outcome is obtained when HSCT is performed early after diagnosis, before 10 years of age, in absence of pre-

existing organ damage (mainly liver disease), and using myeloablative regimens. Superior overall survival is attained in transplants from matched donors (Ferrua *et al*, 2019). However, HSCT still remains associated with risks of transplant-related complications leading to morbidity and mortality, especially in older patients, those with organ damage, and those with poor donor availability. Thus, therapeutic alternatives to safely and more effectively treat patients for whom HSCT is too risky are strongly needed.

The HIGM1 phenotype is mostly consequential to defective CD40L expression by CD4+ T-cells, and may thus benefit from correction of defective cells. Genetically modified T-cells have been employed safely in a number of settings, including allogenic stem cell transplantation for the prevention of graft versus host disease (Ciceri *et al*, 2009), treatment of infectious diseases (Tebas *et al*, 2014), and treatment of malignancies (Doran *et al*, 2019; June & Sadelain, 2018).

CD40LG is a tightly regulated gene, and its unregulated expression can give rise to T- (Brown *et al*, 1998) and B- lymphoproliperative disorders (Sacco *et al*, 2000). Thus, HIGM1 is not amenable to a more traditional gene addition approach, while gene editing aims to restore the physiological regulation of expression of CD40LG from its native locus.

SR-Tiget has been developing an innovative strategy based on a formulation of ex-vivo gene edited autologous CD4+ lymphocytes, hereafter referred to as the drug product (DP). We first showed transplanting wild-type CD4+ T lymphocytes in a syngeneic mouse model of HIGM1 results in a partial rescue of IgG production, especially against pre-experienced antigens, and protection against *Pneumocystis murina* and thus can ameliorate the disease phenotype. We have also demonstrated the feasibility of correcting human CD4+ T lymphocytes with a CRISPR/Cas9 and AAV6-vector based editing technology; gene edited cells were capable of achieving regulated expression of CD40LG and induce IgG class switching in vitro (Vavassori *et al*, 2020). Subsequently, the manufacturing process has been modified to use an integrase defective lentiviral vector (IDLV) instead of AAV6.

Further non-clinical toxicology and biodistribution studies are planned.

The sponsor proposes a first-in-man, Phase I/II clinical study (TIGET EDI-T-HIGM1) to treat patients with HIGM1. The study will generate clinical safety and efficacy data in patients with HIGM1 with the DP through a dose-escalation design.

1.2 Objectives and Endpoints

Objective	Endpoint	Assessment
Primary Safety		
Evaluation of the short-term (28 days) safety and tolerability of the DP	1 Incidence of dose-limiting toxicities (DLT) 2 Establishing the tolerability of the DP dose infused according to the dose escalation protocol	1 Incidence of DP-related adverse events (AE) as defined in the protocol 2 As defined in the protocol
Evaluation of the medium long/term safety of the DP	1 Safety and tolerability of the DP 2 Absence of malignancy related to the DP	1 As measured by AE reporting 2 ACP monitoring and Safety and tolerability as measured by AE reporting
Primary Efficacy		
Evaluation of the engraftment and biological efficacy of the DP	1 Detectable CD4+ edited cells at 7, 14 and 28 days after the each dose of the DP (>L.O.Q.) 2 Increase in CD4+ edited cell engraftment after dose escalation 3 Restoration of CD40L expression and binding to CD40 in edited cells	1 and 2: ddPCR of subjects with edited cells 3 Percentage of cells expressing CD40L after PMA/ionomycin stimulation (by flow cytometry) and MFI intensity of CD40L expression in positive cells after PMA/ionomycin stimulation by FACS OR percentage of cells capable of binding to CD40 after PMA/ionomycin stimulation
Secondary Safety		
Evaluation of the immunogenicity	1 Absence of immune response to CD40L 2 Absence of immune response to ΔNGFR 3 Absence of new, clinically evident autoimmune disorders	1 and 2: To be defined 3 AE reporting

Secondary Efficacy		
Evaluation of manufacturing feasibility	Infusion of the DP	1 Generation of DP that meets the target dose and release criteria 2 Percentage of patients that receive the DP
Exploratory		
Evaluation of the biological efficacy of the DP at 1 month after the last infusion	Specific IgG production in response to vaccination	Specific Antibody titre
Evaluation of sustained engraftment of the DP	1 Long term persistence of detectable CD4+ edited cells >L.O.Q. 2 Restoration of CD40L expression	1 PCR of subjects with edited cells at PCR 2 Percentage of cells expressing CD40L after PMA/ionomycin stimulation (by FACS) and MFI intensity of CD40L expression in positive cells after PMA/ionomycin stimulation by FACS.
Evaluation of the biological efficacy of the DP at 1, 3, 6 and 12 months after the last dose of the DP	1 Evidence of total IgG production (in vivo) 2 Evidence of total IgA/IgE production (in vivo) 3 Reduction of high IgM levels (in those with high level before treatment) 4 Switched memory B-cell count in PB 5 T-cell function	1 Serum IgG levels 2 Serum IgA levels 3 Serum IgM levels 4 Immunophenotype 5 Proliferative response to mitogens and antigens.
Evaluation of the extent of immune reconstitution	1 Cytokine profile and production in vitro (including IFNg, IL12, IL10,	1 ELISA, Elispot/Quantiferon, Bioplex TCR and IgH sequencing (DNA)

	IL4, IL5, BAFF QFT) before and after treatment 2 Study and tracking of TCR repertoire 3 Study and tracking of BCR repertoire 4 Number of edited stem cell memory cells 5 B-cell compartment distribution including number of switched memory IgG-IgA B-cell memory (before and after gene therapy). 6 Treg and Th profile before and after treatment 7 ΔNGFR expression 8 Secretory IgA production 9 Evidence of IgG production (in vitro) 10 Presence of autoantibodies 11 Presence of sCD40LG	2 TBD 3-7: FACS 8 ELISA 9 T-B co-culture 10 Autoantibody titre (in vivo) 11 TBD
Characterisation of editing events	Identification and tracking of genome integrity events including off-target integrations, undesired integrations, deletions/translocations, NHEJ events	100-metaphase karyotype, ddPCR, other TBD
Evaluation of the clinical efficacy of the DP	1 Overall survival 2 Amelioration or resolution of chronic infections (if present before treatment) 3 Infection rate before/after treatment	1-3: Disease-progression related AE monitoring as specified in the protocol 4-5: Disease-progression related concomitant

	4 Frequency of Ig supplementation 5 Requirement for antimicrobial prophylaxis 6 Performance Status 7 Quality of life	medication recording as specified in the protocol 6 Lansky/Karnofski scores 7 Paediatric Quality of Life Inventory (PedsQL) (for participants < 18 years) and SF-36 (for participants ≥ 18 years)

1.3 Overall Design

This Phase I/II single centre clinical trial is designed as a three steps staggered study and follows an adaptive design. The intent of this study is to generate safety and efficacy data for patients with HIGM1 treated with the DP.

The following are foreseen:

1.3.1 Recruitment phase

Signature of informed consent for Steps 2 and 3 will specify whether lymphodepletion will be administered or not, depending on the results of the interim analysis.

Criteria to enroll each additional patient: No DP-related life-threatening event in 2 or more /5 patients within 1 month of DP administration

1.3.2 Screening phase

The patient will be evaluated for their eligibility as defined by the inclusion and exclusion criteria; sequencing of the CD40LG gene may be required if not already available.

1.3.3 The baseline phase will include (equal for all steps):

- Instrumental workup and laboratory analysis
- Vaccination course against rabies (3 doses) and tick-borne encephalitis (2 doses), last dose 2-4 weeks before planned apheresis (exceptions may be granted on clinical grounds, e.g. anaphylaxis).
- Leukocyte apheresis, to be performed 2-4 weeks after last vaccine dose
- Should the administration of lymphodepletion be foreseen, the opportunity of performing a lymphocyte backup will be evaluated

1.3.4 Treatment phase (see Appendix 2 Figure 10)

Step 1: dose escalation (3 patients).
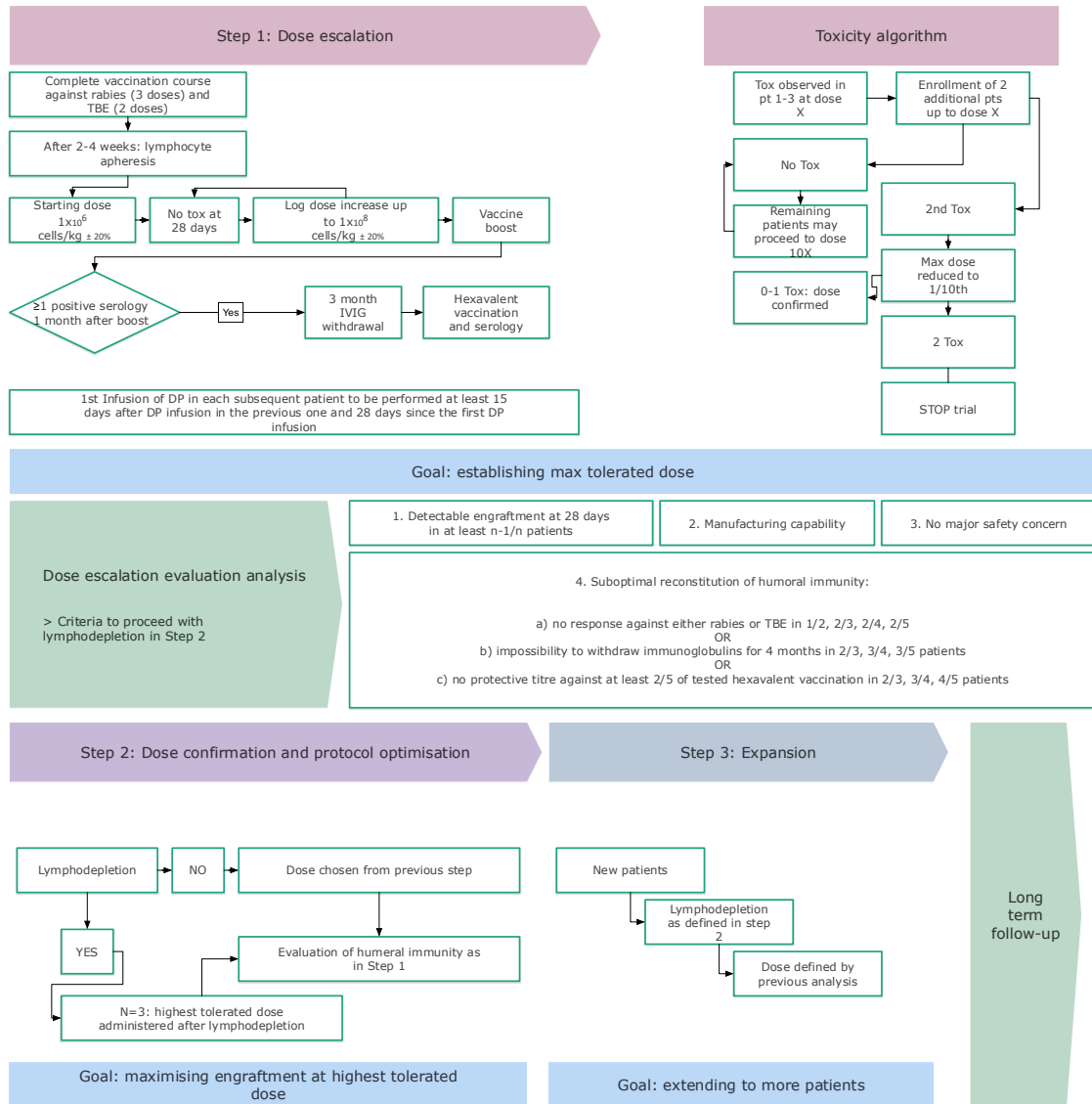
- The starting dose will be 1-2 x10^6 cells/kg ±20%.

- In the absence of grade 4, DP-related adverse effects and safety concerns, 28 days after the first infusion $1\text{-}2\text{x}10^7$ cells/kg ±20% will be administered to the same patient.

- In the absence of grade 4, DP-related adverse effects and safety concerns, 28 days after the second infusion $1\text{-}2\text{x}10^8$ cells/kg ±20% will be administered to the same patient.

- 15 days after the maximum tolerated DP administration, or as soon as the patient is in good clinical conditions, a vaccine boost will be administered, and response will be assessed after 15 days and after 1 month.

- If a measurable IgG response is documented against at least one vaccine, immunoglobulin supplementation may be withdrawn. Serum IgG will be monitored every 15 days; Ig supplementation must be restarted if IgG levels fall below 500 mg/dL or the lower limit for age (if below 500 mg/dL).

- Should IgG levels remain above the aforementioned limits for 3 months without immunoglobulin supplementation, patients will receive hexavalent vaccine and IgG specific response against HbsAg, tetanus, difteria toxoid, Bordetella pertussis and H. influenzae will be measured after 3-4 weeks.

Dose escalation evaluation: the following will be discussed with the advisory board:

- Safety data collected in Step 1
- Treatment dose for Step 2
- Dose schedule for Step 2
- Depending on safety and efficacy data collected in the first step, the PI may discuss with the sponsor the opportunity of administering or withdrawing the administration of lymphodepleting/immunomodulating agents before the infusion in Step 2. The following criteria will need to be met to proceed with the administration of lymphodepletion in Step 2:
  o Adequate manufacturing capability
  o Detectable engraftment of the DP in n-1/n Step 1 patients at 28 days
  o No major safety concerns
  o Suboptimal reconstitution of humoral immunity:
    ▪ Absent response against at least one between TBE and rabies in 1/2, 2/3, 2/4, or 2/5 patients OR
    ▪ Impossibility to withdraw immunoglobulin supplementation for 4 months in 2/3, 3/4, 3/5 patients

- No protective titer against at least 2/5 of tested antigens of hexavalent vaccine in at least 2/3, 3/4, 4/5 patients

Step 1: Dose escalation

- Complete vaccination course against rabies (3 doses) and TBE (2 doses)
- After 2-4 weeks: lymphocyte apheresis
- Starting dose 1×10^6 cells/kg ± 20%
- No tox at 28 days
- Log dose increase up to 1×10^8 cells/kg ± 20%
- Vaccine boost
- ≥1 positive serology 1 month after boost → Yes → 3 month IVIG withdrawal → Hexavalent vaccination and serology
- 1st Infusion of DP in each subsequent patient to be performed at least 15 days after DP infusion in the previous one and 28 days since the first DP infusion

Toxicity algorithm

- Tox observed in pt 1-3 at dose X → Enrollment of 2 additional pts up to dose X
- No Tox
- Remaining patients may proceed to dose 10X
- 0-1 Tox: dose confirmed
- 2nd Tox
- Max dose reduced to 1/10th
- 2 Tox
- STOP trial

Goal: establishing max tolerated dose

Dose escalation evaluation analysis

> Criteria to proceed with lymphodepletion in Step 2

1. Detectable engraftment at 28 days in at least n-1/n patients
2. Manufacturing capability
3. No major safety concern
4. Suboptimal reconstitution of humoral immunity:

a) no response against either rabies or TBE in 1/2, 2/3, 2/4, 2/5
OR
b) impossibility to withdraw immunoglobulins for 4 months in 2/3, 3/4, 3/5 patients
OR
c) no protective titre against at least 2/5 of tested hexavalent vaccination in 2/3, 3/4, 4/5 patients

Step 2: Dose confirmation and protocol optimisation

- Lymphodepletion → NO → Dose chosen from previous step
- YES
- N=3: highest tolerated dose administered after lymphodepletion
- Evaluation of humeral immunity as in Step 1

Goal: maximising engraftment at highest tolerated dose

Step 3: Expansion

- New patients → Lymphodepletion as defined in step 2 → Dose defined by previous analysis

Goal: extending to more patients

Long term follow-up

Figure 10: Block diagram of trial design

Step 2: dose confirmation and protocol optimization (3 patients).

- The lymphodepleting regimen and the DP dose defined in the dose escalation evaluation analysis will be administered in 3 additional patients.
- 15 days after the DP administration, or as soon as the patient is in good clinical conditions, a vaccine boost will be administered, and response will be assessed after 15 days and after 1 month.
- If a measurable IgG response is documented against at least one vaccine, immunoglobulin supplementation may be withdrawn. Serum IgG will be

monitored every 15 days; Ig supplementation must be restarted if IgG levels fall below 500 mg/dL or the lower limit for age (if below 500 mg/dL).

- Should IgG levels remain above the aforementioned limits for 3 months without immunoglobulin supplementation, patients will receive hexavalent vaccine and IgG specific response against HbsAg, tetanus, diphtheria toxoid, Bordetella pertussis and H. influenzae will be measured after 3-4 weeks.
- Additional infusions may be performed depending on the availability of the DP.

Step 3: expansion. Depending on safety and efficacy data, the trial may enroll up to 3 additional patients to consolidate step 2 results. The lymphodepletion regimen and the DP dose may be adjusted if necessary.

Catch-up: patients enrolled in Step 1 may be treated again with the best treatment protocol, if appropriate

1.4 Key elements of study design

Study Design	Open-label, single arm, single centre, staggered dose escalation, Phase I/II
Population	Participants with HIGM1
Sample size (n)	It is anticipated that 9 participants may be enrolled. A minimum of 6 participants are planned to assess the primary endpoints. Additional patients may be enrolled should significant toxicity be observed during Step 1 or Step 2, according to specified criteria.
Key Inclusion Criteria	- Genetic diagnosis of HIGM1 - Absent expression of CD40L on CD4+ T-cells upon in vitro stimulation with PMA/ionomycin - OR Absent binding to CD40 by CD40LG - OR Absent downstream signaling by CD40 upon binding to mutated CD40LG (in this case feasibility of the correction to be demonstrated on patient cells) - Requirement of chronic Ig supplementation - Suitability to perform a leukapheresis - Signature of informed consent

	• No adequate HLA-identical sibling donor available OR no matched unrelated donor (9/10 or 10/10) within 3 months o OR allogeneic unrelated transplant considered to be too high risk • Age >= 3 years

Key Exclusion Criteria	• CD40LG mutation upstream of the vector insertion site in intron 1 or any other mutation not amenable to correction with gene editing
	• Documented HIV RNA, HCV RNA or HBV DNA positivity, presence of total syphilis antibodies and compliance to Tissue Directive (manufacturing issue). Patients with positive Hepatitis C or Hepatitis B antibody due to prior resolved disease may be enrolled, only if repeated confirmatory negative RNA test are obtained
	• End-organ dysfunction or other severe disease or clinical condition which, in the judgment of the investigator, would make the subject inappropriate for entry into this study, including – but not limited to – the following:
	    o End stage liver disease, defined as
	        ▪ patients awaiting liver transplantation
	        ▪ sclerosing cholangitis
	        ▪ PELD(Chang *et al*, 2018) >20 or MELD-Na >20
	        ▪ Severe fibrosis (F3-F4)
	• Clinically evident CNS impairment by neurological examination
	• Previous immune reconstitution inflammatory syndrome (IRIS)
	• Past or current cryptococcal meningitis
	• DLCO or FEV1 <80% of predicted
	    o OR dyspnea on slight activity
	• Clinically relevant cardiac disease
	    o NYHA >=II
	    o Ejection fraction < 50%

	• Malignant neoplasia (except local skin cancer) or a documented history of hereditary cancer syndrome. Participants with a prior successfully treated malignancy and a sufficient follow-up to exclude recurrence (based on oncologist opinion) can be included after discussion and approval by the sponsor. • Leukaemia/lymphoma or other forms of lymphopropliheration, clinically relevant cytogenetic alterations characteristic at CGH array and/or karyotype performed on 100 metaphases, or other serious haematological disorders. • Has previously undergone allogeneic hematopoietic stem cell transplantation and has evidence of residual cells of donor origin. • Patient enrolled in other clinical trials • Patient unable to donate a sufficient number of lymphocytes for drug product manufacturing • Systemic corticosteroid therapy or other immunosuppressive drugs that may interfere with the DP within 4 weeks before the apheresis and/or DP infusion
Cell Harvesting	Cells will be harvested from 1-3 consecutive apheresis of lymphocytes. The residual fraction from the manufacturing process will be stored as a potential backup source.
Manufacture of the DP from Participant Purified Cells	The preparation of the participant's specific edited cells will be performed at _____ under GMP conditions. Purified cells will be edited with CRISPR/Cas9 and IDLV vector technology.
Lymphodepleting/ immunomodulating regimen	The DP administrations in Step 1 of the trial will not be preceded by lymphodepletion in order to assess in vivo tolerability and safety of this approach, without interference with the additional potential preparatory regimen-related toxicity. Based on the dose escalation evaluation analysis performed after Step 1, the use of fludarabine 30 mg/mq/die for 3 days before the following doses in subsequent patients will be determined.

Drug Substance	The drug substance is defined as autologous peripheral blood CD4+ T-cells CRISPR-edited at the CD40LG locus for the treatment of Hyper IgM Type 1 (HIGM1).
Drug Product (DP)	The drug product (final medicinal product: _____) is defined as the cryopreserved drug substance supplied in an ethylene vinyl acetate (EVA) bag.

Dose of the Drug Product	Three doses will be manufactured for each patient
	    -1×10^6 cells/kg ±20% (one dose)
	    -1×10^7 cells/kg ±20% (one dose)
	    -1×10^8 cells/kg ±20% (one or more doses)
	
	The study follows an adaptive design divided in 3 consecutive steps:
	
	Step 1: dose escalation (3 patients)
	    -The starting dose is $1-2 \times 10^6$ cells/kg.
	    -In the absence of dose limiting adverse effects, 28 days after the infusion the patient will receive an additional dose of $1-2 \times 10^7$ cells/kg.
	    -In the absence of dose limiting adverse effects, 28 days after the infusion the patient will receive an additional dose of $1-2 \times 10^8$ cells/kg.
	
	Patient 2 may receive the first infusion 28 days after the first infusion in Patient 1, and Patient 3 may receive the first infusion 15 days after the first infusion in Patient 2.
	
	If one patient experiences a DLT, defined as a grade ≥ 4 AE attributed to the DP he or she shall:
	    -not receive additional DP doses if the DLT occurred after the first dose or the PI/advisory board deem it appropriate
	    -eventually receive log inferior doses in the subsequent steps should the PI/advisory board deem it appropriate AND it is presumed that it may not occur at a lower product dose.
	    If no grade ≥4 adverse event is observed for the highest dose, the dose will be confirmed as the highest dose.

	If 1 patient experiences a grade ≥4 AE for a certain dose, 2 more additional patients will be recruited and escalate up to the same dose. If another patient experiences a grade ≥4 AE for a certain dose, subsequent patients will be limited to receiving a dose equal to 1/10 of the toxic dose.
	
	If no other patient experiences a DLT, the other patients may continue the dose escalation.
	
	Dose escalation evaluation analysis will then take place to define :
	    -DP Dose and Schedule
	    -Lymphodepletion
	
	Step 2: Dose confirmation and protocol optimization (3 additional patients)
	    Up to 3 additional patients will be enrolled for the second phase of the study and receive the DP at the dose defined in the dose escalation evaluation analysis.
	    Depending on safety and efficacy data collected in the first phase, lymphodepletion is expected to be administered in this phase
	
	Step 3: Expansion (up to 3 additional patients)
	    Up to 3 patients of any age may be enrolled. DP dose and lymphodepletion may be adjusted.
	
	    Catch-up: The patients enrolled in Step 1 may catch-up with the rest of the cohort by proceeding to the dose confirmation step and receive additional dose(s) of the DP.

Administration of the Drug Product	The administration of the DP will be performed into a central vein over 30 minutes using an infusion pump. During DP infusion, vital signs (pulse, blood pressure, respiration rate, and body temperature) will be monitored at 30-minute intervals for the first 3 hours and every hour for the following 3 hours.

	The patient will be hospitalised for at least 24 hours after the DP infusion.
Duration of Follow up	All participants treated with the DP will be monitored for a period of 2 years after gene therapy to evaluate safety and efficacy. After completion of the 2 year follow up, participants will be enrolled in a long-term follow up study or a registry, as permitted by local regulations, and followed up for at least 6 additional years or until the DP can no longer be recovered from the blood.

Analyses Plan

The sample size for this study been determined based on feasibility and to gain clinical experience with the DP. It is not based on statistical considerations. If the number of participants is deemed adequate, summary statistics may be presented as appropriate. Preliminary participant data will be collated and reviewed on an ongoing basis throughout the study. The following data reviews are scheduled:

Type	Purpose	Data cut point (data availability):	Parameters for review
Dose escalation evaluation analysis	Defining: Treatment dose Schedule Lymphodepletion	Three subjects have been followed up for 6 months after the last dose of the DP	Primary endpoint and safety data
Interim analysis	Interim analysis	Six subjects have been followed up for 6 months post gene therapy	Primary endpoint and safety data
Final Analysis	Final analysis	All subjects have been followed up for 2 years post gene therapy	All available data